

Combine Multi-predictor of Gas Concentration Prediction Based on Wavelet Transforms

Wu Xiang^{*1,2}, Qian Jian-Sheng¹

¹School of Information and Electrical Engineering, China University Mining & Technology, Xuzhou, Jiangsu 221116, China

²School of Medical Informatics, Xuzhou Medical College, Xuzhou, Jiangsu 221116, China
Email: dmshxzmc@163.com

Abstract

A method of combine multi-predictor is proposed based on wavelet transform to improve the prediction precision of coal mine gas concentration time series. Firstly, the proposed model employ Mallat algorithm to decompose and reconstruct the gas concentration time series to isolate the low-frequency and high-frequency information. Then, ARMA model is built for the prediction of high-frequency information and rectifies deviations of the predicted values by Markov bias correction method while the SVM model is used to fit the prediction of the low-frequency information. At last, these predicted values are superimposed to obtain the predicted values of the original sequence. This method makes an effective separation of the feature information of gas concentration time series and takes full advantage of the features of different prediction models to achieve complementary advantages. The comparison experiment with the single-predictor models (BP, SVM) and single-predictor models based on wavelet decomposition (W-BP, W-SVM) show that the proposed method improves the overall prediction precision. The results show that method has high precision and strong practicability.

Keywords: wavelet transforms, multi-predictor, combine prediction, deviation correction

Copyright © 2014 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

Coal mine gas concentration exceeded discharge standard is one of the most important factors of coalmine security in production. The accurate prediction and real-time monitoring and control of coalmine gas concentration are important measures to prevent gas explosion and gas outburst [1]. Since safety monitoring system has been installed in most coal mine, there is a large number of coal mine safety data. Therefore, it is significant to take advantage of these historical to make an accurate and effective trends prediction the of mine gas concentration.

In recent years, several methods are reported in the literature for gas concentration prediction. Papers [2-4] show that the gas concentration sequence is chaotic time series. In [2], a prediction model was constructed by an adding-weight one-rank local-region method in the reconstruction phase space. In [3], a prediction model was built using time series and adaptive fuzzy reasoning neural system. In [4], it proposed a max Lyapunov index model. In LS-SVM case, the sparseness and robustness may lose, and the estimation of the support values is optimal only in the case of a Gaussian distribution of the error variables. In [5], it proposed the weighted LS-SVM to overcome these draw backs. In[6], the chaotic phase space reconstructive method was used to reconstruct the sample space of gas concentration in multivariate time series and the Gaussian process regression model was used to predict the gas concentration around the work face. These methods provide a good guidance for gas data prediction. However, the gas concentration time series a non-stationary time series with a strong randomness which contains multi-dimensional information and single prediction method cannot fit for information on each dimension.

The wavelet decomposition and reconstruction can decompose the multi-component signal information into a low-frequency approximate signal and a set of high-frequency detail signals. The low-frequency signal react the inherent variation trend of the information while the high-frequency signal react the stochastic disturbance influence of it. In view of the different rules of these two types of signals different models and parameters can be utilized to

independently predict these signals [7]. Based on this idea, this paper proposed a gas concentration prediction combine multi-predictor based on wavelet transform. It makes a combined prediction for the gas concentration time series while different prediction model is utilized for these different signals decomposed by the wavelet decomposition and reconstruction. Based on the research and application in the I1826 Coal Face of Luling coal mine of Huaibei Mining Group Company in Anhui Province, it shows that this method can take advantage of different predictor and effectively predict the gas concentration.

The rest of this paper is organized as follows. Section II analyses corresponding basic theories and methods. The wavelet decomposition and reconstruction algorithm is described in section II-A and the Markov correction method for the ARMA model is designed in section II-B. The proposed gas concentration prediction combine multi-predictor based on wavelet transform is presented in Section III. The proposed prediction model is tested by the gas concentration data and the result is compared with that of other models in Section IV. Section V includes the conclusions of this paper.

2. Analysis of Basic Theories and Methods

2.1. Wavelet Decomposition and Reconstruction

The essence of the wavelet decomposition and reconstruction is to divide a set of primitive sequence containing comprehensive information into several groups of sequences with different characteristics by a group of band pass filters [8]. In this paper, The Mallat algorithm is adopt as the wavelet decomposition and reconstruction method, let $Y = \{y_1, y_2 \cdots, y_N\}$ be the original sequence, where N is the sequence length, the algorithm can be described as follow.

$$\begin{cases} a_{j+1} = H(a_j) \\ d_{j+1} = G(a_j) \end{cases}, \quad j = 0, 1, \cdots, J \quad (1)$$

$H(\cdot)$ and $G(\cdot)$ represent the low-pass filter and high-pass filter. a_{j+1} and d_{j+1} are the components of the original signal in adjacent frequency band under the resolution of $2^{-(j+1)}$ while a_{j+1} represent the low-frequency approximate component and d_{j+1} represent the high-frequency detail component. Let J be the decomposition level. We can get J detail components d_1, d_2, \cdots, d_J and an approximate component a_J . For the length of the decomposed sequence is the half of that of the original one, binary interpolation method was adapted in the reconstruction sequence reconstructing [9].

$$\begin{cases} A_j = (H^*)^j a_j \\ D_j = (G^*)^{j-1} G^* d_j \end{cases}, \quad j = 0, 1, \cdots, J \quad (2)$$

H^* and G^* are the dual operators of H and G . Detail sequences D_1, D_2, \cdots, D_J and approximate sequence A_J are the reconstruction sequences of d_1, d_2, \cdots, d_J and a_J . They have the same length with original sequence. And the original sequence can be represented as the sum of reconstruction sequences.

$$Y = D_1 + D_2 + \cdots + D_J + A_J. \quad (3)$$

2.2. Markov Correction Method (MCM)

The high-frequency detail sequences can be seen as a strong randomness stationary time series and ARMA (Auto-Regressive and Moving Average) model can be applied to the prediction of these sequences. However, these high-frequency detail sequences contain a greater amount of random ingredients, and the higher level the wavelet decomposition reaches, the stronger the randomness of these detail sequences while become. Although the ARMA

model exhibits a good performance for nearly stationary time series, it cannot do better for the time series with strong randomness than statistical model. The specific performance is that most predicted points have good prediction precision, but part of the points appears large prediction deviation for the mutation characteristics. Therefore, considering the single point reliability, a deviation correction method based on the Markov constraints model is proposed to correct the result of the ARMR model. The method can be described as follows. According to utility evaluation of the predicted output of the ARMA model, it decides whether to correct the result. The prediction value will be kept if evaluation result is consistent with the Markov prediction interval, otherwise it will be corrected. Let S_i be the state value of the data predicted by Markov model and $[Low_i, Up_i]$ be the corresponding classification boundary. The state value of the prediction value Y_i of the ARMR model is S_i' and the corrected value is \hat{Y}_i . The correcting method can be described as follow.

$$\hat{Y}_i = \begin{cases} Up_i & S_i' > S_i \\ Y_i & S_i' = S_i \\ Low_i & S_i' < S_i \end{cases} \quad (4)$$

If the state value S_i' is higher than the state value S_i , the prediction value will be corrected to the maximum value of the interval boundary. And if the state value S_i' is lower than the state value S_i , it will be corrected to the minimum value of the interval boundary. Otherwise, it will not be corrected.

The purpose is to check the prediction value of the detail sequence and reconstruction and timely correct the relative large deviation. With the three states segmentation method and the state window parameter W_s and the possibility transition matrix window parameter W_c , the correction steps for the predicted value of a detail sequence of the gas concentration time series $\{y_1, y_2, \dots, y_N\}$ can be described as follow.

Firstly, we transform the detail sequence which has a W_c step size to a state sequence before the time N. With the boundaries $\mu - \sigma$ and $\mu + \sigma$, the items of the detail sequence are transformed to three state E_1, E_2 and E_3 , where $\mu = \frac{1}{W_s} \sum_{i=1}^{W_s} y_{N-i}$ and $\sigma = \sqrt{\frac{1}{W_s} \sum_{i=1}^{W_s} (y_{N-i} - \mu)^2}$.

Let s_t be the state sequence of y_t , where $s_t = \{s_1, s_2, \dots, s_N\}$. The transformation Equation

$$s_t = \begin{cases} E_1 & 0 \leq y_t < \mu - \sigma, \\ E_2 & \text{if } \mu - \sigma \leq y_t < \mu + \sigma, \\ E_3 & y_t \geq \mu + \sigma. \end{cases} \quad (5)$$

Secondly, we can get the Markov transition probability matrix depend on the state sequence.

Thirdly, we get the deviation correction equation and the corrected value \hat{y}_{N+1} .

$$\hat{y}_{N+1} = \begin{cases} \hat{y}_{n+1} & \hat{s}_{N+1} = s_{N+1}, \\ \mu - \sigma & \hat{s}_{N+1} \neq s_{N+1} \wedge s_{N+1} = E_1, \\ \mu - \sigma & \text{if } \hat{s}_{N+1} \neq s_{N+1} \wedge s_{N+1} = E_2 \wedge \hat{s}_{N+1} = E_1, \\ \mu + \sigma & \hat{s}_{N+1} \neq s_{N+1} \wedge s_{N+1} = E_2 \wedge \hat{s}_{N+1} = E_3, \\ \mu + \sigma & \hat{s}_{N+1} \neq s_{N+1} \wedge s_{N+1} = E_3. \end{cases} \quad (6)$$

Where \hat{y}_{N+1} is the predicted value based on the ARMA model, \hat{s}_{N+1} is the corresponding state value based on the predicted value and s_{N+1} is the state value base on the state value sequence and Markov transition probability matrix.

3. Gas Concentration Time Series Multi-predictor Based on Wavelet Transform

Firstly, we use the Mallat algorithm to decompose and reconstruct the gas time series. Then, the different prediction models are established for the low-frequency approximate sequence and high-frequency detail sequences. At last, the final predicted value was calculated by the sum of the results of every prediction model. The block schematic for the workflow of the combine multi-predictor is shown in Figure 1.

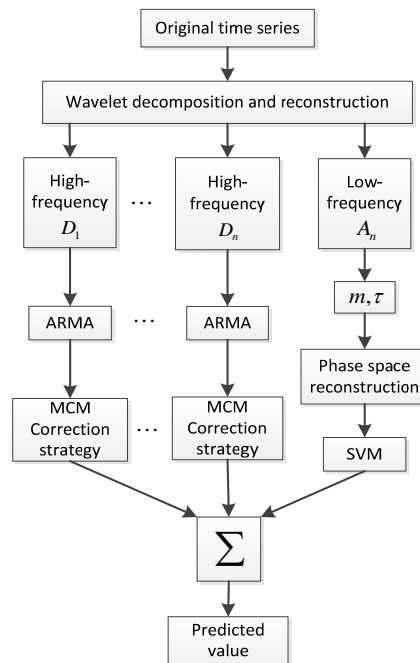


Figure 1. Gas Concentration Multi-predictor Combination Prediction Framework Based on Wavelet Transform

Based on the dyadic wavelet decomposition and reconstruction, the gas concentration time series $Y = \{y_1, y_2, \dots, y_N\}$ can be obtained as follows:

$$Y = D_1 + D_2 + \dots + D_J + A_J,$$

Where $D_1\{d_{1,1}, d_{1,2}, \dots, d_{1,N}\}$, $D_2\{d_{2,1}, d_{2,2}, \dots, d_{2,N}\}$, \dots , $D_J\{d_{J,1}, d_{J,2}, \dots, d_{J,N}\}$ are there constructed high-frequency sequences of every layers and $A_J\{a_{J,1}, a_{J,2}, \dots, a_{J,N}\}$ is the reconstructed low-frequency sequence of the J layer.

The high-frequency sequences acquiring by wavelet decomposition and reconstruction of the time series can be seen as the stationary time series with strong randomness, which reflect the local mutation characteristics of the data. Therefore, we build ARMA prediction model and correct the deviation of the predicted value by the Markov correction method. The specific steps can be described as follows:

(1) Build the ARMA(p, q) model for D_j and use the existing sequence values to estimate the parameters where $1 \leq j \leq J, i \leq M$.

(2) With adaptive testing for the ARMA(p, q) model and AIC criterion method applied for the order determination of the model, the model parameters p and q are obtained.

(3) The predicted value $\hat{d}_{j,N+k}$ of $d_{j,N+k}$ can be obtained by the ARMA(p, q) model.

(4) Grid method is applied to parameter optimization of the weighted Markov correction method where the state divided window $W_s \in [a_s, b_s]$ and the probability calculation window $W_c \in [a_c, b_c]$ and the minimum classification error rate is applied as the criterion to obtain the optimum parameters W_s and W_c .

(5) With the optimum parameters W_s and W_c and the equation (5), the gas concentration time series is transformed to the state sequence.

(6) The predicted state value $s_{j,N+k}$ is estimated depend on the present values of the state sequence and the possibility transition matrix.

(7) Calculate the state value $\hat{s}_{j,N+k}$ for the predicted value $\hat{d}_{j,N+k}$ with equation (5).

(8) The final corrected predicted value $\hat{d}_{j,N+k}$ is obtained by the equation (6).

The low-frequency detail sequence with nonlinearity reflects the basic characteristics of the gas concentration time series. Therefore, we use the SVM model for its prediction. The specific steps can be described as follows:

(1) C-C method is applied to get the optimum delayed time and embedding dimension of $A_J \{a_{J,1}, a_{J,2}, \dots, a_{J,N}\}$.

(2) The input and output vectors of the SVM model is obtained through phase space reconstruction with the delayed time and embedding dimension.

(3) The parameters of SVM are optimized by the PSO (Particle Swarm Optimization) [10, 11].

(4) The predicted value $a_{J,N+k}$ of the detail sequence is obtained by trained SVM model [12].

The final predicted value of the gas concentration time series is obtained by superimposing the predicted values of all components.

4. Experiment Results and Discussions

We test the proposed model with 860 gas concentration samples coming from the I1826 Coal Face of Luling coal mine of Huaibei Mining Group Company in Anhui Province. We select the first 720 data as the experimental samples and the last 140 data as the predictive samples. Firstly, the gas concentration time series is reconstructed to three sequences with 3 level scale wavelet base db3 by the wavelet decomposition and reconstruction of the original data. The reconstructed sequences are shown in Figure 2.

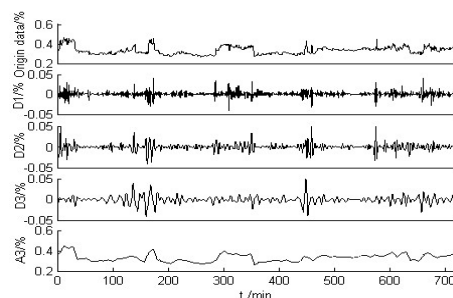


Figure 2. Wavelet Decomposition and Reconstruction of the Original Gas Concentration Time Series

We build the ARMA model for the high-frequency sequences D1, D2, and D3 while the AIC criterion method is used for order determination. The new constructed models are ARMA(2, 9), ARMA(9, 8) and ARMA(7, 10). Then we correct the predicted value by the weighted Markov correction method and obtain the optimum parameters W_s and W_c by the grid method where $W_s \in [1, 100]$ and $W_c \in [1, 200]$. The optimum parameters obtained by the ARMA model with weighted Markov correction method are show as the Table 1.

Table 1. Model Parameters of Detail Sequences

Reconstructe d Sequence	p	q	W_s	W_c
D1	2	9	64	197
D2	9	8	100	169
D3	7	10	72	193

We calculate the largest Lyapunov exponent by small data sets for the low-frequency approximate sequence and the result is $\lambda = 0.1146$. It indicates that the sequence has chaotic characteristic. With the optimum delayed time $\tau = 2$ and the embedding dimension $m = 5$ obtained by C-C method, we reconstruct phase space of the sequence and get the input and output vectors of the SVM model. With radial basis function as the kernel function of the SVM model, the parameters of SVM optimized by the PSO are $c = 241.9441$, $\varepsilon = 0.01$, $\sigma = 0.1$.

For comparison, we choose four evaluation criterions MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error), RMSE (Root Mean Square Error) and NRMSE (Normalized root mean square error). Expressions of these criterions are show as follow:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (7)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right|, \quad (8)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}, \quad (9)$$

$$NRMSE = \sqrt{\frac{1}{n} \times \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})}}, \quad (10)$$

Where i is the index of the data, \hat{y}_i is the predicted value of the data, y_i is the true value, \bar{y} is the average value and n is the total number of the data.

In order to verify the precision of the fitted model, use the algorithm proposed in this paper to predict the last 140 gas concentration values. Table 2 shows the prediction error of each sequence. It shows that the higher level the wavelet decomposition and reconstruction reaches, the stronger the randomness and noise of these detail sequences become and the lower the precision of ARMA prediction model get. And Markov constraint correction method can greatly correct the deviation of the predicted value due to randomness. However, for the D3 sequence with small randomness, ARMA prediction model can get a high predictive precision and Markov correction method do not have a remarkable result. Experiment results show that for the gas concentration time series with strength volatility, the combination of ARMA model

and the Markov modification strategies can take advantage of the predictive ability of the linear model and statistical model.

Table 2. Prediction Error of Wavelet Decomposition Sequence

Reconstructed Sequence	Method	MAE	MAPE	RMSE	NRMSE
D1	ARMA	0.0303	31.4719	0.0632	0.2574
	ARMA-MCM	0.0137	17.0502	0.0232	0.0944
D2	ARMA	0.014	3.8727	0.0295	0.1193
	ARMA-MCM	0.0111	3.0899	0.0249	0.1006
D3	ARMA	0.0076	0.8507	0.0137	0.0431
	ARMA-MCM	0.0078	0.864	0.0133	0.0418
A3	SVM	0.0131	0.0363	0.019	0.0349

Figure 3 shows the chart of the final predicted value superimposed upon every subsequence and the actual data. From the Figure, It can be seen that the final predicted data of the proposed method can fit the actual gas concentration data well.

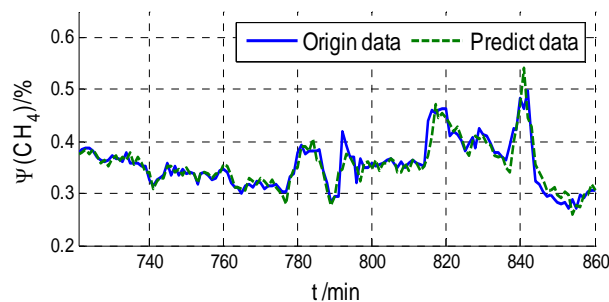


Figure 3. Actual Value (the solid line) and Model Output (the dashed line) of Checking Samples

To verify the effectiveness of the proposed method, routine methods are used to predict the gas concentration samples for comparison. These methods include BP neural network prediction model, SVM prediction model, SVM prediction model based on BP (W-BP) and SVM prediction model based on wavelet decomposition (W-SVM). The predicted results comparison show as the Table 3.

Table 3. Comparison of Different Prediction Methods

	BP	SVM	W-BP	W-SVM	Multi-predictor
MAE	0.0372	0.0317	0.0282	0.0158	0.0131
MAPE	0.1035	0.0681	0.0778	0.0426	0.0363
RMSE	0.0477	0.0279	0.0389	0.023	0.019
NRMSE	0.0878	0.09	0.0716	0.0424	0.0349

In the BP neural network prediction model, we get the average value of ten time independent prediction as the final predicted value. The network hidden layer transfer function is Sigmod function, the transport layer transfer function is Purelin function, the training algorithm is variable learning rate momentum and gradient descent algorithm and the learning rate is 0.1. In the SVM model, we chose radial basis function as the kernel function and optimize the parameters with the PSO method. From Table 3, it can be seen that the precision of the multi-

predictor model proposed in this paper is obviously superior to the single-predictor models, such as BP and SVM. And the prediction error is smaller than that of the super position of the predicted values of the wavelet decomposition and reconstruction sequences based on the single-predictor BP or SVM. The experiment results show that the combination of multi-scale wavelet decomposition and multi-predictor can improve the prediction precision and have a strong practicability.

5. Conclusion

This paper discusses the multi-predictor combination prediction method for gas concentration and builds the multi-predictor prediction model with wavelet transform, Markov correction model, ARMA and SVM. According to the theoretical analysis and experimental results, the following conclusions can be drawn.

(1) The wavelet decomposition and reconstruction can decompose the gas concentration information into different sequences. According to the characteristics of different sequences, we apply different prediction models for different sequences and take advantages of each predictor to effectively improve the prediction precision.

(2) For the high-frequency detail sequences has a strong randomness, we combine the ARMR model and Markov constraints correction method and take full advantage of the linear fitting ability of the ARMR model and the statistical ability of Markov model.

(3) The single-predictor models have some disadvantages, such as a great volatility of the result, susceptible interference caused by uncertainties and unwarranted prediction precision. In this paper, we proposed a multi-predictor combination prediction method based on wavelet transform. Compared with the routine single-predictor model BP and SVM, and wavelet-based single-predictor model W-BP and W-SVM, The proposed method improved the prediction precision.

References

- [1] Zhou SN, Lin BQ. Coalbed gas occurrence and flow theory. Beijing: China Coal Industry Publishing House. 1999.
- [2] Cheng J, Bai JY, Qian JS, et al. Short-term forecasting method of coalmine gas concentration based on chaotic time series. *Journal of China University of Mining & Technology*. 2008; 37(2): 231-235.
- [3] Zhang JY, Cheng J, Hou YH, et al. Forecasting coalmine gas concentration based on adaptive neuro-fuzzy inference system. *Journal of China University of Mining & Technology*. 2007; 36(4): 494-498.
- [4] Cui XY, Zhao GQ. Application of chaos forecast of mine gas density monitoring. *Journal of Liaoning Technical University*. 2008; 27: 184-186.
- [5] Qiao MY, Ma XP, Lan JY, et al. Time series short-term gas prediction based on weighted LS-SVM. *Journal of Mining and Safety Engineering*. 2011; 28(2): 310-314.
- [6] Dong DW, Li SG, Chang XT, et al. Prediction model of gas concentration around working face using multivariate time series. *Journal of Mining and Safety Engineering*. 2012; 29(1): 135-139.
- [7] He GG, Ma SF, Li Y. A study on forecasting for time series based on wavelet analysis. *Acta Automatica Sinica*. 2012; 29(1): 1012-1014.
- [8] Ajay SP, Devender S, Sunil KS. Intelligent Hybrid Wavelet Models for Short-Term Load Forecasting. *IEEE Tran. on power systems*. 2010; 8(3): 1266-1273.
- [9] Mallat SG. A theory for multi resolution signal decomposition: The wavelet representation. *IEEE Tran. on Pattern Analysis and Machine Intelligence*. 1989; 11(7): 674-693.
- [10] Kennedy J, Eberhart RC. *Particle swarm optimization*. Proceedings of IEEE International Conference on Neural Networks. Piscataway, NJ: IEEE Service Center, 1995: 1942-1948.
- [11] Eberhart RC, Shi Yuhui. *Particle swarm optimization: developments, applications and resources*. Proceedings of the 2001 Congress on Evolutionary Computation. Piscataway, NJ, USA: IEEE Press. 2001; 1: 81-86.
- [12] Vapnik VN. The nature of statistical learning theory. New York: Springer. 1995.