IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received May 6, 2022, accepted May 22, 2022, date of publication May 30, 2022, date of current version June 13, 2022.

*Digital Object Identifier 10.1109/ACCESS.2022.3179364*

# A New Hybrid Cryptocurrency Returns Forecasting Method Based on Multiscale Decomposition and an Optimized Extreme Learning Machine Using the Sparrow Search Algorithm

**XIAOXU DU[1], ZHENPENG TANG[2], JUNCHUAN WU[3], KAIJIE CHEN[1], AND YI CAI[1]**
[1]School of Economics and Management, Fuzhou University, Fuzhou 350108, China
[2]School of Economics and Management, Fujian Agriculture and Forestry University, Fuzhou 350002, China
[3]School of Economics and Management, Nanchang University, Nanchang 330031, China

Corresponding author: Zhenpeng Tang (zhenpt@fzu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 71973028 and Grant 71573042.

**ABSTRACT** The return series of cryptocurrencies, which are emerging digital assets, exhibit nonstationarity, nonlinearity, and volatility clustering compared to other traditional financial markets, making them exceptionally difficult to forecast. Therefore, accurate cryptocurrency price forecasting is essential for market participants and regulators. It has been demonstrated that improved data forecasting accuracy can be achieved through decomposition, but few researchers have performed information extraction on the residual series generated by data decomposition. Based on the construction of a ''decomposition-optimization-integration'' hybrid model framework, in this paper, we propose a multi-scale hybrid forecasting model that combines the residual components after primary decomposition for secondary decomposition and integration. This model uses the variational modal decomposition (VMD) method to decompose the original return series into a finite number of components and residual terms. Then, the residual terms are decomposed, and the features are extracted using the completed ensemble empirical mode decomposition with adaptive noise (CEEMDAN) method. The components are predicted by an extreme learning machine optimized by the sparrow search algorithm, and the final predictions are summed to obtain the final results. Forecasts for the returns of Bitcoin and Ethereum, which are significant cryptocurrency assets, are compared with other benchmark models constructed based on different ideas. We find that the proposed quadratic decomposition VMD-Res.-CEEMDAN-SSA-ELM hybrid model demonstrates the optimal and most stable forecasting performance in both one-step and multi-step ahead prediction of the cryptocurrency return series.

**INDEX TERMS** Cryptocurrency, model selection, decomposition-ensemble, extreme learning machine, sparrow search algorithm.

## I. INTRODUCTION

In recent years, the rise of blockchain technology in the context of the increasing integration of finance and the internet has led to the rapid development of cryptocurrency, a new type of virtual asset. Cryptocurrency uses the principles of cryptography to secure transactions on a transaction-by-transaction basis using the encryption of virtual currencies and digital hashing combined with smart contracts. Bitcoin (BTC), the first cryptocurrency, was introduced in 2008 [1]. Since then, the variety of cryptocurrencies is growing by leaps and bounds, such as Ethereum (ETH) and Ripple (XRP). Currently, there are more than 9,000 cryptocurrencies and more than 400 related exchanges, and the terms ''coin wave'' and ''chain wave'' are buzzwords. An important issue regarding cryptocurrencies is price volatility.

As seen in Figure 1, the global cryptocurrency market experienced significant volatility during the selected period,

The associate editor coordinating the review of this manuscript and approving it for publication was Bo Pu.


**FIGURE 1.** Global cryptocurrency market capitalization and trading volume (2014.03-2022.03).

especially in 2017, when a main wave of growth occurred that represented a 3,175% year-on-year increase in market capitalization, and in 2018, when the market fell by 78%. Then, during the global COVID-19 epidemic, the market experienced a nearly two-year oscillation period when the cryptocurrency market was experiencing another major market [2]. Compared to traditional financial assets, whose valuation is based on fundamental information, cryptocurrencies have a decentralized and virtual existence without a physical backing. Moreover, they are neither associated with any commodity nor with a company, and governments have no senior regulatory authority over them [3]. Due to the specificity of cryptocurrencies, the complexity of their price fluctuations can be attributed to the multiple factors and uncertainties that interact in the market, including the economic and political environment, and investor behavior that can lead to price instability. In addition, cryptocurrency trading rules are different from those of traditional financial markets. Since decentralized cryptocurrencies can be traded 24 hours a day and 7 days a week, information and events generated at any time can immediately affect the price of cryptocurrencies, rather than at specific market trading times (such as the stock market) [4]. In summary, one of the most challenging areas of time series is the accurate prediction of the trends in cryptocurrency market quotes [5], [6]. Although the cryptocurrency market is extremely complex and risky, it still represents an emerging alternative investment product with high returns and low correlation to other traditional financial assets [7]. These characteristics make cryptocurrency financial instruments that can be used to hedge against uncertainty [8]–[10]. Therefore, an accurate cryptocurrency price prediction model can deepen the grasp of cryptocurrency market price fluctuation patterns, and provide a reasonable basis for investors' investment decisions in terms of optimal hedging, option pricing, portfolio diversification, etc., as well as provide a reference for the government to formulate relevant regulatory policies [11], [12].

According to existing research, asset price volatility in financial markets is dynamic and highly nonlinear [4]. The price forecasting problem of the cryptocurrency market is similar to that of traditional stock and foreign exchange markets and is also a financial time series forecasting problem. However, due to the special trading time system of the cryptocurrency market, its price volatility is more obvious and different from other financial markets [13]. Currently, the methods involved in financial time series forecasting mainly include the following traditional econometric models, artificial intelligence methods, and hybrid models.

Traditional econometric forecasting models include linear multiple regression models [14], error correction models (ECMs)[15], autoregressive integrated moving average models (ARIMA) [16], [17]and vector autoregressive models (VAR) [18]. But the econometric models have specific assumptions, such as that the time series are trending and repeatable and that the data are stable. For data that meet these assumptions, a good prediction can be achieved. However, such models have limited predictive power for time series with nonlinear, nonstationary and volatile clustering characteristics [19].

With the rapid development of computer technology, many artificial intelligence (AI) techniques have been used in research related to time series forecasting. Common representative models include random forests [20], backpropagation neural networks (BPs) [21], artificial neural networks (ANNs) [22], Bayesian neural networks (BNNs) [23], convolutional neural networks (CNNs) [24], and support vector regressions (SVRs) [25], [26]. Extreme learning machines (ELMs) [27], as an emerging learning framework for feedforward neural networks, can overcome the training dilemma of backpropagation algorithms for single hidden layer feedforward networks (SLFNs). Due to its advantages in learning convergence speed and parameter settings and noise resistance, it has been gradually applied to classification and prediction of various complex sequences, and a series of important results have been achieved [28]–[30]. Artificial intelligence algorithms are data-driven, have addressed some limitations of traditional econometric models in forecasting to a certain extent, and have significantly improved forecasting accuracy. However, such methods are more sensitive to parameters and model settings and are prone to local optima and overfitting problems [31].

In the field of financial time series forecasting, hybrid models have become more popular forecasting methods, and their constructed frameworks have been widely used in many studies and proven to be effective in improving forecasting ability. Many researchers have built hybrid models to achieve effective forecasting for time series, including Bitcoin price forecasting [32], exchange rate forecasting [33], [34], and international crude oil price forecasting [35]. Generally, hybrid models are based on the idea of "decomposition-integration", which is divided into three steps: data decomposition, modal forecasting, and integrated learning. Unlike traditional end-to-end price forecasting methods, hybrid algorithms are used to first decompose the original data through a decomposition algorithm that extracts time-domain features of the time series [36]. Common decomposition algorithms, include empirical mode decomposition (EMD) [37], ensemble empirical mode decomposition (EEMD) [38], and complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) [39].With the development of decomposition algorithm research, the VMD [40] algorithm effectively separated components with similar frequencies, thus enabled more efficient decomposition of the original sequence and demonstrated its superior performance in dealing with complex signals disturbed by noise. After decomposing the data, each decomposed component is predicted separately by the prediction models mentioned above, such as econometric models and artificial intelligence models. Finally, the components are summed to obtain the overall prediction results.

Currently, hybrid models have achieved many results in forecasting research in related fields. Sun *et al.* [41] proposed a new carbon price prediction model based on EEMD-IBA-ELM, tested the validity and stability of the model by examining the historical carbon prices, then concluded that the proposed model could significantly improve the prediction accuracy. CAO *et al.* [42] constructed CEEMDAN to decompose the stock price series and then predicted them by LSTM, empirically obtained that CEEMDAN was more thorough in decomposition than EMD. This hybrid prediction model showed superior performance in predicting the stock price series. Zhu *et al.* [43] decomposed the carbon price into multiple modes using the VMD model, further reconstructed the modes according to the evolutionary clustering algorithm proposed by CCI, and made predictions. The final carbon price prediction results were obtained, proving that the "decomposition-clustering-prediction" method could better predict carbon prices. Jiang *et al.* [44] constructed new two-stage ensemble models by combining EMD (or VMD), ELM, and improved harmony search (IHS) algorithm for stock price prediction. The results show that the proposed model has superior performance in terms of accuracy and stability compared with other models.

The advantages of each model can be maximized through model integration. As a result, the advantages of each model can be used to overcome better the shortcomings of a single model that obtains significant differences in prediction results in different situations and under various prediction evaluation criteria [45], [46]. Therefore, more prediction information is used, which improves the prediction performance. However, the existing hybrid model construction methods still have the following shortcomings. Most of the previous integrated hybrid models decompose the original sequence into a finite number of modal components and residual terms through one decomposition, then the resulting modalities are predicted through the prediction model [47], [48]. The residual terms, which are discarded as general components, will accumulate the prediction errors generated by the decomposition, thus causing a certain degree of data distortion [36], [49], [50]. In fact, the residual term in its complex and nonlinear form may still carry valid predictive information, and further dissection of the residual term is necessary [51]. As a single forecasting model, the ELM model has advantages in terms of convergence speed and parameter settings, but the results obtained by forecasting using an unoptimized single learning machine algorithm are unstable, and the forecasting accuracy is not high. Therefore, advanced related algorithms are needed to optimize the single forecasting model to improve the accuracy and stability of the forecasting part. Many existing time series forecasts only consider one-step ahead forecasting. However, investors are also very concerned about the short-term market in the actual investment process, especially the cryptocurrency market, which is characterized by more significant volatility clustering. Therefore, multi-day-ahead forecasting is more important, and multi-step ahead forecasting can help investors provide a more comprehensive and effective reference basis.

In view of the shortcomings of the existing research and based on the inheritance of the abovementioned model construction idea and overcoming its limitations, a hybrid model is constructed in this paper that is composed of a data decomposition algorithm, an optimization algorithm, and a forecasting model, namely, the proposed VMD-Res.-CEEMDAN-SSA-ELM hybrid model, where Res. represents the residual term after VMD. The innovation of the VMD-Res.-CEEMDAN-SSA-ELM model lies in the following points.

(1) Focusing on the residual terms generated by the decomposition previous studies ignored in constructing the hybrid model, we apply the quadratic decomposition technique to combine VMD and CEEMDAN to a complex cryptocurrency return series. First, the VMD algorithm is adopted, the most effective for processing complex signals [52]. The original series is decomposed, and then the residual terms obtained after primary decomposition are taken into account. Next, the residual terms obtained after the VMD decomposition of the original sequence are decomposed further by using the CEEMDAN algorithm to extract the complex nonlinear information. The overall data characteristics of the original time series can be better understood through the secondary decomposition technique, which is more accurate and complete for the decomposition of the original data. (2) A single prediction model can vary in its predictive effectiveness in different situations. In artificial intelligence algorithms, although the ELM model has some advantages in classification and prediction studies, it depends on the input parameters. Therefore, the hidden layer neuron parameters of the ELM are optimized by introducing the cluster intelligence optimization SSA algorithm [53], which has advantages in terms of search accuracy, convergence rate, stability and avoidance of local optimization. The process achieves better stability of the prediction module and improves the prediction accuracy to compensate for the deficiencies associated with a single

prediction model. (3) While previous studies have focused more on one-step ahead forecasting, this paper applies the proposed VMD-Res.-CEEMDAN-SSA-ELM model to one-step ahead and multistep ahead forecasting of cryptocurrency returns, and the results of multi-step ahead forecasting are more closely related to the market. Moreover, the accuracy and robustness of the model in predicting complex, nonlinear, and volatility-clustered time series is verified by comparing it with benchmark models. Finally, the proposed approach is entirely data-driven and does not require excessive assumptions or consideration of exogenous influences that lead to market sentiment fluctuations, facilitating investors to make appropriate decisions that are more realistic.

The rest of this paper is presented as follows. In Section 2, the individual components and details of the VMD-Res.-CEEMDAN-SSA-ELM model are proposed. In Section 3, the daily closing price data of Bitcoin and Ethereum, which are representative among cryptocurrencies, are obtained from the CoinMarketCap website as empirical samples. The traditional single benchmark model, the integrated benchmark approach, and the VMD-Res.-CEEMDAN-SSA-ELM model constructed in this paper are compared based on one-step-ahead and five-step-ahead forecasting using evaluation metrics to test the performance of the proposed model. This paper is concluded in Section 4, and a plan for future work is presented.

## II. INTRODUCTION TO THE METHODOLOGY

A combined model (VMD-Res.-CEEMDAN-SSA-ELM) was developed based on the idea of "decomposition-ensemble" and the combination of secondary decomposition techniques with machine learning methods, aiming to predict cryptocurrency returns more accurate. Because the prediction model proposed in this paper consists of several models, the components of the model and the overall model are described below: the VMD algorithm, the CEEMDAN algorithm, the extreme learning machine, the sparrow search algorithm, and the whole new hybrid model built in this study.

### A. VARIATIONAL MODE DECOMPOSITION (VMD)

VMD is an adaptive, quasi-orthogonal, and completely nonrecursive decomposition method proposed by Dragomiretskiy and Zosso (2013). In the process of signal decomposition, the optimal center frequency and finite bandwidth of each mode can be matched adaptively by searching and solving to achieve the effective separation and frequency domain division of the characteristic mode components of the signal. Thus, the effective decomposition components of a given signal are obtained, and finally the optimal solution of the problem is obtained. The detailed VMD steps are shown as follows.

Step 1: A Hilbert transform is implemented on every modal signal to obtain a unilateral spectrum. The exponential term of the modal function corresponding to the center frequency is mixed and multiplied by the $e^{-jw_k t}$ phase to adjust the spectrum of each component signal to the fundamental frequency band. Then, the bandwidth is determined by estimating each component using the Gaussian smoothing method. The corresponding constrained variational model can be described as

$$\begin{cases} \min\limits_{\{u_k\},\{w_k\}} \left\{ \sum_k \left\| \partial_t \left[ (\delta(t) + \frac{j}{\pi t})^* u_k(t) \right] e^{-jw_k t} \right\|_2^2 \right\} \\ s.t. \ \sum_k u_k = f \end{cases} \tag{1}$$

In this equation, $\{u_k\} := \{u_1, \ldots, u_K\}$ is the mode component obtained after VMD decomposition, and $\{w_k\} := \{w_1, \ldots, w_K\}$ is the center frequency of the mode components, $VMF$. $\partial_t$ denotes the partial derivative of $t$, $\delta(t)$ refers to the shock function, and $^*$ denotes the convolution sign. $f$ is the original input signal.

Step 2: To make the signal reconstruction accurate, it must be constrained by introducing an incremental Lagrange function to convert the original equation into an unconstrained variational problem. As a result, the optimal solution is derived as follows:

$$\begin{aligned} L(\{u_k\}, \{w_k\}, \lambda) &= \alpha \sum_k \left\| \partial_t \left[ (\delta(t) + \frac{j}{\pi t})^* u_k(t) \right] e^{-jw_k t} \right\|_2^2 \\ &\quad + \left\| f(t) - \sum_k^{u_k(t)} \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \end{aligned} \tag{2}$$

where $\alpha$ is the quadratic penalty factor introduced to guarantee the accuracy of signal reconstruction when it occurs, and $\lambda$ is the Lagrange multiplier used to control the strictness of the constraints.

Step 3: Solve the variational problem by searching for the optimal solution of equation (2) using the alternating direction method of multipliers (ADMM). Equations (3) – (5) are iterated several times to obtain $\hat{u}_k^{n+1}(w)$, $\hat{w}_k^{n+1}$, and $\hat{\lambda}^{n+1}$. Futhermore, the optimal solution of the constrained variational model is obtained until the iterative condition (6) is satisfied.

$$\hat{u}_k^{n+1}(w) = \frac{\hat{f}(w) - \sum_{i \neq k} \widehat{u_i}(w) + \frac{\hat{\lambda}(w)}{2}}{1 + 2\alpha(w - w_k)^2} \tag{3}$$

$$\hat{w}_k^{n+1} = \frac{\int_0^\infty w|\hat{u}_k(w)|^2 dw}{\int_0^\infty |\hat{u}_k(w)|^2 dw} \tag{4}$$

$$\hat{\lambda}^{n+1} = \hat{\lambda}^n + \tau \left[ \hat{f}(w) - \sum_k \hat{u}_k^{n+1}(w) \right] \tag{5}$$

In these equations, $\hat{u}_k^n(w)$, $\hat{f}(w)$ and $\hat{\lambda}^n(w)$ are the Fourier transforms of $\hat{u}_k^n, f(t)$ and $\lambda^n$ respectively.

$$\frac{\sum_k \left\| \hat{u}_k^{n+1} - \hat{u}_k^n \right\|_2^2}{\left\| \hat{u}_k^n \right\|_2^2} < \varepsilon \tag{6}$$

### B. COMPLETE ENSEMBLE EMPIRICAL MODE DECOMPOSITION WITH ADAPTIVE NOISE (CEEMDAN)

Torres *et al.* (2011) proposed a CEEMDAN algorithm based on EMD and EEMD. The algorithm effectively suppresses the mode mixing of EMD by adding finite times of adaptive white noise at each stage. It can achieve a more thorough decomposition of the signal data with a more minor reconstruction error by removing noise residuals with fewer averaging times. The decomposition steps of CEEMDAN are as follows:

Step 1: Let $x(n)$ be the original signal sequence, $\varepsilon$ be the adaptive coefficient, $w_i(n)$ be the noise sequence added for each decomposition, and $x_i(n)$ be the signal sequence after $i$ times of noise is added. The average value of $N$ sub-experiments of EMD decomposition is the first intrinsic mode component $IMF_1$,

$$x_i(n) = x(n) + \varepsilon w_i(n) \tag{7}$$

$$IMF_1(n) = \frac{1}{N}\sum_{i=1}^{N} IMF_{1i}(n) \tag{8}$$

Step 2: Calculate the residual sequence $r_1(n)$ of the first stage and obtain a new $r_1^i(n)$ for the $N$ sub-experiments until the EMD decomposition finishes its work on the $IMF$ component.

$$r_1(n) = x(n) - IMF_1(n) \tag{9}$$

Step 3: Based on Step 2, calculate the second intrinsic mode component $IMF_2$.

$$IMF_2(n) = \frac{1}{N}\sum_{i=1}^{N} IMF_1\{r_1(n) + \varepsilon_1 IMF_1[w_i(n)]\} \tag{10}$$

Step 4: Repeat the calculation to the stage $k+1$. We can obtain the residual sequence $r_k(n)$ at stage and the $k+1$th intrinsic modal component $IMF_{k+1}$.

$$r_k(n) = r_{k-1}(n) - IMF_k(n) \tag{11}$$

$$IMF_{k+1}(n) = \frac{1}{N}\sum_{i=1}^{N} IMF_1\{r_1(n) + \varepsilon_k IMF_k[w_i(n)]\} \tag{12}$$

Step 5: The above steps are repeated. If the number of extreme points of the residual sequence is $\leq 2$, the EMD is stopped, then the final residual sequence $R(n)$ and intrinsic mode component $IMF_k$ are obtained. Finally, the initial signal sequence $x(n)$ is decomposed as

$$x(n) = \sum\nolimits_{i=1}^{K} IMF_k(n) + R(n) \tag{13}$$

### C. EXTREME LEARNING MACHINE OPTIMIZED BY THE SPARROW SEARCH ALGORITHM (SSA-ELM)

#### 1) EXTREME LEARNING MACHINE (ELM)

Huang *et al.* (2004) proposed an ELM algorithm to solve single hidden-layer feedforward neural networks. It mainly uses the generalized inverse theory of matrices. Compared with the traditional neural network learning algorithm, only a unique optimal solution must be generated in ELM. The process is achieved by setting the number of hidden layer nodes of the network. The input weights and biases do not need to be adjusted during execution; therefore, the advantages of ELM are fast learning and good generalization performance. The ELM network model is illustrated in Figure 2.

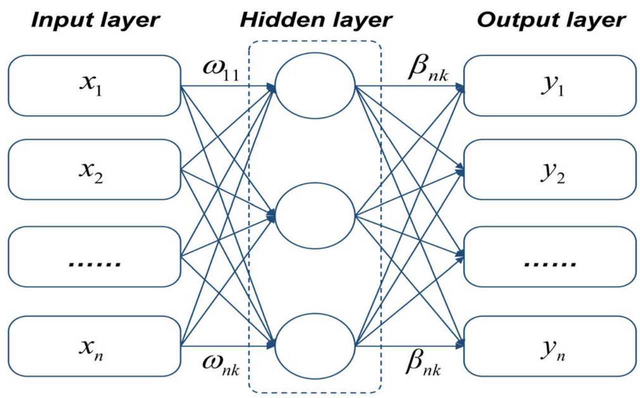


**FIGURE 2.** Extreme learning machine process.

In Figure 2, $x_1 \sim x_n$ are the nodes of the input neuron, $\omega_{11} \sim \omega_{nk}$ are the weights between the input layer and the hidden layer, $g(x)$ is the activation function, $b_1 \sim b_k$ are the hidden layer node thresholds, $\beta_{11} \sim \beta_{nk}$ are the weights between the hidden layer and the output network model layer, and $y_1 \sim y_n$ are the outputs of the model.

Suppose there are $N$ training samples $\{(x_i, y_i)\}_{i=1}^{N}$, and $\boldsymbol{x}_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^{\boldsymbol{T}} \in \boldsymbol{R}^n$ refers to the n-dimensional input data of the training set. $\boldsymbol{t}_i = [t_{i1}, t_{i2}, \ldots, t_{im}]^{\boldsymbol{T}} \in \boldsymbol{R}^m$ is the m-dimensional ideal output value of the training set. The ELM network model expression of the hidden layer nodes $K$ and assuming an activation function of $g_i(x_i)$ can be expressed as

$$y_j = \sum_{i=1}^{K} \beta_i g_i(\omega_i x_j + b_i), \quad j = 1, 2 \ldots, K \tag{14}$$

In this equation, $\omega_i$ is the input weight vector connected between the input layer nodes and the first hidden layer node $i$. $\beta_i$ is the weight matrix between the $i$th hidden layer and the output layer. $b_i$ is the threshold value of the $i$th hidden layer node. $y_j$ is the actual output of the network model. And $g_i\left(\omega_i \boldsymbol{x}_j + b_i\right)$ is the activation function.

If the single-feeder neuron network in the hidden layer can approach any training sample with zero error, it can be expressed as

$$\sum_{i=1}^{N} \|y_i - t_i\| = 0 \tag{15}$$

Then we obtained

$$t_j = \sum_{i=1}^{K} \beta_i g_i(\omega_i x_j + b_i), \quad j = 1, 2 \ldots, N \tag{16}$$

which can be a matrix represented as

$$\boldsymbol{H\beta} = \boldsymbol{T} \tag{17}$$

In this equation,

$$\boldsymbol{H} = \begin{bmatrix} g_1(\omega_1 x_1 + b_1) & \cdots & g_i(\omega_K x_1 + b_K) \\ \vdots & \ddots & \vdots \\ g_i(\omega_1 x_N + b_1) & \cdots & g_i(\omega_K x_N + b_K) \end{bmatrix}_{N\times K} \tag{18}$$

$\boldsymbol{H}$ stands for the hidden-layer output matrix and $\boldsymbol{T}$ for the ideal output vector.

Therefore, the optimal solution of $\boldsymbol{H\beta} = \boldsymbol{T}$ is obtained, which is given by

$$\hat{\boldsymbol{\beta}} = \boldsymbol{H}^{+}\boldsymbol{T} \tag{19}$$

In this equation, $\boldsymbol{H}^{+}$ refers to the augmentation matrix of the matrix $\boldsymbol{H}$.

The entire training process needs to be run only once to obtain the optimal solution, making the ELM's generalization ability very strong.

### 2) SPARROW SEARCH ALGORITHM (SSA)

The sparrow search algorithm (SSA) is a new intelligent optimization algorithm proposed by Xue *et al.* in 2020 that idealizes and formulates the corresponding rules for the predatory behavior of sparrow groups. This algorithm assumes two types of sparrows: discoverers and foragers. Discoverers actively search for food, and foragers obtain food from discoverers. In addition, some predators can grab food.

The role of the discoverer is to guide the entire sparrow population in searching and predating. And its position can be expressed by an equation,

$$X_{i,d}^{t+1} = \begin{cases} X_{i,d}^{t} \cdot \exp\left(\frac{-i}{\alpha \cdot iter_{\max}}\right) & R_2 < ST \\ X_{i,d}^{t} + Q \cdot L & R_2 \geq ST \end{cases} \tag{20}$$

In this equation, $t$ is the current number of iterations; $X_{i,d}^{t+1}$ is the location of the $i$th sparrow in the $t+1$th iteration of the $d$th dimension; $T$is the maximum number of iterations. $\alpha \in [0, 1]$ is a random number; $Q$is a random number that obeys a normal distribution; $R_2$ and $ST$ are the warning and safety values, respectively. And $L$ is a matrix of size $1 \times d$ where each element is 1.

If $R_2 < ST$, this means that predators are not nearby, so the discoverers can perform a wide-area search. If $R_2 \geq ST$, that is, the predators have been found, the rest of the sparrows need to leave their present position.

Foragers should observe the discoverers during this process, and when an abundant food source is noted, the foragers will leave their location to compete for food. If the scramble is successful, they will receive food from the finder; thus, the foragers' positions are updated as

$$X_{i,d}^{t+1} = \begin{cases} Q \cdot exp\left(\frac{X_{worst}^{t} - X_{i,j}^{t}}{i^2}\right) & i > \frac{n}{2} \\ X_P^{t+1} + \left|X_{i,d}^{t} - X_P^{t+1}\right| \cdot \boldsymbol{A}^{+} \cdot L & other\ situation \end{cases} \tag{21}$$

In this equation, $X_{worst}^{t}$ is the worst position of discoverers of the entire process in the $t$th iteration, $X_P^{t+1}$ is the best position of the discoverers in the $t+1$th iteration, and $\boldsymbol{A}$ is an dimensional matrix with the same dimensions as the input. Each element is randomly assigned 1 or $-1$, and $\boldsymbol{A}^{+} = \boldsymbol{A}^{T}(\boldsymbol{A}\boldsymbol{A}^{T})^{-1}$; $n$ is the number of sparrows.

Usually, some foragers will act as warning sparrows to help the discoverer forage, and when in danger, they will counter trap or withdraw, close to other sparrows.

$$X_{i,d}^{t+1} = \begin{cases} X_{best}^{t} + \lambda \cdot \left|X_{i,d}^{t} - X_{best}^{t}\right| & f_i > f_g \\ X_{i,d}^{t} + J\left(\frac{X_{i,d}^{t} - X_{worst}^{t}}{(f_i - f_w) + \varepsilon}\right) & f_i = f_g \end{cases} \tag{22}$$

In this equation, $X_{best}^{t}$ is the best position of warners in the $t$-th iteration; as a step control parameter, $\lambda$ is a random number that obeys a normal distribution with an average of 0 and variance of 1; $J \in [-1, 1]$ is a random number; $f_i$ is the current sparrow's fitness value; $f_g$ and $f_w$ are the best fitness value and the worst fitness value, respectively; $\varepsilon$ is defined as a tiny constant, which is mainly used to prevent the case of the $f_i - f_w = 0$. The main steps of the sparrow algorithm are as follows.

Step 1: Initialize the population, set the total number of sparrow population $n$, number of discoverers, number of warners, the maximum number of iterations $T$, and alarm threshold $R_2$.

Step 2: Use mean squared error (MSE) as the fitness function and then calculate the fitness value of each sparrow. Find and define the best and worst fitness values as $f_g$ and $f_w$, respectively.

Step 3: The new positions of the discoverers, foragers, and warners are calculated using equations (20) – (22), and if the fitness value of the new position is greater than that of the previous position, it is updated.

Step 4: Perform iterations and repeat step 3 to continuously update the positions of sparrows, stopping when the number of iterations is $T$. Therefore, the position of the sparrow with the lowest fitness value in all iterations is the optimal solution.

### 3) THE EXTREME LEARNING MACHINE FOR SPARROW SEARCH ALGORITHM OPTIMIZATION (SSA-ELM)

ELM can be used for nonlinear function fitting and prediction problems with small-sample learning. However, the stability of model training can be affected by its input weights and implied layer thresholds. The SSA algorithm has the advantages of high search accuracy, fast convergence, and good stability, so that the SSA algorithm can optimize the ELM input parameters and weights to improve the prediction efficiency and obtain more stable prediction results. The detailed operation flow of the SSA-ELM model is as follows:

Step 1: The sparrow population is initialized as the discoverer, forager, and predator. The corresponding fitness value $P_b$ of each sparrow is calculated separately, and the best fitness is defined along with the position of the corresponding sparrow as $X_{best}$.

Step 2: Iterations are performed to determine the optimal initial weights and thresholds. These data can be obtained by comparing the value of fitness function MSE. More specifically, when exercising the second iteration, the minimum MSE of the current sparrow generations should be compared with the optimal adaptive value of the previous optimal fitness value $P_b$. If it is less than $P_b$, the optimal fitness value $P_b$ must be updated to the minimum MSE of the current generation of sparrows, and the position of this sparrow should be updated to the optimal position $X_{best}$. Otherwise, the optimal adaptation value $P_b$ and the optimal position $X_{best}$ need not be updated, and the next iteration can be performed.

Step 3: The iteration should be stopped until it reaches the set value $iter_{\max}$. Finally, the optimal weights and thresholds obtained from the model optimized by the sparrow algorithm are used to construct a new ELM model for prediction. The specific steps of the SSA-ELM are shown in figure 3.

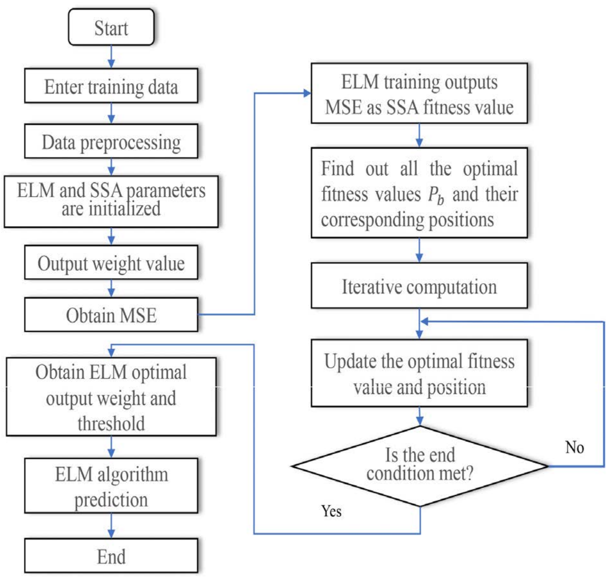


**FIGURE 3.** SSA-ELM process.

### D. VMD-RES.-CEEMDAN-SSA-ELM MODEL

As previously mentioned, cryptocurrency return series have complex characteristics, such as typical nonstationary, nonlinear, and volatility clustering, with limited accuracy when using a single forecasting method. Because the VMD decomposition technique can decompose the complex signal into several mode components with much lower complexity, the prediction accuracy is substantially improved when each modal component obtained from the VMD decomposition is modeled separately through common forecasting methods. Therefore, we adopted the VMD algorithm to decompose the original sequence firstly. And then, based on the first decomposition, a secondary decomposition of the residual terms by CEEMDAN is considered, which in turn enables a more thorough decomposition of the original sequence. Finally, ELM is used to optimize the SSA. The advanced SSA-ELM model can improve the prediction accuracy, convergence speed, and stability of cryptocurrency returns to a certain extent. The specific modeling steps are shown in Figure 4.

Step 1: First decomposition. Decompose the original series into each mode component $\mathbf{VMF}_i$ through the VMD decomposition technique, then subtract each mode component from the original series to obtain the residual series (Res.).

Step 2: Secondary decomposition. CEEMDAN was applied to decompose the residual series further to obtain another set of subseries $\mathbf{IMF}_i$. Then we normalized the decomposed components $\mathbf{VMF}_i$ and $\mathbf{IMF}_i$.

Step 3: Forecast process. The modal components obtained from the decomposition of the original series and the residual term decomposition are predicted by the SSA-ELM.

Step 4: Ensemble. The predictions of the residuals are superimposed with the predictions of each $\mathbf{VMF}_i$ to obtain the final forecast result.

## III. EMPIRICAL ANALYSIS

### A. DATA DESCRIPTION AND EVALUATION CRITERIA

Bitcoin and Ethereum account for nearly 66% of the market capitalization in the cryptocurrency market and enjoy the majority of the daily trading volume, even reaching more than 70% of the whole market on June 30, 2021. Therefore, in this study, the log returns of the daily closing prices of Bitcoin and Ethereum were selected as the prediction objects. The log returns of the $t$th trading day, $r_t = [\ln(p_t) - \ln(p_{t-1})] \times 100$, and the returns of the above two virtual currencies are predicted using the proposed VMD-Res.-CEEMDAN-SSA-ELM to verify the model's effectiveness, and the daily closing prices of BTC and ETH were obtained from the web (http://www.CoinMarketCap.com/). Combined with past data, awareness and interest in cryptocurrencies was not high until 2017, after which cryptocurrency assets such as Bitcoin really caught the attention of investors and academics. Trading volume data also confirmed this trend. Therefore, this study selects the return rates of BTC and ETH from January 1, 2017, to June 30, 2021, with 1,642 returns data each.

In the respective returns datasets of Bitcoin and Ethereum, the training and test sets were divided; the first 1,492 returns data were used as the training set, and the remaining 150 data were used as the test set. Table 1 lists the descriptive information related to Bitcoin and Ethereum returns data. The empirical operation in this study was completed by MATLAB 2019b.

In this study, four evaluation metrics, mean absolute error (MAE), normalized root mean square error (NRMSE), and symmetric mean absolute percentage error (SMAPE), coefficient of determination ($R^2$) were selected to test the prediction effectiveness of the models. In addition, to more concisely compare the differences in evaluation metrics between different benchmark models and the proposed model, we define the following three evaluation metrics relative to the proposed

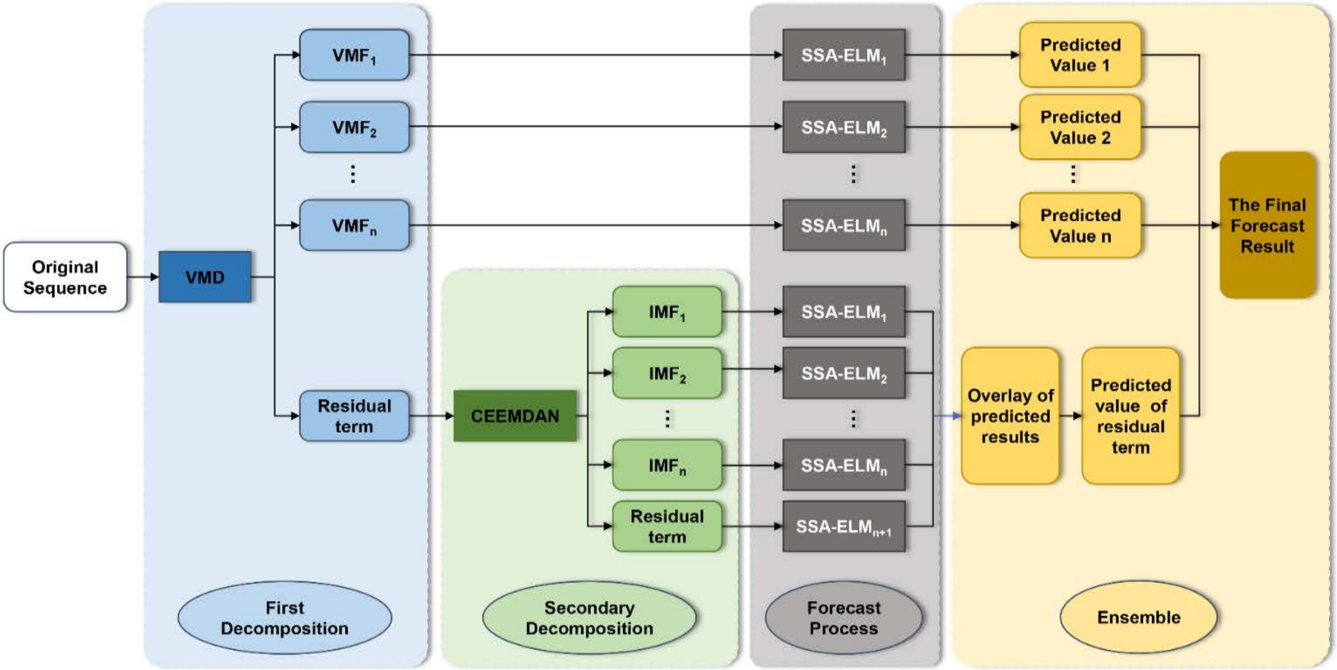


**FIGURE 4.** Structure of the VMD-Res.-CEEMDAN-SSA-ELM model.

**TABLE 1.** Descriptive statistics for Bitcoin and Ethereum return data.

| Cryptocurrency | | Statistics | | | | | |
|---|---|---|---|---|---|---|---|
| | | Mean | Maximum | Minimum | Median | Standard deviation | Sample Length |
| BTC | Train set | 0.2371 | 22.5119 | -46.4730 | 0.2476 | 4.2630 | 1492 |
| | Test set | 0.0377 | 17.1820 | -14.8107 | -0.0143 | 4.7594 | 150 |
| ETH | Train set | 0.3422 | 29.0130 | -55.0714 | 0.1408 | 5.7034 | 1492 |
| | Test set | 0.3653 | 22.5649 | -31.7459 | 0.6860 | 6.4284 | 150 |

model. Table 2 presents the definitions and formulas for the relevant evaluation indicators.

### B. DATA PROCESSING

In practical applications, the optimal number of modal components cannot be directly determined when decomposing the original time series through VMD decomposition because of the admixture of noise in the original time series. Therefore, the average instantaneous frequency observation method was used in this study to determine the optimal K value. For Bitcoin returns series data, the average instantaneous frequency decreases less at the end when the value of K is 11, and that is over decomposition. Therefore, the optimal number of VMD decomposition modes for the Bitcoin returns series was 10. Similarly, the optimal number of components for VMD decomposition of the Ethereum returns series was 12. After determining parameter K, time series data decomposition was performed.

Then, the sum of each mode component generated by the VMD decomposition is subtracted from the original sequence to obtain the residual sequence. Due to its complexity, the residual series using a predictive algorithm is complicated to predict accurately. Hence, previous studies typically neglected this series. To a certain extent, this operation tends to lead to a loss of information. Thus, the secondary decomposition technique was adopted in this study to extract more available information, and the complex residual series was further decomposed using CEEMDAN technology. Taking the decomposition process of the Bitcoin return sequence as an example, the decomposition process is shown in Figure 5.

In addition, due to the large span of values of individual features, the differences in their units and magnitudes led to features that are not comparable with each other. Therefore, the data needed to be normalized for each decomposed subseries of modal component data before being predicted using machine learning methods.

In this study, the data were linearly altered using the minimum-maximum deviation normalization method with the following expressions:

$$\tilde{x} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{23}$$

where $x$ denotes the original feature data, $\tilde{x}$ denotes the standardized subseries data, $x_{max}$ represents the maximum value in the original sequence, and $x_{min}$ denotes the minimum value.

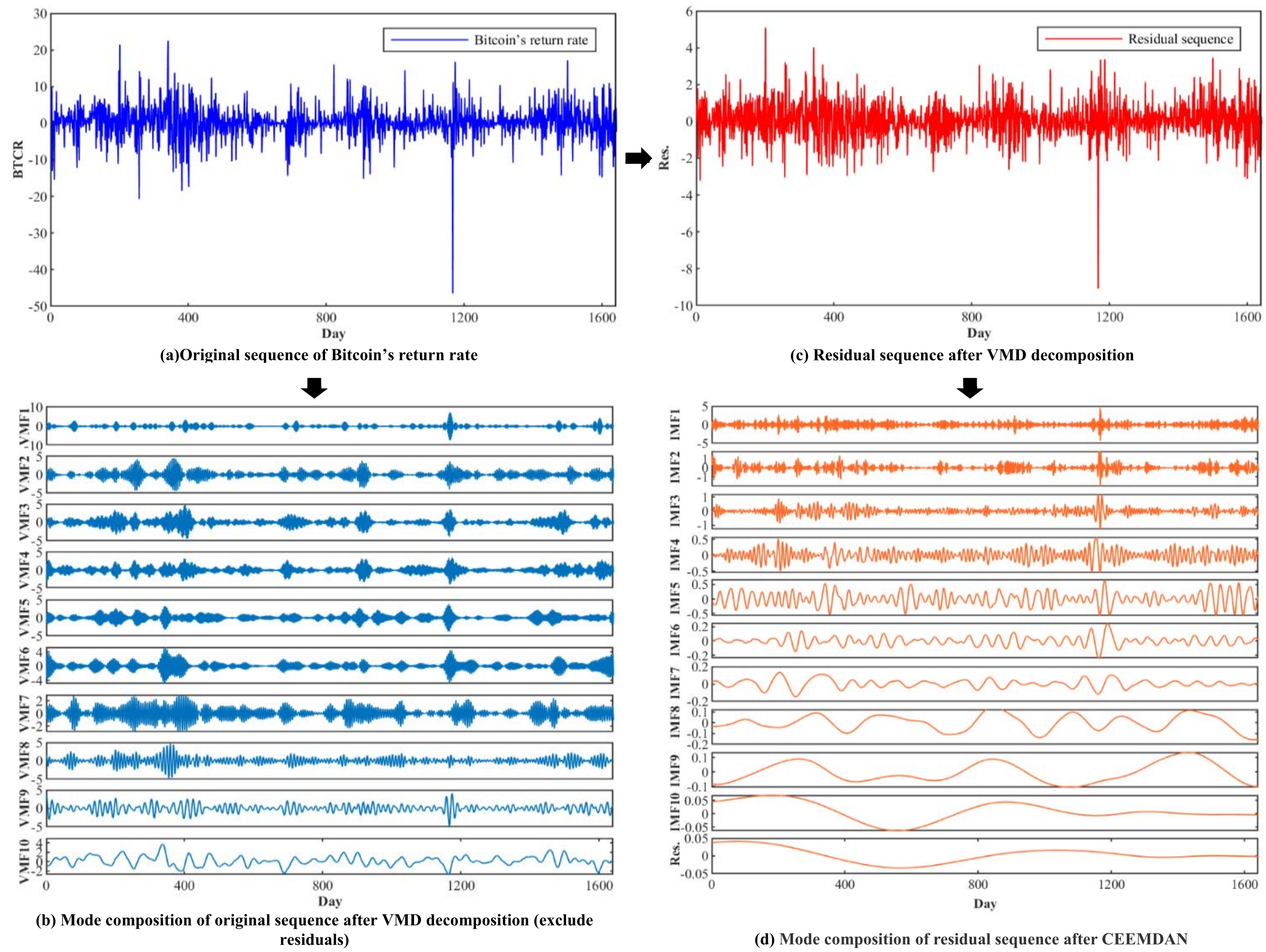


**FIGURE 5.** Mode components of the original bitcoin return series decomposed by VMD-Res.-CEEMDAN.

### C. FORECAST RESULTS

To verify the validity of the proposed model, we compare the predictive validity and stability of the proposed model with those of the basic models.

First, we test the forecasting performance differences between the hybrid and single forecasting models and select the optimal basic model for the forecasting module. We introduce traditional econometric models and artificial intelligence models as basic models to determine the differences in forecasting performance between single forecasting models. The intelligent cluster SSA algorithm is also selected to optimize the above better basic prediction models, so that the model with the best prediction performance can be selected as the main component of the hybrid model.

Second, to examine the difference in prediction performance between the proposed quadratic decomposition model considering residual terms compared with the commonly used methods in the general decomposition-integration framework, we empirically modeled the commonly used methods in each possible combination. Specifically, in the modal decomposition stage, EMD, CEEMDAN, or VMD can be used to decompose the return series. At the same time, due to different modeling ideas, scholars have been able to improve the forecasting accuracy by decomposing the high-frequency components of the primary decomposition quadratically and integrating the forecasts, and we correspondingly incorporate such models. Finally, based on the residual terms generated after considering the primary decomposition proposed in this paper, we construct an integrated forecasting model combining the residual terms.

Based on the above considerations, ten additional benchmark models are constructed in this paper compared with the proposed models. Specifically: (1) single prediction models: ARIMA, BP, SVR, ELM, SSA-ELM; (2) hybrid prediction models: EMD-SSA-ELM, CEEMDAN-SSA-ELM, CEEMDAN-VMD-SSA-ELM, VMD-SSA-ELM, VMD-Res.-SSA-ELM.

In addition, one-step and multistep ahead forecasting was performed in this paper by one-step-ahead and five-step-ahead, i.e., the data of the first six trading days allow forecasting the return of the next 1 and 5 trading days, respectively.

**TABLE 2.** Relevant evaluation indicators.

| Evaluation Indicator | Definition | Equation |
|---|---|---|
| MAE | Mean Absolute Error | $\frac{\sum_{i=1}^{n}\lvert y_i-\hat{y}_i\rvert}{n}$ |
| NRMSE | Normalized Root Mean Squared Error | $\frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}}{y_{\max}-y_{\min}}$ |
| SMAPE | Symmetric Mean Absolute Percentage Error | $\frac{1}{n}\sum_{i=1}^{n}\frac{\lvert y_i-\hat{y}_i\rvert}{(\lvert y_i\rvert+\lvert\hat{y}_i\rvert)/2}$ |
| $R^2$ | Coefficient of Determination | $R^2=1-\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2}$ |
| $R_{MAE}$ | Reducting percentages of Mean Absolute Error | $R_{MAE}=\frac{MAE_b-MAE_p}{MAE_b}$ |
| $R_{NRMSE}$ | Reducting percentages of Normalized Root Mean Squared Error | $R_{NRMSE}=\frac{NRMSE_b-NRMSE_p}{NRMSE_b}$ |
| $R_{SMAPE}$ | Reducting percentages of Symmetric Mean Absolute Percentage Error | $R_{SMAPE}=\frac{SMAPE_b-SMAPE_p}{SMAPE_b}$ |

Note: n denotes the length of the time series, $y_i$ is the actual value at the time i, and $\hat{y}_i$ is the prediction value at the time i. Furthermore, b is the benchmark model, and p is the proposed model. The smaller the values of the three evaluation indicators MAE, NRMSE, and SMAPE mean the better the prediction effectiveness of the model. And the bigger values of the $R^2$ mean a better overall fit. However, the bigger values of $R_{MAE}$, $R_{NRMSE}$, and $R_{SMAPE}$ imply a more significant improvement of the proposed model compared to the basic model.

The evaluation metrics of the prediction results are shown in Table 3.

Figures 6 (a) - (d) show the true values and prediction results of the different models for Bitcoin and Ethereum in one-step and five-step ahead.

## *D. MODEL COMPARISON AND ANALYSIS*

### 1) NONCOMBINED MODELS

First, the common econometric and artificial intelligence models ARIMA, BP, SVR, and ELM are used as benchmark prediction models to predict the return data of BTC and ETH, as shown in Table 3. From the overall performance, under the comparison of the four single models, it can be seen that ELM performs the best in all evaluation metrics in the prediction of Bitcoin and Ethereum 1 day ahead and 5 days ahead. ARIMA, on the other hand, has the worst prediction performance. The possible reason is that ARIMA, as a classical linear model, has difficulty capturing the pattern due to cryptocurrency return data's the non-linear and high volatility characteristics.

Furthermore, we introduce the SSA optimization algorithm combined with ELM. The four evaluation indicators of SSA-ELM improve to a certain extent compared with other single forecasting models when forecasting complex financial time series with high volatility, such as cryptocurrency returns. Therefore, we adopt SSA-ELM as the primary model for the forecasting module in the following construction of the hybrid model.

However, in a comprehensive view, the $R^2$ indicators of the overall fit of the single forecasting model starting from a data-driven approach do not perform well. They cannot effectively capture the complex cryptocurrency return data characteristics.

### 2) COMBINED MODEL WITHOUT CONSIDERING RESIDUAL TERM DECOMPOSITION

By comparing the forecasting results of the common hybrid models EMD-SSA-ELM, CEEMDAN-SSA-ELM, CEEMDAN-VMD-SSA-ELM and VMD-SSA-ELM constructed under the decomposition-based integration framework. It is easy to find that the combined model has significantly improved in four evaluation metrics, MAE, NRMSE, SMAPE and $R^2$, compared with the single prediction model. The prediction accuracy and stability of the combined model with the decomposition technique are better than those of the single prediction model.

Furthermore, these combined models are viewed separately. The evaluation indices of the prediction results of CEEMDAN-SSA-ELM are generally better than those of EMD-SSA-ELM, which shows that the CEEMDAN decomposition is more complete than the EMD decomposition for data decomposition, and thus makes the data features extraction more adequate. The VMD-SSA-ELM achieves the best prediction results in this stage. In this regard, it can be shown that the VMD decomposition technique has stronger decomposition ability for time series with high complexity and high volatility like cryptocurrency prices, and can better extract serial data features and handle complex signals.

### 3) COMBINED MODEL CONSIDERING RESIDUAL TERM DECOMPOSITION

The VMD-Res.-SSA-ELM and the proposed model consider the residual terms after the first decomposition and incorporate the prediction of the resulting residual series. The difference is that the model proposed in this paper further decomposes the residual terms generated from the primary decomposition to obtain the corresponding components, and integrates the predictions to obtain the final prediction results. The prediction results of these two models are shown in Table 3. The predictive evaluation metrics obtained from the model proposed in this paper are further improved than the VMD-Res.-SSA-ELM. It can be seen that the residual series generated after the original return series decomposed by VMD also contains important and complex information, and the direct use of the SSA-ELM model to forecast the residual terms directly has a limited effect. For example,

**TABLE 3.** Comparison of the one-step and five-step ahead prediction results of different models for BTC and ETH returns return data.

| Cryptocurrency | Model | One-step ahead prediction | | | | Five-step ahead prediction | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | MAE ($R_{MAE}$) | NRMSE ($R_{NRMSE}$) | SMAPE ($R_{SMAPE}$) | $R^2$ | MAE ($R_{MAE}$) | NRMSE ($R_{NRMSE}$) | SMAPE ($R_{SMAPE}$) | $R^2$ |
| BTC | ARIMA | 3.5206 (82.49%) | 0.1491 (84.05%) | 0.9260 (66.02%) | -0.0045 | 3.5303 (68.44%) | 0.1493 (71.04%) | 0.9651 (60.22%) | -0.0070 |
| | BP | 3.5040 (82.41%) | 0.1486 (84.00%) | 0.9320 (66.23%) | 0.0022 | 3.5079 (68.24%) | 0.1487 (70.92%) | 0.9481 (59.51%) | 0.0009 |
| | SVR | 3.4692 (82.24%) | 0.1481 (83.94%) | 0.9242 (65.95%) | 0.0095 | 3.4883 (68.06%) | 0.1485 (70.89%) | 0.9440 (59.33%) | 0.0012 |
| | ELM | 3.4581 (82.18%) | 0.1478 (83.90%) | 0.9234 (65.92%) | 0.0134 | 3.4851 (68.03%) | 0.1483 (70.85%) | 0.9389 (59.11%) | 0.0059 |
| | SSA-ELM | 3.4420 (82.09%) | 0.1475 (83.87%) | 0.9035 (65.17%) | 0.0176 | 3.4761 (67.95%) | 0.1482 (70.83%) | 0.9219 (58.36%) | 0.0076 |
| | EMD-SSA-ELM | 2.6954 (77.14%) | 0.1065 (77.67%) | 0.7975 (60.54%) | 0.4873 | 3.6302 (69.31%) | 0.1470 (70.59%) | 0.9172 (58.14%) | 0.0233 |
| | CEEMDAN-SSA-ELM | 2.5622 (75.95%) | 0.1011 (76.48%) | 0.7707 (59.17%) | 0.5379 | 3.4664 (67.86%) | 0.1416 (69.46%) | 0.8715 (55.95%) | 0.0943 |
| | CEEMDAN-VMD-SSA-ELM | 1.0199 (39.57%) | 0.0431 (44.78%) | 0.4207 (25.20%) | 0.9162 | 2.2919 (51.39%) | 0.0949 (54.45%) | 0.6260 (38.67%) | 0.5930 |
| | VMD-SSA-ELM | 0.9922 (37.89%) | 0.0383 (37.98%) | 0.4015 (21.62%) | 0.9336 | 1.1985 (7.03%) | 0.0501 (13.62%) | 0.4121 (6.84%) | 0.8868 |
| | VMD-Res.-SSA-ELM | 0.8844 (30.31%) | 0.0346 (31.32%) | 0.3543 (11.18%) | 0.9458 | 1.1616 (4.08%) | 0.0473 (8.58%) | 0.4083 (5.98%%) | 0.8989 |
| | **VMD-Res.-CEEMDAN-SSA-ELM** | **0.6163** | **0.0238** | **0.3147** | **0.9744** | **1.1142** | **0.0432** | **0.3839** | **0.9155** |
| ETH | ARIMA | 4.7814 (83.10%) | 0.1190 (84.45%) | 0.9394 (70.96%) | -0.0103 | 4.7817 (69.45%) | 0.1189 (70.60%) | 0.9413 (58.66%) | -0.0092 |
| | BP | 4.7438 (82.97%) | 0.1184 (84.37%) | 0.9391 (70.95%) | -0.0005 | 4.7522 (69.26%) | 0.1187 (70.54%) | 0.9311 (58.21%) | -0.0049 |
| | SVR | 4.7375 (82.94%) | 0.1171 (84.20%) | 0.9343 (70.80%) | 0.0220 | 4.7453 (69.22%) | 0.1183 (70.45%) | 0.9554 (59.27%) | 0.0009 |
| | ELM | 4.7032 (82.82%) | 0.1169 (84.17%) | 0.9257 (70.53%) | 0.0242 | 4.7185 (69.05%) | 0.1179 (70.35%) | 0.9213 (57.77%) | 0.0074 |
| | SSA-ELM | 4.6413 (82.59%) | 0.1158 (84.02%) | 0.9001 (69.69%) | 0.0420 | 4.6547 (68.62%) | 0.1174 (70.22%) | 0.9103 (57.26%) | 0.0154 |
| | EMD-SSA-ELM | 3.2330 (75.01%) | 0.0773 (76.07%) | 0.6716 (59.38%) | 0.5738 | 4.4833 (67.42%) | 0.1116 (68.68%) | 0.8572 (54.61%) | 0.1104 |
| | CEEMDAN-SSA-ELM | 3.0361 (73.39%) | 0.0721 (74.34%) | 0.6489 (57.96%) | 0.6287 | 4.4252 (66.99%) | 0.1115 (68.65%) | 0.8445 (53.93%) | 0.1127 |
| | CEEMDAN-VMD-SSA-ELM | 1.2655 (36.15%) | 0.0338 (45.27%) | 0.3628 (24.81%) | 0.9184 | 2.8797 (49.28%) | 0.0727 (51.91%) | 0.6108 (36.30%) | 0.6228 |
| | VMD-SSA-ELM | 1.2027 (32.82%) | 0.0281 (34.16%) | 0.3321 (17.86%) | 0.9437 | 1.5710 (7.03%) | 0.0386 (9.44%) | 0.4014 (3.06%) | 0.8939 |
| | VMD-Res.-SSA-ELM | 1.1557 (30.09%) | 0.0271 (31.73%) | 0.3323 (17.91%) | 0.9477 | 1.5529 (5.94%) | 0.0380 (8.01%) | 0.3962 (1.79%) | 0.8984 |
| | **VMD-Res.-CEEMDAN-SSA-ELM** | **0.8080** | **0.0185** | **0.2728** | **0.9755** | **1.4606** | **0.0350** | **0.3891** | **0.9128** |

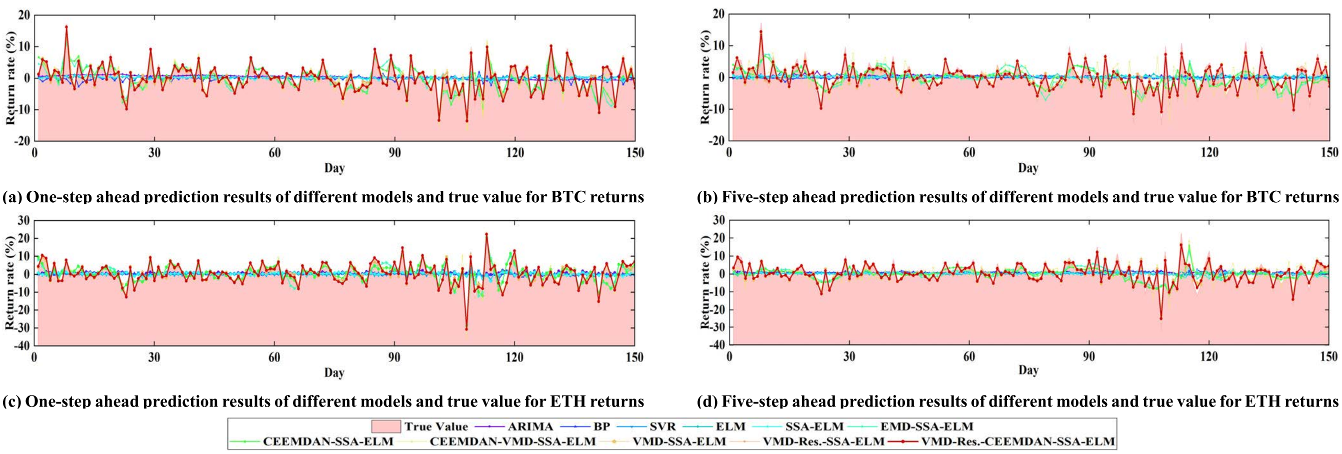


**FIGURE 6.** True values, one-step and five-step ahead prediction results of the different models for BTC and ETH.

for the one-step ahead prediction of bitcoin returns, the proposed model achieves an overall fit evaluation index $R^2$ of 0.9744.

In general, the proposed model performs best in all evaluation metrics compared to all other benchmark models. Therefore, the quadratic decomposition technique proposed

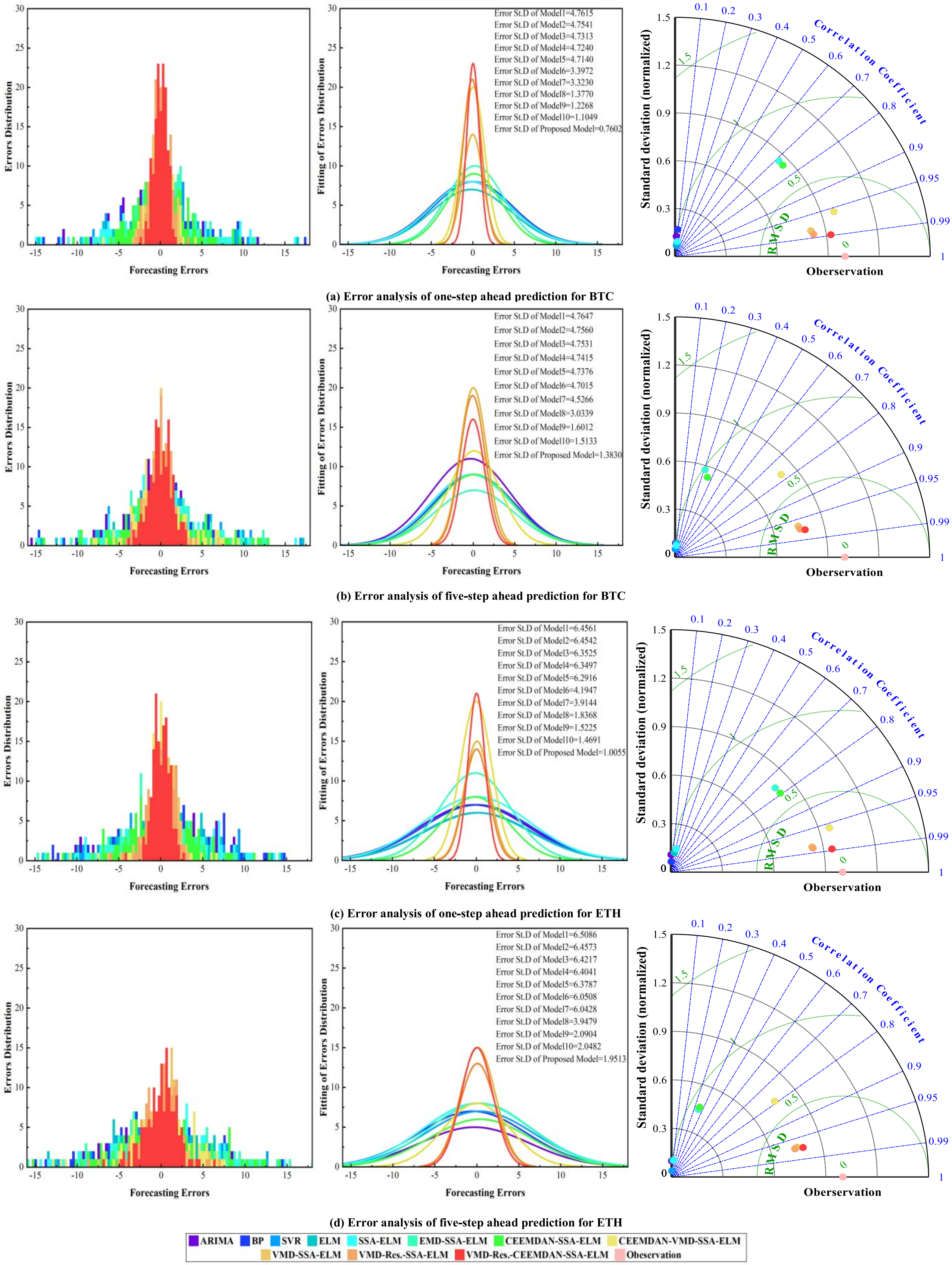


**FIGURE 7.** Error analysis of multi-step predictions for all models in cases of BTC and ETH.

in this paper for the residual term appears to be necessary. The original sequence is decomposed twice by VMD and CEEMDAN, which can effectively combine the advantages of the two algorithms to grasp the original sequence's characteristics better, and thus can be combined with the SSA-ELM prediction module to obtain more accurate prediction results further.

#### 4) ANALYSIS OF FORWARDING MULTISCALE FORECASTING

A longitudinal comparison of multiscale forward forecasting shows that the forecasting performance of both the benchmark model and the proposed multiscale hybrid model gradually decreased as the forecasting scale increased. The result implies that all models outperformed the one-step ahead prediction scenario in terms of predictive power over the five-step ahead prediction. This is mainly because as the forecasting scale increases, the complexity and high volatility of the cryptocurrency return series increase. Therefore, the forecasting accuracy 5 days in advance is lower than forecasting 1 day in advance. In addition, some data information is not trained by the model in actual Bitcoin and Ethereum returns forecasting, which leads to the gradual weakening of the model's forecasting ability and explains the increased difficulty of forecasting cryptocurrency return series multiple steps ahead compared to one day ahead.

#### 5) FORECAST ERROR ANALYSIS

To further understand and compare the forecast error distribution between the models, we analyze the forecast errors generated by the benchmark models and the proposed model in the empirical study of one-step and five-step ahead prediction scenarios for major cryptocurrency returns. Figures 7 (a) - (d) show the forecast error distributions and the corresponding fitted error distribution curves of the benchmark and proposed models, and then depict Taylor plots with the error data. It is worth mentioning that Taylor diagrams are used to facilitate the analysis of correlation measures between different models by presenting statistical information such as standard deviation, correlation coefficient, and root mean square deviation (RMSD) generated by each different forecasting model in a single plot during the actual forecasting process [54]. The Taylor plots plotted in the paper are generated by normalizing the standard deviation and root mean square deviation to make the comparison between models more intuitive.

We combine the results of the one-step and five-step ahead forecasts of the two cryptocurrency assets returns for error analysis. First, the error distribution curves fitting the prediction errors of the different models show that the hybrid model has a more minor error than the single model. In addition, the latter three hybrid models using VMD decomposition techniques have a smaller range of error point fluctuations and run more smoothly. The errors of the proposed models are more distributed around zero, with the slightest corresponding standard deviations. Taylor plots show that the proposed model is closest to the actual values among all empirical results compared to the other models. The proposed model has optimal prediction accuracy and stability by the positions of the three main statistical indicators.

#### 6) BRIEF SUMMARY

Overall, the VMD-Res.-CEEMDAN-SSA-ELM multiscale hybrid model can forecast cryptocurrency returns more accurately. Compared to the benchmark model, the proposed hybrid model performs best in the four evaluation metrics shown in the one-step and five-step ahead forecasting process for two cryptocurrency asset returns. Moreover, compared with the single forecasting model, the hybrid model improves the evaluation indicators significant without considering the residual term. Specifically, compared with the common ARIMA model, the MAE of the proposed model decreases by 82.49%, the NRMSE by 84.05%, and the SMAPE by 66.02% in the context of a one-step ahead forecast of bitcoin returns, and the $R^2$ index reaches 0.9744., We found that the proposed model had a more concentrated prediction error distribution and showed excellent and stable prediction performance in comparing the prediction errors.

## IV. CONCLUSION AND FUTURE WORK

Based on the idea of ''decomposition-integration'', this paper proposed a multi-scale hybrid model containing quadratic decomposition, optimization and prediction algorithms and applies it to the study of daily returns of mainstream cryptocurrencies. The empirical analysis led to the following conclusions.

(1) Based on complex systems methodology, decomposition and integration techniques were used to decompose the cryptocurrency return series into subseries, then predict each subseries individually, and finally integrate and reconstruct the prediction results of each subseries to form the overall prediction results. This process could improve the forecasting accuracy more effectively than a single model.

(2) The VMD technique performed better when dealing with highly complex time series data such as nonstationary and nonlinear data. Adapting VMD decomposition combined with the SSA-ELM algorithm could substantially improve the prediction results compared with CEEMDAN and EMD.

(3) The combined VMD-Res.-CEEMDAN-SSA-ELM quadratic decomposition model had a significantly stronger forecasting ability than the combined single decomposition model. Considering the residual terms, the VMD and CEEMDAN quadratic decomposition techniques could be used to decompose the nonstationary, nonlinear, and highly complex financial time series effectively with clustered fluctuations into several more regular smooth subseries. The combined model had significant advantages over the single and other combined models without considering the residual terms. Information and the optimal prediction results achieved in both the 1-step-ahead and 5-step-ahead prediction studies proved the robustness of the model.

The proposed multiscale hybrid model, VMD-Res.-CEEMDAN-SSA-ELM, conducted an empirical study based on the returns of Bitcoin and Ethereum. We found that the proposed model could be used to improve the accuracy of cryptocurrency return forecasting effectively. These forecasts can help short-term investors in the cryptocurrency market to more accurately understand and grasp the market's price trends. However, the empirical evidence in this paper for cryptocurrency returns series also proves that the trends of price series can be affected by complex multidimensional factors, sharp fluctuations still characterize the data, and the performance of the model is weakened in multistep forecasting. Therefore, in future research, we plan to consider upgrading the proposed model by combining other multidimensional and complex influencing factors, and hence considering higher frequencies (e.g., hours) and longer time horizons (e.g., months or years). We look forward to this work can better grasp the financial time series characteristics and apply them to the actual portfolio strategy design to provide reference for investors and regulators.

## REFERENCES

[1] S. Nakamoto. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[2] S. Aras, "On improving GARCH volatility forecasts for bitcoin via a meta-learning approach," *Knowl.-Based Syst.*, vol. 230, Oct. 2021, Art. no. 107393.

[3] D. Shen, A. Urquhart, and P. Wang, "Does Twitter predict bitcoin?" *Econ. Lett.*, vol. 174, pp. 118–122, Jan. 2019.

[4] H. Guo, D. Zhang, S. Liu, L. Wang, and Y. Ding, "Bitcoin price forecasting: A perspective of underlying blockchain transactions," *Decis. Support Syst.*, vol. 151, Dec. 2021, Art. no. 113650.

[5] R. Chowdhury, M. A. Rahman, M. S. Rahman, and M. R. C. Mahdy, "An approach to predict and forecast the price of constituents and index of cryptocurrency using machine learning," *Phys. A, Stat. Mech. Appl.*, vol. 551, Aug. 2020, Art. no. 124569.

[6] Z. Zhang, H.-N. Dai, J. Zhou, S. K. Mondal, M. M. García, and H. Wang, "Forecasting cryptocurrency price using convolutional neural networks with weighted and attentive memory channels," *Expert Syst. Appl.*, vol. 183, Nov. 2021, Art. no. 115378.

[7] S. F. Cheng, G. De Franco, H. B. Jiang, and P. K. Lin, "Riding the blockchain mania: Public firms' speculative 8-K disclosures," *Manage. Sci.*, vol. 65, pp. 5901–5913, Dec. 2019.

[8] Y. Ma, F. Ahmad, M. Liu, and Z. Wang, "Portfolio optimization in the era of digital financialization using cryptocurrencies," *Technological Forecasting Social Change*, vol. 161, Dec. 2020, Art. no. 120265.

[9] I. Shaikh, "Policy uncertainty and bitcoin returns," *Borsa Istanbul Rev.*, vol. 20, no. 3, pp. 257–268, Sep. 2020.

[10] D. Aggarwal, "Do bitcoins follow a random walk model?" *Res. Econ.*, vol. 73, no. 1, pp. 15–22, Mar. 2019.

[11] J. Li, Y. Yuan, and F.-Y. Wang, "A novel GSP auction mechanism for ranking bitcoin transactions in blockchain mining," *Decis. Support Syst.*, vol. 124, Sep. 2019, Art. no. 113094.

[12] R. Matkovskyy and A. Jalan, "From financial markets to bitcoin markets: A fresh look at the contagion effect," *Finance Res. Lett.*, vol. 31, pp. 93–97, Dec. 2019.

[13] M. Muzammal, Q. Qu, and B. Nasrulin, "Renovating blockchain with distributed databases: An open source system," *Future Gener. Comput. Syst.*, vol. 90, pp. 105–117, Jan. 2019.

[14] N. Uras, L. Marchesi, M. Marchesi, and R. Tonelli, "Forecasting bitcoin closing price series using linear regression and neural networks models," *PeerJ Comput. Sci.*, vol. 6, p. e279, Jul. 2020.

[15] B. Kapar and J. Olmo, "Analysis of bitcoin prices using market and sentiment variables," *World Economy*, vol. 44, no. 1, pp. 45–63, Jan. 2021.

[16] Z. H. Munim, M. H. Shakil, and I. Alon, "Next-day bitcoin price forecast," *J. Risk Financial Manage.*, vol. 12, no. 2, p. 103, Jun. 2019.

[17] D. T. Nguyen and H. V. Le, "Predicting the price of bitcoin using hybrid ARIMA and machine learning," in *Proc. Int. Conf. Future Data Secur. Eng.*, Vietnam, 2019, pp. 696–704.

[18] R. Bohte and L. Rossini, "Comparing the forecasting of cryptocurrencies by Bayesian time-varying volatility models," *J. Risk Financial Manage.*, vol. 12, no. 3, p. 150, Sep. 2019.

[19] P. Jiang, Z. Liu, J. Wang, and L. Zhang, "Decomposition-selection-ensemble forecasting system for energy futures price forecasting based on multi-objective version of chaos game optimization algorithm," *Resour. Policy*, vol. 73, Oct. 2021, Art. no. 102234.

[20] J. Yoon, "Forecasting of real GDP growth using machine learning models: Gradient boosting and random forest approach," *Comput. Econ.*, vol. 57, no. 1, pp. 247–265, Jan. 2021.

[21] S. Qi, K. Jin, B. Li, and Y. Qian, "The exploration of internet finance by using neural network," *J. Comput. Appl. Math.*, vol. 369, May 2020, Art. no. 112630.

[22] O. Sohaib, W. Hussain, M. Asif, M. Ahmad, and M. Mazzara, "A PLS-SEM neural network approach for understanding cryptocurrency adoption," *IEEE Access*, vol. 8, pp. 13138–13150, 2020.

[23] L. Cocco, R. Tonelli, and M. Marchesi, "Predictions of bitcoin prices through machine learning based frameworks," *PeerJ Comput. Sci.*, vol. 7, p. e413, Mar. 2021.

[24] S. Alonso-Monsalve, A. L. Suárez-Cetrulo, A. Cervantes, and D. Quintana, "Convolution on neural networks for high-frequency trend prediction of cryptocurrency exchange rates using technical indicators," *Expert Syst. Appl.*, vol. 149, Jul. 2020, Art. no. 113250.

[25] R. Gupta, S. Tanwar, S. Tyagi, and N. Kumar, "Machine learning models for secure data analytics: A taxonomy and threat model," *Comput. Commun.*, vol. 153, pp. 406–440, Mar. 2020.

[26] M. Poongodi, A. Sharma, V. Vijayakumar, A. P. Sharma, R. Iqbal, and R. Kumar, "Prediction of the price of ethereum blockchain cryptocurrency in an industrial finance system," *Comput. Electr. Eng.*, vol. 81, Jan. 2020, Art. no. 106527.

[27] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, Jul. 2004, pp. 985–990.

[28] G.-B. Huang, D. H. Wang, and Y. Lan, "Extreme learning machines: A survey," *Int. J. Mach. Learn. Cybern.*, vol. 2, no. 2, pp. 107–122, Jun. 2011.

[29] M. Shariati, M. S. Mafipour, P. Mehrabi, Y. Zandi, D. Dehghani, A. Bahadori, A. Shariati, N. T. Trung, M. N. Salih, and S. Poi-Ngian, "Application of extreme learning machine (ELM) and genetic programming (GP) to design steel-concrete composite floor systems at elevated temperatures," *Steel Compos. Struct.*, vol. 33, pp. 319–332, Nov. 2019.

[30] I. Ahmad, M. Basheri, M. J. Iqbal, and A. Raheem, "Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection," *IEEE Access*, vol. 6, pp. 33789–33795, 2018.

[31] S. Sun, S. Wang, and Y. Wei, "A new ensemble deep learning approach for exchange rates forecasting and trading," *Adv. Eng. Informat.*, vol. 46, Oct. 2020, Art. no. 101160.

[32] D. Aggarwal, S. Chandrasekaran, and B. Annamalai, "A complete empirical ensemble mode decomposition and support vector machine-based approach to predict bitcoin prices," *J. Behav. Experim. Finance*, vol. 27, Sep. 2020, Art. no. 100335.

[33] H. Lin, Q. Sun, and S.-Q. Chen, "Reducing exchange rate risks in international trade: A hybrid forecasting approach of CEEMDAN and multilayer LSTM," *Sustainability*, vol. 12, no. 6, p. 2451, Mar. 2020.

[34] S. Sun, S. Wang, Y. Wei, and G. Zhang, "A clustering-based nonlinear ensemble approach for exchange rates forecasting," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 6, pp. 2284–2292, Jun. 2020.

[35] W. Liu, W. D. Liu, and J. Gu, "Forecasting oil production using ensemble empirical model decomposition based long short-term memory neural network," *J. Petroleum Sci. Eng.*, vol. 189, Jun. 2020, Art. no. 107013.

[36] Y. Li, S. Wang, Y. Wei, and Q. Zhu, "A new hybrid VMD-ICSS-BiGRU approach for gold futures price forecasting and algorithmic trading," *IEEE Trans. Computat. Social Syst.*, vol. 8, no. 6, pp. 1357–1368, Dec. 2021.

[37] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London Ser. A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

[38] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise-assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, p. 41, Jan. 2009.

[39] M. E. Torres, M. A. Colominas, G. Schlotthauer, and P. Flandrin, "A complete ensemble empirical mode decomposition with adaptive noise," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2011, pp. 4144–4147.
[40] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.
[41] W. Sun and J. Zhang, "Carbon price prediction based on ensemble empirical mode decomposition and extreme learning machine optimized by improved bat algorithm considering energy price factors," *Energies*, vol. 13, no. 13, p. 3471, Jul. 2020.
[42] J. Cao, Z. Li, and J. Li, "Financial time series forecasting model based on CEEMDAN and LSTM," *Phys. A, Stat. Mech. Appl.*, vol. 519, pp. 127–139, Apr. 2019.
[43] J. Zhu, P. Wu, H. Chen, J. Liu, and L. Zhou, "Carbon price forecasting with variational mode decomposition and optimal combined model," *Phys. A, Stat. Mech. Appl.*, vol. 519, pp. 140–158, Apr. 2019.
[44] M. Jiang, L. Jia, Z. Chen, and W. Chen, "The two-stage machine learning ensemble models for stock price prediction by combining mode decomposition, extreme learning machine and improved harmony search algorithm," *Ann. Oper. Res.*, vol. 309, no. 2, pp. 553–585, Feb. 2022.
[45] Q. Qin, K. Xie, H. He, L. Li, X. Chu, Y.-M. Wei, and T. Wu, "An effective and robust decomposition-ensemble energy price forecasting paradigm with local linear prediction," *Energy Econ.*, vol. 83, pp. 402–414, Sep. 2019.
[46] W. Sun and C. Xu, "Carbon price prediction based on modified wavelet least square support vector machine," *Sci. Total Environ.*, vol. 754, Feb. 2021, Art. no. 142052.
[47] G. Xing, S. Sun, and J. Guo, "A new decomposition ensemble learning approach with intelligent optimization for PM2.5 concentration forecasting," *Discrete Dyn. Nature Soc.*, vol. 2020, pp. 1–11, Mar. 2020.
[48] B. Wang and J. Wang, "Energy futures and spots prices forecasting by hybrid SW-GRU with EMD and error evaluation," *Energy Econ.*, vol. 90, Aug. 2020, Art. no. 104827.
[49] X. Sun, H. Zhang, J. Wang, C. Shi, D. Hua, and J. Li, "Ensemble streamflow forecasting based on variational mode decomposition and long short term memory," *Sci. Rep.*, vol. 12, no. 1, pp. 1–19, Jan. 2022.
[50] Z. Xu, J. Zhou, L. Mo, B. Jia, Y. Yang, W. Fang, and Z. Qin, "A novel runoff forecasting model based on the decomposition-integration-prediction framework," *Water*, vol. 13, no. 23, p. 3390, Dec. 2021.
[51] E. Gulay and O. Duru, "Hybrid modeling in the predictive analytics of energy systems and prices," *Appl. Energy*, vol. 268, Jun. 2020, Art. no. 114985.
[52] H. Huang, J. Chen, X. Huo, Y. Qiao, and L. Ma, "Effect of multi-scale decomposition on performance of neural networks in short-term traffic flow prediction," *IEEE Access*, vol. 9, pp. 50994–51004, 2021.
[53] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: Sparrow search algorithm," *Syst. Sci. Control Eng.*, vol. 8, no. 1, pp. 22–34, Jan. 2020.
[54] K. E. Taylor, "Summarizing multiple aspects of model performance in a single diagram," *J. Geophys. Res., Atmos.*, vol. 106, no. D7, pp. 7183–7192, Apr. 2001.

**XIAOXU DU** is currently pursuing the Ph.D. degree with the School of Economics and Management, Fuzhou University, China. His main research interests include financial market complexity, financial risk management, big data mining, model selection, and price forecasting.

**ZHENPENG TANG** received the Ph.D. degree in management science and engineering from the Wuhan University of Technology, China, in 2003. He is currently a Professor with the School of Economics and Management, Fujian Agriculture and Forestry University, China. He is the author of four books and over 70 journal articles. His research interests include financial risk management, big data analysis, artificial intelligence, and predictive modeling.

**JUNCHUAN WU** received the Ph.D. degree in financial engineering from Fuzhou University, China, in 2020. He is currently a Lecturer with the School of Economics and Management, Nanchang University, China. His research interests include financial risk management, big data analysis, data mining, and predictive modeling.

**KAIJIE CHEN** is currently pursuing the Ph.D. degree with the School of Economics and Management, Fuzhou University, China. His main research interests include financial risk management, big data mining, model selection, and price forecasting.

**YI CAI** is currently pursuing the master's degree with the School of Economics and Management, Fuzhou University, China. His main research interests include financial risk management, big data mining, model selection, and price forecasting.