# Modeling and Forecasting of Fractal Volatility of Highly Volatile Markets

Yuan Yao and Yan Li

*Abstract*—**Characterizing volatility is of great significance for the study of financial markets. This paper analyzes realized volatility (RV), weighted adjusted realized volatility (WARV), multi-fractal volatility (MFV), and weighted adjusted multi-fractal volatility (MVWA) sequences based on closing price sequences of the Shanghai Composite Index. The results show that all four volatility sequences have the characteristics of a peak, fat tail, and long memory. The heterogeneous autoregressive (HAR) model can well depict these peak, fat-tail, and long-memory features. Thus, we construct HAR-RV, HAR-AWRV, HAR-MFV, and HAR-MVWA models and find that medium-term and long-term volatility influences predicted volatility. We also introduce a superior long short-term memory (LSTM) model. To expect volatility, results from HAR-RV, HAR-AWRV, HAR-MFV, and HAR-MVWA, along with daily, weekly, and monthly volatility sequences for RV, AWRV, MFV, and MVWA, are used as inputing variables to construct four HAR-LSTM models: HAR-RV-LSTM, HAR-AWRV-LSTM, HAR-MFV-LSTM, and HAR-MVWA-LSTM. Compared with benchmark HAR model and LSTM model, model confidence set (MCS) test results show that the HAR-LSTM model has the most predictive ability. These MCS test results are related to the selection of benchmark volatility sequences. Two multi-fractal volatility are more suitable for the Chinese stock market, and MVWA works better.**

*Index Terms*—**RV, Fractal Volatility, HAR, LSTM, HAR-LSTM, MCS test.**

## I. Introduction

VOLATILITY is one of the most critical risk management indicators for financial assets. It can be classified into three categories: implied volatility, time-varying volatility, and Realized volatility(RV), RV measured by a non-parametric method based on high-frequency data, can reflect the most information on prices and describe market volatility most accurately [1]-[2].

However, RV follows the efficient market hypothesis, which holds that the current price of an underlying asset reflects all the information available between past and present. The market presents a random walk: regardless of what may have occurred or is currently taking place, those cannot determine what will happen in the future. However, in reality, investors do not immediately respond to information but wait until the trend is evident before taking corresponding measures. The response time is different, leading to biased random walks. The fractal market hypothesis has therefore attracted attention[3].

Manuscript received March 14, 2022; revised October 24, 2022.

This work was supported by Henan Provincial Philosophy and Social Sciences Planning Annual Project under Grant 2022BJJ030 and Henan Provincial Colleges and Universities Philosophy and Social Sciences Fundamental Research Major Project 2021-JCZD-01

Yuan Yao is Professor of Business School, Henan University, Kaifeng, 475004, China, (e-mail: yaoyuan@henu.edu.cn).

Yan Li is postgraduate of Business School, Henan University, Kaifeng, 475004, China, (corresponding author: 104753190936@henu.edu.cn).

Meanwhile, the fractal market hypothesis holds that all stable markets have fractal structures. Based on fractal theory, multi-fractal volatility(MFV) has strong practicability in financial markets [4]. To reduce extremes' influence on MFV, use the standard deviation of a multi-fractal singularity index to improve MFV [5]. Reducing the power of noise on volatility, RV replaces the square of returns in multi-fractal fluctuation [6]. Meanwhile, weight-adjusted realized volatility is used as the correction factor to reduce noise's influence on volatility further and obtain a better method for measuring MFV [7]. Research has proved the advantages of MFV and ameliorative MFV and the feasibility of applying MFV in China's market.

While measuring volatility is continuously improved, volatility modeling methods have also received extensive attention. Heterogeneous autoregressive (HAR) model can better portray RV's long-memory and fat-tail distributions [8]; consequently, based on HAR model, scholars have made a series of optimizations. Introduced the jump component constructs a HAR-RV-J model, the jump can affect predicted next-day volatility [9]. Based on external information shocks constructs a HAR-VRV-hopping model, the results show that HAR-VRV-hopping model outperform HAR-RV-hopping model [10].

Based on HAR model, Some scholars have made improvements by considering jump behavior and leverage effect [11].At the same time, The existing literature has focused on the symbolic price difference [12]. Research devoted to learn modified threshold biopower variation [13].

Recently, based on machine learning, predicting volatility models have evolved enormously. Through out-of-sample volatility prediction, this existing literature has proved the advantage of Long short-term memory(LSTM) [14]. When predict realized volatility through LSTM-RV, the accuracy is significantly improved [15]. Meanwhile, research into hybrid model has also attracted attention among researchers. Whether construct factor-augmented heterogeneous autoregressive model (FAHAR), combining LSTM with FAHAR to predict volatility[16]or combining LSTM with realized GARCH (RGARCH) model to predict volatility[17], results have supported two hybrid models. For example, a hybrid model, combining neural network with GARCH model, is better than a single model [18].

Chinese researchers have conducted a series of studies on the combination of neural networks and GARCH models in which RV is divided into long-term and short-term components by HP filter. The long-term part is fitted using an autoregressive neural network, and the short-term part is fitted by AR (1) model. At the same time, A two-component mixed exchange rate volatility model based on neural network has been proposed [19] in which the parameters and explanatory variables of GARCH family models are added to the neural network's input-layer variables. This mixed model, using LSTM and GARCH [20], shows improved forecast precision for RMB exchange rate fluctuations.

According to the above analysis of extensive literature, researchs on hybrid models have two types: First, through decomposition and reconstruction frame to attain the best prediction results. Second, Accroding to the econometric models fit volatility sequences, fitting parameters and explanatory variables are added as inputs variables to deep learning models. There have been relatively little research devoted to combining LSTM with HAR. This study based on the second type researches, combine LSTM with the HAR model to predict the Shanghai Composite Index volatility.

Based on the above understanding, this paper has several contributions. First, We construct the Shanghai Composite Index's volatility sequences, including RV, WARV, MFV, and MVWA and examined some features of several kinds of volatility sequences. The statistical analysis shows that these sequences have a peak, fat tail, and long memory. Second, In that HAR model can well characterize peak, the fat tail, and long memory, we use the HAR model to predict several kinds of volatility sequences. The in-sample results show that medium-term and long-term volatility influence volatility forecasts. Third, Considering the advantages of LSTM in long memory, we build LSTM to predict each volatility series. And we use results obtained in the second step as augmented input variables.

Fourth, literatures show that hybrid models are better than a single model, and few scholars have combined HAR model and LSTM. This paper constructs the HAR-LSTM model to predict volatility sequences. Last, considering a single loss function index may lead to errors in a model's evaluation, we introduce an MCS test. This test allowes us to evaluate out-of-sample predictability within several models in this paper, including HAR-RV-LSTM, HAR-AWRV-LSTM, HAR-MFV-LSTM, HAR-MVWA-LSTM, RV-LSTM, AWRV-LSTM, MFV-LSTM, MVWA-LSTM, HAR-RV, HAR-WARV, HAR-MFV, and HAR-MVWA. The results show that the HAR-LSTM model has the most potent predictive ability; meanwhile, MCS results are associated to benchmark volatility sequences, and MVWA can enhance the model's accuracy.

## II. Models and Method

### A. HAR Model

HAR, which had a long memory and volatility heterogeneity was proposed by Corsi [8], can predict volatility. It regards daily, weekly, and monthly volatility as proxy variables for short-term, medium-term, and long-term fluctuations. The specific regression relationships are as follows:

$$RV_{t+1} = \beta_0+\beta_d RV_t+\beta_w RV_{t-5,t}+\beta_m RV_{t-22,t}+\varepsilon_{t+1} \quad (1)$$

$$\begin{cases} RV_t = \sum_{i=1}^{n} r_{t,i}^2 \\ RV_{t-5,t} = \frac{1}{5}\sum_{i=0}^{4} RV_{t-i}^2 \\ RV_{t-22,t} = \frac{1}{22}\sum_{i=0}^{21} RV_{t-i}^2 \end{cases} \quad (2)$$

Here, $\varepsilon_{t+1}$is a random error term; $RV_t$, $RV_{t-5,t}$ , and $RV_{t-22,t}$ represent daily, weekly, and monthly RV, respectively.

### B. LSTM Model

Long and short-term memory neural network (LSTM) has the strengths of long memory and can play a better role in volatility prediction. The unit of LSTM, as shown in Fig. 1.

At *t*, the forward propagation formula of LSTM as shown in (8)–(13):

Storage unit candidate state:

$$\widetilde{c}^t = \tanh(W_c x^t + U_c h^{t-1}) \quad (3)$$

Input gate:

$$i^t = \sigma(W_i x^t + U_i h^{t-1}) \quad (4)$$

Forget gate:

$$f^t = \sigma(W_f x^t + U_f h^{t-1}) \quad (5)$$

Output gate:

$$o^t = \sigma(W_o x^t + U_o h^{t-1}) \quad (6)$$

Storage unit update:

$$c^t = i^t \odot \tilde{c}^t + f^t \odot c^{t-1} \quad (7)$$

And finally, the result at *t*:

$$h^t = o^t \odot \tanh(c^t) \quad (8)$$

Here, $h_{t-1}$ is output at *t*-1 moment, $x_t$ is output at *t*, $W_*$ and $U_*$ are the weight and the storage unit, respectively, while $\sigma$ is the sigmoid function; $\odot$ represents the Hadamard product between elements.

### C. Evaluation Indicators and MCS Test

There is no fixed standard for model evaluation indicators. According to Hansen et al. [21], a model is evaluated using as many arrows as possible so that the evaluation results will be as reasonable. Six indicators (MAE, MSE, HMAE, HMSE, QLIKE, and R2LOG) are commonly used, nevertheless, LSTM results may present negative values, we use the four following evaluation indicators:

$$L_1 : MAE = \frac{1}{N}\sum_{t=1}^{N} |y_t - \widehat{y}_t| \quad (9)$$

$$L_2 : MSE = \frac{1}{N}\sum_{t=1}^{N} (y_t - \widehat{y}_t)^2 \quad (10)$$

$$L_3 : HMAE = \frac{1}{N}\sum_{t=1}^{N} \left|1 - \frac{\widehat{y}_t}{y_t}\right| \quad (11)$$

$$L_4 : HMSE = \frac{1}{N}\sum_{t=1}^{N} \left(1 - \frac{\widehat{y}_t}{y_t}\right)^2 \quad (12)$$

Here, *N* is length of prediction sample, $\hat{y}_t$ is prediction result, and $y_t$ is volatility (RV, AWRV, MFV, and MVWA).

However, if a given rating index ($L_i$) is used to evaluate the model's predictive ability, it can only prove that the model has superiority in this period, and according to this evaluation index, with a change in the data, this conclusion may be wrong. We use the MCS test proposed by Hansen et al. [22] to solve this problem. The MCS detects a model's predictive

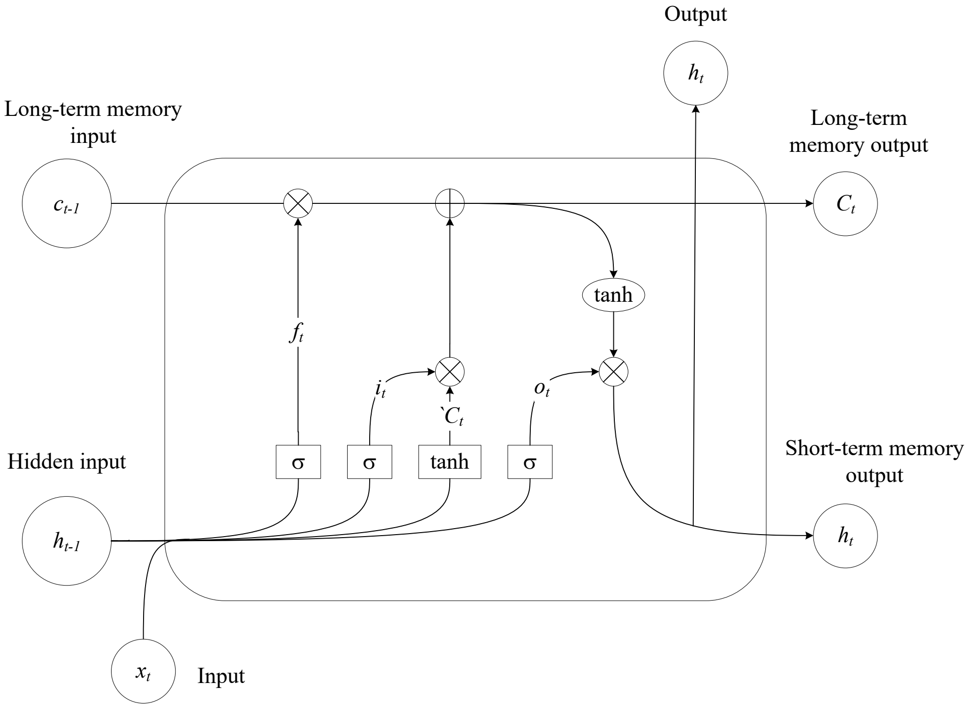


Fig. 1. Unit (LSTM)

ability and finds the best model from a set for the selected confidence level.

The MCS test proceeds follows: Construct a model set $M^0$ (in this paper, $M^0$ contains 12 models, $M^0 = \{1, 2, \cdots, m_0 = 12\}$ ). Each model has $N$ next-day prediction result sets $\widehat{y}_t \, (t = 1, 2, \cdots, N)$, with loss function values of above four evaluation indicators $(L_1, L_2, L_3, L_4)$ denoted as $L_{i,j,t}$, $i = 1, 2, \cdots, m_0$, $j = 1, 2, 3, 4$, and $t = 1, 2, \cdots N$. Relative loss function values $d_{ii+k,j,t} = L_{i,j,t} - L_{i+k,j,t}$ can be calculated for any two volatility models. Denoted as $\mu_{ii+k,j} = E\left(d_{ii+k,j,t}\right)$, if$\mu_{ii+k,j} < 0$, the *i*-th model is better than the *i*+*k*-th model. That is to say, it is denoted as the set of superior objects $M^* = \{i \in M^0 : \mu_{ii+k,j} \leq 0, forallj \in M^0\}$.

The purpose of the MCS test is to find using a series of significant tests to eliminate the model with poorer predictive ability. Both models are assumed to have the same predictive power each time $H_{0,M} : \mu_{ii+k,j} = 0, foralli, i + k \in M$,where $M \subset M^0$.

The MCS test is based on equivalence testing ($\delta_M$) and an elimination rule ($e_M$). Equivalence tests test the null hypothesis for each $M \subset M^0$; the elimination rule is: Delete the models in $M$ which reject the null hypothesis.

The MCS algorithm steps are as follows:

First, set $M = M^0$.

Second, according to equivalence test $\delta_M$ test the null hypothesis $H_{0,M}$ at significance level $\alpha$.

Three, if accepting the null hypothesis $H_{0,M}$, denote it as $\widehat{M}^*_{1-\alpha} = M$; otherwise, remove the model that rejects the null hypothesis according to the elimination rule, repeat steps 2 and 3, and finally get the MCS $\widehat{M}^*_{1-\alpha}$.

In this paper, the MCS algorithm *t* statistics are as follows:

$$T_R = \max_{i,i+k \in M} \frac{\left|\overline{d}_{ii+k,j}\right|}{\sqrt{\operatorname{var}\left(d_{ii+k,j}\right)}} \tag{13}$$

$$T_{SQ} = \max_{i,i+k \in M} \frac{\left(\overline{d}_{ii+k,j}\right)^2}{\operatorname{var}\left(\overline{d}_{ii+k,j}\right)} \tag{14}$$

$$\overline{d}_{ii+k,j} = \frac{1}{N} \sum_{t=1}^{N} d_{ii+k,j,t} \tag{15}$$

The null hypothesis is rejected if the *t* statistic is more significant than a given critical value. We used the bootstrap algorithm to calculate $T_R$, $T_{SQ}$, and the *p* value.

### D. HAR-LSTM Family Models

In this paper, the HAR-LSTM family models include the following parts:

1) Construct volatility sequences using the closing price of five-minute transaction data: RV, AWRV, MFV, and MVWA.

2) The HAR model predict the volatility sequences, which can capture these sequences' long memory and fat tail. The HAR model prediction results are obtained, including HAR-RV, HAR-AWRV, HAR-MFV, and HAR-MVWA.

3) The prediction results of HAR-RV, HAR-AWRV, HAR-MFV, and HAR-MVWA, as well as the daily, weekly, and monthly volatility sequences of RV, AWRV, MFV, and MVWA, are input into the LSTM model to construct the HAR-LSTM model. The HAR-LSTM model prediction results are obtained, including HAR-RV-LSTM, HAR-AWRV-LSTM, HAR-MFV-LSTM, and HAR-MVWA-LSTM.

4) According to the MCS test, we compare and analyze the prediction results of HAR-LSTM family models, HAR family models, and LSTM family models.

The structure of HAR-LSTM is shown in Fig. 2.

## III. Results

### A. Data Selection and Analysistitle

As the research object, the five-minute high-frequency trading data of the Shanghai Composite Index from January

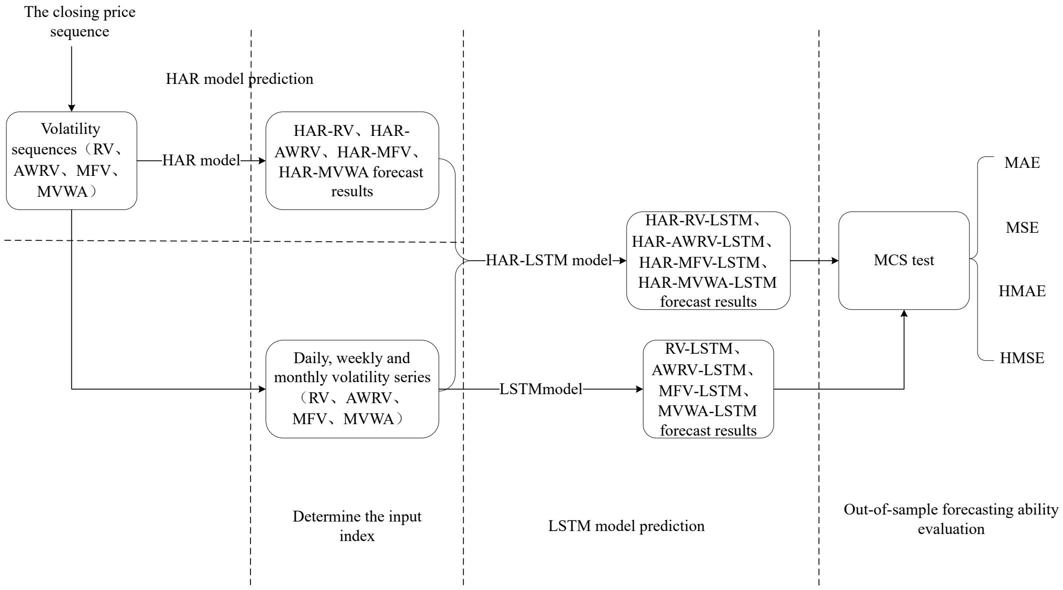


Fig. 2. Family models (HAR-LSTM)

2, 2011, to September 30, 2021, are obtained from the Wind database. To show that the asymmetric fractality of the stock market has a predictive power, four volatility elements are constructed accroding formulas (16)–(19), which are as follows: RV, AWRV, MFV, and MVWA.

$$\begin{cases} RV_t = \sum_{i=1}^{n} r_{t,i}^2 \\ r_{t,j} = \log\left(P_{t,j}/P_{t-1,j}\right) \end{cases} \tag{16}$$

$$\begin{cases} a = \frac{E(WRQ_t)/n}{Var(WRV_t)} \\ b = \frac{Var(WRV_t) - E(WRQ_t)/n}{Var(WRV_t)} \\ WARV_t = aE(WRV_t) + bWRV_t \\ WRV_t = \sum_{j=1}^{n} \varpi_j r_{t-1+\frac{j}{n},\frac{j}{n}}^2 \\ WRQ_t = \frac{2}{3} n \sum_{j-1}^{n} \varpi_j r_{t-1+\frac{j}{n},\frac{j}{n}}^4 \\ \varpi_j = \frac{\sum_t \sum_{j-1}^{n} r_{t,j}^2}{n \sum_t r_{t,j}^2} \end{cases} \tag{17}$$

$$\begin{cases} MFV_t = \lambda_1 \Delta_{\alpha,t}, \lambda_1 = \frac{E(r_t^2)}{E(\Delta_{\alpha,t})} \\ \Delta_\alpha = \alpha_{\max} - \alpha_{\min} \end{cases} \tag{18}$$

$$MVWA_t = \lambda S_{\alpha,t}, \lambda = \frac{E(WARV_t)}{E(S_{\alpha,t})} \tag{19}$$

Here, $P_{t,j}$ marks the close of period *i* in trading day *t*; $r_{t,j}$ represents the return rate during *i* for day *t*, total *n* sampling intervals on *t*-th trading day. With $\alpha$ as the singular index, $S_{\alpha,t}$ is the standard deviation for $\alpha$.

Fig.3 is four volatility sequences, we can see that RV and AWRV exist some abnormal values, and abnormal values in AWRV are smaller than RV, but the number of abnormal values are not reduce. MFV and MVWA can reduce abnormal values, and MVWA sequence seemingly more stable.

Table I shows the data statistics of the Shanghai Composite Index. The standard deviation results show that MVWA has the highest stability. At the same time, skewness and kurtosis results show that RV, AWRV, MFV, and MVWA sequences have a peak characteristic.As a result of the Jarque–Bera test, all series in Table I are not the normal distribution at a 1% significance level, each volatility sequence has fat-tail characteristics. The skewness and kurtosis results are the smallest in the MFV sequence, proving that MFV is closer to normal distribution. The Ljung–Box Q statistics of RV, AWRV, MFV, and MVWA sequences show that the null hypothesis is rejected concerning a 1% significance level, so the samples show autocorrelation, in other words, RV, AWRV, MFV, and MVWA show long memory. The absence of unit root is rejected with 1% significant level as a result of the Augmented Dickey–Fuller (ADF) test. Therefore, all time-series are stationary.

### B. Experimental Process

As previously stated, we use HAR-RV, HAR-AWRV, HAR-MFV, and HAR-MVWA to forecast the volatility series of the Shanghai Composite Index. The results are shown in Table II.

As Table II shows, in addition to the HAR-MFV model, the parameter estimation with positive short-term, medium-term, and long-term fluctuations in each model indicate that the market has strong persistence. The parameters of the realized volatility model are indigenous at the 1% significance level. In the multi-fractal volatility models, the short-term volatility of HAR-MFV was not indigenous at the 10% significance level. Compared with models based on realized volatility, models based on multi-fractal volatility are more affected by medium-term and long-term fluctuations. Therefore, to improve the model's prediction accuracy, in this paper, we consider introducing LSTM, which is superior in long memory. Compared with models based on multi-fractal volatility, these models based on realized volatility, the goodness of fit is notably higher. In addition, the HAR-AWRV model has the highest measure of goodness of fit, showing the more substantial explanatory power of AWRV for market volatility characteristics.

Therefore, considering that long-term volatility has a higher impact on the prediction of multi-fractal volatility, we introduce LSTM with a long memory. The HAR family models' prediction results and volatility sequences are input vari-

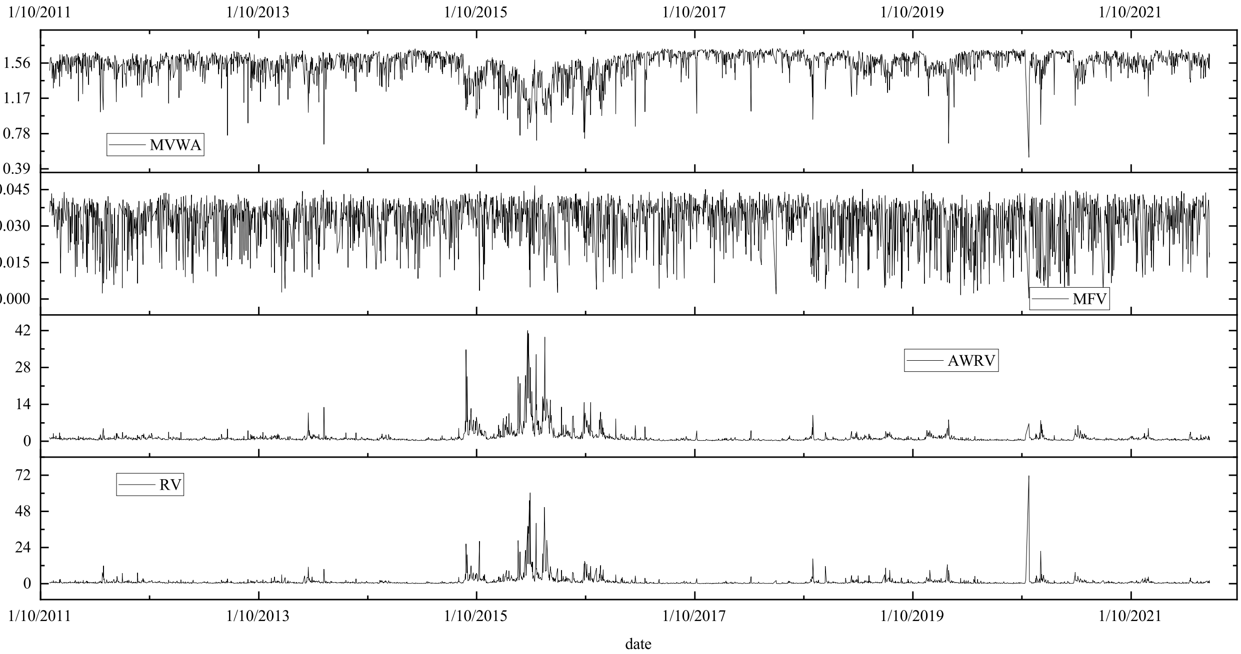


Fig. 3. Several volatility sequences

TABLE I
STATISTICAL ANALYSIS RESULTS FOR VOLATILITY SERIES

| | RV | AWRV | MFV | MVWA |
|---|---|---|---|---|
| mean | 1.544615 | 1.544615 | 1.544615 | 1.544615 |
| std | 3.751937 | 2.932712 | 0.428785 | 0.149851 |
| skewness | 9.520172 | 7.711306 | -1.21006 | -2.06217 |
| kurtosis | 125.0194 | 82.29514 | 3.866211 | 8.724687 |
| Jarque–Bera | 1661118*** | 710744.1*** | 719.6432*** | 5422.11*** |
| Q (5) | 3121.7*** | 5061.9*** | 83.978*** | 2907.9*** |
| Q (10) | 4429.2*** | 7590.8*** | 117.50*** | 5038.9*** |
| Q (22) | 6625.7*** | 12571*** | 199.28*** | 9139.7*** |
| ADF | -11.57991*** | -5.355688*** | -18.03181*** | -5.685381*** |

Note: "*", "* *", and "***" are expressed at the level of 1%, 5%, and 10%, respectively. Q(n) is the Ljung–Box Q statistic with lag order *n*, and ADF is the Augmented Dickey–Fuller unit root test result.

TABLE II
PARAMETER ESTIMATION RESULTS FOR EACH VOLATILITY MODEL

| model | $\beta_0$ | $\beta_d$ | $\beta_w$ | $\beta_m$ | $R^2$ |
|---|---|---|---|---|---|
| HAR-RV | 0.209188*** | 0.406299*** | 0.363457*** | 0.104702*** | 0.492801 |
| | (2.922089) | (18.97775) | (10.64621) | (2.842903) | |
| HAR-AWRV | 0.155912*** | 0.441998*** | 0.294097*** | 0.171093*** | 0.553077 |
| | (2.646618) | (20.58368) | (8.585795) | (4.807281) | |
| HAR-MFV | 0.714464*** | -0.00554 | 0.166553*** | 0.378329*** | 0.032692 |
| | (6.536924) | (-0.25302) | (3.001105) | (4.258418) | |
| HAR-MVWA | 0.148407*** | 0.243831*** | 0.307793*** | 0.351842*** | 0.412301 |
| | (3.873354) | (11.22174) | (7.136988) | (7.23526) | |

Note: "*", "* *", and "***" are expressed at 1%, 5% and 10%, respectively.

ables for LSTM. We construct four hybrid models: HAR-RV-LSTM, HAR-AWRV-LSTM, HAR-MFV-LSTM, and HAR-MVWA-LSTM. The out-of-sample prediction results of RV-LSTM, AWRV-LSTM, MFV-LSTM, and MVWA-LSTM are compared with the prediction results of HAR-RV-LSTM, HAR-AWRV-LSTM, HAR-MFV-LSTM, and HAR-MVWA-LSTM.

### C. LSTM Model Parameters Settings

This paper use a rolling window to predict volatility and predicte the volatility of the eleventh day from the first ten days' daily, weekly, and monthly volatility indicators. Whether the HAR model prediction results are input indicators, the input dimension is 3 or 4, and the output dimension is 1. The mean square error is the loss function during model training, and the model optimizer select the

Adam algorithm. We find the final model parameter with the minor loss function by comparing the results from repeated experiments. The number of hidden layers is 2, the number of remote layer nodes is 128, and the learning rate is 0.0005, while the model's batch processing capacity is 64; the number of iterations is 500. The model's predictive ability is best at these parameters. Meanwhile, to prevent the model from overfitting, we introduce an early stop mechanism into LSTM, stopping training when the validation set's loss function do not change within 30 consecutive cycles.

### D. MCS Test

Based on RV, we compare and evaluate the out-of-sample predictive ability of the proposed models using the MCS test. Use the bootstrap method to calculate the statistics and p-value for the MCS test. We set $d = 2$ (block length), iteration $B = 10000$, and significance level $\alpha = 0.1$. Table III shows MCS test results for each volatility model based on RV. The numbers in the table show the corresponding $p$ values of the $T_R$ and $T_{SQ}$statistics under each loss function of the MCS test. If the $p$ value less than 0.1, the model is weaker than other models; If the $p$ value is 1, the model is the best model.

Table III shows that HAR, LSTM, and HAR-LSTM family models, which based on AWRV, MFV, and MVWA, respectively, $p$ values are less than 0.1. MCS results show that when benchmark volatility is RV, models based on other three volatility sequences: AWRV, MFV, and MVWA are poor. Some $p$ values of HAR, LSTM, and HAR-LSTM family models based on RV are greater than 0.1, Thus, HAR, LSTM, and HAR-LSTM family models based on RV have better prediction accuracy. We believe that the MCS test results are related to the selection of benchmark volatility sequences. In addition, The HAR-RV model result only has two $p$ values are greater than 0.1 and less than 1, and RV-LSTM model has four $p$ values are greater than 0.1 and less than 1, so RV-LSTM is better than HAR-LSTM. HAR-RV-LSTM model $p$ values are all equal to 1, showing HAR-RV-LSTM has the most potent predictive ability. To verify this thought, which the MCS test results are related to the selection of benchmark volatility, and to further explore the effectiveness of multi-fractal volatility in the Chinese market, we use AWRV, MFV, and MVWA as the benchmark volatility for the MCS test, respectively. The test results are shown in Table IV, Table V, and Table VI.

As shown in Table IV, similar to Table III, only HAR-AWRV, AWRV-LSTM, and HAR-AWRV-LSTM models passed the MCS test, which $p$ values are greater than 0.1. HAR-AWRV and AWRV-LSTM models both have two $p$ values are greater than 0.1, therefore HAR-AWRV and AWRV-LSTM have the similar prediction ability. HAR-AWRV-LSTM $p$ values are all equal to 1, so when AWRV is benchmark volatility, HAR-AWRV-LSTM has the most potent predictive ability.

As shown in Table V, also only HAR-MFV, AWRV-MFV, and HAR-MFV-LSTM models passed the MCS test, which $p$ values are greater than 0.1. HAR-MFV and MFV-LSTM models both have four $p$ values are greater than 0.1, therefore HAR-MFV and MFV-LSTM also have similar prediction ability. HAR-MFV-LSTM $p$ values are all equal to 1, so when MFV is benchmark volatility, HAR-MFV-LSTM has the most potent predictive ability.

As shown in Table VI, also only HAR-MVWA, AWRV-MVWA, and HAR-MVWA-LSTM models passed the MCS test, which $p$ values are greater than 0.1. HAR-MVWA has five $p$ value greater than 0.1, and MVWA-LSTM model has four $p$ values are greater than 0.1, therefore HAR-MVWA is better than MVWA-LSTM. HAR-MVWA-LSTM $p$ values are all equal to 1, so when MVWA as the benchmark volatility, HAR-MFV-LSTM has the most potent predictive ability.

Above four table results show that the MCS test results are related to the selection of benchmark volatility. And as shown in Table III to Table VI, the HAR-LSTM model $p$ values are 1, greater than 0.1, further proving that HAR-LSTM model has the highest out-of-sample predictive ability. In addition, from Table IV to Table VI, when two realized volatility are the benchmark, these modles, which are based on realized volatility are more passed the MCS test, also consistent when the base volatility is multi-fractal volatility. Thus multi-fractal volatility sequences are better than realized volatility sequences. In other words, two multi-fractal volatility are more applicable to the Chinese market.

Fig. 4 shows the out-of-sample prediction results for HAR-LSTM family models based on MFV and MVWA, respectively. In our paper, when the prediction curve is very close to the initial volatility sequence, it proved the prediction model has high prediction accuracy. According to Fig. 4, compared with the distance between prediction line of HAR-MFV-LSTM and the MFV line, the prediction line of HAR-MVWA-LSTM is remarkably close to the MVWA line, indicating that the prediction accuracy of HAR-MVWA-LSTM is higher, meaning that MVWA can more describe Chinese stock volatility.

## IV. Conclusion

In this paper, we obtain RV, WARV, MFV, and MVWA sequences using the closing price series from the Shanghai Composite Index. Through the statistical analysis, we find that all four volatility series has the characteristics of peak, fat tail, and long memory. Literatures show that HAR models can well describe volatility characteristics, including peak, fat tail, and long memory. Therefore, we construct HAR-RV, HAR-WARV, HAR-MFV, and HAR-MVWA models based on four kinds of volatility. The sample's relevant results show that AWRV has a more vital ability to explain the market's volatility characteristics; MFV is more susceptible to medium-term and long-term fluctuations. Since the HAR model verifies the crucial role of medium-term and long-term changes in volatility, LSTM with long memory is introduced. The prediction results of HAR-RV, HAR-AWRV, HAR-MFV, and HAR-MVWA combine with the correlation volatility sequences, which contain RV, AWRV, MFV, and MVWA, are input variables of the LSTM model. Through MCS test the out-of-sample prediction accuracy of four family models. Four family models include HAR-RV-LSTM, HAR-AWRV-LSTM, HAR-MFV-LSTM, HAR-MVWA-LSTM, RV-LSTM, AWRV-LSTM, MFV-LSTM, MVWA-LSTM, HAR-RV, HAR-WARV, HAR-MFV, and HAR-MVWA. The MCS test show that the MCS test results are associated with the selection of benchmark volatility. When two multi-fractal volatility curves are the benchmark volatility, prediction models have higher prediction accuracy;

TABLE III
VOLATILITY MODELS' MCS TEST RESULTS (BASED ON RV)

| | MAE | | MSE | | HMAE | | HMSE | |
|---|---|---|---|---|---|---|---|---|
| | $T_R$ | $T_{SQ}$ | $T_R$ | $T_{SQ}$ | $T_R$ | $T_{SQ}$ | $T_R$ | $T_{SQ}$ |
| HAR-RV | 0 | 0 | 0.0001 | 0 | 0.0092 | 0.0067 | **0.2628** | **0.5885** |
| HAR-AWRV | 0 | 0 | 0 | 0 | 0 | 0.0067 | 0 | 0.0247 |
| HAR-MFV | 0 | 0 | 0 | 0 | 0 | 0.0067 | 0.0568 | 0.0247 |
| HAR-MVWA | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0.0641 | 0.0885 |
| RV-LSTM | 0 | 0 | **0.2041** | **0.1927** | 0.0251 | 0.024 | **0.2628** | **0.6247** |
| AWRV-LSTM | 0 | 0 | 0 | 0 | 0.0193 | 0.0118 | 0.0628 | 0.0247 |
| MFV-LSTM | 0 | 0 | 0 | 0 | 0.0001 | 0.0067 | 0.0628 | 0.088 |
| MVWA-LSTM | 0 | 0 | 0 | 0 | 0 | 0.0003 | 0.0638 | 0.0081 |
| HAR-RV-LSTM | **1** | **1** | **1** | **1** | **1** | **1** | **1** | **1** |
| HAR-AWRV-LSTM | 0 | 0 | 0 | 0 | 0.0193 | 0.0118 | 0.0628 | 0.0247 |
| HAR-MFV-LSTM | 0 | 0 | 0 | 0 | 0.0001 | 0.0067 | 0.0641 | 0.0724 |
| HAR-MVWA-LSTM | 0 | 0 | 0 | 0 | 0 | 0.0002 | 0.0141 | 0.0585 |

Note: The numbers in the table show $p$ values for MCS results from 10,000 bootstrap simulations; $p$ values greater than 0.1 appear in boldface.

TABLE IV
VOLATILITY MODELS' MCS TEST RESULTS (BASED ON AWRV)

| | MAE | | MSE | | HMAE | | HMSE | |
|---|---|---|---|---|---|---|---|---|
| | $T_R$ | $T_{SQ}$ | $T_R$ | $T_{SQ}$ | $T_R$ | $T_{SQ}$ | $T_R$ | $T_{SQ}$ |
| HAR-RV-LSTM | 0 | 0 | 0 | 0 | 0 | 0 | 0.0303 | 0.0307 |
| HAR-MFV-LSTM | 0 | 0 | 0 | 0 | 0 | 0 | 0.0736 | 0.0935 |
| MFV-LSTM | 0 | 0 | 0 | 0 | 0 | 0 | 0.0763 | 0.0935 |
| RV-LSTM | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0.0297 | 0.003 |
| HAR-MVWA-LSTM | 0 | 0 | 0 | 0 | 0 | 0 | 0.003 | 0.02 |
| HAR-MFV | 0 | 0 | 0 | 0 | 0 | 0 | 0.0281 | 0.0207 |
| MVWA-LSTM | 0 | 0 | 0 | 0 | 0 | 0 | 0.0028 | 0.0199 |
| HAR-MVWA | 0 | 0 | 0 | 0 | 0 | 0 | 0.0067 | 0.029 |
| HAR-RV | 0 | 0 | 0 | 0 | 0 | 0.0001 | 0.0876 | 0.0307 |
| HAR-AWRV | 0 | 0 | 0 | 0 | 0 | 0.0001 | **0.876** | **0.307** |
| AWRV-LSTM | 0 | 0 | 0 | 0 | 0 | 0.0001 | **0.763** | **0.325** |
| HAR-AWRV-LSTM | **1** | **1** | **1** | **1** | **1** | **1** | **1** | **1** |

Note: The numbers in the table show $p$ values for MCS results from 10,000 bootstrap simulations; $p$ values greater than 0.1 appear in boldface.

compared with MFV, MVWA works better. In addition, whatever benchmark volatility, the HAR-LSTM model has the best effect and the slightest prediction error.

## REFERENCES


[1] Torben, G, Andersen, et al, "Intraday and interday volatility in the Japanese stock market,“ *Journal of International Financial Markets, Institutions and Money*, 2000.

[2] Andersen, Torben G., et al, "Modeling and forecasting realized volatility," “ *Econometrica*, vol.71, no.2, pp.579-625, 2003.

[3] Weron, Aleksander, and Rafał Weron, "Fractal market hypothesis and two power-laws," *Chaos, Solitons & Fractals*, vol.11, pp.289-296, 2000.

[4] Wei, Yu, and Peng Wang. "Forecasting volatility of SSEC in Chinese stock market using multifractal analysis," *Physica A: Statistical Mechanics and its Applications*, vol.387, no.7, pp.1585-1592, 2008.

[5] WANG Peng, WANG Jianqiong, "Multifractal Volatility Measure of China's Stock Market and Its Validity," in *Chinese Journal of Management Science*, vol.16, no.06, pp.9-15, 2008.

[6] Chen H, Wu C, "Forecasting volatility in Shanghai and Shenzhen markets based on multifractal analysis," in *Physica A Statistical Mechanics & Its Applications*, vol.390, no.16, pp.2926-2935, 2011.

[7] YUAN Ying, ZHANG Tonghui, ZHUANG Xintian, "Multifractal volatility forecast of Chinese stock market," in *Systems Engineering Theory & Practice*,vol.40, no.09, pp.2269-2281, 2020.

[8] Corsi F, "A Simple Approximate Long-Memory Model of Realized Volatility," in *Journal of Financial Econometrics*, vol.7, no.2, pp.174-196, 2009.

[9] Andersen T G, Bollerslev T, Diebold F X, "Roughing It Up: Including Jump Components in the Measurement, Modeling, and Forecasting of Return Volatility," in *The Review of Economics and Statistics*, vol.89, no.4, pp.701-720, 2007.

[10] GONG Yizhou, HUANG Ran, " A high- frequency volatility model with signed jumps variation based on external information impact," in *Systems Engineering Theory & Practice*, vol.39, no.09, pp.2189-2202, 2019.

[11] ZHANG Tong-hui, YUAN Ying, ZHUANG Xin-tian, "Multifractal Volatility Modeling of Stock Market Considering Jumps and Leverage Effects," in *Journal of Northeastern University (Natural Science)*, vol.41, no.04, pp.594-598, 2020.

[12] LIU Yi, QU Jianwen, DONG Xugao, ZHANG Lei, "The application of signed range to forecasting the volatility of financial," in *Systems Engineering Theory & Practice*, vol.41, no.09, pp.2256-2270, 2021.

[13] Corsi F, D Pirino, R Renò, "Threshold bipower variation and the impact of jumps on volatility forecasting," in *Journal of Econometrics*,

TABLE V
VOLATILITY MODELS' MCS TEST RESULTS (BASED ON MFV)

| | MAE | | MSE | | HMAE | | HMSE | |
|---|---|---|---|---|---|---|---|---|
| | $T_R$ | $T_{SQ}$ | $T_R$ | $T_{SQ}$ | $T_R$ | $T_{SQ}$ | $T_R$ | $T_{SQ}$ |
| HAR-RV-LSTM | 0 | 0 | 0 | 0 | 0 | 0.0319 | 0.0111 | 0.0551 |
| RV-LSTM | 0 | 0 | 0 | 0 | 0 | 0.0234 | 0.021 | 0.0559 |
| HAR-AWRV-LSTM | 0 | 0 | 0 | 0 | 0.0038 | 0.0632 | 0.03 | 0.0359 |
| AWRV-LSTM | 0 | 0 | 0 | 0 | 0.0001 | 0.0318 | 0.0299 | 0.015 |
| HAR-AWRV | 0 | 0 | 0 | 0 | 0 | 0.0168 | 0.0111 | 0.0559 |
| HAR-RV | 0 | 0 | 0 | 0 | 0 | 0.0184 | 0.0175 | 0.0256 |
| HAR-MFV | 0 | 0 | 0.0003 | 0.0004 | **0.1719** | **0.1318** | **0.2741** | **0.159** |
| HAR-MVWA-LSTM | 0 | 0 | 0 | 0.0001 | 0.0719 | 0.018 | 0.0372 | 0.05053 |
| MVWA-LSTM | 0 | 0 | 0 | 0.0001 | 0.0169 | 0.058 | 0.041 | 0.069 |
| HAR-MVWA | 0 | 0 | 0.0003 | 0.0002 | 0.0142 | 0.0298 | 0.07413 | 0.0326 |
| MFV-LSTM | 0.0851 | 0.0849 | 0.0937 | 0.0949 | **0.5649** | **0.6318** | **0.5658** | **0.559** |
| HAR-MFV-LSTM | **1** | **1** | **1** | **1** | **1** | **1** | **1** | **1** |

Note: The numbers in the table show *p* values for MCS results from 10,000 bootstrap simulations; *p* values greater than 0.1 appear in boldface.

TABLE VI
VOLATILITY MODELS' MCS TEST RESULTS (BASED ON MVWA)

| | MAE | | MSE | | HMAE | | HMSE | |
|---|---|---|---|---|---|---|---|---|
| | $T_R$ | $T_{SQ}$ | $T_R$ | $T_{SQ}$ | $T_R$ | $T_{SQ}$ | $T_R$ | $T_{SQ}$ |
| HAR-RV | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| HAR-AWRV | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| HAR-MFV | 0 | 0 | 0 | 0.0002 | 0 | 0 | 0 | 0.0001 |
| HAR-MVWA | 0 | 0 | **0.448** | **0.291** | **0.277** | **0.1421** | **0.126** | 0.079 |
| RV-LSTM | 0 | 0 | 0 | 0 | 0 | 0 | 0.0003 | 0.0001 |
| AWRV-LSTM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| MFV-LSTM | 0 | 0 | 0.0001 | 0.0002 | 0 | 0 | 0.0002 | 0.0001 |
| MVWA-LSTM | 0.0777 | 0.0845 | **0.4516** | **0.4413** | 0.0784 | 0.0798 | **0.3078** | **0.3175** |
| HAR-RV-LSTM | 0 | 0 | 0 | 0 | 0 | 0 | 0.0003 | 0.0001 |
| HAR-AWRV-LSTM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| HAR-MFV-LSTM | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.0001 |
| HAR-MVWA-LSTM | **1** | **1** | **1** | **1** | **1** | **1** | **1** | **1** |

Note: The numbers in the table show *p* values for MCS results from 10,000 bootstrap simulations; *p* values greater than 0.1 appear in boldface.

vol.159, no.2, pp.276-288, 2008.

[14] CHEN Wei-hua, "Forecasting Volatility of Shanghai Composite Index with Deep Learning," in *Statistics & Information Forum*, vol.33, no.05, pp.99-106,2018.

[15] Liu Guangying, Wu Hongchao, Kong Xinbing, "Deep Learning LSTM Model and VaR Risk Management," in *Statistics & Decision*, vol.37, no.08, pp.136-140, 2021.

[16] Kim D, Baek C, Phillips A, "Factor-augmented HAR model improves realized volatility forecasting," in *Applied Economics Letters*, vol.27, no.12, pp.1002-1009, 2020.

[17] Zhang Y, Huang P, Zhou P, et al, " Forecast Volatility based on Realized GARCH and deep LSTM Neural Network," in *2020 International Conference on Public Health and Data Science (ICPHDS)*, pp.99-103, 2020.

[18] Mademlis D K, Dritsakis N, "Volatility Forecasting using Hybrid GARCH Neural Network Models: The Case of the Italian Stock Market," in *International Journal of Economics and Financial Issues*, vol.11, no.1, pp.49, 2021.

[19] YAO Yuan, LIU Zhen-qing, ZHAI Jia, CAO Yi, "Two-component hybrid volatility models on CNY exchange rates," in *Journal of Management Sciences in China*, vol.22, no.11, pp.91-105, 2019.

[20] Cao Yi, Ren Siru, "Research on RMB exchange rate volatility prediction based on LSTM and GARCH family mixed model," in *Application Research of Computers*, vol.37, pp.79-82, 2020.

[21] Liu S, Oosterlee C W, Bohte S M, "Pricing options and computing implied volatilities using neural networks," in *Risks*,vol.7, no.1, pp.16, 2019.

[22] Hansen P R,, Lunde A, Nason J M, "The Model Confidence Set," in *Econometrica*, vol.79, no.2, pp.453-497, 2011.

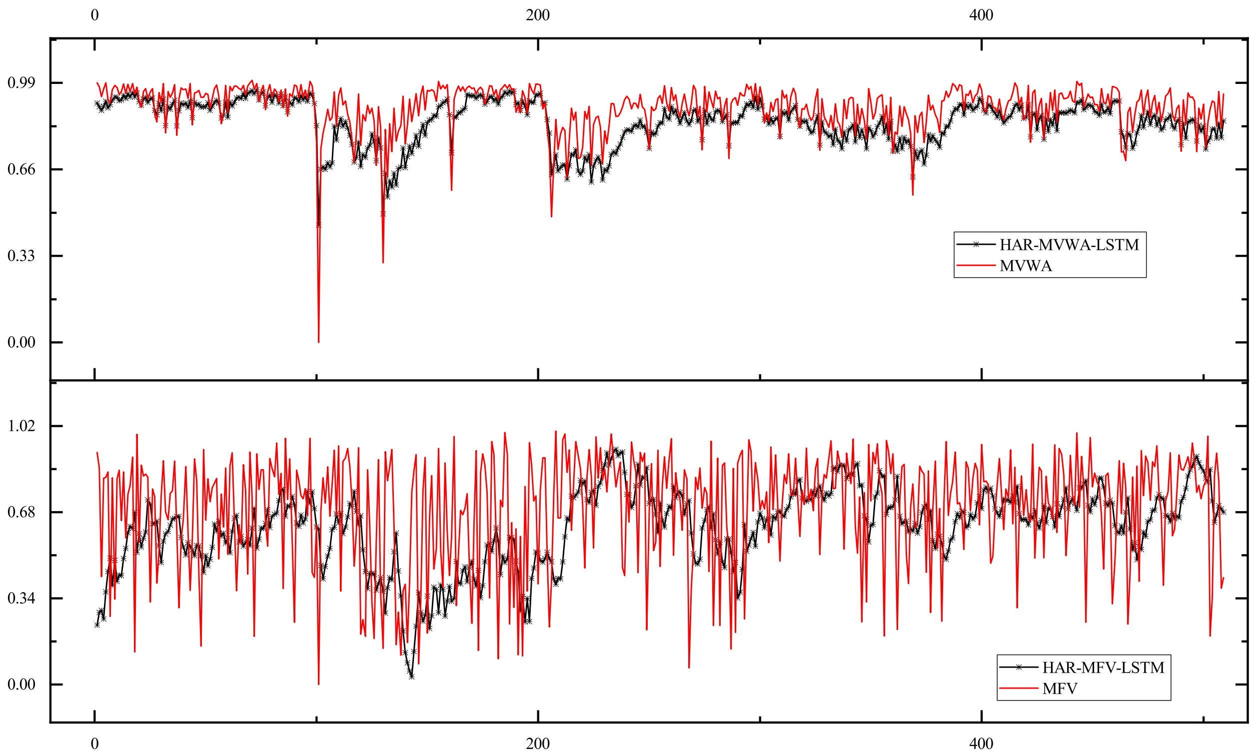


Fig. 4. out-of-sample prediction results