Applied Intelligence (2020) 50:4296–4309
https://doi.org/10.1007/s10489-020-01814-0

A hybrid stock price index forecasting model based on variational mode decomposition and LSTM network

Hongli Niu[1] · **Kunliang Xu**[1] · **Weiqing Wang**[1]

Published online: 17 July 2020

Abstract

Changes in the composite stock price index are a barometer of social and economic development. To improve the accuracy of stock price index prediction, this paper introduces a new hybrid model, VMD-LSTM, that combines variational mode decomposition (VMD) and a long short-term memory (LSTM) network. The proposed model is based on decomposition-and-ensemble framework. VMD is a data-processing technique through which the original complex series can be decomposed into a limited number of subseries with relatively simple modes of fluctuations. It can effectively overcome the shortcomings of mode mixing that sometimes exist in the empirical mode decomposition (EMD) method. LSTM is an improved version of recurrent neural networks (RNNs) that introduces a "gate" mechanism, and can effectively filter out the critical previous information, making it suitable for the financial time series forecasting. The capability of VMD-LSTM in stock price index forecasting is verified comprehensively by comparing with some single models and the EMD-based and other VMD-based hybrid models. Evaluated by level and directional prediction criteria, as well as a newly introduced statistic called the complexity-invariant distance (CID), the VMD-LSTM model shows an outstanding performance in stock price index forecasting. The hybrid models perform significantly better than the single models, and the forecasting accuracy of the VMD-based models is generally higher than that of the EMD-based models.

Keywords Decomposition-and-ensemble · VMD · LSTM · Stock price forecasting · Hybrid model

1 Introduction

Compared with the general market, the financial market is a complex and changeable system accompanied by high volatility and high noise. The stock price index, often featured as a non-linear and non-stationary time series, reflects the overall price level and changes in the financial market. According to the index, investors can judge the macro trend of the stock market and then adjust their investment strategies and portfolios in time. Similarly, multiple social subjects, such as the economists, entrepreneurs and government workers, consider the stock price index as an important reference when measuring and evaluating social development. Therefore, as one of the most typical time series data, the accurate prediction of stock price indices is gradually becoming a hot research area [1, 2]. In recent years, artificial neural networks (ANNs) have become an auxiliary tool for the prediction of financial time series because of their better performance than that of traditional econometric models [3–6]. The ANN, which is composed of many interconnected neurons, is an emulation of the biological system of the human brain. Due to its data-driven characteristics, taking advantage of the historical and current time data to predict the future trends is a crucial application in financial time series. For instance, Adhikari and Agrawal [7] established a model to predict exchange rates and stock price indices by integrating a random walk and an ANN. Kuremoto et al. [8] combined deep belief networks and Boltzmann machines to predict time series. Niu and Wang [9] established a random data-time effective RBF neural network and forecasted crude oil prices and stock price indices. Wang and Wang [10] applied the EMD algorithm and stochastic time neural networks (STNN) to forecast stock price indices, including the NYSE, DAX, FTSE and HSI. Das et al. [11] proposed a novel hybrid SVM-TLBO model to predict the daily closing prices of the COMDEX commodity futures index traded in the

✉ Kunliang Xu
kunliangxu@yeah.net

Hongli Niu
niuhongli@ustb.edu.cn

[1] Donlinks School of Economics and Management, University of Science and Technology Beijing, Beijing 100083, China

Multi Commodity Exchange of India Limited. This model consists of a support vector machine (SVM) and a teaching-learning-based optimization (TLBO) and its better performance has been verified.

Although the above neural networks have great predictive performance, their accuracy is still unsatisfactory when the time series are dynamic and non-linear. With the connections between hidden layer units established, the dependencies of data at different points in time can be explained by recurrent neural networks (RNNs) [12, 13]. The before–after associated data structure makes RNNs especially suitable for the prediction of time series data [14, 15], and the literatures have witnessed their wide application in the forecasting of financial time series [16–18]. However, the historical information kept in an RNN will increase as the time span increases, which will lead to gradient disappearance or gradient explosion, distorting the predictive results of the RNN. The long short-term memory (LSTM) network introduces the "memory cell" structure to replace the hidden layer units of the RNN [19]. Every memory cell consists of a special "gate" mechanism, which makes LSTM retain crucial historical information and filter out that which is unimportant. This structure ensures that the previous inputs can be preserved in the LSTM internal state and achieve internal communication. Based on the superior mechanism, LSTM is widely used in natural language recognition [20–22], time series prediction [23–25], and especially for stock price time series forecasting [26–29], as well as other fields [30–32]. For example, Ding and Qin [26] proposed an associated deep recurrent neural network model with multiple inputs and outputs based on LSTM (associated net) to forecast the multiple prices, including the open lowest and highest of a stock simultaneously.

Considering the deficiencies of single models applied in the prediction of complex data, hybrid models that integrate two or more individual models have been proposed to solve problems such as low accuracy and lagging predictions. In this way, various preprocessing methods are introduced to reduce the noise or further capture different patterns from the original series, and then the processed data are input into the forecasting model to obtain the final result. In addition, there are hybrid models that implement prediction tasks by combining features learned from different representations of the same series [33, 34]. The proposed new approaches are gradually becoming mainstream [35–37]. The decomposition-and-ensemble framework is one of the most common hybrid methods following the "divide-and-conquer" principle [38–41]. The main principle of this framework, which is utilized to predict complex time series, is to decompose the original series data into a limited number of independent subseries, which makes it easier for models to learn the features of fluctuations in each subseries adaptively. In addition, the goal of the ensemble is to aggregate the predictive results of these subseries and then conclude the forecasting consequence of the original series [42, 43]. Empirical mode decomposition (EMD) is one of the most common approaches due to its advantages in decomposing non-linear and non-stationary time series [44]. For example, Awajan et al. [45] used the EMD-HW bagging method to accurately forecast the stock market time series of six countries. Li [46] introduced a new hybrid model combining EMD and an RBF neural network to predict the stock index futures and verified its superior performance. Yang and Lin [47] proposed an integrated forecasting model combining EMD, ARIMA and SVR, through which the original series can be decomposed into linear and non-linear parts and predicted, respectively. Zhang et al. [48] established a stochastic Ising model for each IMF obtained by EMD to forecast the stock market indices. Huang et al. [49] combined EMD and a probabilistic neural network (PNN) to explore the mutation characteristics of plasma electrical signals. Tang et al. [50] introduced complementary ensemble EMD (CEEMD) and the extended extreme learning machine (EELM) as decomposition and forecasting tools, respectively, to predict the crude oil prices.

However, EMD-based methods have generally been proved to have some shortcomings, such as boundary effects, mode overlap, sensitivity to noise and a lack of exact mathematical foundations, which may negatively impact the precision of decomposition and then lead to distorted results. Variational mode decomposition (VMD) is a novel multiresolution technique that originated in signal processing. Significantly different from EMD, VMD is a completely non-recursive model that can decompose the original signal or series data into multiple components with a specific bandwidth in the spectral domain. It has been proved that VMD performs better than models of the same kind in noise robustness and component decomposition accurately [51]. Based on its excellent ability, it has been successfully applied to fault detection [52], biomedical signal analysis [53] and time series forecasting [54–56]. Particularly, the higher prediction accuracy of VMD-based neural networks than that of EMD-based and single neural networks has been verified in stock price [57, 58], non-ferrous metal price [59] and wind speed forecasting [60]. Considering the excellent performance of VMD in the decomposition of the non-linear data series and the advantageous characteristic of LSTM in forecasting, in the present work, a hybrid neural network model based on VMD and LSTM is proposed for forecasting stock price indices from the world financial market. In the VMD-LSTM model, the VMD algorithm is first used to decompose the original stock price series into different IMFs with distinct time frequencies, and then each IMF sequence is trained by the LSTM model, and the corresponding predicted IMF is obtained. Finally, the ensemble forecasting result of the original stock prices is produced by aggregating all the predictive IMF components. By comparison with other prediction methods, the proposed model can indeed improve the overall prediction results.

In brief, the main contribution of this paper includes two aspects. One is that we propose a hybrid VMD-LSTM model that presents excellent applicability in forecasting stock indices with non-linear and non-stationary features. To verify the superiority of the proposed model, we adopt the BPNN, ELM, CNN, LSTM, and their EMD-based and VMD-based hybrid models for comparison. The results indicate that the hybrid models perform significantly better than the single models, and among the hybrid models, the forecasting accuracy of the VMD-based models is generally higher than that of the EMD-based models. The other aspect is that in the performance evaluation of forecasting, a novel statistic called the complexity-invariant distance (CID) is adopted to evaluate the performance of all kinds of models, through which the differences between the predicted values of and the original series can be more intuitively presented.

The rest of the paper is organized as follows. Section 2 introduces the VMD and LSTM methodologies and provides the basic framework of the proposed hybrid model. Section 3 demonstrates in detail the experimental forecasting results of 4 stock price indices obtained by using the proposed model and then compares them with the results of the other models considered. Finally, Section 4 concludes the work.

2 Methodologies

2.1 Variational mode decomposition

Variational mode decomposition (VMD), developed by Dragomiretskiy and Zosso [61], is a novel adaptive decomposition method that adopts a non-recursive signal decomposition optimization algorithm. It assumes that the center frequency bandwidth of each subseries is limited. It requires minimizing the sum of the estimated bandwidth of each subseries while decomposing the original series. This method effectively improves the mode mixing, endpoint effects and other shortcomings of EMD. The subseries or IMFs obtained by VMD have the advantages of high accuracy, strong noise robustness, and fast convergence. Therefore, VMD is more suitable for addressing complex data such as financial time series. It has been widely used to implement the decomposition of multi-frequency signals adaptively, in which the frequency band can be estimated at the same time to effectively balance the errors between different frequency bands.

Suppose the original signal $f(t)$, it can be decomposed into K intrinsic mode functions (IMFs), u_K, and that each mode u_k needs to be close to a center frequency ω_k. The bandwidth of a mode can be estimated by the following steps. First, for each mode u_k, the Hilbert transform is adopted to calculate the correlation analysis signal and then to obtain a unilateral frequency spectrum. Second, for each mode u_k, the frequency spectrum of the mode is transferred to the baseband by mixing with an exponential tuned to the respective estimated center pulsation. Finally, the H^1 Gaussian smoothness of the demodulated signal is adopted to estimate the bandwidth. The constraints of the variational problem can be formulated as follows:

$$\begin{cases} min_{(\omega_k, u_k)} \left\{ \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta_t + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s.t. \quad \sum_{k=1}^{K} u_k = f(t) \end{cases} \tag{1}$$

In the above formula, $\{u_k\} = \{u_1, u_2, \ldots u_K\}$ and $\{\omega_k\} = \{\omega_1, \omega_2, \ldots \omega_K\}$ represent the sets of the kth sub-mode and its corresponding central frequency, respectively. ∂_t denotes the differential processing of t, $\|\cdot\|$ indicates the norm processing, δ_t is the Dirac function, and * is the convolution symbol. K denotes the total number of IMFs; an IMF with a high-order k denotes a low-frequency component.

To solve the optimization problem of constrained variational decomposition, an augmented Lagrangian function L is introduced:

$$L(\{u_k\}, \{\omega_k\}, \lambda) = \alpha \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta_t + \frac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_{k=1}^{K} u_k(t) \right\|_2^2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^{K} u_k(t) \right\rangle \tag{2}$$

where α is the penalty parameter and $\lambda(t)$ is the Lagrangian multiplier. To obtain the saddle point of the above formula, which is also the solution to the original constraint conditions, VMD adopts the alternating direction method of multipliers (ADMM) [62]. The calculation process of this method can be found in Fig. 1.

2.2 Long short-term memory network

The long short-term memory (LSTM) network [19] is an extension of the recurrent neural network (RNN) [12–15]. The structure of the RNN is similar to a multi-layer perception network, and many internal connections are established between nodes in the hidden layer so that information can be transmitted forward or backward conveniently. In the RNN, the state of the history node can be regarded as the input, and the final output can be considered as the feedback of the input layer and the historical state of all hidden layers. However, its performance in the experiment is somewhat worse than expected, which is mainly due to the following two reasons. On the one hand, it is difficult for an RNN to determine the best window length for historical data observation, and thus, the fluctuation characteristics of the data may be insufficiently

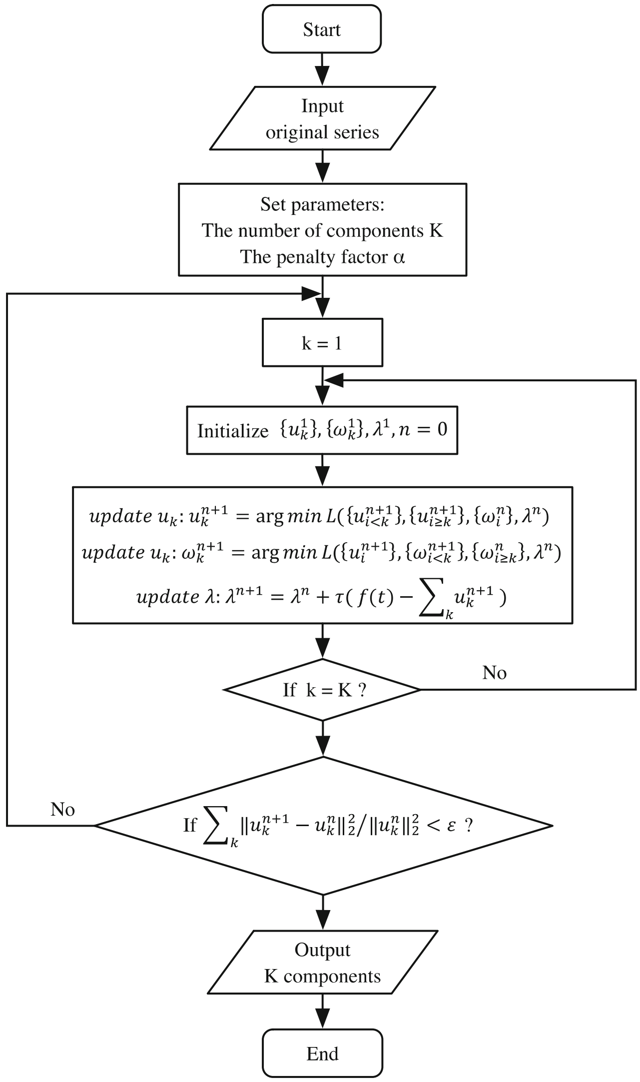

Fig. 1 The structure of the VMD algorithm

extracted. On the other hand, the gradient of an RNN may show an exponential increase or decay when using the method of gradient descent to address long-term data, which will lead to gradient explosion or disappearance. The reason for this phenomenon is that the long connection structure of an RNN keeps much non-critical information without filtering. Hence, the traditional RNN is not an ideal model for processing long-term data.

LSTM creatively introduces the "gate" mechanism to improve the RNN; it replaces the hidden layer nodes of the RNN with special memory cells (blocks) that implement the filtering and processing of historical states and information. The basic structure of the memory cell can be seen in the upper right plot of Fig. 2. Each memory cell contains three gates: input gate i_t, forget gate f_t, and output gate o_t. The input gate determines how much of the latest information will be added to the cell at the current time point. The forget gate determines how much information should be discarded from the previous cell state and how much should be kept, which prevents the internal cell values from growing without boundaries. The output gate is designed to filter the new state and output the filtered information. The long-term connection structure and selective information storage mechanism between the input layer and the output layer ensure that errors will be transmitted in the network as a constant; thus, the problems of gradient disappearance and explosion can be effectively resolved. The basic steps of LSTM are as follows:

First, the input gate i_t falters and extracts new information from the input x_t at the current state (time t) and creates a candidate value $\widetilde{c}_t$ for updating the state.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\widetilde{c}_t = tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{4}$$

Next, the forget gate f_t filters and keeps the historical information that can indicate long-term trends and discards the non-critical information.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{5}$$

By removing part of the information from the old cell and adding the filtered candidate value, the old cell state c_{t-1} is updated to the new cell state c_t.

$$c_t = f_t * c_{t-1} + i_t * \widetilde{c}_t \tag{6}$$

Here, * indicates the dot multiplication between matrices.

Finally, the output gate o_t filters the updated state c_t, and the final output is calculated based on the updated state and the output gate state.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{7}$$

$$h_t = o_t * tanh(c_t) \tag{8}$$

Here, h_t is the hidden layer state, namely, the activation of the memory cell. W_c, W_f, and W_o represent the appropriate weight matrices, b_i, b_c, b_f, and $b_o$denote the corresponding bias vectors. $\sigma(\cdot)$ and tanh($\cdot$) are the sigmoid function and hyperbolic tangent function, respectively.

2.3 VMD-LSTM

In view of the advantages of VMD and LSTM, we construct a hybrid model named VMD-LSTM by combining the two

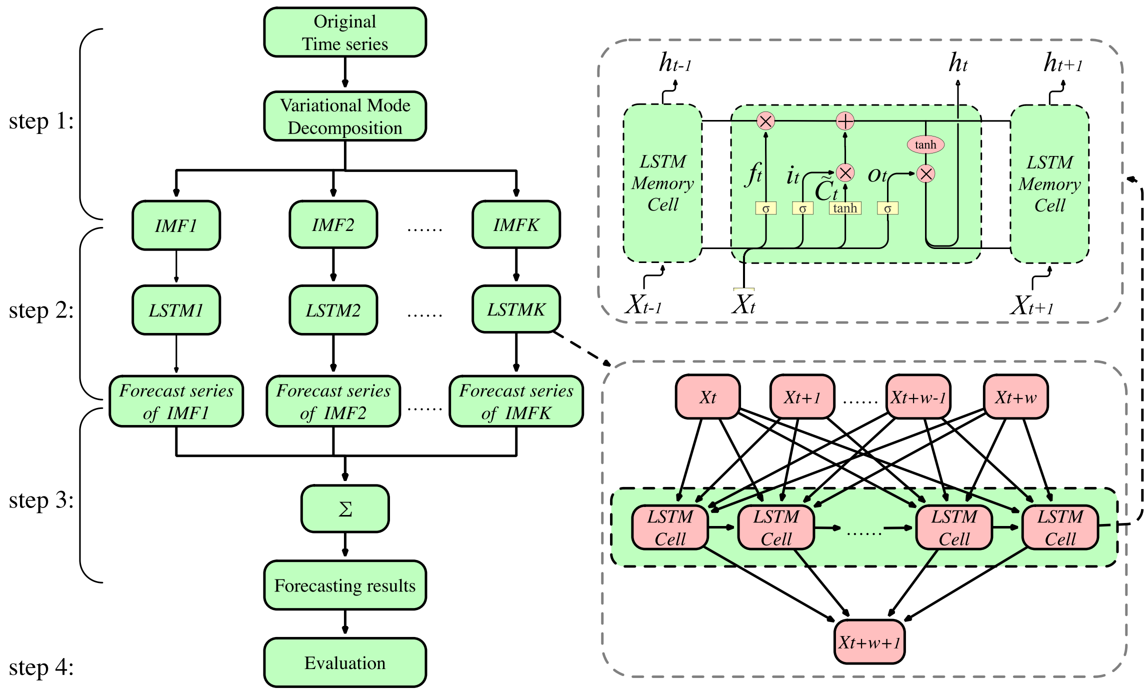

Fig. 2 Flow chart of the VMD-LSTM model

techniques. The flow chart in Fig. 2 depicts its implementation process, in which the VMD-LSTM operation is carried out as follows:

Step 1: The VMD technique is utilized to decompose the original stock price time series ( ), = 1, 2, ⋯ , into K mutually independent subseries, denoted by IMF1, IMF2, ⋯ IMFK, which represent the different local oscillations of prices ranging from high frequency to low frequency, that is, the initial series is reconstructed in terms of the IMFs as:

$$x(t) = \sum_{k=1}^{K} IMF_k(t)$$

The execution of the VMD technique can be based on the algorithm flow of VMD shown in Fig. 1. Each IMF series then becomes the new forecasting sample of the LSTM model.

Step 2: Each component IMF obtained by VMD is split into training and test datasets at a fixed ratio, and the input and output sets are split according to the step size. In this work, the first 80% of timepoints of each component series are used as the training set, and the remaining 20% are used as the test set. The two sets are completely separated so that the generalization capability is ensured. The LSTM network is employed to train and build the forecasting model based on the training dataset, in which parameters including the learning rate, batch size, number of iterations and number of hidden layer units that are critical to ensure the predictive accuracy of the model are required to be pre-set. The step size is set to 4, meaning that in both the training set and test set, the prediction value can be obtained based on the input of the previous 4th-order datapoints. According to the established model, the forecasting result of each IMF subseries for the stock prices is obtained.

Step 3: The final predictive result of the original stock price time series is obtained by summarizing the separate predicted outputs $f_k(t)$, $k = 1, 2, \ldots, K$ of each IMF, that is, $\sum_{k=1}^{K} f_k(t)$. The prediction errors of each IMF with different dimensions result in different effects for the final predictive result; among them, the low-frequency IMFs usually consist of large values that represent the macro trend of the series. From the perspective of level forecasting accuracy, the prediction errors of the low-frequency IMFs will lead to a more significant impact on the final predictive result than will the high-frequency IMFs mainly composed of relatively small values.

Step 4: Multiple performance measures, i.e., the MAE, RMSE, MAPE, and $_{stat}$, and a novel complexity-invariant distance (CID) from information theory are introduced to evaluate the prediction capacity of VMD-LSTM from different perspectives. In addition, other models, including the VMD-ELM, VMD-CNN, VMD-BPNN and EMD-based hybrid models, as well as single models, are considered for comparison.

3 Experimental results and comparative analysis

3.1 Data description and pre-processing

In this work, four representative stock price indices from the world financial market are selected for the case study: the daily closing prices of the Hong Kong Hang Seng Index (HSI), S&P500 Index (SPX), London FTSE Index (FTSE) and Nasdaq Index (IXIC). The time period is from 04 January, 2010, to 04 December, 2019, excluding public holidays. To conduct the experiments, the first 80% of the samples from 2010 to 2017 are used to train the model, and the remaining 20% are used for testing to examine the effectiveness of the proposed model through the performance evaluation measures. Table 1 provides a detailed breakdown of the four selected stock price indices. To reduce the impact of noise and facilitate optimization of the solving process, each subseries IMF obtained by VMD is normalized to the range of [0, 1] by the following maximum and minimum standardized method:

$$x(t)' = \frac{x(t) - \min x(t)}{\max x(t) - \min x(t)} \tag{9}$$

and then, to obtain the real predicted value and compare it with the actual set intuitively, the normalized output $x(t)$' can be reverted to $x(t)$ as follows:

$$x(t) = x(t)'(\max x(t) - \min x(t)) + \min x(t) \tag{10}$$

3.2 Performance evaluation criteria

Usually multiple performance measures are employed to assess the prediction quality of a developed model. In this work, the commonly used mean absolute error (MAE), root mean square error (RMSE) and mean absolute percentage error (MAPE) are adopted to evaluate the accuracy of level forecasting. They are respectively calculated as follows:

$$MAE = \frac{1}{N}\sum_{t=1}^{N}\left|x_t - \widehat{x}_t\right| \tag{11}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(x_t - \widehat{x}_t\right)^2} \tag{12}$$

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{x_t - \widehat{x}_t}{x_t}\right| \tag{13}$$

where x_t denotes the actual value, $\widehat{x}_t$ denotes the predictive value, and N is the number of sample points of the predictive results; the same applies hereinafter. The MAE is used to measure the average absolute error between the predicted value and the actual value, the RMSE is used to measure the deviation between the predicted value and the actual value, which is more sensitive to outliers, and the MAPE is used to measure the average relative errors between the predicted value and the actual value. Generally, the smaller the MAE, RMSE, and MAPE values are, the smaller the difference between the predicted value and the actual value, that is, the higher the prediction accuracy of the model. The value of the MAPE is the standard if the value of the error class evaluation index is different and difficult to define based on the document [63].

Considering the importance of correctly predicting the direction in stock price index predictions from an economic point of view, the directional statistic D_{stat}, which provides the correctness of predicted direction of the time series in terms of the percentage, is employed to measure this capability, which can be calculated as follows:

$$D_{stat} = \frac{1}{N}\sum_{t=1}^{N} d_t \tag{14}$$

$$d_t = \begin{cases} 1, \mathit{if}\ \left(\widehat{x}_t - \widehat{x}_{t-1}\right)(x_t - x_{t-1}) \geq 0, t > 1 \\ 0, \text{otherwise} \end{cases} \tag{15}$$

D_{stat} is the probability of changing in the same direction between the predicted sets and actual sets. A higher D_{stat} corresponds to a better performance of the model.

In addition, a fresh evaluation statistic called the complexity-invariant distance (CID), which was proposed by Batista [64] and originally used to measure the complexity difference between two time series by using information theory, is introduced in this work. The CID is adopted to calculate

Table. 1 The datasets of the four selected stock price indices

Index	Dataset	Total number	Training set	Training number	Test set	Test number
HSI	2010/01/04 ~ 2019/12/04	2443	2010/01/04 ~ 2017/12/06	1954	2017/12/07 ~ 2019/12/04	489
SPX	2010/01/04 ~ 2019/12/04	2498	2010/01/04 ~ 2017/12/07	1998	2017/12/08 ~ 2019/12/04	500
FTSE	2010/01/04 ~ 2019/12/04	2509	2010/01/04 ~ 2017/12/08	2007	2017/12/09 ~ 2019/12/04	502
IXIC	2010/01/04 ~ 2019/12/04	2498	2010/01/04 ~ 2017/12/07	1998	2017/12/08 ~ 2019/12/04	500

the total absolute error between the predicted value and the actual value, in which the original error is adjusted using the information differences of the complexity between the predicted and actual values. For two time series $X = (x_1, x_2 \ldots x_N,)$ and $Y = (y_1, y_2 \ldots y_N,)$ of length N, the statistic $CID(x, y)$ can be calculated as follows:

$$CID(X, Y) = ED(X, Y) \times CF(X, Y) \tag{16}$$

$$ED(X, Y) = \sqrt{\sum_{t=1}^{N} (x_t - y_t)^2} \tag{17}$$

$$CF(X, Y) = \frac{max\{CE(X), CE(Y)\}}{min\{CE(X), CE(Y)\}} \tag{18}$$

$$CE(X) = \sqrt{\sum_{t=1}^{N} (x_{t+1} - x_t)^2} \tag{19}$$

Similar to the common indicators, a smaller CID means less difference between the two time series. It can effectively overcome the subjective similarities directly observed by the human eye.

3.3 Empirical experiments and results

In this section, four experiments involving four stock price indices are conducted. In each experiment, to validate the superior predictive performance of the proposed VMD-LSTM model in financial time series forecasting, eleven other models are employed for comparison, including the single models BPNN [58], ELM [55], CNN [33],and LSTM [19], and the hybrid models EMD-BPNN [4], EMD-ELM [35], EMD-CNN [36], EMD-LSTM [42], VMD-BPNN [58], VMD-ELM [55], and VMD-CNN [52]. In this comparison, the parameters in each model are determined based on the results of repeated experiments. For all utilized forecasting models, the number of input units is set to 4, and that of the output is set to 1. Considering that these models have different mechanisms when implementing the prediction, the hidden layer architectures and other parameters are determined separately. In LSTM, the number of hidden units is set to 50, the number of epochs is 400 and the batch size is 64. In the ELM, it is necessary only to set the number of hidden units to 10. In the one-dimensional convolutional layer of the CNN, the number of filters is set to 10, the kernel size is 1, the pool size in the max-pooling layer is set to 1, the number of epochs is 400 and the batch size is 64. In the BPNN, the number of hidden units is set to 10, the number of epochs is 400, the goal is 0.001 and the learning rate is 0.001. For the hybrid models, EMD can adaptively decompose the original series data without any pre-set parameters, while in VMD, the moderate bandwidth constraint is set to 2000 by default, the noise-tolerance is set to 0, and the tolerance of the convergence criterion is set to 1e-7. In the implementation of the VMD-LSTM model, the raw stock price series is first decomposed into several modes of the IMFs by executing the VMD algorithm, which enables diminishing the non-stationary and non-linear features of the stock price data, making it easy to carry out predictions. For comparison, the number of decomposed components K by VMD is set to be the same as that obtained by the EMD technique, as performed by many scholars [57, 59, 60], which determines the number of components by continuously sifting until there is no extremal point in the newly obtained kth component, named the residue [44]. Figure 3 presents different decomposed subseries of the HSI as an example by EMD and VMD, where the IMFs are sorted from high frequency to low frequency, with the IMFs in the upper order representing the high-frequency components, which also means that it is relatively difficult to build a training model for it and make accurate predictions, but this part is a critical component that measures the fluctuations of the prediction, especially in predicting the direction of fluctuations accurately. Each mode subseries represents a kind of oscillatory factor embedded in the price time series. Compared with EMD, the amplitudes of components decomposed by VMD are more fixed within a smaller range, especially when there are fewer obvious abnormal values or outliers in the relatively high-frequency components. In addition, the high-frequency components of EMD are concentrated in the first few, which easily leads to mode mixing. However, each component obtained by VMD has a relatively independent frequency distribution; thus, each can be predicted accurately. The daily closing prices of HSI, SPX, FTSE and IXIC are decomposed into 9, 9, 9 and 8 modes, respectively, through the VMD algorithm.

Afterwards, the normalized subseries IMFs are trained and predicted by the LSTM model. The number of input samples is set to 4 and the number of outputs is set to 1, that is, the 4th order historical data are used to predict the data of the next period. Figure 4 shows the comparisons between the predicted and actual values of the IMFs for HSI through the LSTM neural network as an example. It is shown visually that LSTM can make accurate predictions of the IMFs under different fluctuation modes. This performance is especially notable in the high-frequency mode decomposed by VMD, but it is slightly inferior in EMD. The final predictive results of HSI can be computed by summing the predicted value of all IMFs linearly.

Figure 5 demonstrates the VMD-LSTM forecasting results of the four selected stock indices in the test set in comparison with those of the other models. To facilitate observation, only

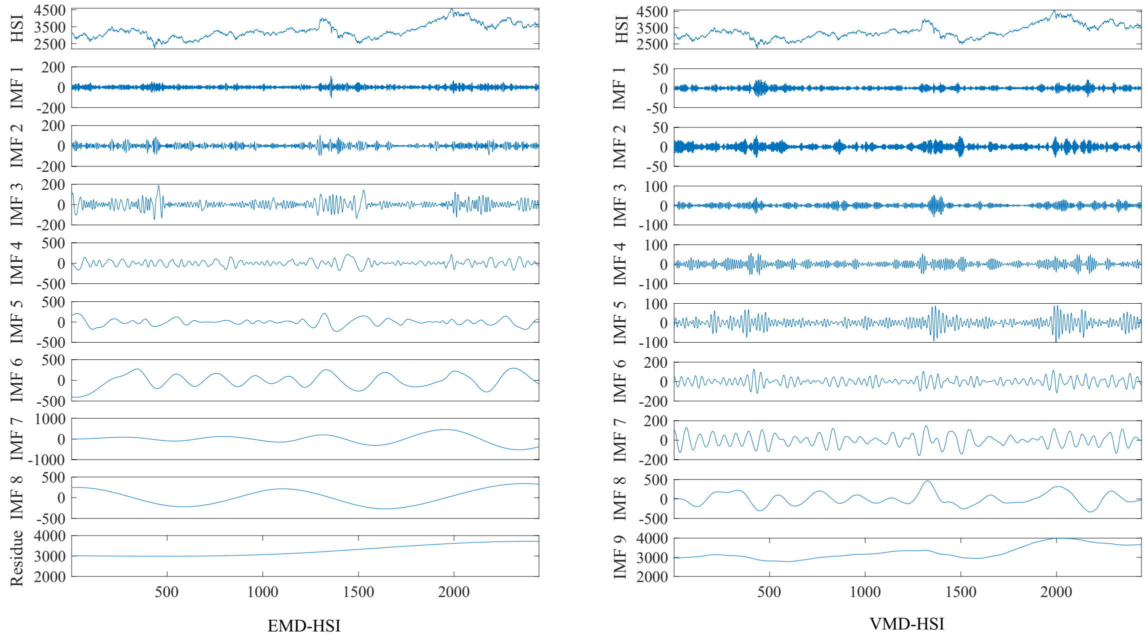

Fig. 3 The decomposition results obtained by the EMD and VMD techniques for HSI

the prediction curves of the hybrid models are shown. The predicted value of the VMD-LSTM model is generally the closest to the actual value, implying the best prediction accuracy in this comparison. A better performance for predicting large fluctuations of stock prices is further seen in the inset plot, where a certain volatile part of the datasets is magnified. Since the price series of stock indices fluctuate widely, to discriminate the performances of the different models, the predicted errors and their corresponding boxplots are displayed in Fig. 6. It can be found that the forecasting errors of VMD-LSTM are evenly distributed around zero and have a relatively smaller fluctuation range than that of the other five models, which also confirms the better predictive performance and adaptability of VMD-LSTM to different datasets.

Tables 2,3,4 and 5 quantitatively display a detailed comparison of the MAE, MAPE, RMSE, D_{stat} and CID criteria among different single models (BPNN, ELM, CNN, and LSTM) and their hybrid forms (EMD-ANNs and VMD-

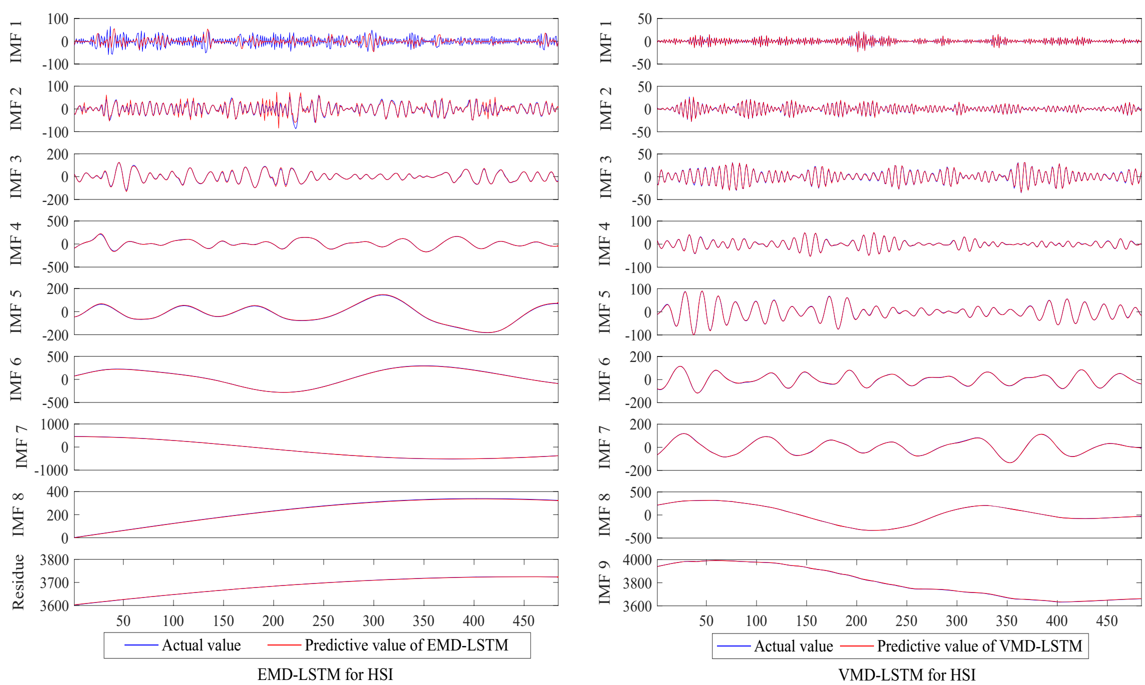

Fig. 4 The LSTM prediction results of different IMFs for HSI in the test set

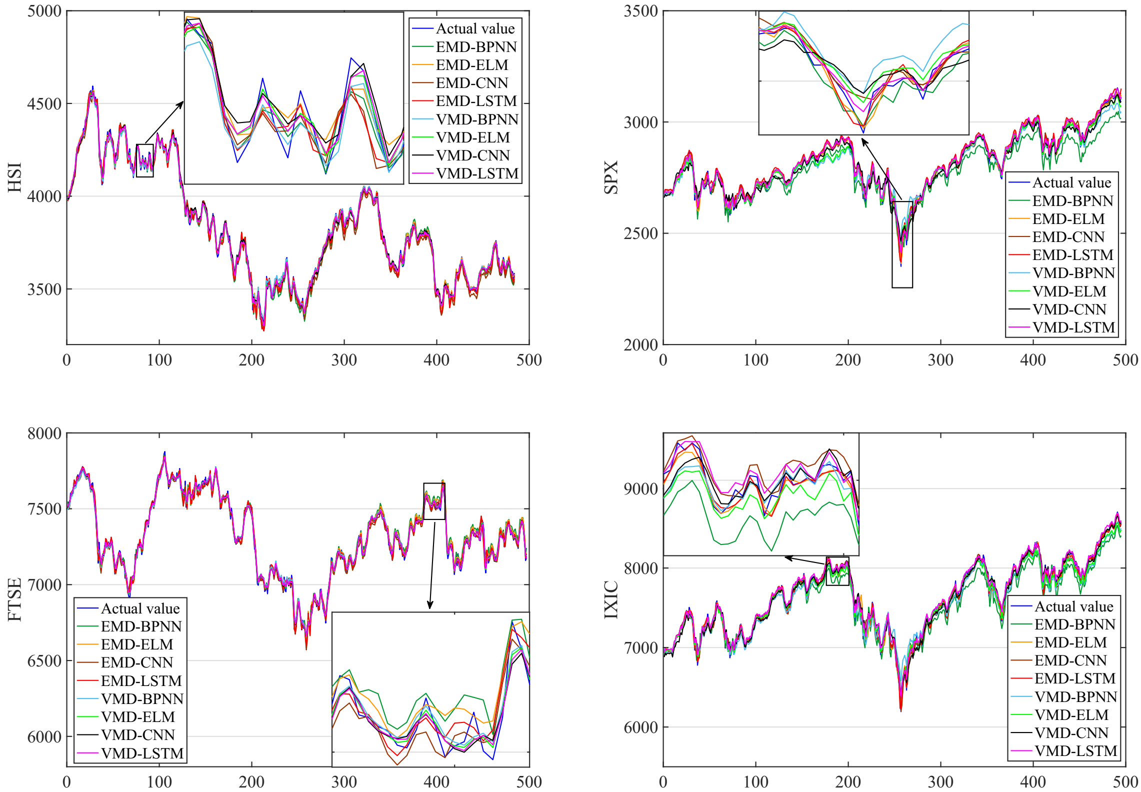

Fig. 5 The predicted results of different models for the selected stock price indices

ANNs) for the four selected stock price indices. In addition, Figs. 7 and 8 display the D_{stat} and CID values, respectively, where the averages are indicated by the blue dotted lines. The following can be concluded from these tables and figures:

(1) For the four stock indices, the predictive accuracy of the hybrid models based on "decomposition-and-ensemble" is on the whole significantly higher than that of the single models according to the various performance evaluation criteria. The level forecasting measures (MAE, MAPE and RMSE) for the hybrid models are obviously smaller than those for the single models, except for the EMD-BPNN model, which exhibits the worst predictive performance compared with the other models

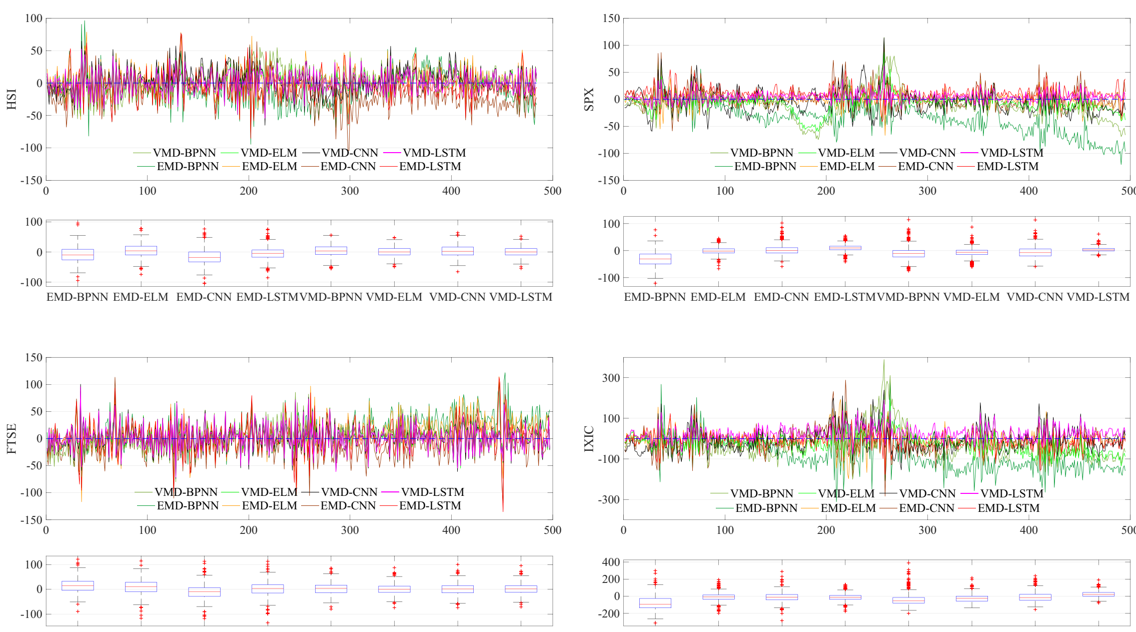

Fig. 6 The forecasting errors of different models for the selected stock price indices

Table 2 Predictive performance evaluation of different models for HSI

Model	MAE	MAPE (%)	RMSE	D_{stat} (%)	CID
VMD-LSTM	12.365	0.324	15.744	**79.089**	522.188
VMD-ELM	**12.223**	**0.321**	**15.325**	**79.089**	**494.341**
VMD-CNN	15.877	0.417	20.058	77.226	688.134
VMD-BPNN	15.480	0.408	19.480	78.675	694.284
EMD-LSTM	16.466	0.430	21.635	77.296	564.699
EMD-ELM	17.280	0.453	21.933	74.741	625.243
EMD-CNN	24.465	0.644	29.665	76.847	802.815
EMD-BPNN	21.311	0.563	26.676	77.019	739.014
LSTM	33.184	0.867	43.690	50.104	982.237
ELM	33.830	0.895	44.355	49.897	1090.735
CNN	39.133	1.016	50.826	50.311	1392.752
BPNN	36.100	0.938	48.018	48.654	1115.632

Table 4 Predictive performance evaluation of different models for FTSE

Model	MAE	MAPE (%)	RMSE	D_{stat} (%)	CID
VMD-LSTM	17.187	0.236	22.485	76.008	801.733
VMD-ELM	**16.651**	**0.229**	**21.716**	75.807	**750.774**
VMD-CNN	18.559	0.255	23.973	75.395	871.849
VMD-BPNN	19.307	0.265	24.515	76.210	864.928
EMD-LSTM	21.054	0.289	27.869	**76.411**	770.592
EMD-ELM	24.337	0.334	31.042	73.992	907.133
EMD-CNN	23.506	0.322	30.440	74.411	884.966
EMD-BPNN	25.685	0.353	32.489	**76.411**	910.515
LSTM	45.592	0.624	58.784	50.605	1334.329
ELM	42.704	0.587	55.603	49.194	1265.068
CNN	44.081	0.607	58.946	51.613	1481.541
BPNN	42.590	0.586	56.620	51.210	1352.156

in the SPX and IXIC price forecasting (see Table 3 and Table 5). However, the directional forecasting measure D_{stat} values for the BPNN, ELM, CNN and LSTM are all smaller than those for the hybrid models, which can also be clearly seen in Fig. 7, suggesting that the EMD-based and VMD-based models greatly enhance the predictive ability compared to the single models. This could be explained by the fact that EMD and VMD can help to efficiently determine the inner patterns of stock price fluctuations, which have complex non-linear and non-stationary features. In terms of the CID, its values for HSI and FTSE further indicate the better forecasting ability of the hybrid models, while for SPX and IXIC, its values for the hybrid models EMD-BPNN (1180.604 for SPX, 2782.508 for IXIC) and VMD-BPNN (949.961 for SPX, 2868.013 for IXIC) are larger than those of the single models (see Tables 3 and 5).

(2) The predictive accuracy of each single model can be improved effectively by combining with decomposition methods, and the VMD-based models usually outperform the EMD-based models. Taking the HSI (shown in Table 2) for discussion, the level forecasting indices (MAE, MAPE and RMSE) for the LSTM model have values of 33.184, 0.00867 and 43.690, respectively, while the indices for the EMD-LSTM model decrease to 16.466, 0.0043 and 21.635, respectively, with a reduction of almost 50% for each measure. The indices for the VMD-LSTM model decrease to 12.365, 0.00324 and 15.744, respectively, with a further reduction of 24.9%, 24.6% and 27.2%. The directional forecasting measure D_{stat} of LSTM is 50.104 (%), which means that making either correct or wrong predictions in a direction is almost meaningless, but that of EMD-LSTM and VMD-LSTM

Table 3 Predictive performance evaluation of different models for SPX

Model	MAE	MAPE (%)	RMSE	D_{stat} (%)	CID
VMD-LSTM	**7.096**	**0.256**	**9.674**	85.425	**309.564**
VMD-ELM	13.035	0.461	18.284	**85.830**	605.386
VMD-CNN	17.658	0.635	22.386	81.579	790.017
VMD-BPNN	21.091	0.748	28.151	83.198	949.961
EMD-LSTM	13.195	0.475	16.592	77.530	516.622
EMD-ELM	10.298	0.370	13.950	75.506	471.086
EMD-CNN	13.722	0.494	19.389	67.814	778.138
EMD-BPNN	36.011	1.252	44.263	73.684	1180.604
LSTM	18.730	0.676	26.129	52.024	646.525
ELM	19.984	0.717	27.407	51.822	642.958
CNN	21.165	0.764	29.194	50.000	849.893
BPNN	25.492	0.915	35.726	49.798	843.668

Table 5 Predictive performance evaluation of different models for IXIC

Model	MAE	MAPE (%)	RMSE	D_{stat} (%)	CID
VMD-LSTM	**32.055**	**0.427**	**42.324**	**85.020**	1417.136
VMD-ELM	42.982	0.560	53.674	82.794	1782.457
VMD-CNN	45.884	0.614	59.216	82.389	2231.005
VMD-BPNN	67.243	0.891	82.880	84.211	2868.013
EMD-LSTM	33.368	0.444	44.205	76.316	**1370.362**
EMD-ELM	34.924	0.467	48.949	72.672	1635.740
EMD-CNN	43.404	0.579	59.654	71.053	2018.640
EMD-BPNN	94.335	1.221	112.427	74.494	2782.508
LSTM	56.627	0.752	77.565	57.315	1975.553
ELM	83.522	1.094	103.792	47.773	2503.118
CNN	71.605	0.957	95.438	48.178	3132.509
BPNN	68.310	0.908	91.649	48.785	2073.950

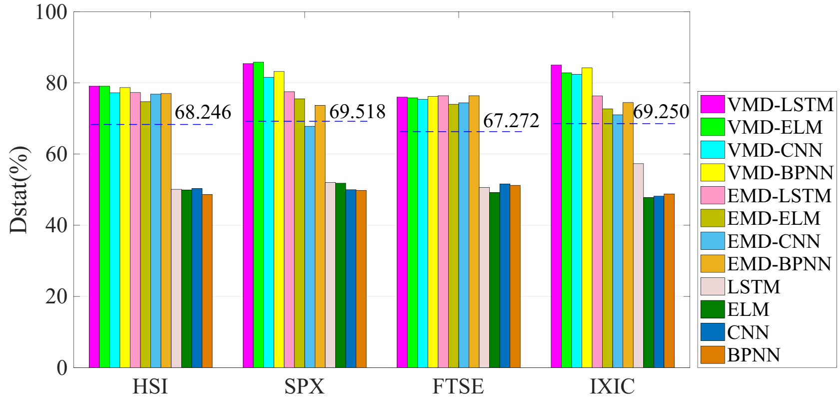

Fig. 7 D_{stat} values for different models on the four selected stock price indices; the dotted line represents the average of D_{stat} for the different models

reaches 77.296 (%) and 79.089 (%), respectively, representing increases of 27.19 (%) and 28.69 (%). This means that the predictive accuracy has been significantly improved from LSTM to EMD-LSTM and further to VMD-LSTM. In addition, the CIDs of LSTM, EMD-LSTM and VMD-LSTM are 982.237, 564.599 and 522.188, respectively, which further validate the fact that the VMD-LSTM model has a better forecasting performance than that of EMD-LSTM for the HSI stock index. Likewise, the performances of the VMD-ELM, VMD-CNN and VMD-BPNN models are superior to those of the EMD-ELM, EMD-CNN and EMD-BPNN models, respectively. The same results are also obtained for the SPX, FTSE and IXIC data (except for EMD-BPNN < BPNN < VMD-BPNN and CNN < VMD-CNN < EMD-CNN for SPX and IXIC according to the level forecasting indices). In addition, among the four single models, the LSTM model performs best for the HSI, SPX and IXIC indices. However, after combining with EMD and VMD, the performance of the ELM-based forecasting models is better than that of the LSTM-based forecasting models in some experiments (VMD-LSTM < VMD-ELM for HSI and FTSE and EMD-LSTM < EMD-ELM for SPX according to the level forecasting indices), but the superiority is not as significant as that of the LSTM-based models over the ELM-based models. Among these models, the VMD-ELM and VMD-LSTM models obtain comparable performances. For the HSI and FTSE datasets, the predictive performance measured by the VMD-ELM is optimal according to the level forecasting indices and the CID, being slightly better than that of VMD-LSTM. However, VMD-LSTM performs significantly better than the VMD-ELM for the SPX and IXIC datasets. The results indicate that VMD-LSTM has an excellent and stable performance that cannot be matched by the other models. In addition, the EMD-CNN (except for HSI) and VMD-CNN models perform worse than other hybrid models of the same class in directional forecasting.

(3) The predictive accuracy of the proposed VMD-LSTM model applied to different datasets is generally better than that of the other compared models, especially in SPX and IXIC forecasting, where VMD-LSTM holds the relatively smallest forecasting errors (the performance evaluation measures are marked in bold in the tables). In terms of the results for HSI, VMD-LSTM performs slightly worse than the VMD-ELM, but the differences in the performance evaluation indices MAE, MAPE, and RMSE between the two models are controlled within a small range, and the D_{stat} values of the two models both reach the same level of 79.089%, which is the highest level among all results. The same results are also shown for the FTSE. This effectively verifies the excellent prediction capability of the proposed VMD-LSTM model.

(4) In terms of the newly introduced CID evaluation index, its variation tendencies for different models are not totally

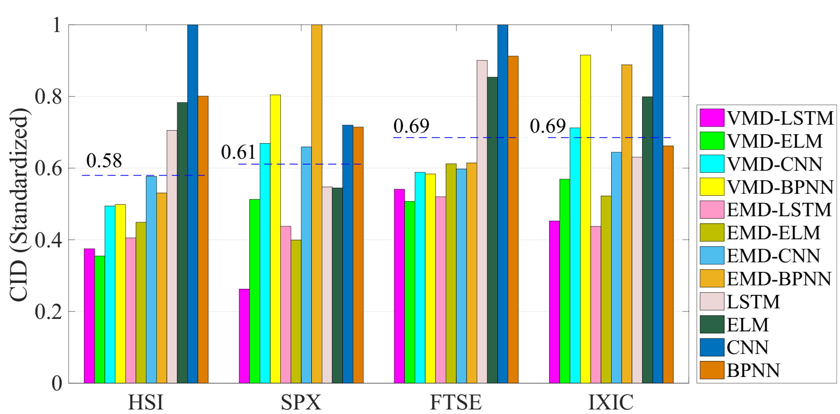

Fig. 8 The standardized CIDs for different models on the four selected stock price indices; the dotted line represents the average *CID* for the different models

consistent with the level and directional criteria, but it can effectively reflect the forecasting performance. The tables and Fig. 8 depict the CIDs for different models of the four stock price indices. It is shown that on the whole, a smaller CID value responds better to the forecasting results. Taking HSI as an example, the variation tendency of the CID is almost the same as that of the MAE, MAPE and RMSE level forecasting errors, except for the VMD-BPNN (CID = 694.284).

4 Conclusion

In this paper, a hybrid neural network VMD-LSTM model is developed to improve the accuracy in predicting stock price indices, which often exhibit complex properties such as nonlinearity, nonstationarity and uncertainty. The VMD algorithm is used to decompose the original time series into a limited number of subseries, and LSTM is used to construct a training and predicting model for each subseries. The final predictive result is obtained by an aggregation of all the predicted subseries. Compared with the EMD technique, VMD has the advantages of decomposing data series with high accuracy, strong noise robustness, and fast convergence. Four stock price indices (HSI, SPX, FTSE, and IXIC) from the world financial market are studied as cases. After a comprehensive comparative analysis with four single models, i.e., the BPNN, ELM, CNN, and LSTM, and seven hybrid models, i.e., the EMD-BPNN, EMD-ELM, EMD-CNN, EMD-LSTM, VMD-BPNN, VMD-EL,M and VMD-CNN, the experimental results show first that the differences between the actual values and predicted values decrease significantly after combining a single model with EMD and VMD, which implies that the hybrid models outperform the single models. Then, the VMD-ANN models are shown to exhibit better prediction ability than that of the popular EMD-ANN models (for example, VMD-LSTM > EMD-LSTM > LSTM), which suggests that the VMD algorithm is a promising tool for financial time series forecasting. In addition, the performance of LSTM is superior to that of the ELM, CNN and BPNN models for all datasets, and VMD-LSTM has a better prediction capability than that of the VMD-ELM, VMD-CNN, and VMD-BPNN for the SPX and IXIC data series. In this work, the level forecasting measures (MAE, RMSE, MAPE) and directional forecasting index D_{stat} as well as a new complexity-invariant distance (CID) from information theory are used to evaluate the predictive accuracies. The experimental results indicate the effectiveness of the CID in assessing the stock price forecasting effect. We hope that our proposed model can contribute to the research on neural networks and the predictions of time series data in other fields.

Acknowledgments The work was partially supported by the Humanities and Social Sciences Foundation of Ministry of Education of China (No. 18YJCZH134, 18YJC790106) and the Fundamental Research Funds for the Central Universities (No. FRF-BR-18-001B).

References

1. Nam K, Seong N (2019) Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market. Decis Support Syst 117:100–112
2. Song Y, Lee J, Lee J (2019) A study on novel filtering and relationship between input-features and target-vectors in a deep learning model for stock price prediction. Appl Intell 49:897–911
3. J. Wang, J. Wang, Forecasting stock market indexes using principle component analysis and stochastic time effective neural networks, Neurocomputing156 (2015) 68–78
4. Zhang C, Pan H, Ma Y, Huang X (2019) Analysis of Asia Pacific stock markets with a novel multiscale model. Physica A 534: 120939
5. Loureiro ALD, Miguéis VL, Da SLFM (2018) Exploring the use of deep neural networks for sales forecasting in fashion retail. Decis Support Syst 114:81–93
6. Herzog S, Tetzlaff C, Wörgötter F (2020) Evolving artificial neural networks with feedback. Neural Netw 123:153–162
7. Adhikari R, Agrawal RK (2014) A combination of artificial neural network and random walk models for financial time series forecasting. Neural Comput Appl 24:1441–1449
8. Kuremoto T, Kimura S, Kobayashi K, Obayashi M (2014) Time series forecasting using a deep belief network with restricted boltzmann machines. Neurocomputing 137:47–56
9. Niu H, Wang J (2014) Financial time series prediction by a random data-time effective RBF neural network. Soft Comput 18:497–508
10. Wang J, Wang J (2017) Forecasting stochastic neural network based on financial empirical mode decomposition. Neural Netw 90:8–20
11. Das SP, Achary NS, Padhy S (2016) Novel hybrid SVM-TLBO forecasting model incorporating dimensionality reduction techniques. Appl Intell 45:1148–1165
12. I. Goodfellow, Y. Bengio, A. Courville, Deep learning: Cambridg, MIT press. 2016
13. Mandic DP, Chambers JA (1999) Exploiting inherent relationships in RNN architectures. Neural Netw 12:1341–1345
14. Deng T, He X, Zeng Z (2018) Recurrent neural network for combined economic and emission dispatch, applied intelligence. Appl Intell 48:2180–2198
15. Guo T, Xu Z, Yao X, Chen H, Aberer K, Funaya K (2016) Robust online time series prediction with recurrent neural networks, in: IEEE international conference on data science and advanced analytics. IEEE.:816–825
16. Hajiabotorabi Z, Kazemi A, Samavati FF, Ghaini FMM (2019) Improving DWT-RNN model via B-spline wavelet multiresolution to forecast a high-frequency time series. Expert Syst Appl 138: 112842
17. Wang Q, Xu W, Huang X, Yang K (2019) Enhancing intraday stock price manipulation detection by leveraging recurrent neural networks with ensemble learning. Neurocomputing 47:46–58
18. Berradi Z, Lazaar M (2019) Integration of principal component analysis and recurrent neural network to forecast the stock Price of Casablanca stock exchange. Procedia Computer Science 148: 55–61
19. Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9:1735–1780

20. Sundermeyer M, Ney H, Schlüter R (2015) From feedforward to recurrent LSTM neural networks for language modeling. IEEE/ACM Trans Audio Speech Lang Process 23:517–529
21. Mahmoudi N, Docherty P, Moscato P (2018) Deep neural networks understand investors better[J]. Decis Support Syst 112:23–34
22. Kraus M, Feuerriegel S (2017) Decision support from financial disclosures with deep neural networks and transfer learning[J]. Decis Support Syst 104:38–48
23. Vo NNY, He X, Liu S, Xu G (2019) Deep learning for decision making and the optimization of socially responsible investments and portfolio. Decis Support Syst 124:113097
24. Wang K, Qi X, Liu H (2019) Photovoltaic power forecasting based LSTM-convolutional network. Energy 189:116225
25. Karevan Z, Suykens JAK (2020) Transductive LSTM for time-series prediction: an application to weather forecasting. Neural Netw 125:1–9
26. Ding G, Qin L (2019) Study on the prediction of stock price based on the associated network model of LSTM. Int J Mach Learn & Cyber 11:1307–1317. https://doi.org/10.1007/s13042-019-01041-1
27. S. Chen, L. Ge, Exploring the attention mechanism in LSTM-based Hong Kong stock price movement prediction, Quantitative Finance19 (2019) 1507–1515
28. Baek Y, Kim HY (2018) ModAugNet: a new forecasting framework for stock market index value with an overfitting prevention LSTM module and a prediction LSTM module. Expert Syst Appl 113:457–480
29. Liang X, Ge Z, Sun L et al (2019) LSTM with wavelet transform based data preprocessing for stock Price prediction. Math Probl Eng 2019:1340174
30. Xu P, Xie F, Su T, Wan Z, Zhou Z, Xin X, Guan Z (2020) Automatic evaluation of facial nerve paralysis by dual-path LSTM with deep differentiated network. Neurocomputing 388:70–77. https://doi.org/10.1016/j.neucom.2020.01.014
31. B. Zhang, H. Zhang, G. Zhao, J. Lian, Constructing a PM2.5 concentration prediction model by combining auto-encoder with Bi-LSTM neural networks, Environmental Modelling & Software 124 (2020) 104600
32. Wang H, Yu L, Tian S, Peng YF, Pei XJ (2019) Bidirectional LSTM malicious webpages detection algorithm based on convolutional neural network and independent recurrent neural network. Appl Intell 49:3016–3026
33. Kim T, Kim HY (2019) Forecasting stock prices with a feature fusion LSTM-CNN model using different representations of the same data. PLoS One 14:e0212320
34. Kim HY, Won CH (2018) Forecasting the volatility of stock price index: a hybrid model integrating LSTM with multiple GARCH-type models. Expert Syst Appl 103:25–37
35. Yang D, Yang K (2016) Multi-step prediction of strong earthquake ground motions and seismic responses of SDOF systems based on EMD-ELM method. Soil Dyn Earthq Eng 85:117–129
36. N. Gu, H. Pan, Bearing fault diagnosis method based on EMD-CNNs, CSMA. (2017)
37. Xu W, Peng H, Zeng X, Zhou F, Tian X, Peng X (2019) A hybrid modelling method for time series forecasting based on a linear regression model and deep learning. Appl Intell 49:3002–3015
38. Yu L, Wang S (2008) Kin Keung Lai, forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Econ 30:2623–2635
39. Zhou LG, Fujita H (2017) Posterior probability based ensemble strategy using optimizing decision directed acyclic graph for multi-class classification. Inf Sci 400-401:142–156
40. Sun J, Li H, Fujita H, Fu BB, Ai WG (2020) Class-imbalanced dynamic financial distress prediction based on Adaboost-SVM ensemble combined with SMOTE and time weighting. Information Fusion 54:128–144
41. J. Sun, H. Fujita, P. Chen, H. Li, Dynamic financial distress prediction with concept drift based on time weighting combined with Adaboost support vector machine ensemble, Knowledge-Based Systems120 (2017) 4–14
42. Cao J, Li Z, Li J (2019) Financial time series forecasting model based on CEEMDAN and LSTM. Physica A 519:127–139
43. Wang J, Tang L, Luo Y, Ge P (2017) A weighted EMD-based prediction model based on TOPSIS and feed forward neural network for noised time series. Knowl-Based Syst 132:167–178
44. Huang NE, Shen Z, Long SR et al (1988) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings: mathematical. Phys Eng Sci 454:903–995
45. Awajan AM, Ismail MT, al Wadi S (2018) Improving forecasting accuracy for stock market data using EMD-HW bagging. PLoS One 13:e0199582
46. Li H (2017) Price forecasting of stock index futures based on a new hybrid EMD-RBF neural network model. Agro Food Industry Hi Tech 28:1744–1747
47. Yang H, Lin H (2016) An integrated model combined ARIMA, EMD with SVR for stock indices forecasting. International Journal on Artificial Intelligence Tools 25:1650005
48. Zhang B, Wang J, Fang W (2015) Volatility behavior of visibility graph EMD financial time series from Ising interacting system. Physica A 432:301–314
49. Huang Y, Hou S, Xu S, Zhao S, Yang L, Zhang Z (2019) EMD-PNN based welding defects detection using laser-induced plasma electrical signals. J Manuf Process 45:642–651
50. Tang L, Wu Y, Yu L (2018) A randomized-algorithm-based decomposition-ensemble learning methodology for energy price forecasting. Energy 157:526–538
51. Upadhyay A, Pachori RB (2015) Instantaneous voiced/non-voiced detection in speech signals based on variational mode decomposition. J Frankl Inst 352:2679–2707
52. Liu C, Cheng G, Chen X, Pang Y (2018) Planetary gears feature extraction and fault diagnosis method based on VMD and CNN. Sensors 18:1523
53. Lahmiri S (2014) Comparative study of ECG signal denoising by wavelet thresholding in empirical and variational mode decomposition domains. Healthc Technol Lett 1:104–109
54. Lahmiri S (2016) A variational mode decomposition approach for analysis and forecasting of economic and financial time series. Expert Syst Appl 55:268–273
55. Abdoos AA (2016) A new intelligent method based on combination of VMD and ELM for short term wind power forecasting. Neurocomputing 203:111–120
56. Shahzad SJH, Nor SM, Kumar RR, Mensi W (2017) Interdependence and contagion among industry-level US credit markets: an application of wavelet and VMD based copula approaches. Physica A 466:310–324
57. Bisoi R, Dash PK, Parida AK (2019) Hybrid Variational mode decomposition and evolutionary robust kernel extreme learning machine for stock price and movement prediction on daily basis. Appl Soft Comput 74:652–678
58. Salim L (2016) Intraday stock price forecasting based on variational mode decomposition. Journal of Computational Science 12(2016):23–27
59. Liu Y, Yang C, Huang K, Gui W (2019) Non-ferrous metals price forecasting based on variational mode decomposition and LSTM network. Knowl-Based Syst 188:105006. https://doi.org/10.1016/j.knosys.2019.105006
60. Liu H, Mi X, Li Y (2018) Smart multi-step deep learning model for wind speed forecasting based on variational mode decomposition, singular spectrum analysis. LSTM network and ELM, Energy Conversion and Management 159:54–64

61. Dragomiretskiy K, Zosso D (2014) Variational mode decomposition. IEEE Trans Signal Process 62:531–544
62. Boyd S, Parikh N, Chu E et al (2011) Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning 3:1–122
63. Makridakis S (1993) Accuracy measures: theoretical and practical concerns. Int J Forecasting 9:527–529
64. G.E.A.P.A. Batista, E.J. Keogh, O.M. Tataw, V.M. Alves de Souza, CID: an efficient complexity-invariant distance for time series, Data Mining and Knowledge Discovery 28 (2014) 634–669