

Yun Ling\*, BaoLong Yue and Hua Zhang

# A New Wrapped Ensemble Approach for Financial Forecast

**Abstract:** The financial market is a highly complex and dynamic system that has great commercial value; thus, many financial elite are drawn to research on the subject. Recent studies show that machine learning methods perform better than traditional statistical ones. In our study, based on the characteristics of financial sequence data, we propose a wrapped ensemble approach using a supervised learning algorithm to predict stock price volatility of China's stock markets. To check our new approach, we developed an intelligent financial forecast system and used the Hushen 300 index data to test our model; it proves that our model performs better than a single algorithm. We also compared our model with the famous ensemble approach of bagging, and the result shows that our model is better.

**Keywords:** Wrapped ensemble approach, stock forecast, intelligent forecast system.

---

**\*Corresponding author: Yun Ling**, Department of Computer Science and Information Engineering, Zhejiang Gongshang University, Hangzhou, Zhejiang 310018, China, e-mail: yling@mail.zjgsu.edu.cn

**BaoLong Yue and Hua Zhang:** Department of Computer Science and Information Engineering, Zhejiang Gongshang University, Hangzhou, Zhejiang 310018, China

## 1 Introduction

Stock market modeling and forecasting are important areas of financial research and application [4], and because of their highly complex and dynamic nature, there is no universal approach available. An approach that works perfectly in the American market may be invalid in the Chinese market, and an approach that operated well on the previous year may not be effective on the current year. However, once a forecast becomes successful in some period, the result will translate to fruitful rewards [24].

Most of the works can be divided into two categories. The first category is based on technical analysis. In actual trading, it is very difficult for people to choose from different kinds of technical indicators [14], and it is also a problem to choose the proper parameters for the chosen indicators [15]. For example, the moving average convergence and divergence (MACD) has three parameters to be chosen depending on different investment vehicles or different periods; however, the parameters are usually chosen based on the experience of the traders [13]; thus, a new trader would feel confused and find it difficult to obtain the best results. Therefore, a number of researchers tried to study how to search for the optimal parameters for technical indicators [9]. In 2006, Chavarnakul and Enke [3] developed an intelligent technical analysis-based equivolume charting in stock. In 2011, Rodríguez-González and García-Crespo [19] used neural networks to improve the financial indicator relative strength index (RSI) and Lin et al. [14] developed an intelligent stock trading system based on improved technical analysis using genetic algorithm (GA) and echo state network (ESN) to promote the technical indicator, such as gold cross and dead cross. Other research studied the profitability of technical analysis based on knowledge [11].

The other category is forecasting the future trend of stocks based on machine learning. An interesting work, which we consider our main reference for this article, was the recent research performed by Teixeira and de Oliveira [21], which combined technical analysis and nearest-neighbor classification to predict future stock trend. Tsai and Hsiao [22] studied in detail the feature selection issue for stock prediction in 2010; they combined multiple feature selection methods such as principal component analysis, GAs, and decision tree to search the perfect feature. Dai et al. [5] made a stock index forecast in 2012 using independent component analysis and back-propagation neural networks. Kwon et al. [11] studied the financial correlation with various stocks and forecasted using recurrent neural networks. In 2007, Kwon and Moon [10] proposed a hybrid approach for stock forecasting, which

distributed loads by parallelization. In 2011, Lin et al. [13] did a forecast using support vector machine (SVM) algorithm, but they used representative prototypes of financial reports to account for the quantitative and qualitative features. Cao and Tay [2] used the SVM algorithm with adaptive parameters to forecast.

Previous studies used either sequence data or indicator data, but with single data, the time series method and learning method cannot be combined. Thus, in our article, we proposed a new ensemble method for stock price forecasting. Our ensemble method is based on the wrap method, which consists of feature selection, model training, and parameter optimization. In feature selection, which is used by many researchers such as Huang and Wang in 2006 [8] and Tsai in 2010 [22], we used GA to make the choice; GA facilitates the choice process using binary coding. For parameter optimization, we used the particle swarm optimization (PSO) algorithm, the same approach used by Cao and Tay [2], to complete the process because it is a simple and quick algorithm with good convergence. To train the model, we used different types of machine learning method and studied which RBF network performs best [10].

The rest of this article is organized as follows: in Section 2, we will introduce the materials and methods used; in Section 3, we will introduce the model we proposed and provide the evaluation indication; our experimental process and results are given in Section 4, where we also give a comparison between different approaches. The conclusions and remarks are provided in the final section.

## 2 Material and Methods

### 2.1 Influence Factors of Stock Price

Stock price means the actual trading price formed through bidding by the buyer and seller in the financial market [15]. Stock price is determined by the law of supply and demand [8]. When supply and demand change, the stock price changes as well, that is, if the supply exceeds the demand, the stock price must fall; if the demand exceeds the supply, the stock price must rise. Therefore, we can see that the supply and demand factors directly affect stock price [23].

The financial market is much too complex and highly dynamic; thus, factors affecting the stock price change are also complex and various. The macroeconomy and stock market operation, the prospect of the industry, and the operating status of a company are the main factors that affect an investor's expectations and decisions, and they all lead to changes in stock price. In addition, political factors, force majeure, the psychological status of investors, national macropolicies, and even manipulative factors contribute to stock price.

### 2.2 Influential Technical Indicator

According to Murphy (1999), technical analysis is the study of market action, primarily using charts, to forecast future price trends. The most important premise of this type of analysis is that market action discounts everything [17], i.e., everything that can affect the market is reflected in the price and volume, more specifically, opening price, closing price, the highest price, the lowest price, and volume; thus, we can predict the future using indicators deduced from the price and volume. However, in this article, we will not judge the trend according to the indicator, but we will use the indicators as inputs for machine learning. Here are four technical indicators are used.

#### 2.2.1 Moving Average

The moving average (MA) is based on the concept of average cost, which is part of the Dow Jones theory. It connects the average price of a period to a line to display the history of the stock price's fluctuations and then reflect the future trends of the stock price. The SMA formula for a period of  $n$  is presented in Formula (1):

$$SMA(t) = \frac{1}{n} \sum_{i=t-n+1}^t x(i). \quad (1)$$

### 2.2.2 Relative Strength Index

The RSI, an indicator proposed by Welles Wilder in 1978 [25], is used to measure the relative strength of securities. It constraints its value to 0 to ~100; generally, if its value is higher than 70, we can speculate that the market is in condition of overbought; conversely, we can get the signal that the market is oversold if its value is lower than 30. The RSI is computed using Formula (2):

$$RSI = \frac{AvgRise}{AvgFall + AvgRise} \times 100, \quad (2)$$

where AvgRise is the average rise value during  $n$  periods and AvgFall is the average fall value during  $n$  periods. Welles Wilder has proved that parameter  $n$  is generally 14. This article also adopts this parameter.

### 2.2.3 Stochastic Oscillator

The term stochastic oscillator, namely, %K and %D, was first used by George C. Lane in 1950 [12]. It was used to forecast time price trend reversal by comparing the range of the close price with the other prices. It is made up with two lines, the fast (%K) and the slow (%D) [Formulas (3) and (4), respectively]. Closing levels that are consistently near the top of the range indicate accumulation (buying pressure) and those near the bottom of the range indicate distribution (selling pressure).

$$\%K = \alpha \cdot RSV_n + (1 - \alpha) \cdot K_{n-1} \quad (3)$$

$$\%D = \alpha \cdot K_n + (1 - \alpha) \cdot D_{n-1} \quad (4)$$

$$RSV = \frac{C_n - L_n}{H_n - L_n} \times 100\% \quad (5)$$

In the above formula, RSV is the raw stochastic value,  $C_n$  is the close price of the  $n^{\text{th}}$  day,  $H_n$  is the highest price of the  $n^{\text{th}}$  day, and  $L_n$  is the lowest price of the  $n^{\text{th}}$  day; in this article, we adopt the most common parameters,  $\alpha = 1/3$  and  $n = 9$ .

### 2.2.4 Moving Average Convergence and Divergence

The MACD indicator, proposed by Gerald Appel in 1970 [1], is a technical tool that is used to judge the intensity, direction, energy, and trend cycle of stock prices and then to support the decision of sell and buy. The MACD often consists of two lines (DEM and DIF) and one chart (an MACD bar). To use the MACD indicator, there are three parameters that have to be set, and in this article, we used 12 and 26 for the calculation of the DIF and 9 for the DEM.

$$DIF = EMA_{(close, low)} - EMA_{(close, fast)} \quad (6)$$

$$DEM = EMA_{(DIF, n)} \quad (7)$$

$$MACD \text{ bar} = DIF - DEM \quad (8)$$

For this indicator, we did not use the MACD bar directly but we took the DIF and the DEM as two input variables.

## 2.3 Stock Price Analysis Methods

### 2.3.1 Genetic Algorithm

The main idea of GA comes from the theory of evolution proposed by Darwin, i.e., in natural selection, only the fittest can survive. It works with a set of candidate solutions called population and generates successive populations of alternate solutions that are represented by a chromosome [7]. If associated with the characteristics of exploitation search, GA can efficiently deal with large search spaces and hence has a less chance of obtaining a local optimal solution compared with other algorithms [8]. In the current work, we adopted the GA algorithm to help obtain the optimal feature combination, as did Huang and Wang [8] in 2006. However, we only used the algorithm for feature selection, not for parameter optimization, because we wanted to use simpler and more rapid methods, such as PSO, to optimize the parameters.

### 2.3.2 Particle Swarm Optimization

PSO is a population-based stochastic optimization technique developed by Eberhart and Kennedy in 1995 [6] and inspired by animal social behavior, as exemplified by a flock of birds or a school of fish. The adjustment toward pbest and gbest by the particle swarm optimizer is conceptually similar to the crossover operation used by GAs. It uses the concept of fitness, as do all evolutionary computation paradigms [20]. Compared with other algorithms, PSO provides faster and better results with lower cost.

The PSO we used adjusts the velocities using the following formula:

$$\begin{aligned} vx[i][j] = & w \times vx[i][j] \\ & + c_1 \times \text{rand}() \times (\text{pbest } x[i][j] - \text{present } x[i][j]) \\ & + c_2 \times \text{rand}() \times (\text{pbest } x[i][\text{gbest}] - \text{present } x[i][j]). \end{aligned} \quad (9)$$

In this article, we used the PSO to achieve better parameters for forecast algorithm, and we set self-learning and group learning coefficients  $c_1$  and  $c_2$  as 0.2, the initial particle number as 60, and the weight coefficient  $w$  as 1.0.

### 2.3.3 RBF Network

The essence of the RBF network is to convert the inputs from one space to another. It is divided into two states, finishing a mapping:  $\xi: R^m \rightarrow R^p$ . The architecture is displayed in Figure 1. Each neuron  $i$  of the hidden layer consists of a center  $\mu_i$  and width  $\sigma_i$ . The Euclidean distance between the input signal  $S$  and the center of the neuron is first evaluated; then an activation function, i.e., the basis function (BF), is applied [8]. The RBF network has the best unique approximation and better convergence speed. In this article, the parameters are wrapped in the model [16].

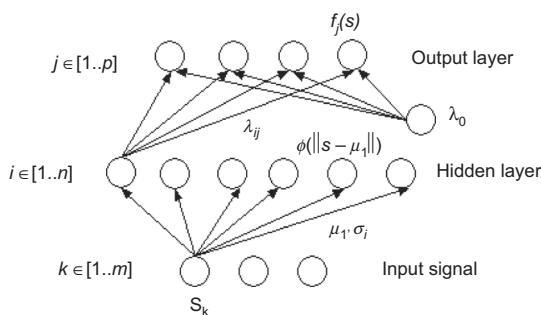


Figure 1. Architecture of an RBF Network.

### 3 Proposed Hybrid Models

#### 3.1 A New Ensemble Method for Stock Data

In previous works, researchers often used the ensemble method, such as bagging, boost, and random forest method, to improve the precision of the single learning method. Yet traditional methods have not considered the characteristics of stock prediction. For that reason, we proposed a new ensemble model considering the specificity of stock forecast.

In a previous study, a single type of market data is often used. Some use only the closing price sequence, which we call sequence data in this work. Others apply all kinds of market indicator data daily as inputs; however, they have not dealt with the single closing price data as a sequence, which we call indicator data in this work. To use both kinds of data fully, we propose a new method to integrate both:

$$\hat{Y}_t = f(\hat{N}_t) + f(\hat{\varepsilon}_t), \quad (10)$$

where  $\hat{Y}_t$  is the yield,  $\hat{N}_t$  is the value of the price, and  $\hat{\varepsilon}_t$  is the residual error.

In our experiment, the indicator data performed well in predicting the direction of stock price change, whereas the sequence data operated well in predicting the yield. Thus, we divided the prediction progress into two parts. Just as in Formula (10), we separated the estimation of yield into two parts, i.e., the estimation of the price value and the estimation of residual error, and in this article,  $f(x)=x$ .

#### 3.2 Data and Inputs

##### 3.2.1 Inputs for Sequence Data

When building the sequence data model, we only use the daily close price sequence data. However, we will not use the close price sequence directly; on the contrary, we will deal with the data using the financial method given in Formula (11) in order to decrease the errors coming from data noise. The method that we adopted was also used by others to make sequence prediction.

$$r_t = \ln(1 + R_{t+1}) = \ln\left(\frac{P_{t+1}}{P_t}\right) = \ln(P_{t+1}) - \ln(P_t) \quad (11)$$

In Formula (11), we did not use the simple return; instead, we used the log return;  $R_{t+1}$  is the simple return of day  $(t+1)$ ,  $P_t$  is the price of day  $t$ . The log return has many advantages compared with the simple return. First, with the log return, it is easier to calculate the multiple-periods return; all we need to do is calculate the simple addition of each period return. Second, the log return has a statistical property that is easy to deal with.

##### 3.2.2 Inputs for Indicator Data

To build the indicator data model, we used the daily open price, highest price, lowest price, and close price data. However, we did not directly use the sequence data because the original data contain a lot of noise; on the contrary, we adopted some technical indicators, which we introduced above. In addition, we also used other indicators adopted in pioneer research. We extracted some features from the price of history data as inputs for each day  $t$ ; some of the inputs were also used in the works of Kwon and Moon [10] and Teixeira and de Oliveira [21].

Parts of the inputs are reported in Table 1, where  $P_o(t)$ ,  $P(t)$ ,  $P_h(t)$ , and  $P_l(t)$  indicates the stock's open, close, high, and low price, respectively, on day  $t$ ;  $V(t)$  is the volume of trades on day  $t$ ;  $Bollinger_{High}$  and  $Bollinger_{Low}$  are the upper and lower Bollinger bands, respectively;  $MA_s$  and  $MA_l$  are the short- and long-term simple moving average applied to price; and  $VMA_s$  and  $VMA_l$  are the simple moving average of short- and long-term forecast applied to volume.

Before using the inputs introduced in Table 1, we also normalized the inputs to avoid the effects between the bigger and the smaller numbers. Here we adopt the Z-score standard method using the following formula:

$$Z_{ij} = \frac{X_{ij} - \bar{X}_i}{\sigma_i}, \quad (12)$$

where  $Z_{ij}$ ,  $X_{ij}$ ,  $\bar{X}_i$ , and  $\sigma_i$  representing the number after normalization, the number before normalization, the average value of the inputs  $i$ , and the standard deviation of inputs  $i$ , respectively.

### 3.2.3 Wrapped Hybrid RBF Network

In a previous study, the forecast and optimization processes are often separated [16], and then, their models are only fitted with the fixed period, i.e., if the periods are changed, their model may fail. Thus, in this article, we propose a wrapped hybrid RBF model (WHR) based on the new ensemble method proposed above. The

**Table 1.** Features Used as Inputs for Indicator Mode.

---

$X_1 = DIF(t)$
$X_2 = DEM(t)$
$X_3 = (DIF(t) - DIF(t-1))/DIF(t-1)$
$X_4 = DEM(t) - DEM(t-1)/DEM(t-1)$
$X_5 = Bollinger_{High}$
$X_6 = Bollinger_{Low}$
$X_7 = (Bollinger_{High}(t) - Bollinger_{High}(t-1))/Bollinger_{High}(t-1)$
$X_8 = (Bollinger_{Low}(t) - Bollinger_{Low}(t-1))/Bollinger_{Low}(t-1)$
$X_9 = (RSI(t) - RSI(t-1))/RSI(t-1)$
$X_{10} = [(P(t) - Po(t)) - (P(t-1) - Po(t-1))]/(P(t-1) - Po(t-1))$
$X_{11} = (P(t) - average(P(t), \dots, P(t-5)))/average(P(t), \dots, P(t-5))$
$X_{12} = (V(t) - average(V(t), \dots, V(t-5)))/average(V(t), \dots, V(t-5))$
$X_{13} = RSI(t)$
$X_{14} = \% K(t)$
$X_{15} = \% D(t)$
$X_{16} = (\% K(t) - \% K(t-1))/\% K(t-1)$
$X_{17} = (\% D(t) - \% D(t-1))/\% D(t-1)$
$X_{18} = (P(t) - P(t-1))/P(t-1)$
$X_{19} = (P(t) - P_l(t-1))/(Ph(t) - P_l(t))$
$X_{20} = (MA_s(t) - MA_s(t-1))/MA_s(t-1)$
$X_{21} = (MA_l(t) - MA_l(t-1))/MA_l(t-1)$
$X_{22} = (MAS(t) - MA_l(t-1))/MA_l(t-1)$
$X_{23} = (P(t) - MA_l(t))/MA_l(t)$
$X_{24} = (P(t) - \min(P(t), \dots, P(t-5)))/\min(P(t), \dots, P(t-5))$
$X_{25} = (P(t) - \max(P(t), \dots, P(t-5)))/\max(P(t), \dots, P(t-5))$
$X_{26} = (V(t) - V(t-1))/V(t-1)$
$X_{27} = (VMA_s(t) - VMA_s(t-1))/VMA_s(t-1)$
$X_{28} = (VMA_l(t) - VMA_l(t-1))/VMA_l(t-1)$
$X_{29} = (VMA_s(t) - VMA_l(t-1))/VMA_l(t-1)$
$X_{30} = (V(t) - VMA_l(t-1))/VMA_l(t-1)$
$X_{31} = (V(t) - \min(V(t), \dots, V(t-5)))/\min(V(t), \dots, V(t-5))$
$X_{32} = (V(t) - \max(V(t), \dots, V(t-5)))/\max(V(t), \dots, V(t-5))$

---

WHR is made up of three parts, the feature selection model, the forecast model, and the parameter optimization model; however, because the three models have a unique object that is subject to features and parameters, the three parts are not independent but are an organic whole. In this model, the features and parameters are all dynamically changed, and each change contributes to a systematic error. The goal in this hybrid model is to search for a relatively optimal global model.

In the feature selection parts shown in Figure 2, we adopted the GA, but unlike other studies that use complex coding, we used common binary coding to represent the features to be included or excluded. Binary coding not only perfectly and simply represents the feature but also improves the efficiency. In our experiment, the chromosome length is 74, which is the number of features, the crossing rate is 0.25, and the mutation rate is 0.01, and it is worth mentioning that we used a mean square error (MSE) as the fitness function of GA.

In the parameter optimization parts shown in Figure 2, the PSO algorithm was used. PSO is a highly efficient search algorithm that has a fast convergence rate and fewer parameters to set, and more importantly, it does not need any encoding or decoding. In our experiment, the self-learning factor and the group learning factor were all set as 2.0, whereas the maximal inertia weight was set as 0.9 and the minimal inertia weight as 0.4. The object function is similar to the fitness function in the GA for the wrapped model with unique training target.

The forecast model in Figure 2 was designed based on the new ensemble method for stock data proposed in Section 3.1. First, we used the indicator RBF model to estimate the future price  $f(\hat{N}_t)$ ; then we took advantage of the time series RBF model to estimate the future residual error; then the final price was obtained using Formula (10).

As a whole, our model uses a uniform evaluation indicator for feature selection and parameter optimization, and the search process consists of two cycling, one for the parameter search and another for the feature search. As for the stop of the search process, it was set at the point when the value of the evaluation indicator does not lessen for a time during the search; the time we used in our experiment is 100.

### 3.2.4 Evaluation Indicator

In a previous study, all types of evaluation indicators were used, and the common evaluation indicators are listed in Formulas (13)–(16), which are mean absolute error (MAE), mean absolute percent error (MAPE), mean square error (MSE), and root mean square error (RMSE). Pai and Lin [18] have applied them to one of their studies on stock price forecast; Lu and Wu [15] also used them in their study on financial forecast in 2011.

$$\text{MAE} = \frac{1}{N} \sum_{t=1}^N |d_t - z_t| \quad (13)$$

$$\text{MAPE} = \frac{100}{N} \sum_{t=1}^N \left| \frac{d_t - z_t}{d_t} \right| \quad (14)$$

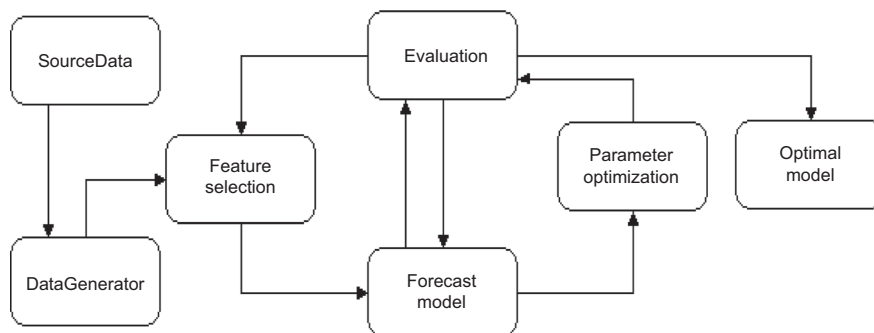


Figure 2. The Wrapped Hybrid Forecast Model.



$$\text{MSE} = \frac{1}{N} \sum_{t=1}^N (d_t - z_t)^2 \quad (15)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (d_t - z_t)^2} \quad (16)$$

However, these traditional evaluation indicators have their limitations; unlike prediction in the traditional field, where we do not need to consider the direction of the prediction result, stock price prediction has its own characteristics and the accuracy of stock price movement direction counts more than traditional prediction. If we ignore the data given above, we may obtain a model that has low system error but is useless. Some researchers also use the direction accuracy indicator and the system error indicator at the same time to overcome the limitations of the traditional indicator; however, their method has changed the problem to a multiobjective optimization problem, thus bringing more problems.

In this article, we designed a new indicator to provide the model with comprehensive evaluation not only to overcome the limitations of the traditional indicator but also to maintain it as a single-objective optimization problem. We call this the weighted mean absolute error (WMAE):

$$\text{WMAE} = \frac{\sum_{\substack{t=1 \\ d_t \times z_t > 0}}^n |d_t - z_t| + \alpha \sum_{\substack{t=1 \\ d_t \times z_t \leq 0}}^{N-n} |d_t - z_t|}{N}, \quad (17)$$

where  $\alpha$  is the penalty factor,  $dt$  is the real value, and  $zt$  is the prediction value. In this article,  $\alpha$  was set as 5.0. Based on the traditional indicator, the WMAE adds a penalty factor to amplify the loss brought about by mistakes in direction judgment.

## 4 Experiment and Comparison

### 4.1 Experiment Designing

#### 4.1.1 Data Set

The data in this article were from the China Shanghai and Shenzhen stock exchange database (<http://chart.yahoo.com/table.csv?s=000300.SS&a=6&b=12&c=1998&d=6&e=9&f=2013&g=d&q=q&y=0&z=000300.SS&x=.csv>). For our experiment, we used 7 years of data (from January 2, 2006, to December 31, 2012).

#### 4.1.2 Sliding Time Window

In our experiment, we used the sliding time window, which is usually used in the analysis of time series data. Tsai and Hsiao [22] have used this method in their study on feature selection. Teixeira and de Oliveira [21] also used this method in stock forecasting, but their window model was static and did not have a dynamic property; thus, their model will only be fit up to a certain period and will not adapt to changes in market situation.

To improve the adaptivity of the model, we tried to improve the traditional window used in most studies. Table 2 displays the detail of the sliding time window, which comprised training data and test data but without validation data. The reason is that the step-forward forecast is different from the multistep-forward forecast, and if we use the data between train data and test data as the validation data, we will unconsciously have some information loss; however, the stock price is usually closely linked with the most recently generated data; thus, information loss is rather great. The approach we proposed, however, avoids that problem. Based on professional experience, the trading data, as reflected by the candlestick chart, will repeat itself,



**Table 2.** Improved Sliding Window.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	
Training set					Test									
	Training set					Test								
		Training set						Test						
				Training set						Test				
					Training set						Test			
					Training set							Test		
						Training set						Test		

and if we consider the train data and validation data as a whole, then we can use the train set to train and check the model. In Table 2, we assumed that the length of the training set  $T$  (which will be automatically determined by the algorithm according to the market) is five, and we can use the data of  $T$  day to train and check the model and make a forecast.

#### 4.1.3 Experimental Procedure and Result Analysis

In our experiment, we used the Hushen 300 index to make a forecast. First, we used sequence data to experiment on the wrapped RBF algorithm and forecast the probable value of yield, and then based on real data, we calculated the residual error, which is defined as MAE [Formula (13)]. Next, we used the indicator data to train and test the model, and to smooth and standardize the data, we used the yield rather than the rate of return as the forecast label. Meanwhile, we constructed the train record for sequence data based on the forecast result of the value of yield, then trained and tested the model, and then forecasted the residual error of indicator forecast. Lastly, the forecast value of the ensemble model was calculated by the forecast data and the forecast residual error. According to Formula (10), we separated the forecast value into two parts. In our experiment, we used the indicator data to forecast the value of yield and the sequence data to forecast the residual error of the earlier indicator forecast.

To make a comparison, we also performed the experiment using a single method, without using the wrapped method or the ensemble method. We used a diagram to visualize the forecast data, and the results are shown in Figure 3, which displays the volatility of real and forecast data.

However, if solely based on Figure 3, we could not determine the better result. Thus, to determine which result is better, we introduced a new concept, the cumulative absolute error (CAE):

$$CAE_t = \sum_{i=1}^t |d_i - z_i| \quad (18)$$

Using Formula (18), we can draw the CAE map (Figure 4), which shows the CAEs of the different approaches. Using the map, we can see that the wrapped ensemble method (WEM) is better than other models.

To see the forecast of the real price, we need to convert volatility to price. The calculation method of price, given in Formula (19), is opposite to the method of volatility:

$$P_{t+1} = e^{[r_t + \ln(P_t)]}, \quad (19)$$

where  $P_{t+1}$  is the future price,  $r_t$  is the volatility, and  $P_t$  is the latest price.

The comparisons of model precision are given in Table 3, which shows that the WEM has the best performance. The sequence wrapped RBF (SWRBF) and the indicator wrapped RBF perform better than the sequence wrapped SVR (SWSVR) and the indicator wrapped MLP (IWMLP), and this is the reason why we used the RBF network algorithm to build the ensemble model. Our experiment proved that the wrapped

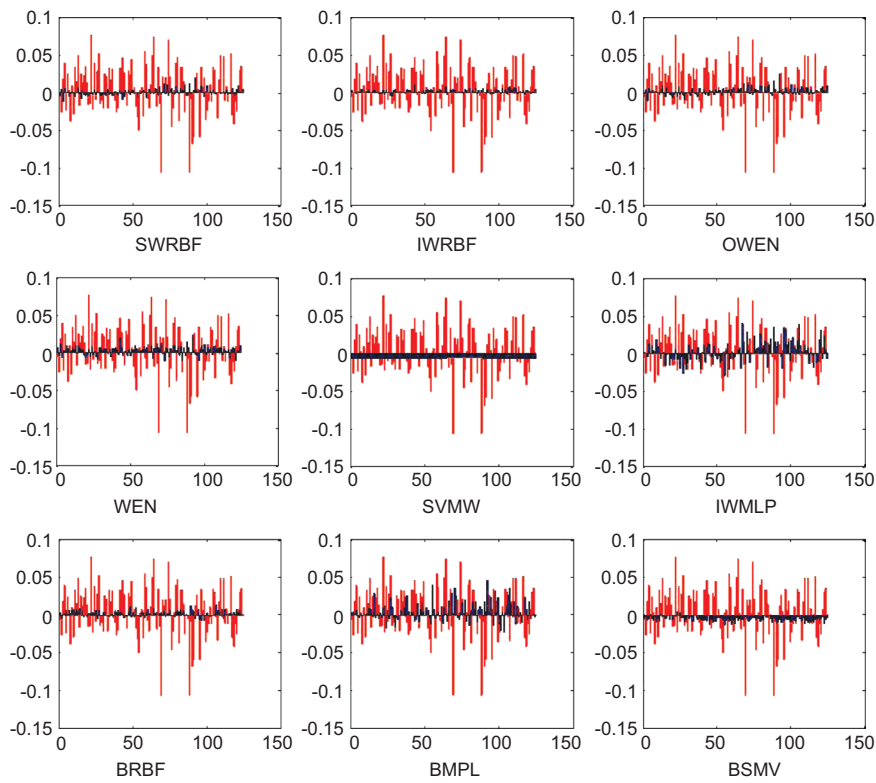


Figure 3. Volatility Forecast Map.

ensemble model performs better than the model without ensemble and that the wrapped model has a better parameter. In addition, in comparison with the bagging model, i.e., the bagging of RBF, MLP, and SVM, we can see that our model is still superior.

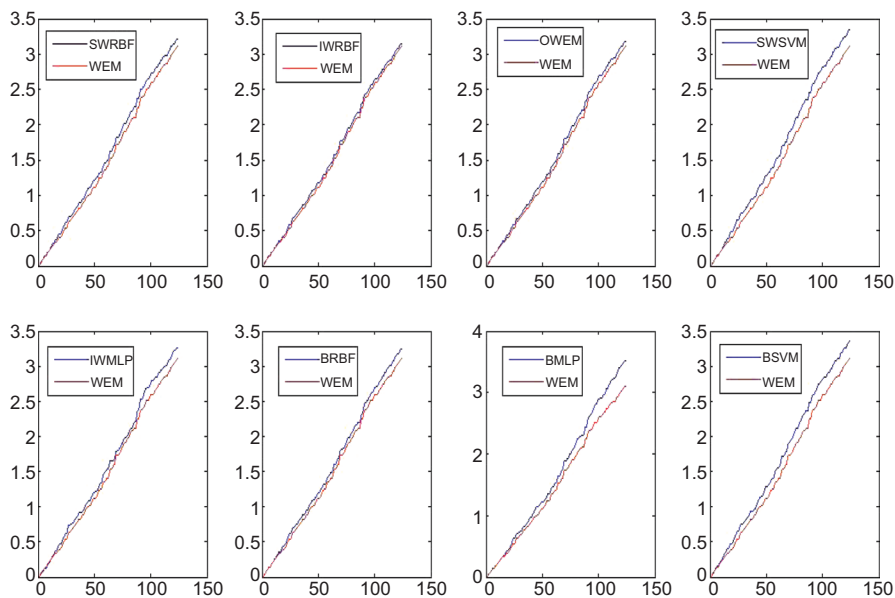


Figure 4. Cumulative Absolute Errors Map.

Table 3. Experiment Result and Comparison.

Number	Model	MAE	MSE	RMSE	MAPE	WMAE
1	SWRBF	0.025916	0.033058	0.001093	120.8282	0.156030
2	IWRBF	0.025379	0.032050	0.001027	108.1462	0.120064
3	OWEM	0.025643	0.032742	0.001072	121.5405	0.155867
4	<b>WEM</b>	<b>0.025065</b>	<b>0.031978</b>	<b>0.001023</b>	<b>106.7072</b>	<b>0.110801</b>
5	SWSVM	0.027007	0.033885	0.001148	157.2586	0.195109
6	IWMLP	0.026320	0.034858	0.001215	205.3762	0.175098
7	BRBF	0.025652	0.032380	0.001048	121.3737	0.146843
8	BMLP	0.028405	0.035110	0.001232	195.1284	0.184346
9	BSVM	0.027041	0.033725	0.001137	155.9387	0.194475

## 5 Conclusions

In this article, we have proposed a new method, the WEM, for financial forecasting. It not only supplies the model with better parameters but also gives the model self-adaptivity. Six groups of experiments on Hushen 300 data proved that our approach performed rather well in forecasting stock price. In addition, we proposed a new systematic error evaluation formula that considers the forecast of direction, and the results of the experiment show that the evaluation indicator can give a reasonably comprehensive evaluation of financial forecast. However, there are still much to do in future work. We will try more algorithm groups and combine more financial business and market characteristics to further improve the performance of our model.

**Acknowledgments:** This work was supported by the National Natural Science Foundation of China (grant no. 61170099), the Science Technology Department of Zhejiang Province (grant no. 2012R10041-18) and the Postgraduate Research Innovation Fund of Zhejiang GongShang University.

Received January 31, 2013; previously published online July 26, 2013.

## Bibliography

- [1] G. Appel, The Moving Average Convergence-Divergence Trading Method[M], Traders Pr, 1985.
- [2] L. J. Cao and F. E. H. Tay, Support vector machine with adaptive parameters in financial time series forecasting, *IEEE Trans. Neural Netw.* **14** (2003), 1506–1518.
- [3] T. Chavarnakul and D. Enke, Intelligent technical analysis based equivolume charting for stock trading using neural networks, *Expert Syst. Appl.* **34** (2008), 1004–1017.
- [4] W. Cui, A. Brabazon and M. O'Neill, Efficient trade execution using a genetic algorithm in an order book based artificial stock market, in: *GECCO '09, Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation Conference: Late Breaking Papers*, pp. 2023–2028, 2009.
- [5] W. Dai, J. Wu and C. Lu, Combining nonlinear independent component analysis and neural network for the prediction of Asian stock market indexes, *Expert Syst. Appl.* **39** (2012), 4444–4452.
- [6] Eberhart and Kennedy, *Swarm Intelligence*[M], Morgan Kaufman, 2011.
- [7] J. H. Holland, *Adaptation in Natural and Artificial Systems*, MIT Press, Cambridge, MA, 1975.
- [8] C. Huang and C. Wang, A GA-based feature selection and parameters optimization for support vector machines, *Expert Syst. Appl.* **31** (2006), 231–240.
- [9] M. Kordos and A. Cwiok, A new approach to neural network based stock trading strategy, *Intell. Data Eng. Autom. Learn. (IDEAL)* **69** (2011), 429–436.
- [10] Y. Kwon and B. Moon, A hybrid neurogenetic approach for stock forecasting, *IEEE Trans. Neural Netw.* **18** (2007), 851–864.
- [11] Y. Kwon, S. Choi and B. Moon, Stock prediction based on financial correlation, in: *GECCO '05, Proceedings of the 2005 Conference on Genetic and Evolutionary Computation*, pp. 2061–2066, 2005.
- [12] G. C. Lane, *Using Stochastics, Cycles & R.S.I. to the Moment of Decision*[M], 1986.

- [13] M.-C. Lin, A. J. T. Lee, R.-T. Kao and K.-T. Chen, Stock price movement prediction using representative prototypes of financial reports, *ACM Trans. Manage. Inform. Syst.* **2.3** (2011), 19.
- [14] X. Lin, Z. Yang and Y. Song, Intelligent stock trading system based on improved technical analysis and echo state network, *Expert Syst. Appl.* **38** (2011), 11347–11354.
- [15] C. Lu and J. Wu, *Predicting stock index using an integrated model of NLICA, SVR and PSO*, vol. 33, pp. 228–237, Springer-Verlag, 2011.
- [16] E. P. Maillard and D. Gueriot, RBF neural network, basis functions and genetic algorithm, in: *International Conference on Neural Networks*, vol. 4, pp. 2187–2192, IEEE, 1997.
- [17] J. J. Murphy, *Technical Analysis of the Financial Markets*[M], Prentice Hall Press, 1999.
- [18] P. Pai and C. Lin, A hybrid ARIMA and support vector machines model in stock price forecasting, *Omega* **24** (2004), 497–505.
- [19] A. Rodríguez-González, Á. García-Crespo, and R. Colomo-Palacios, et al. CAST: using neural networks to improve trading systems based on technical analysis by means of the RSI financial indicator, *Expert Syst. Appl.* **38** (2011), 11489–11500.
- [20] K. V. Sujatha and S. M. Sundaram, A combined PCA-MLP model for predicting stock index, in: *A2CWIC '10, Proceedings of the 1st Amrita ACM-W Celebration on Women in Computing in India*, 2010.
- [21] L. A. Teixeira and A. L. I. de Oliveira, A method for automatic stock trading combining technical analysis and nearest neighbor classification, *Expert Syst. Appl.* **37** (2010), 6885–6890.
- [22] C. Tsai and Y. Hsiao, Combining multiple feature selection methods for stock prediction: union, intersection, and multi-intersection approaches, *Decis. Support Syst.* **50** (2010), 258–269.
- [23] L. Wang and J. Wu, Neural network ensemble model using PPR and LS-SVR for stock market forecasting, *Adv. Intell. Comput.* **6838** (2012), 1–8.
- [24] Q. Wen, Z. Yang, Y. Song and P. Jia, Automatic stock decision support system based on box theory and SVM algorithm, *Expert Syst. Appl.* **37** (2010), 1015–1022.
- [25] J. W. Wilder, *New Concepts in Technical Trading Systems*[M], Trend Research, 1978.