

Article

Forecasting Stock Market Volatility Using CNN-BiLSTM-Attention Model with Mixed-Frequency Data

Yufeng Zhang ^{1,*}, Tonghui Zhang ^{1,*}  and Jingyi Hu ²¹ School of Economics, Ocean University of China, Qingdao 266100, China² SWUFE-UD Institute of Data Science, Southwestern University of Finance and Economics, Chengdu 611130, China

* Correspondence: zth@ouc.edu.cn

Abstract: Existing stock volatility forecasting models predominantly rely on same-frequency market data while neglecting mixed-frequency integration and face particular challenges in incorporating low-frequency macroeconomic variables that exhibit temporal mismatches with financial market dynamics. To address this limitation, this study develops a novel hybrid approach for stock market volatility forecasting, which synergistically combines a deep learning model (CNN-BiLSTM-Attention) with the GARCH-MIDAS model. The GARCH-MIDAS model can fully exploit mixed-frequency information, including daily returns, monthly macroeconomic variables, and EPU. The deep learning model can effectively capture both spatial and temporal patterns of multivariate time-series data, thus effectively improving prediction accuracy and generalization ability in stock market volatility forecasting. The results indicate that the CNN-BiLSTM-Attention model yields the most accurate forecasts compared to the benchmark models. Furthermore, incorporating additional predictors, such as macroeconomic indicators and the Economic Policy Uncertainty Index, also provides valuable information for stock market volatility prediction, notably enhancing the model's forecasting effect.

Keywords: macroeconomic variables; volatility forecasting; GARCH-MIDAS model; deep learning model

MSC: 62R07; 91B84; 91G15



Academic Editor: Anatoliy Swishchuk

Received: 10 April 2025

Revised: 14 May 2025

Accepted: 16 May 2025

Published: 5 June 2025

Citation: Zhang, Y.; Zhang, T.; Hu, J. Forecasting Stock Market Volatility Using CNN-BiLSTM-Attention Model with Mixed-Frequency Data. *Mathematics* **2025**, *13*, 1889. <https://doi.org/10.3390/math13111889>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Market volatility represents a multifaceted and time-varying process that attracts significant analytical attention across financial markets. Accurate volatility forecasting models serve as critical mechanisms for market participants to identify impending regime shifts, thereby enabling proactive portfolio adjustments. Financial market volatility is crucial in investment decisions, option pricing, and risk management. In particular, the outbreak of the global financial crisis and the rapid development of financial derivatives markets have made the study of financial market volatility increasingly important. Research has shown that financial volatility exhibits a range of complex dynamic characteristics, including time-varying volatility, aggregation, asymmetry, and long memory.

Several traditional methods exist for predicting stock market volatility. Bollerslev [1] proposed the generalized autoregressive conditional heteroskedasticity (GARCH) family models that utilize low-frequency data to capture volatility clustering phenomena and generate volatility forecasts. Nevertheless, these models fail to incorporate intraday trading dynamics of financial instruments. Andersen and Bollerslev [2] addressed this limitation

by calculating realized volatility using intraday high-frequency trading data to measure intraday volatility. Corsi [3] introduced the HAR model for volatility forecasting within the framework of the heterogeneous market hypothesis, demonstrating enhanced predictive performance. These traditional methods exhibit strong predictive performance for linear correlation variables, but the inherently nonlinear dynamics of financial markets necessitate considering multiple exogenous variables in operational frameworks—such as macro factors [4], sentiment factors [5], and investor attention factors [6]. The relationships between these variables frequently demonstrate non-stationary temporal evolution and nonlinearity, posing significant challenges for conventional linear frameworks to effectively capture evolving market dynamics, thereby constraining predictive performance. Concurrently, classical econometric models remain fundamentally constrained to stationary processes, whereas financial price series inherently exhibit persistent trends and cyclical patterns that systematically violate covariance–stationarity prerequisites. Therefore, the continued use of these methods may limit further improvements in prediction performance.

In recent years, the accelerated advancement of artificial intelligence coupled with the emergence of big data has driven exponential growth in the application of deep learning techniques within financial time-series prediction. These techniques are able to describe the nonlinear relationship between data and multidimensional influencing factors, which has led many scholars to utilize deep learning models to predict stock market volatility. For example, Chen et al. [7] used a deep learning approach to predict the volatility of SSEC. Nelson et al. [8] employed the LSTM to forecast stock price trends. Experimental results demonstrated that this approach attains an average accuracy of 55.9%. Similarly, Sima et al. [9] utilized BiLSTM for financial time-series prediction and found that it provided more accurate forecasts than ARIMA and LSTM models. Vidal et al. [10] introduced a hybrid model combining CNN and LSTM to forecast gold price volatility. Their study demonstrated that the model effectively extracts time-series features for high-accuracy prediction, outperforming the individual CNN and LSTM models. Wang et al. [11] developed a CNN-BiLSTM architecture by enhancing conventional BiLSTM frameworks with convolutional feature extraction, benchmarking its stock forecasting performance against LSTM, BiLSTM, and CNN-LSTM. Empirical validation demonstrated that the proposed model achieves superior predictive accuracy through joint temporal–spatial feature learning. Furthermore, historical data contribute differently to data points at different moments. To improve model prediction accuracy, attention mechanisms are often integrated into deep learning models. Zhou et al. [12] improved the feature extraction process for precious metal price dynamics through the integration of an RSA mechanism, enabling effective extraction of spatiotemporal features. The existing literature demonstrates that neural networks achieve higher accuracy in volatility prediction, motivating their application in this paper.

Additionally, hybrid models have increasingly been used to forecast volatility with promising results. Building upon neural network hybridization paradigms, Roh [13] systematically merged artificial neural networks with GARCH-family volatility estimators (GARCH, EGARCH, and EWMA) for KOSPI volatility forecasting, where the NN-GARCH/EGARCH model had a lower MAE than the solitary neural network model. Contemporary research extensions have further developed this paradigm by synergistically coupling GARCH architectures with LSTM networks [14–18], with empirical validations confirming that these enhanced hybrid frameworks exhibit statistically significant accuracy improvements. Nguyen et al. [19] proposed the SR-RV model by combining stochastic volatility (SV) with a recurrent neural network (RNN). By integrating various models, not only is forecasting accuracy improved, but interpretability is also enhanced.

However, the aforementioned traditional econometric and deep learning models do not consider the impact of macroeconomic variables. Relevant studies in the literature, such as those of Barro [20], Paye [21], and Nonejad [22], have empirically demonstrated that macroeconomic and financial indicators possess significant predictive capacity for stock market volatility. Macroeconomic variables can ultimately precipitate stock market shocks via a range of transmission mechanisms, including marginal utility, the cost of capital, discount rates, and so forth. The inherent frequency disparity between predominantly monthly macroeconomic indicators (low-frequency) and high-frequency stock trading data poses significant challenges for conventional econometric models in processing mixed-frequency data. To overcome the issue of non-identical frequencies in a dataset, Engle et al. [23] proposed the GARCH-MIDAS model, which was applied to the GARCH model using the mixed-data sampling (MIDAS) regression method based on mixed data proposed by Ghysels [24]. This approach successfully addressed the challenge of volatility fitting based on data with varying frequencies. The GARCH-MIDAS model, which incorporates macroeconomic variables, is currently a widely used tool for forecasting volatility. Fang et al. [25] employed the data of macroeconomic variables in the United States to fit the volatility of the gold future market through the GARCH-MIDAS model. Yi et al. [26] employed the data of the Chinese crude oil future market as a case study, utilizing the GARCH-MIDAS model to examine the explanatory and fitting capabilities of macroeconomic uncertainty. These studies have collectively enhanced the predictive accuracy of volatility.

In recent years, the international situation has become increasingly complex and dynamic, leading to heightened economic policy uncertainty in various countries. China's stock market is in an emerging development phase, exhibiting distinctive characteristics of a "policy market". Fluctuations in the market are closely linked to shifts in economic policies. To quantify the degree of economic policy uncertainty in China, Huang et al. [27] employed textual analysis methods in conjunction with the news reports of numerous Chinese newspapers pertaining to economic and policy matters. This approach was utilized to construct the China Economic Policy Uncertainty (EPU) Index, which can effectively reflect the alterations in economic policies in a prompt and continuous manner. An increasing number of studies have demonstrated a close relationship between EPU and stock market volatility prediction [28–31].

This paper took the SSEC Index as the object of study in the context of volatility prediction. We employed factor analysis to construct macroeconomic indicators through dimensionality reduction and utilized a multifactor GARCH-MIDAS model to address the issue of mixed-frequency data pertaining to macroeconomic and SSEC Index trading data. Subsequently, the CNN-BiLSTM-Attention model systematically captured spatiotemporal dependencies within multivariate chronological datasets through hierarchical feature learning, enabling robust volatility estimation. In this model, the CNN layer is capable of capturing the hierarchical structure of the data, the BiLSTM layer can leverage the advantages of LSTM while addressing the challenges posed by long memory in time series, and the attention mechanism layer is designed to extract pertinent key information while focusing on the trend and change in the time series. Ultimately, the model was evaluated for forecasting accuracy based on the established criteria for multiple forecasting ability. Furthermore, the empirical section concludes with an investigation into the impact of diverse predictors.

In summary, the contributions of this paper include the following two aspects:

- By considering multiple predictive factors, including both macro-level economic indicators and policy uncertainty indicators, as well as micro-level stock market trading information data, a more comprehensive analytical perspective for stock market volatility prediction from both macro and micro information can be obtained. In constructing

the indicators, this study departs from conventional approaches by innovatively integrating LASSO regression with factor analysis. This method enables the selection of key factors from a large pool of raw variables to build macroeconomic fundamental indicators. Additionally, by incorporating trading-based technical indicators and volatility jump components at the micro level, the model establishes a robust foundation for volatility forecasting with improved predictive accuracy and enhanced explanatory power.

- The combination of the multifactor GARCH-MIDAS model and the CNN-BiLSTM-Attention model addresses the issue of disparate frequencies between macro data and trading data. Furthermore, the application of the CNN-BiLSTM-Attention model to stock market volatility prediction effectively depicts the nonlinear characteristics of volatility series, and the model demonstrates superior performance in volatility prediction compared to the benchmark model.

The remainder of the paper is structured as follows: Section 2 introduces the GARCH-MIDAS and CNN-BiLSTM-Attention models, along with a comprehensive summary of the evaluation criteria. Section 3 shows the data utilized in this paper and empirically benchmarks the out-of-sample predictive performance of five competing models across four distinct loss functions in volatility forecasting. Section 4 presents an extension analysis and considers economic value performance, and Section 5 provides the conclusions.

2. Methodology

2.1. Multifactor GARCH-MIDAS Model

Compared to the GARCH model for same-frequency data, the primary distinction of the GARCH-MIDAS model for mixed-frequency data is that it incorporates a component equation setup, separating the financial asset volatility into long-run and short-run volatility. As demonstrated by Engle's study, the volatility of stock returns can be decomposed into a short-term component, denoted as $g_{i,t}$, and a long-term component, represented by τ_t . The stock return series can be expressed as follows:

$$r_{i,t} = \mu + \sqrt{\tau_t g_{i,t}} \varepsilon_{i,t} \quad (1)$$

To ensure the distinct identifiability of these two components, which is crucial for the interpretability and robustness of parameter estimates, a normalization constraint is imposed on the short-term component. The short-term component, $g_{i,t}$, follows a GARCH (1,1) process, with its unconditional mean normalized to unity, i.e., $E[g_{i,t}] = 1$. This is typically achieved by formulating the $g_{i,t}$ process as follows:

$$g_{i,t} = (1 - \alpha - \beta) + \frac{\alpha(r_{i-1} - \mu)^2}{\tau_t} + \beta g_{i-1,t} \quad (2)$$

with $\alpha > 0$ and $\beta > 0$ and $\alpha + \beta < 1$.

Then, we specify the long-term component, τ_t , by smoothing realized volatility in the spirit of MIDAS regression and MIDAS filtering:

$$\tau_t = m + \theta \sum_{j=1}^K \phi_j(\omega_1, \omega_2) RV_{t-j} \quad (3)$$

$$RV_t = \sum_{i=1}^{N_t} r_{i,t}^2 \quad (4)$$

where K is the lag time for smooth realized volatility. $\phi_j(\omega_1, \omega_2)$ in the above specification represents the so-called beta weights, defined as follows:

$$\phi_j(\omega_1, \omega_2) = \frac{(j/K)^{\omega_1-1}(1-j/K)^{\omega_2-1}}{\sum_{i=1}^K (i/K)^{\omega_1-1}(1-i/K)^{\omega_2-1}} \quad (5)$$

The aforementioned single-factor GARCH-MIDAS model solely considers historical volatility data and neglects to incorporate the influence of other exogenous low-frequency variables. To encompass the impact of diverse low-frequency factors on the long-term components of stock market volatility, this paper presents a multifactor GARCH-MIDAS model incorporating exogenous variables, denoted by X_i , representing other exogenous low-frequency variables.

$$\tau_t = m + \theta_1 \sum_{j=1}^K \phi_{1,j}(\omega_{1,1}, \omega_{1,2}) RV_{t-j} + \theta_2 \sum_{j=1}^K \phi_{2,j}(\omega_{2,1}, \omega_{2,2}) X_{1,t-j} + \theta_3 \sum_{j=1}^K \phi_{3,j}(\omega_{3,1}, \omega_{3,2}) X_{2,t-j} \quad (6)$$

where θ_1 , θ_2 , and θ_3 denote the marginal contributions of exogenous low-frequency variables to the long-term component of the SSEC Index volatility.

Finally, the total conditional variance can be defined as follows:

$$h_{i,t} = \tau_t g_{i,t} \quad (7)$$

where $h_{i,t}$ is the total conditional variance.

2.2. Convolutional Neural Network

A Convolutional Neural Network (CNN) represents an effective approach for extracting features from time-series data. The primary benefit of a CNN is the automated identification of crucial spatial characteristics. During the training phase, the convolutional layer is optimized to extract features with a high degree of discriminative properties.

A CNN comprises a convolutional layer and a pooling layer that can efficiently extract global spatial features and circumvent overfitting by reducing the number of parameters and features. The convolution layer is the core of a CNN. A convolutional layer is responsible for extracting spatial features from neighboring data points of the input data, which is achieved through a convolution operation. The parameters of a convolutional layer typically include the number of convolution kernels, the size (or length) of the convolution kernels, and a bias term. During forward propagation, the convolution layer extracts features by computing the dot product between each convolution kernel and a local region of the input data as follows:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} \times k_{ij}^l + b_j^l\right) \quad (8)$$

2.3. Bidirectional Long Short-Term Memory

A Bidirectional Long Short-Term Memory (BiLSTM) network comprises two layers of Long Short-Term Memory (LSTM) units operating in opposite directions. The network's output is influenced by the forward and backward LSTM layers, which process the input data simultaneously in both forward and backward sequences. By integrating bidirectional processing mechanisms, the proposed framework enables enhanced temporal information flow optimization within the neural architecture, as detailed in Figure 1. The updated states of the hidden layers of the forward LSTM, backward LSTM, and the final representations of the BiLSTM are defined as follows:

$$h_t = f(W_1 \times x_t + W_2 \times h_{t-1}) \quad (9)$$

$$h'_t = f(W_3 \times x_t + W_5 \times h'_{t+1}) \quad (10)$$

$$O_t = g(W_4 \times h_t + W_6 \times h'_t) \quad (11)$$

where h_t represents the hidden state vector of the forward-pass LSTM layer at time step t , h'_t denotes the hidden state vector of the backward-pass LSTM layer at time step t , O_t is the final output vector of the BiLSTM layer at time step t , and W_1 – W_6 denote the corresponding weights of the layers.

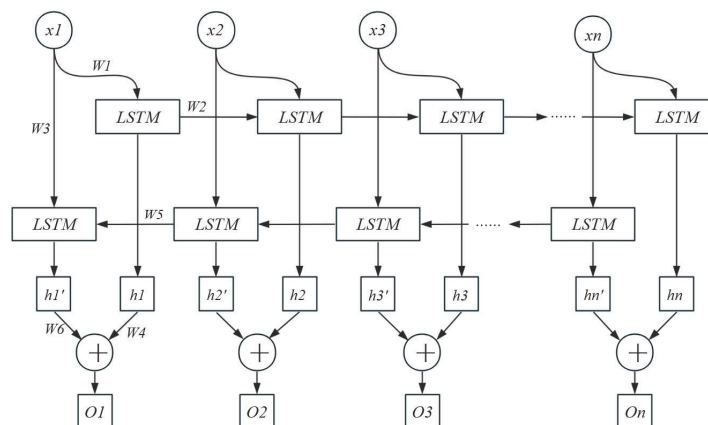


Figure 1. BiLSTM structure.

2.4. Attention Mechanism

Attention mechanisms combined with hybrid modeling have been applied across various fields. The combination of an attention mechanism—which dynamically captures long-range dependencies among elements through relative importance computation—with CNN-BiLSTM has demonstrated successful applications in specific quantitative finance domains. The model of the attentional mechanism is illustrated in Figure 2.

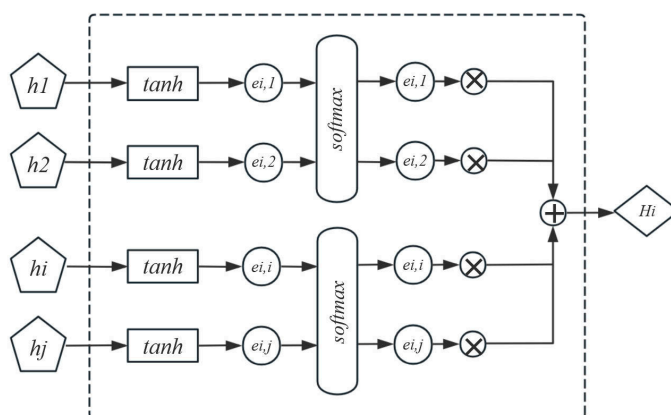


Figure 2. One basic unit of attention.

The attentional mechanism formula is as follows:

$$e_{ij} = \tanh(W_1 \times h_i + W_2 \times h_j + b) \quad (12)$$

$$a_{ij} = \text{softmax}(e_{ij}) = \frac{\exp(e_{ij})}{\sum_j \exp(e_{ij})} \quad (13)$$

$$H_i = \sum_j a_{ij} \times h_j \quad (14)$$

where e_{ij} denotes the relationship between the hidden states, h_i and h_j ; W denotes the weight; b denotes the bias vector; and H_i denotes the final output state after the attention layer.

2.5. CNN-BiLSTM-Attention Model

The prediction of realized volatility for the stock market is a challenging task due to its complex, nonlinear, and time-varying nature. To address this, this paper proposed a hybrid deep learning model combining a Convolutional Neural Network (CNN), a Bidirectional Long Short-Term Memory (BiLSTM) network, and an attention mechanism. The model of CNN-BiLSTM-Attention is illustrated in Figure 3.

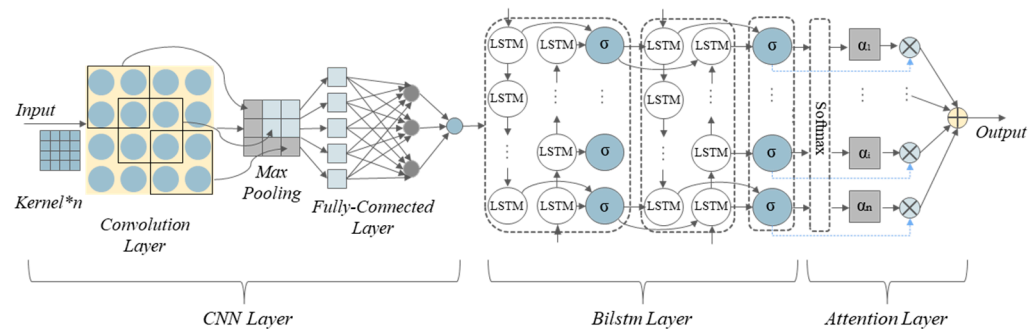


Figure 3. CNN-BiLSTM-Attention framework.

This hybrid model was chosen because of its synergistic ability to address complex features of financial time series as compared to traditional machine learning models and simpler deep learning structures. It incorporates three key components. First, the Convolutional Neural Network (CNN) serves as an effective feature extractor. It automatically identifies crucial local patterns and micro-structural features from the input sequence—patterns that often signal volatility changes. This eliminates the need for extensive manual feature engineering. Second, the Bidirectional Long Short-Term Memory (BiLSTM) network processes these extracted features. The BiLSTM effectively captures the long memory and persistence inherent in stock market volatility. By processing sequence information in both forward and reverse directions, it also achieves a more comprehensive contextual understanding of temporal dependencies. Finally, the integrated attention mechanism allows the model to dynamically assign varying importance weights to the outputs of the BiLSTM layer. This enables the model to focus on the most influential segments of the learned temporal representations when making a prediction. Consequently, this helps refine key signals within complex stock market environments, leading to enhanced predictive accuracy and robustness.

In summary, the synergistic combination of a CNN for feature extraction, a BiLSTM for comprehensive temporal modeling, and an attention mechanism for adaptive focus establishes a powerful framework. This hierarchical design enables the model to learn intricate patterns from SSEC Index data more effectively than could be achieved using these components in isolation or by employing simpler models.

The research idea of the paper is as follows: firstly, using LASSO regression and factor analysis to downscale macroeconomic indicators to obtain a small number of macro factors; then combining high-frequency returns with low-frequency macro factors to construct a multifactor GARCH-MIDAS model to obtain conditional volatility; finally, determining conditional volatility containing macroeconomic information, technical indicator factors obtained by downsizing through principal component analysis, economic policy uncertainty indices (EPU), and jump components as features, which are used as input to the CNN-BiLSTM-Attention model to ultimately predict the volatility of the SSEC Index.

2.6. Evaluation Criteria for Prediction Performance

To assess model performance, this paper employed four evaluation metrics: the mean square error (MSE), the root mean square error (RMSE), and the mean absolute error (MAE). The formulas for these metrics are defined below:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (15)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (16)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (17)$$

where n represents the total number of test samples, y_i represents the true value, and $\hat{y}$ is the corresponding predicted value. For each evaluation metric, smaller values indicate better model performance.

2.6.1. Out-of-Sample R^2

To assess the out-of-sample forecasting performance of the proposed models, we employed the out-of-sample R^2 statistic. The out-of-sample R^2 (R^2_{OOS}) reflects the percentage reduction in the mean squared forecast error (MSFE) of a given model relative to a benchmark model. In this study, the Heterogeneous Autoregressive Realized Volatility (HAR-RV) model was used as the benchmark. A positive R^2_{OOS} indicates that the evaluated model has a lower MSFE than the benchmark, suggesting superior predictive performance.

To assess the statistical significance of the improvement, the Clark and West (CW) test [32] was employed. The null hypothesis of the CW test posits that the MSFE of the benchmark model is less than or equal to that of the evaluated model. The formal definition of R^2_{OOS} is as follows:

$$R^2_{\text{OOS}} = 1 - \frac{\sum_{t=1}^n (\widehat{\text{RV}}_{i,t}^F - \text{RV}_t)^2}{\sum_{t=1}^n (\widehat{\text{RV}}_t^B - \text{RV}_t)^2} \quad (18)$$

where n represents the total number of test samples, RV_t denotes the realized stock market volatility for month t , and $\widehat{\text{RV}}_t^B$ and $\widehat{\text{RV}}_{i,t}^F$ refer to the volatility forecasts produced by the benchmark model and the i -th forecasting model.

2.6.2. Model Confidence Set (MCS)

This paper used the model confidence set proposed by Hansen [33] to test whether there was a statistically significant difference in the out-of-sample predictive effects of a predictive model. In conducting the MCS test, in addition to the selected loss functions, MSE and MAE, three additional loss functions—HMSE, HMAE, and QLIKE—were incorporated, resulting in a total of five loss functions used as performance evaluation metrics for the models.

$$\text{HMSE} = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{y_i} \right)^2 \quad (19)$$

$$\text{HMAE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (20)$$

$$\text{QLIKE} = \frac{1}{n} \sum_{i=1}^n \left(\frac{y_i}{\hat{y}_i} - \log \left(\frac{y_i}{\hat{y}_i} \right) - 1 \right) \quad (21)$$

The basic principle of MCS is as follows: All predictive models constitute a credible set, M_0 , and each test assumes that a certain two models in M_0 have the same predictive ability, establishing the following original hypothesis:

$$H_{0,M} : \forall u, v \in M \subset M_0, E(d_{i,uv}) = 0 \quad (22)$$

Based on the equivalence test and the elimination rule, models that reject the null hypothesis are excluded, and the process continues iteratively until no further models can be removed. The test employs the range statistic and the semi-quadratic statistic:

$$T_R = \max_{u,v \in M} \frac{|\bar{d}_{i,uv}|}{\sqrt{\text{var}(\bar{d}_{i,uv})}} \quad (23)$$

$$T_{SQ} = \max_{u,v \in M} \frac{(\bar{d}_{i,uv})^2}{\text{var}(\bar{d}_{i,uv})} \quad (24)$$

where $\bar{d}_{i,uv}$ denotes the average of the relative loss function values of models u and v . When the statistical values of T_R and T_{SQ} are greater than the critical value, the original hypothesis is rejected and the worse model is eliminated.

2.6.3. Direction-of-Change Test (DoC)

As observed by Degiannakis and Filis [34], the DoC test plays a pivotal role in market timing and asset allocation trading strategies. DoC gauges the extent to which forecasting models accurately anticipate the direction of change in actual volatility. A dummy variable, designated as D_i , is initially assigned a value of 1 if the model accurately forecasts the volatility trend of the stock market on trading day i and 0 otherwise. The dummy variable D_i can be expressed as follows:

$$D_i = \begin{cases} 1, & \text{if } RV_i > RV_{i-1} \text{ and } \widehat{RV}_i > RV_{i-1} \\ 1, & \text{if } RV_i \leq RV_{i-1} \text{ and } \widehat{RV}_i \leq RV_{i-1} \\ 0, & \text{otherwise} \end{cases} \quad (25)$$

The proportion of forecasts accurately capturing the direction of volatility change is defined as follows:

$$DoC = \frac{1}{n} \sum_{i=1}^n D_i \quad (26)$$

Subsequently, we employed the Pesaran–Timmermann nonparametric test [33] to examine the null hypothesis that the target model's DoC ratio does not exceed that of a random walk.

3. Empirical Results and Analysis

3.1. Data Description and Processing

In this paper, 5 min high-frequency SSEC Index data (000001.SH) from 1 February 2012 to 28 June 2024 were selected as the research sample, and the relevant data were all sourced from the Wind database. There was a total of 48 sample points for each trading day in the SSEC dataset with a 5 min sampling frequency (excluding the opening price); we used realized volatility to measure stock market volatility, which effectively reduced the impact of noise and error on volatility estimates compared to squared daily returns, in the following form:

$$RV_t = \sum_{j=1}^{48} r_{t,j}^2, t = 1, 2, \dots, T \quad (27)$$

where $r_{t,j}$ denotes the logarithmic return at the j -th moment of the t -th trading day.

Figure 4 shows the daily realized volatility of the SSEC Index. As the figure illustrates, the RV displays a time-varying and aggregated character, particularly during the 2015–2016 stock crash, when the index exhibited very high volatility.

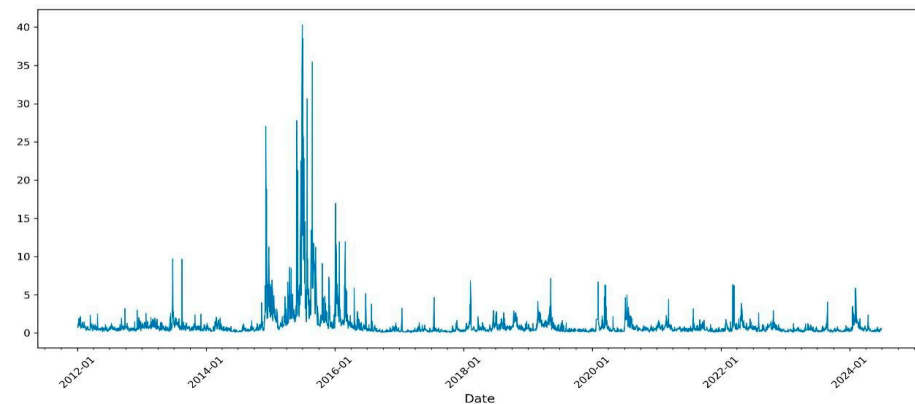


Figure 4. Realized volatility of the SSEC Index.

Four predictors were included in the analytical framework:

(1) Macroeconomic Indicators: To develop a comprehensive set of macroeconomic fundamental indicators, this study, guided by the research of Cakmakli and Dijk [35], employed monthly data for 32 macroeconomic variables spanning five key domains, including economic conditions and finance. The data cover the sample period from February 2012 to June 2024. These variables were first subjected to LASSO (Least Absolute Shrinkage and Selection Operator) regression for feature selection, after which factor analysis was performed to construct the macroeconomic fundamental factors. The data were drawn from the “CSMAR” and are shown in Table 1.

Table 1. Macroeconomic variable descriptions.

	Variable	Meaning of Variable
Economic Situation	MECI	Macroeconomic Coincident Index
	MELAI	Macroeconomic Lagging Index
	MELEI	Macroeconomic Leading Index
	CPI	Consumer Price Index
	RPI	Retail Price Index (% , 2012 = 100)
	CSI	Consumer Satisfaction Index
	CEI	Consumer Expectations Index
	RSSCG	Total Retail Sales of Social Consumer Goods
Industry	PPIRM	Purchasing Price Index of Raw Material (% , 2012 = 100)
	PPI	Producer Price Index (% , 2012 = 100)
	CGPI	Corporate Goods Price Index: Overall (% , 2012 = 100)
Finance	M0	Money Stock: M0
	M1	Money Stock: M1
	M2	Money Stock: M2
	Fldb	Financial Institutions: Deposit Balances
	Fifb	Financial Institutions: Financial Bonds
	Firmbl	Foreign Institutions’ RMB Liabilities

Table 1. Cont.

	Variable	Meaning of Variable
Investment	Vol	Trading Volume
	REDI	Real Estate Development Investment
	CFAI	Cumulative Fixed Asset Investment
	FAI-PI	Fixed Assets Investment: Primary Industry
	FAI-SI	Fixed Assets Investment: Secondary Industry
	FAI-TI	Fixed Assets Investment: Tertiary Industry
Trade	IOP	Total Value of Imports and Exports
	IP	Total Value of Imports
	OP	Total Value of Exports
	AERusd	Average Exchange Rate: America (USD Per CNY)
	AEReuro	Average Exchange Rate: European Union (EUR Per CNY)
	AERhkd	Average Exchange Rate: Hong Kong Dollar (HKD Per CNY)
	AERjpy	Average Exchange Rate: Japanese Yen (JPY Per CNY)
	REER	Real Effective Exchange Rate of the RMB
	NEER	Nominal Effective Exchange Rate of the RMB

This study initially considered a comprehensive set of macroeconomic variables identified from the established literature as pertinent to stock market volatility. This presented a significant challenge to the subsequent volatility forecasting model. To refine this initial selection, the LASSO (Least Absolute Shrinkage and Selection Operator) method was first employed for variable screening. Given the extensive nature of the initially selected variable pool, a considerable number of variables typically remained even after this LASSO procedure. Therefore, factor analysis was subsequently applied to this reduced set of variables. This second step served to further mitigate dimensionality and to construct the final, more parsimonious macroeconomic fundamental indicators utilized in our model.

Each of the selected raw macroeconomic variables comprised 149 monthly observations. For the purpose of dimensionality reduction using LASSO regression, this dataset was partitioned into training and test sets based on a 7:3 ratio. Consequently, the first 105 observations constituted the training set, while the remaining 44 observations formed the test set. The LASSO regression technique was subsequently employed for an initial screening of this comprehensive set of macroeconomic indicators. The optimal regularization parameter, λ , for the LASSO model was determined via a 10-fold cross-validation procedure applied to the training data. This process identified an optimal λ value of 0.0276. Utilizing this λ , the LASSO model selected six macroeconomic variables as most relevant. These variables, detailed in Table 2, encompassed key economic domains, such as consumption, finance, and trade. Collectively, these selected indicators provided a multifaceted representation of China's overall macroeconomic conditions, thereby offering a parsimonious yet informative basis for the subsequent factor analysis intended to construct the final macroeconomic fundamental indicators.

Table 2. Macroeconomic variables screened by LASSO regression.

Variable	Regression Coefficient	Variable	Regression Coefficient
CSI	1.3441	CFAI	0.3368
RSGG	0.2106	IOP	1.0471
PPIRM	−0.4801	AERusd	−0.9624

In order to further reduce the number of macroeconomic variables, composite macroeconomic indicators were constructed, and their applicability was improved in subsequent volatility forecasting; factor analysis was conducted on the six macroeconomic variables

selected through LASSO regression. This dimensionality reduction process yielded two principal macroeconomic factors, F_{m1} and F_{m2} , which together explained over 70% of the total variance. The detailed factor loadings are presented in Table 3.

Table 3. Macroeconomic variable factor load matrix.

Variable	F_{m1}	F_{m2}
CSI	0.12	0.82
RSGG	−0.08	0.78
CFAI	0.05	0.71
IOP	0.18	0.65
PPIRM	0.85	−0.11
AERusd	0.79	0.22

It can be seen from Table 3 that F_{m1} is primarily composed of the PPIRM and the AERusd. This factor can be interpreted as the “Cost and Foreign Exchange Factor”, capturing macroeconomic pressures stemming from production costs and external exchange rate risks. F_{m2} is mainly influenced by the CSI, RSGG, CFAI, and IOP. This factor reflects the expansion of domestic demand, as well as production and investment activity, and can thus be characterized as the “Economic Growth Factor”.

(2) SSEC Index Technical Indicators: Two types of technical indicators were selected in this paper. One consisted of the fundamental trading indicators, specifically including the opening price, the highest price, the lowest price, turnover, and five other characteristic indicators. The other included the main technical indicators, specifically the CCI, DMA, and MACD, and 17 other characteristics of the indicators; the data sample interval for the period was from 1 April 2014 to 28 June 2024, and the data came from the Wind database. The specific indicators are shown in Table 4.

Table 4. SSEC Index technical indicator descriptions.

Variable	Index	Variable	Index
HIGH	Highest Price	CCI	Commodity Channel Index
LOW	Lowest Price	DMA	Different of Moving Average
OPEN	Open Price	EXPMA	Exponential Moving Average
VOLUME	Trading Volume	KDJ	Stochastic Oscillator
AMOUNT	Trading Amount	MA (5)	5-Day Moving Average
TURN	Turnover Rate (%)	MA (20)	20-Day Moving Average
CHANGE	Price Change	MACD	Moving Average Convergence–Divergence
PCT_CHANGE	Percentage price change	PRICEOSC	Price Oscillator
BOLL	Bollinger Bands	ROC	Rate of Change
BBI	Bull and Bear Index	RSI	Relative Strength Index
BIAS	Bias Ratio	SOBV	Selling On-Balance Volume

To reduce the complexity of the volatility forecasting model, this paper performed principal component analysis (PCA) to downscale the technical indicators. Based on the results of the PCA, the variance contribution rates of the first five principal components were 36.37%, 19.96%, 12.47%, 8.44%, and 6.61%, respectively, with a cumulative variance contribution rate of 83.85%. As this cumulative value exceeded 80% and the eigenvalues of these five components were all greater than 1, it was indicated that most of the original information could be effectively retained by selecting these components. Therefore, the first five principal components were retained and utilized as technical indicators of stock market trading in the subsequent volatility modeling and forecasting.

As shown in Table 5, the principal component loading matrix yields five stock market trading technical indicators. F_{t1} represents the overall market trend and price levels, with high loadings on variables such as price indicators (HIGH, LOW, and OPEN), moving averages (MA and EXPMA), trend indicators (BBI, DMA, and MACD), and several oscillator indicators. These components are considered the most significant for capturing the general direction of market movements. F_{t2} reflects momentum and short-term fluctuations, showing strong positive loadings on price change rates (CHANGE and PCT_CHANGE) and several oscillators (BIAS, RSI, SOBV, and CCI), while exhibiting negative loadings on trend and moving average indicators. F_{t3} primarily captures trading activity, with its loadings almost exclusively concentrated on volume and turnover-related measures. F_{t4} and F_{t5} , although contributing less variance compared to the first three components, capture more subtle and complex relationships among indicators. These components may reflect specific market patterns or signal combinations derived from technical indicators.

Table 5. Principal component analysis load matrix.

Variable	F_{t1}	F_{t2}	F_{t3}	F_{t4}	F_{t5}
HIGH	0.2925	0.0288	0.1851	−0.1412	0.0097
LOW	0.2586	−0.0516	−0.0196	−0.3620	0.0668
OPEN	0.1984	−0.1489	0.1862	−0.3826	0.0584
VOLUME	0.1495	0.1276	0.4909	0.1921	0.0136
AMOUNT	0.1745	0.1125	0.4729	0.1529	0.0139
TURN	0.1485	0.1280	0.4908	0.1934	0.0137
CHANGE	0.1066	0.3110	−0.1943	0.3315	−0.0943
PCT_CHANGE	0.1076	0.3195	−0.1968	0.3342	−0.0911
BBI	0.3000	−0.2186	−0.0699	0.0545	0.0477
BIAS	0.2416	0.3027	−0.0855	−0.0627	0.0202
BOLL	0.1458	−0.2094	−0.0804	0.2782	0.5296
CCI	0.1894	0.2428	−0.0662	−0.2735	0.1659
DMA	0.2245	−0.2351	−0.0701	0.1755	−0.2959
EXPMA	0.2971	−0.2293	−0.0950	0.1060	0.0450
KDJ	0.1803	0.2006	−0.1159	−0.2356	0.2176
MA(5)	0.2720	−0.2019	−0.0835	0.0080	−0.0389
MA(20)	0.1553	−0.2441	−0.1039	0.3108	0.3833
MACD	0.3108	−0.1226	−0.0554	−0.0930	−0.2133
PRICEOSC	0.1817	−0.1868	−0.0920	0.0738	−0.4997
ROC	0.2116	0.1785	−0.0293	0.0007	−0.2574
RSI	0.1837	0.3103	−0.1546	−0.0786	0.1191
SOBV	0.1656	0.2478	−0.2103	0.0991	0.0504

(3) Economic Policy Uncertainty Indicator: In recent years, economic policy uncertainty in China has risen significantly, driven by various international factors, including the global financial crisis, interest rate hikes by the U.S. Federal Reserve, and ongoing trade tensions between China and the United States. As a macro-level risk factor, incorporating economic policy uncertainty into volatility forecasting models can enhance the models' ability to capture shifts in market sentiment and risk appetite induced by changes in the policy environment. This, to some extent, helps address the limitations of models that rely solely on historical price information.

In this study, the Economic Policy Uncertainty (EPU) Index developed by Davis and his collaborators was employed as a proxy for China's economic policy uncertainty [36]. The index is constructed using natural language processing techniques, wherein a semantic screening algorithm identifies news articles related to economic policy fluctuations from *People's Daily* and *Guang Ming Daily*. These are then standardized and compiled into a monthly time series. The sample period considered in this study spanned from February

2012 to June 2024. The relevant data were obtained from the China Economic Policy Uncertainty website (<http://www.policyuncertainty.com>) (accessed on 14 August 2024).

(4) Jump Component: The jump component of realized volatility refers to the discontinuous part of total volatility caused by sudden price changes, such as market shocks or extreme events. It exhibits temporal dependence and is particularly effective in capturing abrupt price movements triggered by unexpected news or extreme market conditions, making it a valuable source of information for forecasting stock market volatility. This component is typically analyzed alongside the continuous component, which reflects volatility from regular trading activities. Together, they constitute the total realized volatility.

In this study, the jump component of the realized volatility for the Shanghai Composite Index was calculated using the bipower variation method. The sample period for the jump component spanned from 1 April 2014 to 28 June 2024. The detailed calculation method is presented as follows:

The realized continuous volatility can be defined as follows:

$$BV_t = \frac{\pi}{2} \frac{n}{n-1} \sum_{i=2}^n |r_{t,i}| |r_{t,i-1}| \quad (28)$$

The modified Z test statistic of jump volatility is expressed as follows:

$$Z_t = \frac{\sqrt{n}(\ln RV_t - \ln BV_t)}{\sqrt{\left(\frac{\pi^2}{4} + \pi - 5\right) \max\left(1, \frac{TQ_t}{BV_t^2}\right)}} \quad (29)$$

$$TQ_t = n\mu_{4/3}^{-3} \left(\frac{n}{n-2}\right) \sum_{i=3}^n |r_{t,i}|^{4/3} |r_{t,i-1}|^{4/3} |r_{t,i-2}|^{4/3} \quad (30)$$

Given a significance level, α , $z_t > \Phi_\alpha$, a jump is identified when the test statistic z_t exceeds the critical value Φ_α , where Φ_α denotes the upper α quantile of the standard normal distribution. In this case, the indicator function for a jump takes the value 1; otherwise, it is set to 0, indicating no jump. The jump component is then constructed as follows:

$$RJV_t = I_t(Z_t > \Phi_\alpha)(RV_t - BV_t) \quad (31)$$

Table 6 gives the descriptive statistics of the SSEC Index returns and other predictors. From Table 6, it can be seen that the distribution of the SSEC Index return shows significant left-skewed and leptokurtic characteristics. The results of the Jarque–Bera statistic indicate that none of the variable data obeys the normal distribution. The ADF test statistics reject the null hypothesis of unit root presence at the 1% significance level, confirming stationarity across all variables. Thus, further econometric modeling analysis can be carried out directly.

Table 6. Descriptive statistical analysis.

	Mean	Std	Min	Max	Skew	Kurt	J-B	ADF
SSEC	0.0165	1.3029	−8.4907	5.7636	−0.8787	6.6071	5189.71 ***	−14.193 ***
F_{m1}	0.0313	0.9784	−2.8481	1.4216	−0.8968	−0.0828	357.57 ***	−6.309 ***
F_{m2}	−0.0045	1.0004	−3.0727	2.2558	−0.2953	−0.2128	43.60 ***	−5.420 ***
EPU	0.0162	0.2579	−0.9912	1.3680	0.2576	1.7215	285.78 ***	−12.906 ***
F_{t1}	0.0000	2.6036	−24.6145	9.8868	−1.1875	8.6198	9717.1 ***	−11.34 ***
F_{t2}	0.0000	1.9201	−10.6931	19.7875	1.5822	12.2962	19,593 ***	−10.569 ***
F_{t3}	0.0000	1.7815	−6.9085	13.2545	0.6037	2.4255	893.11 ***	−12.648 ***
F_{t4}	0.0000	1.4112	−11.8985	12.4343	−0.0357	9.3538	10,636 ***	−10.711 ***
F_{t5}	0.0000	1.1922	−15.1047	26.6686	4.137	113.9692	1,586,535 ***	−12.932 ***
Jump	0.1073	0.3793	0.0000	9.5464	12.4978	230.5110	4,752,081 ***	−9.357 ***

Note: The significance level presented *** 1%.

3.2. Estimation of Multifactor GARCH-MIDAS Model

To incorporate macro low-frequency data in the forecasting of stock market volatility, full-sample multifactor GARCH-MIDAS model estimation was first carried out, in which the macro factors, F_{m1} and F_{m2} , as well as monthly EPU, were used as the monthly low-frequency economic variables. The daily returns of the SSEC were used as the high-frequency variables to construct the multifactor GARCH-MIDAS model. The parameter estimates and corresponding statistics are reported in Table 7. In the model estimation, the lag parameter $K = 24$ was chosen for the low-frequency variables. The analysis assumed a 24-month lag for macroeconomic effects on market volatility. From the results of parameter estimation, all estimated parameters demonstrated statistical significance at the 1% level except for the mean parameters of short-term volatility μ , which were not significant. The estimates of α and β were significantly non-zero at 0.0815 and 0.9010, respectively, and the sum of the two was close to 1, indicating that the GARCH model was smooth and well-fitted.

Table 7. Multifactor GARCH-MIDAS model parameter estimate results.

μ	α	β	m	θ_1	θ_2	θ_3	ω_1	ω_2	ω_3
0.0110 (1.3095)	0.0815 *** (5.5068)	0.9010 *** (51.4857)	0.5250 *** (3.5473)	0.6050 *** (3.3989)	0.5700 *** (3.3529)	−0.7150 *** (−3.4878)	6.50 *** (3.6111)	8.20 *** (3.9048)	1.85 *** (3.3636)

Notes: The table gives the estimation results of multifactor GARCH-MIDAS from Table 7. The numbers in parentheses are the t -values corresponding to the above parameter estimates. *** denote rejection of the null hypothesis at the 1% significance level respectively.

Regarding the long-term component, the model incorporates three low-frequency variables: the EPU Index and two macroeconomic factors (F_{m1} and F_{m2}). The parameter θ represents the long-term influence of these variables on the volatility of the SSEC Index. Specifically, θ_1 , associated with the EPU Index, was estimated at 0.6050 and was significantly positive. This indicates that rising policy uncertainty significantly increases market volatility. In an environment of uncertain policy outlooks and frequent shifts in economic signals, investor risk aversion tends to increase, thereby amplifying volatility. This finding confirms that incorporating EPU as a long-term explanatory variable enhances the model's ability to reflect market risks in the face of major policy changes and unexpected events. Parameters θ_2 and θ_3 capture the effects of the macroeconomic fundamentals F_{m1} and F_{m2} , respectively. The estimate for θ_2 was 0.5700 and was significantly positive, suggesting that increases in F_{m1} —comprising the PPIRM and AERUSD—lead to heightened volatility. This reflects that rising production costs and currency depreciation increase uncertainty in corporate profitability and inflation pressure, thereby intensifying stock market fluctuations. In contrast, θ_3 was estimated at −0.7150 and was significantly negative, implying that F_{m2} , interpreted as the economic growth factor, contributes to market stability. Stronger macroeconomic performance tends to reduce volatility in the SSEC Index. Finally, ω denotes the optimal weighting parameter. The estimation results show that ω exceeded 1 for all predictors, indicating that the influence of lagged observations on long-term volatility diminishes over time. In other words, more recent information exerts a stronger effect on volatility, which is consistent with the actual dynamics of economic behavior.

Based on the estimation results of the multifactor GARCH-MIDAS model, we can obtain the short-term component (g_t) and the long-term component (τ_t), which are multiplicatively combined to obtain conditional volatility (h_t).

3.3. Forecasting Results

Based on the previous analysis and indicator construction, seven input features were selected for volatility forecasting using the CNN-BiLSTM-Attention deep learning model. These include five technical indicator components (F_{t1} – F_{t5}) derived from principal component analysis, the conditional volatility component (h_t) incorporating macroeconomic and economic policy uncertainty information, and the jump component ($Jump$). All features are obtained through dimensionality reduction and serve as comprehensive inputs for the model's predictive framework. These features collectively form the predictive input matrix for volatility forecasting. To further validate the superiority of the proposed CNN-BiLSTM-Attention model in forecasting stock market volatility, this study adopted the HAR-RV model as a benchmark and compared it with traditional machine learning models (SVR, Random Forest, and XGBoost) as well as other deep learning models (LSTM, BiLSTM, CNN-LSTM, and CNN-BiLSTM). The forecasting performance of the CNN-BiLSTM-Attention model was comprehensively evaluated against these alternatives.

Given the multifactor GARCH-MIDAS model's lag parameter specification, the conditional volatility prediction timeline commenced on 1 April 2014. Correspondingly, the CNN-BiLSTM-Attention model's input data spanned from 1 April 2014 to 28 June 2024. Following standard data partitioning protocols, we allocate 80% of observations to the training set and 20% to the test set. To mitigate overfitting, 20% of the training subset is reserved for validation, with parameter calibration guided by validation performance metrics. To mitigate the effects of dimensionality across features, feature scaling of the sample set is required.

To standardize the original data, we apply max–min normalization, scaling the values to a range between 0 and 1. This normalization helps prevent overfitting and enhances model accuracy. The min–max scaling process is implemented as follows:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (32)$$

where X' is the normalized data, X is the original data, and $X_{\max}$ and $X_{\min}$ are the maximum and minimum values.

In this paper, the main hyperparameters involved in the CNN-BiLSTM-Attention model were comprehensively considered in terms of both prediction accuracy and training time. The CNN layer was configured to extract local features using 64 filters and a convolution kernel of size 3 while enhancing the nonlinear expression ability through the ReLU activation function. The number of neuron nodes in the BiLSTM hidden layers was 32 and 16. This prevented overfitting by the BiLSTM structure, which captures forward and backward information in the time series. Additionally, a dropout layer rate of 0.2 was included to avoid overfitting. The attention layer improves the model's focus on features at key time steps by learning the weighting coefficients of each time step of the input sequence. The model was trained using the Adam optimizer. The hyperparameters for the proposed CNN-BiLSTM-Attention model were optimized using a grid search strategy. The selected training configuration included a batch size of 32 and 60 epochs. At this point, the model demonstrated enhanced fitting and generalization capabilities while maintaining high stability, which is particularly important for volatility forecasting.

After selecting appropriate parameters, the seven-dimensional time-series dataset was transformed into a supervised learning dataset in input–output mode. The values of the past five days were used as input features to predict the volatility of the SSEC Index on the subsequent trading day. The resulting predictions are presented in Figure 5.

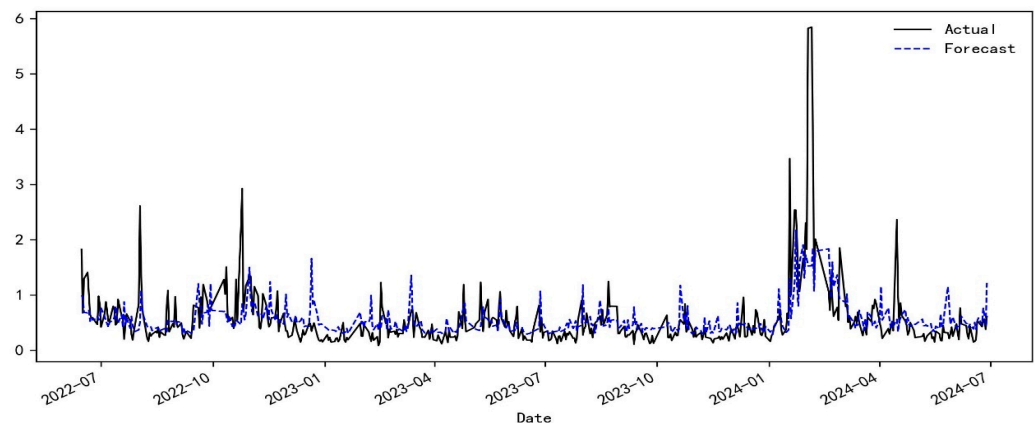


Figure 5. CNN-BiLSTM-Attention model's volatility forecasting results.

In general, the model demonstrated superior performance in identifying turning points and undulating trends, with more accurate predictions that aligned more closely with the actual outcome. Overall, the model's predictive efficacy was enhanced.

3.4. Out-of-Sample Forecasting Performance Evaluation

The out-of-sample predictive performance is more critical than the in-sample performance, as market participants prioritize a model's ability to forecast future outcomes over its capacity to analyze historical data. In this section, we evaluate the out-of-sample forecasting performance of the nine models.

3.4.1. Out-of-Sample Forecasting Loss Function Values

The results of the forecasting loss function evaluation of different models on the test dataset are shown in Table 8. The results presented in the table indicate that the predictive performance of the benchmark HAR-RV model is relatively limited. With a mean squared error (MSE) of 0.4367, it performs worse than all other models across the evaluated loss functions, as reflected in consistently higher error metrics. Traditional machine learning models such as SVR and XGBoost improve forecasting accuracy to a certain extent by leveraging nonlinear modeling techniques. Both models outperform the HAR-RV benchmark in terms of prediction error indicators.

Table 8. Comparison of evaluation error indices of nine models.

Model	MSE	RMSE	MAE
HAR-RV	0.4367	0.6608	0.3801
SVR	0.2960	0.5441	0.2667
Random Forest	0.2843	0.5332	0.3453
XGBoost	0.2871	0.5359	0.2726
LSTM	0.2947	0.5428	0.3200
BiLSTM	0.2260	0.4754	0.2772
CNN-LSTM	0.2065	0.4544	0.2531
CNN-BiLSTM	0.1940	0.4405	0.2690
CNN-BiLSTM-Attention	0.1913	0.4373	0.2405

Note: The numbers in bold in the table are the optimal values corresponding to each evaluation criterion.

The deep learning model LSTM achieves an MSE of 0.2947, which is slightly higher than that of Random Forest and XGBoost. This suggests that although LSTM effectively captures temporal dependencies, its prediction accuracy remains marginally inferior to some traditional machine learning approaches. In contrast, the BiLSTM model improves

out-of-sample forecasting performance by processing input sequences in both forward and backward directions, thereby reducing prediction errors.

Further enhancements were observed in the CNN-LSTM and CNN-BiLSTM models, which integrate convolutional layers to extract local temporal features, leading to improved forecasting accuracy and lower MSE values. The proposed CNN-BiLSTM-Attention model further advances performance by dynamically assigning attention weights to key time steps. It achieved the lowest loss values, except for MAPE, among all the models, with MSE, MAE, RMSE, and MAPE recorded at 0.1913, 0.4373, 0.2405, and 56.7565%. These results clearly demonstrate the superior capability of the CNN-BiLSTM-Attention model in stock market volatility forecasting.

3.4.2. Out-of-Sample R^2 Results

Table 9 displays the out-of-sample R^2 statistics for several forecasting models evaluated relative to the HAR-RV benchmark model. As shown in Table 9, the out-of-sample R^2 values of all the models are positive, indicating that both traditional machine learning and deep learning models exhibit improved predictive performance relative to the benchmark HAR-RV model. Except for LSTM, all deep learning models outperform traditional machine learning models in terms of R^2_{oos} . Notably, the CNN-BiLSTM-Attention model achieved the highest R^2_{oos} value of 0.5619, representing a 56.19% improvement in predictive accuracy over the benchmark model. These results highlight the superior forecasting capability of the CNN-BiLSTM-Attention model within the entire set of evaluated models.

Table 9. Out-of-sample R^2 results.

Model	R^2_{oos}	CWstat	CWp_Value
HAR-RV	-	-	-
SVR	0.3222	4.1764	0.0000
Random Forest	0.3489	3.891	0.0000
XGBoost	0.3425	3.8102	0.0001
LSTM	0.3253	4.0108	0.0000
BiLSTM	0.4823	4.4539	0.0000
CNN-LSTM	0.5288	4.46	0.0000
CNN-BiLSTM	0.5565	4.6635	0.0000
CNN-BiLSTM-Attention	0.5619	4.6535	0.0000

3.4.3. Robustness Test

(1) DoC test

As illustrated in Table 10, the results of the DoC test indicate that all forecasting models rejected the null hypothesis of no directional predictability at the 1% significance level. This confirms that each model is effective in forecasting the direction of volatility movements. The benchmark model, HAR-RV, achieved a DoC ratio of 0.5538, indicating a directional prediction success rate of 55.38%. In contrast, the CNN-BiLSTM-Attention model recorded a DoC ratio of 0.6981, representing a success rate of 69.81%—the highest among all the models evaluated. This corresponds to an improvement of 14.43 percentage points over the benchmark model. Moreover, the CNN-BiLSTM-Attention model outperformed traditional machine learning models, including SVR, Random Forest, and XGBoost, in directional accuracy. These results demonstrate the CNN-BiLSTM-Attention model's superior capability in capturing the directional dynamics of volatility, further affirming its robustness in predictive performance.

Table 10. DoC test results.

Model	DoC	PT _{stat}	<i>p</i> -Value
HAR-RV	0.5538	2.3840	0.0086
SVR	0.6827	8.0947	0.0000
Random Forest	0.6558	6.9048	0.0000
XGBoost	0.6596	7.0711	0.0000
LSTM	0.6692	7.4959	0.0000
BiLSTM	0.6866	8.2662	0.0000
CNN-LSTM	0.6769	7.8349	0.0000
CNN-BiLSTM	0.6827	8.0947	0.0000
CNN-BiLSTM-Attention	0.6981	8.7781	0.0000

(2) MCS test

To evaluate the model's forecasting performance more objectively, we performed an MCS test on the prediction results of various models. The MCS test-derived *p*-values are presented in Table 11. As illustrated in Table 11, the SVR model exhibited the strong performance under the HMSE loss function, but its effectiveness declined across other loss metrics. In contrast, the CNN-BiLSTM-Attention model achieved a *p*-value of 1 in the MCS test across multiple loss functions, including MSE, MAE, QLIKE, and HMAE, indicating statistically superior predictive performance. These results underscore the model's robustness in forecasting the volatility of the SSEC Index relative to alternative approaches. This conclusion is further supported by consistent findings from the out-of-sample R^2 and DoC tests.

Table 11. MCS test results.

Model	MSE	MAE	QLIKE	HMSE	HMAE
HAR-RV	0.1078	0.0200	0.0000	0.0526	0.0078
SVR	0.1078	0.4742	0.0358	1.0000	0.0098
Random Forest	0.1078	0.0012	0.1216	0.0010	0.0000
XGBoost	0.4148	0.4742	0.1414	0.0716	0.0098
LSTM	0.1078	0.0024	0.1216	0.0336	0.0020
BiLSTM	0.4148	0.3314	0.0000	0.0716	0.0098
CNN-LSTM	0.4148	0.4742	0.9254	0.0716	0.0098
CNN-BiLSTM	0.8200	0.4742	0.9254	0.0716	0.0098
CNN-BiLSTM-Attention	1.0000	1.0000	1.0000	0.0716	1.0000

Note: This table presents the *p*-values from the MCS test, derived from 10,000 bootstrap simulations. *p*-values equal to 1 are bolded, signifying that the corresponding models exhibit the best out-of-sample forecasting performance.

(3) Different forecasting windows

The selection of the forecasting window can affect prediction accuracy [37]. Therefore, to evaluate the robustness of our proposed model, we conducted a sensitivity analysis by altering the training–test split. The training set size varied from the baseline 80% of the full sample to 70% and subsequently 60%, with the test set correspondingly constituting 30% and 40% of the full sample.

As demonstrated in Table 12, the performance of the models remained consistent with previous findings when the length of the forecasting window was varied. The deep learning models—particularly the CNN-BiLSTM-Attention model—consistently exhibited superior predictive accuracy, highlighting their robustness across different forecasting horizons.

Table 12. Out-of-sample forecasting performance for different forecasting windows.

Model	MSE	RMSE	MAE	R^2_{oos}	CWstat	CWp_Value
HAR-RV	0.4303	0.656	0.3799	-	-	-
SVR	0.3611	0.6009	0.302	0.1604	4.3865	0.0000
Random Forest	0.3550	0.5958	0.2925	0.1750	4.5591	0.0000
XGBoost	0.3580	0.5983	0.2980	0.1680	4.4512	0.0000
LSTM	0.3565	0.5971	0.2952	0.1715	4.5038	0.0000
BiLSTM	0.3556	0.5964	0.3737	0.1731	4.6508	0.0000
CNN-LSTM	0.2953	0.5434	0.2806	0.3134	4.9267	0.0000
CNN-BiLSTM	0.2831	0.5321	0.3196	0.3426	5.4615	0.0000
CNN-BiLSTM-Attention	0.2687	0.5184	0.2683	0.3753	5.2689	0.0000
HAR-RV	0.4287	0.6547	0.3870	-	-	-
SVR	0.3960	0.6293	0.3100	0.0757	4.5903	0.0000
Random Forest	0.3791	0.6157	0.3886	0.1151	5.3329	0.0000
XGBoost	0.3796	0.6161	0.3193	0.1140	5.0365	0.0000
LSTM	0.4283	0.6545	0.3239	0.0006	3.1427	0.0013
BiLSTM	0.3940	0.6277	0.3085	0.0810	4.6500	0.0000
CNN-LSTM	0.3960	0.6293	0.3100	0.0757	4.5903	0.0000
CNN-BiLSTM	0.2836	0.5325	0.3258	0.3389	6.1202	0.0000
CNN-BiLSTM-Attention	0.2883	0.5369	0.2914	0.3272	5.8539	0.0000

(4) Replacement sample data

To further assess the robustness and generalizability of the previous findings, which were based on the Shanghai Composite Index, this study replaced the target series with the CSI 300 Index. This allowed for a more comprehensive evaluation of the model's forecasting performance in a different market context. The 5 min high-frequency data of the CSI 300 Index, along with its associated technical indicators, were sourced from the Wind database. The sample period remained consistent with that of the Shanghai Composite Index, spanning from 1 April 2014 to 28 June 2024. The data were divided into training and testing sets in an 8:2 ratio.

Table 13 presents the out-of-sample forecasting results based on the CSI 300 Index. The results indicate that the forecasting accuracy is comparable to that observed for the Shanghai Composite Index. Compared with the benchmark HAR-RV model, all alternative models exhibited positive R^2_{oos} values, confirming that both traditional machine learning and deep learning models incorporating multiple predictors can enhance the volatility forecasting accuracy for the CSI 300 Index.

Table 13. CSI 300 out-of-sample forecast accuracy evaluation results.

Model	MSE	RMSE	MAE	R^2_{oos}	CWstat	CWp_Value
HAR-RV	0.4115	0.6415	0.3702	-	-	-
SVR	0.2805	0.5296	0.2588	0.3183	4.0511	0.0000
Random Forest	0.2691	0.5187	0.3305	0.346	3.7552	0.0000
XGBoost	0.2713	0.5209	0.2651	0.3407	3.7108	0.0001
LSTM	0.2793	0.5285	0.3104	0.3213	3.9885	0.0000
BiLSTM	0.2135	0.4621	0.2688	0.4812	4.3812	0.0000
CNN-LSTM	0.1955	0.4422	0.2455	0.5249	4.3501	0.0000
CNN-BiLSTM	0.1833	0.4281	0.2601	0.5545	4.5599	0.0000
CNN-BiLSTM-Attention	0.1798	0.424	0.2315	0.563	4.5905	0.0000

Specifically, the CNN-BiLSTM-Attention model achieved MSE, RMSE, MAE, and MAPE values of 0.1798, 0.4240, 0.2315, and 54.6521%, respectively, which remained the lowest among all competing models. These results confirm that the CNN-BiLSTM-Attention model, which integrates a range of predictive factors, continues to demonstrate superior forecasting performance in the context of the CSI 300 Index, thereby validating its robustness and effectiveness in modeling stock market volatility.

4. Extension Analysis and Economic Value Performance

4.1. Extension Analysis

4.1.1. Forecasting Performance During Periods of High and Low Volatility

Liang et al. [38] demonstrated that model predictive capabilities vary according to divergent market volatility scenarios. This study further investigated the predictive performance of the HAR-RV model, traditional machine learning models, and deep learning models, including the CNN-BiLSTM-Attention model, under distinct market volatility conditions. To achieve this, market states within the out-of-sample interval were classified as either high volatility or low volatility, using the mean stock volatility during this period as the delineation criterion. The specific definitions for these volatility levels are as follows:

$$\begin{cases} \text{High Volatility : } RV_t > n^{-1} \sum_{t=1}^n RV_t \\ \text{Low Volatility : } RV_t \leq n^{-1} \sum_{t=1}^n RV_t \end{cases} \quad (33)$$

Table 14 presents the out-of-sample prediction performance of different models under two subsample periods: high and low volatility. The results indicate that deep learning models generally maintain robust predictive capabilities across both regimes. Notably, the CNN-BiLSTM-Attention model consistently achieved lower loss function values and higher R^2_{oos} values compared to most competing models. However, during the high volatility period, the CNN-BiLSTM model outperformed the CNN-BiLSTM-Attention model, achieving an R^2_{oos} of 0.5157, slightly exceeding the CNN-BiLSTM-Attention model's 0.4728. This deviation may be attributed to the increased model complexity introduced by the attention mechanism. Specifically, in highly volatile and noisy environments, the attention mechanism may struggle to learn stable and meaningful attention weights, increasing the risk of overfitting or misinterpreting short-lived fluctuations. Moreover, attention mechanisms that weigh information across the entire historical sequence may inadvertently dilute the influence of recent market movements, which are often more critical during periods of heightened volatility. In contrast, the BiLSTM model—without the additional attention layer—can more directly transmit recent temporal information, thus achieving better forecasting performance under such conditions.

Table 14. Forecasting performance during periods of high and low volatility.

Model	MSE	RMSE	MAE	R^2_{oos}	CWstat	CWp_Value
HAR-RV	1.0011	1.0006	0.5961	-	-	-
SVR	0.9068	0.9523	0.6235	0.0942	2.5616	0.0052
Random Forest	0.6444	0.8028	0.4115	0.3563	2.9206	0.0017
XGBoost	0.8340	0.9132	0.5035	0.1669	2.5660	0.0051
LSTM	0.7249	0.8514	0.4689	0.2759	2.8246	0.0024
BiLSTM	0.5636	0.7507	0.4059	0.4371	3.1649	0.0008
CNN-LSTM	0.5382	0.7336	0.4264	0.4624	3.1480	0.0008
CNN-BiLSTM	0.4849	0.6963	0.3903	0.5157	3.4281	0.0003
CNN-BiLSTM-Attention	0.5278	0.7265	0.4240	0.4728	3.2055	0.0007
HAR-RV	0.1799	0.4242	0.2819	-	-	-
SVR	0.0177	0.1330	0.1037	0.9017	5.0471	0.0000
Random Forest	0.1205	0.3472	0.3160	0.3300	3.5072	0.0002
XGBoost	0.0380	0.1949	0.1674	0.7889	4.2773	0.0000
LSTM	0.0988	0.3144	0.2524	0.4507	3.6390	0.0001
BiLSTM	0.0724	0.2691	0.2192	0.5974	4.0248	0.0000
CNN-LSTM	0.0544	0.2332	0.1729	0.6977	4.3302	0.0000
CNN-BiLSTM	0.0611	0.2472	0.2136	0.6605	4.0879	0.0000
CNN-BiLSTM-Attention	0.0380	0.1950	0.1573	0.7886	4.4674	0.0000

In the low-volatility regime, the SVR model exhibited superior performance relative to the CNN-BiLSTM-Attention model. This advantage can be explained by two factors. First, SVR's ϵ -insensitive loss function allows it to disregard small deviations, thereby focusing on capturing the general trend without being unduly affected by minor market fluctuations. Second, the SVR model with a radial basis function (RBF) kernel possesses strong generalization ability, making it particularly effective in environments characterized by smooth, non-extreme nonlinear relationships. In contrast, the CNN-BiLSTM-Attention model, with its higher parameter complexity, may be more susceptible to overfitting in low-volatility contexts where the signal-to-noise ratio is relatively low and informative features are less pronounced.

4.1.2. Examining the Validity of Different Predictors

While the evaluation of the prediction results on the test dataset shows the effectiveness of the CNN-BiLSTM-Attention model in forecasting the SSEC Index volatility, it does not clearly explain the contribution of different predictors to volatility prediction. Therefore, the predictors were classified into four categories: (A) micro-trading behavioral data, specifically technical indicators; (B) macroeconomic indicators; (C) economic policy uncertainty (EPU); and (D) jump components.

Subsequently, the potential of the various predictors in enhancing the accuracy of the CNN-BiLSTM-Attention model in volatility prediction was investigated. Figure 6 illustrates the mean loss (calculated as the mean of the four evaluation metrics) of the CNN-BiLSTM-Attention model when utilizing different combinations of predictors. As illustrated in Figure 6, the average loss of the CNN-BiLSTM-Attention model when utilizing solely technical indicators was the highest, while the average loss of the model incorporating four predictors was the lowest. Incorporating macroeconomic factors, EPU indicators, and the jump component can potentially mitigate the average loss of the model prediction to a certain extent. Furthermore, the average loss of the combination of predictors comprising macroeconomic indicators was less than that of the other combinations of predictors. This suggests that macroeconomic indicators contain valuable information and play an instrumental role in volatility forecasting.

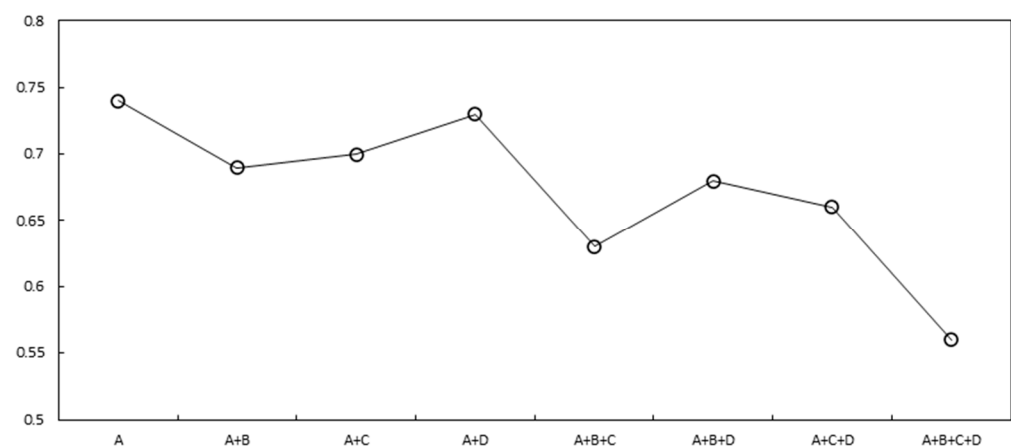


Figure 6. Mean loss of CNN-BiLSTM-Attention model with different predictors.

4.2. Economic Value Performance

In addition to statistical evaluation, both market participants and policymakers are concerned with the economic value of forecasting models. To better assess the practical applicability of these models, this study adopted the realized utility framework for economic value analysis [39]. This framework exclusively relies on RV forecasts and is designed to quantify the utility gains of investors with mean-variance preferences when allocating

assets characterized by time-varying volatility and a constant Sharpe ratio. In this paper, we applied this framework to evaluate the average expected utility associated with precious metal futures, calculated as follows:

$$U(\widehat{RV}_{t+1}) = \frac{1}{n+1} \sum_{t=1}^n \frac{SR^2}{\gamma} \left(\frac{\sqrt{RV_t}}{\sqrt{\widehat{RV}_{i,t}}} - \frac{1}{2} \frac{RV_t}{\widehat{RV}_{i,t}} \right) \quad (34)$$

where γ and SR represent the risk aversion coefficient and Sharpe ratio, respectively. According to Guo et al. [40], this study set the SR to 0.4 and considered three different levels of risk aversion, specifically, $\gamma = 2$, $\gamma = 3$, and $\gamma = 6$. Under the assumption that the model can perfectly forecast realized volatility, the corresponding maximum average expected utilities were 4.000%, 2.667%, and 1.333%.

Table 15 presents the average expected realized utility generated by each model under different levels of risk aversion, thereby reflecting the economic significance of each model's performance in volatility forecasting.

Table 15. Economic value performance results.

Model	$\gamma = 2$	$\gamma = 3$	$\gamma = 6$
HAR-RV	2.9456	1.9637	0.9819
SVR	3.0622	2.0415	1.0208
Random Forest	3.5641	2.3761	1.1881
XGBoost	3.5159	2.3440	1.1720
LSTM	3.1534	2.1023	1.0512
BiLSTM	3.6983	2.4655	1.2328
CNN-LSTM	3.7004	2.4669	1.2335
CNN-BiLSTM	3.7068	2.4712	1.2356
CNN-BiLSTM-Attention	3.6952	2.4635	1.2317

Note: The table reports the realized utility values of each model based on the mean-variance utility function. All values are expressed as percentages.

From Table 15, it is evident that, with the exception of the LSTM model, the realized utility values of deep learning models are generally higher than those of other models. This indicates that market participants with varying levels of risk aversion can obtain greater economic gains when using models such as BiLSTM, CNN-LSTM, CNN-BiLSTM, and CNN-BiLSTM-Attention for volatility forecasting. The likelihood of achieving higher investment returns is thus enhanced.

Under the scenario where $\gamma = 2$, both traditional machine learning models and deep learning models exhibited higher realized utility values compared to the benchmark HAR-RV model. Given known risk aversion coefficients and Sharpe ratios, the CNN-BiLSTM-Attention model achieved a realized utility that was 0.7496 percentage points higher than that of the HAR-RV model. This implies that, within a fixed risk-preference framework, an investor would require an additional cost of 0.7496% to elevate the utility of the HAR-RV model to the level achieved by the CNN-BiLSTM-Attention model. This cost difference can be interpreted as an additional return potential for asset management institutions.

In summary, the realized utility outcomes across different forecasting models not only validate the superior predictive accuracy of deep learning models—particularly the CNN-BiLSTM-Attention model—in modeling the volatility of the Shanghai Composite Index, but also highlight their practical value in delivering economic benefits to investors in real-world financial decision-making.

5. Conclusions

This paper investigates the impact of macro and micro predictors, such as macroeconomic indicators and micro-trading technical indicators, indices used to forecast stock market volatility, using the GARCH-MIDAS model and the CNN-BiLSTM-Attention model. Firstly, we apply LASSO and factor analysis to condense macroeconomic predictors and PCA to synthesize technical indicators through dimensionality reduction techniques. Subsequently we integrate monthly macroeconomic indicators and EPU with daily stock market index data through a GARCH-MIDAS model to estimate conditional volatility, effectively bridging the mixed-frequency data structure. In the final prediction phase, the conditional volatility (h_t) is combined with other predictors to form the input tensor for our CNN-BiLSTM-Attention model for volatility forecasting. To evaluate model performance, we conduct comprehensive empirical analyses using daily returns from the SSEC Index, employing MSE, RMSE, MAE, and MAPE as evaluation metrics. The results demonstrate that the CNN-BiLSTM-Attention framework achieves superior predictive accuracy compared to benchmark models. These findings are further substantiated through MCS tests and DoC tests, confirming the model's robustness in volatility forecasting. Finally, through extension analysis, the results show that deep learning models perform better during high-volatility periods, while the SVR model achieves better results in low-volatility periods and macroeconomic indicators contain valuable information and play an instrumental role in volatility prediction.

The out-of-sample prediction results of the CNN-BiLSTM-Attention model demonstrate substantial economic value across different levels of risk aversion. Other deep learning models also exhibit superior performance, further confirming that the application of deep learning techniques to stock market volatility forecasting not only improves predictive accuracy in statistical evaluations but also yields meaningful economic benefits for investors. These findings underscore the strong practical utility and application potential of deep learning models in real-world investment decision-making contexts.

This paper develops a methodological framework for analyzing volatility dynamics in China's finance markets, integrating predictive modeling techniques with empirical market data. In future research, the model architecture could be extended through the incorporation of option-implied, forward-looking information and investor sentiment to enhance volatility forecasting accuracy, or more adaptive model fusion techniques, which could include dynamic weight allocation mechanisms for ensemble models based on prevailing market conditions or economic cycle stages, could be developed. Furthermore, the model's practical applications in financial risk management systems and derivative pricing mechanisms warrant dedicated exploration.

Author Contributions: Conceptualization, Y.Z. and T.Z.; methodology, Y.Z.; software, Y.Z.; validation, Y.Z., T.Z. and J.H.; formal analysis, Y.Z.; investigation, Y.Z., T.Z. and J.H.; data curation, Y.Z. and J.H.; writing—original draft preparation, Y.Z.; writing—review and editing, Y.Z., T.Z. and J.H.; funding acquisition, T.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant no. 72271224), the National Social Science Foundation of China (grant no. 21BJY263, 22BJL018), the Shandong Social Science Planning Fund Program (grant no. 24DJJJ26), and the Fundamental Research Funds for the Central Universities (grant no. 842451014).

Data Availability Statement: Data used in the paper are available upon request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SSEC	Shanghai Composite Index
CNN	Convolutional Neural Network
BiLSTM	Bidirectional Long Short-term memory
EPU	Economic Policy Uncertainty Index

References

1. Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327. [\[CrossRef\]](#)
2. Andersen, T.G.; Bollerslev, T. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *Int. Econ. Rev.* **1998**, *39*, 885–905. [\[CrossRef\]](#)
3. Corsi, F. A simple approximate long-memory model of realized volatility. *J. Financ. Econom.* **2009**, *7*, 174–196. [\[CrossRef\]](#)
4. Qiu, Y.; Song, Z.; Chen, Z. Short-term stock trends prediction based on sentiment analysis and machine learning. *Soft Comput.* **2022**, *26*, 2209–2224. [\[CrossRef\]](#)
5. Lei, B.; Zhang, B.; Song, Y. Volatility forecasting for high-frequency financial data based on web search index and deep learning model. *Mathematics* **2021**, *9*, 320. [\[CrossRef\]](#)
6. Li, D. Forecasting stock market realized volatility: The role of investor attention to the price of petroleum products. *Int. Rev. Econ. Financ.* **2024**, *90*, 115–122. [\[CrossRef\]](#)
7. Chen, W. Forecasting Volatility of Shanghai Composite Index with Deep Learning. *Stat. Inf. Forum* **2018**, *33*, 99–106.
8. Nelson, D.M.; Pereira, A.C.; De Oliveira, R.A. Stock market's price movement prediction with LSTM neural networks. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1419–1426.
9. Siami-Namini, S.; Tavakoli, N.; Namin, A.S. The performance of LSTM and BiLSTM in forecasting time series. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9 December 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 3285–3292.
10. Vidal, A.; Kristjanpoller, W. Gold volatility prediction using a CNN-LSTM approach. *Expert Syst. Appl.* **2020**, *157*, 113481. [\[CrossRef\]](#)
11. Wang, H.; Wang, J.; Cao, L.; Li, Y.; Sun, Q.; Wang, J. A stock closing price prediction model based on CNN-BiSLSTM. *Complexity* **2021**, *2021*, 5360828. [\[CrossRef\]](#)
12. Zhou, J.; He, Z.; Song, Y.N.; Wang, H.; Yang, X.; Lian, W.; Dai, H.N. Precious metal price prediction based on deep regularization self-attention regression. *IEEE Access* **2019**, *8*, 2178–2187. [\[CrossRef\]](#)
13. Roh, T.H. Forecasting the volatility of stock price index. *Expert Syst. Appl.* **2007**, *33*, 916–922.
14. Kristjanpoller, W.; Fadic, A.; Minutolo, M.C. Volatility forecast using hybrid neural network models. *Expert Syst. Appl.* **2014**, *41*, 2437–2442. [\[CrossRef\]](#)
15. Bhattacharya, S.; Ahmed, A. Forecasting crude oil price volatility in India using a hybrid ANN-GARCH model. *Int. J. Bus. Forecast. Mark. Intell.* **2018**, *4*, 446–457. [\[CrossRef\]](#)
16. Kim, H.Y.; Won, C.H. Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. *Expert Syst. Appl.* **2018**, *103*, 25–37. [\[CrossRef\]](#)
17. Hu, Y.; Ni, J.; Wen, L. A hybrid deep learning approach by integrating LSTM-ANN networks with GARCH model for copper price volatility prediction. *Phys. A Stat. Mech. Its Appl.* **2020**, *557*, 124907. [\[CrossRef\]](#)
18. Kakade, K.; Mishra, A.K.; Ghate, K.; Gupta, S. Forecasting commodity market returns volatility: A hybrid ensemble learning garch-lstm based approach. *Intell. Syst. Account. Financ. Manag.* **2022**, *29*, 103–117. [\[CrossRef\]](#)
19. Nguyen, T.N.; Tran, M.N.; Gunawan, D.; Kohn, R. A statistical recurrent stochastic volatility model for stock markets. *J. Bus. Econ. Stat.* **2023**, *41*, 414–428. [\[CrossRef\]](#)
20. Barro, R.J. Rare disasters and asset markets in the twentieth century. *Q. J. Econ.* **2006**, *121*, 823–866. [\[CrossRef\]](#)
21. Paye, B.S. 'Déjà vol': Predictive regressions for aggregate stock market volatility using macroeconomic variables. *J. Financ. Econ.* **2012**, *106*, 527–546. [\[CrossRef\]](#)
22. Nonejad, N. Forecasting aggregate stock market volatility using financial and macroeconomic predictors: Which models forecast best, when and why? *J. Empir. Financ.* **2017**, *42*, 131–154. [\[CrossRef\]](#)
23. Engle, R.F.; Ghysels, E.; Sohn, B. Stock market volatility and macroeconomic fundamentals. *Rev. Econ. Stat.* **2013**, *95*, 776–797. [\[CrossRef\]](#)
24. Ghysels, E.; Sinko, A.; Valkanov, R. MIDAS regressions: Further results and new directions. *Econom. Rev.* **2007**, *26*, 53–90. [\[CrossRef\]](#)

25. Fang, L.; Yu, H.; Xiao, W. Forecasting gold futures market volatility using macroeconomic variables in the United States. *Econ. Model.* **2018**, *72*, 249–259. [[CrossRef](#)]
26. Yi, A.; Yang, M.; Li, Y. Macroeconomic uncertainty and crude oil futures volatility—evidence from China crude oil futures market. *Front. Environ. Sci.* **2021**, *9*, 636903. [[CrossRef](#)]
27. Huang, Y.; Luk, P. Measuring economic policy uncertainty in China. *China Econ. Rev.* **2020**, *59*, 101367. [[CrossRef](#)]
28. Alqahtani, A.; Bouri, E.; Vo, X.V. Predictability of GCC stock returns: The role of geopolitical risk and crude oil returns. *Econ. Anal. Policy* **2020**, *68*, 239–249. [[CrossRef](#)] [[PubMed](#)]
29. Li, T.; Ma, F.; Zhang, X.; Zhang, Y. Economic policy uncertainty and the Chinese stock market volatility: Novel evidence. *Econ. Model.* **2020**, *87*, 24–33. [[CrossRef](#)]
30. Li, X.; Ye, C.; Bhuiyan, M.A.; Huang, S. Volatility forecasting with an extended GARCH-MIDAS approach. *J. Forecast.* **2024**, *43*, 24–39. [[CrossRef](#)]
31. Gong, X.; Zhang, W.; Xu, W.; Li, Z. Uncertainty index and stock volatility prediction: Evidence from international markets. *Financ. Innov.* **2022**, *8*, 57. [[CrossRef](#)]
32. Clark, T.E.; West, K.D. Approximately normal tests for equal predictive accuracy in nested models. *J. Econom.* **2007**, *138*, 291–311. [[CrossRef](#)]
33. Hansen, P.R.; Lunde, A.; Nason, J.M. The model confidence set. *Econometrica* **2011**, *79*, 453–497. [[CrossRef](#)]
34. Degiannakis, S.; Filis, G. Forecasting oil price realized volatility using information channels from other asset classes. *J. Int. Money Financ.* **2017**, *76*, 28–49. [[CrossRef](#)]
35. Çakmaklı, C.; van Dijk, D. Getting the most out of macroeconomic information for predicting excess stock returns. *Int. J. Forecast.* **2016**, *32*, 650–668. [[CrossRef](#)]
36. Davis, S.J.; Liu, D.; Sheng, X.S. Economic policy uncertainty in China since 1949: The view from mainland newspapers. In Proceedings of the Fourth Annual IMF-Atlanta Fed Research Workshop on China's Economy, Atlanta, GA, USA, 19–20 September 2019; Volume 19, pp. 1–37.
37. Inoue, A.; Jin, L.; Rossi, B. Rolling window selection for out-of-sample forecasting with time-varying parameters. *J. Econom.* **2017**, *196*, 55–67. [[CrossRef](#)]
38. Liang, C.; Wei, Y.; Li, X. Uncertainty and crude oil market volatility: New evidence. *Appl. Econ.* **2020**, *52*, 2945–2959. [[CrossRef](#)]
39. Bollerslev, T.; Hood, B.; Huss, J.; Pedersen, L.H. Risk Everywhere: Modeling and Managing Volatility. *Rev. Financ. Stud.* **2018**, *31*, 2729–2773. [[CrossRef](#)]
40. Guo, X.; Huang, D.; Li, X.; Liang, C. Are categorical EPU indices predictable for carbon futures volatility? Evidence from the machine learning method. *Int. Rev. Econ. Financ.* **2023**, *83*, 672–693. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.