Prediction of the prices of digital cryptocurrencies in a high-frequency data world

Chenggang Li, Xuan Wang, Yan Gong, Liang Wu*, Huiyang Li

Abstract

This study predicts the prices of digital cryptocurrencies, which are non-linear, highly volatile, multi-scale, noisy, and, therefore, considerably more difficult to forecast compared to other financial products. Despite the proliferation of cryptocurrency markets, research on digital cryptocurrency prices remains scarce, while the existing research often fails to elucidate the characteristics of digital cryptocurrency prices. We developed CEEMD-LSTM—a hybrid prediction model based on the decomposition-reconstruction-integration framework—to predict digital cryptocurrency price movements. The CEEMD method is applied to decompose prices into high-frequency components, low-frequency components, and trend sequences, and reconstruct the combined IMF components. Thereafter, we apply the LSTM model to predict each component separately and obtain the sorted predicted values of prices by adding and integrating the predicted values of each component. Using four time-intervals of high-frequency data of Bitcoin and Ether, we conduct a systematic comparison of the prediction accuracy between CEEMD-LSTM and other models, single or combined. Results reveal that the prediction effect of the CEEMD-LSTM model consistently outperforms that of other models in the high-frequency data world. These findings enhance the understanding and prediction of digital cryptocurrency price movements. Additionally, they enable investors to manage risks and promote market competitiveness, offer reliable support for decision-making, address uncertainty arising from volatility, facilitate timely strategy adjustments to exploit opportunities and identify potential risks, and provide new methods for risk mitigation.

Keywords: *Digital cryptocurrencies, Complementary Ensemble Empirical Mode Decomposition, Long Short-Term Memory Model, High-frequency data*

JEL Classification: G10, G12, G14

Article history: Received: November 2024; Accepted: August 2025; Published: March 2026

1 INTRODUCTION

Since the launch of Bitcoin in 2009, cryptocurrencies have experienced rapid global growth, attracting widespread attention and concern. Crypto market total capitalization achieved a historic breakthrough in 2021, peaking in November, close to $3 trillion, and ending the year with $2.25 trillion, an annual increase of nearly 200% (TokenInsight, 2022). According to CoinGecko data,1 the global market value of cryptocurrencies is approximately $2.58 trillion,

[1] CoinGecko is an application of crypto market ranking charts designed to provide cryptocurrency investors with the indicators to help them make decisions. It comprehensively analyzes digital currency-related transactions,

https://doi.org/10.7441/joc.2026.01.09

surpassing the amount of U.S. dollars in circulation. Bitcoin, the most widely recognized cryptocurrency, was first proposed by Nakamoto (2008), who described its algorithm in 2008. Bitcoin has the largest market size and user base in the global cryptocurrency market. It also has the greatest volatility among digital cryptocurrencies, with its maximum one-day price fluctuation reaching up to 336.84%. The high degree of uncertainty associated with digital cryptocurrencies and their potential risks to the real economy have attracted significant academic attention. Goczek and Skliarov (2019) argue that Bitcoin prices are highly volatile and driven by numerous factors. Therefore, accurately predicting cryptocurrency prices has become an important research issue. Reliable forecasting techniques can mitigate investors' risks and protect their assets.

However, predicting cryptocurrency prices remains challenging. The fluctuations of cryptocurrency indices are induced by complex factors such as global economic level, country-specific regulatory policies, international speculative capital, and investor sentiment (Wu & Huang, 2009; Yeh et al., 2010; Xie & Shi, 2015). With the development of big data, machine learning has significantly advanced and has been widely applied in financial forecasting. Deep learning, introduced by Hinton and Salakhutdinov (2006), utilizes multi-layer neural networks to establish non-linear reversible mappings from underlying signals to higher-level semantics by simulating the multi-layer abstraction mechanism of the human brain in the learning process. Deep learning demonstrates significant capability in extracting the sample characteristics. However, owing to the presence of substantial noise, high complexity, and multi-dimensionality in financial time series, employing machine learning for forecasting may lead to overfitting. This may hinder the deep learning model from capturing the essential characteristics of the series, declining its performance. Many researchers have attempted to solve this problem by extracting the features with the signal decomposition techniques. The EMD, proposed by Huang et al. (1998), along with its improved variants, can decompose complex signals into the intrinsic mode functions (IMFs). The IMFs effectively characterize complex features such as non-linearity and multi-scale characteristics of time series data and enable deep learning models to improve the forecasting accuracy continuously.

This study combines deep learning models with signal decomposition models and develops a combined model known as the CEEMD-LSTM model based on the decomposition-reconstruction-integration framework. The goal is to capture key features of a digital cryptocurrency price, mitigate potential risks associated with digital cryptocurrencies, and address existing research gaps in cryptocurrency price forecasting. The contributions of this study are threefold. First, we introduce the deep learning algorithm into the prediction of digital cryptocurrency indices, and decompose the original sequence through the CEEMD method to better fit the non-linear and high-noise characteristics of cryptocurrencies prices. Second, we propose a new intrinsic mode function (IMF) recombination method. Specifically, the low-frequency components are restructured and recombined into a new sequence, while the high-frequency components and trend terms are applied by deep learning model to predict cryptocurrency prices. Our approach demonstrates significantly superior predictive performance in cryptocurrency price prediction than other IMF recombination methods. Third, we present a practical and reliable modeling scheme for the prediction of the digital

market capitalization, and other data, and organizes development activities, communities, etc. It ranks digital currencies by development activity, community activity, and trading liquidity.

https://doi.org/10.7441/joc.2026.01.09

cryptocurrency indices. The LSTM model, which exhibits robust generalization capability, is selected as the framework. It is combined with the CEEMD model to construct an effective hybrid prediction model for digital cryptocurrencies prices.

The remainder of this article is structured as follows: Section 2 presents a review of the relevant literature and summarizes existing research findings. Section 3 discusses the construction of the proposed CEEMD-LSTM combined model and introduces the data sources. Section 4 presents the results, utilizing the proposed model to predict the digital cryptocurrencies prices and comparing it with other single models and combined models to highlight the effectiveness of the proposed model. Section 5 is the discussion. Section 6 concludes the study.

2 THEORETICAL BACKGROUND

For financial time series forecasting, traditional econometric models, such as the GARCH, HAR-RV, and ARFIMA models, which are developed from theoretical frameworks (Li & Feng, 2014; Nelson et al., 2017; Mohammadi, 2017; Kim & Fan, 2019), have been widely applied to analyze and forecast stock, bond, futures, and oil markets (Hinton & Salakhutdinov, 2006; Lei et al., 2021; Pan et al., 2019; Salisu et al., 2022; Torres et al., 2011). However, these traditional models have several drawbacks as they require data to satisfy certain statistical assumptions and are not well-suited for handling high-dimension and noisy financial time series. Moreover, when dealing with non-linear, noisy, and high-scale time series such as digital cryptocurrencies prices, the traditional econometric models often yield unsatisfactory forecasting performance because of their limited ability to learn and fit non-linear, high-noise, and high-scale characteristics of such time series (Sarkodie et al., 2022).

In the era of big data, the exponential expansion of data volume, constant development of computing capability, and rise of artificial intelligence have rendered new research methods necessary for time series prediction (He & Chang, 2001; Arel et al., 2010; Hsu, 2016; Herwartz, 2017; Jing et al., 2021). Research shows that machine learning methods have better prediction abilities compared to traditional econometric methods (Huang et al., 1998; Yang & Yang, 2005; Yang & Wang, 2021; Salles et al., 2022). Although the traditional machine learning methods possess robust generalization ability, its prediction effectiveness depends largely on artificial feature design. This presents some challenges in time series prediction, such as over-fitting, under-fitting, slow convergence, etc., which may limit their effectiveness in practical financial applications. Deep learning models, which are based on multi-layer neural network architectures, have attracted substantial attention from the industry because of their suitability for solving complex issues. Common deep learning models include convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks (Bengio, 2009; Bengio, 2013).

High-frequency data analysis has become increasingly important in various fields, including finance, meteorology, and environmental monitoring. For instance, Gao et al. (2024) forecast short-term ride-hailing demand using a CNN-ATTBiLSTM model, demonstrating the effectiveness of deep learning approaches in traffic flow prediction. Similarly, Cui and Dou (2018) and Wu et al. (2024) predicted regional carbon emission trading prices using GA-VMD and CNN-BiLSTM-attention models, showcasing the potential of these techniques in handling complex economic data. Additionally, Long et al. (2024) applied the IPSO-EGA-LSTM model

https://doi.org/10.7441/joc.2026.01.09

to predict Dongting Lake water levels, further illustrating the applicability of deep learning in environmental monitoring.

Signal decomposition techniques are crucial for enhancing the accuracy of predictive models. Chen et al. (2022) developed a short-term wind speed forecasting model based on multi-mode decomposition and sparrow search optimized residual networks, demonstrating the advantages of combining signal decomposition with deep learning. Ding et al. (2024) integrated deep learning with signal decomposition to make reliable predictions of river ice breakup dates in northern cold regions. Moreover, Xue et al. (2009) built an online learning eye movement cognitive hierarchy intelligent recognition model using LSTM neural networks, highlighting the versatility of signal decomposition technology in analyzing complex visual data.

Deep learning models have been extensively applied across various domains. Li et al. (2019) employed deep learning to forecast the short-term power output of wind farm clusters, demonstrating its potential in renewable energy sectors. Zhang and Ding (2021) showed the effectiveness of BERTopic and LSTM models in predicting emerging topics. Qian et al. (2024) further expanded the applications of deep learning by proposing a multidimensional satisfaction quantification approach for public library users through a hybrid Word2Vec and LSTM model. Furthermore, Yao et al. (2017) studied global stock market extreme risk spillovers using Transformer-LSTM, while Xu (2025) predicted AI research theme evolution in the LIS discipline by integrating BERTopic and LSTM.

In summary, studies in high-frequency data analysis, signal decomposition technology, and deep learning models are rapidly evolving and mutually reinforcing. The integration of these three aspects holds significant potential for addressing more complex issues and improving prediction accuracy (Su et al., 2017; Werner & Esteban, 2017; Tang et al., 2021; Tian & Qin, 2024; Jiang & Wang, 2024; Liu et al., 2024; Ji et al., 2025; Zhai et al., 2025). Researchers could consider combining these technologies to develop more sophisticated models capable of handling vast and intricate datasets. Relevant examples include Su et al. (2024), who researched sandstorm weather prediction models in Northwest China using GCN-LSTM integrated with spatial-temporal features, and Jiang et al. (2023), who constructed financial time series forecasting models utilizing integrated decomposition and deep learning techniques.

Owing to its embedded feedback and loop structure, the RNN can partially address the dependence between short time-series data (Graves et al., 2013). However, as data volume increases, RNN experiences significant gradient disappearance or gradient explosion issue. The LSTM model, proposed by Hochreiter and Schmidhuber (1997), can address these challenges. As an enhancement to the RNN model, the LSTM model outperforms the traditional RNN model in terms of time series prediction. The issue of gradient scattering or gradient blast is resolved by incorporating memory units and forgetting units into the network architecture. Numerous studies have demonstrated the advantages of the LSTM model in financial price prediction. Nelson et al. (2017) proposed an LSTM-based stock price prediction model and suggested that it was less risky than other models in the trading simulation. Yang and Wang (2019) applied the LSTM, support vector regression (SVR), multilayer perceptron (MLP), and autoregressive integrated moving average (ARIMA) models to predict and analyze the stock indices of 30 countries worldwide. They indicated that the LSTM model had wide applicability and prediction superiority. Jing et al. (2021) utilized the LSTM neural network method to predict the stock price and found that this method had better effectiveness than the compared

https://doi.org/10.7441/joc.2026.01.09

models. As the transaction price of digital cryptocurrency is also financial time series data, the LSTM model has certain superiorities in predicting complex time-series data. Therefore, we chose the LSTM model as the theme framework in this study.

As financial markets are sophisticated systems influenced by numerous exogenous factors and financial time series are inherently noisy, a growing body of literature suggests that a single deep learning model may struggle to extract key characteristics of complex sequences. Combining the existing forecasting methods and decomposing the time series data to enhance forecasting accuracy remains challenging. Therefore, researchers integrate financial measurement, signal engineering, and other disciplines into machine learning models, hoping to achieve accurate financial time series prediction. Fourier transform, wavelet transform, and other analysis methods have been applied to financial time series (Herrera et al., 2019; Shen et al., 2015; Yu & Chen, 2017). However, Fourier transform struggles to deal with data that changes abruptly, and wavelet analysis has non-adaptive defects. Given the high volatility and noise characteristics of digital cryptocurrency prices, Fourier transform and wavelet analysis are not productive in accurately predicting such financial time series.

Huang et al. (2015) proposed empirical mode decomposition (EMD), which can adaptively declare non-linear signals as intrinsic mode functions (IMFs) and efficiently damp out continuous noise. However, EMD cannot adequately handle intermittent or mixed noise. Therefore, Wu and Huang (2009) proposed an improved EMD method, which is ensemble empirical mode decomposition (EEMD). EEMD mitigates the issue induced by noise mixing. With the help of EEMD, white noise is overlaid on the original signal and then disaggregated into multiple IMF sets. The average of the respective IMF sets is assumed to be the proper result. EEMD removes noise mixing by separating noise from primitive signal components in different IMF organizations. However, the EEMD method requires plentiful averaging time to decrease the Gaussian white noise, which significantly increases the computing time.

To resolve the challenges of EEMD, Yeh et al. (2010) introduced complementary EEMD (CEEMD). CEEMD is a sophisticated approach, involving the addition of a pair of contrasting white noise signals to the initial signal, followed by separate EMD factorization. CEEMD ensures that the factorization is at least equivalent to EEMD and reduces reconfiguration errors induced by white noise. Torres et al. (2011) further improved CEEMD and proposed the CEEMD with adaptive noise (CEEMDAN), which eliminated the issues of incomplete decomposition and large reconstruction error of EEMD. Compared with the CEEMDAN method, CEEMD has fewer high-frequency components, which can significantly reduce the forecasting difficulty of the LSTM model. To mitigate noise-induced interference noise in practical applications, many researchers have combined empirical mode decomposition with other models to increase the effectiveness of the original single model. Li and Feng (2014) reconstructed the IMF components and decomposed the stock index price series and the investor sentiment based on EEMD, thus confirming that investor sentiment and stock index price fluctuation show different fluctuation relationships at different time scales. Yang and Lin (2017) applied EMD to decompose the exchange rate and input the obtained IMF into the extreme learning machine prediction model, thus improving the accuracy of exchange rate prediction. To provide better forecasting results to stock market participants and researchers, Lin et al. (2021) proposed a modified modelling procedure by combining the EEMD with

https://doi.org/10.7441/joc.2026.01.09

multidimensional KNN time series prediction with invariance (EEMD-MKNN-TSPI) and identified that the EEMD-MKNN-TSPI model performs well in forecasting the stock price.

Generally, most existing studies focus on financial time series prediction in traditional markets, such as stock market, bond market, futures market, and commodity market (Fang & Li, 2015; Jang & Lee, 2019; Yang & Wang, 2019); research on digital cryptocurrency price prediction remains scarce. By adopting the decomposition-reconstruction-integration framework, we combine EMD with the LSTM model and propose a combined CEEMD-LSTM model. This method first decomposes digital cryptocurrencies price using CEEMD, recombines IMFs by reconstruction, then builds LSTM models to predict each subsequence in turn, and integrates all predictions to obtain the ultimate prediction results. In this study, the high-frequency trading prices of Bitcoin and Ethereum—the most popular digital cryptocurrencies—are selected as test objectives. The high-frequency data with four-time intervals are selected, and the existing mainstream machine learning prediction models and combined models based on different EMD methods are compared to test the validity of the CEEMD-LSTM model.

3 RESEARCH OBJECTIVE, METHODOLOGY, AND DATA

3.1. CEEMD model

CEEMD is a decomposition method optimized by Yang and Wang (2021) based on EMD and EEMD. EMD may experience modal aliasing. Even after adding self-noise to EEMD and decomposing it, residual noise will appear. CEEMD combines the advantages of EMD and EEMD, while addresses their shortcomings. CEEMD first adds two types of self-generated noise (positive and negative) to the original sequence to form two new sequences. Thereafter, the new sequence is decomposed into different IMFs and residual components using EMD, and averaged to cancel out the added self-generated noise. The steps for CEEMD decomposition are as follows:

Step 1: Add two sets of self-noise sequences ($w_i^+(t)$ and $w_i^-(t)$) to the original signal $x(t)$ to form new sequences ($h_i^+(t)$ and $h_i^-(t)$):

$$\begin{cases} h_i^+(t) = x(t) + w_i^+(t) \\ h_i^-(t) = x(t) + w_i^-(t) \end{cases} \tag{1}$$

Step 2: Use EMD method to decompose the new sequences ($h_i^+(t)$ and $h_i^-(t)$) and obtain two sets of IMF components ($\widetilde{imf_j}$) and residual component (res).

Step 3: Average the two sets of IMF components after EMD decomposition, eliminate the added self-noise, and obtain the final IMF component and residual component after CEEMD decomposition:

$$\widetilde{imf_j} = \frac{1}{2T}\sum_{i=1}^{2T} imf_j^i \tag{2}$$

https://doi.org/10.7441/joc.2026.01.09

$$X = \widetilde{imf_1} + \widetilde{imf_2} + \cdots + \widetilde{imf_j} + res$$
$$= \sum_{k=1}^{j} imf_k + res \tag{3}$$

Where, X is the original time series.

3.2. LSTM model

The LSTM model introduces gated self-circulation to ensure that the gradients can be transmitted for a long time, thereby solving the issue of gradient disappearance. It is suitable for continuous prediction. As Figure 1 depicts, one neuron includes one cell state and three gate mechanisms. The cell state changes over time, and the information in the cell state is either introduced or abandoned through gate mechanisms.

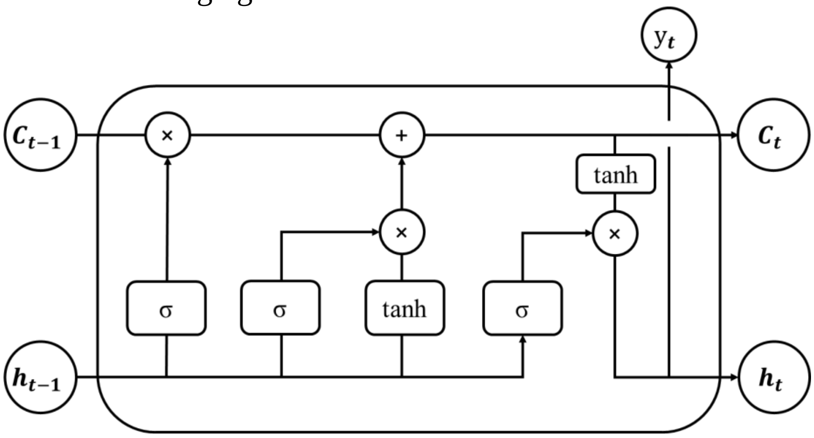

Fig.1 Internal structure of LSTM model neurons
Source: Compiled by the author

LSTM mainly comprises three gate structures: forgetting gate, input gate, and output gate, to protect and control cell states. All gate control units have sigmoid non-linear characteristics. The forgetting gate determines which information in the cell state will be discarded. The input gate determines which new input information will be retained, comprising a sigmoid neural network layer and a tanh neural network layer. The output gate determines the output value based on the cell state and applies a sigmoid neural network layer to determine the portion of the cell state output.

The processing of a neuron is achieved by the above gating mechanisms. Multiple neurons are connected in series to form an LSTM model, which has selective memory ability to extract long-term dependency features of learning sequences. It is widely applied in time series prediction.

3.3. Research hypotheses and model construction of CEEMD-LSTM for digital cryptocurrencies

In the context of integrating CEEMD (complementary ensemble empirical mode decomposition) and LSTM (long short-term memory networks), the research hypotheses for the model construction of CEEMD-LSTM include:

(1) Improvement in prediction accuracy: By decomposing the original time series into multiple scale components using CEEMD and modeling each intrinsic mode function (IMF) and residual component separately with LSTM, it is expected that this approach can achieve higher prediction accuracy compared to using LSTM alone or other traditional methods.

(2) Enhancement of feature extraction capability: Compared to directly feeding raw data into an LSTM, decomposing the data through CEEMD first allows for better capture of information at different frequency components. This enables the LSTM to learn data characteristics at a finer level, thus improving the overall performance of the model.

(3) Reduction of overfitting risk: As CEEMD decomposes complex original signals into several relatively simpler sub-series, this simplifies the learning process for the subsequent LSTM network, thereby reducing the risk of overfitting.

(4) Adaptability to various application scenarios: The proposed CEEMD-LSTM model is not only suitable for specific fields such as weather forecasting and financial market analysis but also demonstrates good versatility and adaptability, enabling its wide application in other scenarios requiring time series prediction.

These hypotheses outline the anticipated benefits and improvements of integrating CEEMD with LSTM for time series analysis and prediction tasks across various domains.

According to the hypotheses and the previous analysis, the time series data of digital cryptocurrency price are non-linear, highly complex, multi-scale, highly noisy, and with high randomness. Therefore, we apply the CEEMD-LSTM model to forecast the digital cryptocurrencies price. Based on the decomposition-reconstruction-integration framework, the CEEMD method decomposes the digital cryptocurrency price series into many characteristic modal components, which reflect the characteristics of digital cryptocurrency on different time scales. We conduct the single-sample T-test to differentiate the high- and low-frequency components and reconstruct the low-frequency components. Thereafter, we apply the LSTM model to identify and learn the characteristics of each component and fit its law to achieve accurate forecast of each part. Finally, the predicted results of all components are integrated to obtain the overall price forecast. Figure 2 illustrates the forecasting procedure of the CEEMD-LSTM model.

https://doi.org/10.7441/joc.2026.01.09

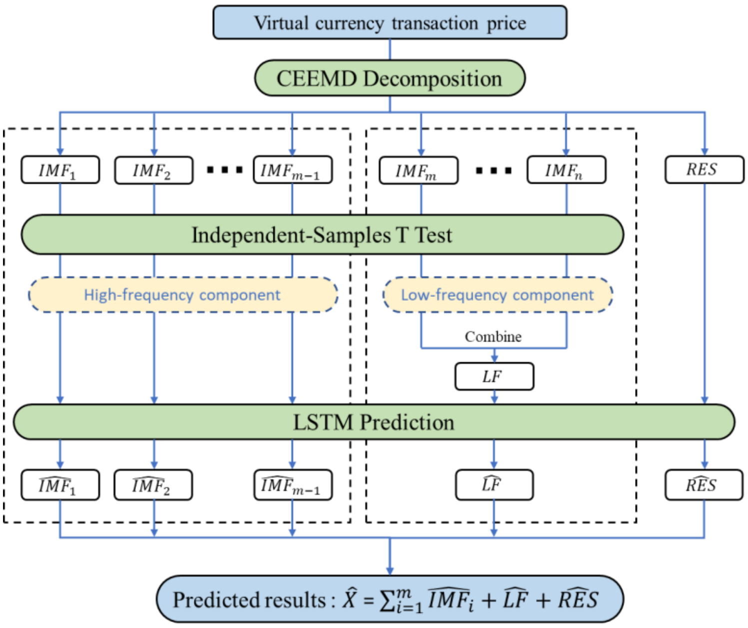

Fig.2 CEEMD-LSTM network structure
Source: Compiled by the author

The forecasting steps are as follows:

Step 1: The CEEMD method is utilized to decompose the digital cryptocurrencies price series to n intrinsic mode function components (IMF1, IMF2, …, IMFn) and the trend term (RES).

Step 2: The single sample T-test is performed for all IMFs. The first component of $\alpha > 0.05$ (set as IMFm) and its subsequent components are determined as the low-frequency components (IMFm, IMFm+1, ..., IMFn), representing the periodic trends of digital cryptocurrencies prices. We restructure the low-frequency components into a new component (LF). The components before IMFm are high-frequency components (IMF1, IMF2, ..., IMFm-1), which represent the short-term random volatility of cryptocurrency prices. RES is a trend term that represents the long-term fluctuations in the original sequence.

Step 3: Model predictions for IMF1, IMF2, …, IMFm-1, LF, and RES subsequences using LSTM neural networks, respectively.

Step 4: Combine prediction results $\widehat{\mathrm{IMF}}_1$, $\widehat{\mathrm{IMF}}_2$, ……, $\widehat{\mathrm{IMF}}_{m-1}$, $\widehat{\mathrm{LF}}$, $\widehat{\mathrm{RES}}$ in Step 3, and obtain the prediction results of the original sequence.

3.4. Data sources

We collect high-frequency trading data for Bitcoin (BTC) and Ethereum (ETF), which represent the two cryptocurrencies with highest market capitalization and are denominated in USDT, to elaborate the predictive modeling process for cryptocurrencies. Those high-frequency trading data are from Binance Cryptocurrency Exchange, which can be downloaded by calling the API provided by the exchange using the Python 3.x language. The selected data covers the period over January 1, 2018, to December 31, 2020, with prices recorded at one-minute intervals. The downloaded minute-level data are manually sorted to extract 15-minute, 30-minute, 1-hour, and

https://doi.org/10.7441/joc.2026.01.09

2-hour prices (denoted by BTC15min, BTC30min, BTC1h, BTC2h, ETH15min, ETH30min, ETH1h, and ETH2h, respectively).

4 RESULTS

4.1. CEEMD-based cryptocurrency price decomposition

As the model construction methods of the two cryptocurrencies are similar, the 2-hour high-frequency data of Bitcoin are selected as the representative research objective. In this experiment, the EMD, EEMD, CEEMD, and CEEMDAN methods are operated in MATLAB 2020a. Referring to a series of experiments conducted by Gao et al. (2024) to adjust the key parameters of the model, the optimal model parameters obtained are as follows: the number of CEEMD ensemble members (NE) was set to 100, the standard deviation of the added noise relative to the standard deviation of the original sequence was set to 0.2, and the maximum number of iterations for each component was set to 1000. Figure 3 presents the decomposition results of CEEMD.

As illustrated in Figure 3, the first sequence is the raw data of Bitcoin transaction prices, which is decomposed into 13 sequences through the CEEMD decomposition. The first 12 sequences are the IMF components. We identify that the fluctuation frequencies of IMF1 to IMF12 decrease successively, and the last sequence is the trend item. The first part of the IMF component sequence is the high-frequency component, which has a complex image of irregular high-frequency fluctuations, and the means fluctuate around 0. The high-frequency components reflect the impact of the drivers such as investors' behavior and investment, policy intervention, and other factors on the short-term fluctuation of Bitcoin prices. When the IMF series fluctuation is at the back, it is called the low-frequency component, indicating that its frequency is low. It shows that factors such as macroeconomic policies and economic cycles, represented by those components, have substantial impacts on the Bitcoin market. It also reflects the patterns of change in long-term time series. Regarding the trend item series, Figure 4 depicts that the RES component first declines and then gradually rises, capturing the long-term volatility trend of Bitcoin prices over the three-year sample period.

https://doi.org/10.7441/joc.2026.01.09

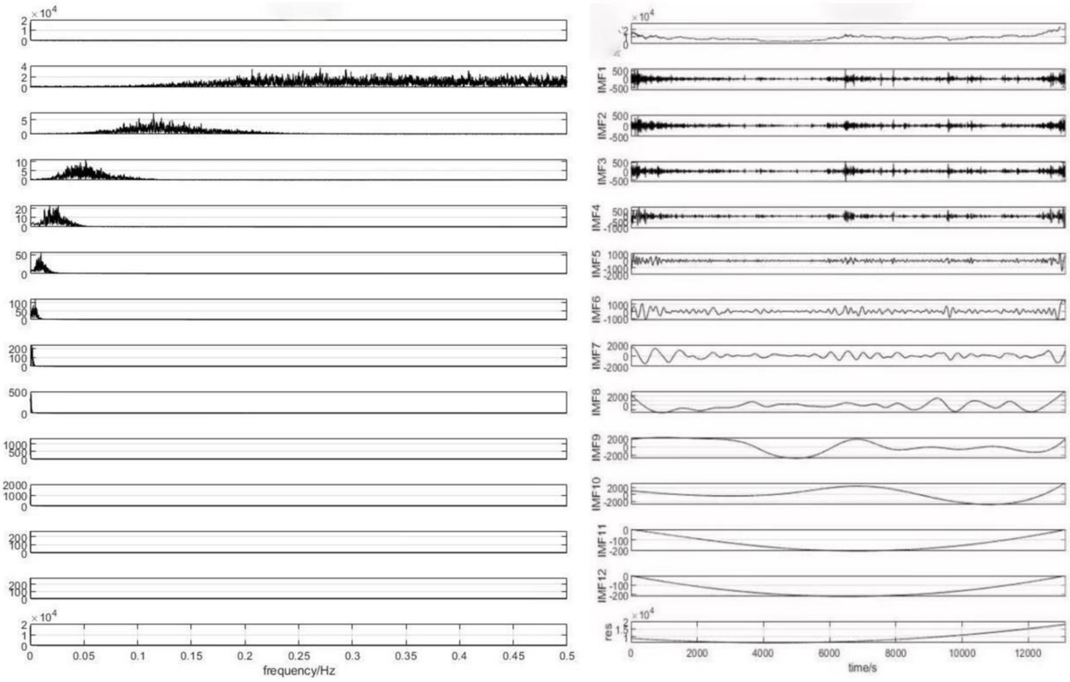

Fig.3 BTC2h decomposed by CEEMD decomposition of each sequence after decomposition of each component sequence and spectrogram[2]
Source: Compiled by the author

To intuitively analyze the effect of each IMF on the original Bitcoin price, we apply the Fourier transform method to draw the spectrum of each component as illustrated in Figure 3. We identify that the frequency of IMF1 to IMF12 is decreasing, while the amplitude value is increasing, indicating that the influence of the low-frequency component is more prominent than that of the high-frequency component, and it has the characteristics of periodic fluctuations. Following the research of Li and Feng (2015), we assume that the high-frequency components fluctuate around the mean of 0. The one-sample T test with mean 0 for each IMF is conducted, and the first IMF sequence with a deviation from the mean 0 is selected to identify high- and low-frequency components.

According to Table 1, the t-value of T test of the IMF1 sequence is -0.167, the mean difference is -0.0789, and the p-value is 0.867, greater than 0.05 significance level. The results show that the IMF1 sequence can be classified as high-frequency component. Similarly, we identify the IMF2 and IMF3 as high-frequency components. As the p-value of IMF4 sequence is less than 0.05, indicating that the mean value of IMF4 is significantly different from 0, we consider IMF4 as a low-frequency component. The IMF5 and IMF7 sequences are also identified as the low-frequency components.

Table 1. BTC2h One-sample t-test for each component after CEEMD-based decomposition
Source: Compiled by the author

Component	**Z Z = 0**				
	T-value	**df**	**Sig.**	**Mean**	**Difference 95%**

[2] For space consideration, we take BTC2h as an example.

https://doi.org/10.7441/joc.2026.01.09

			(2-tailed)	difference	confidence interval	
					Lower Bound	Upper Bound
IMF1	-0.167	13104	0.867	-0.0789	-1.0052	0.8473
IMF2	-0.076	13104	0.939	-0.0329	-0.8805	0.8147
IMF3	0.103	13104	0.918	0.0632	-1.1401	1.2665
IMF4	-2.031	13104	0.042	-4.2567	-4.4346	-0.0787
IMF5	-1.646	13104	0.1	-2.7789	-6.0884	0.5306
IMF6	5.079	13104	0	13.3115	8.174	18.4491
IMF7	-1.588	13104	0.112	-6.544	-14.6196	1.5316
IMF8	-23.124	13104	0	-161.0498	-174.7014	-147.3981
IMF9	10.73	13104	0	133.3802	109.0149	157.7454
IMF10	-4.989	13104	0	-65.982	-91.9047	-40.0593
IMF11	-245.558	13104	0	-134.1505	-135.2213	-133.0796
IMF12	-255.959	13103	0	-145.0906	-146.2017	-143.9794
RES	315.426	13104	0	9025.0559	8968.9717	9081.1401

4.2. Modeling of LSTM network models

The evidence in this part of the study is provided by Keras + TensorFlow 2.5 and implemented on Python 3.7. BTC2h is employed to decompose and reconstruct, and the LSTM model is developed for predicting each sub-sequence respectively. Rolling forecasts are performed for each sub-sequence including RMSE, MAPE, MAE, and R2, shown as follows.

RMSE:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \tag{4}$$

MAPE:

$$\mathrm{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{5}$$

MAE:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{6}$$

R2:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\bar{y}_i - y_i)^2} \tag{7}$$

Where, y_i is the original value of the sequences. $\hat{Y}_i$ is the forecast value of the sequences. $\overline{Y}_i$ is the mean value of the sequences. N is the total number of predicted data.

https://doi.org/10.7441/joc.2026.01.09

Before the prediction, to ensure the removal of the impact induced by different dimensions of different sequences, the data of each series are Z-score normalized. The training and test sets includes the first 90% and last 10% of the data, respectively. The rolling prediction modeling is applied to predict the bitcoin trading price in the next period by the former 24h trading data, and the lookback setting is the amount of data within 24h (e.g., the lookback is set to 12 in BTC2h), and the LSTM model is predicted for the normalized sequences respectively. Taking BTC2h as an example, after repeatedly adjusting the parameters for comparison, we set the number of hidden layers as 1, set the number of nodes in the input layer as 29, set the number of samples per training of (batch size) as 150 (the specific setting logic is shown in the following formula), and set the number of iterations epoch as 300. The optimizer is selected as Adam.

$$\begin{gathered} \mathrm{Batch_{size}} = 150 * \frac{x}{15\mathrm{min}} \\ x = 2\mathrm{h}, 1\mathrm{h}, 30\mathrm{min}, 15\mathrm{min} \end{gathered} \tag{8}$$

4.3. IMF sequence restructuring and optimization

Previous research has shown that high-frequency components contain substantial noise and short-term stochastic effects, while the low-frequency components are relatively smoother. This suggests that if the high-frequency components are applied to forecast, substantial noise would not be cancelled out, but would be utilized to expand the impacts of the stochastic factors.

Therefore, we forecast the high-frequency components line-by-line, while the low-frequency components are aggregated into a new sequence (denoted as LF). The reorganized components are illustrated in Figure 4. The processing of the remaining three different time span data is similar to BTC2h. Owing to space limitation, it is not explained in detail. The specific equations are as follows:

$$\text{CEEMD-LSTM} = \sum_{i=1}^{m} \widehat{\mathrm{IMF}}_i + \widehat{\mathrm{LF}} + \widehat{\mathrm{res}} \tag{9}$$

Where, $\widehat{\mathrm{IMF}}_i$ represents the forecast value of each HF, $\widehat{\mathrm{LF}}$ represents the forecast value of the LF after restructuring, and $\widehat{\mathrm{res}}$ represents the forecast RES value.

We adopt the LSTM model to predict the three high-frequency components of BTC2h decomposed by CEEMD, the new sequence LF after recombination, and trend term (res), and sum the predictions of each sequence to obtain the final prediction result. We name the prediction result as CEEMD-LSTM1, that is:

$$\text{CEEMD-LSTM1} = \widehat{\mathrm{IMF}}_1 + \widehat{\mathrm{IMF}}_2 + \widehat{\mathrm{IMF}}_3 + \widehat{\mathrm{LF}} + \widehat{\mathrm{res}} \tag{10}$$

To evaluate the effectiveness of the IMF restructuring approach in this study, we compare it with two previous approaches:

(1) The high-frequency component is recombined, and the new recombined sequences are named HF. The LSTM model is applied to forecast HF, LF, and the trend term (res) respectively. We name this approach as CEEMD-LSTM2.

https://doi.org/10.7441/joc.2026.01.09

(2) The LSTM model is applied to forecast HF, the nine low-frequency components (IMF4 ~ IMF12), and the trend term (res). We name this combined forecasting approach as CEEMD-LSTM3.

The two approaches are shown as follows:

$$\text{CEEMD-LSTM2} = \widehat{\text{HF}} + \widehat{\text{LF}} + \widehat{\text{res}} \tag{11}$$

$$\text{CEEMD-LSTM3} = \widehat{\text{HF}} + \sum_{i=4}^{12} \widehat{\text{IMF}}_i + \widehat{\text{res}} \tag{12}$$

Fig. 4 BTC2h decomposition recombination diagram
Source: Compiled by the author

With the model parameters assigned earlier and the high-frequency IMF identified based on the T test, the prediction results of the three approaches above for predicting BTC2h are shown in Table 2. The CEEMD-LSTM1 is the optimal combination, and the LSTM prediction effect is the best. The RMSE, MAPE, MAE, and R^2 of the CEEMD-LSTM1 increase by 56.93%, 53.59%, 54.3%, and 0.068% respectively, compared with the LTSM2, and increase by 60.76%, 59.81%, 60.24%, and 0.085% respectively, compared with LTSM3.

Tab.2 Evaluation results of BTC2h under different combinations
Source: Compiled by the author

Different combination models	**RMSE**	**MAPE**	**MAE**	**R^2**
CEEMD-LSTM1	59.0546	0.224917	38.6078	0.999845
CEEMD-LSTM2	137.114	0.484589	84.48618	0.999164
CEEMD-LSTM3	150.505	0.559636	97.09343	0.998993

4.4 Prediction results of CEEMD-LSTM model

https://doi.org/10.7441/joc.2026.01.09

The reconstruction process and forecasting models of BTC15min, BTC30min, BTC1h, ETH15min, ETH30min, ETH1h, and ETH2h cryptocurrency prices are identical to those applied for BTC2h. In this study, we construct two single benchmark models: SVR and RNN, to test the advantages of the LSTM models in predicting the digital cryptocurrency prices based on the high-frequency data.

To test the effectiveness of the CEEMD decomposition methods in predicting the digital cryptocurrency high-frequency data, we apply the EMD, EEMD, and CEEMDAN methods to decompose the original series of the digital cryptocurrency price. We uniformly set the decomposition integration number (NE) as 100 and set the ratio of additional noise standard deviation to the original sequence standard deviation as 0.2. The prediction effect comparison results of six combined models (EMD-LSTM, EEMD-LSTM, CEEMD-LSTM, CEEMDAN-LSTM, CEEMD-SVR, CEEMDAN-SVR) as well as three single models are presented in Table 3.

Tab.3 Comparison of the prediction effect of each model for two cryptocurrencies
Source: Compiled by the author

Cryptocurrency Types / Modeling Method		**BTC**				**ETH**			
2H	**Evaluation Indicators**	**RMSE**	**MAPE**	**MAE**	**R^2**	**RMSE**	**MAPE**	**MAE**	**R^2**
Single model	RNN	1031.8954	2.2534	493.1619	0.952881	5.8782	0.7963	3.9066	0.997273
	SVR	215.1512	1.1123	149.9307	0.997877	22.5746	3.8802	14.7054	0.952508
	LSTM	214.7447	0.6705	124.1319	0.997949	5.7398	0.7565	3.7245	0.997391
Combination Model	EMD-LSTM	95.2110	0.3754	64.1094	0.999597	3.2533	0.4844	2.3555	0.99916
	EEMD-LSTM	79.5071	0.3814	59.5661	0.999719	3.7813	0.6647	3.0250	0.998868
	CEEMD-LSTM	59.0546	0.2249	38.6078	0.999845	2.0676	0.3146	1.5061	0.999661
	CEEMDAN-LSTM	68.1847	0.2958	48.0665	0.999793	2.6572	2.6572	2.0802	0.999441
	CEEMD-SVR	448.0140	2.3515	292.2702	0.990154	14.8013	2.7558	11.3414	0.981582
	CEEMDAN-SVR	518.0155	3.2105	445.1707	0.985857	18.8322	3.3518	13.6968	0.966345
1H	**Evaluation Indicators**	**RMSE**	**MAPE**	**MAE**	**R^2**	**RMSE**	**MAPE**	**MAE**	**R^2**
Single model	RNN	640.8625	1.3971	301.0667	0.981809	5.9695	0.8622	4.3312	0.997189
	SVR	152.7230	0.7075	96.9472	0.998958	22.4604	3.7664	14.4313	0.954043
	LSTM	151.4348	0.5145	92.0974	0.998982	4.5167	0.6410	3.0658	0.998388
Combination Model	EMD-LSTM	81.5150	0.2790	49.0196	0.999705	2.5133	0.3453	1.7117	0.999501
	EEMD-LSTM	70.3977	0.3488	53.8063	0.99978	3.5447	0.6267	2.8168	0.999006
	CEEMD-LSTM	47.8958	0.1765	30.6502	0.999898	1.6368	0.2381	1.1500	0.999788
	CEEMDAN-LSTM	56.4226	0.2505	40.4279	0.999859	2.4167	0.4061	1.8770	0.999538
	CEEMD-SVR	421.1338	2.3118	315.6599	0.991295	14.6474	2.6671	11.1982	0.981977
	CEEMDAN-SVR	458.5310	2.8424	419.8839	0.989483	7.7185	1.3620	6.1874	0.995317
30min	**Evaluation Indicators**	**RMSE**	**MAPE**	**MAE**	**R^2**	**RMSE**	**MAPE**	**MAE**	**R^2**
Single model	RNN	815.6243	2.0763	442.5075	0.970564	10.9575	2.1506	9.8530	0.990534
	SVR	139.1725	0.7360	102.2101	0.999123	22.3402	3.6890	14.1987	0.954684
	LSTM	91.7324	0.3802	58.1297	0.999627	3.3416	0.4847	2.3216	0.999119
Combination Model	EMD-LSTM	50.9303	0.1904	32.5850	0.999885	1.9015	0.2749	1.3256	0.999715
	EEMD-LSTM	66.1308	0.3407	51.6727	0.999807	3.4514	0.6064	2.7247	0.999062
	CEEMD-LSTM	30.5753	0.1162	19.6364	0.999959	1.1266	0.1623	0.7776	0.9999
	CEEMDAN-LSTM	43.8454	0.2069	32.2471	0.999915	2.0730	0.3528	1.6165	0.999661
	CEEMD-SVR	425.4745	3.1053	283.8442	0.991096	17.1798	2.8615	11.9567	0.974836
	CEEMDAN-SVR	574.8010	3.7208	519.1311	0.983014	15.4443	2.6697	11.1182	0.979637
15min	**Evaluation Indicators**	**RMSE**	**MAPE**	**MAE**	**R^2**	**RMSE**	**MAPE**	**MAE**	**R^2**
Single model	RNN	3040.2608	7.7454	1647.1723	0.590931	12.0365	2.3499	11.1988	0.988579
	SVR	125.8602	0.6409	92.8621	0.999285	22.2149	3.6171	25.0730	0.955214
	LSTM	89.5605	0.4717	49.5635	0.999645	2.5112	0.4565	1.8033	0.999503
Combination Model	EMD-LSTM	40.9220	0.1448	24.9776	0.999926	1.3011	0.1795	0.8817	0.999866
	EEMD-LSTM	57.9615	0.3074	45.7293	0.999851	3.3319	0.5924	2.6557	0.999126
	CEEMD-LSTM	29.6382	0.1116	19.6234	0.999961	0.8831	0.1332	0.6326	0.999938
	CEEMDAN-LSTM	38.9289	0.1909	29.3415	0.999933	2.3110	0.3829	1.6795	0.999235
	CEEMD-SVR	398.5344	2.1416	291.0015	0.992563	17.8576	2.9217	12.2788	0.972854

https://doi.org/10.7441/joc.2026.01.09

	CEEMDAN-SVR	225.8824	1.0915	163.1229	0.997645	9.8839	1.8297	8.2434	0.99266

Table 3 shows that, according to the four evaluation indicators, for both cryptocurrencies, the CEEMD-LSTM model performs the best in forecasting. Taking BTC2H as an example, the forecast effect of LSTM is the best in a single model, but there is no difference between LSTM and SVR. Compared with the combined models, the CEEMD-LSTM model demonstrates better predictive performance, and its RMSE, MAPE, MAE, and R^2 is improved by 72.5%, 66.46%, 68.9%, and 0.19%, respectively, compared with the single LSTM model; by 37.98%, 40.1%, 39.78%, and 0.0248%, respectively, compared with the EMD-LSTM model; by 25.72%, 41.02%, 35.19%, and 0.0126%, respectively, compared with the EEMD-LSTM model; by 13.39%, 23.96%, 19.68%, and 0.005%, respectively, compared with the CEEMDAN-LSTM model.

For the CEEMD-LSTM model, the prediction effect of 15-minute high-frequency data is better than 30-minute, 1-hour, and 2-hour high-frequency data. The reason is that data of the next period are predicted based on data of the previous 24 hours, and reduction of period and increase of data volume improve the prediction effect of the model. So, the smaller the time interval period, the higher the prediction effect.

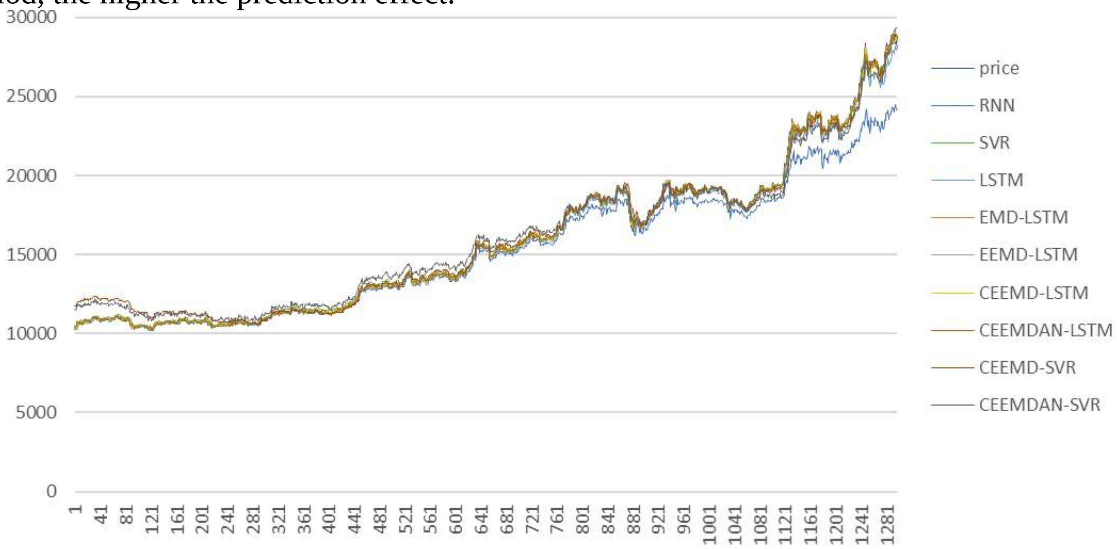

Fig.5 Comparison of prediction results of 9 models for BTC2h
Source: Compiled by the author

5 DISCUSSION

This study proposes the decomposition-reconstruction-integration framework to extract the intrinsic features and applies the CEEMD-LSTM model to forecast the digital cryptocurrencies price in a high-frequency world. This study enriches the literature on a deep learning framework for predicting financial time series, and proves that a deep learning framework has robust predictive performance in the high-frequency digital cryptocurrency market.

Although the prediction effect of the CEEMD-LSTM model in forecasting the digital cryptocurrencies price is best, the combined model of SVR performs worse than the corresponding single model. This may be due to the white noise added during the decomposition processes of the CEEMD and CEEMDAN methods, which induces the SVR

https://doi.org/10.7441/joc.2026.01.09

model to over-fit the white noise. However, increasing the penalty parameter may improve the forecasting performance of the SVR-based combined models. For the RNN-based prediction model, Figure 5 illustrates that the price of cryptocurrencies significantly fluctuates in the latter-half of the period, leading to a decrease in the predictive performance of RNN. This is primarily because RNN has limited ability to solve the long-time prediction problem. For the combined model of EEMD, CEEMDAN, the forecast data of Ethernet shows that the prediction effect of EEMD and CEEMDAN is far inferior to that of CEEMD. The reason is that this research adopts a uniform number of decomposition and integration (NE=100). The increase in data volume implies an increase in decomposition difficulty. For the EEMD and CEEMDAN, insufficient processing iterations often leave white noise at levels that are not negligible. Therefore, under the same number of decomposition and integration, CEEMD exhibits substantially higher decomposition efficiency than the other methods.

Notably, Xie and Cao (2022) identify that among four major cryptocurrencies in the Chinese market: Bitcoin, Ethereum, Ripple, and Litecoin. Ethereum exhibits the highest market efficiency, whereas Bitcoin shows the weakest. In contrast, countries such as Singapore and Japan adopt more open policies, actively formulating regulations conducive to the development of blockchain technology and cryptocurrencies. In the European and American markets, certain nations hold relatively positive attitudes toward cryptocurrencies, encouraging innovation while exploring effective regulatory frameworks. According to Baur and Hoang (2021), within the European market, cryptocurrencies can occasionally serve as weak safe-haven assets. Mnif et al. (2023) indicate that the investigated energy-efficient cryptocurrencies reacted negatively to the pandemic but positively to the Russia–Ukraine conflict. Separate evaluations tailored to the unique contexts of the United States and China revealed distinct outcomes. For China, the impact of economic policy uncertainty (EPU) was found to be negative and statistically insignificant. Conversely, the situation in the United States differs, given its position as the world's most developed economy. American policies wield significant global influence, positioning cryptocurrencies as potentially effective hedging instruments according (He et al., 2024). This underscores the importance of considering regional nuances when assessing the role and potential of cryptocurrencies in different parts of the world. Research on digital cryptocurrency price prediction remains scarce, and existing studies often fail to fully capture the distinctive characteristics of digital cryptocurrency prices.

To address this gap, we develop a combined prediction model-referred to as the CEEMD-LSTM model-based on the decomposition-reconstruction-integration framework, to predict digital cryptocurrency price movements. We apply the CEEMD method to decompose prices into high-frequency components, low-frequency components, and trend sequences, and reconstruct the combined IMF components. Thereafter, we utilize the LSTM model to predict each component separately and obtain sorted predicted values of prices by adding and integrating the predicted values of each component. Using high-frequency Bitcoin and Ethereum data across four time-intervals, we conduct a systematic comparison of the prediction accuracy between the CEEMD-LSTM model and other models, whether single or combined. We identify that the CEEMD-LSTM model consistently outperforms competing approaches in high-frequency forecasting. This refined analysis highlights the complex interplay between national policies, market dynamics, and the evolving landscape of cryptocurrencies, offering valuable insights for investors, policymakers, and stakeholders navigating this dynamic domain.

https://doi.org/10.7441/joc.2026.01.09

As blockchain-based cryptocurrencies, Bitcoin and Ethereum demonstrate robust liquidity and value storage functions on a global scale. Their price dynamics are primarily driven by factors such as the supply–demand relationship in the global market, investor sentiment, technological advancements, and macroeconomic environments. These factors transcend geographical boundaries, leading to a high degree of similarity in price trends worldwide, with both prices mutually influencing (Yermack, 2013; Huang & Huang, 2023). Specifically, owing to the acceleration of information dissemination, especially in the financial sector, any significant news or technological innovations regarding these two cryptocurrencies rapidly spread worldwide, inducing almost simultaneous reactions from market participants almost everywhere. Moreover, the facilitation of international investment channels has further narrowed the price differences among different regions.

However, despite this universality, local laws and regulations, tax policies, and varying levels of acceptance toward cryptocurrencies can, nonetheless, exert unique short-term impacts on local markets. Therefore, while predicting the prices of Bitcoin and Ethereum, although reliance on global indicators is possible, it is also necessary to consider specific regional variables to achieve more accurate analysis. This not only aids in enhancing the accuracy of investment decisions but also provides new perspectives for understanding the operational mechanisms of the cryptocurrency market. Against the backdrop of increasing global economic integration, studying the market price dynamics of cryptocurrencies such as Bitcoin and Ethereum holds immense significance for exploring the development of future forms of currency. Regulatory stances toward Bitcoin and Ethereum vary widely across countries. Some adopt open and supportive approaches, while others enforce stringent regulatory measures or outright bans. In Asian markets, China maintains a cautious approach toward cryptocurrencies, having already restricted related activities. Neither Bitcoin nor Ethereum is considered legal tender in China and, thus, not protected by law. In contrast, cryptocurrencies such as Bitcoin are rapidly developing in the United States because the country supports cryptocurrencies trading.

However, this study has some limitations. First, there are numerous digital cryptocurrencies in the market, while this study only took two digital cryptocurrencies as examples. Second, external factors such as macroeconomic policies and investor sentiment are not explicitly incorporated. These limitations must be addressed in future studies to further elucidate the cryptocurrency market dynamics and to better inform the exploration of future currency forms. Overcoming these constraints may lead to deeper insights into the trends and behaviors of the cryptocurrency market, thereby contributing to the broader discourse on the evolution of money.

6 CONCLUSION

In response to the non-linear, high-noise, multi-scale, and unstructured characteristics of digital cryptocurrencies price series, we adopt the framework of decomposition-reconstruction-integration, apply a unique IMF reconstruction method, and construct the CEEMD-LSTM model to forecast the digital cryptocurrencies price. In this framework, the CEEMD was first applied to decompose the high-frequency digital cryptocurrencies trading price, and a new IMF restructuring approach was proposed for the CEEMD decomposed IMF series. The LSTM model was then implemented to simultaneously forecast the new sequence LF, high-frequency

https://doi.org/10.7441/joc.2026.01.09

components, and trend terms, and the forecast data were summed to a final forecast set. Finally, we combined the forecast results of each sequence to obtain the final results.

By using high-frequency trading price data for Bitcoin and Ethereum across four different time intervals, we evaluate the prediction effectiveness of the CEEMD-LSTM model against various indicators and compare it with other single and hybrid models. We conclude as follows: (1) According to the results of the four evaluation indicators, the IMF restructuring method proposed in this study significantly outperforms existing IMF reconstruction approaches. (2) In the LSTM model framework, the prediction error declines as the time interval decreases, and, therefore, as the computing power allows, we increase the prediction effectiveness of the model by decreasing the time interval of the data. (3) For both digital cryptocurrencies, the CEEMD-LSTM model constructed in this study consistently outperforms the other hybrid models across all four indicators.

Funding
This study is supported by Guizhou Provincial Program on Commercialization of Scientific and Technological Achievements (QKHCG[2025]ZD010), Guizhou Provincial Education Department Natural Science Research Project (Qian Jiao Ji [2023] No.033), and Guizhou Province Key Laboratory of Sovereign Blockchain (No.: ZSYS[2024]003).

References

1. Arel, I., Rose, D. C., & Karnowski, T. P. (2010). Deep machine learning – A new frontier in artificial intelligence research. *IEEE Computational Intelligence Magazine*, *5*(4), 13-18. https://doi.org/10.1109/MCI.2010.938364
2. Baur, D. G., & Hoang, L. T. (2021). A crypto safe haven against Bitcoin. *Finance Research Letters*, *38*, 101431. https://doi.org/10.1016/j.frl.2020.101431
3. Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and Trends in Machine Learning*, *2*(1), 1-127. https://doi.org/10.1561/2200000006
4. Bengio, Y. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *35*(8), 1798-1828. https://doi.org/10.1109/TPAMI.2013.50
5. Chen, C. P., et al. (2022). Short-term wind speed prediction model based on multi-mode decomposition and sparrow search optimized residual networks. *Power System Technology*, *46*(8), 2975-2985. https://doi.org/10.13335/j.1000-3673.pst.2021.2510
6. Cui, H. Y., & Dou, X. S. (2018). Carbon price forecasts in Chinese carbon trading market based on EMD-GA-BP and EMD-PSO-LSSVM. *Operations Research and Management Science*, *27*(7), 133-143. https://doi.org/10.12005/orms.2018.0166
7. Ding, H., et al. (2024). Forecasting river ice breakup dates in northern cold regions using deep learning and signal decomposition. *Journal of Hydraulic Engineering*, *55*(5), 577-585. https://doi.org/10.13243/j.cnki.slxb.20230549

8. Fang, Z. B., & Li, Y. (2015). Denoising and predicting of exchange rate data based on HHT-SVR model. *Journal of Applied Statistics and Management*, *34*(5), 900-909. https://doi.org/10.13860/j.cnki.sltj.20150922-061
9. Gao, Y., Zong, W., Hu, K., et al. (2024). Short-term ride-hailing demand prediction based on CNN-ATTBiLSTM. *Operations Research and Management Science*, *33*(11), 211-217. https://doi.org/10.12005/orms.2024.0375
10. Goczek, U., & Skliarov, I. (2019). What drives the Bitcoin price? A factor augmented error correction mechanism investigation. *Applied Economics*, *51*(59), 1-18. http://dx.doi.org/10.1080/00036846.2019.1619021
11. Graves, A., Mohamed, A. R., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. *IEEE International Conference on Acoustics*, 6645-6649. https://doi.org/10.1109/ICASSP.2013.6638947
12. He, C., Li, Y., Wang, T., & Shah, S. A. (2024). Is cryptocurrency a hedging tool during economic policy uncertainty? An empirical investigation. *Humanities & Social Sciences Communications*, *11*(1), 1-12.
13. He, J., & Chang, S. (2001). Chinese stock market's multifractal walk and forecasting. *Chinese Journal of Management Science*, *10*(3), 12-18. https://doi.org/10.3321/j.issn:1003-207X.2002.03.003
14. Herrera, G. P., et al. (2019). Long-term forecast of energy commodities price using machine learning. *Energy*, *179*, 214-221. https://doi.org/10.1016/j.energy.2019.04.077
15. Herwartz, H. (2017). Stock return prediction under GARCH – An empirical assessment. *International Journal of Forecasting*, *33*(3), 569-580. https://doi.org/10.1016/j.ijforecast.2017.01.002
16. Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786), 504-507. https://doi.org/10.1126/science.1127647
17. 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
18. Hsu, M. W., et al. (2016). Bridging the divide in financial market forecasting: Machine learners vs. financial economists. *Expert Systems with Applications*, *61*, 215-234. http://dx.doi.org/10.1016/j.eswa.2016.05.033
19. Huang, N. E., Shen, Z., & Long, S. R. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London Series A*, *454*(1971), 903-995. https://doi.org/10.1098/rspa.1998.0193
20. Huang, Q., & Huang, W. (2023). Influencing factors and interrelationship of cryptocurrency price volatility – Evidence from Ethereum and Bitcoin. *Fujian Forum*, *3*, 187-200.
21. Huang, R. P., Zuo, W. M., & Bi, L. Y. (2015). Predicting the stock market based on microblog mood. *Journal of Industrial Engineering and Engineering Management*, *29*(1), 47-52. https://doi.org/10.3969/j.issn.1004-6062.2015.01.006

22. Jang, H., & Lee, J. (2019). Machine learning versus econometric jump models in predictability and domain adaptability of index options. *Physica A: Statistical Mechanics and Its Applications*, *513*, 74-86. https://doi.org/10.1016/j.physa.2018.08.091
23. Ji, X. F., et al. (2025). Prediction of tourism passenger flows during golden weeks using CNN-LSTM-Attention combined models: A case study of Dali Prefecture. *Journal of Arid Land Resources and Environment*, *39*(3), 200-208. https://doi.org/10.13448/j.cnki.jalre.2025.056
24. Jiang, Y. Y., Shao, J., Chen, M. K. & Wang F. Y. (2023). Financial time series forecasting model construction using integrated decomposition and deep learning techniques. *Statistics and Decision*, *39*(24), 152-156. https://doi.org/10.13546/j.cnki.tjyjc.2023.24.027
25. Jiang, Z. R., & Wang, R. R. (2024). Research on cash flow prediction of listed companies based on MF-LSTM. *Journal of Applied Statistics and Management*, *43*(3), 512-526. https://doi.org/10.13860/j.cnki.sltj.20240509-008
26. Jing, N., Wu, Z., & Wang, H. (2021). A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction. *Expert Systems with Applications*, *178*, 115019. https://doi.org/10.1016/j.eswa.2021.115019
27. Kim, D., & Fan, J. Q. (2019). Factor GARCH-It models for high-frequency data with application to large volatility matrix prediction. *Journal of Econometrics*, *208*(2), 395-417. https://doi.org/10.1016/j.jeconom.2018.10.003
28. Lei, B., Zhang, B., & Song, Y. (2021). Volatility forecasting for high-frequency financial data based on web search index and deep learning model. *Mathematics*, *9*(4), 320. https://doi.org/10.3390/MATH9040320
29. Li, C., et al. (2022). Short-term power forecasting for wind farm clusters based on SDAE deep learning and multiple integration. *High Voltage Engineering*, *48*(2), 504-512. https://doi.org/10.13336/j.1003-6520.hve.20210130
30. Li, H. L., Feng, C. E. (2014). Relationship between investor sentiment and stock indices fluctuation based on EEMD. *Systems Engineering-Theory & Practice*, *34*(10), 2495-2503. hhttps://d.wanfangdata.com.cn/periodical/xtgcllysj201410003
31. Lin, G., Lin, A., & Cao, J. (2021). Multidimensional KNN algorithm based on EEMD and complexity measures in financial time series forecasting. *Expert Systems with Applications*, *168*, 114443. https://doi.org/DOI: 10.1016/j.eswa.2020.114443
32. Liu, W. L., Zhang, Y., & Yang, X. Y. (2024). Online margin trading combination strategies based on LSTM predictive information. *Systems Engineering - Theory & Practice*, *44*(8), 2493-2511. https://doi.org/10.12011/SETP2022-0819
33. Long, Y., et al. (2024). Research on Dongting Lake water level prediction scheme based on IPSO-EGA-LSTM model. *Resources and Environment in the Yangtze Basin*, *33*(6), 1262-1272. https://d.wanfangdata.com.cn/periodical/cjlyzyyhj202406011
34. Mnif, E., Mouakhar, K., & Jarboui, A. (2023). Energy-conserving cryptocurrency response during the COVID-19 pandemic and amid the Russia-Ukraine conflict. *Journal of Risk Finance*, *24*(2), 169-185. https://doi.org/10.1108/JRF-06-2022-0161

https://doi.org/10.7441/joc.2026.01.09

35. Mohammadi, M. (2017). Prediction of α-stable GARCH and ARMA-GARCH-M models. *Journal of Forecasting*, *36*(7), 859-866. https://doi.org/10.1002/for.2477

36. Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3440802

37. Nelson, D., Pereira, A., & Oliveira, R. (2017). Stock market's price movement prediction with LSTM neural networks. *International Joint Conference on Neural Networks*. https://ieeexplore.ieee.org/document/7966019

38. Pan, Z., et al. (2019). Improving volatility prediction and option valuation using VIX information: A volatility spillover GARCH model. *Journal of Futures Markets*, *39*(6), 744-776. https://doi.org/10.1002/fut.22003

39. Qian, M. H., Fan, A. Y., & Tian, T. (2024). Research on multidimensional satisfaction quantification methods for public library users based on Word2Vec and LSTM hybrid models. *Library Construction*, *6*, 89-100. https://doi.org/10.19764/j.cnki.tsgjs.20240377

40. Salisu, A. A., Gupta, R., & Demirer, R. (2022). Global financial cycle and the predictability of oil market volatility: Evidence from a GARCH-MIDAS model. *Energy Economics*, *108*, 105934. https://doi.org/10.1016/j.eneco.2022.105934

41. Salles, R., et al. (2022). TSPred: A framework for nonstationary time series prediction. *Neurocomputing*, *467*, 197-202. https://doi.org/10.1016/j.neucom.2021.09.067

42. Sarkodie, S. A., Ahmed, M. Y., & Owusu, P. A. (2022). COVID-19 pandemic improves market signals of cryptocurrencies - Evidence from Bitcoin, Bitcoin Cash, Ethereum, and Litecoin. *Finance Research Letters*, *44*, 102049. https://doi.org/10.1016/j.frl.2021.102049

43. Shen, F., Chao, J., & Zhao, J. (2015). Forecasting exchange rate using deep belief networks and conjugate gradient method. *Neurocomputing*, *167*, 243-253. https://doi.org/10.1016/j.neucom.2015.04.071

44. Su, J., Li, G., & Zhang, X. (2024). Research on sandstorm weather prediction models in Northwest China using GCN-LSTM integrated with spatial-temporal features. *Journal of Arid Land Resources and Environment*, *38*(5), 111-120. https://doi.org/10.13448/j.cnki.jalre.2024.103

45. Su, Z., Lu, M., & Li, D. X. (2017). Deep learning in financial empirical applications: Dynamics, contributions and prospects. *Journal of Financial Research*, *5*, 111-126. https://d.wanfangdata.com.cn/periodical/jryj201705008

46. Tang, Y., et al. (2021). Building trend fuzzy granulation based LSTM recurrent neural network for long-term time series forecasting. *IEEE Transactions on Fuzzy Systems*, *30*(6), 1599-1613. https://doi.org/10.1109/TFUZZ.2021.3062723

47. Tian, L. X., & Qin, H. Q. (2024). Research on groundwater level prediction methods based on multivariate LSTM neural networks. *Journal of Arid Land Resources and Environment*, *38*(9), 138-146. https://doi.org/10.13448/j.cnki.jalre.2024.192

48. TokenInsight 2021 Crypto Trading Industry Annual Review. https://s2.tokeninsight.com/static/levelPdf/TokenInsight_2021_Yearly_Crypto_Trading_Industry_Report_compressed.pdf

https://doi.org/10.7441/joc.2026.01.09

49. Torres, M. E., Colominas, M. A., & Schlotthauer, G., & Flandrin, P. (2011). A complete ensemble empirical mode decomposition with adaptive noise. In *Proceedings of 2011 IEEE International Conference on Acoustics, Speech and Signal Processing*. https://doi.org/10.1109/ICASSP.2011.5947265

50. Werner, K. R., & Esteban, H. P. (2017). Volatility of main metals forecasted by a hybrid ANN-GARCH model with regressors. *Expert Systems with Applications*, *84*, 290–300. https://doi.org/10.1016/j.eswa.2017.05.024

51. Wu, L. L., Tai, Q. R., Bian, Y., & Li, Y. H. (2024). Research on the prediction of regional carbon price: A GA-VMD and CNN-BiLSTM-attention approach. *Operations Research and Management Science*, *33*(9), 134-139. http://www.jorms.net/CN/10.12005/orms.2024.0296

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

53. Xie, P., & Shi, W. G. (2015). A literature review of cryptocurrency. *Journal of Financial Research*, *1*, 1-15. http://dx.doi.org/10.15408/etk.v20i1.169111

54. Xie, W., & Cao, G. (2022). Research on multifractal cross-correlation between cryptocurrencies and the Chinese stock market based on the MFXDMA method. *China Management Science*, *30*(10), 72-84. https://doi.org/10.16381/j.cnki.issn1003-207x.2020.0402

55. Xu, H. Q. (2025). Analysis and trend prediction of AI research themes evolution in LIS discipline by integrating BERTopic and LSTM. *Information Science*, *43*(05), 105-116. http://doi.org/10.13833/j.issn.1007-7634.2025.05.011

56. Xue, Y. F., Chen, Z., Qiu, Y. S., Liu, J. H. (2024). Online learning eye movement cognitive hierarchy intelligent recognition model based on long short-term memory neural networks. *Modern Distance Education*, *5*, 70-78. http://doi.org/10.3969/j.issn.1001-8700.2024.05.008

57. Yang, H. L., & Lin, H. C. (2017). Applying the hybrid model of EMD, PSR, and ELM to exchange rates forecasting. *Computational Economics*, *49*(1), 99-116. https://doi.org/10.1007/s10614-015-9549-9

58. Yang, Q., & Wang, C. W. (2019). A study on forecast of global stock indices based on deep LSTM neural network. *Statistical Research*, *36*(3), 65-77. https://doi.org/10.19343/j.cnki.11-1302/c.2019.03.006

59. Yang, Y. W., & Yang, Z. J. (2005). Financial time series forecasting based on support vector machine. *Journal of Systems & Management*, *14*(2), 176-181. https://doi.org/10.3969/j.issn.1005-2542.2005.02.017

60. Yang, Y., & Wang, J. (2021). Forecasting wavelet neural hybrid network with financial ensemble empirical mode decomposition and MCID evaluation. *Expert Systems With Applications*, 166, 114097. https://doi.org/10.1016/S0925-2312(03)00372-2

61. Yao, Y. H., Wang, X. X., Chen, W., & Chen, Z. S. (2025). Research on global stock market extreme risk spillover using Transformer-LSTM quantile regression. *Chinese*

Journal of Management Science, *33*(8), 1-17. http://doi.org/10.16381/j.cnki.issn1003-207x.2024.1601

62. Yeh, J. R., Shieh, J. S., & Huang, N. E. (2010). Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method. *Advances in Adaptive Data Analysis*, *2*(2), 135–156. https://doi.org/10.1142/S1793536910000422
63. Yermack, D. (2013). Is Bitcoin a real currency? An economic appraisal. In Chuen, D. L. K. (ed.). *Handbook of digital currency*. Academic Press. https://doi.org/10.1016/B978-0-12-802117-0.00002-3
64. Yu, L. W., & Chen, T. (2017). Long-term memory in stock market and its dynamic structure. *Journal of Applied Statistics and Management*, *36*(6), 1106-1118. https://doi.org/10.13860/j.cnki.sltj.20170417-003
65. Zhai, Y. K., Jia, Q. S., Qiao, Y., & Zhao, J. (2025). Prediction of remote consultation service duration using ATT-FC-LSTM model. *Chinese Journal of Management*, *22*(3), 568-576. https://doi.org/10.3969/j.issn.1672-884x.2025.03.017
66. Zhang, J. H., & Ding, J. D. (2025). Research on emerging theme prediction based on BERTopic and LSTM models. *Information Studies: Theory & Application*, *43*(1), 1-18. https://doi.org/10.13833/j.issn.1007-7634.2025.01.011

Contact information

Prof. Chenggang Li
Guizhou University of Finance and Economics
School of Applied Economics
Guizhou Province Key Laboratory of Sovereign Blockchain
China
E-mail: lichenggang603@126.com

Xuan Wang
State Taxation Administration
Chongqing Beibei District Tax Bureau
China
E-mail: wx873868564@mail.gufe.edu.cn

Yan Gong
Guizhou University of Finance and Economics
School of Applied Economics
China
E-mail: sanshui33@126.com

Liang Wu (Corresponding Author)
Guizhou Normal University
School of Economics and Management
China

https://doi.org/10.7441/joc.2026.01.09

E-mail: wuliang@gznu.edu.cn

Huiyang Li
Christian Brothers University
School of Business
United States
E-mail: hli@cbu.edu

https://doi.org/10.7441/joc.2026.01.09