Sains Malaysiana 53(9)(2024): 3229-3239
http://doi.org/10.17576/jsm-2024-5309-25

Enhanced Foreign Exchange Volatility Forecasting using CEEMDAN with Optuna-Optimized Ensemble Deep Learning Model

(Ramalan Kemeruapan Tukaran Asing yang Dipertingkatkan menggunakan CEEMDAN dengan Model Pembelajaran Mendalam Ensembel Dioptimumkan Optuna)

REHAN KAUSAR[1], FARHAT IQBAL[2,3,*], ABDUL RAZIQ[2], NAVEED SHEIKH[4] & ABDUL REHMAN[4]

[1]*Department of Statistics, Sardar Bahadur Khan Women's University, Quetta, Pakistan*
[2]*Department of Statistics, University of Balochistan, Quetta, Pakistan*
[3]*Department of Mathematics, Imam Abdulrahman Bin Faisal University, Saudi Arabia*
[4]*Department of Mathematics, University of Balochistan, Quetta, Pakistan*

Received: 9 December 2023/Accepted: 15 July 2024

ABSTRACT

Foreign Exchange (FX) is the largest financial market in the world, with a daily trading volume that significantly exceeds that of stock and futures markets. The prediction of FX volatility is a critical financial concern that has garnered significant attention from researchers and practitioners due to its far-reaching implications in the financial markets. This paper presents a novel hybrid ensemble forecasting model integrating a decomposition strategy and three deep learning (DL) models: Long Short-Term Memory (LSTM), Bidirectional LSTM (BiLSTM), and Convolutional Neural Network (CNN). This combination addresses individual models' limitations and further improves the accuracy and stability of FX volatility forecasting. The proposed approach utilizes the CEEMDAN technique to decompose volatility into multiple distinct intrinsic mode functions (IMFs) and merges these IMFs with GARCH and EGARCH volatilities to form the input dataset for the DL models. In addition, we employed an attention mechanism to improve the effectiveness of the DL techniques. Furthermore, the hyperparameters for the DL models are optimized using the Optuna algorithm. Finally, a hybrid ensemble model for forecasting exchange rate volatility is developed by combining the predictions of three distinct DL models. The proposed approach is evaluated against various benchmark models using evaluation measures such as MSE, MAE, HMSE, HMAE, RMSE, Q-LIKE, and the model confidence set (MCS) approach. The results demonstrate that our proposed approach provides accurate and reliable forecasts of FX volatility under different forecasting regimes, making it a valuable tool for financial practitioners and researchers.

Keywords: Currency exchange rate volatility; deep learning; ensemble; CEEMDAN; Optuna

ABSTRAK

Tukaran Asing (FX) merupakan pasaran kewangan terbesar di dunia dengan volum dagangan harian yang jauh melebihi pasaran saham dan pasaran hadapan. Ramalan turun naik FX merupakan kebimbangan kewangan yang kritikal serta telah mendapat perhatian daripada penyelidik dan pengamal kerana implikasinya yang meluas dalam pasaran kewangan. Kajian ini membentangkan model ramalan ensembel hibrid baharu yang menyepadukan strategi penguraian dan tiga model pembelajaran mendalam (DL): Memori Jangka Pendek Panjang (LSTM), LSTM Dwiarah (BiLSTM) dan Rangkaian Neural Konvolusi (CNN). Gabungan ini menangani had model individu dan meningkatkan lagi ketepatan dan kestabilan ramalan turun naik FX. Pendekatan yang dicadangkan menggunakan teknik CEEMDAN untuk menguraikan turun naik kepada pelbagai fungsi mod intrinsik (IMF) yang berbeza dan menggabungkan IMF ini dengan turun naik GARCH dan EGARCH untuk membentuk set data input bagi model DL. Di samping itu, kami menggunakan mekanisme perhatian untuk meningkatkan keberkesanan teknik DL. Tambahan pula, hiperparameter untuk model DL dioptimumkan menggunakan algoritma Optuna. Akhir sekali, model ensembel hibrid untuk meramalkan turun naik kadar pertukaran dibangunkan dengan menggabungkan ramalan tiga model DL yang berbeza. Pendekatan yang dicadangkan dinilai berdasarkan pelbagai model penanda aras menggunakan ukuran penilaian seperti MSE, MAE, HMSE, HMAE, RMSE, Q-like dan pendekatan set keyakinan model (MCS). Keputusan menunjukkan bahawa pendekatan yang dicadangkan dalam kajian ini menyediakan ramalan turun naik FX yang tepat dan boleh dipercayai di bawah rejim ramalan yang berbeza, menjadikannya alat yang berharga untuk pengamal dan penyelidik kewangan.

Kata kunci: CEEMDAN; ensembel; kemeruapan kadar pertukaran mata wang; Optuna; pembelajaran mendalam

INTRODUCTION

Volatility signifies the measure of unanticipated fluctuations in asset returns during a specific time frame. It is typically quantified by the standard deviation of the associated asset returns. Increased volatility corresponds to higher levels of risk. Volatility in financial asset returns is a crucial component within contemporary financial sectors. The precise application of suitable methods for estimating and predicting volatility is very important for risk management, portfolio optimization, derivatives valuation, and more. Nonetheless, such data exhibit complex patterns and nonlinearity, as well as frequently pronounced temporal variations. Consequently, accurately estimating volatility has persistently presented a significant hurdle for financial analysts and researchers.

Until now, various methodologies have been introduced to tackle the challenges of modeling and forecasting volatility. These approaches can be broadly classified into three broad categories: traditional generalized autoregressive conditional heteroscedasticity (GARCH)-type models (Bollerslev 1986; Engle 1982), stochastic volatility models (Kastner, Frühwirth-Schnatter & Lopes 2017; Taylor 1994; Zahid & Iqbal 2020), and techniques based on machine learning (Gamboa 2017; Henrique, Sobreiro & Kimura 2019). Recently, deep learning has also found widespread applications in predicting financial time-series data.

GARCH-type models utilize historical volatility data to forecast future volatility, assuming conditional heteroscedasticity. These models have found applications in diverse domains, including forecasting the FX rate volatility (Dhamija & Bhalla 2010), stock index returns (Hajizadeh et al. 2012) and cryptocurrency Value-at-Risk (Iqbal, Zahid & Koutmos 2023), among others. However, GARCH-class models do not exhibit a persistent capacity to predict long-term observations, consistently presenting challenges when predicting prolonged trends and patterns within financial markets.

Recent advancements in artificial neural network (ANN) models have led to an increased adoption of ANN techniques in the field of volatility prediction (Pradeepkumar & Ravi 2018). Recurrent neural network (RNN), a type of enhanced ANN model, can retain time-related information in the network. However, problems like gradient explosion and gradient vanishing are common during the training of RNNs.

Hochreiter and Schmidhuber (1997) proposed an LSTM network to enhance the RNN. LSTM can effectively solve the problem of RNN in the training process and perform better for longer time series data. BiLSTM combines forward LSTM and backward LSTM to fit data from both forward and reverse directions of the sequence, resulting in improved prediction accuracy. LSTM has found application in financial time series forecasting, as demonstrated by Kim and Won (2018) and Zahid, Iqbal and Koutmos (2022). Jung and Choi (2021) introduced an innovative autoencoder LSTM model in the context of FX volatility forecasting. Abedin et al. (2021) introduced a novel BiLSTM model combining bagging ridge regression and BiLSTM neural networks as base regressors that outperformed other exchange rate prediction models. CNN is well known for extracting and generating features using convolutional and pooling layers. More recently, it has displayed remarkable effectiveness when applied to time series prediction (Dauphin et al. 2017; Kalchbrenner et al. 2016).

In recent years, ensemble learning has attracted significant attention from researchers. This technique involves constructing multiple models that forecast target values through various algorithms or distinct subsets of training and testing datasets (Granata & Di Nunno 2021). These individual models' outcomes are then merged within the ensemble to yield a final prediction for the target values. Researchers have particularly emphasized that the diversity among the components of an ensemble leads to significant improvements when compared to traditional models (Sobri, Koohi-Kamali & Rahim 2018). Despite their effectiveness in reducing the inherent variance within complex and volatile financial markets, ensemble methods have received comparatively little attention in the existing literature on forecasting FX markets.

Researchers developed an innovative ensemble forecasting method known as 'Decomposition and Ensemble' to tackle the challenge of forecasting irregular and nonstationary data (Kausar et al. 2023; Risse 2019). In recent research, empirical mode decomposition (EMD) has emerged as a prevalent approach within the field, as Santhosh, Venkaiah and Kumar (2019) observed. However, EMD encounters mode-mixing issues. In response, more advanced versions of EMD have been introduced, including ensemble EMD (EEMD) (Wu & Huang 2009) and complete ensemble EMD with adaptive noise (CEEMDAN) (Torres et al. 2011). Among these, CEEMDAN stands out for its capacity to avoid mode mixing and reduce noise in the modes, setting it apart from other versions (Colominas, Schlotthauer & Torres 2014). Moreover, variational modal decomposition (VMD) exhibits superior noise robustness and accuracy in component decomposition when compared to EMD (Dragomiretskiy & Zosso 2014). Nevertheless, while the 'Decomposition and Ensemble' approach can improve the accuracy of financial time series forecasting, it has some issues with hybrid methodology. In the prediction phase for each component, the aggregation of actual and predicted values may lead to an accumulation of estimation errors, potentially undermining the precision of the forecasting.

In response to these issues, this study introduces a hybrid model that combines CEEMDAN with three DL neural network models. Instead of aggregating the predictive values of individual modes during the prediction stage, our proposed approach simultaneously inputs various modes into DL neural networks, directly providing the predicted value for the target variable. This design is intended to circumvent the accumulation of estimation errors. The RNNs can grasp the temporal relationships between successive data points, rendering them well-suited for addressing time-series forecasting tasks. This study recommends a hybrid model combining various financial time-series models with distinct neural networks instead of using a single econometric model and a single neural network, as in previous studies. The GARCH and Exponential GARCH (EGARCH) models provide volatility estimates as inputs to the three deep learning models in this approach. We believe this approach will enhance the accuracy of financial market volatility predictions.

Furthermore, choosing the appropriate values of the hyperparameters is another consideration, particularly in deep learning. The tuning of these parameters often determines the effectiveness of a neural network model. Various optimization algorithms have been suggested for tuning the hyperparameters of DL techniques. Examples include Bayesian optimization, genetic algorithm and particle swarm optimization (PSO). The study investigates how Optuna (Akiba et al. 2019), a new optimization algorithm, can be used to effectively automate the iterative trial-and-error process of hyperparameter optimization, improving the effectiveness of an ensemble-based approach for predicting FX volatility.

Improving predictions for exchange rate volatility is critical, emphasizing the importance of the current work. Our research makes a significant contribution to the realms of modeling exchange rate volatility and forecasting by introducing an innovative methodology known as CEEMDAN-GE-OPT-LBC. To elaborate, the improved forecasts generated by our model hold practical value for multinational corporations, financial institutions, and traders seeking effective currency risk hedging strategies.

The remaining sections of the study are organized as follows: Next section provides a detailed discussion of the methodological aspects of the research. Subsequent section describes the data and performance evaluation criteria. The section that follow captures the results and discussion, and last section provides the concluding remarks.

CONTENT AND METHODS

In this section, we explain the models utilized in constructing our proposed model and outline the development of our innovative model CEEMDAN-GE-OPT-LBC, which integrates an attention mechanism.

COMPLETE ENSEMBLE EMPIRICAL MODE DECOMPOSITION WITH ADAPTIVE NOISE

The EEMD algorithm improves signals by introducing white noise based on EMD, ensuring even distribution and mitigating mode mixing effects. CEEMDAN refines EEMD by using adaptive white noise, overcoming its limitations. The algorithm's steps include:

Step 1 Take the original signal $S(t)$ and introduce white noise $v^i(t)$ with a standard normal distribution. The ith signal is given as $S^i(t) = S(t) + v^i(t)$. The experimental signal $S^i(t)$ was decomposed by the EMD to get IMF_1^i, so $IMF_1 = \frac{1}{I}\sum_{i=1}^{I} IMF_1^i$, and residual $r_1(t) = S(t) - IMF_1$.

Step 2 Add white noise signals $v^i(t)$ to the residual signal $r_1(t)$, performing i experiments ($i = 1,2,\ldots,I$), and applying EMD to each experiment in order to decompose the signal $r_1^i(t) = x(t) + v^i(t)$ into its first-order component IMF_1^i. $IMF_2^i = \frac{1}{I}\sum_{i=1}^{I} IMF_1^i$, and residual $r_2(t) = S(t) - IMF_2$.

Step 3 Continue the decomposition procedure as previously explained to get the *IMF* component that satisfies the conditions and their respective residuals. The program ends if the residual has monotonous functions and cannot be decomposed via EMD. $S(t) = \sum_{i=1}^{n} IMF_i + r_n(t)$ can be used to represent the original signal.

RECURRENT NEURAL NETWORK MODELS

The memory-equipped neural network LSTM is well suited for processing and forecasting important time-series events. It addresses the issue of losing long-term historical information by introducing gate mechanisms into the RNN architecture. The BiLSTM merges forward and backward LSTMs, accommodating data from both directions and concatenating predictions. While LSTM handles data sequentially in one direction, BiLSTM incorporates reverse-directional LSTMs, enhancing its ability to capture patterns often overlooked by standard LSTMs. The CNN combines convolutional and pooling layers. Convolutional layers extract key features using kernels, while Max Pooling retains strong features, reducing complexity and overfitting. Parameter sharing in convolution reduces optimization parameters, enhancing training efficiency and scalability and making it beneficial for tasks like time-series prediction and image recognition (Gu et al. 2018).

ATTENTION MECHANISM

Human visual observation employs the attention mechanism to filter valuable data and disregard irrelevant information. This study employs the attention mechanism following LSTM, BiLSTM, and CNN layers. These layers initially process the input sequence, after which the attention mechanism assigns weights to hidden states based on their significance.

HYPERPARAMETER OPTIMIZATION

Optuna is an open-source Python library that streamlines hyperparameter optimization, thereby improving machine learning model performance (Akiba et al. 2019). Optuna excels at parallel trial execution, efficiently exploring the hyperparameter space, and supports early stopping to save time, resources and prevent overfitting. Optuna's flexibility, efficiency, and user-friendly nature drove its selection for automating optimal hyperparameter search.

THE ARCHITECTURE AND WORKFLOW OF THE PROPOSED APPROACH

We employ the novel CEEMDAN-GE-OPT-LBC method to forecast FX realized volatility. The comprehensive forecasting procedure is outlined as follows:

Step 1. The realized volatility is decomposed into multiple modes or IMFs using the CEEMDAN technique.
Step 2 The GARCH and EGARCH models will be estimated using logarithmic returns. Subsequently, to create the three hybrid models, namely LSTM, BiLSTM, and CNN neural networks, each will be separately fed with the provided GARCH/EGARCH volatility estimates and IMFs.
Step 3 Efficient Optuna optimization will be employed to fine-tune the hyperparameters of these models through a series of experiments. The hyperparameter ranges are set as [16, 128] for the number of neurons, [32, 128] for batch size, and [0.1, 0.3] for dropout probability.
Step 4 The final prediction results are derived by averaging the predictions generated by the three hybrid models.
Step 5 Finally, we examine the efficacy of the suggested model on the test dataset by comparing the forecasted volatility with the realized volatility values. Evaluate the forecasting performance across different time horizons, 1-day, 5-day, and 10-day, for both the proposed model and benchmark models.

The effectiveness of our approach represented as CEEMDAN-GE-Optuna-ALSTM-ABiLSTM-ACNN (CEEMDAN-GE-OPT-LBC) is validated through comprehensive comparisons with a variety of models including single DL models (LSTM, BiLSTM, CNN), hybrid DL models (GE-ALSTM, GE-ABiLSTM, GE-ACNN), hybrid ensemble DL models (GE-ALSTM-ABiLSTM-ACNN (GE-LBC), CEEMDAN-GE-ALSTM-ABiLSTM-ACNN (CEEMDAN-GE-LBC)), and decomposed hybrid ensemble optimized DL models (CEEMDAN-GE-PSO-ALSTM-ABiLSTM-ACNN (CEEMDAN-GE-PSO-LBC)).

The experiments are implemented on a personal computer with an Intel® Core™ i3-8130U CPU, 8 GB RAM and Windows 11 64-bit operating system. Python 3.9.12 is used as the programming language throughout the analysis. The Python libraries utilized in these experiments include Pandas, NumPy, Keras, TensorFlow, and Optuna (Figure 1).

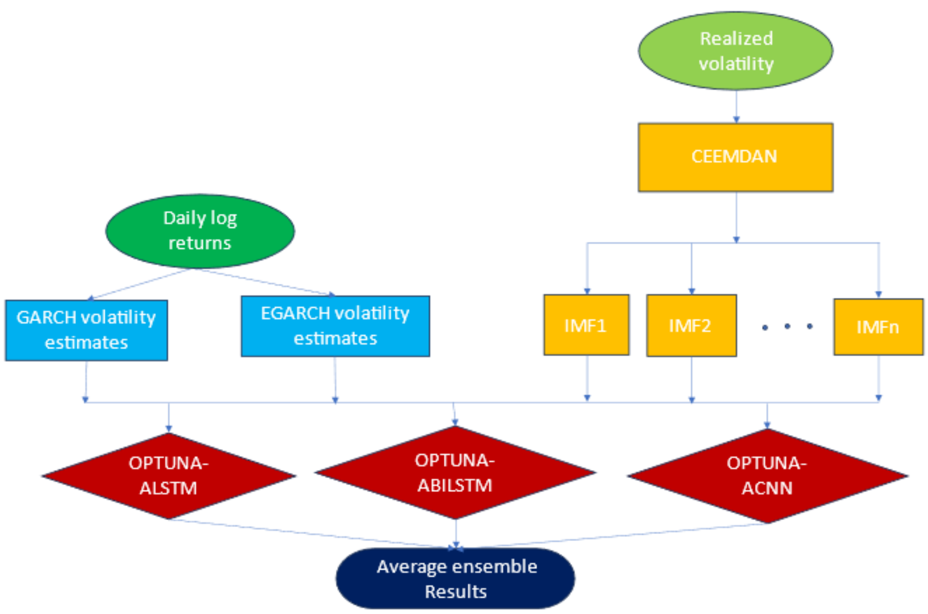

FIGURE 1. CEEMDAN-GE-OPT-LBC ensemble framework

DATA DESCRIPTION AND PERFORMANCE EVALUATION CRITERIA

We used Pakistan rupee to US dollar (PKR/USD) price data from (https://www.investing.com/), covering the period from January 4, 2000 to June 30, 2023. In a daily currency exchange rate series, let P_t represent the price then the log return r_t is defined as the first difference in the logarithm of the exchange rate.

$$r_t = log\, P_t - \log P_{t-1} \tag{1}$$

Realized volatility is a measure that quantifies the fluctuations in returns over a specific trading period. It provides a more accurate assessment of the variability in an asset's returns. To evaluate the performance of our forecasts against an expected target, we will use the realized volatility as the target feature for comparing the predicted output.

$$\nu_t = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(r_t - \bar{r}_t)^2} \tag{2}$$

Here r_t denotes the daily logarithmic returns of the asset, while n represents the number of days for calculating volatility (set at 7 past days in this study). $\bar{r}_t$ denotes the mean return of the asset over n days and v_t refers to the realized volatility of the asset. To reduce the influence of noise and enhance the optimization process, we normalize the data within the range of [0,1] as illustrated in Equation (3):

$$\tilde{X}_t = \frac{X_t - X_{min}}{X_{max} - X_{min}} \tag{3}$$

where $\tilde{X}_t$ denotes the normalized value at time t. X_{max} and X_{min} are the maximum and the minimum true value of the time series, respectively.

LOSS FUNCTIONS

Prediction model performance will be evaluated using various error measures, including mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and quasi-likelihood (Q-LIKE), along with the MCS method. The Q-LIKE loss function, robust to microstructure, incorporates an asymmetric property that penalizes under-prediction more than over-prediction, enhancing risk management considerations. Besides, nonlinear assessment necessitates nonlinear loss metrics aligned with the volatility series characteristics. To this end, due to suitability, heteroscedasticity-adjusted mean absolute error (HMAE) and heteroscedasticity-adjusted mean squared error (HMSE) are chosen.

$$MSE = \frac{1}{N}\sum_{t=1}^{N}(\nu_t - \hat{\sigma}_t^2)^2 \qquad HMSE = \frac{1}{N}\sum_{t=1}^{N}\left(1 - \frac{\hat{\sigma}_t^2}{\nu_t}\right)^2$$

$$MAE = \frac{1}{N}\sum_{t=1}^{N}|\nu_t - \hat{\sigma}_t^2| \qquad HMAE = \frac{1}{N}\sum_{t=1}^{N}\left|1 - \frac{\hat{\sigma}_t^2}{\nu_t}\right|$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\nu_t - \hat{\sigma}_t^2)^2} \qquad QLIKE = \frac{1}{N}\sum_{t=1}^{N}(\log(\hat{\sigma}_t^2) + \frac{\nu_t}{\hat{\sigma}_t^2})$$

where $\hat{\sigma}_t^2$ is the predicted realized volatility at time t, v_t is the actual realized volatility at time t, and N denotes the total count of observations in the forecasts.

Recent research has focused on evaluating models to determine their relative performance, following the framework introduced by Diebold and Mariano (1995). In contrast, the MCS test proposed by Hansen, Lunde and Nason (2011) for comparing volatility models does not require a predefined benchmark. The MCS test systematically evaluates models until the chosen confidence level cannot reject the null hypothesis of equal performance. Each examined model receives p-values from the MCS procedure.

RESULTS AND DISCUSSION

Table 1 displays the descriptive statistics, which encompass measures such as the mean, standard deviation, skewness, and kurtosis. Additionally, it provides the results of the stationary test of augmented Dickey-Fuller (ADF) and the normality test of Jarque-Bera (JB) for the data. The JB test shows that the daily log returns and volatility of the FX rates in this study display tails that are heavier than those of a normal distribution, suggesting a non-normal distribution of model errors. Furthermore, both time series exhibit stationarity at the 1% significance level when tested using the ADF test.

Figure 2 shows the graph of time-series data for daily prices, returns, squared returns and realized volatility of PKR/USD data. Figure 2(a) depicts the daily PKR/USD prices, highlighting an upward trend in the exchange rate. Figure 2(b) displays the log-returns of PKR/USD, showing high returns clustering near increasing returns and low returns gathering near low returns - a phenomenon known as 'volatility clustering' in financial time series. Figure 2(c) presents squared log returns of PKR/USD, showing significant spikes in volatility through large peaks. Lastly, Figure 2(d) depicts the realized volatility of PKR/USD returns, computed using Equation (2). Notably, the pronounced spike in variance starting in 2008 corresponds to the financial crisis, aligning with the increased clustering of returns during that period. In essence, the behavior of logarithmic returns, squared log returns, and realized volatility all exhibit similar patterns.

Figure 3 demonstrates the decomposition of the realized volatility of PKR/USD into ten intrinsic modes functions (IMFs), ranging from high to low frequency. These ten intrinsic modes are denoted as IMF_1 to IMF_{10}, with IMF_1 representing the mode with the highest frequency and IMF_{10} representing the mode possessing the lowest frequency. It can be observed that the decomposed intrinsic modes all exhibit different periodicity.

The experimental procedure consists of two parts, outlined as follows: In the first part, we evaluate the effectiveness of CEEMDAN and GARCH/EGARCH estimates in enhancing the forecasting performance of neural networks. We compare the predictability of realized volatility between individual models and a hybrid model for *n*-day-ahead predictions. In the second part, we assess the efficacy of the hyperparameter tuning method Optuna in improving the forecasting results for realized volatility by comparing the prediction performances of CEEMDAN-GE-LBC, CEEMDAN-GE-OPT-LBC, and CEEMDAN-GE-PSO-LBC.

Table 2 provides an overview of the forecast errors observed in out-of-sample testing for both individual deep learning models and hybrid deep learning models. The experimental results consistently demonstrate that our proposed ensemble hybrid model, CEEMDAN-GE-OPT-LBC, outperforms all benchmark models across all assessed scenarios. Table 2 shows that incorporating GARCH and EGARCH volatility forecasts as inputs in DL models has noticeably improved the predictive performance of individual DL models. Additionally, the attention mechanism can automatically learn the importance of variables in the training process and make the most critical input variables play a leading role by weighing each input variable. Besides, ensemble results indicate enhanced forecasting performance compared to standalone models. For example, when comparing the RMSE values of GE-LSTM and GE-BiLSTM (0.02728 and 0.02631) with those of GE-LBC (0.02529), it is evident that GE-LBC improves the results. However, GE-CNN provides the lowest RMSE values compared to GE-LBC.

To further analyze the impact of decomposed IMFs on the model's prediction performance, we conducted a performance comparison by introducing decomposed intrinsic modes into DL models. As displayed in Table 2, including IMFs as input in the model led to a substantial improvement in prediction performance across all evaluation criteria, enhancing the model's fitting performance.

Our final experiment enhanced the best model's performance by utilizing the Optuna framework to fine-tune hyperparameters for LSTM, BiLSTM, and CNN networks. Optuna produced the most favorable evaluation metrics, as demonstrated in Table 2. Our Optuna-driven approach surpassed the utilization of PSO for hyperparameter tuning within the hybrid ensemble model, as exemplified by a reduced RMSE value of 0.01528 compared to the RMSE value of 0.01741 of PSO. This difference highlights the superior performance of Optuna over PSO. When we compared the time taken by each optimization method, we observed that PSO consumed more time than Optuna, with a time difference of 947.23 seconds noted between the two optimization approaches. As a result, Optuna delivers promising results and exhibits efficiency in terms of time consumption, making it a preferable choice for hyperparameter tuning.

Our proposed model achieved the lowest values for all evaluation metrics, including MAE, MSE, RMSE, HMAE, HMSE, and Q-LIKE. As a result, our innovative approach significantly enhances the accuracy of predicting FX volatility. The graph in Figure 4 displays the volatility forecasted by the hybrid ensemble models compared to the realized volatility, which serves as our research's target value. The graphs provide visual confirmation of the proposed model's superior predictive performance compared to other benchmark models. It is evident that the proposed model brings substantial enhancements to volatility peaks, successfully projecting more consistent patterns of growth and decline during both pre and post-peak periods.

Additionally, we conducted predictive experiments for different time horizons (1, 5, and 10 days ahead) to assess the variation in forecasting accuracy at different prediction intervals. The outcomes of these experiments are briefly presented in Table 3. Remarkably, our proposed CEEMDAN-GE-OPT-LBC models consistently exhibit the lowest errors, demonstrating their robustness across varying prediction horizons (1, 5, and 10 days ahead). Table 3 illustrates the consistent superiority of our suggested method over other benchmark forecasting models, as evidenced by metrics such as MSE, MAE, and Q-LIKE. The analysis shows that a standalone CNN model outperforms LSTM and BiLSTM models when forecasting one-day-ahead volatility. However, when predicting volatility five days and ten days in advance, the CNN model did not perform best compared to the LSTM and BiLSTM models. This observation suggests that CNNs excel at capturing short-term, localized patterns within volatility data, rendering them particularly effective for one-day-ahead forecasts. In contrast, their performance may decrease when faced with longer-term dependencies. Furthermore, the results show a general trend of increasing errors in all existing models as the forecasting horizon extends.

TABLE 1. Descriptive statistics for input variables

	Count	Mean	SD	Skewness	Kurtosis	ADF	JB
Log returns	5857	0.020065	0.268964	0.376989	5.640385	-10.53958***	7902.66***
Realized volatility	5857	0.194972	0.156235	1.394222	1.789718	-6.07102***	384.43***

***denote a rejection of the null hypothesis at the 1% significance level

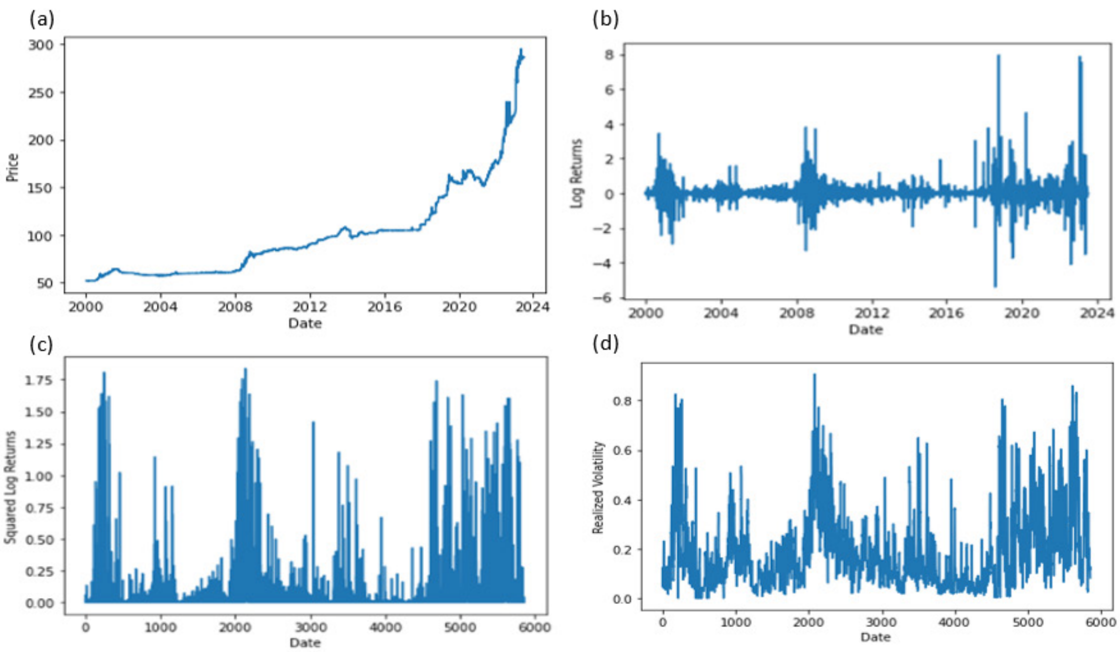

FIGURE 2. The graph of time-series data for daily prices (a), log returns (b), squared log returns (c), and realized volatility (d) of PKR/USD data from 2000 to 2023

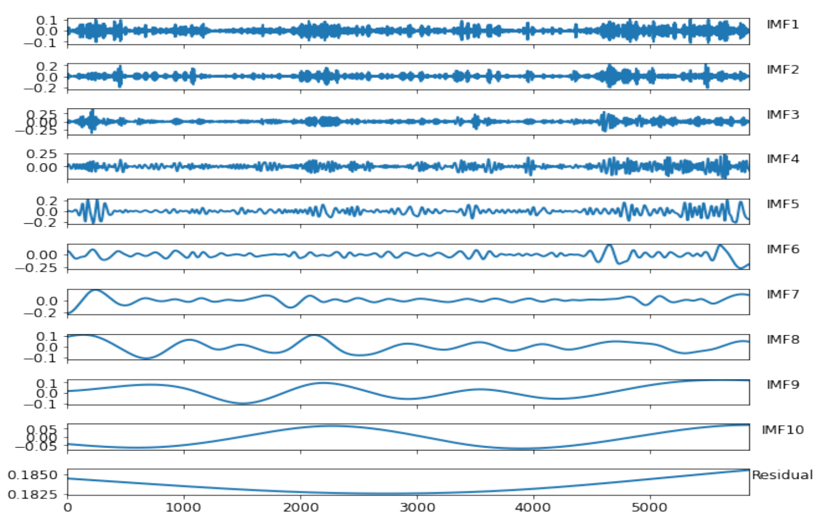

FIGURE 3. Realized volatility decomposition using CEEMDAN

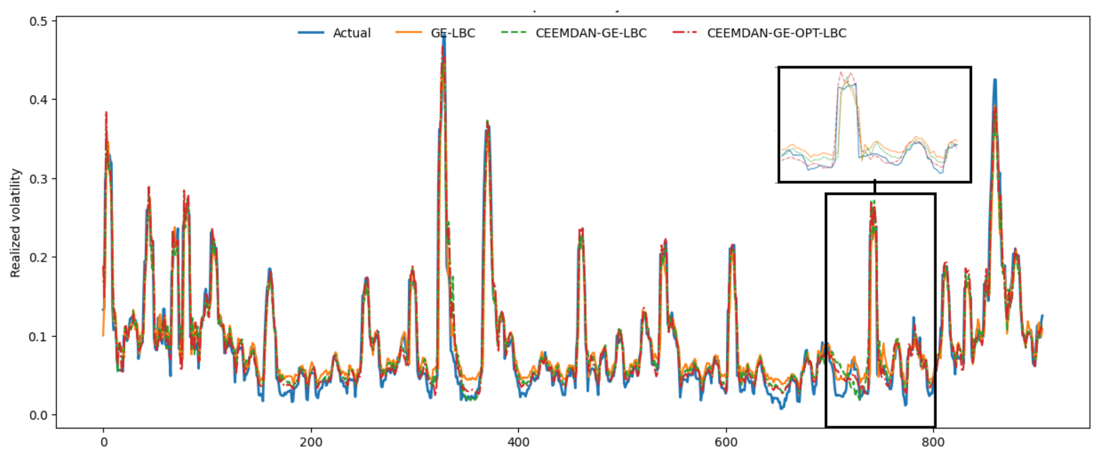

FIGURE 4. A comparison of realized volatility forecast results: Proposed model VS benchmark models

TABLE 2. Results of the performance evaluation for different forecasting models utilizing varied error functions

Models	MAE	MSE	RMSE	HMAE	HMSE	Q-LIKE
LSTM	0.02383	0.00094	0.03065	0.28448	0.13934	-1.57231
BiLSTM	0.03308	0.00150	0.03875	0.34676	0.17344	-1.53284
CNN	0.01520	0.00057	0.02386	0.21658	0.10812	-1.61761
GE-LSTM	0.01927	0.00074	0.02728	0.24467	0.12672	-1.59669
GE-BiLSTM	0.01664	0.00069	0.02631	0.21745	0.13076	-1.60940
GE-CNN	0.01607	0.00059	0.02421	0.21626	0.11647	-1.60836
GE-LBC	0.01675	0.00064	0.02529	0.22233	0.12134	-1.60601
CEEMDAN-GE-LBC	0.01341	0.00037	0.01914	0.17592	0.05932	-1.62740
CEEMDAN-GE-OPT-LBC	**0.01067**	**0.00023**	**0.01528**	**0.14392**	**0.03968**	**-1.63882**
CEEMDAN-GE-PSO-LBC	0.01239	0.00030	0.01741	0.16492	0.04890	1.63183

Bold values highlight the results of the best-performing model

We further investigate how well the competing models perform in forecasting out-of-sample data using the MCS test over alternative multi-step forecast horizons of 1, 5, and 10 days. A larger 'MCS *p*-value' means higher prediction accuracy for a given model. Table 4 reports the empirical findings from the MCS test, confirming the robustness of our conclusions. Across all loss function criteria, it is evident that our proposed model successfully passes the MCS test. When compared to alternative models, our proposed model has the highest *p*-value of 1, indicating that it is superior to all other models. Based on the MAE criterion, the proposed model seems to be the best, followed by the CEEMDAN-GE-LBC, GE-LBC, and CNN models. The MCS *p*-value for CEEMDAN-GE-OPT-LBC consistently remains at 1 across all forecasting horizons. This result validates that the multi-scale hybrid model, which combines CEEMDAN-GE with ensemble DL models optimized using Optuna, demonstrates greater robustness compared to the other models.

In summary, our study represents an initial exploration into the integration of Optuna with an ensemble model and the CEEMDAN decomposition method. The approach we propose demonstrates superior performance over all benchmark models across various forecasting horizons (1, 5, and 10-day ahead). Additionally, our findings suggest that Optuna is a more effective algorithm for hyperparameter tuning when compared to PSO. These outcomes further validate the efficacy of our approach in predicting future realized PKR/USD volatility.

Tripathi, Kumar and Inani (2020) introduced an ensemble method for predicting daily exchange rates using a combination of linear and nonlinear time-series

TABLE 3. Comparison of out-of-sample forecast errors for PKR/USD exchange rate volatility across different forecast horizons

	1-day ahead prediction			5-day ahead prediction			10-day ahead prediction		
Models	MAE	MSE	QLIKE	MAE	MSE	QLIKE	MAE	MSE	QLIKE
LSTM	0.0238	0.0009	-1.5723	0.0557	0.0047	-1.4297	0.1352	0.0312	-0.2060
BiLSTM	0.0330	0.0015	-1.5328	0.0541	0.0045	-1.4370	0.1350	0.0314	-0.2038
CNN	0.0152	0.0005	-1.6176	0.0558	0.0048	-1.4333	0.1346	0.0313	-0.2037
GE-LSTM	0.0192	0.0007	-1.5966	0.0521	0.0044	-1.4435	0.1334	0.0299	-0.2254
GE-BiLSTM	0.0166	0.0006	-1.6094	0.0495	0.0043	-1.4510	0.1334	0.0303	-0.2155
GE-CNN	0.0160	0.0005	-1.6083	0.0534	0.0044	-1.4401	0.1330	0.0298	-0.2198
GE-LBC	0.0167	0.0006	-1.6060	0.0514	0.0043	-1.4458	0.1330	0.0298	-0.2221
CEEMDAN-GE-LBC	0.0134	0.0003	-1.6274	0.0421	0.0030	-1.4949	0.1306	0.0285	-0.2382
CEEMDAN-GE-OPT-LBC	**0.0106**	**0.0002**	**-1.6388**	**0.0345**	**0.0026**	**-1.5291**	**0.1294**	**0.0274**	**-0.2550**
CEEMDAN-GE-PSO-LBC	0.0124	0.0003	-1.6318	0.0373	0.0026	-1.5192	0.1295	0.0284	-0.2417

Bold values highlight the results of the best-performing model

TABLE 4. MCS test results with 1-day, 5-day and 10-day ahead forecast horizons

	1-day	5-days	10-days
Models	MAE	MAE	MAE
LSTM	0.954	0.557	0.981
BiLSTM	0.950	0.510	0.981
CNN	0.967	0.557	0.985
GE-LSTM	0.961	0.712	0.985
GE-BiLSTM	0.967	0.764	0.985
GE-CNN	0.967	0710	0.989
GE-LBC	0.967	0.712	0.989
CEEMDAN-GE-LBC	0.987	0.900	0.989
CEEMDAN-GE-OPT-LBC	1.000	1.000	1.000
CEEMDAN-GE-PSO-LBC	0.987	0.903	1.000

The numbers are the MCS *p*-values

forecasting methods, including mean forecast, ARIMA and neural networks. Their results suggest that this approach outperformed individual component models. Additionally, ensemble methods, which combine multiple base LSTM models, have demonstrated reduced variance and enhanced performance in related stock market prediction work by Borovkova and Tsiamas (2019). Our research findings align with these studies, as we constructed an ensemble model by combining LSTM, BiLSTM, and CNN models with various input variables.

Baffour, Feng and Taylor (2019) examined five major currency pairs employing a hybrid model that integrated a neural network model with an econometric model. Their results indicated that the hybrid model exhibited superior performance compared to standard GARCH and other GARCH-type models, significantly enhancing accuracy in forecasting exchange rate volatility. Existing literature also supports the idea that integrating neural network models with econometric models yields favorable outcomes. As a result, we developed our model by merging econometric and deep learning models to improve FX volatility prediction.

Li et al. (2021) introduced the VMD-BiLSTM model for crude oil forecasting, using VMD to break down historical data and BiLSTM for prediction. Their results show that this approach outperforms other models. Our research found that applying CEEMDAN to decompose PKR/USD volatility substantially improves our model's forecasting accuracy. Pradeepkumar and Ravi (2017) introduced the PSOQRNN model, a Quantile Regression Neural Network trained using PSO, for forecasting volatility in financial time series. They conducted a comparative analysis of PSOQRNN against various models and found that, for the majority of the eight financial time series examined, PSOQRNN outperformed the other models. This underscores the role of optimization methods in enhancing prediction model performance.

There is a noticeable absence of research on FX time series volatility forecasting models using the Optuna optimization method in the existing literature. In comparison to the PSO approach, our results clearly indicate that Optuna outperforms the PSO method, highlighting its remarkable efficacy. Flexible optimization settings, smooth result integration, early low-performing trial pruning for computational efficiency, and efficient hyperparameter space exploration are all advantages of Optuna. Our research marks the first attempt to forecast PKR/USD volatility, achieved through an ensemble of models that consider a diverse set of input variables and the novel optimization method known as Optuna.

CONCLUSION

This study explores the use of an innovative hybrid neural network model to forecast volatility in the PKR/USD currency exchange rates. Our proposed methodology will utilize the CEEMDAN-GE-OPT-LBC ensemble deep learning method. This method combines three distinct deep learning models, LSTM, BiLSTM, and CNN, with an attention mechanism, as well as the CEEMDAN decomposition method and Optuna for hyperparameter tuning. The attention mechanism has demonstrated the ability to dynamically select the most useful features and accelerate model convergence. The CEEMDAN approach divides FX volatility into distinct subsequences defined by different IMFs. We integrated the GARCH and EGARCH models to enhance FX volatility predictions with three deep learning models. Our results show that a hybrid model incorporating IMFs and GARCH/EGARCH estimates beats the performance of a single DL model. We employed the Optuna algorithm to identify the hyperparameters for optimizing the DL models. Utilizing Optuna for hyperparameter tuning led to improved prediction accuracy compared to the benchmark models.

Additionally, our findings indicate that the Optuna approach surpasses the PSO method in improved accuracy and reduced time consumption. We then introduce an ensemble of three deep learning models as the predictive tool, amalgamating diverse quantitative model inputs such as GARCH/EGARCH estimates and IMFs. The results affirm that the ensemble model surpasses individual models when exposed to varied data inputs. Overall, the innovative combination model presented in this paper demonstrates superiority and robustness in FX volatility prediction. Furthermore, the proposed model consistently yields the lowest loss function values across various prediction time horizons (1, 5, and 10 days). The proposed model can enhance monetary policy decision-making, improve exchange rate stability, and provide valuable support for risk management in international trade and investments.

REFERENCES

Abedin, M.Z., Moon, M.H., Hassan, M.K. & Hajek, P. 2021. Deep learning-based exchange rate prediction during the COVID-19 pandemic. *Annals of Operations Research*. https://doi.org/10.1007/s10479-021-04420-6

Akiba, T., Sano, S., Yanase, T., Ohta, T. & Koyama, M. 2019. Optuna: A next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. pp. 2623-2631. https://doi.org/10.1145/3292500.3330701

Baffour, A.A., Feng, J. & Taylor, E.K. 2019. A hybrid artificial neural network-GJR modeling approach to forecasting currency exchange rate volatility. *Neurocomputing* 365: 285-301. https://doi.org/10.1016/j.neucom.2019.07.088

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

Borovkova, S. & Tsiamas, I. 2019. An ensemble of LSTM neural networks for high-frequency stock market classification. *Journal of Forecasting* 38(6): 600-619. https://doi.org/10.1002/for.2585

Colominas, M.A., Schlotthauer, G. & Torres, M.E. 2014. Improved complete ensemble EMD: A suitable tool for biomedical signal processing. *Biomedical Signal Processing and Control* 14: 19-29. https://doi.org/10.1016/j.bspc.2014.06.009

Dauphin, Y.N., Fan, A., Auli, M. & Grangier, D. 2017. Language modeling with gated convolutional networks. *ICML '17: Proceedings of the 34th International Conference on Machine Learning* 70: 933-941.

Dhamija, A.K. & Bhalla, V.K. 2010. Financial time series forecasting: Comparison of various arch models. *Global Journal of Finance and Management* 2(1): 159-172. http://www.ripublication.com/gjfm.htm

Diebold, F.X. & Mariano, R.S. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* 13(3): 253-263. https://doi.org/10.1198/073500102753410444

Dragomiretskiy, K. & Zosso, D. 2013. Variational mode decomposition. *IEEE Transactions on Signal Processing* 62(3): 531-544. https://doi.org/10.1109/TSP.2013.2288675

Engle, R.F. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50(4): 987-1008. https://doi.org/10.2307/1912773

Gamboa, D.A. 2017. A comparative study of volatility forecasting methods. *The European Journal of Finance* 23(7-9): 663-681.

Granata, F. & Di Nunno, F. 2021. Forecasting evapotranspiration in different climates using ensembles of recurrent neural networks. *Agricultural Water Management* 255: 107040. https://doi.org/10.1016/j.agwat.2021.107040

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J. & Chen, T. 2018. Recent advances in convolutional neural networks. *Pattern Recognition* 77: 354-377. https://doi.org/10.1016/j.patcog.2017.10.013

Hajizadeh, E., Seifi, A., Zarandi, M.F. & 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

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

Henrique, B.M., Sobreiro, V.A. & Kimura, H. 2019. Literature review: Machine learning techniques applied to financial market prediction. *Expert Systems with Applications* 124: 226-251. https://doi.org/10.1016/j.eswa.2019.01.012

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

Iqbal, F., Zahid, M. & Koutmos, D. 2023. Cryptocurrency trading and downside risk. *Risks* 11(7): 122. https://doi.org/10.3390/risks11070122

Jung, G. & Choi, S.Y. 2021. Forecasting foreign exchange volatility using deep learning autoencoder-LSTM techniques. *Complexity* https://doi.org/10.1155/2021/6647534

Kalchbrenner, N., Espeholt, L., Simonyan, K., Oord, A.V.D., Graves, A. & Kavukcuoglu, K. 2016. Neural machine translation in linear time. *ArXiv preprint ArXiv*: 1610.10099. https://doi.org/10.48550/arXiv.1610.10099

Kastner, G., Frühwirth-Schnatter, S. & Lopes, H.F. 2017. Efficient Bayesian inference for multivariate factor stochastic volatility models. *Journal of Computational and Graphical Statistics* 26(4): 905-917. https://doi.org/10.1080/10618600.2017.1322091

Kausar, R., Iqbal, F., Raziq, A. & Sheikh, N. 2023. A hybrid approach for accurate forecasting of exchange rate prices using VMD-CEEMDAN-GRU-ATCN model. *Sains Malaysiana* 52(11): 3293-3306. http://doi.org/10.17576/jsm-2023-5211-20

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: 25-37. https://doi.org/10.1016/j.eswa.2018.03.002

Li, Y., Jiang, S., Li, X. & Wang, S. 2021. The role of news sentiment in oil futures returns and volatility forecasting: Data-decomposition based deep learning approach. *Energy Economics* 95: 105140. https://doi.org/10.1016/j.eneco.2021.105140

Pradeepkumar, D. & Ravi, V. 2017. Forecasting financial time series volatility using particle swarm optimization trained quantile regression neural network. *Applied Soft Computing* 58: 35-52. https://doi.org/10.1016/j.asoc.2017.04.014

Pradeepkumar, D. & Ravi, V. 2018. Soft computing hybrids for FOREX rate prediction: A comprehensive review. *Computers & Operations Research* 99: 262-284. https://doi.org/10.1016/j.cor.2018.05.020

Risse, M. 2019. Combining wavelet decomposition with machine learning to forecast gold returns. *International Journal of Forecasting* 35(2): 601-615. https://doi.org/10.1016/j.cor.2018.05.020

Santhosh, M., Venkaiah, C. & Kumar, D.V. 2019. Short-term wind speed forecasting approach using ensemble empirical mode decomposition and deep Boltzmann machine. *Sustainable Energy, Grids and Networks* 19: 100242. https://doi.org/10.1016/j.segan.2019.100242

Sobri, S., Koohi-Kamali, S. & Rahim, N.A. 2018. Solar photovoltaic generation forecasting methods: A review. *Energy Conversion and Management* 156: 459-497. https://doi.org/10.1016/j.enconman.2017.11.019

Taylor, S.J. 1994. Modeling stochastic volatility: A review and comparative study. *Mathematical Finance* 4(2): 183-204. https://doi.org/10.1111/j.1467-9965.1994.tb00057.x

Torres, M.E., Colominas, M.A., Schlotthauer, G. & Flandrin, P. 2011. A complete ensemble empirical mode decomposition with adaptive noise. *Proceedings of IEEE International Conference on Acoustics.* https://10.1109/ICASSP.2011.5947265

Tripathi, M., Kumar, S. & Inani, S.K. 2020. Exchange rate forecasting using ensemble modeling for better policy implications. *Journal of Time Series Econometrics* 13(1): 43-71. https://doi.org/10.1515/jtse-2020-0013

Wu, Z.H. & Huang, N.E. 2009. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis* 01(01): 1-41. https://doi.org/10.1142/S1793536909000047

Zahid, M. & Iqbal, F. 2020. Modeling the volatility of cryptocurrencies: An empirical application of stochastic volatility models. *Sains Malaysiana* 49(3): 703-712. http://dx.doi.org/10.17576/jsm-2020-4903-25

Zahid, M., Iqbal, F. & Koutmos, D. 2022. Forecasting bitcoin volatility using hybrid GARCH models with machine learning. *Risks* 10(12): 237. https://doi.org/10.3390/risks10120237

*Corresponding author; email: fsmuhammad@iau.edu.sa