

The Prediction of Back Titration Based on Kernel Principal Component Analysis and Radial Basis Function Neural Network

Wu Tiebin

Ddepartment of Energy and Electrical Engineering,
Hunan University of Humanities, Science and
Technology
Loudi, China
e-mail: wutiebin81@163.com

Liu Yunlian

Ddepartment of Energy and Electrical Engineering,
Hunan University of Humanities, Science and
Technology
Loudi, China
e-mail: liuyunlian85@163.com

Long Wen*

Guizhou Key of Economics System Simulation,
Guizhou University of Finance and Economics
Guiyang, China
e-mail: lw770457@163.com
* Corresponding Author

Li xinjun

Ddepartment of Energy and Electrical Engineering,
Hunan University of Humanities, Science and
Technology
Loudi, China
e-mail: lixinjun80@163.com

Abstract- A sampling method on the basis of imitation orthogonalization is proposed to ensure the typicality and ergodicity of samples. Considering the characteristics of back titration (BT) during cobalt removal with arsenic salt, such as many influencing factors and strong coupling, kernel principal component analysis (KPCA) is applied at first. Through KPCA, the effective characteristics of data can be extracted to reduce the dimensions of variables and to eliminate the coupling between variables. Then the extracted characteristic components are utilized as the input of radial basis function (RBF) neural network. Owing to there are many parameters in RBF neural network, which means that it is difficult to obtain the global optimal parameters, rival penalized competitive learning (RPCL) algorithm is adopted first to determine the original values of hidden nodes. On this basis, the improved particle swarm optimization (IPSO) is employed to select the parameters of RBF neural network. It is proved by the simulation results of industrial data that the BT prediction model is effective.

Keywords- Sampling method; KPCA ; RBF; IPSO; BT

I. INTRODUCTION

Zinc hydrometallurgy is the main technology of zinc smelting to reduce the concentration of cobalt ions in zinc sulphate solution to the required technology range by adding moderate arsenic salt, zinc powder and sulfuric acid into the solution [1]. During cobalt removal with arsenic salt, zinc ions are likely to be hydrolyzed to generate basic zinc carbonates. If there are too many basic zinc carbonates, the oxidation-reduction reaction between cobalt ions and zinc powders is expected to be prevented; on the contrary, too few basic zinc carbonates will induce the direct reaction between the hydrogen ions in solution and zinc powders to waste a large amount of zinc powders. In the process of production, back titration (BT) [2] is applied to measure the pH of solution. While the efficiency

of cobalt removed by cementation can be influenced if the BT is too high or too low. Although the BT of the solution in the reactor for cobalt removal with arsenic salt can be adjusted by adding waste sulfuric acid after electrolysis, the BT in the cobalt removal reactor is tested every two hours. Besides, the added waste acid fails to immediately influence the evaluation indexes of solution, including BT. Moreover, there are 2 hours of lag in BT test. But if the waste acid is added unreasonably, the BT of a lot of solution will fail to meet the requirements. Thereby, it is necessary to predict whether the BT in reactor is reasonable or not. Neural network is widely applied to predict the key parameters which are difficult to be detected online or with long time delay during industrial process [3,4]. Among them, radial basis function (RBF) neural network has some advantages, including simple structure, fast convergence rate, favorable generalization capability and strong nonlinear approximation ability, so it has been widely used. RBF neural network is based on the regression modeling of data, which means that the computing speed and precision of the model can be influenced if there are too many data characteristics or the data characteristics are with little correlation. Due to the following characteristics, including many influencing factors on BT, high redundancy, and strong nonlinear coupling relation, to effectively extract the data characteristics and to reduce the dimensions of variables plays an important role in improving the performance of prediction model. Therefore, rival penalized competitive learning (RPCL) [5] is utilized first to extract the characteristic components including data information. And then the extracted characteristic components are served as the input of RBF neural network. Meanwhile, the performance of RBF neural network is greatly influenced by some parameters, including the network structure,

namely the number of hidden nodes, the center vector of Gaussian function, base width vector, the weight vector of network, etc. But it is difficult to obtain the global optimum values of these parameters. Thus, RPCL algorithm [5] is adopted to determine the initial values of the hidden nodes. On this basis, the improved particle swarm optimization (IPSO) is employed to select the other parameters of RBF neural network. Finally, the effectiveness of the proposed BT prediction model based on kernel principal component analysis and radial basis function (KPCA-RBF) is verified by industrial data.

II. EXTRACTING THE DATA CHARACTERISTICS BASED ON KPCA

KPCA is a feature extraction method based on kernel method and maps the original variables into characteristic space by kernel function to minimize the sum of the distances from the original variables to nonlinear principal components. The method can overcome the limitation produced during processing nonlinear problems by using linear methods to extract the effective features and reduce the dimensions of sample [6].

Mapping sample $x_i (i=1, 2, \dots, M), x_i \in R^N$ to feature space $\phi(x_i)$ by the nonlinear mapping function ϕ ,

Abbreviated as F, the covariance matrix of the mapping data is C.

$$C = \frac{1}{M} \sum_{i=1}^M \phi(x_i) \phi(x_i)^T \quad (1)$$

Eigenvalue decomposition of covariance matrix.

$$\lambda V = CV \quad (2)$$

In the formula, λ is the characteristic value of the covariance matrix C, and V is the corresponding feature vector.

Linear combination of V for $\phi(x_1), \phi(x_2), \dots, \phi(x_M)$

$$V = \sum_{i=1}^M \alpha_i \phi(x_i) \quad (3)$$

$$\lambda(\phi(x_i) V) = \phi(x_i) CV \quad (4)$$

Define a $M \times M$ matrix K, Where D is

$$K_{ij} = K(x_i, x_j) = \phi(x_i) \bullet \phi(x_j) \quad (5)$$

$$M \lambda \alpha' = K \alpha' \quad (6)$$

Normalized feature vector V^k , Can obtain a sample x corresponding to the main element $p(k)$

$$p(k) = \sum_{i=1}^M \alpha_i'^k (\phi(x_i), \phi(x_j)) = \sum_{i=1}^M \alpha_i'^k K(x_i, x) \quad (7)$$

The principal components extracted by using KPCA are utilized to replace the independent variables of original sample and correspond to the dependent variables of original sample so as to serve as the input of RBF neural network.

III. RBF NEURAL NETWORK

RBF neural network was proposed by J. Moody and C. Darken in the late 1980s, and has been proved that it can approximate arbitrary continuous function at arbitrary accuracy. The network structure of RBF neural network is shown in Fig. 1.

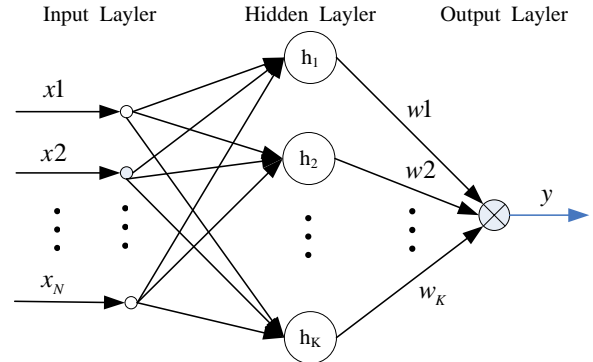


Figure 1. RBF neural network

RBF neural network, $X = [x_1, x_2, \dots, x_N]^T$ is input, $Y = [y]^T$ is output. Radial basis vector $H = [h_1, h_2, \dots, h_k, \dots, h_K]^T$ [7].

$$h_k = \exp \left[-\frac{\|X - C_k\|^2}{2b_k^2} \right] \quad (k=1, 2, \dots, K) \quad (8)$$

$C_k = [c_{k1}, c_{k2}, \dots, c_{kN}]$ ($k=1, 2, \dots, K$) for the center vector of the k node, $B = [b_1, b_2, \dots, b_K]^T$ is wide vector.

ω is the weight vector of RBF

$$\omega = [w_1, w_2, \dots, w_K]^T \quad (9)$$

y is the output of RBF

$$y_m(k) = w_1 h_1 + w_2 h_2 + \dots + w_K h_K \quad (10)$$

The performance of RBF neural network is largely affected by the parameters, including the number of hidden nodes, the center vector of Gaussian function, base width vector, the weight vector of network, etc. Besides, it is difficult to obtain the global optimum values of these parameters. Thereby, RPCL method is applied at first to determine the number of the suboptimal hidden nodes. Then the IPSO [8] is utilized to optimize and set the parameters of RBF neural network.

IV. THE APPLICATION OF KPCA-RBF NEURAL NETWORK METHOD IN BT PREDICTION

A. The influencing factors of BT

The BT in the reactor of cobalt removal with arsenic salt is influenced by many factors or variables. The main factors obtained by using partial correlation analysis are as follows: the flow rate of the liquid after copper removal, the concentration of copper, cobalt and nickel ions in

solution, the flow rate of adding arsenic salt, the molar concentration of NaOH and arsenic element in arsenic salt solution, the BT of thickener underflow and return flow, the centigrade temperature of solution, the concentration of sulphuric acid in waste acid solution, the oxidation-reduction potential (ORP) of No. 1 reactor, the alloy content in the solution of cobalt removal reactor, the additive volume of zinc powders and the flow rate of adding waste acid.

These influencing factors are not independent with each other but most of them have strongly nonlinear correlation, for example, under the same condition, the greater the flow rate of the liquid after copper removal is, the greater the flow rate of adding waste acid and the return flow of underflow will be. Besides, the greater the ion concentration is, the more the zinc powders are added will be within the unit time with the same condition. The strongly nonlinear relationship among these influencing factors, namely variables, is able to impact the performance of prediction model. Hence, KPCA method is applied at first to extract the principal elements to eliminate the strongly nonlinear correlation between samples and to reduce the dimensions of samples.

B. The sampling method based on imitation orthogonalization

There are many influencing factors of BT and they are with great fluctuations. In order to guarantee the prediction effect of the model, it is needed to ensure that the selected learning samples are representative and ergodic within the scope of the parameters for the entry working conditions. Thereby, the authors proposed a sampling method based on imitation orthogonalization, and the specific procedures are as follows:

Step 1. If there are N parameters in each sample, each parameter is divided into a number of segments within the fluctuation range according to the importance and the fluctuation range of parameters. It suggests that the greater the influence of the parameter and the fluctuation range, the more segments the parameter can be divided.

Step 2. The idea of orthogonalization^[9] is introduced to combine the segments of each parameter. Each combination includes 20-100 samples. If numerous combinations appear in the practical production process, more samples will be selected for each combination; on the contrary, fewer samples will be selected.

Step 3. The similarity of samples in each segment is calculated [10]. If the similarity of samples is more than 98%, only the latest one is reserved. In this way, the typicality and ergodicity of sample can be effectively ensured.

C. Simulation analysis

The KPCA-RBF prediction model (K_RBF) is compared with the prediction model based on the basic RBF neural network (B_RBF). Samples are added in the interval of two hours to record the BT in 100 hours continuously, as indicated in Fig. 2.

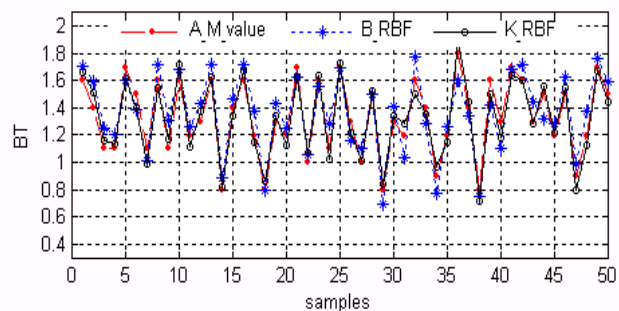


Figure 2. The forecasted results obtained by using these two kinds of prediction models

Table 1 The performance statistics of the results of these two kinds of prediction models

	HK_RBF	G_RBF
Average relative error	4.89%	8.38%
Maximum relative error	12.81%	17.08%
The proportion for the results with relative error less than 10%	96.0%	76.0%

As shown in Fig. 2 and Table 1, A_M_value is the actual measurement value, the performance statistics of KPCA-RBF are the following: the average relative error is 4.89%; the maximum relative error is 12.81%; and the results with relative error less than 10% account for 96%. It can be seen that its prediction accuracy is significantly greater than that of the prediction model based on the basic RBF neural network. The simulation results of industrial data demonstrated that BT forecasting method proposed in this research can completely meet the requirements of industrial application.

V. CONCLUSION

In this research, a KPCA-RBF prediction model is proposed. On this basis, the BT prediction during cobalt removal with arsenic salt is applied to verify the effectiveness of the proposed BT prediction model through industrial data.

VI. ACKNOWLEDGEMENTS

This work was partially supported by The project supported by National Natural Science Foundation of China (NO. 61463009), and science and Technology Department of Hunan Province(NO.2014GK3033), and science and Technology Department of Loudi city, and Scientific Research Fund of Hunan Provincial Education Department(NO.14B097)

*Corresponding Author: Long Wen, E-mail: lw770457@163.com

REFERENCES

- [1] A.Nelson, W. Wang, G. P. Demopoulos, G. Houlachi. The removal of cobalt from zinc electrolyte by cementation: A critical review[J]. Mineral Processing and Extractive Metallurgy Review.2011, 20(4), 325-326
- [2] Wu T B, Yang C H, Li Y G, et al. spent acid adding control in cobalt removal using arsenic in zinc hydrometallurgy. Journal of Central South University (Science and technology) [J], 2013, 44(12):4881-4886
- [3] Bao Y P, Ma J Y, Song Q. Based on grey neural network of sinter basicity

- [4] combination prediction [J]. Control theory and applications,2008, 25(4): 791-793
- [5] L Xu, Krzyzak A, Oja E. Rival Penalized Competitive learning for Clustering Analysis,RBF net,and Curve Detection. IEEE Trans on Neural Networks. 1993,4(4):636-649
- [6] B Schölkopf, A Smola, K R Müller. Nonlinear component analysis as a kernel eigenvalue problem [J]. Neural Computation. 1998, 10(5): 1299-1319
- [7] Yu W, Morales A. Neural networks for the optimization of crude oil blending[J]. International journal of neural systems, 2005, 15(05): 377-389.
- [8] Wu T B. Particle Swarm Algorithm Based on Boundary Buffering-Natural Evolution and its Application in Constrained Optimization[J].Applied Mechanics and Materials, 2014, 670-671, 1517-1521.
- [9] Wang J J. Based on the orthogonal experiment to optimize the parameters of airborne laser radar measurement accuracy design [J]. Chinese laser, 2013,40 (2): 226-233
- [10] WU T B, YANG C H, LI Y G. Improved operational pattern-PID based control of ORP in cobalt removal process using arsenic salt. Journal of Jilin University (Engineering and Technology Edition), 2014, 44(5): 1385-1391]