Hindawi
Complexity
Volume 2019, Article ID 7458961, 15 pages
https://doi.org/10.1155/2019/7458961

*Research Article*

# Chinese Currency Exchange Rates Forecasting with EMD-Based Neural Network

**Jying-Nan Wang**,[1,2] **Jiangze Du**,[3,4] **Chonghui Jiang**,[3,4] **and Kin-Keung Lai**[5]

[1] *College of International Finance and Trade, Zhejiang Yuexiu University of Foreign Languages, Shaoxing, Zhejiang, China*
[2] *Research Institute for Modern Economics and Management, Zhejiang Yuexiu University of Foreign Languages, Shaoxing, Zhejiang, China*
[3] *School of Finance, Jiangxi University of Finance and Economics, Nanchang, China*
[4] *Research Centre of Financial Management and Risk Prevention, Jiangxi University of Finance and Economics, Nanchang, China*
[5] *Department of Industrial and Manufacturing Systems Engineering, The University of Hong Kong, Pok Fu Lam, Hong Kong*

Correspondence should be addressed to Jiangze Du; jiangze.du@hotmail.com

Received 27 June 2019; Revised 1 September 2019; Accepted 19 September 2019; Published 30 October 2019

Guest Editor: Benjamin M. Tabak



The Chinese currency, RMB, is developing as an international currency. Therefore, the effective strategy for trading RMB exchange rates would be attractive to international investors and policymakers. In this paper, we have constructed hybrid EMD-MLP models to forecast RMB exchange rates and developed a trading strategy based on these models. Empirical results show that the proposed hybrid EMD-MLP* model always performs best based on both NMSE and $D_{\text{stat}}$ criteria when the forecasting period is greater than five days. Moreover, we compare the models' performance using different horizons and find that accuracy will increase with the growth of the forecasting horizons; however, the NMSE will become larger. Lastly, we adopt the best performing model to develop trading strategies with longer forecasting horizons when considering the number of profitable trading activities. If we consider a 0.3% transaction cost, the developed strategy will bring an annual return exceeding 10%, as well as enough trading opportunities.

## 1. Introduction

As the Chinese government continues the process of RMB internationalization, RMB currency trading becomes increasingly important in personal investment, corporate financial decision-making, governments' economic policies, and international trade and commerce. The RMB has therefore attracted an increasing attention from policymakers, investment institutions, and entrepreneurs worldwide. One reflection of this is the 2015 finding of the Society for Worldwide Interbank Financial Telecommunication (SWIFT) that the RMB has overtaken the Japanese yen to become the fourth ranking world payment currency. The first milestone in the process of RMB internationalization is the establishment of an RMB offshore center in Hong Kong in 2004. In December 2008, the Chinese Premier announced the pilot program for cross border trade settlement in RMB, making the offshore RMB officially deliverable in Hong Kong. From July 19, 2010, the offshore market for RMB officially commenced. Recent literature [1–4] has found that RMB movements have an impact on regional currencies.

Due to the importance of the RMB, the purpose of this study is to create a reliable RMB forecasting model and provide a possible application for designing trading strategies. Two kinds of current rates for exchanging the US dollar (USD) into Chinese RMB will be considered. If the RMB is traded onshore (in mainland China), it is referred to as CNY, whereas if traded offshore (mainly in Hong Kong), it is designated as CNH. Since the onshore and offshore markets might respond differently to changes depending on financial markets, the CNY and CNH are not always at the same price level. Craig et al. [5] and Funke et al. [6] study CNY-CNH pricing differentials in detail.

It is well documented that the exchange rate series is considered as nonlinear and nonstationary time series and interactively influenced by many factors, which makes the accurate forecasting of exchange rate rather challenging. During past decades, traditional econometric and statistical techniques, including autoregressive integrated moving average (ARIMA), cointegration analysis, vector autoregression (VAR), and error correction (ECM) models, have been widely used in foreign exchange rates forecasting. However, in the real-world financial markets, exchange rate series are nonlinear and rarely form purely linear combinations [7–10]. Thus, the above traditional models always provide unreliable forecasting if one continue applying these traditional econometric and statistical models. The main reason of this deficiency is that traditional econometric and statistical models are constructed based on linear assumptions, which will be unable to capture the nonlinear features hidden in the exchange rate series.

Considering the limitation of traditional econometric and statistical models, many nonlinear artificial intelligence models (AI), such as artificial neural networks (ANN) [11, 12], feedforward neural networks (FNN) [13, 14], support vector regression (SVR) [15–17], and genetic programming (GP) [18, 19], have been applied to investigate the forecasting ability of financial time series. Yu et al. [20] provide a complete review of foreign exchange rate forecasting with ANN and also introduce SVR and GP. However, many AI-based models also have their own shortcomings. For instances, ANN usually suffers from overfitting and local minima, while other models, including GP and SVR, are sensitive to parameter estimation. Recently, due to the complex characteristics of exchange rate series, several studies show that a single model is unable to capture all the features and make accurate forecasts [21].

To overcome above shortcomings, many researchers start to rely on the hybrid model to forecast exchange rate series accurately. The composite forecasting method allows us to obtain an approach to the dynamics underlying the data by combining the predictions obtained from different individual techniques. Álvarez-Díaz and Álvarez [22] attempt to exploit the nonlinear structure by constructing the genetic programming and neural networks composite method and apply this composite method to forecast Yen/US$ and Pound Sterling/US$ exchange rates. Other scholars also try to improve the forecasting accuracy of exchange rate series by applying hybrid forecasting models. For example, Zhao and Yang [23] and Wong et al. [24] use fuzzy clustering and ANN to forecast financial time series. Yu et al. [25] propose an online big data-driven oil consumption forecasting model based on Google trend by combining the relationship investigation and prediction improvement. Yu et al. [14] design a novel ensemble forecasting approach for complex time series by coupling sparse representation (SR) and feedforward neural network (FNN), i.e., the SR-based FNN approach.

The empirical mode decomposition (EMD) technique is first proposed by Huang et al. [26]. From the theoretical views, EMD is suitable for time series data in terms of decomposing the original data into components, which could break the forecasting task down into simpler forecasting subtasks. These decompositions consist of a finite and often small number of intrinsic mode functions (IMFs) and one residual. From the perspective of financial practice, various economic activities may cause a different impact periodicity on the exchange rate; for example, the company's financial report is released quarterly, but government economic indicators are usually published once a year. EMD is potentially helpful in decomposing these effects and leads to improve forecasting works. There are several studies employing the EMD technique in hybrid models. For instance, Yu et al. [27] forecast crude oil prices with an EMD-based neural network model; Chen et al. [28] also combine EMD and the ANN approach to forecast tourism demand, and Lin et al. [29] propose a hybrid model using EMD and SVR for foreign exchange rate forecasting. In this empirical experiment, the original CNY or CNH price series, with characteristics of nonlinearity and nonstationarity, are divided into several independent subseries by the EMD technique and then partial or all IMFs and one residual are used to forecast. Tang et al. [30] combine the ensemble empirical mode decomposition (EEMD) with random vector functional link (RVFL) network and find the proposed EEMD-based RVFL network performs significantly better in terms of prediction accuracy than not only single algorithms such as RVFL network, extreme learning machine (ELM), kernel ridge regression, random forest, backpropagation neural network, least square support vector regression, and autoregressive integrated moving average, but also their respective EEMD-based ensemble variants. We summarize in Table 1 the main characteristics of these studies on individual and hybrid forecasting methods in the recent literature.

This study will focus on using ANN to forecast the RMB with different forecasting horizons of 1, 5, 10, 20, and 30 days. Nevertheless, Huang et al. [32] show that ANN performs better than the random walk while the forecasting horizon is less than five days, but for longer horizons, such as 10 and 30 days, the general performance of ANN is worse than the random walk model. In this study, a type of feedforward artificial neural network model called multilayer perceptron (MLP) is adopted. Empirical studies consist not only of the pure MLP model but also a hybrid model with EMD to improve forecasting performance.

In this study, the hybrid forecasting model combining MLP and EMD is similar to the work of Yu et al. [27]. However, this study further considers the influence of IMFs that have different levels of frequency. In concrete terms, higher frequency IMFs could be regarded as noise components, when our forecasting horizon is longer. Based on this concept, two hybrid models are proposed and named EMD-MLP and EMD-MLP$^{*}$. Comparing these to pure MLP, the former adjusts the original time series data by subtracting higher frequency IMFs and then uses MLP to compute the one-day ahead predictions. The EMD-MLP$^{*}$ model applies the "divide-and-conquer" principle to construct a novel forecasting methodology, in which partial or

Table 1: Literature on individual and hybrid forecasting models.

| Study | Type | Algorithm | Data | Time range | Criteria |
|---|---|---|---|---|---|
| Aladag et al. [11] | Individual | ANN | Exchange rates (TL/EUR, LEU/EUR) | 2005–2012 | RMSE |
| Alvarez-Diaz and Alvarez [18] | Individual | GP | Exchange rates (various pairs) | 1971–2000 | $R$-square |
| Álvarez-Díaz and Álvarez [22] | Individual | GP, ANN | Exchange rates (GBP/USD, JPY/USD) | 1973–2002 | NMSE, SR |
| Álvarez Díaz [31] | Individual | GP | Exchange rates (GBP/USD, JPY/USD) | 1973–2002 | U-Theil value, SR |
| Huang et al. [32] | Individual | ANN | Exchange rates (USD/GBP, USD/JPY) | 1997–2002 | RMSE |
| Kajitani et al. [13] | Individual | ANN | Canadian lynx data | 1821–1934 | RMSE |
| Neely et al. [19] | Individual | GP | Exchange rates (various pairs) | 1974–1995 | Excess return |
| Nikolsko-Rzhevskyy and Prodan [33] | Individual | Markov switching | Exchange rates (various pairs) | 1983–2008 | MSE |
| Wong et al. [24] | Individual | Novel ANN | Simulated time series data | – | MAPE, NMSE |
| Yu et al. [16] | Individual | Least squares SVM | Crude oil | 2008–2015 | MAPE |
| Yu et al. [25] | Individual | Artificial intelligence | Global oil consumption | 2004–2015 | PCC, AUC, RMSE, MAPE |
| Zhang [10] | Individual | ANN | Simulated time series data | — | MSE, MAPE |
| Zhao and Yang [23] | Individual | CRPSO-based neuron model | Simulated time series data, electroencephalogram data | — | MSE |
| Cao [15] | Hybrid | SVM + SOM | Sunspot data, Santa Fe datasets A, C, and D, and the two building datasets | — | NMSE, RMSE, CV |
| Chen et al. [28] | Hybrid | EMD + ANN | Tourism demand | 1971–2009 | MAPE, RMSE, MAE |
| Khashei et al. [7] | Hybrid | ARIMA + ANNs + fuzzy | Exchange rates (USD/Iran rials), gold price | 2005–2006 | MAE, MSE |
| Lin et al. [29] | Hybrid | EMD + SVM | Exchange rates (various pairs) | 2005–2009 | MAPE, RMSE, MAE, DS, CP, CD |
| Tang et al. [30] | Hybrid | EEMD + RVFL | Crude oil | 1986–2010 | MAPE, RMSE, DS |
| Yu et al. [12] | Hybrid | ANN + WD | Patient visits | 2010–2015 | RMSE, MAPE |
| Yu et al. [27] | Hybrid | EMD + ANN | Crude oil | 1986–2006 | NMSE, DS |
| Yu et al. [17] | Hybrid | OPL + SVMQR | Ten publicly available datasets | — | Empirical risk, quantile property |
| Yu et al. [14] | Hybrid | Sparse representation + ANN | Crude oil | 1986–2013 | RMSE MAPE |
| Zhang [21] | Hybrid | ARIMA + ANN | Wolf's sunspot data, Canadian lynx data, and GBP/USD | — | MSE, MAE |

ANN: artificial neural network; SVM: support vector machine; GP: genetic programming; CRPSO: cooperative random learning particle swarm optimization; RVFL: random vector functional link; OPL: orthogonal pinball loss; SVMQR: SVM quantile regression; WD: wavelet decomposition; SOM: self-organization feature map; RMSE: root mean squared error; NMSE: normalized mean squared error; MAE: mean absolute error; MAPE: mean absolute percentage error; CV: coefficient of variation; PCC: percentage correctly classified accuracy; AUC: area under the receiver operating curve; SR: success ratio; DS: directional symmetry; CP: correct uptrend; CD: correct downtrend.

all IMFs and one residual are used. According to the empirical evidence in this study, both types of hybrid models are superior to the pure MLP model, whether with 1-, 5-, 10-, 20-, or 30-day forecasting horizons. Moreover, the application of trading strategies is proposed. Although the transaction costs can make the profits practically disappear or become negative, as proved by Álvarez Díaz [31], this empirical analysis shows that, even considering a 0.3% transaction cost in each trade, the trading strategy based on EMD-MLP, on average, produces an annualized profit exceeding 10%.

In this paper, we will investigate RMB exchange rate forecasting and develop the relevant trading strategies based on the constructed models. The main contributions come from three aspects. First, this study focuses on the RMB, including the onshore RMB exchange rate (CNY) and offshore RMB exchange rate (CNH). We will use daily data to forecast RMB exchange rates with different horizons based on three types of models, i.e., MLP, EMD-MLP, and EMD-MLP*. Second, we will consider not only the MLP model but also the hybrid EMD-MLP and EMD-MLP* models to improve forecasting performance. It should be noted that the EMD-MLP* is different from the methodology proposed by Yu et al. [27]. We regard some IMF components as noise factors and delete them to reduce the volatility of the RMB. Finally, we will choose the best forecasting models to construct the trading strategies by introducing different critical numbers and considering different transaction costs.

The remainder of the paper is organized in the following manner. Section 2 provides a brief description of ANN and EMD. The overall forecasting process and model notations of three kinds of forecasting models are also included in this part. In Section 3, two currency exchange rates of USD to RMB, CNY, and CNH, are used to test the effectiveness of the proposed methodology, and the selected models are applied in trading strategies. The conclusions drawn from this study are presented in Section 4.

## 2. Methodology

*2.1. Artificial Neural Network (ANN).* This study considers MLP based on an error backpropagation algorithm. Figure 1 shows a simple MLP structure with three input nodes ($X1$, $X2$, and $X3$). One hidden layer consists of four hidden nodes and one output node ($Y$). The nodes are organized in layers and are usually fully connected by weights, which indicate the effects of the corresponding nodes. In each node of the hidden and output layers, all data are firstly processed by the integration function (also called the summation function), which combines all incoming signals, and secondly processed by the activation function (also called the transfer function), which transforms the output of the node. In general, the amount of the hidden layer is less than 3, due to the converging restriction.

Particularly, the MLP model is fitted by the training data, and then the testing data and an out-of-sample dataset are used to verify its forecasting performance. Through supervised learning algorithms, the parameters (weights and node intercepts) are adjusted iteratively by a process of minimizing the forecasting error function. Formally, an MLP with an input layer with $n$ nodes, one hidden layer consisting of $J$ hidden nodes, and an output layer with one output node calculates the following function:

$$\begin{aligned} o(x) &= f\left(w_0 + \sum_{j=1}^{J} w_j \cdot f\left(w_{0j} + \sum_{i=1}^{n} w_{ij} x_i\right)\right) \\ &= f\left(w_0 + \sum_{j=1}^{J} w_j \cdot f\left(w_{0j} + \mathbf{w}_j^T \mathbf{x}\right)\right), \end{aligned} \quad (1)$$

where $\mathbf{x} = (x_1, \ldots, x_n)$ is the vector of all input variables, $w_0$ is the intercept of the output node, $w_{0j}$ is the intercept of the $j$th hidden node, $w_j$ denotes the weight corresponding to the node starting at the $j$th hidden node to the output node, and $w_{ij}$ denotes the weight corresponding to the node starting at $i$th input node to the $j$th hidden node. Therefore, all hidden and output nodes calculate the function $f(g(\mathbf{z}))$, where $g(\cdot)$ denotes the integration function, which is defined as $g(z) = w_0 + \mathbf{w}^t \mathbf{z}$, and $f(\cdot)$ denotes the activation function, which is often a bounded nondecreasing nonlinear and differentiable function. In this study, the logistic function ($f(u) = 1/(1 + e^{-u})$) is used as the activation function.

Given inputs $\mathbf{x}$ and the current weights, which are initialized with random values from a standard normal distribution, the MLP produces an output $o(\mathbf{x})$. Then, an error function is defined. This study selects the mean square error as follows:

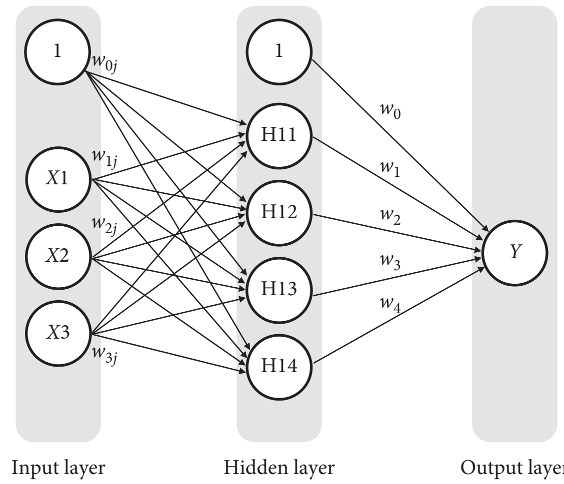


Figure 1: An artificial neural network structure with three input neurons ($X1$, $X2$, and $X3$), one hidden layer consisting of four hidden neurons (H11, H12, H13, and H14), and one output neuron ($Y$).

$$E = \frac{1}{N} \sum_{h=1}^{N} \left(o_h(\mathbf{x}) - y_h\right)^2, \quad (2)$$

where $N$ is the number of data samples and $y_h$ is the observed output. During the iterative training process, the above steps are repeated to adapt all weights until a prespecified criterion is fulfilled. In order to find a local minimum of the error function, the resilient backpropagation algorithm modifies the weights in the opposite direction of partial derivatives. According to Riedmiller and Braun [34], the weights are adjusted by the following rule:

$$w_k^{(t+1)} = w_k^t - \eta_k^{(t)} \cdot \operatorname{sign}\left(\frac{\partial E^{(t)}}{\partial w_k^{(t)}}\right), \quad (3)$$

where $t$ and $k$ index the iteration steps and the weights, respectively. In order to speed up convergence, the learning rate $\eta_k^{(t)}$ increases if the corresponding partial derivative keeps the same sign; otherwise, it will decrease.

*2.2. Empirical Mode Decomposition (EMD).* In this study, we further apply EMD to decompose the time series into several IMFs and remaining residues. These IMFs usually satisfy two conditions: the first is that the number of extrema and zero crossings must be equal or different by not more than one. Second, the mean value of the envelopes, which include both local maxima and minima, must be zero at all points.

Given the time series data $x(t), t = 1, 2, \ldots, T$, Huang et al. [26] propose a sifting process to decompose $x(t)$. The first step is to identify all local maxima and local minima of $x(t)$. Then, the upper and lower envelopes are defined by connecting all local extrema by a spline line. Next, for all points at the envelope, the mean value $m_1^1(t)$ from upper and lower envelopes is calculated. Then, it follows computation of the first IMF of $x(t)$:

$$h_1^1(t) = x(t) - m_1^1(t). \quad (4)$$

If $h_1^1(t)$ does not meet the above two conditions, this study then takes $h_1^1(t)$ as a new data series and repeats procedure (4). Thus, we calculate

$$h_2^1(t) = h_1^1(t) - m_2^1(t). \quad (5)$$

In this calculation, $m_2^1(t)$ is the mean value of the upper and lower envelopes of $h_1^1(t)$.

Repeating the same procedure until meeting both conditions, we get the first IMF component of $x(t)$, $c_1(t)$, that is

$$c_1(t) = h_q^1(t). \quad (6)$$

The stopping rule indicates that absolute values of the envelope mean must be less than the user-specified tolerance level. In this study, the tolerance level is denoted as the standard deviation of $x(t)$ times 0.01. Other interesting stopping rules can be found in the works of Huang et al. [26] and Huang and Wu [35].

After extracting the component $c_1(t)$ from $x(t)$, we denote another series $r_1(t) = x(t) - c_1(t)$, which contains all information except $c_1(t)$. Huang et al. [36] suggest a sifting stop criterion that is $r_n(t)$ becomes a monotonic function or cannot extract more IMFs. Finally, the time series $x(t)$ can be expressed as

$$x(t) = \sum_{i=1}^{n} c_i(t) + r_n(t), \quad (7)$$

where $n$ is the number of IMFs and $r_n(t)$ is the final residue, which represents the central tendency of the data series $x(t)$. These IMFs are nearly orthogonal to each other and all have means of nearly zero. According to the above properties, it is possible to forecast all decompositions and summarize these estimations to predict $x(t)$.

*2.3. Overall Forecasting Process and Model Notations.* Considering the time series $p_t$, $t = 1, 2, \ldots, T$, we would like to predict $l$-day ahead, which is denoted as $\hat{p}_{t+l}$. In this study, the input variables (past observations) include $p_{t-59}$, $p_{t-19}$, $p_{t-9}$, and $p_{t-4}$ to $p_t$, which represent the past price levels of 60 days, 20 days, 10 days, and the last 5 days, respectively. The output is the $l$-day ahead prediction. Formally, it could be shown as follows:

$$\hat{p}_{t+l} = \varphi(p_t, \ldots, p_{t-4}, p_{t-9}, p_{t-19}, p_{t-59}, \mathbf{w}), \quad (8)$$

where $\varphi(\cdot)$ is a function determined by neural network training and $\mathbf{w}$ is a weight vector of all parameters of MLP.

Three kinds of forecasting models are adopted in this study. The first is the pure MLP model with one and two hidden layers. The second is an EMD-MLP model, which uses EMD to subtract some volatile IMFs from the original data series and then uses the new data series to calculate the final prediction by MLP technology. Figure 2 indicates the procedure of the EMD(-1)-MLP model, which means the first IMF component is ignored. The third kind of model is named EMD-MLP*, which generally consists of the following four steps:

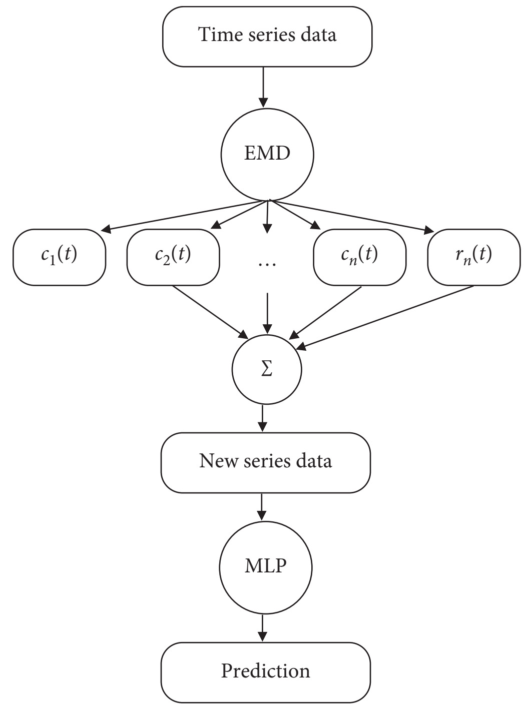


Figure 2: An example of the EMD(-1)-MLP model. We decompose a time series data by the EMD and generate $n$ IMF components, $c_1(t), c_2(t), \ldots, c_n(t)$, and one residue $r_n(t)$. We sum all decompositions except $c_1(t)$ to produce a new time series data. Then, the MLP is applied to compute the prediction.

(1) Decompose the original time series into IMF components and one residual component via EMD
(2) Determine how many IMF components are used, which depends on the length of the forecasting period
(3) For each chosen IMF and residual component, the MLP model is used as a forecasting tool to model these components and to make the corresponding prediction
(4) Add all prediction results to one value, which can be seen as the final prediction result for the original time series.

As an example, Figure 3 represents the above procedure of the EMD(-1)-MLP* model. Time series data are decomposed via EMD and generates $n$ IMF components, $c_1(t), c_2(t), \ldots, c_n(t)$, and one residue $r_n(t)$. In addition to $c_1(t)$, we forecast each decomposition and then sum them as the prediction results for the original time series.

## 3. Empirical Experiments

*3.1. Data.* Two kinds of currency exchange rates of USD to RMB are considered in this study. If the RMB is traded onshore (in mainland China), it is referred to as CNY, and if traded offshore (mainly in Hong Kong), it is named CNH.

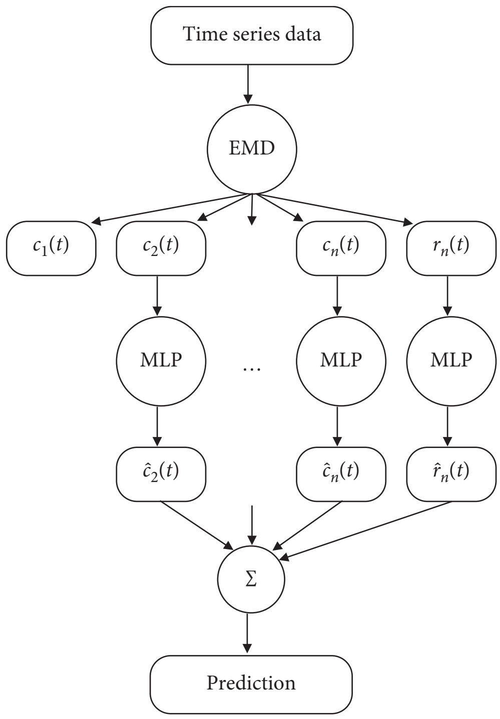


FIGURE 3: An example of the EMD(-1)-MLP* model. We decompose a time series data by the EMD and generate $n$ IMF components, $c_1(t), c_2(t), \ldots, c_n(t)$, and one residue $r_n(t)$. Besides $c_1(t)$, we forecast each decomposition and them sum them as the prediction.

Both time series data are downloaded from Bloomberg. For CNY, daily data from January 2, 2006, to December 21, 2015, with a total of 2584 observations are examined in this study. Since the CNH officially commenced on July 19, 2010, its sample period is from January 3, 2011, to December 21, 2015, with a total of 1304 observations. For training the neural network models, two-thirds of the observations are randomly assigned to the training dataset and the remainder is used as the testing dataset. In addition, it should be noted that the time series in price levels is used in the following analysis, rather than in return levels.

According to descriptions shown in Section 3.2, the EMD technique is used to decompose both CNY and CNH price series into several independent IMFs and one residual component. The decomposed results of CNY and CNH are represented in Figures 4 and 5. Comparing these two figures, the original data series and decompositions seem dissimilar, which is due to the different lengths of the sample periods. In fact, their IMFs have some similar characteristics. First, focusing on the sample period 2011–2015, their residual components have the same trend. Starting at about 6.5, they drop slightly to 6.2 and then rise slightly to 6.5. Second, the swing period of high-frequency IMFs is shorter and also the mean of absolute values of high-frequency IMFs is smaller when compared with low frequency. For example, in the series of CNY, the absolute values of $c_1$, $c_3$, and $c_5$ are 0.00298, 0.00596, and 0.02002, respectively. In our estimation model, it is not necessary to include all the high-frequency IMFs, as considering high-frequency IMFs may reduce the forecasting accuracy when the forecasting period is long. Therefore, in the model EMD-MLP, the high-frequency IMF is removed to get the denoised time series. In the model EMD-MLP*, not only are all the decompositions chosen to forecast the exchange rate, but consideration is also given to using part of the decompositions to conduct forecasting, discarding the high-frequency IMF series.

*3.2. Experimental Results.* Two main criteria are considered in this empirical experiment: the normalized mean squared error (NMSE), and the directional statistic ($D_{\text{stat}}$), to evaluate the levels of prediction and directional forecasting, respectively. Typically, following [37], the NMSE is defined by

$$\text{NMSE} = \frac{1}{\sigma_\Psi^2} \frac{1}{N_t} \sum_{s \in \Psi} \left(\widehat{p}_{s+l} - p_{s+l}\right)^2, \tag{9}$$

where $\Psi$ refers to the test dataset containing $N_t$ observations, $\widehat{p}_{s+l}$ is the $l$-day ahead prediction, $p_{s+l}$ is the actual value, and $\sigma_\Psi^2$ is the variance of $p_{s+l}$. Clearly, the NMSE is one of the most important criteria for measuring the validity of the forecasting model, but from a business perspective, improving the accuracy of directional predictions can support decision-making, so as to generate greater profits. Furthermore, the Diebold-Mariano (DM) test [38] is adopted to investigate whether adding EMD could improve the forecasting performance of the MLP model. Specifically, the null hypothesis is that the EMD-MLP (or EMD-MLP*) model has a lower forecast accuracy than the corresponding MLP model. For example, in the case of the EMD(-1)-MLP(5,3), the corresponding model infers the MLP(5,3) model.

Besides the NMSE evaluation, the directional statistic ($D_{\text{stat}}$) is applied to measure the ability to predict movement direction [27, 39], which can be expressed as

$$D_{\text{stat}} = \frac{1}{N_t} \sum_{s \in \Psi} a_s \times 100\%, \tag{10}$$

where $a_s = 1$ if $(\widehat{p}_{s+l} - p_s)(p_{s+l} - p_s) \geq 0$ and $a_s = 0$ otherwise. Pesaran and Timmermann [40] provide the directional accuracy (DAC) test to examine predictive performance. In this case, the DAC test is used to determine whether $D_{\text{stat}}$ is significantly larger than 0.5. We have also adopted the excess profitability test proposed by Anatolyev and Gerko [41] in every related empirical study. Since both tests produce the same outcome, we only report the results of DAC tests in the following empirical studies.

Table 2 compares the forecasting performance, in terms of the NMSE, for the MLP, EMD-MLP, and EMD-MLP* models. The NMSE is reported as the percentage for $l$-day ahead predictions where $l = 1, 5, 10, 20$, and 30. For each prediction period, the minimum NMSE is designated in bold font. Panel A of Table 2 displays the experimental results of the MLP models. It is clear that, with the increase in forecasting period, the NMSE becomes larger, suggesting reduced forecasting accuracy for a longer forecasting period. Also, the more complex MLP models with two hidden layers

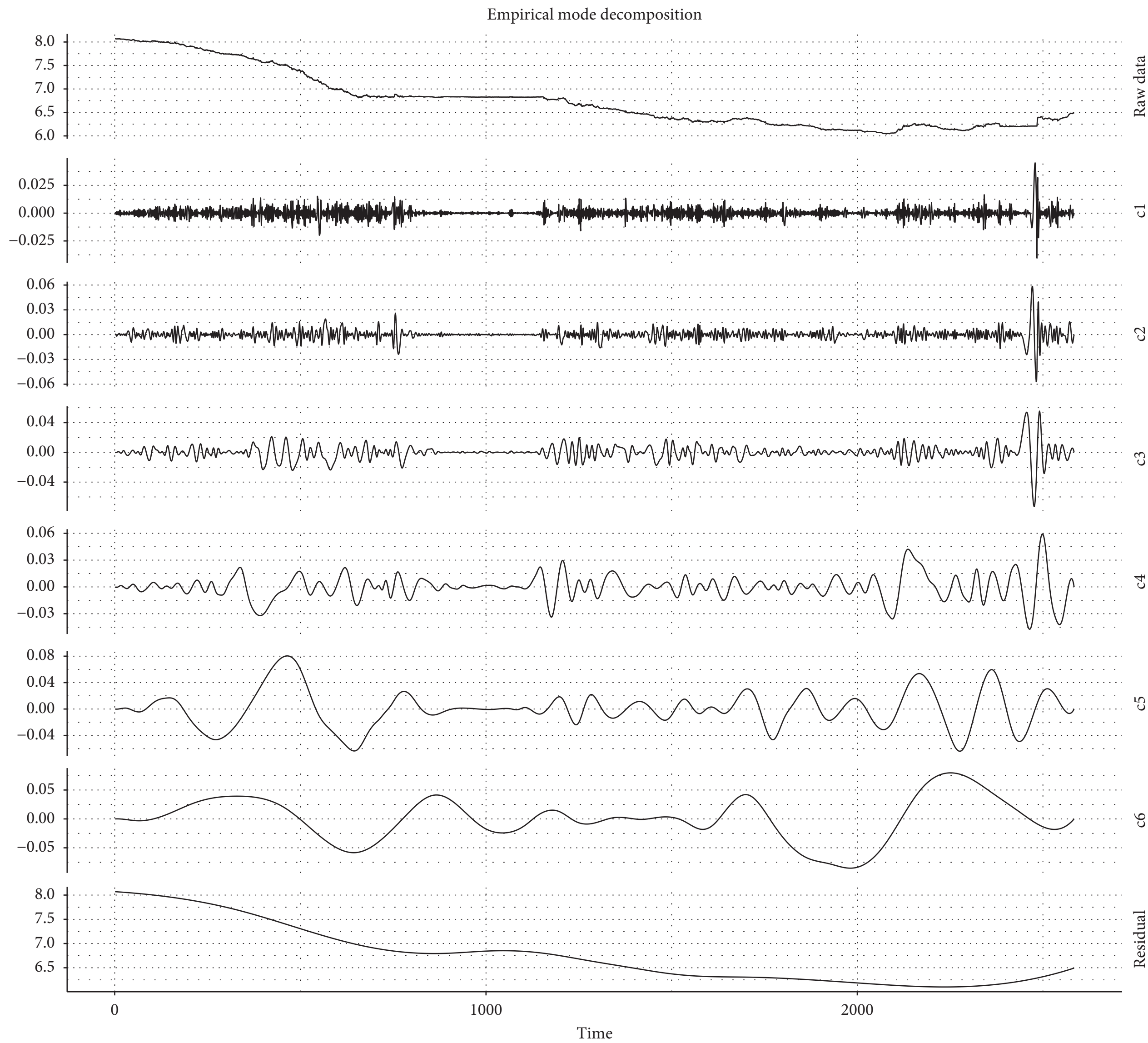


FIGURE 4: EMD decomposition of CNY from 2006 to 2015 is shown in this figure including the raw data series, 6 IMFs, and one residual.

do not offer better forecasting accuracy and are even worse for 1-day ahead prediction.

The empirical results of EMD-MLP are discussed in Panel B of Table 2. EMD(-1)-MLP treats the highest frequency IMF $c_1$ as noise and applies the MLP model after subtracting the $c_1$ series from the original time series. Empirical results show that the NMSE of 1-day ahead forecasting dropped dramatically; however, improvement of 5-day ahead forecasting with EMD(-1)-MLP models is limited. EMD(-2)-MLP models, which deduct the two highest frequency IMF series, perform better than EMD(-1)-MLP models in the 5-day ahead forecasting experiment. In the same way, the EMD(-3)-MLP models perform better when the forecasting period is longer than 10 days. The results of EMD-MLP* models are shown in Panel C of Table 2. Although the calculation process is more complicated, the forecasting performance of EMD-MLP* is always better than MLP or EMD-MLP models, according to the criterion of NMSE. In addition, the random walk models with and without drifts (the random walk model without drifts predicts that all future values will equal the last observed value. Given a variable $Y$, the $k$-step-ahead forecast from period $t$ of $Y$ is $\widehat{Y}_{t+k} = Y_t$. For the random walk model with drifts, the $k$-step-ahead forecast from period $t$ of $Y$ is $\widehat{Y}_{t+k} = Y_t + k\widehat{d}$, where $\widehat{d}$ is estimated by the average change of $Y_t$ in the past 60 days) are used to examine the same data. The related results in Panel D of Table 2 indicate that, for the MLP-EMD and MLP-EMD* models, their performance is obviously better than the random walk models in most situations.

Table 3 is based on the forecasting performance of the above models with $D_{\mathrm{stat}}$, which shows similar experimental results to the NMSE criterion. The table shows that if EMD(-2)-MLP(5,3)* is applied to forecasting the direction of the CNY series 30 days later, the hitting rate can be as high as 86.39%. In general, although the calculation of EMD-MLP*

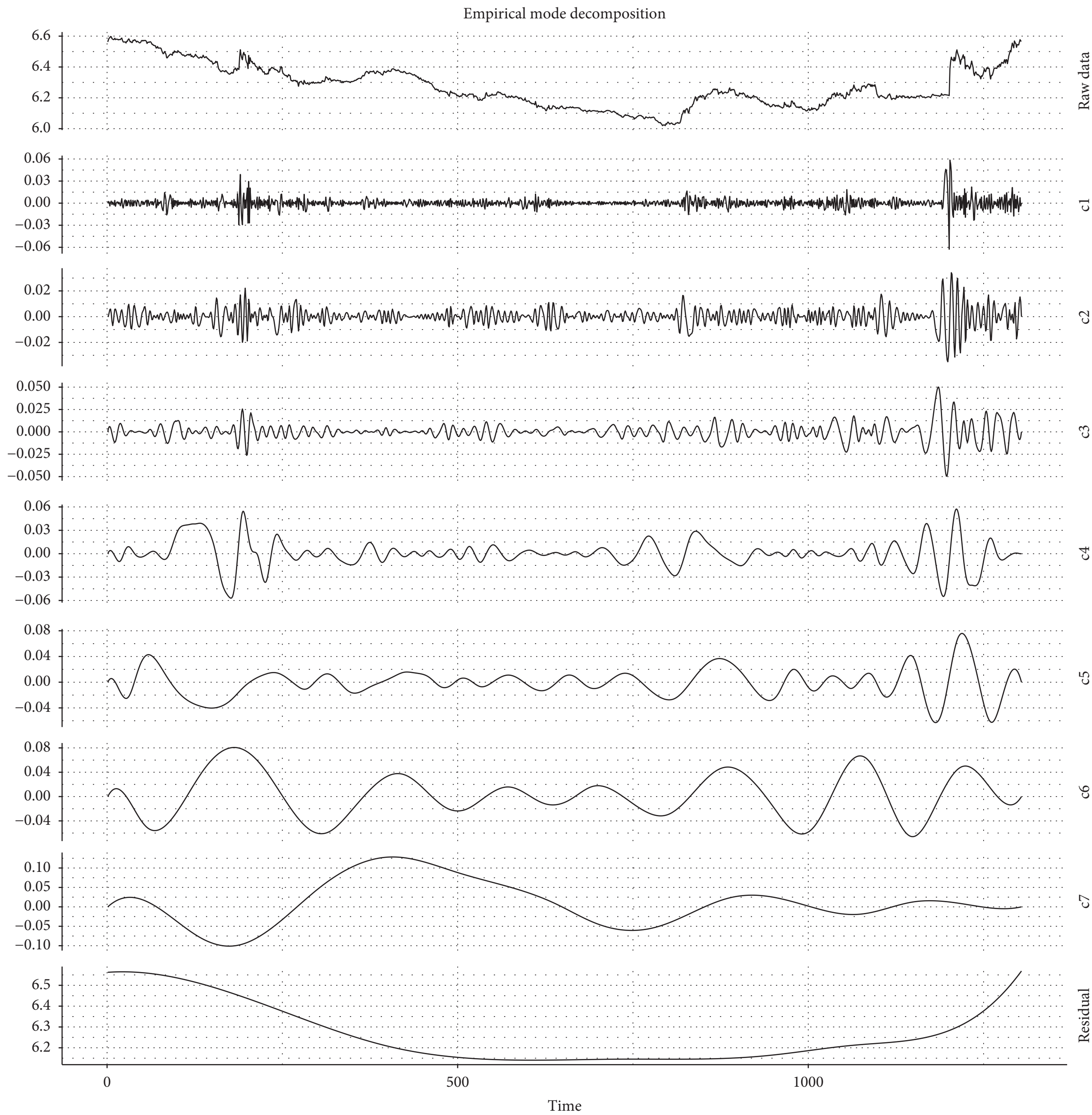


Figure 5: EMD decomposition of CNH from 2011 to 2015 is shown in this figure including the raw data series, 7 IMFs, and one residual.

models is more complicated, the forecasting ability appears to be the best of all the proposed models. The highest frequency IMF is therefore treated as noise and EMD(-1)-MLP* is used for the forecast. In this way, the performance of the selected model would be relatively better, based on both NMSE and $D_{\text{stat}}$ criteria. Besides, it should be noted that performance of predicting over longer periods is worse than shorter ones intuitively. In terms of NMSE, the experimental results are consistent with above argument, i.e., the NMSE performance of predicting 20 and 30-days ahead is poor. But the $D_{\text{stat}}$ performance is totally different; our results show that predictions over longer periods often have higher $D_{\text{stat}}$ (in order to verify the robustness, we also used three-fourths of the observations as a training dataset to reexamine the results of Tables 2 and 3. The related outcomes are shown in Tables 4 and 5. In addition, we considered a hyperbolic tangent function as the activation function and reported related results in Tables 6 and 7).

In addition to the above experiments, the dataset of CNY and CNH series from 2011 to 2015 is considered. We use the same method of dividing the data into the training set and testing set. According to the experimental results in Tables 2 and 3, we do not report MLP models and only adopt a number of EMD-MLP and EMD-MLP* models to conduct predictions. The statistical results of NMSE and $D_{\text{stat}}$ are shown in Tables 8 and 9, respectively.

Table 2: NMSE comparisons for different models.

| | 1-day | 5-day | 10-day | 20-day | 30-day |
|---|---|---|---|---|---|
| Panel A: MLP model | | | | | |
| MLP(3) | 0.0231 | 0.0874 | 0.2118 | 0.3698 | 0.5888 |
| MLP(5) | 0.0204 | 0.0854 | 0.2088 | 0.3939 | 0.5723 |
| MLP(5,3) | 0.0285 | 0.0943 | 0.2088 | 0.3623 | 0.4746 |
| MLP(6,4) | 0.0376 | 0.0896 | 0.2021 | 0.3613 | 0.5453 |
| Panel B: EMD-MLP model | | | | | |
| EMD(-1)-MLP(3) | 0.0145*** | 0.0823 | 0.2120 | 0.3894 | 0.5217* |
| EMD(-1)-MLP(5,3) | 0.0143*** | 0.0821** | 0.2111 | 0.3836 | 0.5447 |
| EMD(-2)-MLP(3) | 0.0207 | 0.0450*** | 0.1613** | 0.3760 | 0.5258 |
| EMD(-2)-MLP(5,3) | 0.0236** | 0.0505*** | 0.1583** | 0.3667 | 0.5258 |
| EMD(-3)-MLP(3) | 0.0434 | 0.0436*** | 0.0739** | 0.2162** | 0.4355*** |
| EMD(-3)-MLP(5,3) | 0.0419 | 0.0456*** | 0.0694** | 0.2144** | 0.3892** |
| Panel C: EMD-MLP* model | | | | | |
| EMD(0)-MLP(3)* | 0.0087*** | 0.0217*** | 0.0410*** | 0.0806*** | 0.1244*** |
| EMD(0)-MLP(5,3)* | 0.0088*** | 0.0222*** | 0.0404*** | 0.0695*** | 0.1048*** |
| EMD(-1)-MLP(3)* | **0.0075*** | **0.0200*** | 0.0385*** | 0.0812*** | 0.1136*** |
| EMD(-1)-MLP(5,3)* | 0.0083*** | 0.0314*** | **0.0368*** | **0.0681*** | 0.1051*** |
| EMD(-2)-MLP(3)* | 0.0172* | 0.0214*** | 0.0444*** | 0.0760*** | 0.1152*** |
| EMD(-2)-MLP(5,3)* | 0.0190** | 0.0256*** | 0.0423*** | 0.0759*** | **0.1034*** |
| Panel D: random walk model | | | | | |
| No drift | 0.0144 | 0.0861 | 0.2349 | 0.5323 | 0.9357 |
| With drift | 0.0148 | 0.0973 | 0.2834 | 0.6981 | 1.2954 |

Consider the CNY from January 2, 2006, to December 21, 2015, with a total of 2584 observations. This table compares the forecasting performance, in terms of the NMSE, for the MLP, EMD-MLP, and EMD-MLP* models. We report the NMSE as percentage for $l$-day ahead predictions where $l = 1, 5, 10, 20$, and 30. The DM test [38] is used to compare the forecast accuracy of EMD-MLP (EMD-MLP*) model and the corresponding MLP model. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. For each length of prediction, we mark the minimum NMSE as bold.

Table 3: $D_{stat}$ comparisons for different models.

| | 1-day | 5-day | 10-day | 20-day | 30-day |
|---|---|---|---|---|---|
| Panel A: MLP model | | | | | |
| MLP(3) | 51.74 | 53.49 | 63.49*** | 69.65*** | 70.72*** |
| MLP(5) | 51.98 | 55.17 | 62.17*** | 68.44*** | 71.08*** |
| MLP(5,3) | 50.18 | 54.93 | 65.78*** | 69.29*** | 73.39*** |
| MLP(6,4) | 48.98 | 55.41 | 62.65*** | 69.89*** | 71.20*** |
| Panel B: EMD-MLP model | | | | | |
| EMD(-1)-MLP(3) | 61.94*** | 59.74*** | 66.99*** | 69.04*** | 72.54*** |
| EMD(-1)-MLP(5,3) | 62.42*** | 57.69*** | 65.54*** | 68.92*** | 71.32*** |
| EMD(-2)-MLP(3) | 62.30*** | 69.47*** | 69.64*** | 69.41*** | 72.42*** |
| EMD(-2)-MLP(5,3) | 60.02*** | 66.71*** | 70.24*** | 70.25*** | 72.54*** |
| EMD(-3)-MLP(3) | 57.50*** | 72.96*** | 78.67*** | 75.33*** | 73.63*** |
| EMD(-3)-MLP(5,3) | 56.78*** | 69.59*** | 80.48*** | 74.37*** | 73.27*** |
| Panel C: EMD-MLP* model | | | | | |
| EMD(0)-MLP(3)* | 69.99*** | 80.65*** | 81.45*** | 82.59*** | 86.39*** |
| EMD(0)-MLP(5,3)* | 68.31*** | 79.45*** | 80.36*** | 82.95*** | 85.78*** |
| EMD(-1)-MLP(3)* | **71.67*** | **81.01*** | 81.08*** | 81.86*** | 84.08*** |
| EMD(-1)-MLP(5,3)* | 70.35*** | 80.17*** | **82.89*** | **82.95*** | 86.39*** |
| EMD(-2)-MLP(3)* | 63.51*** | 80.89*** | 80.72*** | 82.10*** | 86.15*** |
| EMD(-2)-MLP(5,3)* | 63.51*** | 78.73*** | 82.41*** | 81.98*** | **86.39*** |

Consider CNY from January 2, 2006, to December 21, 2015, with a total of 2584 observations. This table compares the forecasting performance, in terms of the $D_{stat}$, for the MLP, EMD-MLP, and EMD-MLP* models. We report $D_{stat}$ as percentage for $l$-day ahead predictions where $l = 1, 5, 10, 20$, and 30. Moreover, we examine the ability of all models to predict the direction of change by the DAC test [40]. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. For each length of prediction, we mark the maximum $D_{stat}$ as bold.

Tables 8 and 9 indicate the EMD-MLP* models perform better than the EMD-MLP models on both NMSE and $D_{stat}$ criteria. Comparing the results of the CNY and CNH series, the NMSE of CNY is found to be less than that of CNH on average, no matter how long the forecasting period. The main reason for this is the original volatility of CNY series is smaller than that of CNH series. For the $D_{stat}$ test part, forecasting performance improves with the growth of

Table 4: Robust tests for NMSE comparisons with different subsamples.

| | 1-day | 5-day | 10-day | 20-day | 30-day |
|---|---|---|---|---|---|
| Panel A: MLP model | | | | | |
| MLP(3) | 0.0175 | 0.0740 | 0.2010 | 0.3638 | 0.5954 |
| MLP(5) | 0.0201 | 0.0699 | 0.2082 | 0.3663 | 0.5158 |
| MLP(5,3) | 0.0212 | 0.0761 | 0.2117 | 0.3614 | 0.5561 |
| MLP(6,4) | 0.0221 | 0.0694 | 0.2066 | 0.3586 | 0.5746 |
| Panel B: EMD-MLP model | | | | | |
| EMD(-1)-MLP(3) | 0.0137*** | 0.0705 | 0.1444** | 0.3827 | 0.5885 |
| EMD(-1)-MLP(5,3) | 0.0117*** | 0.0713 | 0.1982*** | 0.3251** | 0.4948*** |
| EMD(-2)-MLP(3) | 0.0192 | 0.0371*** | 0.1434** | 0.3568 | 0.5725* |
| EMD(-2)-MLP(5,3) | 0.0222 | 0.0380*** | 0.1433** | 0.3380* | 0.5417 |
| EMD(-3)-MLP(3) | 0.0423 | 0.0468*** | 0.0691*** | 0.2003** | 0.4676** |
| EMD(-3)-MLP(5,3) | 0.0442 | 0.0410*** | 0.0647*** | 0.1965** | 0.4551** |
| Panel C: EMD-MLP* model | | | | | |
| EMD(0)-MLP(3)* | **0.0079***** | 0.0201*** | 0.0391*** | 0.0798*** | 0.1302*** |
| EMD(0)-MLP(5,3)* | 0.0084*** | 0.0213*** | 0.0461*** | **0.0614***** | 0.1160*** |
| EMD(-1)-MLP(3)* | 0.0083*** | **0.0192***** | 0.0413*** | 0.0717*** | 0.1298*** |
| EMD(-1)-MLP(5,3)* | 0.0091*** | 0.0220*** | 0.0425*** | 0.0655*** | 0.1140*** |
| EMD(-2)-MLP(3)* | 0.0191 | 0.0222*** | 0.0456*** | 0.0801*** | 0.1294*** |
| EMD(-2)-MLP(5,3)* | 0.0201 | 0.0240*** | **0.0381***** | 0.0660*** | **0.1050***** |

Consider the CNY from January 2, 2006, to December 21, 2015, with a total of 2584 observations. For training the neural network models, three-fourths of the observations are randomly assigned to the training dataset and the remainder is used as the testing dataset. This table compares the forecasting performance, in terms of the NMSE, for the MLP, EMD-MLP, and EMD-MLP* models. We report the NMSE as percentage for $l$-day ahead predictions where $l = 1, 5, 10, 20$, and 30. The DM test [38] is used to compare the forecast accuracy of EMD-MLP (EMD-MLP*) model and the corresponding MLP model. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. For each length of prediction, we mark the minimum NMSE as bold.

Table 5: Robust tests for $D_{\text{stat}}$ comparisons with different subsamples.

| | 1-day | 5-day | 10-day | 20-day | 30-day |
|---|---|---|---|---|---|
| Panel A: MLP model | | | | | |
| MLP(3) | 49.76 | 56.98 | 63.43*** | 70.13*** | 71.15*** |
| MLP(5) | 45.96 | 56.98 | 66.61*** | 69.81*** | 72.12*** |
| MLP(5,3) | 53.88** | 55.56 | 64.07*** | 69.49*** | 70.67*** |
| MLP(6,4) | 49.92 | 59.68** | 63.12*** | 68.85*** | 71.47*** |
| Panel B: EMD-MLP model | | | | | |
| EMD(-1)-MLP(3) | 62.12*** | 57.14* | 71.54*** | 69.17*** | 71.79*** |
| EMD(-1)-MLP(5,3) | 62.76*** | 56.98 | 67.25*** | 70.93*** | 74.04*** |
| EMD(-2)-MLP(3) | 64.03*** | 71.59*** | 72.18*** | 68.05*** | 70.99*** |
| EMD(-2)-MLP(5,3) | 62.92*** | 72.70*** | 72.02*** | 69.97*** | 72.12*** |
| EMD(-3)-MLP(3) | 55.94*** | 70.16*** | 79.01*** | 75.88*** | 74.36*** |
| EMD(-3)-MLP(5,3) | 55.63*** | 75.24*** | 80.60*** | 76.04*** | 73.08*** |
| Panel C: EMD-MLP* model | | | | | |
| EMD(0)-MLP(3)* | 69.89*** | 79.37*** | 81.08*** | 80.35*** | 86.86*** |
| EMD(0)-MLP(5,3)* | **72.11***** | **82.54***** | 80.60*** | **82.43***** | 86.38*** |
| EMD(-1)-MLP(3)* | 67.83*** | 81.90*** | 81.40*** | 81.31*** | **87.50***** |
| EMD(-1)-MLP(5,3)* | 68.46*** | 81.59*** | 82.19*** | 80.67*** | 87.18*** |
| EMD(-2)-MLP(3)* | 61.81*** | 80.63*** | 81.24*** | 80.35*** | 87.18*** |
| EMD(-2)-MLP(5,3)* | 63.23*** | 80.32*** | **82.83***** | 81.47*** | 86.70*** |

Consider the CNY from January 2, 2006, to December 21, 2015, with a total of 2584 observations. For training the neural network models, three-fourths of the observations are randomly assigned to the training dataset and the remainder is used as the testing dataset. This table compares the forecasting performance, in terms of $D_{\text{stat}}$, for the MLP, EMD-MLP, and EMD-MLP* models. We report $D_{\text{stat}}$ as percentage for $l$-day ahead predictions where $l = 1, 5, 10, 20$, and 30. Moreover, we examine the ability of all models to predict the direction of change by the DAC test [40]. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. For each length of prediction, we mark the maximum $D_{\text{stat}}$ as bold.

forecasting periods. Also, no significant difference is found for the results of CNY and CNH series based on $D_{\text{stat}}$ criteria. For forecasting over 20 days ahead, the hitting rate exceeds 89%. Even for forecasting 1-day ahead, $D_{\text{stat}}$ can achieve more than 79%.

*3.3. Applying the Trading Strategy.* This part introduces an application for designing a trading strategy for CNY (since the cases of CNY and CNH, from 2011 to 2015, produce the similar results, we do not report the latter in this paper). According to the results in Tables 2 and 3, in terms of NMSE

Table 6: Robust tests for NMSE comparisons with different activation functions.

| | 1-day | 5-day | 10-day | 20-day | 30-day |
|---|---|---|---|---|---|
| Panel A: MLP model | | | | | |
| MLP(3) | 0.0281 | 0.0877 | 0.2123 | 0.3729 | 0.5579 |
| MLP(5) | 0.0245 | 0.0820 | 0.2075 | 0.3729 | 0.5149 |
| MLP(5,3) | 0.0260 | 0.0859 | 0.2079 | 0.3491 | 0.4450 |
| MLP(6,4) | 0.0222 | 0.0887 | 0.1996 | 0.3436 | 0.4733 |
| Panel B: EMD-MLP model | | | | | |
| EMD(-1)-MLP(3) | 0.0120*** | 0.0830 | 0.2045 | 0.3717 | 0.5285 |
| EMD(-1)-MLP(5,3) | 0.0137*** | 0.0862 | 0.2097 | 0.3662 | 0.5282 |
| EMD(-2)-MLP(3) | 0.0204** | 0.0354*** | 0.1466*** | 0.3654 | 0.5571 |
| EMD(-2)-MLP(5,3) | 0.0182** | 0.0351*** | 0.1636** | 0.3665 | 0.4772 |
| EMD(-3)-MLP(3) | 0.0417 | 0.0450*** | 0.0705*** | 0.2140** | 0.5782 |
| EMD(-3)-MLP(5,3) | 0.0393 | 0.0441*** | 0.0711*** | 0.2198** | 0.3996 |
| Panel C: EMD-MLP* model | | | | | |
| EMD(0)-MLP(3)* | **0.0077***** | 0.0226*** | 0.0491*** | 0.0820*** | 0.1246*** |
| EMD(0)-MLP(5,3)* | 0.0084*** | **0.0198***** | **0.0382***** | **0.0669***** | **0.1106***** |
| EMD(-1)-MLP(3)* | 0.0095*** | 0.0223*** | 0.0482*** | 0.0826*** | 0.1528*** |
| EMD(-1)-MLP(5,3)* | 0.0084*** | 0.0258*** | 0.0460*** | 0.0845*** | 0.1210*** |
| EMD(-2)-MLP(3)* | 0.0212** | 0.0239*** | 0.0469*** | 0.0892*** | 0.1316*** |
| EMD(-2)-MLP(5,3)* | 0.0196** | 0.0229*** | 0.0414*** | 0.0784*** | 0.1497*** |

Consider the CNY from January 2, 2006, to December 21, 2015, with a total of 2584 observations. In this table, a hyperbolic tangent function is used as the activation function. This table compares the forecasting performance, in terms of the NMSE, for the MLP, EMD-MLP, and EMD-MLP* models. We report the NMSE as percentage for $l$-day ahead predictions where $l = 1, 5, 10, 20$, and 30. The DM test [38] is used to compare the forecast accuracy of EMD-MLP (EMD-MLP*) model and the corresponding MLP model. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. For each length of prediction, we mark the minimum NMSE as bold.

Table 7: Robust tests for $D_{\text{stat}}$ comparisons with different activation functions.

| | 1-day | 5-day | 10-day | 20-day | 30-day |
|---|---|---|---|---|---|
| Panel A: MLP model | | | | | |
| MLP(3) | 50.06 | 55.17 | 62.41*** | 68.56*** | 71.81*** |
| MLP(5) | 48.98 | 56.85* | 63.25*** | 69.17*** | 72.54*** |
| MLP(5,3) | 48.14 | 54.81 | 65.30*** | 66.99*** | 72.17*** |
| MLP(6,4) | 49.94 | 56.25** | 61.45*** | 68.68*** | 72.42*** |
| Panel B: EMD-MLP model | | | | | |
| EMD(-1)-MLP(3) | 69.99*** | 58.41*** | 67.95*** | 70.01*** | 71.45*** |
| EMD(-1)-MLP(5,3) | 67.95*** | 57.81*** | 64.46*** | 70.25*** | 73.39*** |
| EMD(-2)-MLP(3) | 64.59*** | 73.20*** | 70.36*** | 70.37*** | 71.08*** |
| EMD(-2)-MLP(5,3) | 63.51*** | 75.84*** | 69.40*** | 69.89*** | 74.00*** |
| EMD(-3)-MLP(3) | 57.26*** | 72.48*** | 78.80*** | 75.21*** | 68.89*** |
| EMD(-3)-MLP(5,3) | 57.38*** | 71.88*** | 80.00*** | 73.40*** | 73.39*** |
| Panel C: EMD-MLP* model | | | | | |
| EMD(0)-MLP(3)* | 71.07*** | **78.97***** | 80.96*** | 82.71*** | 86.76*** |
| EMD(0)-MLP(5,3)* | **72.03***** | 78.85*** | 80.36*** | 82.10*** | **87.48***** |
| EMD(-1)-MLP(3)* | 68.19*** | 77.52*** | 80.72*** | 81.62*** | 84.81*** |
| EMD(-1)-MLP(5,3)* | 70.11*** | 77.28*** | 80.12*** | **83.07***** | 86.03*** |
| EMD(-2)-MLP(3)* | 61.82*** | 78.97*** | **81.08***** | 80.77*** | 85.18*** |
| EMD(-2)-MLP(5,3)* | 64.71*** | 78.85*** | 78.80*** | 82.59*** | 83.35*** |

Consider the CNY from January 2, 2006, to December 21, 2015, with a total of 2584 observations. In this table, a hyperbolic tangent function is used as the activation function. This table compares the forecasting performance, in terms of $D_{\text{stat}}$, for the MLP, EMD-MLP, and EMD-MLP* models. We report $D_{\text{stat}}$ as percentage for $l$-day ahead predictions where $l = 1, 5, 10, 20$, and 30. Moreover, we examine the ability of all models to predict the direction of change by the DAC test [40]. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. For each length of prediction, we mark the maximum $D_{\text{stat}}$ as bold.

and $D_{\text{stat}}$, the best model for predicting $l$-day ahead can be determined. The forecasting model can be trained based on past exchange rate series and produce $l$-day ahead predictions. According to the forecasting results, trading strategies are set as follows:

$$\begin{aligned} &\text{long: if } \widehat{p}_{s+l} > p_s \times (1+\tau), \\ &\text{short: if } \widehat{p}_{s+l} < p_s \times (1-\tau), \end{aligned} \tag{11}$$

where $\widehat{p}_{s+l}$ is the $l$-day ahead predictions at time $s$ and $\tau$ is a critical number. We long (short) the CNY if $l$-day ahead

Table 8: NMSE comparisons for CNY and CNH.

| | 1-day | 5-day | 10-day | 20-day | 30-day |
|---|---|---|---|---|---|
| Panel A: CNY | | | | | |
| EMD(-1)-MLP(5,3) | 0.2078* | 0.7330*** | 4.9711 | 11.4496 | 14.8922 |
| EMD(-2)-MLP(5,3) | 0.3417 | 0.6292*** | 2.7752** | 8.4836 | 8.5811 |
| EMD(-3)-MLP(5,3) | 0.6015 | 0.6098*** | 1.2525*** | 3.9844*** | 7.2250 |
| EMD(0)-MLP(3)* | 0.2171* | 0.4204*** | 0.6304*** | 1.6314*** | 2.1985*** |
| EMD(0)-MLP(5,3)* | 0.1711** | 0.4124*** | 0.6912*** | 1.5200*** | 1.6964*** |
| EMD(-1)-MLP(3)* | **0.1497**** | **0.3928***** | 0.7120*** | 1.8627*** | 2.1755*** |
| EMD(-1)-MLP(5,3)* | 0.1610* | 0.4774*** | **0.5999***** | 1.2579*** | 1.5669*** |
| EMD(-2)-MLP(3)* | 0.2847 | 0.4485*** | 0.8156*** | 1.3609*** | 1.9762*** |
| EMD(-2)-MLP(5,3)* | 0.2960 | 0.4846*** | 0.8578*** | **1.2355***** | **1.4642***** |
| Panel B: CNH | | | | | |
| EMD(-1)-MLP(5,3) | 0.3814*** | 2.2893 | 6.4050 | 12.8439 | 20.9652 |
| EMD(-2)-MLP(5,3) | 0.4732** | 1.1208** | 4.7997* | 12.2240 | 22.9451 |
| EMD(-3)-MLP(5,3) | 0.9658 | 0.9270*** | 2.3329*** | 7.6194** | 19.4264 |
| EMD(0)-MLP(3)* | **0.2317***** | 0.8170*** | 1.2729*** | 2.5680** | 3.5197*** |
| EMD(0)-MLP(5,3)* | 0.2600*** | 0.6797*** | 1.3611*** | 2.4448** | 3.2915** |
| EMD(-1)-MLP(3)* | 0.2699*** | 0.6471*** | 1.4162*** | **2.2991***** | 3.8647*** |
| EMD(-1)-MLP(5,3)* | 0.2379*** | 0.7165*** | 1.3509*** | 2.3486** | **3.0931**** |
| EMD(-2)-MLP(3)* | 0.4297* | **0.6346***** | 1.2586*** | 2.4497** | 4.4241*** |
| EMD(-2)-MLP(5,3)* | 0.4173** | 0.7528*** | **1.1915***** | 2.5860** | 3.1497** |

Consider both the CNY and CNH from January 3, 2011, to December 21, 2015, with a total of 1304 observations. This table compares the forecasting performance, in terms of the NMSE, for several forecasting models. We report the NMSE as percentage for $l$-day ahead predictions where $l = 1, 5, 10, 20$, and 30. The DM test [38] is used to compare the forecast accuracy of EMD-MLP (EMD-MLP*) model and the corresponding MLP model. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. For each length of prediction, we mark the minimum NMSE as bold.

Table 9: $D_{\text{stat}}$ comparisons for CNY and CNH.

| | 1-day | 5-day | 10-day | 20-day | 30-day |
|---|---|---|---|---|---|
| Panel A: CNY | | | | | |
| EMD(-1)-MLP(5,3) | 73.65*** | 75.00*** | 64.27*** | 60.90*** | 64.90*** |
| EMD(-2)-MLP(5,3) | 67.49*** | 77.97*** | 69.23*** | 68.67*** | 77.78*** |
| EMD(-3)-MLP(5,3) | 61.82*** | 78.71*** | 83.13*** | 77.44*** | 78.28*** |
| EMD(0)-MLP(3)* | 75.37*** | 79.70*** | 83.62*** | 86.22*** | 84.60*** |
| EMD(0)-MLP(5,3)* | **80.05***** | 82.43*** | 84.86*** | **89.47***** | **88.13***** |
| EMD(-1)-MLP(3)* | 75.12*** | 83.17*** | 84.37*** | 87.22*** | 84.09*** |
| EMD(-1)-MLP(5,3)* | 74.88*** | **83.42***** | **85.36***** | 88.72*** | 86.11*** |
| EMD(-2)-MLP(3)* | 68.47*** | 82.43*** | 82.88*** | 88.72*** | 83.59*** |
| EMD(-2)-MLP(5,3)* | 69.70*** | 82.67*** | 83.37*** | 86.97*** | 87.88*** |
| Panel B: CNH | | | | | |
| EMD(-1)-MLP(5,3) | 73.24*** | 62.10*** | 62.50*** | 62.38*** | 65.84*** |
| EMD(-2)-MLP(5,3) | 70.32*** | 76.28*** | 66.67*** | 63.37*** | 64.84*** |
| EMD(-3)-MLP(5,3) | 61.07*** | 79.46*** | 78.92*** | 75.00*** | 75.31*** |
| EMD(0)-MLP(3)* | **79.81***** | 82.64*** | 84.07*** | 84.65*** | 89.03*** |
| EMD(0)-MLP(5,3)* | 76.64*** | 83.13*** | 83.09*** | 85.40*** | **89.78***** |
| EMD(-1)-MLP(3)* | 76.64*** | 82.15*** | 84.80*** | 84.65*** | 84.79*** |
| EMD(-1)-MLP(5,3)* | 78.83*** | 83.37*** | **85.05***** | 84.65*** | 89.53*** |
| EMD(-2)-MLP(3)* | 71.53*** | **83.86***** | 84.31*** | **87.87***** | 85.29*** |
| EMD(-2)-MLP(5,3)* | 72.99*** | 84.60*** | 83.82*** | 83.42*** | 88.78*** |

Consider both the CNY and CNH from January 3, 2011, to December 21, 2015, with a total of 1304 observations. This table compares the forecasting performance, in terms of $D_{\text{stat}}$, for several forecasting models. We report $D_{\text{stat}}$ as percentage for $l$-day ahead predictions where $l = 1, 5, 10, 20$, and 30. Moreover, we examine the ability of all models to predict the direction of change by the DAC test [40]. ***, **, and * denote statistical significance at 1%, 5%, and 10%, respectively. For each length of prediction, we mark the maximum $D_{\text{stat}}$ as bold.

predictions are greater (smaller) than the present price multiplied by $1+\tau$. So, unless $\widehat{p}_{s+l} = p_s(1+\tau)$, people always conduct a transaction at time $s$ and hold it for $l$ days. For instance, consider the case with $l = 10$ and $\tau = 0$. If $p_s = 5$ and we use the EMD-MLP model to find $\widehat{p}_{s+l} = 5.2$, we immediately long CNY and hold it for 10 days. On the contrary, if $p_s = 5$ and $\widehat{p}_{s+l} = 4.9$, we short it.

The reason of setup $\tau$ comes from the following several aspects: first, when we set $\tau = 0$, the number of transactions must be very large. Since high frequency trading comes with

Table 10: Performance of trading strategy based on the best EMD-MLP.

| | | $N_s$ | SD% | Return with transaction cost (%) | | | |
|---|---|---|---|---|---|---|---|
| | | | | 0.0 | 0.1 | 0.2 | 0.3 |
| Panel A: $\tau = 0$ | | | | | | | |
| 1-day | EMD(-1)-MLP(3)* | 812 | 1.46 | 13.59 | −11.41 | −36.41 | −61.41 |
| 5-day | EMD(-1)-MLP(3)* | 822 | 1.53 | 7.12 | 2.12 | −2.88 | −7.88 |
| 10-day | EMD(-1)-MLP(5,3)* | 830 | 1.70 | 6.19 | 3.69 | 1.19 | −1.31 |
| 20-day | EMD(-1)-MLP(5,3)* | 823 | 1.76 | 4.98 | 3.73 | 2.48 | 1.23 |
| 30-day | EMD(-2)-MLP(5,3)* | 823 | 1.73 | 4.59 | 3.76 | 2.92 | 2.09 |
| Panel B: $\tau = 0.5\%$ | | | | | | | |
| 5-day | EMD(-1)-MLP(3)* | 32 | 3.65 | 36.78 | 31.78 | 26.78 | 21.78 |
| 10-day | EMD(-1)-MLP(5,3)* | 112 | 2.80 | 19.18 | 16.68 | 14.18 | 11.68 |
| 20-day | EMD(-1)-MLP(5,3)* | 220 | 2.03 | 12.15 | 10.90 | 9.65 | 8.40 |
| 30-day | EMD(-2)-MLP(5,3)* | 364 | 1.75 | 8.30 | 7.47 | 6.64 | 5.80 |
| Panel C: $\tau = 1\%$ | | | | | | | |
| 5-day | EMD(-1)-MLP(3)* | 5 | 4.93 | 85.02 | 80.02 | 75.02 | 70.02 |
| 10-day | EMD(-1)-MLP(5,3)* | 17 | 4.72 | 40.20 | 37.70 | 35.20 | 32.70 |
| 20-day | EMD(-1)-MLP(5,3)* | 71 | 2.29 | 18.42 | 17.17 | 15.92 | 14.67 |
| 30-day | EMD(-2)-MLP(5,3)* | 131 | 1.72 | 13.08 | 12.24 | 11.41 | 10.58 |

In terms of NMSE and $D_{\text{stat}}$, we select the best models to forecasting $l$-day ahead predictions where $l = 1, 5, 10, 20$, and 30. The third column shows the sample size in each model and is denoted as $N_s$. By the trading strategy rule as (11) with $\tau = 0$, we generate $N_s$ $l$-day returns. The fourth column reports the standard deviation of the annualized returns without transaction cost. The last four columns represent the averages of $N_s$ annualized returns with different transaction costs.

extremely high transaction costs, the erosion of profits should not be ignored and it usually makes the trading strategy fail in reality. Therefore, letting $\tau$ be larger than zero can reduce the number of trades. More importantly, $\tau$ could be regarded as a confidence level indicator, which produces long and short trading thresholds, i.e., $p_s(1+\tau)$ and $p_s(1-\tau)$. Only when the $l$-day ahead prediction $\widehat{p}_{s+l}$ is larger (smaller) than the thresholds, we long (short) the CNY. Specifically, when we set $\tau = 1\%$, denoting that if the forecast price exceeds the present price by more than 1%, we buy it; on the contrary, if the forecast price exceeds the present price by less than 1%, we short it. To fit the practical situation, this empirical analysis also considers annual returns with transaction costs of 0.0%, 0.1%, 0.2%, and 0.3%, respectively.

Table 10 discusses the performance of the above trading strategy with $\tau = 0, 0.5\%, 1\%$, based on best models, selected according to Tables 2 and 3. It displays the best performance models with different forecasting periods. The returns shown in the last four columns have been adjusted to annual returns using different transaction costs. Firstly, it can be seen that the larger the $l$ is, which means the forecasting period is longer, the larger the standard deviation will be. If there is no transaction cost, the return decreases with the increase of the forecasting period $l$. For instance, in Panel A, choosing a forecasting period of 1 day with $\tau = 0$ and zero transaction cost, the annual return will reach 13.59% by choosing the best performance model EMD(-1)-MLP(3)*. However, extending the forecasting period to 30 days, the return will become 4.59% with the best performance model.

However, if transaction costs are considered, it is clear that the return increases as the forecasting period becomes longer, which is opposite to the case where there are no transaction costs. The annual return suffers significant falls in short forecasting periods after considering the transaction cost since the annual return of a short period is offset by the significant cost of high-frequency trading. For instance, with a 0.3% trading cost, the annual return of the best performance model for 30-day ahead forecasting is positive and is only 2.09%, but the annual return of EMD(-1)-MLP(3)*, which is the best performance for forecasting 1-day ahead, reaches as low as −61.41%. The annual return decreases with increasing transaction costs for every forecasting period.

Panels B and C, respectively, display trading strategy performance based on the best model when $\tau = 0.5\%$ and $\tau = 1\%$. The 1-day forecasting case is ignored in this experiment since the predicted value of the 1-day period does not exceed $\tau$. The standard deviation in the fourth column decreases with the growth of forecasting periods. Another finding in these two panels is that the annual return decreases as the transaction cost becomes larger in the same forecasting period, which is similar to Panel A. However, with the same transaction cost, the annual return decreases with the growth of the forecasting period, which is contrary to the situation in Panel A. This may be the result of setting $\tau = 0.5\%$ or $\tau = 1\%$; the annual return will not be eroded by trading costs.

When we set $\tau = 0.5\%$ for 5-day ahead forecasting, although the average annual return with different trading costs is at least 21.78%, there are only 32 trading activities in a total of 832 trading days. However, in choosing long period forecasting, such as 20 days ahead, trading activity is 220 with a 8.4% average annual return. In Panel C, the trading activities are less than in Panel B. If a 0.3% trading cost with $\tau = 1\%$ is considered, the annual return of 30-day ahead forecasting will exceed 10%, and there are more than 130 trading activities. It can be seen that trading activities reduce with the growth of critical number $\tau$. To sum up, applying EMD-MLP* to forecast the CNY could produce some useful trading strategies, even considering reasonable trading costs.

## 4. Conclusions

In this paper, RMB exchange rate forecasting is investigated and trading strategies are developed based on the models constructed. There are three main contributions in this study. First, since the RMB's influence has been growing in recent years, we focus on the RMB, including the onshore RMB exchange rate (CNY) and offshore RMB exchange rate (CNH). We use daily data to forecast RMB exchange rates with different horizons based on three types of models, i.e., MLP, EMD-MLP, and EMD-MLP*. Our empirical study verifies the feasibility of these models. Secondly, in order to develop reliable forecasting models, we consider not only the MLP model but also the hybrid EMD-MLP and EMD-MLP* models to improve forecasting performance. Among all the selected models, the EMD-MLP* performs best, in terms of both NMSE and $D_{\text{stat}}$ criteria. It should be noted that the EMD-MLP* is different from the methodology proposed by Yu et al. [27]. We regard some IMF components as noise factors and delete them to reduce the volatility of the RMB. The empirical experiments verify that the above process could clearly improve forecasting performance. For example, Table 1 shows that the best models are EMD(-1)-MLP(3)*, EMD(-1)-MLP(5,3)*, and EMD(-2)-MLP(5,3)*. Finally, we choose the best forecasting models to construct the trading strategies by introducing different critical numbers and considering different transaction costs. As a result, we abandon the trading strategy of $\tau = 0$ since the annual returns are mostly negative. However, given $\tau = 0.5\%$ or 1%, all annual returns are more than 5%, even including the transaction cost. In particular, by setting $\tau = 1\%$ with 0.3% transaction cost, the annual return is more than 10% in different forecasting horizons. Thus, this trading strategy can help investors make decisions about the timing of RMB trading.

As the Chinese government continues to promoting RMB internationalization, RMB currency trading becomes increasingly important in personal investment, corporate financial decision-making, government's economic policies, and international trade and commerce. This study is trying to apply the neural network into financial forecasting and trading area. Based on the empirical analysis, the forecasting accuracy and trading performance of RMB exchange rate can be enhanced. Considering the performance of this proposed method, the study should be attractive not only to policymakers and investment institutions but also to individual investors who are interested in RMB currency or RMB-related products.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Authors' Contributions

All authors have contributed equally to this article.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (NSFC no. 71961007, 71801117, and 71561012). It was also supported in part by Humanities and Social Sciences Project in Jiangxi Province of China (JJ18207 and JD18094) and Science and Technology Project of Education Department in Jiangxi Province of China (GJJ180278, GJJ170327, and GJJ180245).

## References

[1] M. Fratzscher and A. Mehl, "China's dominance hypothesis and the emergence of a tri-polar global currency system," *The Economic Journal*, vol. 124, no. 581, pp. 1343–1370, 2014.

[2] C. R. Henning, "Choice and coercion in East Asian exchange-rate regimes," *Power in a Changing World Economy, Routledge*, pp. 103–124, 2013.

[3] C. Shu, D. He, and X. Cheng, "One currency, two markets: the renminbi's growing influence in Asia-Pacific," *China Economic Review*, vol. 33, pp. 163–178, 2015.

[4] A. Subramanian and M. Kessler, "The renminbi bloc is here: Asia down, rest of the world to go?," *Journal of Globalization and Development*, vol. 4, no. 1, pp. 49–94, 2013.

[5] R. Craig, C. Hua, P. Ng, and R. Yuen, "Chinese capital account liberalization and the internationalization of the renminbi," *IMF Working Paper*, 2013.

[6] M. Funke, C. Shu, X. Cheng, and S. Eraslan, "Assessing the CNH-CNY pricing differential: role of fundamentals, contagion and policy," *Journal of International Money and Finance*, vol. 59, pp. 245–262, 2015.

[7] M. Khashei, M. Bijari, and G. A. Raissi Ardali, "Improvement of auto-regressive integrated moving average models using fuzzy logic and artificial neural networks (ANNs)," *Neurocomputing*, vol. 72, no. 4–6, pp. 956–967, 2009.

[8] Z.-X. Wang and Y.-N. Chen, "Nonlinear mechanism of RMB real effective exchange rate fluctuations since exchange reform: empirical research based on smooth transition autoregression model," *Financial Theory & Practice*, vol. 6, p. 4, 2016.

[9] G. Zhang, B. Eddy Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: the state of the art," *International Journal of Forecasting*, vol. 14, no. 1, pp. 35–62, 1998.

[10] G. P. Zhang, "An investigation of neural networks for linear time-series forecasting," *Computers & Operations Research*, vol. 28, no. 12, pp. 1183–1202, 2001.

[11] C. H. Aladag and M. M. Marinescu, "Tl/Euro and Leu/Euro exchange rates forecasting with artificial neural networks," *Journal of Social and Economic Statistics*, vol. 2, no. 2, pp. 1–6, 2013.

[12] L. Yu, G. Hang, L. Tang, Y. Zhao, and K. Lai, "Forecasting patient visits to hospitals using a WD&ANN-based decomposition and ensemble model," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 13, no. 12, pp. 7615–7627, 2017a.

[13] Y. Kajitani, K. W. Hipel, and A. I. McLeod, "Forecasting nonlinear time series with feed-forward neural networks: a case study of Canadian lynx data," *Journal of Forecasting*, vol. 24, no. 2, pp. 105–117, 2005.

[14] L. Yu, Y. Zhao, and L. Tang, "Ensemble forecasting for complex time series using sparse representation and neural networks," *Journal of Forecasting*, vol. 36, no. 2, pp. 122–138, 2017.

[15] L. Cao, "Support vector machines experts for time series forecasting," *Neurocomputing*, vol. 51, pp. 321–339, 2003.

[16] L. Yu, H. Xu, and L. Tang, "LSSVR ensemble learning with uncertain parameters for crude oil price forecasting," *Applied Soft Computing*, vol. 56, pp. 692–701, 2017b.

[17] L. Yu, Z. Yang, and L. Tang, "Quantile estimators with orthogonal pinball loss function," *Journal of Forecasting*, vol. 37, no. 3, pp. 401–417, 2018.

[18] M. Alvarez-Diaz and A. Alvarez, "Forecasting exchange rates using genetic algorithms," *Applied Economics Letters*, vol. 10, no. 6, pp. 319–322, 2003.

[19] C. Neely, P. Weller, and R. Dittmar, "Is technical analysis in the foreign exchange market profitable? A genetic programming approach," *The Journal of Financial and Quantitative Analysis*, vol. 32, no. 4, pp. 405–426, 1997.

[20] L. Yu, S. Wang, and K. K. Lai, *Foreign-xchange-ate Forecasting with Artificial Neural Networks*, vol. 107, Springer Science & Business Media, Berlin, Germany, 2010.

[21] G. P. Zhang, "Time series forecasting using a hybrid ARIMA and neural network model," *Neurocomputing*, vol. 50, pp. 159–175, 2003.

[22] M. Álvarez-Díaz and A. Álvarez, "Genetic multi-model composite forecast for non-linear prediction of exchange rates," *Empirical Economics*, vol. 30, no. 3, pp. 643–663, 2005.

[23] L. Zhao and Y. Yang, "PSO-based single multiplicative neuron model for time series prediction," *Expert Systems with Applications*, vol. 36, no. 2, pp. 2805–2812, 2009.

[24] W. K. Wong, M. Xia, and W. C. Chu, "Adaptive neural network model for time-series forecasting," *European Journal of Operational Research*, vol. 207, no. 2, pp. 807–816, 2010.

[25] L. Yu, Y. Zhao, L. Tang, and Z. Yang, "Online big data-driven oil consumption forecasting with google trends," *International Journal of Forecasting*, vol. 35, no. 1, pp. 213–223, 2019.

[26] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[27] L. Yu, S. Wang, and K. K. Lai, "Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm," *Energy Economics*, vol. 30, no. 5, pp. 2623–2635, 2008.

[28] C.-F. Chen, M.-C. Lai, and C.-C. Yeh, "Forecasting tourism demand based on empirical mode decomposition and neural network," *Knowledge-Based Systems*, vol. 26, pp. 281–287, 2012.

[29] C.-S. Lin, S.-H. Chiu, and T.-Y. Lin, "Empirical mode decomposition–based least squares support vector regression for foreign exchange rate forecasting," *Economic Modelling*, vol. 29, no. 6, pp. 2583–2590, 2012.

[30] L. Tang, Y. Wu, and L. Yu, "A non-iterative decomposition-ensemble learning paradigm using RVFL network for crude oil price forecasting," *Applied Soft Computing*, vol. 70, pp. 1097–1108, 2018.

[31] M. Álvarez Díaz, "Speculative strategies in the foreign exchange market based on genetic programming predictions," *Applied Financial Economics*, vol. 20, no. 6, pp. 465–476, 2010.

[32] W. Huang, K. Lai, Y. Nakamori, and S. Wang, "An empirical analysis of sampling interval for exchange rate forecasting with neural networks," *Journal of Systems Science and Complexity*, vol. 16, no. 2, pp. 165–176, 2003b.

[33] A. Nikolsko-Rzhevskyy and R. Prodan, "Markov switching and exchange rate predictability," *International Journal of Forecasting*, vol. 28, no. 2, pp. 353–365, 2012.

[34] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: the RPROP algorithm," in *Proceedings of the IEEE International Conference on Neural Networks*, pp. 586–591, IEEE, San Francisco, CA, USA, March-April 1993.

[35] N. E. Huang and Z. Wu, "A review on Hilbert-Huang transform: method and its applications to geophysical studies," *Reviews of Geophysics*, vol. 46, no. 2, 2008.

[36] N. E. Huang, M.-L. C. Wu, S. R. Long et al., "A confidence limit for the empirical mode decomposition and Hilbert spectral analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 459, no. 2037, pp. 2317–2345, 2003a.

[37] J. Yao and C. L. Tan, "A case study on using neural networks to perform technical forecasting of forex," *Neurocomputing*, vol. 34, no. 1–4, pp. 79–98, 2000.

[38] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *Journal of Business & Economic Statistics*, vol. 20, no. 1, pp. 134–144, 2002.

[39] L. Yu, S. Wang, and K. K. Lai, "A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates," *Computers & Operations Research*, vol. 32, no. 10, pp. 2523–2541, 2005.

[40] M. H. Pesaran and A. Timmermann, "A simple nonparametric test of predictive performance," *Journal of Business & Economic Statistics*, vol. 10, no. 4, pp. 461–465, 1992.

[41] S. Anatolyev and A. Gerko, "A trading approach to testing for predictability," *Journal of Business & Economic Statistics*, vol. 23, no. 4, pp. 455–461, 2005.