CTU Journal of Innovation and Sustainable Development
ISSN 2588-1418 | e-ISSN 2815-6412

DOI:10.22144/ctujoisd.2024.324

EMD combined with ensemble of machine learning predictors for foreign exchange rate forecasting

Duong Tuan Anh[1,2*], and Tran Van Xuan[2]

[1]Department of Information Technology, Ho Chi Minh City University of Foreign Languages and Information Technology, Viet Nam

[2]Faculty of Computer Science and Engineering, Ho Chi Minh City University of Technology, Viet Nam

*Corresponding author (anhdt@huflit.edu.vn)

Article info.

Received 11 Jun 2024
Revised 29 Aug 2024
Accepted 30 Sep 2024

Keywords

Empirical mode decomposition, foreign exchange rates forecasting, k-nearest neighbors, long short term memory

ABSTRACT

Forecasting foreign exchange rates is a critical financial challenge. In this paper, we build on recent trends and address the limitations of prior research by proposing a novel approach. Our method combines empirical mode decomposition (EMD) with ensemble of machine learning predictors in foreign exchange rate forecasting. To demonstrate that our proposed method (called EMD-ML) is effective, we used the new approach to forecast six foreign exchange rate time series at a specific time. The first experiment was implemented to compare the proposed forecasting model EMD–LSTM, which combines empirical mode decomposition (EMD) with ensemble of Long Short-Term Memory (LSTM) models, and the single LSTM model. The results indicate that the proposed EMD–LSTM model is more effective than the single LSTM. Besides, to aim at comparing deep-learning models against shallow machine learning models in combination with the EMD decomposition, the second experiment compared EMD-LSTM with the approach which combines EMD with an ensemble of k-nearest neighbors' predictors (called EMD-KNN method) and the results of the second experiment show that EMD-LSTM cannot outperform EMD-KNN in foreign exchange rates forecasting.

1. INTRODUCTION

Forecasting foreign exchange rates is an important financial problem which is receiving increasing interest because of its intrinsic difficulty and practical applications.

For many years, researchers have developed various methods for foreign exchange rate forecasting problems such as statistical time series models, shallow artificial neural networks, and support vector machines. However, the obtained performances are still limited due to the inherently noisy and non-stationary behavior of the data. With the strong development of deep learning neural networks in recent years, there are more and more approaches and algorithms to solve foreign exchange rate forecasting problems such as convolutional neural network (CNN) (Galeshchuk & Mukherjee, 2017), deep belief network (DBN) (Chao et al., 2011; Shen et al., 2015), long short-term memory (LSTM) (Escudero et al., 2021), etc. These methods have proven effective in forecasting complex problems from many fields. The main characteristic of these models is the ability to extract and learn abstract features of data. Recently, some researchers have also applied autoencoder in combination with LSTM for foreign exchange rate prediction (Hoa et al., 2021).

With advances in time series forecasting research, there has been a growing trend of hybridizing some machine learning methods with data pre-processing techniques in order to improve the predictive performance in financial time series forecasting. Various ensemble models which are based on decomposition approach were proposed and used to handle the nonlinearity and non-stationarity present in financial time series data. By combining Empirical mode decomposition (EMD) with various forecasting models, we can build a hybrid forecasting approach. EMD can decompose a time series into a collection of intrinsic mode functions (IMFs) and a residue. Each IMF is much easier to analyze than the original time series. In other words, with EMD decomposition, a difficult forecasting problem can be divided into an ensemble of simpler forecasting problems. So, the main idea of EMD-based forecasting method is “decomposition-and-ensemble” which has the spirit of “divide and conquer” principle.

EMD has been widely used to preprocess non-stationary original time series which can improve the forecasting accuracy. EMD combined with an ensemble of ANN (artificial neural network) predictors or KNN (k-nearest neighbors) predictors yields a good performance in crude oil price forecasting (Yu et al., 2008), in wind speed forecasting (Liu et al., 2012) and in financial time series forecasting (Lin et al., 2012). Moreover, EMD combined with an ensemble of DBN predictors or LSTM predictors shows a good performance in the prediction of electric load (Qiu et al., 2017), in drought forecasting (Agana & Homaifar, 2018), in wave height forecasting (Zhou et al., 2021), in wind power prediction (Zhang et al., 2020), and in metro passenger flow forecasting (Chen et al., 2019). However, there have been very few research works which apply EMD decomposition in foreign exchange rate forecasting.

To address the non-stationary and nonlinearity of foreign exchange rate data, this study proposes an ensemble method for foreign exchange rates prediction which consists of two steps. At first, time series is decomposed by EMD method into some IMFs and a residue. Then, with different IMFs and residue, machine learning predictors are used in forecasting step. This ensemble method is called EMD-ML

In this experiment, we use Euro/China Yuan (EUR_CNY), Euro/US dollar (EUR_USD), British Pound/US dollar (GBP_USD), US dollar/Japanese Yen (USD_JPY), US dollar/Turkey Lira (USD_TRY), and US dollar/Swiss franc (USD_CHF) exchange rates datasets and two evaluation criteria to evaluate the performances. The first experiment was implemented to compare the proposed forecasting model EMD–LSTM, which combines empirical mode decomposition (EMD) with ensemble of Long Short-Term Memory (LSTM) models, and the single LSTM model. The results indicate that the proposed EMD–LSTM model is more effective than the single LSTM. Besides, to aim at comparing deep-learning models against shallow machine learning models in combination with the EMD decomposition method, the second experiment compared EMD-LSTM with the approach which combines EMD with an ensemble of KNN predictors (called EMD-KNN) and the results show that EMD-LSTM cannot outperform EMD-KNN in foreign exchange rates forecasting. The surprising finding of the second experiment indicates that EMD decomposition can benefit forecasting done by shallow machine learning methods but not by deep learning methods.

The remainder of the paper is organized as follows. Section 2 provides background information on EMD decomposition method, LSTM networks and the KNN algorithm. Section 3 outlines the proposed EMD-ML approach for forecasting foreign exchange rates. Section 4 presents the experimental results and analysis. Finally, Section 5 offers conclusions and suggestions for future research.

2. BACKGROUND

2.1. Empirical Mode Decomposition

Widely used as a time series decomposition technique, EMD decomposes original time series into a small number of oscillatory modes based on their characteristic time scales in the data empirically. Each oscillatory mode is expressed by an intrinsic mode function (IMF) (Huang et al., 1998). To ensure model effectiveness, all IMFs must satisfy the two following conditions: Firstly, for a given sequence of data values, the number of extrema (maxima and minima) and the number of zero crossings must be equal or different at most by one. Secondly, at any point, the mean value of the envelope defined by the local maxima and minima must be zero.

For an original time series $x(t)$, EMD decomposes $x(t)$ through a process, which is described as follows:

Step 1. For time series $x(t)$, find all the local maxima and local minima of $x(t)$.

Step 2. Using a cubic spline interpolation, generate upper $u(t)$ and lower $l(t)$ envelopes of signal $x(t)$.

Step 3. Compute the mean of the two envelopes to get the mean envelope $m(t)$:

$$m(t) = (u(t) + l(t))/2$$

Step 4. To extract an IMF candidate, subtract $m(t)$ from $x(t)$:

$$c(t) = x(t) – m(t)$$

Step 5. Check whether $c(t)$ is an IMF:

(i) If $c(t)$ satisfies the two IMF conditions, $c(t)$ becomes an IMF and meantime replace $x(t)$ with the residue $r(t) = x(t) – c(t)$;

(ii) If $c(t)$ is not an IMF, replace $x(t)$ with $c(t)$.

Repeat Steps 1) - 5) until the residue satisfies the stop criterion.

The EMD extracts the next IMF by applying the above procedure to the residual term $r_1(t) = x(t) – c_1(t)$, where $c_1(t)$ is the first IMF.

Stop criterion. The above-mentioned process can be repeated until the last residue $r_n(t)$ has the number of extrema less than or equal to 1 or becomes a monotonic function, such that no more IMFs can be derived. Due to the stop criterion of the sifting procedure, n is not a predefined parameter of the EMD algorithm since it can be determined by the algorithm itself.

At the end of the above algorithm, the time series $x(t)$ can finally be presented as

$$x(t) = \sum_{j=1}^{n} c_j(t) + r_n(t)$$

where n is the number of IMF components, $r_n(t)$ the final residue of the time series $x(t)$ and $c_j(t)$ ($j = 1,2,...,n$) are the IMF components. All these components have the mean approximating to zero.

Figure 1 shows all the five IMFs and r_n (residue) of a stock time series.

Relative to wavelet decomposition, EMD is a heuristic method that is based on the characteristics of the time series on a local scale. It decomposes the time series without setting a predefined basis function in which the time series is expressed (Huang et al., 1998).

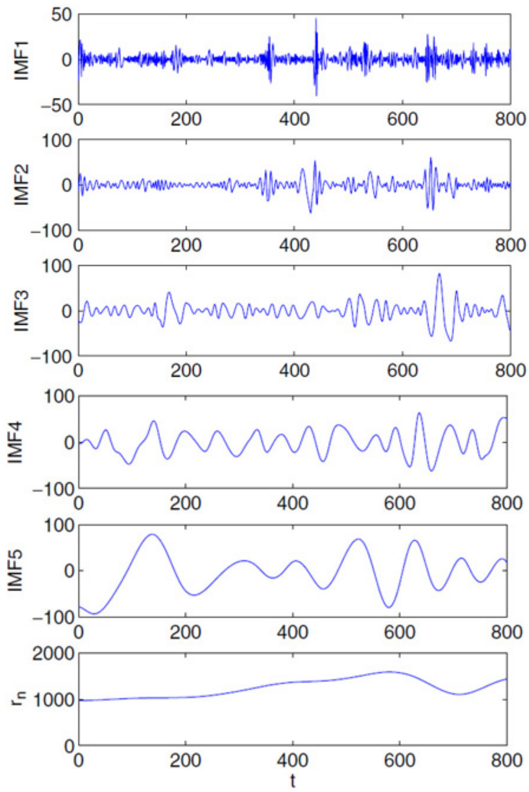

Figure 1. The EMD decomposition of a stock time series (Lin et al., 2012)

2.2. Long Short-Term Memory (LSTM) network

Recurrent Neural Networks (RNN) were developed as a variant of ANN for sequential data. For training RNNs, the modified version of backpropagation algorithm, Backpropagation Through Time (BPTT) algorithm, was developed and used. However, the application of BPTT incurs two main problems: exploding gradient problem or vanishing gradient problem. Long short-term Memory (LSTM) is an improved version of RNN, proposed by Hochreiter and Schmidhuber (1997) to solve the weakness of RNN in dealing with long-term dependencies, i.e. exploding gradient problem and vanishing gradient problem. Each LSTM unit is a generalization of RNN unit, such that part of information about previous time series data points is stored into the network.

Each LSTM unit (also called LSTM cell) has three gates:

– Forget gate, which is responsible for deciding which part of information from the previous state should be saved or discarded.

– Output gate, which is responsible for selecting how much information should be output.

– Input gate, which is responsible for obtaining new information.

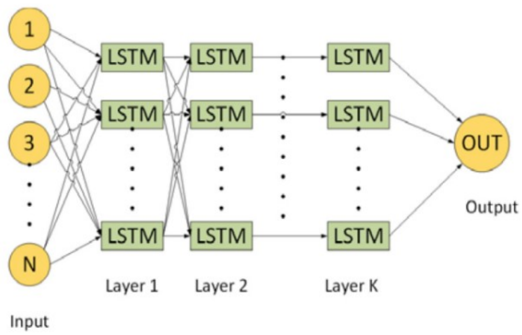

Figure 2. A stacked LSTM network

The deep LSTM neural network contains more than one hidden layer. It consists of many layers of LSTM cells in which the outputs of the previous layer become the inputs of the next layer. The structure of a stacked LSTM network which can be utilized in time series forecasting is depicted in Figure 2.

As for forecasting in complicated time series, such as foreign exchange rate data, among deep neural network models, LSTM is the most suitable deep neural network to deal with this challenging forecasting problem (Hoa et al., 2021). This is the main reason why LSTM is selected to be used as a predictor in this study.

2.3. KNN

In this study, besides LSTM, we use k-nearest neighbors (KNN) method as a predictor in forecasting foreign exchange rates forecasting. This method is a nonparametric method in time series forecasting. Nonparametric prediction does not require any prior knowledge about the process to be modeled. Among nonparametric methods, the KNN algorithm has been shown to be appealing and has been successfully applied in various forecasting studies because of its ability to handle high-dimension and incomplete data (Lin et al., 2012). The forecasting ability and simplicity of KNN make its adaptation to foreign exchange rate data suitable. Due to all these reasons, KNN is selected to be used as a typical shallow machine learning predictor in our study in order to compare against LSTM, a deep neural network predictor.

The application of KNN for time series forecasting is outlined as follows; given a series $T = (t_1, t_2, \ldots, t_n)$, the purpose is to forecast t_{n+h}, where h is the forecast horizon. For predicting a pattern Q with target t_{n+1}, which contains $(t_{n-l}, \ldots, t_{n-1}, t_n)$, the KNN regressor algorithm searches for the k most similar subsequences to Q. When the k most similar subsequences are identified, the target of Q is derived by averaging the targets of the k found subsequences. The l parameter in KNN is called the *pattern size*.

3. THE METHOD

In this paper, the proposed hybrid model for foreign exchange rate forecasting, namely EMD-ML, combines EMD with ensemble of machine learning predictors. In EMD-ML, there are three stages: (1) EMD stage, (2) machine learning model building stage, and (3) machine learning prediction stage. The first stage (EMD stage) decomposes the original time series into a small number of IMFs. In the second stage, a specific method is used to determine the suitable parameter settings of the machine learning model for each decomposed IMF and residue. The third stage applies the trained machine learning model as a forecasting tool to perform the prediction task. The prediction results of all extracted IMFs and the residue are combined to generate an aggregated output which can be seen as the final prediction result for the original time series. The overall framework of the proposed model is illustrated in Figure 3.

The flow of the proposed model is described as follows:

Step 1: *Normalization*. Because the foreign exchange rates time series are highly nonlinear and time-varying, the original time series must be normalized into the range [0, 1] to avoid large fluctuations during the training process. The formula for normalization is given as follows:

$$x'_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \times (x'_{max} - x'_{min}) + x'_{min} \qquad (1)$$

where x_i is the actual value at time point i, x_{max} is the largest value and x_{min} is the smallest value of the time series, x'_i is the normalized value and $[x'_{min}, x'_{max}]$ is the new range we want to convert into. Since we want to normalize the time series to [0, 1], i.e. $x'_{max} = 1$, $x'_{min} = 0$, the formula of the normalization in Eq. (1) becomes:

$$x'_i = (x_i - x_{min})/(x_{max} - x_{min}) \qquad (2)$$

Step 2: *Decomposition*. The original daily foreign exchange rates series $X(t)$ $(t = 1, 2, \ldots, n)$ is decomposed into many components, IMFi(t) (i = 1, 2, ..., k), and the residue. The extracted IMF components are derived from high frequency to low frequency. The details of the EMD algorithm are described in Section 2.1.

Step 3: *Building machine learning ensemble*. After the decomposition process, some specific method is used to determine the suitable parameter setting for the machine learning model for each of the IMFs and the residue.

Step 4: *Forecasting*. The machine learning models, built in Step 3, are applied as predictors to perform the corresponding prediction tasks for each extracted IMF.

Step 5: *Summation*. The prediction results of all extracted IMFs and the residue are combined to generate an aggregated output which can be seen as the final prediction result for the original time series.

Step 6: *Denormalization*. The forecasting result is converted back to the range of the original time series.

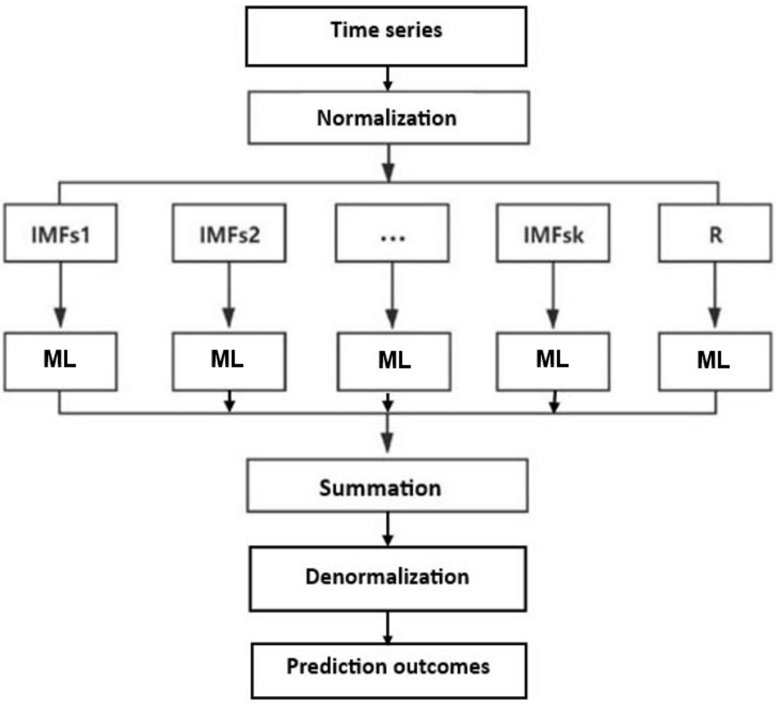

Figure 3. The hybrid EMD-ML model for foreign exchange rates forecasting

Notice that EMD-ML is a general framework and there are many options for selecting the prediction model in Step 4. That means each predictor in the ensemble can be a deep neural network model or a shallow machine learning model.

4. RESULTS AND DISCUSSION

To assess the prediction results of the proposed approach EMD-ML, we implement the two following experiments.

In Experiment 1, EMD-LSTM, which combines empirical mode decomposition (EMD) with an ensemble of LSTM models, is compared to the single LSTM model in foreign exchange rate forecasting.

In Experiment 2, we compare the performances of the four foreign exchange rates forecasting methods: EMD-LSTM, the single LSTM model, the EMD-KNN model which combines EMD decomposition with an ensemble of k-nearest-neighbors (KNN) predictors, and the single KNN model. Experiment 2 aims to compare the performance of EMD combined with a deep neural network to that of EMD combined with a shallow machine learning method.

We implemented EMD decomposition method by using the library PyEMD, a Python library for EMD and some of its variants (Quinn et al., 2021). We implemented LSTM models by using Keras library (the website: http://keras.io). As for k-nearest-neighbors algorithm, we adopt an improved variant of KNN given by Lin et al. (2012). We conducted the experiments on a 4 Intel(R) Xeon(R) CPU@ 2.30GHz RAM26GB PC (with GPU Tesla P100-PCIE-16GB).

Finally, only one-day ahead forecasting (a form of short-term forecasting) is considered in this research.

4.1. The datasets

We tested the comparative forecasting methods on six foreign exchange rates datasets. These datasets are from the website https://www.investing.com/currencies/. The datasets are collected from 03 January 2000 to 23 April 2023. The datasets are described as follows.

1. EUR_CNY dataset: the daily foreign exchange rates between Euro and China yuan.
2. EUR_USD dataset: the daily foreign exchange rates between the Euro and the US dollar.
3. GBP_USD dataset: the daily foreign exchange rates between the British pound and the US dollar.

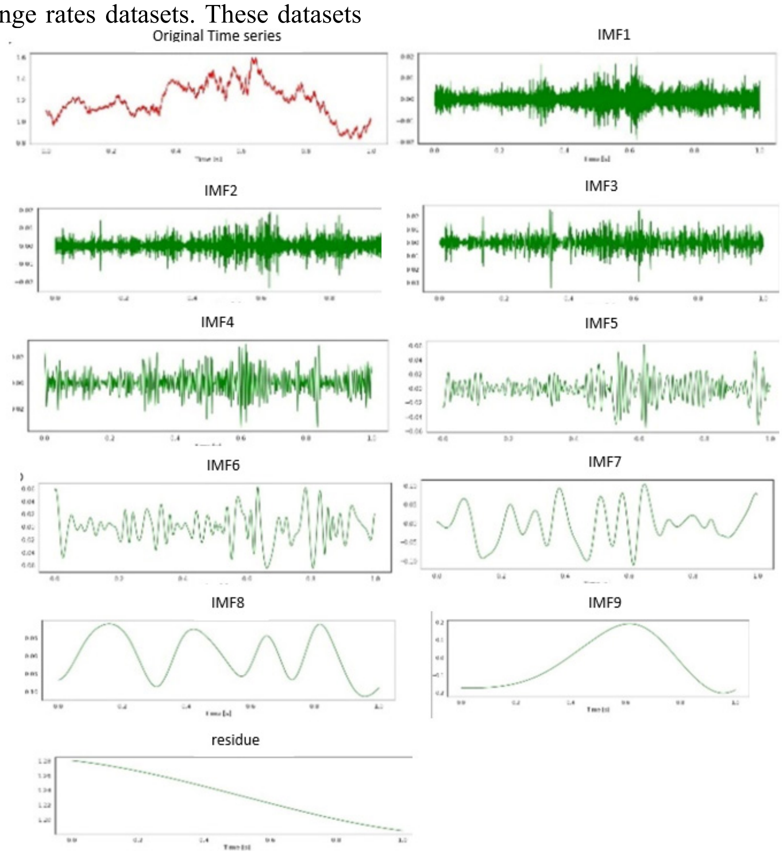

Figure 4. The extracted IMF components of the time series EUR_USD

4. USD_JPY dataset: the daily foreign exchange rates between the US dollar and Japanese yen.
5. USD_TRY dataset: the daily foreign exchange rates between the US dollar and the Turkey lira.
6. USD_CHF dataset: the daily foreign exchange rates between the US dollar and Swiss franc.

Using pyEMD library, we can get graphical representations of the decomposition results for all foreign exchange rates time series. Figure 4 shows the decomposition results for EUR_USD time series which consists of the original time series, nine IMFs and one residue.

Each of the six datasets is divided into training dataset, validation dataset and testing dataset. The training dataset contains 70%, validation dataset

15% and testing datasets 15% of the entire dataset. We normalize the time series to [0, 1], using the formula given in Eq. (2).

4.2. Accuracy evaluation criteria

In this study, to measure the forecasting performance, two main criteria: the root mean squared error (RMSE), and the mean absolute percentage error (MAPE) are used. The calculation equations of RMSE and MAPE are defined as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (3)$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{|\hat{y}_t - y_t|}{y_t} \times 100\% \quad (4)$$

where y_t is the actual value at time point t, $\hat{y}_t$ is the forecast value at time point t and n is the number of observations

These two measures represent different perspectives in evaluating forecasting methods. The first (RMSE) is an absolute performance metric while the second one (MAPE) is a relative metric. The MAPE is a scale-invariant statistic that expresses error as a percentage. The method has better predictive accuracy when RMSE or MAPE is much closer to 0.

4.3. Experiment 1

To determine the best hyper-parameters (the number of units in the input layer, the number of hidden layers, the number of units in hidden layers) for the LSTM network which is used in each LSTM of the ensemble or in the single LSTM model for each foreign exchange rates dataset, we applied Grid Search (Buslim et al., 2021) on the validation set of each dataset.

Grid Search is a widely used hyper-parameter optimization technique in machine learning. It systematically searches for the optimal combination of hyperparameters by exhaustively evaluating the model performance across a predefined grid of hyperparameter values. The process of Grid Search involves specifying a set of hyperparameter values to explore for each hyperparameter in the model. These values are typically chosen based on prior knowledge or domain expertise. The model is then trained and evaluated for each combination of hyperparameters on a validation dataset. The equation for Grid Search can be defined as follows:

$$BestHyperparameters = argmax_{hyperparameters} AccuracyMetric \quad (5)$$

where *BestHyperparameters* represents the combination of hyperparameters that maximizes the chosen accuracy metric.

Grid Search allows us to evaluate the model's performance systematically across various hyper-parameter configurations, enabling us to identify the optimal set of hyperparameters for our specific task. By exhaustively searching the hyperparameter space, Grid Search helps us avoid the bias of manually selecting hyperparameters and provides a more objective and thorough evaluation. Figure 5 illustrates the use of Grid Search in tuning the three parameters of LSTM for EUR_USD dataset.

The best fit hyper-parameter values of LSTM for all the six datasets are given in Table 1. Notice that the number of units in input layer is also called *time-lag*.

In Experiment 1, we compare the performance of EMD-LSTM to that of the single LSTM model in foreign exchange rate forecasting. The MAPE errors (with Eq. (4)) and the RMSE errors (with Eq. (3)) of the two methods are given in Table 2 and Table 3. In Table 2 and Table 3, the smallest errors are in bold.

Based on both the MAPE indicator and RMSE indicator, the EMD-LSTM approach performs better than single LSTM in all datasets.

Table 1 Hyperparameters results of the LSTM

Dataset	Number of hidden layers	Number of units in each hidden layer	Number of units in input layer
EUR_CNY	5	500	8
EUR_USD	2	500	10
GBP_USD	4	500	16
USD_JPY	2	300	6
USD_TRY	1	400	26
USD_CHF	4	400	14

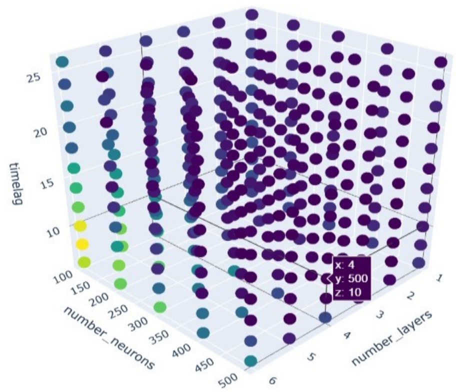

Figure 5. Illustration of grid search to tune the three hyperparameters of LSTM for EUR_USD dataset

As for EMD-LSTM, the average improvement based on MAPE in comparison to the single LSTM on the six datasets is 15.41% and the average improvement based on RMSE in comparison to the single LSTM on the six datasets is 25.14%. The reason for this improvement is that pre-processing the original time series with EMD reduces the complexity in the data, allows noise reduction and therefore improves prediction accuracy of EMD-LSTM in comparison to the single LSTM.

Table 2. MAPE comparisons for the two methods

Dataset	Single LSTM	EMD-LSTM
EUR_CNY	0.3429%	**0.2944%**
EUR_USD	0.3731%	**0.2530%**
GBP_USD	0.4674%	**0.4472%**
USD_JPY	0.3903%	**0.3369%**
USD_TRY	0.6025%	**0.5608%**
USD_CHF	0.3760%	**0.2961%**

Table 3. RMSE comparisons for the two methods

Dataset	Single LSTM	EMD-LSTM
EUR_CNY	0.0337	**0.0278**
EUR_USD	0.0054	**0.0034**
GBP_USD	0.0080	**0.0076**
USD_JPY	0.7355	**0.5900**
USD_TRY	0.1948	**0.1433**
USD_CHF	0.0048	**0.00394**

4.4. Experiment 2

Applying Grid Search on the validation set, we determined the best parameters (k: the number of nearest neighbors, and l: the pattern size) for the KNN predictor which is used in each KNN of the ensemble or the single KNN model for each foreign exchange rates dataset.

Table 4. Parameters tuning results of the KNN

Dataset	Pattern size l	Parameter k
EUR_CNY	6	10
EUR_USD	9	19
GBP_USD	3	22
USD_JPY	6	13
USD_TRY	72	10
USD_CHF	33	40

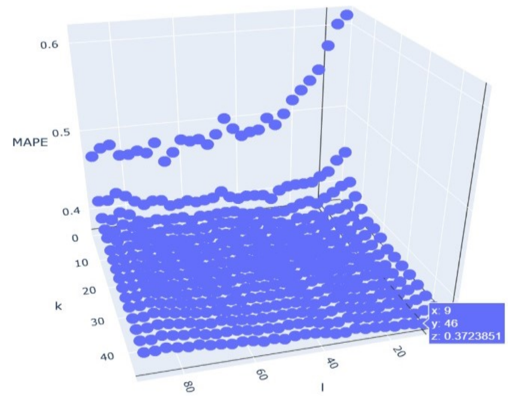

Figure 6. Illustration of grid search to tune the two parameters of KNN for EUR_USD dataset

The best-fit parameter values of the KNN for all the six datasets are given in Table 4. Figure 6 illustrates the use of grid search in tuning the two parameters of KNN for EUR_USD dataset.

In Experiment 2, we compare the predictive accuracy of the four methods: EMD-LSTM, the single LSTM, the EMD-KNN model which combines EMD decomposition with an ensemble of k-nearest-neighbors (KNN) predictors, and the single KNN model. The MAPE comparisons of the four methods are given in Table 6 (the smallest errors are in bold). The RMSE comparisons of the two methods are given in Table 7 (the lowest errors are in bold).

From the experimental results in Table 6 and Table 7, we come with the following observations:

– Based on both the MAPE indicator and RMSE indicator, the single LSTM approach performs better than single KNN in 4 out of 6 datasets. That means the single LSTM performs slightly better than the single KNN.

– Based on both the MAPE indicator and RMSE indicator, the EMD-KNN approach performs better than single KNN in all six datasets.

– Based on the MAPE indicator, EMD-KNN performs better than EMD-LSTM in 4 out of 6 datasets. Based on the RMSE indicator, EMD-KNN performs better than EMD-LSTM in 3 out of 6 datasets. That means EMD-KNN performs slightly better than EMD-LSTM.

Here, we come with a surprising finding: EMD-LSTM cannot outperform EMD-KNN in foreign exchange rate forecasting. The reason of this observation is that after decomposition with EMD, all the components of the original time series become not complicated, stationary and noise-free and then a deep neural network, such as LSTM, cannot bring its great advantages to forecasting these sub-series while a simple machine learning method, such as KNN, can forecast well on these simple sub-series. In other words, EMD decomposition is especially useful in decomposing a complicated time series into a limited set of simple and noise-free time series components. Therefore, EMD decomposition of time series can benefit forecasting done by shallow machine learning methods but not by deep neural network methods.

4.5. Another comparison

Among the six foreign exchange rate datasets used in our experiments, there is only one dataset, EUR_USD, in common with one of the four tested datasets used in the work by Hoa et al. (2021) in foreign exchange rate forecasting. Therefore, we can compare the performances of our two EMD-based methods (EMD-LSTM and EMD-KNN) to those of the three deep learning forecasting methods used in (Hoa et al., 2021) on the EUR_USD dataset. The prediction error comparisons between our two EMD-based methods and the three deep learning methods in (Hoa et al., 2021) on the EUR_USD dataset are shown in Table 5 (the smallest errors are in bold). The results in Table 5 indicate that:

– Our two EMD-based forecasting methods bring out better predictive performances than those of the three deep neural network methods used in (Hoa et al., 2021): convolutional neural network (denoted as Single CNN), LSTM (denoted as Single LSTM), and the combination of AutoEncoder deep neural network and LSTM (denoted as Single AELSTM).

Table 5. The prediction error comparisons between our two EMD-based and the three methods in Hoa et al. (2021)

Methods	RMSE	MAPE
Single CNN	0.01114	0.77927%
Single LSTM	0.00783	0.53735%
Single AELSTM	0.00464	0.29531%
EMD-LSTM	**0.00340**	**0.25300%**
EMD-KNN	**0.00350**	**0.22830%**

Table 6. The MAPE comparisons of the four methods

Dataset	KNN	LSTM đơn	EMD-KNN	EMD-LSTM
EUR_CNY	0.3842%	0.3429%	**0.2074**%	0.2944%
EUR_USD	0.3895%	0.3731%	**0.2284**%	0.2530%
GBP_USD	0.4603%	0.4674%	**0.2899**%	0.4472%
USD_JPY	0.3976%	0.3903%	**0.2533**%	0.3369%
USD_TRY	0.5729%	0.6025%	0.5664%	**0.5608**%
USD_CHF	0.3699%	0.3760%	0.3232%	**0.2961**%

Table 7. The RMSE comparisons of the four methods

Dataset	KNN	LSTM	EMD-KNN	EMD-LSTM
EUR_CNY	0.0376	0.0337	**0.0197**	0.0278
EUR_USD	0.0055	0.0054	0.0035	**0.0034**
GBP_USD	0.0079	0.0080	**0.0049**	0.0076
USD_JPY	0.7386	0.7355	**0.4719**	0.5900
USD_TRY	0.2367	0.1948	0.2351	**0.1433**
USD_CHF	0.0047	0.0048	0.00406	**0.00394**

5. CONCLUSIONS

This paper mainly proposes the hybrid EMD–ML model for short-term foreign exchange rate forecasting. The EMD–ML method is a new forecasting approach by combining EMD decomposition method with machine learning model. The effectiveness of the new approach has been evaluated by performing the experiments on six datasets: EUR_CNY, EUR_USD, GBP_USD, USD_JPY, USD_TRY, and USD_CHF. It was found out that the performance of EMD–LSTM, which combines empirical mode decomposition (EMD) with ensemble of LSTM models, is superior to the single LSTM model in forecasting foreign exchange rates.

However, the EMD-LSTM model cannot outperform EMD–KNN, which combines EMD decomposition method with k-nearest-neighbors method, a shallow machine learning model, to forecast foreign exchange rate time series. These two findings indicate that EMD is especially useful in decomposing a complicated time series into a small set of simple and noise-free time series components.

As for future work, we plan to apply EMD-ML model in forecasting river runoff time series. Moreover, we intend to compare EMD–ML approach to another ensemble approach which combines Seasonal and Trend Decomposition using Loess (STL) with machine learning models (Yin et al., 2020; Chen et el., 2023) in forecasting foreign exchange rate time series.

REFERENCES

Agana, N. A., & Homaifar, A. (2018). EMD-based predictive deep belief network for time series prediction: An application to drought forecasting. *Hydrology*, *5*(1), 18.

Buslim, N., Rahmatullah, I. L., & Setyawan, B. A. (2021). Comparing bitcoin's prediction model using GRU, RNN, and LSTM by hyperparameter optimization grid search and random search. In *Proceedings of the 9th International Conference on Cyber and IT Service Management* (CITSM 2021), Bengkulu, Indonesia, September 22-23.

Chao, J., Shen F., & Zhao, J. (2011). Forecasting exchange rate with deep belief networks. In *Proceedings of IEEE 2011 International Joint Conference on Neural Networks*, San Jose, CA, USA, 31 July-05 August (pp. 1259–1266).

Chen, Q., Wen, D., Li, X., Chen, D., Lv, H., Zhang, J., & Gao, P. (2019). Empirical mode decomposition based long short-term memory neural network forecasting model for the short term metro passenger flow. *PLoS ONE*, *14*(9), e0222365.

Chen, N., Su, C., Wu, S., & Wang, Y. (2023). El Nino Index Prediction based on deep learning with STL decomposition. *Journal of Marine Science and Engineering*, *11*(8), 1529.

Escudero, P., Alcocer, W., & Paredes, J. (2021). Recurrent Neural Networks and ARIMA Models for Euro/Dollar Exchange RateForecasting. *Appl. Sci. MDPI*, *11*(12), 5658.

Galeshchuk, S., & Mukherjee, S. (2017). Deep networks for predicting direction of change in foreign exchange rates. *Intelligent Systems in Accounting, Finance and Management*, *24*(4), 100–110.

Hoa, T. V., Anh, D.T., & Hieu, D. N. (2021). Foreign Exchange Rate Forecasting using Autoencoder and LSTM Networks. In *Proceedings of International Conference on Intelligent Information Technology* (ICIIT), Feb. 25-28, Ho Chi Minh City, Vietnam, ACM (pp. 22-28).

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N. C., Tung, C. C., & Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of Mathematical, Physical and Engineering Sciences*, *454*(1971), 903– 995.

Liu, H., Chen, C., Tian, H., & Li, Y. (2012). A hybrid model for wind speed prediction using empirical mode decomposition and artificial neural networks. *Renewable Energy*, *48*, 545–556.

Lin, A., Shang, P., Feng, G., & Zhong, B. (2012). Application of empirical mode decomposition combined with k-nearest neighbors approach in financial time series forecasting. *Fluctuation and Noise Letters*, *11*(02), p. 1250018.

Qiu, X., Ren, Y., Suganthan, P. N., & Amaratunga, G. A. J. (2017). Empirical mode decomposition based ensemble deep learning for load demand time series forecasting. *Applied Soft Computing*, 54, 246–255.

Quinn, A., Lopes-dos-Santos, V., Dupret, D., Nobre, A., & Woolrich, M. (2021). EMD: Empirical mode decomposition and Hilbert-Huang spectral analyses in Python. *Journal of Open Source Software*, *6*(2977).

Shen, F., Chao, J., & Zhao J. (2015). Forecasting exchange rate using deep belief networks and conjugate gradient method. *Neurocomputing*, *167*, 243-253.

Yin, H., Yin, D., Gu, Y. H., Park, C. J., Han, S. K. & Joo, S. H. (2020). STL-ATTLSTM: Vegetable price forecasting using STL and attention mechanism-based LSTM. *Agriculture*, *10*(12), 612.

Yu, L., Wang, S., & Lai, K.K. (2008). Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics*, *30*(5), 2623–2635.

Zhang, F., Gao, Z., Sun, X. & Xi, J. (2020). Short-term wind power prediction based on EMD-LSTM combined model. *IOP Conf. Ser.:Earth Environ. Sci.*, *514*, 042003.

Zhou, S., Bethel, B.J., Sun, W.; Zhao, Y.; Xie, W. & Dong, C. (2021). Improving significant wave height forecasts using a joint empirical mode decomposition–long short-term memory network. *J. Mar. Sci. Eng*., *9*(7), 744.