Applied Mathematics & Information Sciences
An International Journal

Forecasting Model of Coal Mine Water Inrush Based on Extreme Learning Machine

Zuopeng Zhao[1], *Pei Li*[2] *and Xinzheng Xu*[1]

[1]School of Computer Science and Technology, China University of Mining and Technology, 221008 Xuzhou, China
[2]School of Information and Electrical Engineering, China University of Mining Technology, 221008 Xuzhou, China

Received: 27 Aug. 2012, Revised: 13 Dec. 2012, Accepted: 14 Jan. 2013
Published online: 1 May 2013

Abstract: In order to satisfy the real-time requirement of the coal mine water inrush, comprehensively considering the master influencing factors filtered out by using principal component analysis (PCA) of coal mine water inrush, a forecasting model of coal mine water inrush based on extreme learning machine (ELM) is proposed in this paper by combining with the characteristics of single hidden layer feedforward networks (SLFNs). The model is used to test the samples, and then compare the experimental results of ELM with back-propagation (BP) and support vector machine (SVM). The experimental results show that, compared with BP and SVM, this method is not only learns fast but also has good generalization performance , and thus it can satisfy the real-time requirements of coal mine water inrush effectively. The feasibility of ELM for coal mine water inrush forecast and the availability of the algorithm were validated through experiments.

Keywords: Extreme Learning Machine (ELM), Single-hidden-layer feedforward neural networks (SLFNs), Coal Mine Water Inrush, Forecasting Model, Back-propagation (BP), Support Vector Machine (SVM), Principal Component Analysis (PCA)

1. Introduction

Coal mine water inrush forecast is a complex problem which involve with hydrogeology, engineering geology, mining conditions, rock mechanics and many other factors, and need to be solved in the current coal production. It is difficult to use classical mathematical theory to build forecasting models [1,2].

In general, the existing forecasting analysis methods of mine water inrush are divided into two categories: engineering geomechanics theory and pan-decision-making theory. Yang et al [3] detailed mine water inrush prediction model based on engineering geomechanics, however, the application of the certainty engineering geomechanics method is restricted due to the height of the seam of mining process with pressure (water) system complexity and uncertainty. With the development of computational intelligence, domestic and foreign scholars applied BP algorithm, SVM algorithm and many other algorithms to the coal mine water inrush forecast, and have achieved good effect.

From a mathematical point of view, coal mine water inrush and its influencing factors form an extremely complex nonlinear dynamic systems. By means of the advantages of artificial neural networks in dealing with nonlinear and non-structural and based on a large number of sample instances of coal mine water inrush, Wu [4] developed a neural network forecast model of water inrush by using BP and then use the model to forecast actual coal mine water inrush.Later,the idea that using particle swarm optimization of neural network model to forecast water inrush for the coal mines was suggested by Xue et al [5], avoiding the defects that neural network is easy to fall into local minimum and slow convergence.

Although artificial neural network is widely used, it often cannot control the promotion of the trained network which maybe over-trained and difficult to reach the global optimum, while SVM is a machine learning tool that can solve the multi-dimensional function prediction based on statistical learning theory, so many scholars combined SVM with other methods to build coal mine water inrush forecast model.

A support vector machine using a reduced set (SVM-RS) was presented as a model for predicting water inrush in coal mines by Yan [6]. Continuous valued attributes were discretized by a linear SVM. Preprocessed

* Corresponding author e-mail: zzpeng@cumt.edu.cn, lipeicumt@163.com

data were analyzed by RS, so the method is novel in integrating the advantages of SVM and RS, thereby offsetting their individual deficiencies. In addition, based on the degree of membership of fuzzy theory and SVM,Cao [7] proposed the fuzzy support vector machine model which can be used to assess the water inrush risk from coal floor.

However,the parameters of SVM, such as nuclear function, the error control parameters and the penalty coefficient etc. need to be tuned ,it is not only difficult to determine the parameters but also consume a lot of time to adjust them.

A novel learning algorithm for SLFNs called ELM [8–11] was proposed recently. In ELM, the input weights (linking the input layer to the hidden layer)and hidden biases are randomly chosen, and the output weights(linking the hidden layer to the output layer)are analytically determined by using MooreCPenrose (MP) generalized inverse. ELM not only learns much faster with higher generalization performance than the traditional algorithms but also has good generalization performance .

This paper applied ELM to the coal mine water inrush forecast, and proposed the method of coal mine water inrush prediction model by using ELM based on samples of coal mine water inrush historical data, finally, compared the forecast results of this method with BP and SVM, the experimental results show that the ELM has the advantages of learning fast and good generalization performance in the coal mine water inrush forecast.

2. Extreme learning machine (ELM)

In this section, we describe the essence of ELM [8–11]. This is a unified single hidden layer feedforward network which randomly chooses hidden nodes and analytically determines the output weights of SLFNs.

For N arbitrary distinct samples (x_i,t_i), where $x_i=[x_{i1}, x_{i2}\cdots,x_{in}]^T \in R^n$ and standard SLFNs with N hidden nodes and activation function $g(x)$ are mathematically modeled as

$$\sum_{i=1}^{\tilde{N}}\beta_i g(x_j)=\sum_{i=1}^{\tilde{N}}\beta_i g(w_i x_j+b_i)=o_j, j=1,\cdots,N \quad (1)$$

where $w_i=[w_{i1},w_{i2},\cdots w_{in}]^T$ is the weight vector connecting the i th hidden node and the input nodes, $\beta_i=[\beta_{i1},\beta_{i2},\cdots\beta_{im}]^T$ is the weight vector connecting the i th hidden node and the output nodes, and b_i is the threshold of the i th hidden node. $w_i\cdot x_i$ denotes the inner product of w_i and x_j. The output nodes are chosen linear in this paper.

That standard SLFNs with $\tilde{N}$ hidden nodes with activation function $g(x)$ can approximate these N samples with zero error means that $\sum_{j=1}^{\tilde{N}}\left\|o_j-t_j\right\|=0$, *i.e*, there exist β_i,w_i, and b_i such that

$$\sum_{i=1}^{\tilde{N}}\beta_i g(w_i x_j+b_i)=t_j, j=1,\cdots,N \quad (2)$$

The above N equations can be written compactly as

$$H\beta=T \quad (3)$$

where $H(w_1,\cdots,w_{\tilde{N}},b_1,\cdots,b_{\tilde{N}},x_1,\cdots,x_{\tilde{N}})$

$$=\begin{vmatrix} g(w_1\cdot x_1+b_1) & \cdots & g(w_{\tilde{N}}\cdot x_1+b_{\tilde{N}}) \\ \vdots & \cdots & \vdots \\ g(w_1\cdot x_N+b_1) & \cdots & g(w_{\tilde{N}}\cdot x_{\tilde{N}}+b_{\tilde{N}}) \end{vmatrix}_{N\times\tilde{N}}$$

$$\beta=\begin{vmatrix}\beta_1^T\\ \vdots \\ \beta_{\tilde{N}}^T\end{vmatrix}_{\tilde{N}\times m} \qquad T=\begin{vmatrix}t_1^T\\ \vdots \\ t_N^T\end{vmatrix}_{N\times m}$$

As named in Huang et al.[8,9], H is called the hidden layer output matrix of the neural network.the i th column of H is the i th hidden node output with respect to inputs $x_1,x_2,\cdots x_N$.

Traditionally, in order to train an SLFNs, one may wish to find specific $\hat{w}_i$, $\hat{b}_i$, $\hat{\beta}_i$ ($(i=1,\cdots,\tilde{N})$), such that

$$\begin{aligned}&\left\|H(\hat{w}_1,\cdots,\hat{w}_{\tilde{N}},\hat{b}_1,\cdots,\hat{b}_{\tilde{N}})\hat{\beta}-T\right\|\\ &=\min_{w_i,b_i,\beta_i}\|H(w_1,\cdots,w_{\tilde{N}},b_1,\cdots,b_{\tilde{N}})\beta-T\|\end{aligned} \quad (4)$$

When H is unknown gradient-based learning algorithms are generally used to search the minimum of $\|H\beta-T\|$ ..

According to the ELM theories proved by Huang et al.[8], for fixed input weights and the hidden layer biases seen from Equation (3), to train a SLFNs is simply equivalent to finding a least -squares solution:

$$\begin{aligned}&\left\|H(\hat{w}_1,\cdots,\hat{w}_{\tilde{N}},\hat{b}_1,\cdots,\hat{b}_{\tilde{N}})\hat{\beta}-T\right\|\\ &=\min_{\beta}\|H(w_1,\cdots,w_{\tilde{N}},b_1,\cdots,b_{\tilde{N}})\beta-T\|\end{aligned} \quad (5)$$

Thus, the determination of the output weights is as simple as finding the Least-Square(LS) solution to the given linear system. The mini-mum norm LS solution to the linear system is

$$\hat{\beta}=H^{+}Y \quad (6)$$

where H^{+} is the Moor-Penrose generalized inverse [12] of the hidden layer output matrix H.

3. Using the ELM to construct the coal mine water inrush forecasting model

The existing methods that can construct the forecasting model of water inrush include SVM, BP neural network

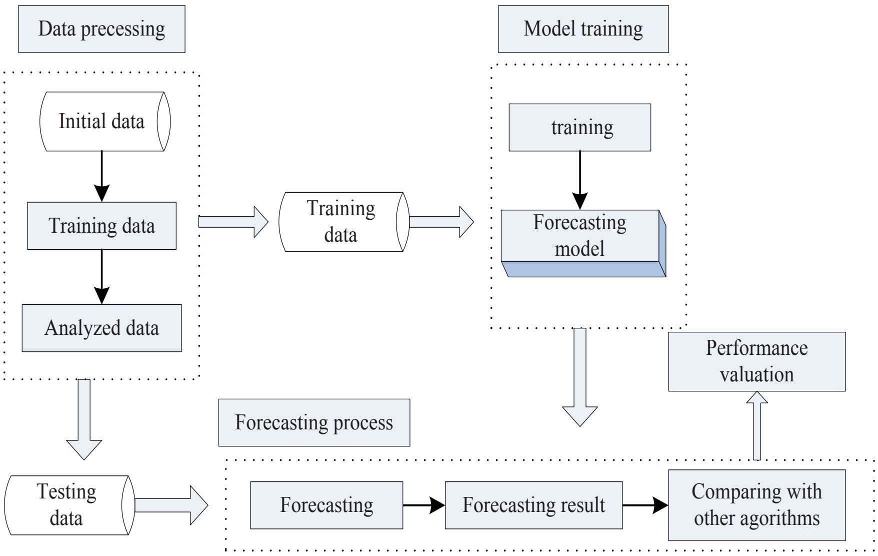

Figure 1 Flow chart of the forecasting model of coal mine water inrush

etc. However, according to the requirements of real-time of coal mine water inrush, these methods which need a long time for training are not appropriate. Therefore, according to the ELM theory described in the previous section, we proposed the forecasting model of coal mine water inrush based on ELM, Figure.1 shows the detailed procedure. It is mainly divided into three main processes: preprocessing data, training model and forecasting.In this paper, we used the PCA method to filter out the main controlling factors that play an important role in the coal mine water inrush among the lots of influencing factors of coal mine water inrush, and then divided the sample data that only contain the main controlling factors into training samples and testing samples. In the training process, we used ELM to construct the coal mine water inrush forecasting model.while in the forecasting process, we used well-established model to predict the testing samples.

3.1. Analysis of the influencing factors of coal mine water inrush [14–16]

There are lots of factors that affect coal mine water inrush, meanwhile the relationship between various factors is very complex.By means of statistical analysis of the large number of cases of coal mine water inrush, the most important factors that control the coal mine water inrush are introduced as follows:

(1)Watery. Watery of the aquifer is the basic condition that determines the size of water of coal mine water inrush and whether the water inrush point can keep water gushing.

(2)Water pressure. In general, when other conditions being equal, the higher the water pressure is, the greater the likelihood of occurrence of coal mine water inrush becomes.The units water inflow of the aquifer is one of the basic factors that evaluate the watery of the aquifer.

(3)Water-resisting layer. Water-resisting layer is the inhibited condition of coal mine water inrush, and its impedance capability depends on the its weight and strength which mainly depends on its thickness and rock properties. Water inrush always occurred in the regions where water-resisting layer thickness is thin.

(4)Fault structure. Geologic structure, especially fault, is one of the main reasons that cause coal mine water inrush. Practice statistics show that water inrush caused by the geological structure account for about69% of all the water inrush accidents, while water inrush occurred along the fault account for 75%.

(5)Karst. It is easy to form karst fissures, caves or subsided column in limestone aquifer which can provide a good living space for the karst water. Water inrush occurs when excavating to the fissure, fracture or other channels and connecting karst water sources.

(6)Mining activities. Mining activities is a predisposing factor of coal mine water inrush and plays a trigger role for coal mine water inrush.Practices proved that the influencing factors which have a pronounced effect on the water inrush including the working thickness, working face dimensions and mining depth.

3.2. Analyzing the impact factors of coal mine water inrush by using PCA

There are many factors that affect coal mine water inrush, but the influencing degree that various factors affect the regional coal mine water inrush is certainly different in the coal mines of different areas, it will not only affect the accuracy of the results but also increase the running time of the machine when using a large number of collected data of a particular coal mine to forecast water inrush.

For the above questions, in this paper, we adopted PCA method to reduce the dimensionality of the original high-dimensional data with many influencing factors. The PCA method not only maximum retain the original information, integrate and simplify the multi-dimensional variable efficiently, but also abandon the problem that the traditional empirical method and regression method determine the right weight insufficiently [17].

According to the principle of PCA, the detailed procedure that use PCA to filter master factors that influence coal mine water inrush is presented as follows [18]:

(1) Raw data standardization. The main purpose is to exclude the impact caused by difference of the order of magnitude and dimension.

$$x_{ij}^{*}=(x_{ij}-\bar{x}_j)/\sigma_i\,(i,j=1,\cdots,n) \tag{7}$$

where x_{ij} is the j th sample of the original data of the i th coal mine water inrush factors,$\bar{x}_i$ and σ_i are the sample mean and the standard of the first coal mine water inrush factors of the i th coal mine water inrush factor respectively.

(2)According to the standardized data sheet $\left(x_{ij}^{*}\right)_{p\times n}$, calculating the correlation coefficient matrix $R=(r_i)_{p\times n}$

$$r_{ij}=\frac{1}{n}\sum\frac{(x_{ij}-\bar{x}_i)(x_{ij}-\bar{x}_j)}{\sigma_i\sigma_j} \tag{8}$$

(3)Calculating the eigenvalues, the eigenvalue contribution rate and the accumulated contribution rate of R.

(4)Determining the principal component number, expression and eigenvectors, and then calculating the impact factor weights. We got the main controlling factors of the coal mine water inrush by using SPSS software to achieve principal component analysis, and selected m principal components in accordance with the principle that characteristic root is greater than 1.After summing the coefficient of the factor i in the m principal components and the products of the variance contribution of each principal component,the absolute value is taken as weights of the factor, that is setting the variance contribution rate of main component q as g_q, and the coefficients of factor q as a_{qi}, so the weight of the factor i is W_i:

$$W_i=\left|\sum_{a=1}^{m}g_q\cdot a_{qi}\right| \tag{9}$$

We can obtain the main controlling factors that determine the coal mine water inrush by comparing the impacting degree of various influencing factors on the coal mine water inrush via the right weight of each factor.

The experiment below shows that this method is simple and effective, and greatly improves the predictive ability of the model.

3.3. Training coal mine water inrush forecasting model by using extreme learning machine

By means of PCA, we can select the watery of the aquifer, water pressure, the thickness of the water-resisting layer, karst, whether the fault, mining thick that a total of six influencing factors as the input parameters of the coal mine water inrush. The principle of parameter values is that the parameter will be expressed in quantitative data if the parameter can be quantitatively, otherwise expressed in binary mode.

Thus we can determine the network structure of water inrush prediction model as follows: 5 input nodes and 1 output node, the numbers of hidden nodes related to the number of input layers and output layers.

According to ELM theory that has been introduced above, it can be drawn the steps that ELM trains the forecast model. Table.1 shows the detail procedure:

Table 1 Training prediction model by using ELM

For N arbitrary distinct samples (x_i,t_i),i=1,$\cdots$,N ,and $\tilde{N}$ hidden nodes and activation function $g(x)$:
Step 1:Randomly assigning the input weights w_i and hidden layer biases b_i where i=1,$\cdots$,$\tilde{N}$;
Step 2: Calculating the output matrix H according to Equation (3);
Step 3: Calculating the output weights β :$\hat{\beta}=H^{+}T$,where $t_i=[t_{i1},t_{i2}\cdots,t_{in}]^T$.

Known by the ELM theory, the number of hidden nodes and the activation function have been assigned at the beginning of learning. The information indicate that there is no better way to auto-complete the hidden node, but finalizing an optimal choice after several experiments. The specific experiments was introduced in the section 4.2.

Once the parameter β_i,w_i, and b_i are determined, It will get the coal mine water inrush forecasting model, and then we can apply this model to predict the testing samples.

4. Experimental evaluation

4.1. Selecting the sample data

In order to construct the model, we can use water inrush data of a mining area face to experiment, and predict the status of coal mine water inrush by using ELM,BP and SVM. All the three algorithms modeling with the same dataset that a total of 100 data which is divided into learning samples and testing samples, of which 80% of the data (80 training samples) for modeling, 20% of the data (20 testing samples) for testing model.

Each sample contains 6 inputs: the watery of the aquifer, water pressure, the thickness of the water-resisting layer, karst, whether the fault, mining thick, and one output: whether the coal mine burst water. Various factors that affect water inrush should be normalized before constructing a predictive model because that water inrush data types and physical dimensions of the influencing factors are different. In our experiments, all the inputs have been normalized into the range [0,1]; while the outputs have been normalized into [-1,1].

4.2. Parameter Selection

In this section, the parameters that the ELM, BP and SVM used modeling are introduced. All experiments on ELM, BP and SVM are carried out in the MATLAB 7.6 environment. It has a very efficient implementation of SVM provided by Libsvm package which has been used in our simulations for SVM, the kernel function used in SVM is radial basis function, it is easy to achieve classification.BP algorithm has been integrated in the MATLAB Neural Network Toolbox, so it can be used directly, The activation function used in BP algorithms is sigmoid function.

Whereas the ELM only need to select the number of hidden nodes under the condition of fixed activation function, the forecasting performance of the model will be different significantly if the number of hidden nodes is different. Therefore, we chose different activation functions (Sine function, Sigmiodal functions, Hardlim function, Triangular basis functions and Radial basis function). For five activation functions, the number of hidden nodes was gradually increased from 5 to 125 with the interval 5, and then the optimal number of nodes can be selected.Figure.2 shows the specific analysis of the results.

As observed from Figure.2, with the increase of hidden nodes, except hardlim, the testing accuracy increases at first and then decreases when the activation function is any of the other four functions. When the number of hidden nodes is 25, the testing accuracy corresponding to the four activation functions reaches maximum, and is higher than the testing accuracy when

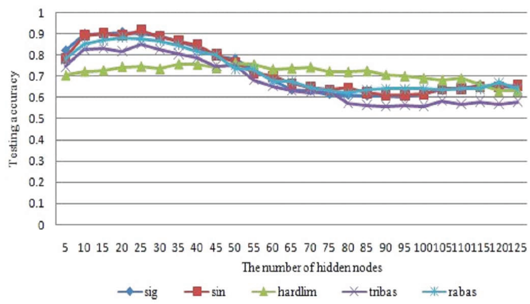

Figure 2 The testing accuracy corresponding to different activation functions

the hardlim as the activation function. When the number of nodes is 80, the testing accuracy reaches minimum, and then tend to a stable surface.When choosing number of hidden nodes around 25 and selecting sine as activation function would be the best, and the testing accuracy reaches up to 91.6%.

In the end, we chose sine function as the activation function and 25 as the number of hidden nodes and calculate average training time, average training accuracy, average testing time, average testing accuracy and number of hidden nodes which were selected as the performance evaluation of the established water inrush model based on ELM of the 50 trials of experiments.

4.3. Comparative experiments that before and after analyzing the sample dataset by using PCA

Training performance of the model would greatly reduce if the water inrush data collected in the site is directly used to establish the water inrush forecasting model.

To tackle the issue mentioned above and improve the performance of model, we used PCA to reduce the dimensionality of the original data before the data was used for training the method. The experiment results verified the effectiveness of this method.

As illustrated in Figure.3, with the increase of the number of nodes, the accuracy before and after dimensionality reduction changing constantly, the training accuracy before dimensionality reduction is always lower than the accuracy after dimensionality reduction, and the accuracy difference even reaches up to ten percentage points when the number of nodes is 35.

Table.3 presents that no matter which kind of algorithms, the testing time before dimensionality reduction cost is always longer than the testing time after dimensionality reduction, which shows that redundant data in the dataset before dimensionality reduction takes up the running time of the machine and slows down the testing speed while the testing accuracy is also reduced. Therefore, it is necessary to filter out the main controlling

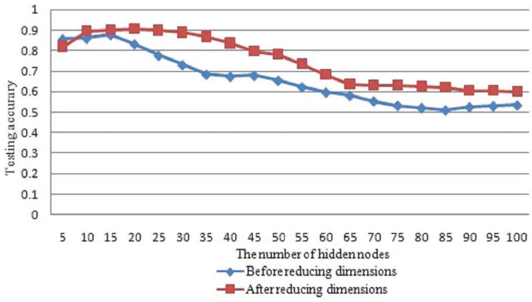

Figure 3 Testing accuracy comparison before and after dimensionality reduction

Table 2 Testing results comparison for ELM and BP, SVM before and after dimensionality reduction

	Before		After	
	Time(s)	Accuracy(%)	Time(s)	Accuracy(%)
ELM	0.0029	76.71	0.0013	91.6
BP	0.0358	58.45	0.0154	65.79
SVM	0.0065	80.00	0.0032	90.00

factors of the original coal mine water inrush data by using PCA method.

4.4. Comparison experiments of training model by using different algorithms

Based on the optimal parameters selected in section 4.2, we trained the samples by using ELM, BP and SVM, and recorded the training time, the training accuracy of three algorithms, the number of hidden nodes, and the number of support vectors of SVM in the experiments.

Table 3 Performance comparisons for ELM, BP and SVM before dimensionality reduction

	Time(s)		Accuracy(%)		No of
	Training	Testing	Training	Testing	SVs/nodes
ELM	0.0097	0.0029	74.35	76.71	25
BP	0.6478	0.0358	74.89	58.45	25
SVM	0.0084	0.0065	75.25	85.00	62

Table.3 and Table.4 respectively give the performance comparison results of the ELM, BP and SVM before and after reducing dimensions of coal mine water inrush data sets by using PCA. As shown in tables, no matter training time or testing time, the time before dimensionality reduction is longer than the time after dimensionality reduction, and the accuracy before dimensionality reduction is much lower than the accuracy after dimensionality reduction.

Table 4 Performance comparison for ELM, BP and SVM after dimensionality reduction

	Time(s)		Accuracy(%)		No of
	Training	Testing	Training	Testing	SVs/nodes
ELM	0.0028	0.0013	84.35	91.60	25
BP	0.5822	0.0154	84.98	65.79	25
SVM	0.0057	0.0032	87.50	90.00	53

The reason why the performance has so big disparity between the different established models of the three algorithms is that there are lots of redundant influencing factors in the water inrush data set before the dimension reduction, but only the master factors play a decisive role in the coal mine water inrush, so filtering out the controlling factors after principal component analysis and then building the model by using the main controlling factors of the sample data set can enhance the prediction performance.

As observed from table.3, compared with BP, the training speed of ELM model is much faster (BP's training time is as 208 times as ELM's training time), it is because that the hidden nodes learning parameters in ELM are randomly assigned and they remain unchanged during the training phase, the learning parameters of ELM are randomly assigned independently in the beginning of learning and not necessarily tuned. The network output weights can be analytically determined by solving a linear system using the least-square method. The training phase can be efficiently completed without time-consuming learning iterations. Although training accuracy of them is similar, all about 85%, the testing accuracy of BP is much lower than ELM. Thus the forecasting performance of ELM model is better than BP.

Comparing ELM with SVM algorithm, no matter training accuracy or testing accuracy, the values both of them are relatively high, especially that the testing accuracy of them is reach up to more than 90%, but the testing accuracy of ELM is higher than SVM even though using less nodes of hidden layers. Both the training time and testing time spent by ELM are shorter than SVM, which all of above show that the overall performance of ELM is better than the SVM.

5. Conclusions

In this paper, we presented a forecasting model of coal mine water inrush based on ELM and verified the feasibility and validity of coal mine water inrush forecast by using the established model. Comparison results of

ELM, BP and SVM show that, the testing accuracy of ELM is much higher than BP,the learning performance of ELM is similar to SVM even though using less number of hidden nodes, and in the process of learning the sample, ELM consumes very little computational time, which all of above illustrate that this method not only learns faster but also has higher generalization performance and prediction accuracy than traditional algorithms. Therefore, this method can well satisfy the real-time requirement of the coal mine water inrush and has actual promotional value.

Acknowledgement

This research was supported by Natural Science Foundation of Jiangsu Provincial (BK2012129).
The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

[1] Biaocan Ling, Discussion on Content of Safety Assessment for Colliery Flood, China Safety Science Journal. **14**, 64-67 (2004).

[2] Shaohe Luo, Jiazhang Qian, Weidong Zhao. Model on water inrush from lateral aquifer into mining pit in. Progress in Mining Science and Safety Technology. 2769-2773 (2007).

[3] Tianhong Yang, Chunan Tang, Zhihong Tan, Wancheng Zhu, Qiyan Feng, State of the Art Inrush Model in Rock Mass Failure and Developing Trend for Prediction and Forecast of Groundwater Inrush, Chinese Journal of Rock Mechanics and Engineering. **26**, 268-277 (2007).

[4] Jinglong Wu, Ju Wang, Chunshan Wang, The Applications of artificial neural network system in the prevention of coal mine water inrush. Coal mine security. 55-57 (2008).

[5] Ru Xue, Guoyong Yang. Forecast for Coal Mine Water Inrush Based on Article Swarm Neural Network, Computer Engineering. **26**, 268-277 (2010).

[6] Zhigang Yan, Peijun Du, Hairong Zhang. Processing Predictors of Water Inrush in Coal Mines Using a SVM-RS Model, Journal of China University of Mining ω Technology. **37**, 295-299 (2008).

[7] Qingkui Cao, Fei Zhao. Risk evaluation of water inrush from coal floor based on fuzzy-support vector machine, JOURNAL OF CHINA COAL SOCIETY. 36, 633-637, 2011.

[8] Huang G B, Zhu Q Y and Siew C K. Extreme Learning Machine: Theory and Applications Neurocomputing. **70**, 489-501 (2006).

[9] Huang G B, Chen L, Convex Incremental Extreme Learning Machine,Neurocomputing. **70**, 3056-3062 (2007).

[10] Huang G B, Chen L. Enhanced Random Search Based Incremental Extreme Learning Machine. Neurocomputing. **71**, 3460-3468 (2008).

[11] Qin-Yu Zhu, A.K. Qin, P.N. Suganthan, Guang-Bin Huang. Evolutionary extreme learning machine, Pattern Recognition. 1759-1763 (2005).

[12] D. Serre, Matrices: Theory and Applications, Springer.New York(2002).

[13] GY Li, WF Zhou, Impact of karst water on coal mining in North China, Environmental Geology. 449-457 (2006).

[14] M.Sun, H S.Yao Li, Analysis and research on main influence factors of deep-well floor water bursting, China Mining Magazine. **18**, 90-92(2009).

[15] Su Gao, ning Zhao, Xiangru iMeng, Water Bursting Mechanism of Coal Floor with Pattern Dynamics. Advanced Materials Research. **18**, 255-260(2011).

[16] Wei Wang, Qinzhong Ma, The Primary Component Analysis Method and Its Application in Earthquake Prediction, EARTHQUAKE RESEARCH IN CHINA. **21**, 409-416 (2005).

[17] Nana Li, Longqing Shi, Zhongjian Li, Zhichao Qu, Yunzhang Guan, Application of Principal-component Analysis in Influence Factor Evaluation of Water-burst from Floor, Coal mining Technology. **16**, 26-28(2011)

[18] JiuwenCao Zhiping Li, Guang-Bin Huanga, Nan Liub, Voting based extreme learning machine. Information Sciences. **185**, 66-77 (2012).

[19] Guorui Feng, Zhenxing Qian, Ningjie Dai. Reversible watermarking via extreme learning machine prediction. Neurocomputing. **82**, 62-68 (2012).

[20] S. Rashid Minhas, Abdul Adeel Mohammed, Q.M. Jonathan Wu. A fast recognition framework based on extreme learning machine using hybrid object information, Neurocomputing. **73**, 1831-1839 (2010).

Zuopeng Zhao received the MS degree in Communication and Information System from China University of Mining and Technology in 2005, and the PhD degree in Cartography and Geographic Information Systems from the Beijing University in 2011. He is currently a vice professor in the China University of Mining Technology. His research interests are in the areas of information technology of coal mine safety, wireless sensor network(WSN), and geographic information system(GIS).

Pei Li received the BS degree in electric engineering and automation from Qingdao Technological University in 2007. She is currently a master at school of Control Theory and Control Engineering, China University of Mining Technology. Her research interests are in the areas of machine learning, intelligent information processing, and coal mine safety information technology.

Xinzheng Xu received his MS degree in Computer Application Technology from Xiamen University in 2005, and his PhD degree in Computer Application Technology from China University of Mining and Technology in 2012. He is currently a lecturer at School of Computer Science and Technology, China University of Mining and Technology. His research interests are in the areas of intelligent information processing, pattern recognition, machine learning, and granular computing.