Combination forecasting method based on the fractal dimension weight

Jiran Zhu[1], Yuancan Xu[2], Hua Leng[1], Haiguo Tang[1], Hanyang Gong[1], Zhidan Zhang[1] and Pei Ao[3,a]

[1]*Electric Power Corporation Research Institute, The State Grid of Hunan, China*
[2]*Electrical and Information Engineering Institute, Changsha University of Science and Technology, China*
[3]*College of Electronic and Electrical Engineering, Shanghai University of Engineering, China*

Abstract. In order to improve the prediction accuracy, a combined forecasting method based on fractal dimension weight is proposed in this paper. Firstly, since the amount of the original data will affect the accuracy of forecasting, the three spline interpolation method is used to increase the amount of data. Secondly, historical data fitting values is obtained by the unbiased grey forecasting model, the SVM regression forecasting model and the BP neural network model. According to these fitting values, the box dimension of every single forecasting model is calculated. The box dimension normalization results are taken as the weights of single forecasting model. Finally, the results of single forecasting models are combined by using the weighted average method. Verified by an example, the proposed combined forecasting method has higher accuracy than the single forecasting models.

Keywords: fractal dimension; unbiased grey forecasting model; SVM regression forecasting model; BP neural network forecasting model; combination forecasting model.

1 Introduction

Water quality prediction is using a certain prediction model or method to calculate the future possible values of indicators, or to predict the trend of water quality indicators over time based on historical data of water quality [1, 2]. The commonly used models and methods include time series model, regression analysis, grey system model and neural network model. Because of the different modeling mechanism and the different conditions of the single models, these models will have limitations. In 1969, Bates proposed combination forecasting idea [3]. A comprehensive prediction model is obtained by combining several forecasting models and methods in order to improve the prediction effects. In recent years, many scholars have also introduced the combination forecasting method to predict water quality problems. Hybrid neural network and ARIMA model was used by Durdu Omer Faruk to predict water quality of Menderes River in Turkey and the higher prediction accuracy was achieved [4]. Combined with wavelet transform and BP neural network, Xu established short-term wavelet neural network model to predict the water quality of freshwater pearl cultivation pond [5]. The principal component analysis (PCA), genetic algorithm (GA) and BP neural network are combined to predict the water quality of Taihu Lake by Ding [6]. Combined with gray model and fuzzy neural network prediction model, Zhang Ying predicted the water quality of Taihu Lake [7]. The results

---

[a] Corresponding author : 34204339@qq.com

show that, the proposed method is feasible to predict the change trend of lake water quality, and is superior to the BP neural network model.

The so-called combination forecasting is to combine the different models, use various information provided by different prediction methods comprehensively, and obtain combination forecasting model through the appropriate weighted average form [8-10]. The most concerned about is how to calculate the weighted average coefficient in combination forecasting. This can more effectively improve the prediction accuracy. According to the fitting values obtained by the single forecasting models, the box dimension of every single forecasting model is calculated based on fractal theory in this paper. The box dimension normalization results are taken as the weights of single forecasting model to improve the prediction accuracy. The proposed method is used to predict the concentration of conventional pollutants DO, COD, and NH3-Nat. Compare with the results of the single forecasting models, the proposed method can more effectively improve the prediction accuracy.

2 Single forecasting model

2.1 Unbiased grey forecasting algorithm

Unbiased grey forecasting model is an improvement based on the traditional grey forecasting model. It eliminates the defects that the traditional model is invalid when the original data sequence growth rate becomes larger.

Assume that the original data sequence is $X^{(0)} = \left\{ x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)} \right\}$, where $x_k^{(0)} \geq 0 (k = 1,2,\ldots,n)$, the modelling steps of unbiased grey forecasting model are shown as follows.

Step 1: The original data sequence is accumulated and $X^{(1)} = \left\{ x_1^{(1)}, x_2^{(1)}, \ldots, x_n^{(1)} \right\}$ is generated, where $x_k^{(1)} = \sum_{i=1}^{k} x_i^{(0)}, k = 1,2,\ldots,n$. The $X^{(1)}$ can be obtained by solving the differential equation Eq.(1).

$$\frac{dx_k^{(1)}}{dt} + ax_k^{(1)} = u \tag{1}$$

Step 2: B and Y_n are determined by $B = \begin{bmatrix} -\frac{1}{2}\left(x_1^{(1)} + x_2^{(1)} \right) & 1 \\ -\frac{1}{2}\left(x_2^{(1)} + x_3^{(1)} \right) & 1 \\ \vdots & \vdots \\ -\frac{1}{2}\left(x_{n-1}^{(1)} + x_n^{(1)} \right) & 1 \end{bmatrix}$ and $Y_n = \begin{bmatrix} x_2^{(0)} \\ x_3^{(0)} \\ \vdots \\ x_n^{(0)} \end{bmatrix}$.

Step 3: The parameters a and u is obtained by using the least squares method according to Eq.(2).

$$\begin{bmatrix} a \\ u \end{bmatrix} = \left(B^T B \right)^{-1} B^T Y_n \tag{2}$$

Step 4: The parameters b and A of unbiased grey model is calculated by $b = \ln \frac{2-a}{2+a}$ and $A = \ln \frac{2u}{2+a}$.

Step 5: The forecasting model is built according to Eq.(3).

$$\hat{x}_1^{(0)} = x_1^{(0)}, \quad \hat{x}_{k+1}^{(0)} = Ae^{bk} \tag{3}$$

When $k = 1,2,\ldots,n-1$, $\hat{x}_{k+1}^{(0)}$ is the fitting value of the original data sequence and $\hat{x}_{k+1}^{(0)} (k \geq n)$ is the predicted value of the original data sequence.

2.2 SVM regression forecasting algorithm

The basic idea of SVM (Support Vector Machine, SVM) regression forecasting algorithm is that the data x of input space are mapped into a high-dimensional feature space through a nonlinear mapping ϕ. Linear regression is done to realize the function fitting in this feature space, and extrapolation forecast can be realized based on this fitting function. The prediction steps of SVM regression forecasting algorithm are shown as follows.

Step 1: Assume that $\{(x_i, y_i), i = 1,2,...,l\}$ is the known training sample sets and $y_i \in R$ is expected output, where $x_i \in R^n$. The appropriate parameter $\varepsilon > 0$, $C > 0$ and the kernel function $K(x, x')$ are selected. The optimal solution $\hat{\alpha}^{(*)} = (\hat{\alpha}_1, \hat{\alpha}_1^{(*)}, \hat{\alpha}_2, \hat{\alpha}_2^{(*)}, ..., \hat{\alpha}_l, \hat{\alpha}_l^{(*)})^T$ is calculated by Eq. (4).

$$\min_{\alpha^{(*)} \in R^{2l}} \frac{1}{2} \sum_{i,j=1}^{l} (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(x_i, x_j) + \varepsilon \sum_{i=1}^{l} (\alpha_i^* + \alpha_i) - \sum_{i=1}^{l} y_i (\alpha_i^* - \alpha_i)$$

$$s.t. \sum_{i=1}^{l} (\alpha_i^* - \alpha_i) = 0$$

$$0 \le \alpha_i, \alpha_i^* \le \frac{C}{l}, i = 1,2,...,l$$

(4)

Step 2: A positive component $\hat{\alpha}$ is selected, where $0 < \hat{\alpha}_j < C/l$. Calculation is done according to Eq.(5).

$$\hat{b} = y_i - \sum_{j=1}^{l} (\hat{\alpha}_j^* - \hat{\alpha}_j) K(x_i \bullet x_j) + \varepsilon$$

(5)

Or a positive component $\hat{\alpha}^{(*)}$ is selected, where $0 < \hat{\alpha}_j^* < C/l$. Calculation is done is done according to Eq.(6).

$$\hat{b} = y_i - \sum_{j=1}^{l} (\hat{\alpha}_j^* - \hat{\alpha}_j) K(x_i \bullet x_j) - \varepsilon$$

(6)

Step 3: A linear regression function is constructed according to Eq.(7).

$$f(x) = \sum_{i=1}^{l} (\hat{\alpha}_i^* - \hat{\alpha}_i) K(x_i \bullet x) + \hat{b}$$

(7)

Based on this regression function, the function of extrapolation prediction can be realized.

2.3 BP neural network forecasting algorithm

The basic idea of the BP neural network prediction is that the coming M point data are predicted by using the bypast N point data. The L sample data are divided into K overlapping data segments that the length of every segment is $N + M$. $X_k, X_{k+1}, ..., X_{k+N-1}$ are inputs, $X_{N+k}, X_{N+k+1}, ..., X_{N+M+k+N-1}$ are outputs and $k = 1,2,...,K$. $L - (N+M) - 1$ samples are obtained. Sample data of M-dimensional space are mapped into N-dimensional space by a non-linear mapping. And the network can be used to training and prediction. The prediction steps are shown as follows.

Step 1: The input data are standardized.

Step 2: Related node parameters are designed in the hidden layer, the input layer and output layer.

Step 3: The original data are divided into two parts. One is training data and the other is testing data.

Step 4: Neural network learning is done by using the training data.

Step 5: The trained network is used to predict the real data and the predict results are used to compare the prediction error.

3 Determination of the combined forecasting weights based on fractal dimension

Fractal theory is a law science that it describes many irregular things. It makes people to deal with problems by the new concept in the field of nonlinear. It can reveal the law behind the complex phenomenon, intrinsic correlation between the local and global nature. As the basic parameters and quantitative characterization of fractal, fractal dimension [11-15] is the important principle of fractal theory. The common fractal dimensions are similar dimension, Hausdorff dimension, box dimension, information dimension and correlation dimension.

In the box counting method, the r size box is used to cover the data sets. The counts of non-empty box are C_r for sufficiently small r and the correlation fractal dimension is shown in Eq (8).

$$D = \frac{\partial}{\partial \ln r} \left(\ln \sum_i C_{r,i}^2 \right), r \in [r_1, r_2]$$

(8)

Where $C_{r,i}$ is the number of data points fell into the lattice and $[r_1, r_2]$ is non-scaling space for the data sets. Due to the data sets has a finite number and they are only in $[r_1, r_2]$. They have the statistical self-similarity. The interval $[r_1, r_2]$ is non-scaling space. Dimension is the logarithmic rate of the increasing C_r when $r \to 0$. Therefore, the slope of double logarithm curve can be used to approximate fractal dimension. That is, $\lg r \to \lg C_r$ diagram is drawn through exchanging r to obtain a series of different C_r. Where $\lg r$ is the horizontal coordinate of the diagram and $\lg C_r$ is the vertical coordinate of the diagram. A straight line through these points is fitted and the slope of this line is obtained. Opposite number of slope is the box dimension.

In the combination forecasting, weights have an important impact on the results of the prediction. Considering the characteristics of fractal dimension, the fitting values of historical data are obtained by forecasting methods and the fractal dimension can be calculated. The normalization result of the fractal dimension is taken as the weight of each forecasting method. Assuming that there are l forecasting methods, D_i is the box dimension for the i th forecasting method , w_i is the weight for the i th forecasting method, then w_i can be determined according to Eq.(9).

$$w_i = D_i / \sum_{i=1}^{l} D_i$$

(9)

4 Combination forecasting method based on the fractal dimension weight

In forecasting, the recent historical data have stronger correlation with the future development trend and the long-term historical data have weaker correlation with the future development trend. Here, the idea of equal-dimension new information is combined with combination forecasting method. Namely after a predictive value is obtained by using combination forecasting method, this predictive value is taken as known value to add into the historical data. The flow chart of combination forecasting method is shown in Figure.1. The steps of combination forecasting method based on the fractal dimension weight are shown as follows.
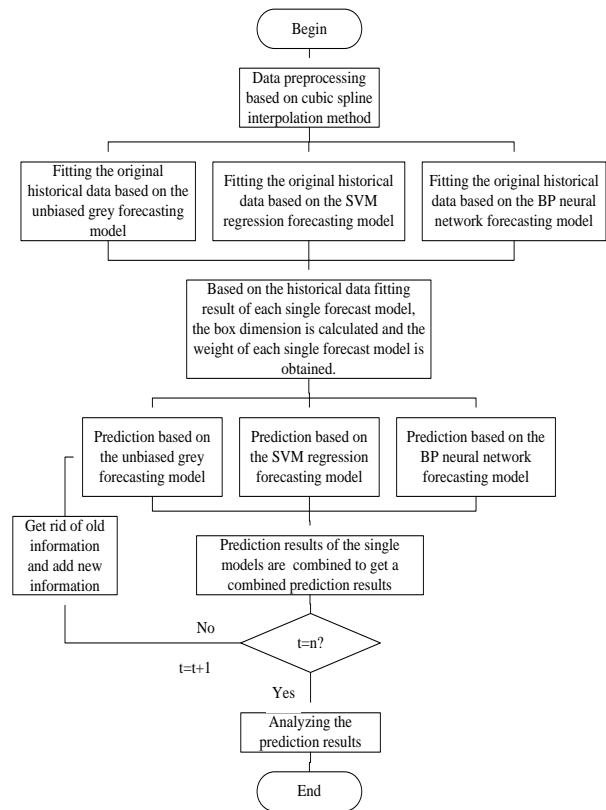
Figure 1. The flow chart for combination forecasting

Step 1: Data are pre-processed. In the actual process of forecasting, the number of the original historical data is often less. In order to improve the accuracy of single model, cubic spline interpolation method can be used to interpolate data into the original data to increase the amount of data. The discrete sampling points $i(t_0)$, $i(t_1)$,…, $i(t_n)$ of function $i(t)$ at the interval $[a,b]$ are known. $a = t_0 < t_1 < \cdots < t_n = b$ are the corresponding sampling time points. h_0, h_1, …, h_{n-1} are the sampling interval. Cubic spline interpolation polynomial $s_j(t) = a_j t^3 + b_j t^2 + c_j t + d_j$ is taken as the approximation of $i(t)$, where the undetermined coefficients a_j, b_j, c_j, d_j can be determined according to the interpolation conditions and the boundary conditions.

Step 2: The single forecasting model is used to fit the original historical data. The unbiased grey forecasting model, the SVM regression forecasting model and the BP neural network forecasting model described in section 1 are used to fit the original historical data.

Step 3: Weights of combination forecasting are determined. Based on the historical data fitting result of each single forecast model, the box dimension is calculated according to Eq.(8) and the weight of each single forecast model is obtained according to Eq.(9).

Step 4: The single forecasting model is used for one step prediction.

Step 5: The weights of combination forecasting is calculated based on Step 3. Then one step prediction results of the single models will be combined to get a combined prediction results based on the weighted average method.

Step 6: Prediction condition is judged. If the condition is not reached, the prediction values obtained from Step 5 are taken as the known data to add into the historical data series and to remove an earliest historical data. Then it goes to Step 4, predict values on each prediction time point are calculated circularly. If the condition is reached, it goes to Step 7.

Step 7: The prediction is completed and the prediction results are analysed. The prediction is over.

5 Combination forecasting for water quality based on the fractal dimension weight

Prediction for river water quality has an important meaning to river water environmental planning, evaluation and management. In this paper, the concentration values of the conventional pollutants DO, COD, NH3-N at the outlet section of Guanting reservoir of Yongding River are taken as the research objects. The original historical concentration values from 21st week to the 40th week in 2010 are selected to build forecasting model. The concentration values from 41st week to the 50th week are taken as prediction samples.

According to the steps of combination forecasting based on the fractal dimension weight in Section 3, firstly, cubic spline interpolation method is used to pre-process the original data and get 400 points concentration values of conventional pollutants DO, COD, NH3-N. They are shown in Fig.2-Fig.4. "○" is used to represent the original data.
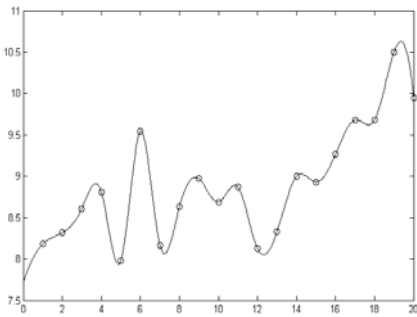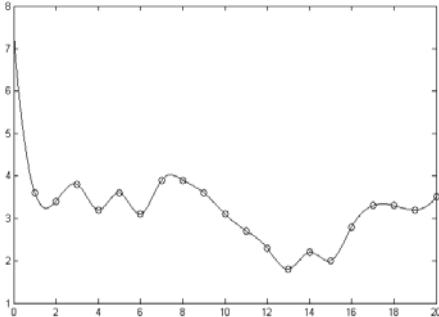
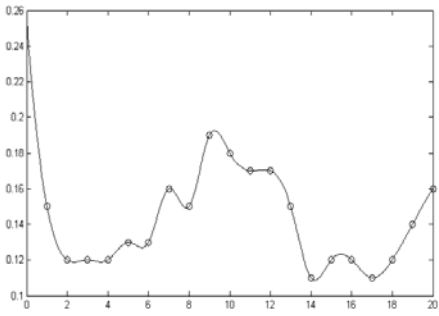**Figure 2.** Preprocessed DO data   **Figure 3.** Preprocessed COD data

Figure 4. Preprocessed NH3-N data

The unbiased grey forecasting model (model1) and BP neural network forecasting model (model2) and SVM regression forecasting model (Model3) are respectively used to fit the original historical data. The box dimension of the single forecasting model is obtained according to Eq. (8) based on the fitting results for prediction. According to Eq. (9), the weight of the single forecasting model is obtained. The box dimensions and the weights are shown in Tab. 1 and Tab. 2.

The prediction from 41st week to the 50th week is completed according to Step 4- 6 in Section 3. The prediction results are shown in Figure.5- Figure.7.

Table 1. Box dimensions of the single forecasting models

Box dimension of the single forecasting model / Index	Box dimension		
	Model1	Model2	Model3
DO	1.3612	1.3208	1.2454
COD	1.3612	1.3208	1.2454
NH3-N	1.0931	1.0931	1.0931

Table 2. Weights of the single forecasting models

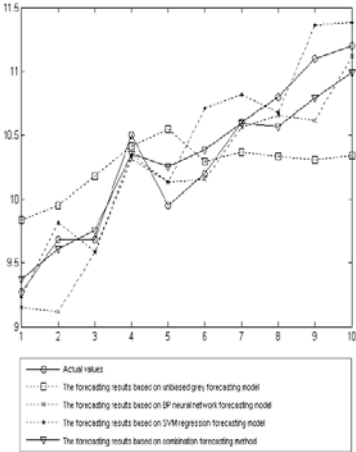| Weight of the single forecasting model / Index	weight		
	Model1	Model2	Model3
DO	0.2866	0.3567	0.3567
COD	0.2926	0.3537	0.3537
NH3-N	0.3050	0.3475	0.3475

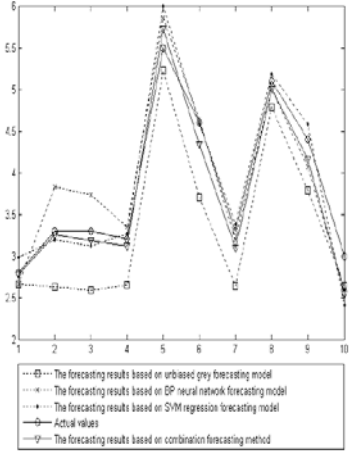**Figure 5.** Forecasting results for DO data
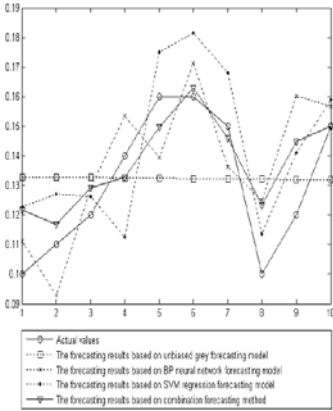

Figure 6. Forecasting results for COD data

Figure 7. Forecasting results for NH3-N data

The sum of the squared errors (SSE) is used to measure the prediction accuracy and shown in Tab. 3. From Tab. 3, it shows that the concentration value SSE of DO, COD and NH3-N are respectively 0.3620, 0.4509 and 0.0019. These results are higher than the other single forecasting models.

Table 3. SSE for the single forecasting models

Forecasting model / Index	Model1	Model2	Model3	Combination Forecasting
DO	2.6588	0.6726	0.5102	0.3620
COD	3.2026	0.9260	0.7195	0.4509
NH3-N	0.0052	0.0033	0.0033	0.0019

6 Conclusion

Since weights have an important impact on the results of the prediction in the combination forecasting, a combination forecasting method based on the fractal dimension weight is proposed to improve the predictive accuracy in this paper. According to the fitting values of the single forecasting model, the box dimension of every single forecasting model is calculated and taken as the weights of single forecasting models. Using the weighted average method, the results of single forecasting models are combined. Through predicting the concentration of conventional pollutants DO, COD and NH3-N, the prediction results are showed that SSE of the proposed method in this paper is lower than the other single forecasting models. So the effectiveness of the proposed method in this paper is verified.

Acknowledgment

References

1. THOMANN R V. The future "Golden Age" of predictive models for surface water quality and ecosystem management [J]. Journal of environmental engineering, **124**(2):94-103, (1998).
2. HIRSCH R M, SLICK J R, SMITH R A. Techniques of trend analysis for monthly water quality data [J]. Water resources research, **18**(1):107-121, (1982).
3. GRANGER C W J, BATES J. The combination of forecasts [J]. Operations research quarterly, **20**(4):451-468, (1969).
4. FARUK D O. A hybrid neural network and arima model for water quality time series prediction [J]. Engineering applications of artificial intelligence, **23**(4):586-594, (2010).
5. XU L Q, LIU S Y. Study of short-term water quality prediction model based on wavelet neural network [J]. Mathematical and computer modelling, **58**(3/4):807-813, (2013).
6. DING Y R, CAI Y J, SUN P D, et al. The user of combined neural networks and genetic algorithms for prediction of river water quality [J]. Journal of applied research and technology, **12**(3):493-499, (2014).
7. ZHANG Ying, GAO Qian –qian. Comprehensive prediction model of water quality based on Grey Model and Fuzzy Neural Network. Chinese Journal of Environmental Engineering, **9**(2):537-545, (2015).
8. S Yin, L Liu, J Hou A multivariate statistical combination forecasting method for product quality evaluation. Information Sciences, (2016), 355-356:229-236.
9. JX Che. Optimal sub-models selection algorithm for combination forecasting model. Neurocomputing, **151**:364-375, (2015).
10. R Adhikari. A neural network based linear ensemble framework for time series forecasting. Neurocomputing, **157**(C):231-242, (2015).
11. Zähle, M. The average fractal dimension and projections of measures and sets in $bold Rsp n$bold Rsp n. Annales De L Institut Henri Poincaré Probabilités Et Statistiques, **2015**(4):747-754.
12. B Ghanbarian, H Daigle. Fractal dimension of soil fragment mass-size distribution: A critical analysis. Geoderma, **245**:98-103, (2015).
13. Y Feng, Y Liu. Fractal dimension as an indicator for quantifying the effects of changing spatial scales on landscape metrics. Ecological Indicators, **53**:18-27, (2015).
14. MJ Lawrence, A Sabra, P Thomas, et al. Fractal dimension: A novel clot microstructure biomarker use in ST elevation myocardial infarction patients.Atherosclerosis, **240**(2):402-407, (2015).

15. Y Zhang, W Zhou, S Yuan, et al. Seizure detection method based on fractal dimension and gradient boosting. Epilepsy & Behavior, **43**(43C):30-38, (2015).