

Modeling and Prediction of Coal Ash Fusion Temperature based on BP Neural Network

Suzhen Miao¹, Qingyin Jiang¹, Hua Zhou¹, Jia Shi¹, Zhikai Cao¹

¹ Department of Chemical and Biochemical Engineering, National Engineering Laboratory for Green Chemical Productions of Alcohols–Ethers–Esters, College of Chemistry & Chemical Engineering, Xiamen University, Xiamen, China

Abstract. Coal ash is the residual generated from combustion of coal. The ash fusion temperature (AFT) of coal gives detail information on the suitability of a coal source for gasification procedures, and specifically to which extent ash agglomeration or clinkering is likely to occur within the gasifier. To investigate the contribution of oxides in coal ash to AFT, data of coal ash chemical compositions and Softening Temperature (ST) in different regions of China were collected in this work and a BP neural network model was established by XD-APC PLATFORM. In the BP model, the inputs were the ash compositions and the output was the ST. In addition, the ash fusion temperature prediction model was obtained by industrial data and the model was generalized by different industrial data. Compared to empirical formulas, the BP neural network obtained better results. By different tests, the best result and the best configurations for the model were obtained: hidden layer nodes of the BP network was setted as three, the component contents (SiO₂, Al₂O₃, Fe₂O₃, CaO, MgO) were used as inputs and ST was used as output of the model.

Key words: XD-APC, BP neural network, compositions of coal ash, coal ash fusion temperature.

1 INTRODUCTION

Ash fusion temperature (AFT), as we know, one of the most widely applied in coal combustion and gasification processes, is of great significance in the gas production process. To our best knowledge, AFT had no definite melting point but a temperature range^[1] and included four kinds of temperature: Deformation Temperature (DT), Softening Temperature (ST), Hemispherical Temperature (HT) and Flow Temperature (FT). As the main indicators of coal ash fusion characteristics, ST was widely applied to represent the AFT in industry. Various kinds of chemical composition in coal ash and the corresponding effects on the AFT were different. Ash clinkering and slagging inside the gasifier was a complex physical and chemical process, which is related to the mineral species, nature, quantity and the interaction between each other. And the primary task is to seek a reasonable and effective coal ash fusion temperature model both qualitatively and quantitatively.

In order to get a reliable and accurate AFT prediction model, numerous methods were investigated in previous publications^[2-11]. For instance, in literature^[2], an AFT prediction model was established by MATLAB according to the relationship between coal ash fusion temperature and chemical composition of ashes. 0.929, the correlation coefficient, was obtained by using the linear least square method. And an improved linear programming method was adopted to establish the model of steam coal blending by Li^[3]. In addition, a regression equation of

relation^[4] between coal ash fusion temperature and significant variables was obtained by 43 sets of ash analyzed data from Rujigou coalmine on the basis of stepwise regression analysis. A polynomial function of AFT was regressed by Sun^[5]. What's more, coal parameters (such as ash content, specific gravity, hard grove index) were investigated in the model of AFT and non-linear and linear correlations was obtained in literature^[6].

BP neural network and support vector machine (SVM), an alternative method, were also applied to predict the ash fusion temperature. Taking YIN's^[7] work for instance, 160 ash samples by the BP network with momentum item were adopted to estimate the AFT. And the maximum prediction deviation and average deviation were 15.08%, 4.93%, respectively, and the overall deviation was larger. Seven components (SiO₂, Al₂O₃, Fe₂O₃, CaO, MgO, TiO₂, Na₂O + K₂O) were introduced as inputs and ST was applied as output by Li^[8]. What's more, SVM method was also used to set up the model^[9] with the maximum prediction deviation (7.4%) and average deviation (0.678%). Based on the fundamental of ant colony feed forward neural network method, the maximum training deviation of the AFT model established by Liu^[10] was 1.78%, the minimum training deviation was 1.39% and the average training deviation was 1.55%. Most scholars believed that the influence of the coal ash components were most remarkable. In the present study, the prediction accuracy was not ideal as most AFT prediction models were linear and great

difference among different areas. However, the BP neural network had strong nonlinear mapping ability and generalization ability, and a certain tolerance for deviation. In conclusion, this study adopted the advanced industrial process control configuration and simulation software XD-APC^[11] to conduct the BP neural network. With the nearly 200 sets of coal ash data came from different regions of China, the quantitative relationship between ST and the ash compositions (including SiO₂, Al₂O₃, Fe₂O₃, CaO, MgO, TiO₂, Na₂O, K₂O, P₂O₅, SO₃) was obtained.

2 Simulation by BP network model

Nearly 200 sets of coal ash data collected from different regions of China and some coal blending were prepared and analyzed. For all of coal ash, the five oxides of SiO₂, Al₂O₃, Fe₂O₃, CaO, MgO were all included, and most of coal ash contained TiO₂. However, Na₂O, K₂O, P₂O₅, SO₃ were not contained in all the coal ash according to different regions. On the other hand, the content of the same composition were widely varied from different parts, for example, SiO₂ was in the range of 0% to 60%. To investigate the influence weight of components of coal ash and obtain an AFT model which can be applicatively and widely used, compositions were discussed and classified, and three models were estimated based on the contents of coal ash compositions and ST, the models were “five inputs and one output”, “six inputs and one output” and “ten inputs and one output”, respectively.

2.1 The “five inputs and one output” model

The so-called “five inputs and one output” was to set the contents of five coal ash compositions, including SiO₂, Al₂O₃, Fe₂O₃, CaO, MgO as inputs, and set the ST of coal ash as output to train the model. The 32 sets of coal ash data for training the model came from literature^[12] and the results of training were exhibited in Figure 1..

After training, the model had certain ability to memory, and can recall according to the memory, it was defined to be the generalization ability of the BP model. The generalization ability was an important index to observe and study the learning effect of network^[13]. During generalization, the parameters adopted in the training model were constants, and the data for generalization were exchanged into another 19 sets of coal ash. The data for generalization was from Ref^[14], and the results of generalization were exhibited in Figure 2..

Conclusions can be obtained from Figure 1 and Figure 2 that the training results of the “five inputs and one output” model were better, which in return reflected the parameters of the model were reasonable. Although the result of generalization was not as good as training, the trend of AFT was nearly identical. The average deviation of generalization was only 4% and the effect was good.

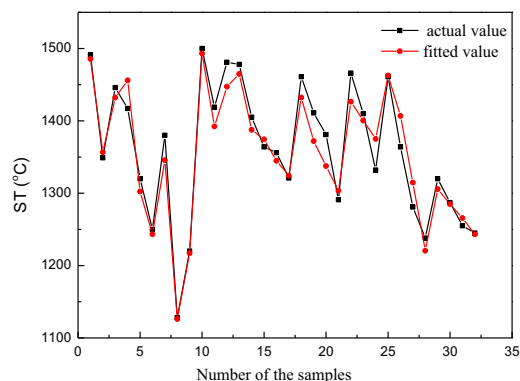


Figure 1. Training results of “five inputs and one output”

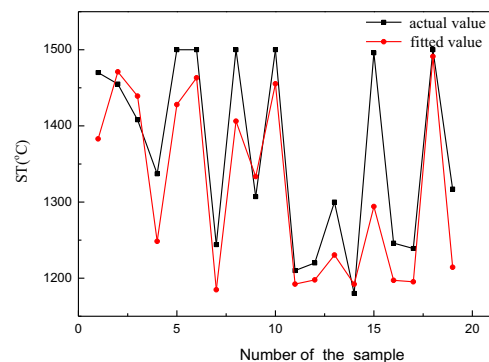


Figure 2. Generalization mapping of “five inputs and one output”

2.2 The “six inputs and one output” model

The so-called “six inputs and one output” was to serve the content of six coal ash compositions, including SiO₂, Al₂O₃, Fe₂O₃, CaO, MgO and TiO₂ as inputs, and serve the ST of coal ash as output to train the model. The 30 data for training were the first 15 sets of literatures^[15,9]. And the rest data were used for generalization. The results were showed in Figure 3 and Figure 4.

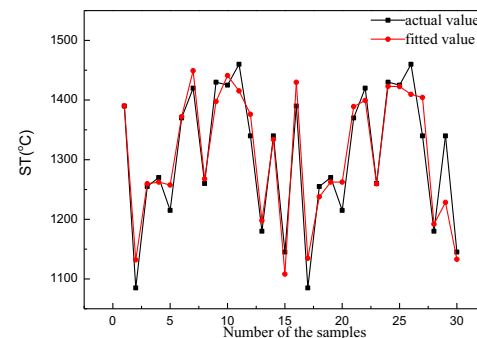


Figure 3. Training results of “six inputs and one output”

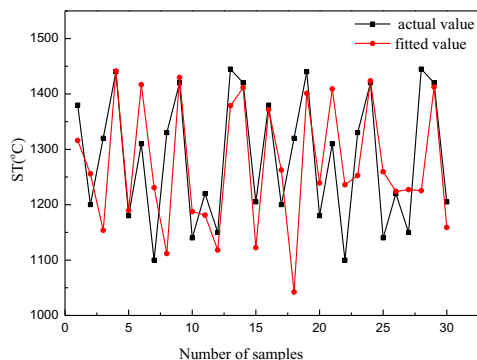


Figure 4. Generalization mapping of “six inputs and one output”

In the AFT model, the selection of data which was a mixture of different areas had a certain impact on the ability of generalization. Because in the process of training, the model studied the data characteristic of different regions, so it had a comprehensive effect. The results of training and generalization showed in Figure 3 and Figure 4 were better.

2.3 The “ten inputs and one output” model

Compared with the “five inputs and one output”, the inputs of the “ten inputs and one output” had the contents of SiO₂, Al₂O₃, Fe₂O₃, CaO, MgO, TiO₂, K₂O, Na₂O, P₂O₅, SO₃. The 30 sets of data for training were from literature [16] and 18 sets of data for generalization was obtained from literature [17]. The results for training and generalization were showed in Figure 5 and Figure 6.

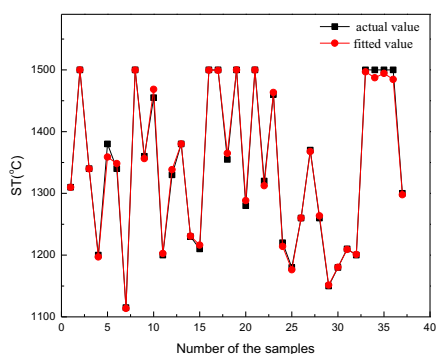


Figure 5. Training results of “ten input and one output”

In Figure 5 it can be seen intuitively that the results were the best, whereas the generalization in Figure 6 were compared, individual point was far from the actual value. That maybe because the training results were so good that made the demand of data for generalization become strict and result in the decrease of its generalization effect.

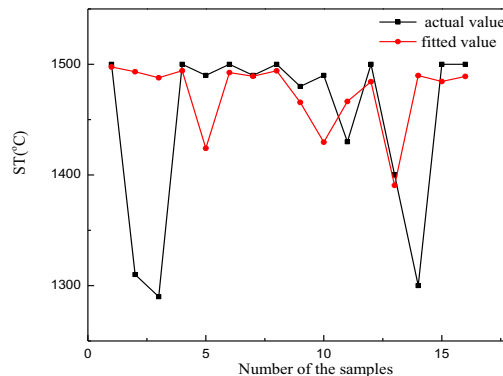


Figure 6. Generalization mapping of “ten inputs and one output”

3. Analysis of the BP neural network

3.1 Analysis of the results of models

The three models above were established based on the characteristic of coal ash data. From training and generalization figures, it can be seen that the accuracy of training was higher but the accuracy of generalization was varied widely. The study analyzed the characteristic of the models from the viewpoint of deviation. The calculation formula of relative deviation was as follows:

$$Deviation = \frac{|ST_a - ST_f|}{ST_a} \quad (1)$$

In the formula, ST_a was the actual value of coal ash softening temperature, ST_f was the fitted value of coal ash softening temperature. And the results of deviations were listed in Table 1.

As shown in Table 1, it’s obviously that deviations in different models were large. The study served the average deviation of generalization as a judgment standard, the model “five inputs and one output” and “ten inputs and one output” were better, but the data used in the former were much more than the later, comparing with the figures of generalization, the model “five inputs and one output” was the best.

3.2 The influence of inputs in a model

In the “ten inputs and one output” model, the contents of coal ash compositions were complete except for the five base stock (SiO₂, Al₂O₃, Fe₂O₃, CaO, MgO), it also contained TiO₂, K₂O, Na₂O, P₂O₅, SO₃. The effect weight of the later five oxides via exchanging the inputs of the “ten inputs and one output” model were investigated. When the number of inputs was five, the composition was the five base stocks. When the number of inputs was six, the composition was the five base stock and TiO₂.

By the analysis of deviations in different inputs in a model, it can conclude that deviations decreased slightly

when the number of inputs decreased, and the more the number decreased, the less of the generalization deviations was. Finally, when the number of inputs was five, the effect of generalization was the best. Besides, XD-APC software was adopted to conduct multivariate analysis and gained the weights of every composition. Results indicated that the weights of oxides TiO_2 , K_2O , Na_2O , P_2O_5 , SO_3 approach to zero which can be ignored. And this was agreed with the analysis in Table 2.

On the base of above, the number of inputs in a model was not the more the better. When some independent variables that had little impact on the dependent variable were conducted into the model for training, the accuracy for generalization decreased, and the prediction ability of a model will become worse. Therefore, the number and quality of variables required "fewer and better". As a result, the model finally identified "five inputs and one output".

3.3 The influence of the number of hidden layer nodes

The influence of numbers of hidden layer nodes was crucial to the model. More less number of hidden layer nodes, less information will be obtained, which is disadvantage for the problem. With the increasing number, the training time will be increased. What's more, it maybe resulted in the problem of excessive anastomosis, and the increasing of training deviations would led to the decrease of the generalization ability. Therefore, choosing the number of hidden layer nodes is exceptionally important for the model. In this paper, the establishment of the number of hidden layer nodes was as follows: At the beginning, set less number of nodes to train the model and analyze the deviations, and then gradually increase the number of nodes, and choose the number of hidden layer nodes with the least deviations of training and generalization.

Table 1 Deviation analysis of different models

Models	Five inputs and one output		Six inputs and one output		Ten inputs and one output	
	training	generalization	training	generalization	training	generalization
D_{ave}	1.58%	4.11%	2.10%	5.99%	0.32%	3.81%
D_{max}	3.24%	13.49%	8.34%	21.04%	1.02%	15.33%
D_{min}	0.11%	0.58%	0.04%	0.09%	0.00%	0.05%

Table 2 Number of different inputs deviation analysis of "ten inputs and one output" model

Number of Inputs	5		6		10	
	train	generalization	Train	generalization	train	generalization
D_{ave}	1.00%	3.59%	1.93%	3.70%	0.32%	3.81%
D_{max}	3.70%	16.02%	10.78%	15.90%	1.02%	15.33%
D_{min}	0.00%	0.02%	0.00%	0.04%	0.00%	0.05%

Table 3 Deviation analysis under different hidden nodes

Hidden Layer Nodes	2		3		4	
	training	generalization	training	generalization	training	generalization
D_{ave}	1.42%	10.11%	1.58%	4.11%	0.75%	9.22%
D_{max}	4.06%	22.25%	3.24%	13.49%	2.46%	22.80%
D_{min}	0.01%	0.09%	0.11%	0.58%	0.03%	0.01%

The data and parameters except nodes were the same with that "five inputs and one output" model mentioned in chapter 2.1. And the nodes were set to 2, 3, 4 and the corresponding deviations were listed in Table 3.

The quality judgment standard for a model was mainly depended on the generalization ability. So when the deviations for generalization became the least, an optimal model can be obtained. According to the deviations exhibited in Table 3, when the node was 2, the deviation for generalization was the largest, with the increasing of

nodes, the deviation decreased first and then increased again. When the node was 3, the deviation reached the

minimum. In addition, although the minimum generalization deviation was small when the node was 2 or 4, the maximum deviation was also large. However the maximum deviation was only half while node was 3. So, the optimizing hidden layer node of the model was 3.

3.4 The influence of choices of range

The choices of range included the maximum and minimum value of composition content, every content of composition in one kind of coal ash was different and the same oxide in different coal ash was also inequable.

The influence of range in the "five inputs and one output" model was investigated. And the maximum value

was always a gradient setting, which was exhibited in Table 4. The study also diminished and expanded the range just like MAXI and MAXIII. After the range was

changed, the models were trained and generalized. The deviations were analyzed and the results were shown in Table 5.

Table 4 Range setting of "five inputs and one output"

Range	SiO ₂	Al ₂ O ₃	Fe ₂ O ₃	CaO	MgO	ST
MIN	0	0	0	0	0	1000
MAXI	30	30	30	30	30	1500
MAXII	60	40	30	20	5	1500
MAXIII	60	60	60	60	60	1500

Table 5 Deviation analysis of models with different ranges

Range	MAXI		MAXII		MAXIII	
	training	generalization	training	generalization	training	generalization
D _{ave}	1.32%	10.05%	1.58%	4.11%	1.61%	6.85%
D _{max}	3.69%	23.78%	3.24%	13.49%	4.96%	16.79%
D _{min}	0.00%	0.04%	0.11%	0.58%	0.02%	0.02%

According to figure 3.5, the effect of range for model cannot be ignored. For MAXII, average deviations were less than those of MAXI and MAXIII. The influence was relatively smaller when the range was expanded, while the deviations became clearly larger if the range was diminished. The reason was that the data which beyond the range were changed when the range was diminished, so the accuracy decreased. Because of the indeterminacy of the compositions of coal ash, there would be a part of data exceed the range, which in reverse brought the model deviation. So the appropriate range for the AFT model was of great significant.

3.5 The Comparison between BP network model and empirical formulas

The "double temperature coordinate graph", which is supplied by Liu [18], is adopted for predicting ST and can be expressed by the following expression:

(I) Calculate the coefficients: b, c, e, which is related to the slag components and they can be defined as follows, respectively

$$b = Fe_2O_3 + CaO + MgO + KNaO (b < 55\%) \quad (2)$$

$$c = 1126.4 \cdot Al_2O_3 + 123.2055 \cdot b + 5217.4 \quad (3)$$

$$e = 587.6594 \cdot SiO_2 + 185.5634 \cdot b + 5374.5 \quad (4)$$

(II) Choose different equations with different coefficient values of c and e

When c was not larger than 23142, the temperature T2' can be expressed by

$$T2' = 963.2804 - 0.0183 \cdot c \quad (5)$$

while if c was larger than 23142, then

$$T2' = 0.0183 \cdot c + 136.7196 \quad (6)$$

When e was less than 33273, the temperature T2' can be calculated by

$$T2'' = 956.148 - 0.0161 \cdot e \quad (7)$$

if e was larger than 33273, then

$$T2'' = 0.0151 \cdot e - 103.0041 \quad (8)$$

(III) Finally, the value of ST can be expressed by

$$ST = T2' + T2'' \quad (9)$$

While if the "ten inputs and one output" model with the same data was adopted to investigate the performance of the regression formulas, 20.58% of the maximum deviation can be obtained with the average deviation was 5.77%. Both of which were larger than those with BP model. And the main reason may be attributed to the fact that the empirical formulas were only a rough approximator of linear relationships between the chemical composition of coal ash and its fusion temperature based on current samples. However, the real relationships were so complex that cannot be replaced by the simple empirical equation. And the well-trained neural network can work better than the empirical formulas.

4. Conclusions

The study adopted the advanced industrial process control configuration and simulation software XD-APC to conduct the BP neural network. Based on fundamental of BP neural network, an universal prediction model of coal ash fusion temperature was established and the optimal parameters were obtained and conclude as follows:

1. For the research of the AFT model based on BP neural network, the effect was best with the inputs were the content of SiO₂, Al₂O₃, Fe₂O₃, CaO, MgO, so the model was confirmed "five inputs and one output" model.

2. For the model, the number of hidden layer node was 3. Under the circumstances, the information obtained from this model won't be too less or too much. It has the most flexible ability to predict data changes.

3. The effect of range cannot be ignored, although the influence was not as important as nodes, reasonable choose of range can improve the ability of generalization.
4. The BP model was superior to empirical formulas.

ACKNOWLEDGMENT

We would like to thank China NSF for the Youth for support of the project (No. 21106120).

References

1. Li W D, Li M, Li W F et al (2010) Study on the ash fusion temperatures of coal and sewage sludge mixtures. *Fuel* 89:1566-1572.
2. Li J B, Shen B X, Li H X et al (2009) Effect of coal blending on the ash fusion temperature and the prediction model for the ash fusion temperature. *Clean Coal Technology* 15:66-70.
3. Gao H G, Li B Y (2001) Mathematical model and its solution for steam coal blending. *Journal of China University of Mining & Technology* 30:627-629.
4. Ma ZH H, Zhao ZH G, Huang W H (2010) Stepwise regression analysis of ash fusion in Rujigou Coalmine, Ningxia. *Coal Geology of China* 22:7-9.
5. Sun Q Y, Zhu X D, Tang L H, et al (2005) Analysis of coal ash fusion temperature using polynomial partial regression functions. *Journal of East China University of Science and Technology* 1:005.
6. G. Özbayoğlu, M. Evren Özbayoğlu (2006) A new approach for the prediction of ash fusion temperatures: A case study using Turkish lignites. *Fuel* 85:545-552.
7. Yin CH G, Luo Z Y, Ni M J, et al (2006) Predicting coal ash fusion temperature with a back-propagation neural network model. *Fuel* 77:1777-1782.
8. Li J ZH, Zhou H, Wang CH L, et al (2007) Employing support vector machine to predict the ash fusion temperature of coal blends. *Journal of China Coal Society* 32:81-84
9. Wang CH L, Zhou H, Li G N, et al (2007) Combining support vector machine and genetic algorithm to predict ash fusion temperature. *Proceedings of the CSEE* 27:11-15.
10. Liu Y P, Zhong Y F, Qian J X, et al (2003) Application of ant colony algorithm and BP neural network in prediction of coal ash fusion point. *Thermal Power Generation* 36:23-26
11. Jiang Q Y (2002) Haitong advanced control configuration software (XD-APC) v2002, 019005[p]. Xiamen University Haitong Control co., LTD
12. M Seggiani (1999) Empirical correlations of the ash fusion temperatures and temperature of critical viscosity for coal and biomass ashes. *Fuel* 78:1121-1125.
13. Wang H, He X G (2001) Study of improving generalization of BP network. *Systems Engineering and Electronics* 23:85-87.
14. Li Q Y, Zhou H, Lin A P, et al (2001) Prediction of ash fusion temperature based on grid search and support vector machine. *Journal of Zhejiang University* 45:2181-2187.
15. Li Y (2002) Modeling of power coal blending with application of some non-linear theories. *Zhejiang University, Hangzhou*.
16. Liu X B, Chen CH (1995) Research on ash fusion temperature. *Coal Chemical Industry* 2:48-52
17. Wen X Q, Xu ZH M, Sun L F, et al (2001) Modeling and prediction of the softening temperature of coal ash. *Journal of China Coal Society* 36:861-866.
18. Liu T X, Zhang J Y, Zhang Z X (1992) A new method for the detection of coal and power coal blending 10:177.