A fuzzy combined forecasting model of coal spontaneous combustion

Pengtao Jia[1*], Jun Deng[2], Shuhui Liang[1]

[1]*School of Computer Science and Technology, Xi'an University of Science and Technology, Xi'an Shaanxi 710054, China*

[2]*School of Energy Engineering, Xi'an University of Science and Technology, Xi'an Shaanxi 710054, China*

Received 25 July 2014, www.tsi.lv

Abstract

This paper focuses on the effective analysis of the coal spontaneous combustion monitoring data, so as to realize the accurate and reliable coal spontaneous combustion limit parameter prediction. Firstly, a weighted multimember fuzzy operation model was constructed. When the additive generator of the model changes, this model can generate new operation clusters. Based on it, a new combined forecasting model of coal spontaneous combustion limit parameter is proposed. The new model can use linear and nonlinear models as its single forecasting models. Its combination is variable and has good generalization ability. Then, the BP neural network model and the support vector machine were used as the single forecasting models of the new model. Finally, for realizing the optimal combination of single models, genetic algorithm and least square method were used to evaluate parameters of new model. The experimental analysis shows that the new model leads to less error and better performance than single models. It can be concluded that the new combined forecasting model is suitable for coal spontaneous combustion.

Keywords: coal spontaneous combustion, limit parameters, combined forecasting, genetic algorithm, least square method

1 Introduction

Coal spontaneous combustion is an important problem in its mining, long distance transportation, and storage, in terms of both mine safety and economics. It has large proportion and wide coverage in China mine. According to statistics, the coal spontaneous combustion fire accounted for more than 94% of the total number of coal mine fire in China [1]. The coal spontaneous combustion is a physical-chemical process, which is extremely complex, dynamic changing and automatic accelerating. Virtually, it is a slow and automatic process of oxidation, heating, and then burning. Coal spontaneous combustion requires certain external conditions, of which the limiting ones are called limit parameters. The limit parameters are considered as the strong basis for estimating the dangerous areas of coal spontaneous combustion. There are some main parameters, such as the lower oxygen concentration, the ceiling air leakage strength and the minimum float coal thickness [2]. Because influencing factors of the limit parameters are relatively complex [3], and it is meaningful to establish appropriate models to forecast the limit parameters of coal spontaneous combustion [4]. At present, there are basically two kinds of forecasting models of limit parameters as follows:

One, mathematical models based on numerical simulation [2, 5-7]. To simplify the calculation, in the process of calculation, only the effects of main factors are considered, the secondary factors are ignored or taken as constant values. Then the limit parameters can be estimated approximately. So there often has large deviation between its calculation results and the actual ones.

Two, forecasting models based on data mining methods. In order to overcome the deficiency of mathematical models, some scholars establish the prediction model through data mining methods, such as the neural network prediction model [8, 9], rough set neural network model [10], support vector machine (SVM) model [4, 11], rough set support vector machine forecasting model [12], regression analysis [13].

As we know the generalization abilities of single forecasting models are poor. Although many scholars perform some efforts on single forecasting models for improving forecasting capability, these investigations mainly focus on exploring new forecasting algorithm and improving old forecasting algorithm, and do not overcome disadvantage of single forecasting models. One of the important directions in improvement of the forecasting ability is the integration of multiple single forecasting models. Several effective methods have been proposed to combine the results of the single forecasting models, such as product operator, mean operator, median operator, max operator, min operator and majority vote method. But these integration methods always have not good capability in different datasets. In this paper, we put forward a new fuzzy combined forecasting model to predict limit parameters of coal spontaneous combustion based on a weighted multimember operation model.

*****Corresponding author** e-mail: pengtao.jia@gmail.com

2 The weighted multimember fuzzy operation model

Many existing fuzzy operation models were binary operators with the same weight. But the various factors in the actual complex systems usually have more than two factors and different weights. So weighted parameters and many factors are introduced into the fuzzy operation model, we construct a weighted multimember fuzzy operation model.

Theorem 1: Assume that function $f(x)$ is an additive generator, then

$$G(x_1, x_2, \cdots, x_n, \alpha_1, \alpha_2, \cdots, \alpha_n) = f^{-1}(\max(f(0), \sum_{i=1}^{n} \alpha_i f(x_i) - 1)) \quad (1)$$

is a weighted multimember fuzzy operation model. Where $x_i \in R^+ \; (i = 1, 2, \cdots, n)$, α_i is the weight of $f(x)$, $\alpha_i \in [0,1]$ and $\sum_{i=1}^{n} \alpha_i = 1$.

When f(x) changes, this model can generate new operation clusters. According to this variability, the model gets good generalization ability. For example, set $f(x) = x^p$, Equation (1) can generate an operator cluster as follows:

$$G(x_1, x_2, ..., x_n, \alpha_1, \alpha_2, ..., \alpha_n) = f^{-1}(\sum_{i=1}^{n} \alpha_i x_i^p - 1) = \sqrt[p]{\sum_{i=1}^{n} \alpha_i x_i^p - 1}, \quad (2)$$

where, p is the parameter of the generator and $p \in (-\infty, 0) \cup (0, +\infty)$. Later, the combined forecast model will be constructed based on Equation (2).

3 Combined forecasting of coal spontaneous combustion limit parameters

3.1 FORECASTING MODEL BASED ON BP NEURAL NETWORK

Limit parameters of coal spontaneous combustion are influenced by many factors. There is a nonlinear relationship between influence factors and limit parameters, so the BP neural network can be used to predict the limit parameters [8]. In this paper, BP neural network with three layers is used as the single forecasting model to predict limit parameters of coal spontaneous combustion. In the input layer, five nodes are used to input the impact factors. If the impact factors need more comprehensive consideration, additional nodes should be added to the input layer. The predictive value of the limit parameter is considered as the only node in the output layer. After training and comparing the forecasting model several times, it found that the training effect is better when only use a hidden layer with ten nodes, and select the logarithmic function *sigmod* as the excitation function. Then set the initial weights in (-1, 1) and the convergence error to 1e-6. Take the ceiling air leakage strength for instance, the structure of BP neural network forecasting model is shown in Figure 1.

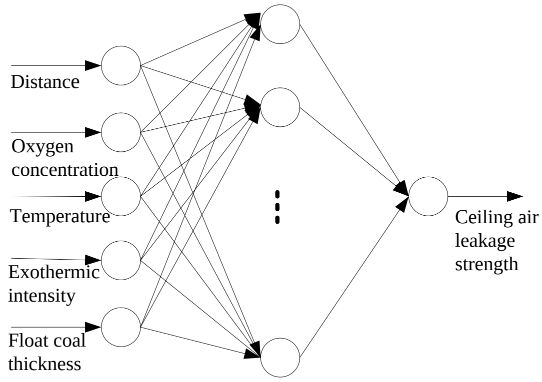

FIGURE 1 The BP forecasting model for limit parameters of coal spontaneous combustion

3.2 FORECASTING MODEL BASED ON SVM

Support Vector Machine (SVM) is a learning machine based on minimum structural risk and statistical learning theory. It has unique advantages in solving small sample, nonlinear and high dimensional pattern recognition problems. The basic idea of using the SVM algorithm to estimate the regression function is that firstly map the input space data x into a high-dimensional feature space through a nonlinear mapping, then proceed the linear regression in the high-dimensional space.

Least Squares Support Vector Machines (LS-SVM) is one kind of the support vector machine. It uses the least squares linear system instead the traditional support vector, that is, using the quadratic programming method to solve the pattern recognition problems, transforming the quadratic optimization of the original SVM algorithm into solving linear equations. So it can effectively reduce the computational complexity. In this paper, the LS-SVM is used as the single forecasting model to predict limit parameters of coal spontaneous combustion.

The modelling steps of LS-SVM forecasting model are as follows:

1) To describe the function fitting problem as an optimization problem;

2) To solve the optimization problem by using the Lagrange method, and convert it to solving the systems of linear equations. The training of the LS-SVM model is mainly to solve the systems of linear equations;

3) To get the LS-SVM fitting model, as follows:

$$y(x) = \sum_{k=1}^{N} \alpha_k K(x_k, x) + b, \quad (3)$$

where, α_k is the support vector, $b \in R$ is the offset value. α_k and b can be obtained according to the training sample data. $K(x_k, x)$ is the kernel function which is a symmetric function and satisfies the Mercer condition.

Predicting by using the LS-SVM model simply needs to calculate the kernel function $K(x_k, x)$ between each training sample and the sample under test.

4) To select kernel function. Selecting the kernel function is an important part for building a model. Consider that the Gaussian radial basis kernel function (RBF) has good learning ability and wide domain of convergence, RBF is selected as the kernel function here, so:

$$K(x_k, x) = \exp(-\| x_i - x_j \|^2 / \sigma^2), \quad (4)$$

where, σ is the kernel parameter.

3.3 CONSTRUCT THE COMBINED FORECASTING MODEL

Modelling method of combined forecasting is a portfolio of predicting the same objects by using two or more predict methods. Theoretical research and practical application show that the combined forecasting has high ability to adapt to the change of the future predict environment, and it can enhance the stability of forecast, so as to achieve the purpose of improving the prediction precision [15,16]. The key of the combination forecasting model lies in its generalization ability. The model can be described as follows:

Assume that the actual observations of a certain prediction problem at time t is $y(t)$ $(t = 1, 2, ..., m)$, there are n feasible forecast methods, the corresponding prediction models respectively are $f_1, f_2, ..., f_n$, their predictive values respectively are: $\hat{y}(t)$ $(t = 1, 2, ..., m; i = 1, 2, ..., n)$, i.e. $\hat{y}(t) = f_i(t)$, and the weighted combination forecasting problem can be described as:

$$\hat{y}(t) = F(\hat{y}_1(t), \hat{y}_2(t), ..., \hat{y}_n(t), \alpha_1, \alpha_2, ..., \alpha_n) \quad (5)$$

where, the combination forecasting values are $y(t)$ $(t = 1, 2, \cdots, m)$, F is the way of combination. Using Equation (5) aims to make the combination forecasting values better than the single prediction effects.

The combined forecasting model of this paper is based on the theory of BP neural network and support vector machine (SVM), the predictive values are $\hat{y}_{BP}(t)$ and $\hat{y}_{SVM}(t)$.

According to the characteristics and advantages of each single forecast model, different weights, such as α and $1-\alpha$ are assigned to each single one in the combined model. Consider Equation (2) as the F of Equation (5), then the combination forecasting model (CFM) can be described as:

$$\hat{y}(t) = \sqrt[p]{\alpha \hat{y}_{BP}(t)^p + (1-\alpha)\hat{y}_{SVM}(t)^p - 1} \quad (6)$$

The parameters are estimated by combining genetic algorithm with least squares method. Due to the objectivity and inevitability of the prediction error, there are errors between the predictive values $\hat{y}(t)$ and the actual ones $y(t)$. Set:

$$E = \sqrt{(\hat{y}(t) - y(t))^2}. \quad (7)$$

Minimizing E is used as the evaluation of the objective function in genetic algorithm, and then the parameters in the combination model can be obtained.

Genetic algorithm (GA) is an adaptive global optimization search algorithm, which is formed by simulating the genetic and evolutionary process of organisms in the natural environment. Given its global optimization ability, GA is used as the parameter estimation module of CFM.

Set the following parameters of GA for parameter optimization:

1) The initial population is 20;

2) Use binary coding with eight numbers;

3) Select operation by using the uniform distribution random model;

4) Do crossover operation by the disperse cross;

5) Mutate operation by using gauss function.

4 Experiment and result analysis

4.1 EXPERIMENTAL DESIGN

Taking the ceiling air leakage strength prediction of coal spontaneous combustion limit parameters as an example, some main influence factors of the ceiling air leakage strength are selected as the input data, such as the exothermic intensity of coal, coal temperature, the measured oxygen concentration, distance from working face to the goaf prediction area and the float coal thickness. And the output data is the ceiling air leakage strength. The dataset from Xinzhou mine of China [8] is shown in Table 1.

Here, the data with number 1 to 20 was used as training samples, and that with number 21 to 25 was used as testing samples.

TABLE 1 Dataset of the combined forecasting model

No.	Input data					Output data
	Distance /m	Oxygen concentration /%	Coal temperature /°C	Exothermic intensity $/10^5 J \cdot s^{-1} \cdot cm^{-3}$	Float coal thickness / m	Ceiling air leakage strength / $cm^3 \cdot cm^{-2} \cdot s^{-1}$
1	1.70	20.60	19.60	0.87	7.0	0.70
2	2.50	20,04	20.30	1.04	6.0	0.83
3	4.70	19.88	22.00	1.27	5.0	1.08
4	7.60	19.03	22.50	1.34	4.0	1.56
5	16.30	18.21	24.20	1.43	2.0	2.35
6	20.50	17.99	25.60	1.51	3.0	2.58
7	25.20	17.60	26.70	1.58	4.0	2.88
8	29.10	17.36	26.80	1.58	3.0	3.17
9	36.40	16.90	27.50	1.62	2.0	3.43
10	43.90	15.74	28.30	1.67	3.0	3.87
11	44.30	15.68	28.60	1.69	4.0	3.92
12	47.00	14.91	28.10	1.66	5.0	4.12
13	53.70	13.77	25.13	1.49	7.0	5.60
14	56.40	13.09	24.80	1.47	6.0	5.77
15	59.00	12.44	24.30	1.43	5.0	6.00
16	61.20	11.93	23.60	1.40	4.0	6.53
17	70.60	10.78	24.67	1.46	2.0	5.39
18	74.30	9.81	26.30	1.55	2.0	4.18
19	78.00	8.85	27.80	1.61	3.0	3.76
20	89.20	7.14	30.40	1.79	3.0	2.89
21	11.00	18.59	23.40	1.39	3.0	2.04
22	39.70	16.50	27.90	1.64	2.0	3.54
23	50.40	14.36	28.20	1.66	6.0	3.86
24	66.80	11.18	24.20	1.43	3.0	6.12
25	83.50	7.97	28.10	1.66	4.0	4.17

Experimental steps are listed as follows:

1) To facilitate comparison, firstly the samples should be standardized to [0, 1]. The standardized formula is as follows:

$$norm(x_i) = \frac{x_i - \min(X)}{\max(X) - \min(X)}. \tag{8}$$

2) Training the single forecasting method BP and SVM on the training sets, getting the prediction results $\hat{y}_{BP}(t)$ and $\hat{y}_{SVM}(t)$ on the test sets.

3) The CFM model is now used. The genetic algorithm is used to estimate parameters on the training sets, and then get the CFM prediction results $\hat{y}(t)$ on the test sets.

4) Normalizing all the prediction results, the normalized formula is as follows:

$$y_i = \hat{y} \times (\max(X) - \min(X)) + \min(X) \tag{9}$$

5) Using evaluation index to evaluate.

4.2 RESULTS AND ANALYSIS

The prediction results of CFM model are shown in Table 2, the parameters of the Equation (6), namely the combination forecasting model, obtained based on GA are α=0.261 and p=2.926.

TABLE 2 Contrast of the test sample's expected outputs and the predicted results

Sample number	Actual value	BP	LS-SVM	CFM
1	1.88	1.81977	1.93764	1.87373
2	3.52	3.66935	3.56403	3.61498
3	4.24	4.12589	4.37216	4.24232
4	6.01	6.01739	5.99676	6.00860
5	3.76	3.79867	3.61776	3.70324

The contrast results on the error evaluation index are shown in Table 3.

TABLE 3 Comparison of error percentage of prediction results of different models (%)

Sample number	BP	Improved BP [8]	LS-SVM	SVM [4]	CFMSPT
1	3.20	8.51	3.07	3.67	0.33
2	4.24	0.57	1.25	1.94	2.70
3	2.69	8.96	3.12	3.04	0.05
4	0.12	1.83	0.22	0.35	0.02
5	1.03	10.90	3.783	2.88	1.51
average error	2.26	6.15	2.29	2.38	0.92

It can be seen from Table 3 that when forecasting the ceiling air leakage strength by using various forecasting models, the result of BP model in this paper is 2.26%, the result of the LS-SVM model is 2.29%, while the result of CFM is 0.92%, it is the lowest.

The prediction results show that BP neural network and SVM can all be used to forecast limit parameters of coal spontaneous combustion, and there is not much difference between their results. But for small samples, different number choice of hidden layer for BP neural network will lead to large difference. For example, the average relative error of improved BP [8] is 6.15%, it is very different with the BP neural network in this paper, because they have different training methods and number of hidden layer. But there is no theory evidence for selecting the number of hidden layer, they are mostly chosen based on experience, and the training time is too much. Thus, excessive fitting situation may appear easily. So, in view of the small sample data, support vector machine (SVM) is more suitable, such as the SVM in reference [4] and the LS-SVM in this paper, their results are basically identical. However, no matter BP model or

SVM model, given the limitations of single model itself, it is hard to work in all cases, so its generalization ability is weak. Gratifyingly, the combined forecasting model makes up for the inadequacy of single model, complements itself by the advantage of single model, and obtains the best prediction effect.

Any linear or nonlinear model can be used in the single forecasting model of CFM. So CFM has variability, it can always establish a combined prediction model, and CFM is most suitable for predicting the characteristics of time series data.

This method is also suitable for predicting the lower oxygen concentration and the minimum float coal thickness.

5 Conclusions

According to this study, it can be concluded as follows:

1) A combined forecasting model is proposed to predict limit parameters of coal spontaneous combustion.

2) Compared with the single forecasting model, the error of the combined forecasting model is lower, and the prediction effect is better.

3) CFM model has variability and strong generalization ability. It can be trained to find the most suitable combined forecasting model for the characteristics of the dataset, thus ensuring the predicted results of CFM model at least as good as the ones of other single forecasting models.

Acknowledgments

The project is supported by the key project of national natural science foundation of China (Program No. 51134019), the natural science basic research plan in Shaanxi province of China (Program No.2012JQ8035) and the special scientific research of the department of education in Shaanxi province (Program No. 2013JK0870).

References

[1] He Q L, Wang D M 2004 Numerical simulation of spontaneous combustion process in goaf areas by fully-mechanized and caving roof coal *Journal of CUMT* **33**(1) 11-4 (*in Chinese*)

[2] Xu J C, Wen H, Deng J, Zhang X H 2000 Study limit parameter of coal self-ignite *Fire Safety Science* **9**(2) 15-7 (*in Chinese*)

[3] Luo D Y, Zhang G S 2010 Impact factors of survived coal spontaneous combustion in goaf *Coal Technology* **29**(11) 83-84 (*in Chinese*)

[4] Meng Q, Wang H Q, Wang Y S, Zhou Y 2009 Predicting limit parameters of coal self-ignition based on support vector machine *Journal of China Coal Society* **34**(11) 1489-93 (*in Chinese*)

[5] Zhang X H, Xi G, Chen X K, Deng J, Wen H 2005 Determining spontaneous combustion danger zones and predicting spontaneous combustion during mining near-neighbored coal seams *Journal of China Coal Society* **30**(6) 733-6 (*in Chinese*)

[6] Wen H 2002 Dynamic numeric simulation of coal self-ignite in goaf in fully mechanized caving face *Journal of China Coal Society* **27**(1) 54-8 (*in Chinese*)

[7] Zhang C, Ti Z Y, Li Z X 2012 Three-dimension heterogeneity dynamic numerical simulation of top coal spontaneous combustion in limit equilibrium zone *China Safety Science Journal* **22**(5) 37-42 (*in Chinese*)

[8] Xu J C, Wang H 2002 The neural network prediction method for the limit parameters of coal ignition *Journal of China Coal Society* **27(**4) 366-70 (*in Chinese*)

[9] Wang H 2011 Principal component neural network prediction model for coal self-ignition duration *Computer Engineering and Applications* **47**(26) 242-5 (*in Chinese*)

[10] Hou Y B 2004 An RSNN-based prediction method for the coal mine spontaneous combustion *Information and Control* **33**(1) 93-6 (*in Chinese*)

[11] Gao Y, Qin M G, Li M J 2010 Analysis on prediction of residual coal spontaneous combustion in goaf based on support vector machine *Coal Science and Technology* **38**(2) 50-4 (*in Chinese*)

[12] Meng Q, Wang Y S, Zhou Y 2010 Prediction of spontaneous combustion in caving zone based on rough set and support vector machine *Journal of China Coal Society* **35**(12) 2100-4 (*in Chinese*)

[13] Deng J, Xing Z, Ma L 2011 Application of multiple regression analysis in coal spontaneous combustion prediction *Journal of Xian University of Science and Technology* **31**(6) 645-8 *(in Chinese*)

[14] Mizumoto M 1989 Pictorial Representation of Fuzzy Connectives, Part 1: Cases of T-norms, T-conorms and Averaging Operators *Fuzzy Sets and Systems* **31** 217-42

[15] Zeng K S, Hu N L 2008 Model of system safety forecasting and combination forecasting *Journal of China Coal Society* **33**(10) 1123-5 (*in Chinese*)

[16] Ao P, Mou L H 2011 Load combination forecasting based on power grid with environmental characteristics *Journal of China Coal Society* **36**(9) 1575-80 (*in Chinese*)

Authors

Pengtao Jia, born in 1977, Xinzheng, Henan, China

Current position, grades: Associate Professor
University studies: M.E. degree in computer application technology at Xian University of Science and Technology, China in June 2002. Doctor's degree in computer science and technology at Northwestern Polytechnical University, China in June 2008.
Scientific interest: Computer Science, data mining, application of artificial intelligence theory.
Publications: more than 20.

Jun Deng, born in 1970, Dazhu, Sichuan, China

Current position, grades: Professor at School of Energy Engineering, Xi'an University of Science and Technology.
University studies: Xi'an University of Science and Technology (China).
Scientific interest: mine safety, coal spontaneous combustion.
Publications: more than 80.

Shuhui Liang, born in 1988, Shuozhou, Shanxi, China

Current position, grades: Postgraduate at School of Computer Science and Technology, Xi'an University of Science and Technology.
University studies: Xi'an University of Science and Technology (China)
Scientific interest: computer science, ensemble learning.
Publications: 3.