Oxygen Supply Prediction Model Based on IWO-SVR in Bio-oxidation Pretreatment

Xin Cai, Xin-yuan Nan, Bin-peng Gao

***Abstract*—Oxygen supply is a key parameter important for oxidization rate of pulp and the energy consumption. Therefore, how to increase the use ratio of oxygen and reduce energy consumption has become an important issue in the metallurgical industry. An IWO-SVR prediction model is proposed to overcome the mentioned problem by taking advantages of both Invasive weed optimization (IWO) and support vector regression machine (SVR). IWO is applied to optimize the model's parameters. Then, the sample data is classified by discriminant functions to predict oxygen supply. The experimental results show the IWO-SVR approach obtains better prediction accuracy than the standard SVR approach. So, the proposed model is well suited to predict the oxygen supply in the process of bio-oxidation pretreatment.**

***Index Terms*—oxygen supply, support vector regression, Invasive weed optimization, Prediction**

I. Introduction

PROCESS for bio-oxidation of gold is currently one of the most important pretreatment on refractory gold ores with strong competitive power and bright prospect. In metallurgical industries and mines, power consumption of air compressor system has reached about 10%~20% of the total consumption[1], even to 20% in biological oxidation pretreatment[2]. The process of biological oxidation pretreatment is to make gold exposed through bacteria oxidizing and decomposing pulp with enough oxygen for further extraction of gold. Based on the experience, oxygen supply of each oxidation tank is adjusted manually by reason of complex biological oxidation process and variable composition in the tank. In this way, it leads to low dissolved oxygen and high energy consumption of the system. So, how to increase the use ratio of oxygen and reduce energy consumption has become an important issue.

To overcome this problem, an efficient way is proposed to predict the oxygen supply of every tank using process parameters' data of the pretreatment process. However, the complexity of the biochemical processes has been a great barrier to modeling the bio-oxidation pretreatment accurately. Since the study on bio-oxidation pretreatment is relatively less, only the experience of literature from relative industry can be referred. Least squares support vector machine (LSSVM) was used to predict gas emissions with any change in geological conditions, production processes or the environment of the working face[3]. In[4], support vector machine based K-means cluster was applied to predict temperature of molten iron in blast furnace. Compared with mechanism model concerned silicon content in molten iron, the proposed method had a higher accuracy. For multiple correlations and variability by time among alumina evaporation process parameters, Yang [5] used LSSVM with adaptive weight for the MIMO system to predict export material liquid density online. Li[6] proposed the model of LSSVM improved by artificial bee colony to predict oxidation reduction potential in the biological oxidation pretreatment process, and received a better accuracy prediction. All of above researches lay the foundation for setting up the intelligent prediction model of oxygen tanks.

Invasive weed optimization (IWO) [7]is a novel ecologically inspired algorithm, proposed by Mehrabian, which mimics the process of weeds colonization and distribution, and has very strong robustness and adaptability. The algorithm is simple, easy to implement, and can effectively converge to the optimum solution of problems. Since its inception in 2006, IWO has found successful applications like broadband matching network design[8], unit commitment problem[9], stock price prediction[10], fault diagnosis of analog circuits [11] and so on. In this paper, a combination method of IWO and SVR is presented and used for oxygen supply prediction.

The remainder of this paper is organized as follows: Section II gives the basic principle and mathematical formulation of SVR. Section III discusses the theory and method of IWO. Then the IWO-SVR approach for oxygen supply prediction is proposed in section IV. Finally, conclusions are drawn in Section V.

II. THEORETICAL BACKGROUND

A. Support Vector Regression

Support Vector Regression (SVR)[12] is the extension of SVM to solve regression and prediction problems. The regression problem is to infer output y according to the given new input sample x. A simple description of the SVR algorithm is provided as follows.

Consider a training dataset $\{(x_1, y_1), \cdots, (x_n, y_n)\}$, $x_i \in R^n \times R$, in which n is the total number of training samples. x_i is n dimension input vectors and y_i represents output. The basic idea of SVR is to map the data into a higher-dimensional feature space via a

Manuscript received January 31, 2015; revised March 11, 2015.

Xin Cai is the School of Electrical Engineering, Xinjiang University, Urumqi, CO 830047 China (phone: +86 15739574733; e-mail: mystery.lian@163.com).

Xin-yuan Nan is with the School of Electrical Engineering, Xinjiang University, Urumqi, CO 830047 China (e-mail: xynan@xju.edu.cn).

Bin-peng Gao is with the School of Electrical Engineering, Xinjiang University, Urumqi, CO 830047 China (e-mail: gbp_xd@sina.com).

nonlinear mapping ϕ and then to do linear regression in this space. So, the goal is to find a function $f(x)$ that gives a deviation ε from the actual output y. The optimal decision function is written as

$$f(x) = (w \cdot \phi(x)) + b \tag{1}$$

where w is an adjustable weight vector, $\phi(x)$ is the data in features space and b is scalar threshold. They can be estimated by minimizing the regularized risk function

$$\min(\frac{1}{2}\|w\|^2 + C\frac{1}{n}\sum_{i=1}^{n} L(y_i - f(x_i))) \tag{2}$$

Subjected to $-\varepsilon \le y_i - f(x_i) \le \varepsilon$

Where $\frac{1}{2}\|w\|^2$ is regularized risk which controls the function capacity; $C\frac{1}{n}\sum_{i=1}^{n} L(y_i - f(x_i))$ is empirical risk, the parameter C is a regularization constant tuning the trade-off between the training error and the generalization performance. $L(y_i - f(x_i))$ is loss function which impacts minimize effect of empirical risk and decides how to punish the deviation $y_i - f(x_i)$. The error (ε) -insensitive loss function $L_\varepsilon(y, f(x_i, w))$ defined as Eq.(3)

$$L_\varepsilon(x, y_i, f(x)) = \begin{cases} \varepsilon & if\ |y_i - f(x)| \le \varepsilon \\ |y_i - f(x)| - \varepsilon, & otherwise \end{cases} \tag{3}$$

The problem of finding w and b to reduce the empirical risk with respect to an ε -insensitive loss function is equivalent to the convex optimization problem that minimizes the margin (w) and slack variables (ξ_i , ξ_i^*) as

$$\min(\frac{1}{2}\|w\|^2 + C\frac{1}{n}\sum_{i=1}^{n}(\xi_i + \xi_i^*)) \tag{4}$$

Subjected to $\begin{cases} y_i - f(x_i) \le \varepsilon + \xi_i \\ f(x_i) - y_i \le \varepsilon + \xi_i^* \\ \xi_i \xi_i^* \ge 0 \quad i = 1, 2, \cdots, n \end{cases}$

The above optimization problem is solved by Lagrange multipliers and its solution is given by

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\langle x_i \cdot x \rangle + b \tag{5}$$

where $b = -\frac{1}{2} w \cdot (x_r + x_s)$; α_i and α_i^* are Lagrange multipliers; and n is the number of support vectors.

Nonlinear transformation of SVR is realized by defining the appropriate kernel function

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \tag{6}$$

The function of kernel function is to replace inner product in high dimensional feature space as a bridge between linear and nonlinear. Thus it can avoid dimension disaster generated by complex high dimensional operation and model dimension raise.

The form of $f(x)$ is similar to a RBF neural network, which is expressed as Fig.1. Output is the linear combination of intermediate nodes. The basic difference between SVR and RBF is theoretical bases. The center of every basis function of SVR corresponding to a support vector, which and output weight are decided by algorithm automatically.

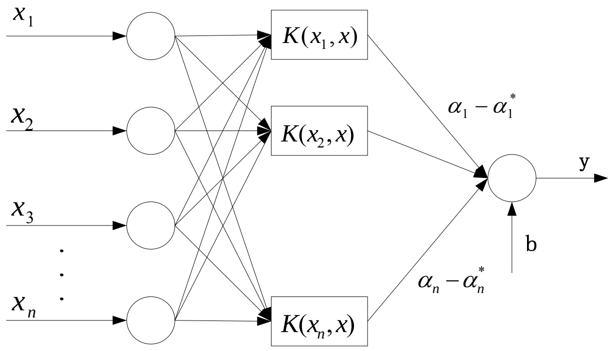

Fig.1 Prediction model based SVR

B. Invasive Weed Optimization

Invasive weed optimization was developed by Mehrabian and Lucas in 2006[7]. IWO is a population-based meta-heuristic algorithm that mimics the colonizing behavior of weeds. Compared with other algorithms, IWO is simpler and has appropriate capability and convergence rate to the global optimal point of objective function. Some of the distinctive properties of IWO in comparison with other evolutionary algorithms are the way of reproduction, spatial dispersal, and exclusive competition.

The process is addressed in detail as follows:

(1) Initialization

A certain number of weeds are randomly spread over the D-dimensional search space. This initial population of each generation will be termed as $X = \{x_1, x_2, \cdots, x_n\}$.

(2) Reproduction

The higher the weed's fitness, the more seeds it produces. Each member of the population X is allowed to produce seeds within a specified region centered at its own position. The number of seeds produced by $x_i, i \in \{1, 2, \cdots, n\}$ depends on its own and the colony's lowest and highest fitness: the number of seeds produced by a plant varies linearly from minimum to maximum possible amounts of produced seeds. The procedure is shown in Fig.2. Namely, the number of produced seeds for the ith plant in every repeat can be calculated with the following equation:

$$S_i = \frac{F_i - F_{\max}}{F_{\min} - F_{\max}}(S_{\max} - S_{\min}) + S_{\min} \tag{7}$$

where F_i is the current weed's fitness. $S_{\max}$ and $S_{\min}$ respectively represent the maximum and the least value of a weed. $F_{\max}$ and $F_{\min}$ respectively represent the maximum and the least fitness of the current population.

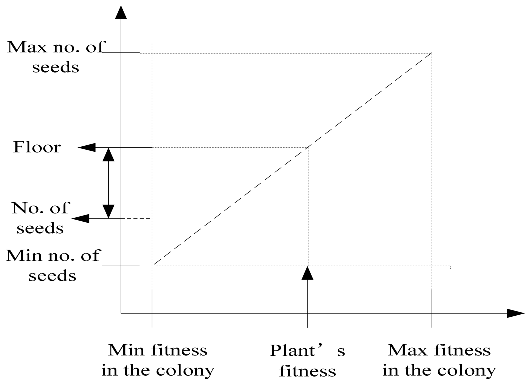

Fig.2 Seed production procedure in a colony of weeds

(3) Spatial dispersal

The generated seeds are randomly distributed over the D-dimensional search space by normally distributed random numbers with mean equal to zero, but with a varying variance σ^2. This means that seeds will be randomly distributed so that they abide near to the parent plant. The position of new seed is given according to

$$\sigma_t = \sigma_{\min} + (\frac{t_{\max} - t}{t_{\max}})^{pow} \cdot (\sigma_{\max} - \sigma_{\min})$$

(8)

where $t_{\max}$ is the maximum number of iterations allowed, t is the current iteration number and *pow* represents the non-linear modulation index.

(4) Competitive exclusion

If a plant leaves no offspring then it would go extinct, otherwise they would take over the world. Thus, there is a need of some king of competition between plants to limit the maximum number of plants in a population. Initially, the number of weeds in a colony will reach its maximum (*pop_max*) by fast reproduction. In this situation, every plant is permitted to produce seeds by in accordance with reproduction method. However, it is expected that by this time the fitter plants have reproduced more than undesirable plants. From then on, only the fittest plants, among the existing ones and the reproduced ones, are taken in the colony and the steps 1 to 4 are repeated until the maximum number of iterations (or function evaluations) have been reached. So, in every generation the population size must be less than or equal to *pop_max*. This method is known as competitive exclusion and is the selection procedure of IWO.

III. IWO-SVR IMPLEMENTATION FOR PREDICTION OF OXYGEN SUPPLY

The flow chart of bio-oxidation pretreatment is showed in Fig.3. Oxidation tank 1 to 3 are used for the first stage oxidation treatment, tank 4 and tank 5 are respectively used for second and third one. Oxygen supply of oxidation tank at different levels is decided by bacterial activity, so it is affected by temperature, PH, pulp density, oxidation reduction potential (ORP) and so on which influence the bacterial activity[13]. And the data in the collection is nonlinear, high dimension, nonstationary and so forth. Standard support vector machine has features to overcome nonlinear, high dimension, but it is hard to solve other problems. So we propose a prediction model of oxygen supply based IWO-SVR in this paper, considering the following issues:

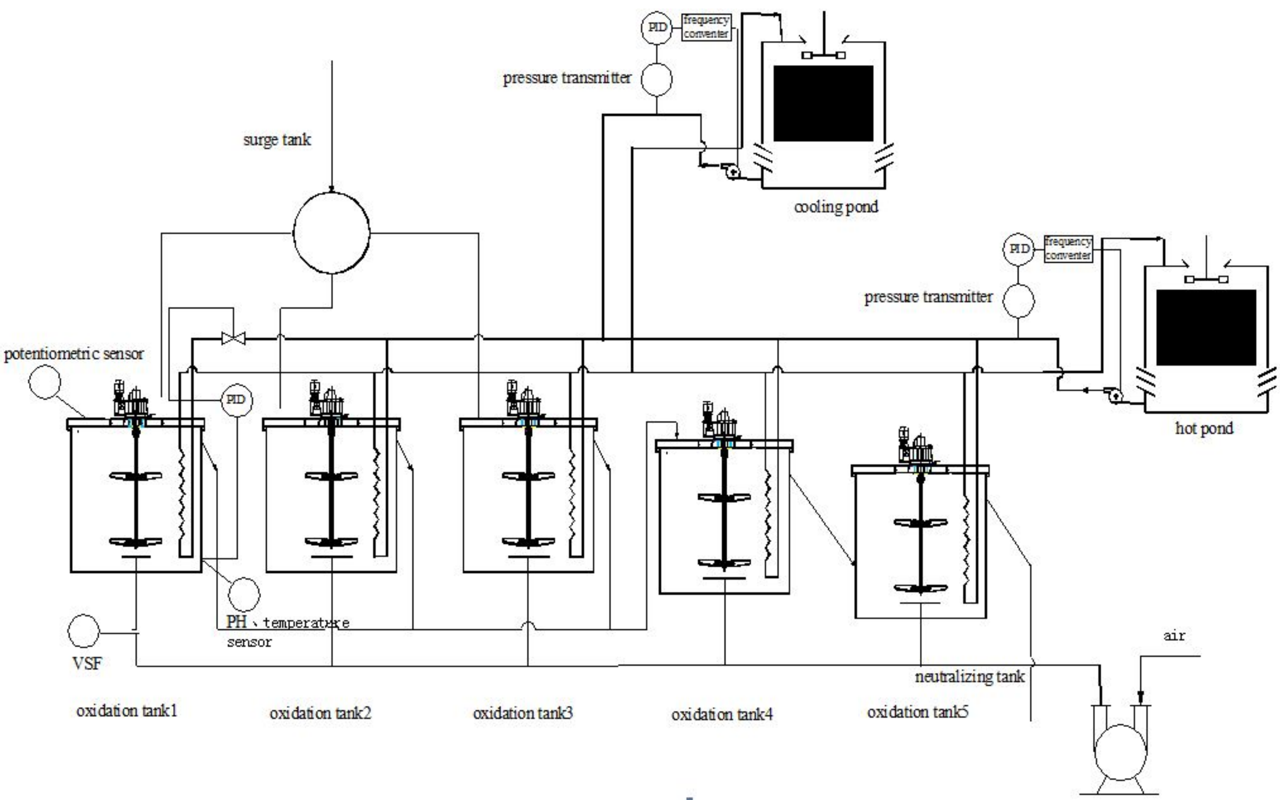

Fig.3 The process of bio-oxidation pretreatment

A. Data Processing

In order to improve prediction precision of the model, data collected in oxidation tank need to be processed in advance.

(1) Outlier processing

Component in biological oxidation tank is complex, oxidation reactions between bacteria and pulp occur under strong acid condition. So data acquired by sensors deviates from the real value due to external disturbance. It may extremely affect data structure and distribution of the whole system, if outlier exists. Wiping out the outlier is an important way to improve accuracy of the model.

According to the error processing criterion in statistical discriminants, given sample data $X=(x_1,x_2,\cdots,x_n)$, average is $\bar{x}$, deviation is $v_i=x_i-\bar{x}$ $(i=1,2,\cdots,n)$, calculating standard deviation by Bayes formula $S=\sigma=[\sum v_i^2/(n-1)]^{\frac{1}{2}}$.

If the deviation v_i $(1\le i\le n)$ of one sample data x_i satisfying: $|v|>3\sigma$, x_i is considered as a outlier which should be eliminated.

(2) Normalizing parameters

In order to improve the calculation efficiency, and prevent individual data from overflowing during the calculation, input parameters should be normalized as follows:

$$x_i^j=\frac{\overline{x_i^j}-\min\overline{x^j}}{\max\overline{x^j}-\min\overline{x^j}}\quad j=1,2,\cdots,n \tag{9}$$

where $\overline{x_i^j}$ represents each input parameter, x_i^j accordingly represents normalized input parameter.

(3) Correlation analysis

Data measured in real time of biological oxidation tank includes temperature, PH, pulp density, oxidation reduction potential. From the analysis of production process, the above variables have multiple correlations, strong coupling relationship. Influences of every factor to oxygen supply are not neglected.

There are principal component analysis (PCA) and correlation analysis to deal with multiple correlations. In this paper, support vector machine is proposed to model, it only need to know the degree of relationship between variables. Using Pearson correlation analysis[14], it's showed as follow:

Table 1. Correlation coefficient

	Y	X1	X2	X3	X4
Y	1				
X1	-0.357	1			
X2	0.468	-0.344	1		
X3	0.018	0.002	-0.3	1	
X4	0.043	-0.073	-0.374	-0.093	1

where X1…X4 respectively represents PH, oxidation reduction potential, temperature, pulp density, Y represents oxygen supply. From Table 1, oxygen supply is positively correlated to oxidation reduction potential, temperature and pulp density, is negatively correlated to PH. According to the analysis, oxidation reduction potential, temperature, pulp density, PH are chosen as input, oxygen supply is output.

(4) Cluster

Collecting a mass of data affecting the oxygen supply, still hundreds of data is obtained after data preprocessing. It not only increases hard to train SVM, but also influences accuracy of prediction model. So it needs to classify the data and K-Medoid clustering[15] is as follows:

a) select k initial categorical objects from the dataset and assign them as initial medoid of k-cluster;
b) compute the distance between every object and cluster center, distribute the object to the cluster represented by center according to the distance.
c) when all objects are allocated to clusters. Recheck the similarity of all categorical objects with current medoids assigned to all clusters. If a categorical object is found which is nearest to some other cluster medoid then de-allocate from the current cluster and allocate categorical object to new cluster & recalculate medoid.
d) repeat c) until no object has changed clusters after a full cycle test of whole data set.

(5) Discriminant analysis

Discriminant analysis is to select variable which can provide more information from the classification of the known observation objects. Set up discrimination function to minimize false rate of identifying the category.

In accordance with the truth of variables effecting oxygen supply, Fisher discriminant analysis[16] is adopted to build the function and analyzed which class predicting data belong to. Then the SVM model is applied to predict oxygen supply corresponding to the class.

B. IWO-SVR Prediction Model

Penalty parameter C, kernel parameter σ of RBF and insensitive loss parameter ε have a great effect on prediction results of SVR model. In order to get the SVR model with better prediction performance, C, σ, ε are optimized by other algorithm.

Because of global and local search of IWO, competitive exclusion between child and parent to avoid premature convergence and the local minimum, IWO is proposed to optimize the above parameters. Concrete steps are as follows:

Step1 Collect data from a gold mine, preprocess the data and divide it into training and prediction sample.

Step2 Classify the training sample to 3 classes based on craft and expertise, get final cluster centers.

Step3 The training process according to the theory mentioned in section 1 is applied to SVR with the subset for training, and its optimal decision function can be obtained. Then the prediction process is performed with the subset for testing. The error of prediction will be obtained.

Step4 Optimization for the parameters C, σ, ε are performed according to the principle of IWO algorithm introduced in section 2. The error of prediction is used as the fitness function of IWO algorithm. The process of Step3 and Step4 is repeated until termination conditions are achieved. Then a prediction of oxygen supply can be conducted with the optimized and trained SVM.

Step5 Build discriminant function with Fisher discriminant analysis, differentiate class of prediction sample with the discriminant function.

Step6 Use the corresponding prediction model of the class to predict oxygen supply, and obtain the output. The prediction model is showed in Fig.4.

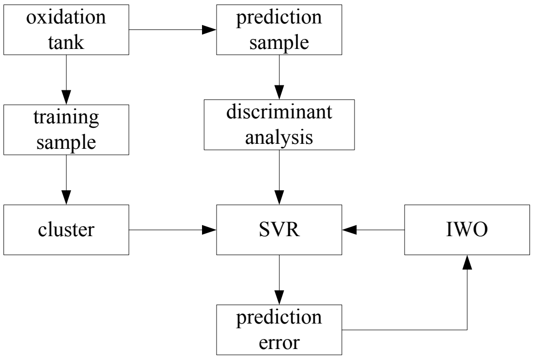

Fig.4 IWO-SVR prediction model

C. Result and Discussions

Choosing online data from a gold mine, 481 groups of data are finally obtained after preprocessing. Former 380 groups are used for training and others for prediction. Training sample is classified into three categories by K-Medoid cluster, and cluster centers are showed as follows:

Table 2. Final class center

	Class 1	Class 2	Class 3
PH	1.83068	1.8831	2.5458
ORP(mV)	468.78041	562.67103	497.99721
Temperature(°C)	39.50311	38.43624	35.40294
Oxygen supply (m^3/h)	1514	784	217

Establish the discriminatory function of classification on the known observation object, get linear discriminant coefficient by using de-normalized Fisher discrimination, showed in Table 3.

Table 3. Fisher linear discriminant factor

	Class 1	Class 2	Class 3
PH	953.953	18080.88	54.121
ORP	24.968	-1.859	-2.276
Temperature	83.72	106.069	168.546
Pulp density	2553.496	1235.921	732.691
(Constant)	-8560.016	-18933.336	-2584.547

Table 4. Classification results

Classification	Prediction members for class 1	Prediction members for class 2	Prediction members for class 3
1	200	0	0
2	0	135	5
3	9	2	130

According to Table 3, the discriminatory functions are:

$$F_1 = 953.953X_1 + 24.968X_2 + 83.72X_3 + 2553.496X_4 - 8560.016$$

$$F_2 = 18080.88X_1 - 1.859X_2 + 106.069X_3 + 1235.921X_4 - 18933.336$$

$$F_3 = 54.121X_1 - 2.276X_2 + 168.546X_3 + 732.691X_4 - 2584.547$$

From the Table 4, the accuracy of classification is 96.3% to show feasibility and effectiveness of discriminatory function.

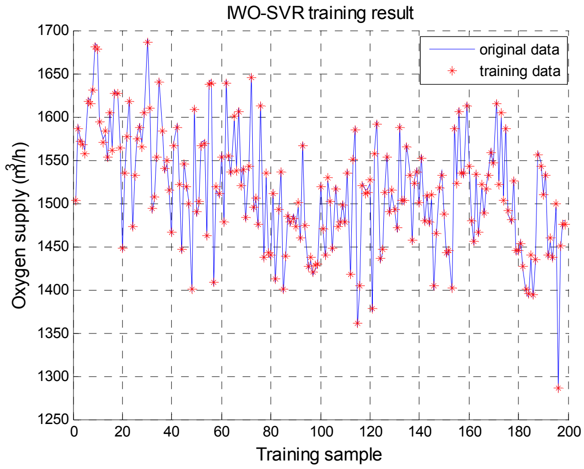

Fig.5. IWO-SVR training results of the class 1

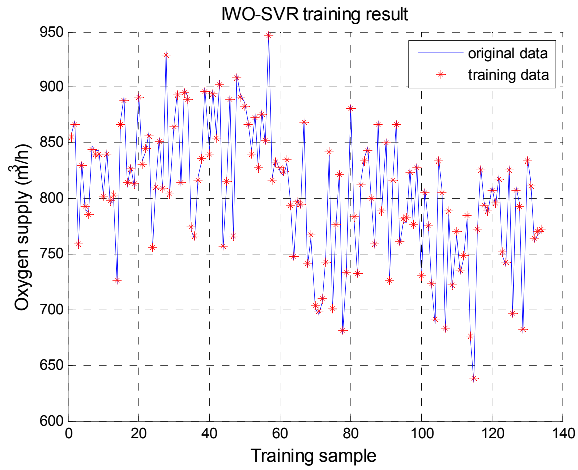

Fig.6. IWO-SVR training results of the class 2

Fig.7. IWO-SVR training results of the class 3

Modeling IWO-SVR for every class, the parameters of C, σ , ε are needed to optimized by IWO, the result obtained is tabulated in Table 5.

Table 5. Parameters of C, σ , ε

	Class 1	Class 2	Class 3
C	0.330	0.189	0.574
σ	0.57	9.19	48.50
ε	0.022	0.010	0.014

From these parameters, the IWO-SVR matched curve of every class with training the sample, respectively showed in Fig.5, Fig.6, Fig.7. From above figures, the prediction model of IWO-SVR has a strong learning ability and good training result.

181 predicting samples are assigned to the class which belongs to by discriminant analysis. Then the sample is applied to predict oxygen supply by means of IWO-SVR model of the class. This prediction method compares to SVR model[17] with the same data, takes 100 samples of class 1 for example, the result is showed as follows:

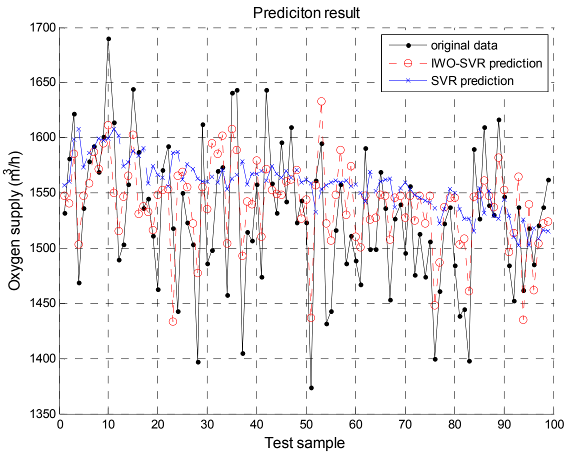

Fig.8. Comparison of the prediction effect

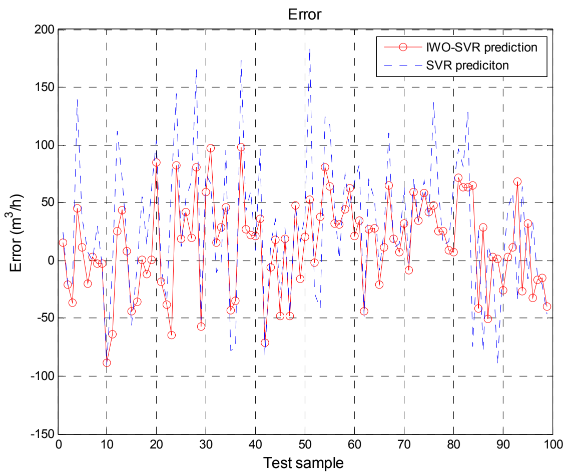

Fig.9. Comparison of the error effect

Figure 8 gives a more significant and comprehensive view of the performance of the IWO-SVR by showing the oxygen supply trend of class 1. Moreover, IWO-SVR model obviously has the lower error of oxygen supply prediction than SVR model in Fig.9. The performance evaluation of the basic and the improved SVR algorithm is showed in the Table 6. The IWO- SVR model is superior to SVR indeed.

Table 6. The performance with IWO-SVR, SVR

Method	Train		Test	
Accuracy	MAPE	RMSE	MAPE	RMSE
IWO-SVR	0.265	4.0394	3.5572	67.4175
SVR	2.9963	58.0529	3.8378	73.0448

IV. CONCLUSION

In this paper we propose an IWO-SVR model, which combines IWO and SVR, for predicting oxygen supply of oxidation tanks. From experimental results, the IWO-SVR model can achieve a better prediction performance compared to the SVR. We consider the intelligent model contributes to adjust oxygen supply automatically to ensure high oxygen rate and reduce energy consumption

ACKNOWLEDGMENT

The authors would like to acknowledge the support of the Natural Science Foundation of China (N0. 61463047).

REFERENCES

[1] Y. Zhao, X. Dai. "Discussions on Energy-saving in the Air Compress System of Underground Mining," *Yunnan Metallurgy*,Vol.35 , pp.3-5,2006.

[2] R. Saidur, M. Hasanuzzaman, N. A. "Rahim. Energy, Economic, and Environmental Analysis of the Malaysian Industrial Compressed-air Systems," *Clean Technologies and Environmental Policy*, Vol.14, pp.195-210, 2012.

[3] X.Y. Gong, et al. "The Research of Gas Prediction Based on LS-SVM and ICA," *Journal of Computational Information Systems*, Vol.10, pp.8273-8283, 2014.

[4] G.M. Cui, T. Sun, Y. Zhang. "Application of Support Vector Machine (SVM) in Prediction of Molten Iron Temperature in Blast Furnace," *Control Engineering of China*, Vol.20, pp.809-812, 2013.

[5] C.H. Yang, et al. "Alumina Evaporation Concentration Prediction Based on Adaptive Weighted LS-SVR," *Control Engineering of China*, Vol.19, pp.187-190, 2012.

[6] W. Li, X.Y Nan. "Research on Prediction of Oxidation Reduction Potential Based on Improved ABC and LSSVM Algorithm," *Computer Measurement & Control*, Vol.22, pp. 3915-3918, 2014.

[7] A. R. Mehrabian, C. Lucas. "A novel numerical optimization algorithm inspired from weed colonization," *Ecological Informatics*, Vol.1, pp.355-366, 2006.

[8] H. Wu, C. Liu. "Broadband matching network design for antennas using invasive weed optimization algorithm," in *Cyber-Enabled Distributed Computing and Knowledge Discovery, 2013 International Conference on, IEEE*, 2013, pp.500-503.

[9] B. Saravanan, E.R. Vasudevan, D.P. Kothari. "A Solution to Unit Commitment Problem Using Invasive Weed Optimization Algorithm," in *the 2013 International Conference on Power, Energy and Control, ICPEC 2013,* February 6, 2013.

[10] H. Liu, X. Lv. "Stock Price Prediction Model Based on Modified IWO Neural Network and its Applications," *Bio Technology: An Indian Journal*, Vol.10, pp. 3918-3930, 2014.

[11] S. H. Cai, et al. "Application of IWO-SVM approach in fault diagnosis of analog circuits," in *Control and Decision Conference (CCDC), 2013 25th Chinese*. Guizhou, May, pp.4786-4791, 2013.

[12] N. Y. Deng, Y. J. Tian. *A novel method of data mining: Support vector machine*, CN: Science Press, 2004.
[13] S. R. Yang. “Basic and Engineering Research of Biological Oxidation of Refractory Gold Ore with Arsenic,” Central South University, Changsha, China, 2004.
[14] B. Y. Zhao, Y. L. Chen, Y. Y. Li, “Correlation Analysis of Fight Test Based SPSS,” *Science & Technology Information*, Vol.15, pp.372, 2009.
[15] 14. L. Kaufman, P. Rousseeuw, “Clustering by Means of Medoids,”1987.
[16] H. J. Chen, et al. “Identifying of Mine Water Inrush Sources by Fisher Discriminant Analysis Method,” *Journal or Central South University: Science and technology*, Vol.40, pp.1114-1120, 2009.
[17] H. Mori, K. Daisuke, “Application of Support Vector Regression to Temperature Forecasting for Short-term Load Forecasting,” in *Neural Networks, 2007. IJCNN 2007. International Joint Conference on. IEEE*, Orlando, FL, Aug., pp.1085-1090, 2007.