Journal of Engineering Managment and Soft Computing (JEMSC)

Forecasting Stock Market Volatility: A Wavelet-Enhanced Hybrid GARCH-Deep Learning Approach

Masoumeh Molaei[1], Hamid Moradi-Kamali[2] and Reza Entezari-Maleki[3]✉

1. MSc., School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran. Email: molaei_m@comp.iust.ac.ir
2. MSc., School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran. Email: ha_moradi@comp.iust.ac.ir
3. Corresponding author, Assistant Prof., School of Computer Engineering, Iran University of Science and Technology, Tehran, Iran. Email: entezari@iust.ac.ir

Article Info

Article type:
Research Article

Article history:
Received 4 August 2025
Received in revised form 5 February 2026
Accepted 8 March 2026
Published online 1 April 2026

Keywords:
Volatility prediction, Deep Learning, Statistical Modeling, Wavelet Transforms.

ABSTRACT

Accurate volatility forecasting is vital for effective decision-making in financial markets, yet remains a complex challenge due to the noisy, non-stationary nature of financial time series and the diversity of volatility definitions. This study proposes a novel hybrid framework that combines statistical, signal processing, and machine learning techniques to enhance volatility prediction. The approach begins with wavelet transformations to extract multi-scale features from raw financial data, effectively addressing non-stationarity. These features are then evaluated using multiple volatility estimators to determine their predictive relevance. The framework integrates GARCH models, wavelet-derived inputs, and deep learning architectures, with Particle Swarm Optimization (PSO) employed for optimal parameter tuning. Leveraging S&P 500 data from 2000 to 2024 and incorporating multi-source inputs, the model achieves a more holistic representation of market dynamics. Empirical results demonstrate that the hybrid method significantly reduces prediction errors and consistently outperforms baseline models and established benchmarks. To validate its practical utility, we developed trading strategies based on the predicted volatility. Backtesting results indicate substantial performance gains: the top-performing strategy tripled net profit growth and reduced maximum drawdown by 39% compared to a no-forecast baseline. These findings underscore the potential of advanced hybrid volatility modeling to improve trading outcomes and mitigate financial risk.

Cite this article: Entezari-Maleki, R. & et al, (2026)., Forecasting Stock Market Volatility: A Wavelet-Enhanced Hybrid GARCH-Deep Learning Approach. *Engineering Management and Soft Computing,* 12 (2). 56-74.

DOI: https://doi.org/10.22091/jemsc.2026.13545.1291

DOI: https://doi.org/10.22091/jemsc.2026.13545.1291

Publisher: University of Qom

1) Introduction

Financial volatility, as a measure of the magnitude of price fluctuations in financial markets, plays a central role in economic decision-making processes. Accurate volatility forecasting, particularly in dynamic and rapidly evolving financial markets, is of critical importance for a wide range of stakeholders, including investors, algorithmic traders, and risk managers. Such forecasts not only provide the foundation for optimal asset allocation and the pricing of derivative instruments, such as options, but also play a key role in the design of hedging strategies and portfolio risk management frameworks (Poon & Granger, 2003). Despite its importance, volatility forecasting remains a challenging task due to the complex and dynamic nature of financial markets. On the one hand, financial data are typically highly noisy, making it difficult to distinguish meaningful signals from random fluctuations and to identify genuine underlying patterns. On the other hand, statistical non-stationarity in financial time series, particularly time-varying behavior in both the mean and variance, poses a significant obstacle to reliable modeling (Li et al., 2022). In addition, volatility often exhibits long-memory properties and clustering behavior, such that periods of high or low volatility tend to persist over time (Poon & Granger, 2003). Moreover, it is important to recognize that numerous methods have been developed not only for volatility forecasting, but the very concept of volatility itself admits multiple definitions, such as conditional volatility, historical volatility, and implied volatility, each of which requires its own appropriate estimators and modeling techniques (Olubusola et al., 2024).

To address these challenges, recent research has increasingly shifted toward the use of more sophisticated and hybrid modeling frameworks, in which the strengths of statistical models are combined with the ability of machine learning methods to capture nonlinear patterns (Namdari Birgani et al., 2024). Within this context, models belonging to the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) family have emerged as one of the most widely used traditional approaches, successfully capturing key structural characteristics of volatility, such as clustering effects. However, limitations such as restrictive distributional assumptions and limited capability in modeling complex nonlinear relationships have motivated the growing adoption of machine learning and, in particular, deep learning models (Di Persio et al., 2023; Koo & Kim, 2022; Mademlis & Dritsakis, 2021; Zhao et al., 2024). Deep learning models, by leveraging complex network architectures, are capable of extracting high-level and nonlinear features from large-scale and noisy datasets. In recent years, an increasing number of studies have explored hybrid approaches that combine GARCH-type models with neural networks, Long Short-Term Memory (LSTM) networks, and other deep learning architectures (Dessie et al., 2025; Kumar et al., 2025; Koo & Kim, 2022; Manogna et al., 2025). By exploiting the nonlinear modeling capacity of deep learning, these approaches aim to improve forecasting accuracy and enhance robustness under varying market conditions.

Alongside model hybridization, the role of data preprocessing and feature extraction has received growing attention. In this regard, signal transformation techniques, such as wavelet transforms, enable the decomposition of financial signals into joint time–frequency components. This capability is particularly valuable for analyzing the multi-scale structure of financial data and for effectively handling non-stationarity in time series. Despite these advances, a review of the existing literature indicates that important gaps remain in the design of volatility forecasting frameworks, particularly with respect to the development of hybrid architectures that can simultaneously and synergistically integrate multi-scale preprocessing, statistical modeling, and deep learning. To address these shortcomings, this study proposes a novel, multi-stage hybrid framework for financial volatility forecasting. The proposed framework begins with wavelet-based feature extraction to manage non-stationarity and capture multi-scale information. These features are then combined with models from the GARCH family and deep learning architectures to identify complex latent patterns in financial data. Finally, PSO is employed to fine-tune model parameters. To evaluate the effectiveness of the proposed framework, a series of backtesting experiments are conducted under realistic trading scenarios, demonstrating substantial improvements in both forecasting accuracy and trading performance compared to baseline models. Building upon prior studies, such as Di Persio et al. (2023) and Koo and Kim (2022), this paper introduces a comprehensive hybrid framework whose primary contribution lies in its hierarchical and multi-stage design. Specifically, the proposed approach integrates three key analytical layers:

1. Preprocessing and Feature Extraction Layer: Unlike related studies that primarily rely on raw data or simple wavelet transformations, this research employs multi-scale Continuous Wavelet Transform (CWT) to simultaneously address non-stationarity and reveal rich time–frequency patterns.
2. Hybrid Modeling Layer (Statistical–Deep Learning): Rather than treating statistical models as competitors to deep learning models, the proposed approach utilizes the outputs of statistical models as complementary features for constructing deep nonlinear predictors. The outputs of statistical models, after wavelet transformation, are incorporated as inputs to deep learning architectures.
3. Integrated Optimization and Performance Evaluation Layer: Given the high dimensionality and complexity of the proposed framework, particle swarm optimization is used for efficient parameter tuning. Model evaluation is performed not only using conventional error metrics but also through the backtesting of three trading strategies under diverse market conditions.

Together, these three layers form a unified framework that systematically overcomes the limitations of standalone statistical or neural models, leading to improved forecasting accuracy and superior financial performance.

The remainder of this paper is organized as follows. Section 2 reviews the theoretical and technical background required to understand the proposed framework. Section 3 provides a comprehensive review of related work and positions the present study within the existing literature. Section 4 describes the experimental setup and model inputs in detail, while Section 5 evaluates the performance of the proposed models in volatility forecasting and trading strategy backtesting. Finally, Section 6 concludes the paper and discusses implications and directions for future research.

2) Literature Review

Predicting volatility in financial markets is considered one of the most challenging tasks in economics and data science. Volatility, defined as a measure of dispersion or uncertainty in asset returns, plays a crucial role in financial decision-making. However, the inherent characteristics of financial data, such as noise, non-stationarity, long-memory effects, and volatility clustering, make accurate forecasting highly complex. To address these challenges, numerous methods have been proposed for estimating and modeling volatility, which are introduced and contextualized in this section.

2-1) Volatility Estimation Metrics

The first step in forecasting is selecting an appropriate metric to quantify volatility. Several approaches have been proposed, utilizing features such as opening, high, low, and closing prices. Seven key methods for volatility estimation are summarized below. In the following formulas, (H_t), (O_t), (C_t), and (L_t) denote the high, open, close, and low prices, respectively, and (n) represents the number of observed days.

- Historical Volatility: One of the most commonly used measures is historical volatility (HV), which is calculated based on the standard deviation of logarithmic returns over a fixed time window (Schwert, 1999). Equation (1) presents the formula for this estimator, where (r_t) denotes the logarithmic return at day (t), defined in Equation (2) and (μ) represents the mean return. This approach relies solely on closing prices and therefore cannot capture intraday volatility.

$$HV = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(r_t - \mu)^2} \tag{1}$$

$$r_t = ln\left(\frac{C_t}{C_{t-1}}\right) \quad (2)$$

- Close-to-Close Volatility: The close-to-close estimator (CCV) is based on daily logarithmic returns computed from the open and close prices of each day (Stoll & Whaley, 1990). Similar to Equation (3), it attempts to partially control for drift effects.

$$CCV = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(rc_t - \mu)^2} \quad (3)$$

- The key difference lies in the computation of (rc_t), defined as:

$$rc_t = ln\left(\frac{C_t}{O_t}\right) + ln\left(\frac{O_t}{C_{t-1}}\right) \quad (4)$$

- Parkinson Volatility: The Parkinson Estimator (PV) is a classic estimator based on the daily price range (difference between high and low) (Parkinson, 1980). As shown in Equation (5), by incorporating intraday information, it generally provides higher accuracy than close-to-close estimators. However, due to the assumption of zero drift and neglecting opening jumps, PV may underestimate volatility, particularly in turbulent markets.

$$PV = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\frac{1}{4ln\,(2)}\,.\left(ln\,ln\left(\frac{H_t}{O_t}\right) - ln\,ln\left(\frac{L_t}{O_t}\right)\right)^2} \quad (5)$$

- Garman-Klass Volatility: The Garman-Klass estimator (GKV) utilizes all available daily price information (Garman & Klass, 1980). As expressed in Equation (6), it is more comprehensive than the Parkinson estimator and provides a more dynamic measure of volatility using both opening and closing prices. Nevertheless, it may still underestimate volatility in volatile markets due to neglecting overnight jumps.

$$GKV = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\frac{1}{2}(ln\,(\frac{H_t}{O_t}))^2 - (2ln\,(2) - 1) \times (ln\,(\frac{C_t}{O_t}))^2} \quad (6)$$

- Rogers-Satchell Volatility: The Rogers-Satchell estimator (RSV) was developed to address the limitations of Parkinson and Garman-Klass estimators (Rogers & Satchell, 1991). It incorporates all daily price information and accounts for non-zero drift, offering higher accuracy in trending markets. However, by ignoring opening jumps, it may still underestimate actual volatility in certain cases.

$$RSV = \sqrt{\frac{1}{n}\sum_{t=1}^{n}((ln\left(\frac{H_t}{O_t}\right) \times ln\left(\frac{H_t}{C_t}\right)) + (ln\left(\frac{L_t}{O_t}\right) \times ln\left(\frac{L_t}{C_t}\right)))} \quad (7)$$

- Yang-Zhang Volatility: Yang-Zhang volatility (YZV) is a hybrid estimator designed to provide an unbiased, drift-independent measure while accounting for opening jumps. It delivers higher accuracy in capturing intraday volatility compared to previous estimators (Yang & Zhang, 2000). Equation (8) presents its computation, where (V_{OC}) and (V_{CO}) denote open-to-close and close-to-open variances, respectively; (VRS) is the Rogers-Satchell volatility, and (k) is a window-length-dependent weighting factor.

$$YZV = \sqrt{V_{CO} + (k\,.V_{OC}) + ((1-k)\,.V_{RS})} \quad (9)$$

$$V_{CO} = \frac{1}{n}\sum_{t=1}^{n}(ln\,(\frac{O_t}{C_{t-1}}) - \frac{1}{n}\sum_{t=1}^{n} ln\,(\frac{O_t}{C_{t-1}}))^2$$

$$V_{OC} = \frac{1}{n}\sum_{t=1}^{n}(ln\,(\frac{C_t}{O_t}) - \frac{1}{n}\sum_{t=1}^{n} ln\,(\frac{C_t}{O_t}))^2 \quad (10)$$

- Garman-Klass–Yang-Zhang Volatility: The Garman-Klass-Yang-Zhang volatility estimator (GKYZV) was developed to combine the advantages of the Garman-Klass and Yang-Zhang estimators, thereby capturing both intraday and overnight volatility simultaneously (Fałdziński & Osińska, 2016). Equation (11) presents the formal definition of this estimator. By incorporating overnight jumps and utilizing all available price information, the GKYZV provides higher accuracy compared to traditional estimators. However, the assumption of zero drift can still lead to volatility overestimation under certain market conditions.

$$GKYZV = \sqrt{\frac{1}{n}\sum_{t=1}^{n}\frac{1}{2}(ln\,(\frac{O_t}{C_{t-1}}))^2 + \frac{1}{2}(ln\,(\frac{H_t}{L_t}))^2 - (2ln\,(2)-1)\,.\,(ln\,(\frac{C_t}{O_t}))^2} \quad (11)$$

It is worth noting that historical volatility is generally adopted as the default estimator in related studies. However, each of the aforementioned formulas serves specific purposes in market analysts' decision-making processes.

2-2) Traditional Statistical Models

Conditional heteroscedasticity models are among the primary tools for modeling volatility in financial markets. A foundational model in this domain is the GARCH(p, q) model, proposed by Bollerslev (1986). Compared to the ARCH model (Engle, 1982), which models variance solely based on past residuals, GARCH offers greater flexibility by expressing the conditional variance as a combination of past squared residuals and its own lagged values. The GARCH model is described in Equation (12):

$$\sigma_t^2 = \omega + \sum_{i=1}^{q}\alpha_i\epsilon_{t-i}^2 + \sum_{j=1}^{p}\beta_j\sigma_{t-j}^2 \quad (12)$$

where (σ_t^2) denotes the conditional variance at time (t), (ω) is the baseline variance, (α_i) captures the effect of past residuals, (β_j) reflects the persistence of prior conditional variances, and (ϵ_t) is the error term represented in (Equation 13):

$$\epsilon_t = \sqrt{\sigma_t z_t} \quad (13)$$

Extending the framework of conditional variance models, the FIGARCH model, introduced by Baillie et al. (1996), was developed to capture long-memory effects in volatility. The defining feature of FIGARCH is the substitution of the integer differencing operator in GARCH with a fractional differencing operator, allowing the conditional variance to reflect persistent long-term behavior (Equation 14):

$$\sigma_t^2 = \omega + \left(1 - \sum_{i=1}^{q}\beta_i L^i\right)^{-1}\left(1-(1-L)^d\right)\left(\alpha_0 + \sum_{j=1}^{p}\alpha_i\epsilon_{t-j}^2\right) \quad (14)$$

$$L^i = \sigma_{t-i} \quad (15)$$

Here, d (0 < d < 1) represents the long-memory parameter, and $(1-L)^d$ functions as a long-memory filter (Equation 15). FIGARCH is particularly effective at modeling volatility over longer horizons, capturing partial and long-term persistence in financial time series.

In addition to these core models, several extensions have been developed to accommodate more complex market behaviors. EGARCH (Nelson, 1991) and APARCH (Laurent, 2004) allow for modeling asymmetric effects and various irregularities in financial time series, while HARCH (Bollerslev et al., 1992) employs a multi-frequency approach to analyze volatility across different temporal horizons. These models collectively provide a robust framework for capturing both short-term fluctuations and long-term dependencies in financial market volatility. With the advent of large and complex datasets, deep learning models have emerged as complementary or alternative approaches to traditional statistical methods. Architectures such as LSTM networks (Hochreiter & Schmidhuber, 1997), Gated Recurrent Units (GRU) (Chung et al., 2014), and Bidirectional LSTMs (BiLSTM) (Graves et al., 2005) have gained significant attention due to their ability to learn long-term temporal dependencies, particularly in sequential data (Olubusola et al., 2024). Echo State Networks (ESN) (Jaeger, 2007), with their randomly initialized reservoir structure, provide computationally efficient modeling of dynamic systems. On the other hand, Bayesian Neural Networks (BNN) (Bate et al., 1998) and Quantile Regression Neural Networks (QRNN) (Zhang et al., 1998) enable probabilistic and multi-quantile forecasting, allowing analysts to assess uncertainty. Although these models require larger datasets and more complex tuning, they offer substantial advantages in identifying nonlinear and hidden relationships in financial data (Zhao et al., 2024).

Previous studies in the field of financial market volatility forecasting can be broadly classified into three main categories: approaches based on statistical models, deep learning models, and hybrid approaches, which have attracted increasing attention in recent years.

2-3) Statistical Model–Based Approaches

Statistical model–based approaches have mainly focused on the GARCH family of models and their extensions. Studies such as Gong and Lin (2018) and Kim et al. (2021) employed these models to forecast volatility in stock markets and cryptocurrency markets, sometimes incorporating leverage effects or using information criteria, such as AIC for model selection. Some studies have also investigated stochastic volatility models or combined statistical models with signal transformation techniques, such as wavelet transform, to improve forecasting accuracy for assets, such as leading U.S. stocks (Rubio et al., 2023). Moreover, more recent models, including MF2-GARCH, have been introduced to flexibly model long-term volatility cycles, particularly during highly turbulent periods such as the COVID-19 crisis (Conrad & Engle, 2025). Nevertheless, these approaches often face limitations when applied independently, particularly in capturing the full complexity and nonlinear patterns of financial markets. A review of prior studies indicates that statistical models provide a solid theoretical foundation for volatility forecasting and are commonly used as benchmark methods in the literature.

2-4) Deep Learning–Based Approaches

In response to the limitations of statistical methods, deep learning models have received growing attention. Early studies in this area, such as Hamid and Iqbal (2004), employed simple neural networks to forecast S&P 500 volatility and compared their performance with implied volatility measures. Reinforcement-based approaches have also been used to forecast gold price volatility using macroeconomic variables (Pierdzioch et al., 2016). In addition, the impact of incorporating unstructured data, such as news, alongside price data to improve stock volatility forecasting has been examined using multimodal deep learning models (Sardelich & Manandhar, 2018).

Recurrent models, such as LSTM, due to their ability to model long-term dependencies, have been applied to forecast the volatility of Chinese stock indices (Yu & Li, 2018), often in comparison with traditional statistical models. However, several studies have shown that deep learning models do not always outperform statistical models or even simple linear regression, particularly when evaluated on out-of-sample test data (Miura et al., 2019) or in highly unstable markets such as Bitcoin (Shen et al., 2021). Comparisons between LSTM and SVR models for different indices have also yielded mixed results depending on the asset under consideration (Liu, 2019). Studies such as Mahajan et al. (2022) reported a slight superiority of statistical models under certain conditions for the NIFTY50 index, while

Zhang et al. (2023), using an XGBoost-based approach, demonstrated the superiority of neural networks in forecasting intraday volatility of the S&P 500. Furthermore, reinforcement neural network models have shown improved performance over baseline models in forecasting S&P 500 volatility using macro-financial variables (Ciner, 2025). Overall, the review of previous studies suggests that deep learning–based approaches perform better than statistical methods in capturing and forecasting nonlinear volatility relationships under complex market conditions and with more complex input data. However, the application of these models requires larger datasets, heavier computational resources, and higher implementation complexity.

2-5) Hybrid Approaches

Given the advantages and limitations of both statistical and deep learning methods, several previous studies have proposed hybrid approaches that combine these two paradigms in order to mitigate their respective weaknesses. Hybrid models based on statistical and deep learning techniques have gained substantial popularity, as they aim to integrate the theoretical foundations of statistical models with the powerful pattern-learning capabilities of deep learning models. Studies such as Hajizadeh et al. (2012) combined EGARCH models with artificial neural networks to forecast S&P 500 volatility. Similarly, Kristjanpoller and Minutolo (2015), Kristjanpoller et al. (2014), and Kristjanpoller and Minutolo (2016) proposed ANN–GARCH hybrid models for forecasting volatility in Latin American stock markets, gold prices, and oil prices, demonstrating that such combinations improve forecasting accuracy compared to standalone GARCH models. Lu et al. (2016) compared the performance of ANN–GARCH hybrid models in the Chinese energy market and showed the superiority of the EGARCH–ANN model. Kim and Won (2018) combined various GARCH models (including EGARCH and EWMA) with LSTM and deep feedforward networks to forecast KOSPI200 volatility, concluding that combining multiple GARCH models with neural networks yields better results.

Stacked models, in which the outputs of several base models (such as GARCH or ANN) are used as inputs to a meta-model, were examined by Ramos-Pérez et al. (2019) for forecasting S&P 500 volatility and were shown to outperform simpler hybrid models. In cryptocurrency markets, Seo and Kim (2020) combined GARCH models with ANN and HONN and incorporated Google Trends data and the VIX index to improve forecasting accuracy. In the field of precious metals, Kristjanpoller and Hernández (2017) used GARCH-based forecasts as inputs to neural networks to predict the volatility of gold, silver, and copper. Other studies, such as Mademlis and Dritsakis (2021), compared multiple hybrid models (ANN–EGARCH, ANN–ARMA–GARCH, and ARMA–GARCH–ANN) for the FTSE index and reported the superiority of the ARMA–GARCH–ANN model. Trierweiler Ribeiro et al. (2021) proposed a novel HAR–PSO–ESN hybrid model for forecasting NASDAQ stock volatility, which demonstrated favorable performance. Finally, recent studies such as García-Medina and Aguayo-Moreno (2023), Amirshahi and Lahmiri (2023), and Ulu (2025) have focused on combining GARCH models with LSTM architectures for high-frequency cryptocurrency volatility forecasting. Di Persio et al. (2023) investigated hybrid architectures combining BiLSTM and GRU networks with GARCH models for global stock markets and portfolio risk management. In addition, Hu et al. (2020) and Siddiraju and Hasan (2025) focused on forecasting copper price volatility and stock market volatility during crisis periods, respectively, using LSTM–GARCH hybrid models.

A review of prior studies indicates that despite significant progress, particularly in hybrid approaches, there remain opportunities for further improvement. Many studies have not fully exploited the potential of advanced preprocessing techniques, such as wavelet transform for multi-scale feature extraction and handling non-stationarity. Moreover, model evaluation has predominantly focused on numerical error metrics, while practical applications in trading strategies have received less attention. In addition, despite the widespread use of recurrent models such as LSTM and GRU, comprehensive comparisons with a broader range of deep learning architectures and the consideration of diverse volatility estimators remain limited. Motivated by these gaps, this study proposes a novel hybrid framework that integrates wavelet transform, a diverse set of statistical and deep learning models, and comprehensive evaluation under realistic scenarios.

3) Methodology

To forecast stock market volatility, this study proposes a novel hybrid framework, illustrated in detail in Figure 1. The primary objective of this framework is to leverage the synergy between classical statistical models, diverse deep learning architectures, and intelligent parameter optimization to achieve more accurate and reliable volatility predictions. The overall process of the framework, as depicted in Figure 1, consists of the following stages:

1. Data Collection and Preliminary Statistical Modeling: Initially, a comprehensive set of market-related data is collected and preprocessed. This dataset includes market indices, currency exchange rates, economic indicators, Google Trends data, and energy and metals market data, alongside a time series of target volatilities (calculated using various estimators) and preliminary forecasts derived from the GARCH-family models. As shown in the upper part of Figure 1 under "Features for Wavelet Transform," these data serve as inputs for the subsequent stage.
2. Multiscale Feature Extraction via Wavelet Transform: The prepared data are then subjected to wavelet transformation to extract multiscale features. This stage, labeled "Wavelet Transform" in Figure 1, decomposes the signals and identifies hidden patterns across different time-frequency scales, which is particularly effective in addressing the non-stationarity of financial data.
3. Correlation Analysis and Feature Selection: After wavelet-based feature extraction, a comprehensive correlation analysis is performed (illustrated in Figure 1 as "Correlation Analysis") to identify and select the most informative features for volatility forecasting. This step reduces dimensionality and focuses the modeling process on the most relevant information.
4. Volatility Forecasting with Hybrid Deep Learning Models: The selected features are subsequently fed into a set of hybrid deep learning models, shown in Figure 1 under "Deep Learning Models." These models include architectures suitable for time series analysis, as well as specialized networks such as ESN, BNN, and QRNN. All models are trained in combination with GARCH-based forecasts and wavelet features, and their hyperparameters are optimized using PSO.

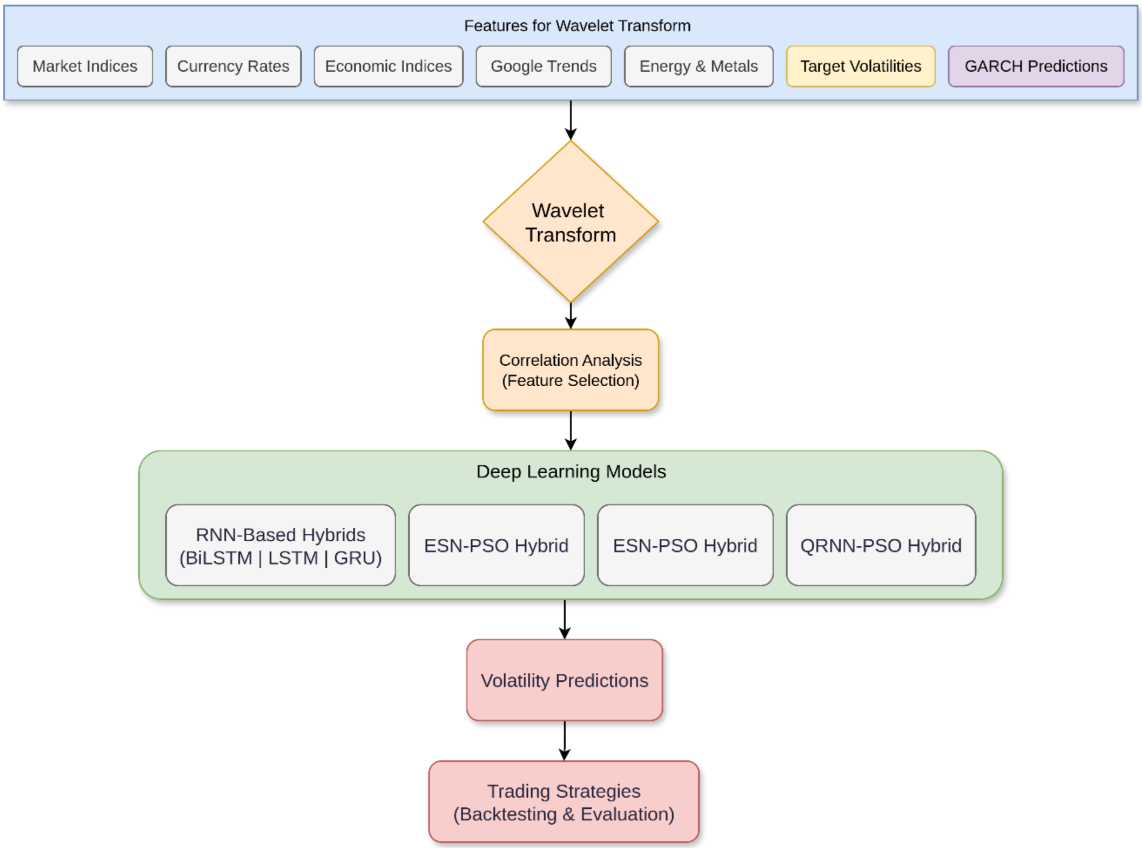

Figure 1. Proposed Multi-Stage Framework for Volatility Forecasting

5. Performance Evaluation and Practical Application: Finally, the volatility predictions obtained from the hybrid models (depicted as "Trading Strategies" in Figure 1) are evaluated both in terms of statistical accuracy and practical applicability. The latter is assessed through backtesting in various trading strategies, as illustrated in the lower section of Figure 1 under "Trading Strategies."

In the following sections, each of these stages is described in detail. In this study, to develop an accurate and robust model for forecasting S&P 500 volatility, a comprehensive and diverse dataset of financial and economic indicators was compiled. The S&P 500 index, representing the overall U.S. equity market, was chosen as the underlying asset due to its high liquidity and broad market coverage. The study period spans from January 2000 to August 2024, encompassing major market conditions including bull and bear markets, as well as significant events such as the 2008 financial crisis and the COVID-19 pandemic.

The collected data include daily prices and trading volumes of the S&P 500, complementary stock market indices such as the DJIA, currency exchange rates, macroeconomic indicators (e.g., GDP growth, unemployment rate, inflation, and interest rates), Google Trends data as proxies for public sentiment, and commodity prices such as oil and gold, serving as alternative assets during periods of market uncertainty. Additionally, the time series of realized volatility for the index, computed using seven established estimators (introduced in Section 2.1), were employed as the target variables for model training. Outputs from preliminary GARCH-family models were also included as inputs to the advanced hybrid models, particularly deep learning architectures. All collected time series were standardized to daily frequency and missing values were handled via interpolation or appropriate imputation methods where necessary. This comprehensive data preparation forms the foundation for subsequent analyses, enabling the design of multi-purpose, data-driven forecasting models.

Initial volatility forecasts for the S&P 500 were generated using a set of ARCH and GARCH-family models, implemented through a rolling window approach. In this approach, a 21-day window (corresponding to the number of trading days in a month) was used to estimate model parameters, and volatility was forecasted for the subsequent day. The window was then advanced by one day, and the process repeated. This method allows continuous updating of the models with new data and improves the modeling of dynamic market behaviors. The statistical models employed in this study include standard GARCH, EGARCH, HARCH, APARCH, and FIGARCH. Each model was selected to capture specific aspects of volatility, such as clustering, leverage effects, long-memory behavior, or multi-frequency decomposition. To enhance robustness, these models were estimated under various error distributions, including normal, Student's t, skewed, and skewed Student's t distributions, and considered asymmetric innovation lags (parameter (O) set to 0 or 1). For simplicity and to mitigate overfitting, the lag orders (p) and (q) in all GARCH models were set to 1. The resulting volatility forecasts from this diverse set of statistical models were subsequently used as inputs for the deep learning models.

To decompose financial signals into time-frequency components and extract multiscale features that enhance the understanding of market patterns and address the challenge of non-stationarity, wavelet transforms were employed. As illustrated in Figure 1, all collected input data were transformed using both Continuous Wavelet Transform (CWT) and Discrete Wavelet Transform (DWT):

1. Continuous Wavelet Transform (CWT): Using the CMor mother wavelet and a decomposition scale, ranging from 1 to 20, the energy of the wavelet coefficients at each scale was computed for each time series. These energies, representing the intensity of fluctuations at different frequencies, were used as input features for deep learning models, such as ESN and BNN, which leverage a continuous and detailed time-frequency representation.
2. Discrete Wavelet Transform (DWT): Using the Haar mother wavelet and decomposition up to three levels, approximation and detail coefficients were extracted. These coefficients, which capture both general trends and sudden changes across multiple scales, were reconstructed and aligned with the original series and used as input features for RNN-based deep learning models (e.g., LSTM and GRU). This approach preserves temporal structure while reducing computational complexity during multiscale feature extraction.

The multiscale features obtained from both types of wavelet transforms played a key role in enriching the input dataset and improving the predictive capabilities of the models in capturing complex volatility patterns.

3-1) Correlation Analysis and Feature Selection

Following multiscale feature extraction and preliminary GARCH forecasts, a comprehensive correlation analysis was conducted to identify and select the most informative features for final volatility prediction. The objective of this step was to reduce the dimensionality of the feature space and focus on information most strongly related to the target volatility time series (computed using seven different estimators). The following statistical methods were applied:

- Pearson and Spearman Correlation: Linear and rank correlations were computed between each candidate feature, including wavelet features, GARCH forecasts, and other market and economic variables, and each of the seven target volatility series. Features with an absolute correlation coefficient exceeding 0.5 were selected as initially relevant.
- Cross-Correlation Analysis: To account for potential time-lagged relationships, cross-correlation up to a maximum lag of 30 days was computed between features and target volatilities. Features with absolute cross-correlation values above 0.5 at significant lags were retained.

- Granger Causality Test: Conducted with a maximum lag of five periods and a significance level of 0.05, this test identified whether past values of a feature could help predict future volatility.

Features identified as relevant by at least one of these methods constituted the final set of input variables for the deep learning models. Detailed results, including correlations between continuous wavelet features and GARCH/FIGARCH forecasts under various distributions, as well as correlations between discrete wavelet features and macroeconomic indicators, exchange rates, and commodity prices, were used to guide the selection of features for each volatility estimator and deep learning model.

3-2) Hybrid Deep Learning Forecasting and Backtesting

The core of the proposed framework consists of a set of hybrid deep learning models that leverage the selected features from previous stages to forecast seven types of target volatilities. All deep learning models were trained using a rolling window approach (25 days of input data to forecast the next 5 days), the Adam optimizer, and a mean squared error loss function. The data were split into training (90%) and validation (10%) sets for the period from January 2000 to January 2022, and a test set covering the last two years. This diverse ensemble of hybrid models allows for the evaluation of various aspects of volatility forecasting, including point accuracy, uncertainty quantification, and conditional distribution prediction. The implemented architectures include:

- RNN-Based Hybrid Architectures: Five models combining LSTM, GRU, and BiLSTM were developed. These models used features from the DWT along with GARCH and FIGARCH outputs. Each architecture consists of two recurrent layers (with 48 and 16 neurons, respectively) and a dense output layer. Key training hyperparameters (batch size: 32, epochs: 80, learning rate: 0.01, dropout rate: 0.1) were set to allow comparison with benchmark results (Di Persio et al., 2023).
- Specialized Hybrid Architectures Optimized with PSO:
    - ESN Hybrid: Utilizes features from CWT, which align well with the reservoir computing structure of ESN. Key reservoir parameters (number of neurons, spectral radius, sparsity) were optimized using PSO.
    - BNN Hybrid: Designed to quantify predictive uncertainty, this model also leverages CWT features. Its architecture and training hyperparameters were optimized via PSO.
    - QRNN Hybrid: Provides probabilistic forecasts and was implemented on DWT features to predict the conditional distribution of volatility. QRNN hyperparameters were similarly optimized using PSO.

To assess the practical applicability of the volatility forecasts, three distinct trading strategies for the S&P 500 were designed and backtested over the test period. The risk-free rate was assumed to be 4% per annum for Sharpe ratio calculations. Each strategy was executed under two scenarios:

- Baseline: Trading decisions based solely on strategy signals (e.g., moving average crossovers) without incorporating volatility forecasts.
- Enhanced: Incorporating hybrid model volatility forecasts, specifically, the median forecast across the seven volatility estimators, to guide entry and exit decisions.

The strategies are as follows:

- Strategy 1: Based on 5-day and 20-day moving average crossovers.
    - Baseline: Buy when MA5 crosses above MA20; sell when MA5 crosses below MA20.
    - Enhanced: Buy condition same as baseline; sell condition requires both the crossover signal and the forecasted volatility exceeding the 50th percentile of the historical volatility distribution.
- Strategy 2: Based on 10-day and 50-day moving average crossovers.

    - Enhanced sell condition: Forecasted volatility exceeds a dynamic threshold (10-day moving average of volatility plus two standard deviations).
- Strategy 3: Based on 50-day and 200-day moving average crossovers.
    - Enhanced buy condition: Forecasted volatility is below the 60th percentile of historical volatility.
    - Enhanced sell condition: Forecasted volatility exceeds the 85th percentile of historical volatility.

The performance of these strategies was evaluated using metrics such as the Sharpe ratio, annualized net returns, and maximum drawdown.

4) Findings and Discussion

In this section, we present the results obtained from the proposed framework and compare them with the approach introduced by Di Persio et al. (2023). The analysis is conducted in two main parts. The first part focuses on evaluating the accuracy of volatility forecasts, while the second part examines the performance of various trading strategies using commonly used financial metrics. All forecasting results are based on the test dataset, which covers the last two years of the collected data, from June 1, 2022, to August 30, 2024. The baseline approach from Di Persio et al. (2023) is also evaluated over the same period. These experiments are conducted for all seven volatility estimators described in chapter 2. In the first part, forecasting accuracy is assessed using the Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Squared Error (MSE), and the results are compared with those reported in the baseline study (Di Persio et al., 2023). In the second part, the performance of the three trading strategies is evaluated using the Sharpe ratio, maximum drawdown, and annualized net returns. Since the baseline study does not include an analysis of trading strategies, this section also enables a practical comparison of the proposed framework's performance. Table 1 provides an overview of the abbreviations used for the different methods throughout this section, which facilitates readability and reference in the subsequent discussion.

4-1) Forecasting Results Using the Proposed Methods

We now present the forecasting results of market volatility obtained using the proposed methods listed in Table 1. Tables 2–4 provide a comprehensive comparison between the baseline approaches from Di Persio et al. (2023), including G-LSTM, G-GRU, and G-LSTM-GRU, and the proposed methods, highlighting the consistent superiority of the latter. Volatility forecasts for the S&P 500 index were generated using both the baseline and proposed methods over the test dataset, covering the last two years of data from June 1, 2022, to August 30, 2024. All results are reported using standard forecasting performance metrics commonly used in volatility prediction studies. The inclusion of statistical features, from GARCH and FIGARCH models, is expected to expand the informative feature space for the proposed methods compared to the baseline approaches. Additionally, the use of wavelet transforms to extract multi-scale features enhances stationarity and memory in the time series, further improving the quality of input features for volatility prediction.

Table 1. Abbreviations and Symbols Used in the Proposed Framework

Abbreviation	Full Name
GFWL	GARCH-FIGARCH-Wavelet-LSTM
GFWBL	GARCH-FIGARCH-Wavelet-BiLSTM
GFWG	GARCH-FIGARCH-Wavelet-GRU
GFWLG	GARCH-FIGARCH-Wavelet-LSTM-GRU
GFWBLG	GARCH-FIGARCH-Wavelet-BiLSTM-GRU

GFWPE	GARCH-FIGARCH-Wavelet-PSO-ESN
GFWPB	GARCH-FIGARCH-Wavelet-PSO-BNN
GFWPQ	GARCH-FIGARCH-Wavelet-PSO-QRNN

Table 2. Root Mean Square Error (RMSE) in Volatility Forecasting

Proposed Methods								Baselines			
GFWPQ	GFWPB	GFWPE	GFWBLG	GFWLG	GFWG	GFWBL	GFWL	G-LSTM-GRU	G-GRU	G-LSTM	Volatility Estimator
2.5×10^{-6}	1.7×10^{-4}	1.1×10^{-6}	**6.0×10^{-7}**	2.4×10^{-5}	1.1×10^{-5}	8.1×10^{-6}	7.2×10^{-6}	1.9×10^{-2}	1.7×10^{-2}	2.3×10^{-2}	Historical
2.7×10^{-6}	2.4×10^{-3}	**7.0×10^{-7}**	5.1×10^{-6}	7.6×10^{-5}	1.5×10^{-4}	1.7×10^{-5}	1.9×10^{-6}	2.2×10^{-2}	1.9×10^{-2}	2.1×10^{-2}	Close-to-Close
7.0×10^{-8}	1.1×10^{-3}	5.1×10^{-7}	4.7×10^{-6}	5.8×10^{-7}	2.8×10^{-6}	1.0×10^{-5}	2.6×10^{-6}	6.5×10^{-2}	5.9×10^{-2}	6.3×10^{-2}	Parkinson
7.5×10^{-7}	2.0×10^{-3}	9.6×10^{-7}	1.9×10^{-6}	1.6×10^{-5}	1.1×10^{-4}	2.9×10^{-6}	1.3×10^{-6}	7.6×10^{-2}	4.7×10^{-2}	6.1×10^{-2}	Garman-Klass
8.8×10^{-6}	5.8×10^{-4}	8.8×10^{-6}	7.7×10^{-6}	2.3×10^{-4}	1.1×10^{-5}	9.0×10^{-6}	1.6×10^{-4}	1.1×10^{-1}	7.5×10^{-2}	8.5×10^{-2}	Yang-Zhang
7.5×10^{-7}	4.1×10^{-4}	4.6×10^{-6}	2.1×10^{-6}	2.7×10^{-6}	1.3×10^{-6}	7.2×10^{-6}	3.8×10^{-6}	9.4×10^{-2}	8.4×10^{-2}	1.0×10^{-1}	G-K-Y-Z
5.7×10^{-6}	4.3×10^{-3}	**4.0×10^{-8}**	2.1×10^{-7}	5.3×10^{-6}	2.6×10^{-6}	3.8×10^{-6}	2.3×10^{-6}	1.1×10^{-2}	8.4×10^{-2}	9.5×10^{-2}	Rogers-Satchell

Table 3. Mean Absolute Error (MAE) in Volatility Forecasting

Proposed Methods								Baselines			
GFWPQ	GFWPB	GFWPE	GFWBLG	GFWLG	GFWG	GFWBL	GFWL	G-LSTM-GRU	G-GRU	G-LSTM	Volatility Estimator
1.3×10^{-6}	1.1×10^{-4}	6.0×10^{-7}	**3.1×10^{-7}**	1.4×10^{-5}	8.2×10^{-6}	5.7×10^{-6}	5.4×10^{-6}	1.2×10^{-2}	1.0×10^{-2}	1.4×10^{-2}	Historical
1.4×10^{-6}	1.6×10^{-3}	**4.9×10^{-7}**	3.7×10^{-6}	3.5×10^{-5}	1.1×10^{-4}	9.6×10^{-6}	1.5×10^{-6}	1.4×10^{-2}	1.0×10^{-2}	1.2×10^{-2}	Close-to-Close
4.6×10^{-8}	6.1×10^{-4}	3.3×10^{-7}	2.7×10^{-6}	4.2×10^{-7}	2.0×10^{-6}	8.2×10^{-6}	1.8×10^{-6}	4.3×10^{-2}	3.6×10^{-2}	3.9×10^{-2}	Parkinson
3.9×10^{-7}	1.2×10^{-3}	6.2×10^{-7}	1.2×10^{-6}	1.2×10^{-5}	9.2×10^{-5}	2.1×10^{-6}	1.0×10^{-6}	4.9×10^{-2}	3.4×10^{-2}	4.3×10^{-2}	Garman-Klass
6.6×10^{-6}	4.1×10^{-4}	6.6×10^{-6}	**5.4×10^{-6}**	1.3×10^{-4}	7.3×10^{-6}	6.4×10^{-6}	8.6×10^{-5}	8.6×10^{-2}	5.3×10^{-2}	5.8×10^{-2}	Yang-Zhang
4.9×10^{-7}	2.7×10^{-4}	3.3×10^{-6}	1.5×10^{-6}	2.0×10^{-6}	9.7×10^{-7}	6.1×10^{-6}	2.5×10^{-6}	7.6×10^{-2}	5.9×10^{-2}	7.4×10^{-2}	G-K-Y-Z
3.9×10^{-6}	2.9×10^{-3}	**2.7×10^{-8}**	1.7×10^{-7}	3.8×10^{-6}	1.7×10^{-6}	2.5×10^{-6}	1.9×10^{-6}	9.3×10^{-2}	5.5×10^{-2}	6.1×10^{-2}	Rogers-Satchell

Table 4. Mean Squared Error (MSE) in Volatility Forecasting

Proposed Methods								Baselines			
GFWPQ	GFWPB	GFWPE	GFWBLG	GFWLG	GFWG	GFWBL	GFWL	G-LSTM-GRU	G-GRU	G-LSTM	Volatility Estimator
1.7×10^{-4}	4.9×10^{-3}	1.1×10^{-4}	**9.8×10^{-5}**	3.4×10^{-4}	2.6×10^{-4}	2.4×10^{-4}	2.2×10^{-4}	5.9×10^{-1}	5.7×10^{-1}	6.5×10^{-1}	Historical
1.5×10^{-4}	6.3×10^{-3}	**8.3×10^{-5}**	2.1×10^{-4}	9.1×10^{-4}	3.3×10^{-3}	2.2×10^{-4}	1.9×10^{-4}	7.5×10^{-1}	6.1×10^{-1}	6.9×10^{-1}	Close-to-Close
6.1×10^{-6}	2.3×10^{-3}	7.6×10^{-5}	3.2×10^{-4}	1.1×10^{-4}	3.1×10^{-4}	2.7×10^{-4}	2.9×10^{-4}	9.7×10^{-1}	8.2×10^{-1}	9.8×10^{-1}	Parkinson
7.0×10^{-5}	3.9×10^{-3}	1.1×10^{-4}	2.4×10^{-4}	7.1×10^{-4}	6.3×10^{-4}	2.3×10^{-4}	2.2×10^{-4}	1.02	7.2×10^{-1}	8.9×10^{-1}	Garman-Klass
3.3×10^{-4}	1.7×10^{-3}	3.5×10^{-4}	**2.3×10^{-4}**	9.4×10^{-4}	3.4×10^{-4}	3.9×10^{-4}	6.7×10^{-4}	1.48	9.9×10^{-1}	1.12	Yang-Zhang
1.1×10^{-4}	2.2×10^{-3}	1.8×10^{-4}	2.0×10^{-4}	2.2×10^{-4}	1.3×10^{-4}	2.4×10^{-4}	2.9×10^{-4}	1.37	1.16	1.41	G-K-Y-Z
3.2×10^{-4}	4.5×10^{-3}	**2.0×10^{-5}**	1.2×10^{-4}	5.7×10^{-4}	2.4×10^{-4}	2.8×10^{-4}	3.3×10^{-4}	1.56	1.02	1.13	Rogers-Satchell

As observed in Table 2, all proposed methods outperform the baseline approaches, consistent with our expectations. Notably, the GFWPQ method, when combined with four out of the seven volatility estimators, delivers the best performance among all forecasting approaches, making it the prime candidate for optimal performance in this experiment. The lowest error observed (7.0×10^{-8}) corresponds to the Parkinson estimator using the GFWPQ method, demonstrating its potential for highly accurate volatility prediction. In contrast, the relatively uniform performance of the baseline methods indicates potential underfitting, which may result from lower-quality time series features, simpler model structures, or a limited feature space due to the absence of statistical features. Results reported in Table 3 are consistent with those in Table 2, further confirming the superiority of the proposed approaches. Table 4, which reports Mean Squared Error (MSE), also demonstrates significant performance improvements over the baseline methods (Di Persio et al., 2023), with the GFWPQ, GFWPE, and GFWBLG methods consistently showing the best performance across different combinations.

These findings suggest that deep learning models benefit substantially from the inclusion of statistical features and the application of wavelet transforms to the time series, regardless of the specific volatility estimator used. Moreover, the wider performance range observed for the proposed methods indicates that these enhancements lead to better fitting and more robust deep learning models. In summary, the key innovation of this study lies not in a single model but in the design of a systematic, multi-stage framework that integrates three distinct analytical and computational strengths: (i) the capability of statistical models to capture conditional variance structures, (ii) the ability of wavelet transforms to extract multi-scale features and handle nonstationarity, and (iii) the high capacity of deep learning architectures to detect complex nonlinear patterns. The quantitative results presented in this section clearly demonstrate the success of this synergy. The superior performance of the combined models, particularly when used with various volatility estimators, indicates that the framework leverages the intelligent integration of multiple sources of information to provide a more comprehensive understanding of market dynamics, moving beyond conventional hybrid approaches.

4-2) Backtesting Results

Volatility forecasts are only practically valuable if they enhance trading decision-making. High statistical accuracy translates into actionable benefits when it improves risk-adjusted returns or mitigates losses and drawdowns. Accordingly, the three trading strategies introduced in Section 4.4 were backtested under three scenarios:

1. Baseline: No volatility estimates are used.
2. Ideal: Realized volatility is used as complete information.
3. Forecast-based: Volatility is predicted using the proposed hybrid models.

All backtests were conducted using daily S&P500 data from January 2018 to December 2023. Performance was evaluated with three common financial metrics: the Sharpe ratio (risk-adjusted returns), net annualized returns, and maximum drawdown (as a measure of downside risk and capital preservation).

Strategy 1, based on short- and long-term moving average crossovers, triggers entries upon bullish crossovers and exits when volatility exceeds a certain threshold. As illustrated in Figure 2, the Sharpe ratio in the baseline scenario is consistently lower than the other two scenarios, indicating that even basic incorporation of volatility information improves risk-return performance. Using realized volatility achieves the best performance in most cases, confirming the value of accurate volatility estimates in trading decisions. Forecasts from the proposed hybrid models generally result in Sharpe ratios close to those obtained with realized volatility. However, differences are observed depending on the volatility estimator; for example, using historical volatility, GFWL outperforms other models, while for the Parkinson estimator, GFWBL provides superior results.

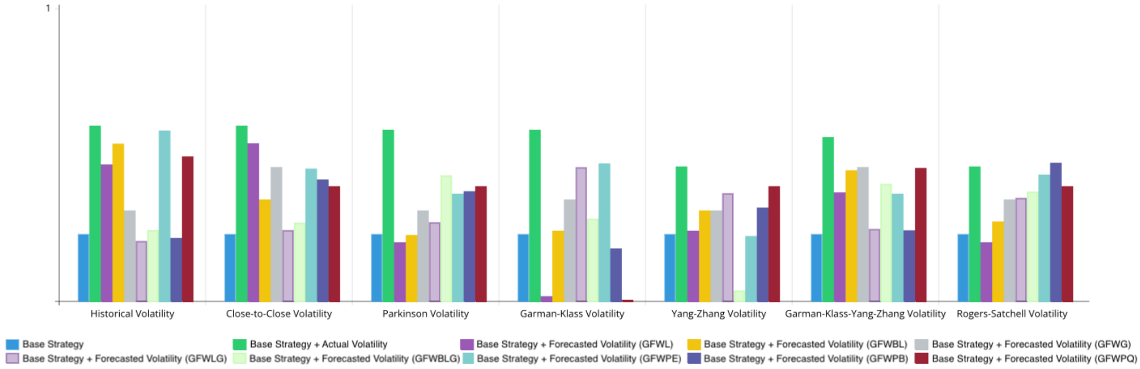

Figure 2. Performance of the First Trading Strategy Based on Short-Term Moving Average Crossover

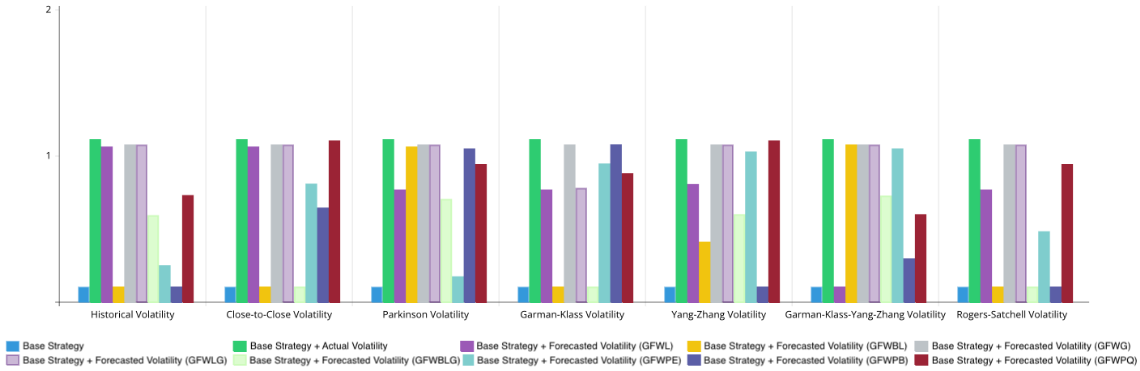

Figure 3. Performance of the Second Trading Strategy with Dynamic Stop-Loss Logic

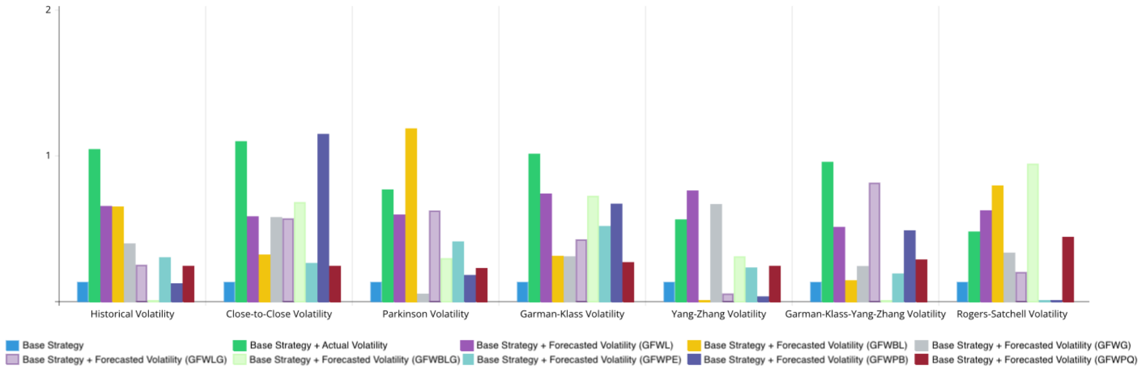

Figure 4. Performance of the Third Trading Strategy with Fully Volatility-Conditioned Entry and Exit

These variations emphasize that no single architecture universally dominates due to market dynamics and scenario-specific conditions. Strategy 2 employs a simpler entry logic but a dynamic exit mechanism tied to forecasted volatility. Analyzing the Sharpe ratios in Figure 3 confirms trends observed in Strategy 1: The baseline scenario underperforms, the ideal scenario performs best, and the hybrid forecasts provide reliable and generally near-ideal performance. Interestingly, the range of performance across different models is narrower in this strategy, likely due to the more conservative risk control logic, which reduces sensitivity to forecast errors.

Strategy 3 incorporates two independent thresholds for entry and exit, providing precise risk management by trading primarily during low-volatility periods and closing positions when volatility rises. As shown in Figure 4, the impact of forecasted volatility is even more pronounced here. Hybrid models, particularly GFWPE combined with multiple volatility estimators, demonstrate stable and robust performance. Conversely, models with lower predictive accuracy exhibit noticeable performance declines, highlighting the importance of architecture selection and model tuning for practical trading applications. Overall, the backtesting results demonstrate that leveraging volatility information, especially forecasted by the proposed hybrid models, significantly improves trading performance. Across all scenarios, the inclusion of volatility, whether realized or predicted, enhances metrics such as the Sharpe ratio. The hybrid models often approach the ideal performance scenario, validating their accuracy and practical applicability. In conclusion, the proposed framework not only excels in statistical volatility forecasting but also enhances decision-making and risk management in practical trading environments.

5) Conclusion

This study proposed a multi-stage framework aimed at delivering a novel and comprehensive hybrid approach for accurate stock market volatility forecasting. The proposed framework leverages the synergistic integration of wavelet transforms for multi-scale feature extraction, statistical models from the GARCH family for capturing well-known volatility characteristics, and a diverse range of advanced deep learning architectures, including LSTM, BiLSTM, GRU, ESN, BNN, and QRNN, together with parameter optimization via PSO. The S&P 500 index was used as the underlying asset over the period from 2000 to 2024, incorporating a wide array of input data, including market indices, exchange rates, macroeconomic variables, Google Trends data, and energy and metal prices. Empirical results demonstrated that the proposed hybrid framework, particularly models combining wavelet-transformed features, GARCH forecasts, and PSO-optimized deep learning architectures, significantly reduced volatility forecasting errors compared to baseline models and prior approaches, including the benchmark study by Di Persio et al. (2023). These improvements were observed across multiple error metrics, including MSE, RMSE, and MAE, for a wide range of volatility estimators.

Beyond statistical accuracy, the practical applicability of the proposed framework was assessed using three distinct trading strategies. Backtesting results revealed that incorporating predicted volatility into trading decisions led to substantial improvements in strategy performance. Specifically, in the most successful strategy, net returns increased threefold, while the maximum drawdown decreased by up to 39% compared to the scenario without volatility forecasting. These findings underscore the practical value of accurate volatility predictions and highlight the potential of the proposed framework to assist market participants in enhancing profitability while mitigating investment risk. Overall, this study demonstrates that combining advanced preprocessing techniques, classical statistical models, and modern deep learning approaches can lead to more robust and precise solutions for the challenging problem of financial volatility forecasting. Despite the promising outcomes, several avenues for future research and further development of the proposed framework remain. One important direction is the expansion of datasets and input features. Examining the framework's performance using higher-frequency intraday data, such as minute-level or tick data, could provide new insights into short-term volatility dynamics. Incorporating features derived from sentiment analysis of financial news, social media, and economic reports could further improve predictive accuracy. Additionally, exploring alternative data sources, such as satellite imagery or supply chain information, may help models identify less-known factors influencing volatility. Another research direction involves the advancement of deep learning models within the framework. Graph neural networks, for instance, could effectively capture complex inter-asset relationships and volatility spillovers across markets. Opportunities also exist in model optimization and interpretability. Exploring alternative metaheuristic optimization algorithms for tuning complex hybrid models could lead to improved forecasting performance. Simultaneously, applying interpretable AI techniques can provide deeper understanding of deep learning decision-making processes and identify the most influential features for volatility forecasting, thereby increasing trust in the models and extracting actionable knowledge. Pursuing these research directions can contribute to the development of more accurate and reliable tools for analyzing and forecasting financial market volatility, ultimately supporting market participants in better risk management and more informed investment decisions.

References

Amirshahi, B., & Lahmiri, S. (2023). Hybrid deep learning and GARCH-family models for forecasting volatility of cryptocurrencies. *Machine Learning with Applications*, *12*(1), 100465. https://doi.org/10.1016/j.mlwa.2023.100465

Andrés García-Medina, & Aguayo-Moreno, E. (2023). LSTM–GARCH hybrid model for the prediction of volatility in cryptocurrency portfolios. *Computational Economics*, *1*(1). https://doi.org/10.1007/s10614-023-10373-8

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *74*(1), 3-30. https://doi.org/10.1016/S0304-4076(95)01749-6

Bate, A., Lindquist, M., Edwards, I. R., Olsson, S., Orre, R., Lansner, A., & De Freitas, R. M. (1998). A Bayesian neural network method for adverse drug reaction signal generation. *European Journal of Clinical Pharmacology*, *54*, 315-321. https://doi.org/10.1007/s002280050466

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

Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, *52*(1-2), 5-59. https://doi.org/10.1016/0304-4076(92)90064-X

Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*. https://doi.org/10.48550/arXiv.1412.3555

Ciner, C. (2025). Forecasting the aggregate market volatility by boosted neural networks. *Finance Research Letters*, *72*(1), 106505. https://doi.org/10.1016/j.frl.2024.106505

Conrad, C., & Engle, R. F. (2025). Modelling volatility cycles: The MF2-GARCH model. *Journal of Applied Econometrics*, *40*(4), 438–454. https://doi.org/10.1002/jae.3118

Dessie, E., Birhane, T., Mohammed, A., & Walelign, A. (2025). A hybrid GARCH, convolutional neural network and long short term memory methods for volatility prediction in stock market. *Journal of Combinatorial Mathematics and Combinatorial Computing*, *124*(461), 476. https://doi.org/10.61091/jcmcc124-30

Di Persio, L., Garbelli, M., Mottaghi, F., & Wallbaum, K. (2023). Volatility forecasting with hybrid neural networks methods for risk parity investment strategies. *Expert Systems with Applications*, *229(1)*, 120418. https://doi.org/10.1016/j.eswa.2023.120418

Di Persio, L., Garbelli, M., Mottaghi, F., & Wallbaum, K. (2023). Volatility forecasting with hybrid neural networks methods for Risk Parity investment strategies. *Expert Systems with Applications*, *229*(1), 120418. https://doi.org/10.1016/j.eswa.2023.120418

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the Econometric Society*, 987-1007. https://doi.org/10.2307/1912773

Fałdziński, M., & Osińska, M. (2016). Volatility estimators in econometric analysis of risk transfer on capital markets. *Dynamic Econometric Models*, *16*, 21-35. https://doi.org/10.12775/DEM.2016.002

Garman, M. B., & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, 67-78. https://doi.org/10.1086/296072

Gong, X., & Lin, B. (2018). Modeling stock market volatility using new HAR-type models. *Physica a Statistical Mechanics and Its Applications*, *516*(1), 194–211. https://doi.org/10.1016/j.physa.2018.10.013

Graves, A., Fernández, S., & Schmidhuber, J. (2005, September). Bidirectional LSTM networks for improved phoneme classification and recognition. In *International Conference on Artificial Neural Networks* (pp. 799-804). Berlin, Heidelberg. https://doi.org/10.1007/11550907_126

Hajizadeh, E., Seifi, A., Fazel Zarandi, M. H., & Turksen, I. B. (2012). A hybrid modeling approach for forecasting the volatility of S&P 500 index return. *Expert Systems with Applications*, *39*(1), 431–436. https://doi.org/10.1016/j.eswa.2011.07.033

Hamid, S. A., & Iqbal, Z. (2004). Using neural networks for forecasting volatility of S&P 500 Index futures prices. *Journal of Business Research*, *57*(10), 1116–1125. https://doi.org/10.1016/s0148-2963(03)00043-2

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*(1), 124907. https://doi.org/10.1016/j.physa.2020.124907

Jaeger, H. (2007). Echo state network. *Scholarpedia*, *2*(9), 2330. https://doi.org/10.4249/scholarpedia.2330

Kim, J. M., Jun, C., & Lee, J. (2021). Forecasting the volatility of the cryptocurrency market by GARCH and Stochastic Volatility. *Mathematics*, *9*(14), 1614. https://doi.org/10.3390/math9141614

Kim, H. Y., & Won, C. H. (2018). Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Systems with Applications*, *103*(1), 25–37. https://doi.org/10.1016/j.eswa.2018.03.002

Koo, E., & Kim, G. (2022). A hybrid prediction model integrating garch models with a distribution manipulation strategy based on lstm networks for stock market volatility. *IEEE Access*, *10(1)*, 34743-34754. https://doi.org/10.1109/ACCESS.2022.3163723

Kristjanpoller R, W., & Hernández P, E. (2017). Volatility of main metals forecasted by a hybrid ANN-GARCH model with regressors. *Expert Systems with Applications*, *84*(1), 290–300. https://doi.org/10.1016/j.eswa.2017.05.024

Kristjanpoller, W., & Minutolo, M. C. (2015). Gold price volatility: A forecasting approach using the Artificial Neural Network–GARCH model. *Expert Systems with Applications*, *42*(20), 7245–7251. https://doi.org/10.1016/j.eswa.2015.04.058

Kristjanpoller, W., & Minutolo, M. C. (2016). Forecasting volatility of oil price using an artificial neural network-GARCH model. *Expert Systems with Applications*, *65*(1), 233–241. https://doi.org/10.1016/j.eswa.2016.08.045

Kristjanpoller, W., Fadic, A., & Minutolo, M. C. (2014). Volatility forecast using hybrid Neural Network models. *Expert Systems with Applications*, *41*(5), 2437–2442. https://doi.org/10.1016/j.eswa.2013.09.043

Kumar, S., Rao, A., & Dhochak, M. (2025). Hybrid ML models for volatility prediction in financial risk management. *International Review of Economics & Finance*, *98*, 103915. https://doi.org/10.1016/j.iref.2025.103915

Laurent, S. (2004). Analytical derivates of the APARCH model. *Computational Economics*, *24(1)*, 51-57. https://doi.org/10.1023/B:CSEM.0000038851.72226.76

Li, Y., Liu, G., & Zhang, Z. (2022). Volatility of volatility: Estimation and tests based on noisy high frequency data with jumps. *Journal of Econometrics*, *229*(2), 422–451. https://doi.org/10.1016/j.jeconom.2021.02.007

Liu, Y. (2019). Novel volatility forecasting using deep learning - Long short term memory recurrent neural networks. *Expert Systems with Applications*, *1*(1). https://doi.org/10.1016/j.eswa.2019.04.038

Lu, X., Que, D., & Cao, G. (2016). Volatility forecast based on the hybrid artificial Neural Network and GARCH-type models. *Procedia Computer Science*, *91*(1), 1044–1049. https://doi.org/10.1016/j.procs.2016.07.145

Mademlis, D. K., & Dritsakis, N. (2021). Volatility forecasting using hybrid GARCH neural network models: The case of the Italian stock market. *International Journal of Economics and Financial Issues*, *11*(1), 49. https://doi.org/10.32479/ijefi.10842

Mademlis, D. K., & Dritsakis, N. (2021). Volatility forecasting using hybrid GARCH neural network models: The case of the Italian stock market. *International Journal of Economics and Financial Issues*, *11*(1), 49–60. https://doi.org/10.32479/ijefi.10842

Mahajan, V., Thakan, S., & Malik, A. (2022). Modeling and forecasting the volatility of NIFTY 50 using GARCH and RNN models. *Economies*, *10*(5), 102. https://doi.org/10.3390/economies10050102

Manogna, R. L., Dharmaji, V., & Sarang, S. (2025). A novel hybrid neural network-based volatility forecasting of agricultural commodity prices: Empirical evidence from India. *Journal of Big Data*, *12*(1), 85. https://doi.org/10.1186/s40537-025-01131-8

Miura, R., Lukáš Pichl, & Taisei Kaizoji. (2019). Artificial neural networks for realized volatility prediction in cryptocurrency time series. *Lecture Notes in Computer Science*, *1*(1), 165–172. https://doi.org/10.1007/978-3-030-22796-8_18

Namdari Birgani, S., Sedighi, A. H., & Molaalizadeh Zavardehi, S. (2024). Portfolio optimization using deep reinforcement learning. *Engineering Management and Software Computing*, *10*(2), 1–22. https://doi.org/10.22091/jemsc.2025.11158.1192

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, 347-370. https://doi.org/10.2307/2938260

Olubusola, O., Mhlongo, N. Z., Daraojimba, D. O., Ajayi-Nifise, A. O., & Falaiye, T. (2024). Machine learning in financial forecasting: A US review: Exploring the advancements, challenges, and implications of AI-driven predictions in financial markets. *World Journal of Advanced Research and Reviews*, *21*(2), 1969-1984. https://doi.org/10.30574/wjarr.2024.21.2.0444

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 61-65. https://www.jstor.org/stable/2352357

Pierdzioch, C., Risse, M., & Rohloff, S. (2016). A boosting approach to forecasting the volatility of gold-price fluctuations under flexible loss. *Resources Policy*, *47*(1), 95–107. https://doi.org/10.1016/j.resourpol.2016.01.003

Poon, S. H., & Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, *41*(2), 478-539. https://doi.org/10.1257/002205103765762743

Ramos-Pérez, E., Alonso-González, P. J., & Núñez-Velázquez, J. J. (2019). Forecasting volatility with a stacked model based on a hybridized Artificial Neural Network. *Expert Systems with Applications*, *129*(1), 1–9. https://doi.org/10.1016/j.eswa.2019.03.046

Rogers, L. C. G., & Satchell, S. E. (1991). Estimating variance from high, low and closing prices. *The Annals of Applied Probability*, 504-512. https://doi.org/10.1214/aoap/1177005835

Rubio, L., Adriana Palacio P., Adriana Mejía C., & Ramos, F. (2023). Forecasting volatility by using wavelet transform, ARIMA and GARCH models. *Eurasian Economic Review*, *13*(3-4), 803–830. https://doi.org/10.1007/s40822-023-00243-x

Rubio, L., Palacio Pinedo, A., Mejía Castaño, A., & Ramos, F. (2023). Forecasting volatility by using wavelet transform, ARIMA and GARCH models. *Eurasian Economic Review*, *13*(3), 803-830. https://doi.org/10.1007/s40822-023-00243-x

Sardelich, M., & Manandhar, S. (2018, December 25). Multimodal deep learning for short-term stock volatility prediction. *ArXiv.org*. https://doi.org/10.48550/arXiv.1812.10479

Schwert, G. W. (1990). Stock market volatility. *Financial Analysts Journal*, *46*(3), 23-34. https://doi.org/10.2469/faj.v46.n3.23

Seo, M., & Kim, G. (2020). Hybrid forecasting models based on the Neural Networks for the volatility of bitcoin. *Applied Sciences*, *10*(14), 4768. https://doi.org/10.3390/app10144768

Shen, Z., Wan, Q., & Leatham, D. J. (2021). Bitcoin return volatility forecasting: A comparative study between GARCH and RNN. *Journal of Risk and Financial Management*, *14*(7), 337. https://doi.org/10.3390/jrfm14070337

Siddiraju, N., & Hasan, R. (2025). A Hybrid LSTM-GARCH Network for stock market volatility prediction. In *2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC), 1*(1), 00427–00433. https://doi.org/10.1109/ccwc62904.2025.10903807

Stoll, H. R., & Whaley, R. E. (1990). Stock market structure and volatility. *The Review of Financial Studies*, *3*(1), 37-71. https://doi.org/10.1093/rfs/3.1.37

Trierweiler Ribeiro, G., Alves Portela Santos, A., Cocco Mariani, V., & dos Santos Coelho, L. (2021). Novel hybrid model based on echo state neural network applied to the prediction of stock price return volatility. *Expert Systems with Applications*, *184*(1), 115490. https://doi.org/10.1016/j.eswa.2021.115490

Ulu, Y. (2025). Forecasting stock volatility via hybrid deep learning and GARCH family models: A case study from BIST30. *Journal of Applied Mathematics and Computation*, *9*(4). http://dx.doi.org/10.26855/jamc.2024.12.001

Yang, D., & Zhang, Q. (2000). Drift-independent volatility estimation based on high, low, open, and close prices. *The Journal of Business*, *73*(3), 477-492. https://doi.org/10.1086/209650

Yu, S., & Li, Z. (2018). Forecasting stock price index volatility with LSTM deep Neural Network. *Recent Developments in Data Science and Business Analytics*, *1*(1), 265–272. https://doi.org/10.1007/978-3-319-72745-5_29

Zhang, C., Zhang, Y., Cucuringu, M., & Qian, Z. (2023). Volatility forecasting with machine learning and intraday commonality. *Journal of Financial Econometrics*, *1*(1). https://doi.org/10.1093/jjfinec/nbad005

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, *14*(1), 35-62. https://doi.org/10.1016/S0169-2070(97)00044-7

Zhao, P., Zhu, H., Siu, W., & Lee, D. L. (2024). From GARCH to neural network for volatility forecast. *Proceedings of the AAAI Conference on Artificial Intelligence*, *38*(15), 16998–17006. https://doi.org/10.1609/aaai.v38i15.29643