**Journal of Theoretical and Applied Information Technology**
30th November 2012. Vol. 45 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

# MULTI-CLASSIFICATION OF BALL MILL LOAD STATUS USING VIBRATION FREQUENCY SPECTRUM

**LIJIE ZHAO, DECHENG YUAN**

College of Information Engineering, Shenyang University of Chemical Technology, Shenyang 110142, Liaoning, China

**ABSTRACT**

On-line monitoring and recognition of mill load status has significant effect on the operating efficiency, product quality and energy consumption for the milling circuit. Due to low reliability to recognize the operating states near the boundary region, a multi- classification model is built to identify the operating status of ball mill load. Spectrum features of shell vibration signals are extracted using kernel principal component analysis as input of the multi-classification model. Partial least square-based extreme learning machine model predicts the output coding of ball mill load status. Bayesian decision theory further enhances the reliability and accuracy of the classification model. The proposed method is compared with one-against-one multi-classification strategy and verified with the experimental ball mill. Experimental results show the accuracy and stability of the proposed multi-classification model of ball mill load status outperforms multi-classification with one-against-one strategy.

**Keywords:** *Ball Mill, Mill Load State, Multi-class Classification, Vibration Spectrum*

## 1. INTRODUCTION

Ball mills are large energy-consuming equipments for grinding, which are widely applied in the dry and wet grinding like ores, chemicals, ceramic raw materials and paints [1]. Ball mill load is one of the key factors for the monitoring, control and optimization of the grinding process. It is difficult to online measure the internal load of ball mills because of a series of complex impact and grinding among steel balls and materials, steel balls and lining [2]. Due to the lack of effective ball mill load measurements, a blockage in the ball mill occurs quite often under over load conditions, which results in the "swollen belly" accident and spoiled the normal performance of the ball mill. Reversely, ball mill under the low-load condition may result in very high unit electricity consumption and steel consumption, low production efficiency and high operating costs, even damage of the grinding devices. Therefore, ball mill should be run under optimal operating condition, avoiding the incident of over-load and low-load operating states.

Ball mill load status has significant effect on the operating efficiency, product quality and energy consumption for the grinding process. Measurements on mill load mainly depend on the skilled operators in the industrial fields so that some economic benefits lost in order to ensure equipment safety and process continuity [3].

The unsupervised learning and supervised learning methods are usually used for recognizing the operating conditions of complex industrial process. The unsupervised learning methods don't make full use of the guide of the labeled patterns so that the specific operational states are difficult to be located. However, the supervised learning methods are widely applied in the pattern classification. Pérez et al. (2009) proposed a probabilistic discriminant partial least squares (p-DPLS) method to improve the reliability of the classification by integrating density methods and Bayes decision theory [4]. Zhao et al. (2011) proposed a binary probabilistic extreme learning machine (p-ELM ) classification method to enhance the reliability of classification and avoid the misclassification caused by the uncertainty of extreme learning machine (ELM) predictions [5]. ELM simply generates neural network models using parameterized linear summation of random basis functions [6,7]. Such the idea originated in random vector of the functional-link (RVFL) nets [8,9,10]. However, p-DPLS and p-ELM belonging to the binary classification methods only deal with the classification with two classes.

Multi-classification problem is classifying the real field data into different class libraries. Multi-classification can be solved by dividing the classification problem into several binary classifications. Pérez et al (2010) proposed a multi-classification based on binary probabilistic

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

discriminant partial least squares (p-DPLS) models, developed with the strategy one-against-one (OAO) and the principle of winner-takes-all [11]. Zhao et al. (2012) proposed an OAO probabilistic ELM multi-classification method to identify the operating state of ball mill [12]. Zong et al. (2012) studied the performance of the one-against-all (OAA) and one-against-one (OAO) ELM for classification in multi-label face recognition applications [13]. Though the strategy one-against-one (OAO) overcomes some of the problems of PAQ and OAA, such as misclassification and incompatible classes for the imbalanced samples among the different classes, it increase complexity of model and computational load. Zhao et al. (2012) proposed a partial least square based extreme learning machine (called PLS-ELM) modeling method to enhance the estimate performance in terms of accuracy and reliability [14]. The one-against-one (OAO) strategy overcomes some problems of PAQ and OAA, such as misclassification and incompatible classes for the imbalanced samples among the different classes, However, it increases model complexity and computational load.

Shell vibration signals have potential advantages of process monitoring, control and optimization for high sensitivity and strong anti-interference. However, the unobvious time-domain feature, vibration frequency spectrum with high dimensionality and colinearity may worsen the model performance. In the study, a probabilistic OAO multi-classification strategy using PLS-ELM and KPCA is proposed to identify the status of ball mill load and enhance recognition stability of process states near the adjoining region. Spectrum features of shell vibration signals are extracted using kernel principal component analysis as input of the multi-classification model. Partial least square-based extreme learning machine model predicts the output coding of ball mill load states. Bayesian decision theory further enhances the reliability and accuracy of the classification model. The proposed method is compared with one-against-one multi-classification strategy and verified with the experimental ball mill. Experimental results show the effectiveness of the proposed method.

## 2. KERNEL FEATURE EXTRACTION OF VIBRATION FREQUENCY SPECTRUM

High dimensionality and colinearity of the vibration frequency spectrum are unfavorable to build the effective mill load model in the wet ball mill [15]. KPCA is usually used for nonlinear feature extraction [16]. Features of shell vibration frequency spectrum signals are extracted using kernel principal component analysis (KPCA) before the multi-classification.

Let $X(N\times n)$ be vibration frequency spectrum, where $\boldsymbol{x}_i=[x_{i1},x_{i2},...,x_{in}]^{\mathrm{T}}\in\mathrm{R}^{\mathrm{n}}$ $i=1,\cdots,N$. Feature extraction of vibration frequency spectrum is formulated as the diagonalization of the estimate of the covariance matrix in the high dimensional feature spaces:

$$C=\frac{1}{N-1}\sum_{i=1}^{N}\Phi(X_i)\Phi(X_i)^T \quad (1)$$

where $\Phi(X_i)$ are centered nonlinear mapping of the input variables. An eigenvalue-decomposition of $C$ is computed:

$$Cu_k=\lambda_k u_k \quad , k=1,2,\cdots,N \quad (2)$$

where $\lambda_k$ and $u_k$ represent the $k$th eigenvalue-eigenvector pair of $C$. KPCA circumvents the calculation of mapping function $\Phi(\mathrm{x})$ by the eigenvalue-decomposition of the centered Gram matrix $G=\overline{\Phi}^T(X)\overline{\Phi}(X)\in R^{N\times N}$.

$$\overline{\Phi}^T(X)\overline{\Phi}(X)\mathrm{v}_{\mathrm{k}}=\xi_{\mathrm{k}}\mathrm{v}_{\mathrm{k}} \quad (3)$$

where $\xi_{\mathrm{k}}$ and $\mathrm{v}_{\mathrm{k}}$ represent the $k$th eigenvalue-eigenvector pair of $G$. Assuming the kernel definition $K_{i,j}=\left\langle\Phi(X_i),\Phi(X_j)\right\rangle=\Phi(X_i)^T\Phi(X_j)$, Gram matrix can be computed from the kernel matrix $K$.

$$G=K-\frac{1}{N}KE_N-\frac{1}{N}E_NK+\frac{1}{N^2}E_NKE_N \quad (4)$$

After constructing the KPCA model in the feature space, the kernel score vector for a new sample $\mathrm{x}\notin X(N\times n)$ is given by

$$t=A^T\left(K(X,\mathrm{x})-\frac{1}{N}K_N\right) \quad (5)$$

where $K(X,x)==\left[K(X_1,x),\cdots,K(X_M,x)\right]^T$ represents the kernel vectors and $A=\left[I-\frac{1}{N}E_N\right]V$.

## 3. PREDICTION MODEL OF BALL MILL LOAD STATUS CODING

According to the filed operation experience, there are mainly three kinds of ball mill operation statuses: overload, normal load and low load. In order to facilitate ball mill load status classification model, the output data are coded into integer and sorted according to the classification labels on the operating statuses. Each number of output labels

Journal of Theoretical and Applied Information Technology
30th November 2012. Vol. 45 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

corresponds to a type of the operating status. Prediction model between independent and integer coding based on the ELM-PLS algorithm can reduce the output coding prediction uncertainty to a certain degree. Assume H- and Y- space have been standardized to zeros mean and unit variance. PLS is used for the linear modeling of the relationship between a set of response variables $Y(N\times m)$ and the hidden layer feature space $H(N\times L)$. PLS-ELM model is expressed as

$$Y = H\beta_{PLS} + e \tag{6}$$

where $H$ is the output vector of the hidden layer with respect to input $X$, $e, \beta_{PLS}$ are noise and coefficient matrix, respectively. Coefficient matrix $\beta_{PLS}$ are solved by bilinear decomposition of both the hidden layer feature space $H(N\times L)$ and the response variables $Y(N\times m)$ as

$$\begin{cases} H = TP^T + E = \sum_{k=1}^{h} t_k p_k^T + E \\ Y = UQ^T + F = \sum_{k=1}^{h} u_k q_k^T + F \end{cases} \tag{7}$$

where $T = [t_1, ..., t_h] \in R^{N\times h}$, $U = [u_1, ..., u_h] \in R^{N\times h}$ are latent score vectors with the extracted $h$ principal components in the H- and Y- space, respectively; $P = [p_1, ..., p_h] \in R^{m\times h}$ and $Q = [q_1, ..., q_h] \in R^{m\times h}$ represent the loadings vectors in H-space and Y-space, respectively; $\boldsymbol{E}$ and $\boldsymbol{F}$ are residuals in the H- and Y- space, respectively. If enough latent variables are remained in the PLS-ELM model, residual $\boldsymbol{E}$ and $\boldsymbol{F}$ can equal to zeros. Coefficient matrix $\beta_{PLS}$ is expressed as follows

$$\hat{\beta}_{PLS} = W\left(P^{\mathrm{T}}W\right)^{-1} BQ^{\mathrm{T}} \tag{8}$$

where $W$ is the weight matrix and $B$ is the diagonal coefficient matrix of inner model.

Given $\{ (x_i, y_i) \mid x_i \in R^n, y_i \in R^m, i = 1, \cdots N \}$, the hidden node output function $G(\boldsymbol{a}_i, \boldsymbol{b}_i, \boldsymbol{x}_j)$ and the number of hidden nodes $L$. PLS-ELM algorithm is as follows:

Step 1 Randomly assign the input weights.

Step 2 Calculate the hidden layer output matrix $H$.

Step 3 Scale the hidden layer output matrix $H$ and output $Y$ to zero mean and unit variance.

Step 4 Initialize $E_0 = H, F_0 = Y$, and $k = 0$. Let $k = k + 1$ and take the output scores $u_k$ as some column of $Y_{k-1}$.

Step 5 Compute input weights $w_k$ by regressing $E_{k-1}$ on $u_k$, normalize $w_k$ to unit length and calculate the input scores $t_k$ in the $H_{k-1}$ block:

$$\begin{cases} \omega_k = E_{k-1}^T u_k / u_k^T u_k \\ t_k = E_{k-1}\omega_k / \left\| E_{k-1}\omega_k \right\| \end{cases} \tag{9}$$

Step 6 Compute output loadings $q_k$ by regressing $Y$ on $t_k$, normalize $q_k$ to unit length, and calculate the new output score $u$ in the $Y$ block.

$$\begin{cases} q_k = Y_{k-1} t_k^T / \left\| Y_{k-1} t_k^T \right\| \\ u_k = Y_{k-1} q_k \end{cases} \tag{10}$$

Step 7 Check convergence on $u_k$. If yes, go to step 8, else go to step 5.

Step 8 Compute the input loadings $p_k$ by regressing $E$ on $t_k$, and normalize $p_k$ to unit length in the $H$ block.

$$\begin{cases} p_k^T = t_k^T E_{k-1} \\ p_k = p_k / \left\| p_k \right\| \end{cases} \tag{11}$$

Step 9 Compute inner model regression coefficients.

$$b_k = u_k^T t_k \tag{12}$$

Step 10 Calculate the residuals $E$ of the deflated $H$ and the residuals $F$ of deflated $Y$ in the residual space:

$$\begin{cases} E_k = E_{k-1} - t_k p_k^T \\ F_k = F_{k-1} - b_k t_k q_k^T \end{cases} \tag{13}$$

Step 11 Repeat steps 4 to 11 until all principal factors are calculated.

## 4. PROBABILISTIC DIRECT MULTI-LABEL CLASSIFICATION BASED ON PLS-ELM

Classification is to construct a model to make a map of samples with a peculiar class label. Binary classification model based on PLS-ELM can improve the computational speed due to the easy of accomplishment, low computational cost and high robustness performance. In the binary PLS-ELM model, output Y-block is firstly coded with the integer 1 if the sample belongs to the class of interest (class $\omega_1$) or 0 otherwise (class $\omega_0$). PLS-ELM can not only deal with regression problem, but can be easily generalized to classification problems. The class of the unlabelled samples is judged by

$$Y = sign\left(H\hat{\beta}_{PLS}\right). \tag{14}$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

For the multi-label classification, the output variables of multi-classification model are discrete class labels. The class is classified as index ID corresponding to the maximum output value of the output node of PLS-ELM model. Let $y_i$ be the $i$th output variable of PLS-ELM model, the class label on the sample $x$ is judged by

$$label(\boldsymbol{x}) = \arg \max_{i\in\{1,\cdots,m\}} y_i(x). \tag{15}$$

In the study, binary probabilistic PLS-ELM is directly extended a multi-classification based on PLS-ELM, called the direct multi-classification. It is applied to the multi-label classification of the operating status on the ball mill load.

The standard error of prediction (SEP) is used to account for the prediction uncertainty of the PLS-ELM model. $SEP_i$ for the $i$-th sample is calculated by

$$SEP_i = \sqrt{(1+h_i)\times MSEC_{bc}} \tag{16}$$

where $h_i$ is the leverage for the $i$-th sample and $MSEC_{bc}$ is the bias corrected mean squared error of calibration. The Leverage value is calculated by $h_i = \mathrm{x}_i^T(\mathrm{X}^T\mathrm{X})^{-1}\mathrm{x}_i$. The bias-corrected $MSEC_{bc}$ is calculated as:

$$MSEC_{bc} = \frac{1}{N-n-2}\sum_{i=1}^{N}(\hat{y}_i - y_i - bias_c)^2 \tag{17}$$

$$bias_c = \frac{1}{N_c}\sum_{i=1}^{N_c}(\hat{y}_i - y_i)^2 \tag{18}$$

The potential functions of the training samples for each class are averaged to obtain the class probability density function (PDF)

$$p(\hat{y}|\omega_c) = \frac{1}{N_c}\sum_{i=1}^{N_C} g_i(\hat{y}) \tag{19}$$

where $g_i(\hat{y})$ is probability density function of each calibration sample $i$ for classes $\omega_i$ with the shape of a Gaussian curve, centred at $\hat{y}_i$ and standard deviation $SEP_i$. Parameters of probability density function are estimated by nonlinear least squares[5].

$$g_i(\hat{y}) = \frac{1}{SEP_i\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\hat{y}-\hat{y}_i}{SEP_i}\right)^2} \tag{20}$$

Suppose that the prior probabilities $P(\omega_i) = N_i/N, i = 1,\cdots C$ and the conditional probabilistic densities $p(y|\omega_i)$. For an unknown sample, the probability with prediction $\hat{y}_u$ for the class $\omega_i$ is given by the Bayes formula :

$$P(\omega_i|\hat{y}_u) = \frac{p(\hat{y}_u|\omega_i)\times P(\omega_i)}{p(\hat{y}_u)} \tag{21}$$

Bayes formula shows that the prior probability $p(\omega_i)$ is converted into a posterior probability $p(\omega_i|\hat{y}_u)$ by prediction $\hat{y}_u$.

## 5. RESULTS AND DISCUSSION

### 5.1 Description Of Experimental System

The experiments on the classification of the mill load operating statuses were accomplished on a laboratory scale lattice-type ball mill. The mill has a diameter of 460 mm, a length of 460 mm in length, and driven by a motor. The vibration signals are acquired by the accelerometer with sampling frequency 51,200Hz. The sensor is amounted on the external surface of the mill shell. It has maximum ball load of 80 kg, pulverizing capacity of 10 kg per hour and a rated revolution of 57 per minute. The ball mill was fed with the different steel balls, copper ore and water addition. The feeds were homogenized and each grinding last for one minute. The experimental data are recorded at the different loads. There are three different sizes of the steel balls, i.e. large, medium and small ball (diameter of 30mm, 20mm and 15mm). The particle size and mass of the copper ores varied over a wide range from 1M to more than 8M and 10kg to 50kg, separately. The mass of the water addition varied from 2kg to 50kg.

### 5.2 Kpca Feature Extraction And Coding Prediction Model

In the study, three types of the operating statuses of ball mill load are considered. Class $\omega_1$ denotes low load status, $\omega_2$ normal status and $\omega_3$ overload status. In order to verify the proposed method, experimental data included 51 low load, 48 normal load, 15 overload for training and 51 low load, 48 normal load, 14 overload for testing. Vibration signals in the time domain are transformed into frequency spectrum by Fast Fourier Transform (FFT). Kernel type of KPCA model is RBF_kernel. According to the percent variances captured by KPCA model, 42 principal components are retained and the rest are dropped from the model.

Prediction performance of ball mill load status coding model is compared and analyzed. Fifty runs were independently conducted and the average of the root-mean-square error of prediction (RMSE) and variance ($R^2$) were calculated for each parameter. ELM model has 1 parameter ($L$), i.e. the number of hidden nodes. PLS-ELM model has 2 parameters ($L$, $h$), including the number of hidden

Journal of Theoretical and Applied Information Technology
30th November 2012. Vol. 45 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

nodes and the number of latent variables. ELM model based on KPCA (KPCA_ELM) has 3 parameters (*L*, *h*, *Par*), including the number of hidden nodes, the number of principal components, and kernel parameter. PLS-ELM model based on KPCA (KPCA_PLS-ELM) has 4 parameters (*L*, *h*, *PC*, *Par*), including the number of hidden nodes, the number of latent variables, the number of principal components and kernel parameter. For ELM model and PLS-ELM model, activation function of hidden layer node use sigmoid function, the hidden node number is set 500, initial value of latent variables numbers of PLS-ELM model is set 42.

Performance comparison of ELM and PLS-ELM model based on KPCA feature extraction is illustrated in Table I. It can be observed that prediction performance of KPCA_PLS-ELM model is better than ELM model, PLS-ELM model and KPCA_ELM model. Therefore, KPCA_PLS-ELM model is used as ball mill load status coding model.

*Table I: Performance Comparision Of Elm And Pls-Elm Model Based On Kpca Feature Extraction*

| Prediction Model | Model Parameters | Training | | Testing | |
|---|---|---|---|---|---|
| | | RMSE | $R^2$ | RMSE | $R^2$ |
| ELM | （500） | 0 | 1 | 0.3516 | 0.7992 |
| | （5000） | 0 | 1 | 0.2334 | 0.8924 |
| PLS-ELM | （500, 42） | 0.0007 | 1 | 0.2981 | 0.8569 |
| | （5000, 42） | 0 | 1 | 0.1691 | 0.9571 |
| KPCA_ ELM | （500, 42,500） | 0 | 1 | 0.2165 | 0.9024 |
| | （5000, 42,500） | 0 | 1 | 0.2023 | 0.9146 |
| KPCA_PLS-ELM | （500, 42, 10,500） | 0.0036 | 1 | 0.1434 | 0.9576 |
| | （5000, 42, 10,500） | 0.0035 | 1 | 0.1387 | 0.9607 |

### 5.3 Multi- Classification Models

In this section, the proposed direct multi-label classification and OAO multi-classification strategy are comparatively evaluated. Prior probabilities of the three classes are $P(\omega_1)=0.4474$, $P(\omega_2)=0.4211$ and $P(\omega_3)=0.1316$, respectively. Figure 1 shows that the probability curves of direct multi-classification model based on KPCA_PLS-ELM. It includes the probability density function, the probability density function multiplied by the prior probability and the posterior probability. Posterior probabilities of three classes are calculated for the unknown testing samples according to the Bayes formula.

Enhanced prediction of output coding based on probabilistic KPCA_PLS-ELM model is as shown in Figure 2. It can be observed that the predictions of the samples for class $\omega_1$ are around the coding integer value 1, class $\omega_2$ around the integer 2, and class $\omega_3$ around the integer 3. Probabilistic KPCA_PLS-ELM model improved reliability and generalization of the output coding prediction.

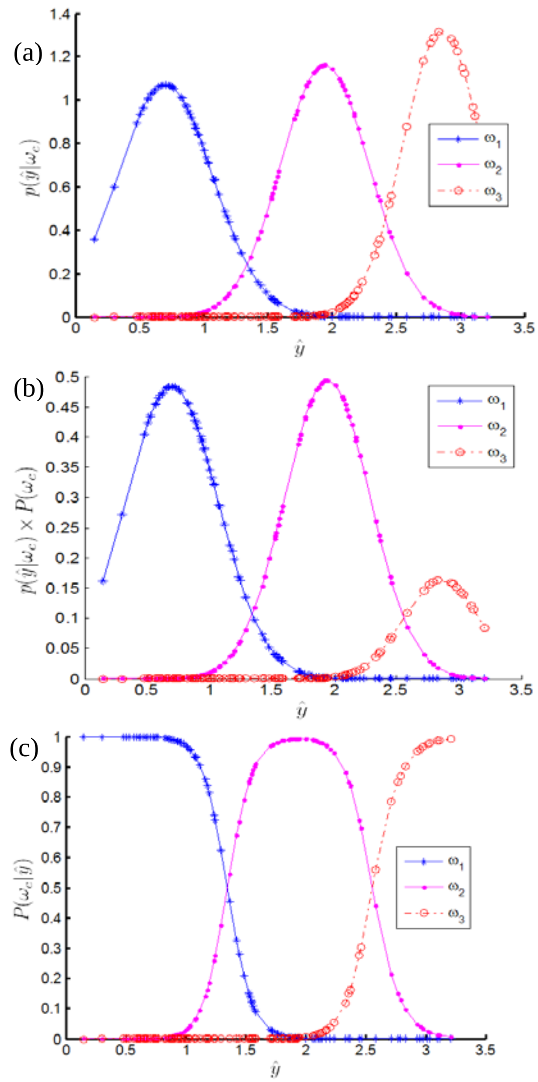


*Figure 1: Prediction Curves Of Probabilistic PLS-ELM Model (A) PDF $p(\hat{y}_u|\omega_c)$ ; (B) Production Function $p(\hat{y}_u|\omega_c)\times P(\omega_c)$ ; (C) Posterior Probabilities $P(\omega_c|\hat{y}_u)$*

Figure 3 shows direct multi-classification results for four testing samples.Figure 3(a) shows probability density function distribution of three classes for three testing samples, where three values

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

in each bracket denote probability density function value of three sample points belonging to classes $\omega_1$ 、 $\omega_2$ and $\omega_3$ , respectively. Figure 3(b) gives the posterior probability distribution curve of three samples, where percentage of three sample points belonging to three classes ( $\omega_1$ 、 $\omega_2$ and $\omega_3$ ) are marked on a side of line.

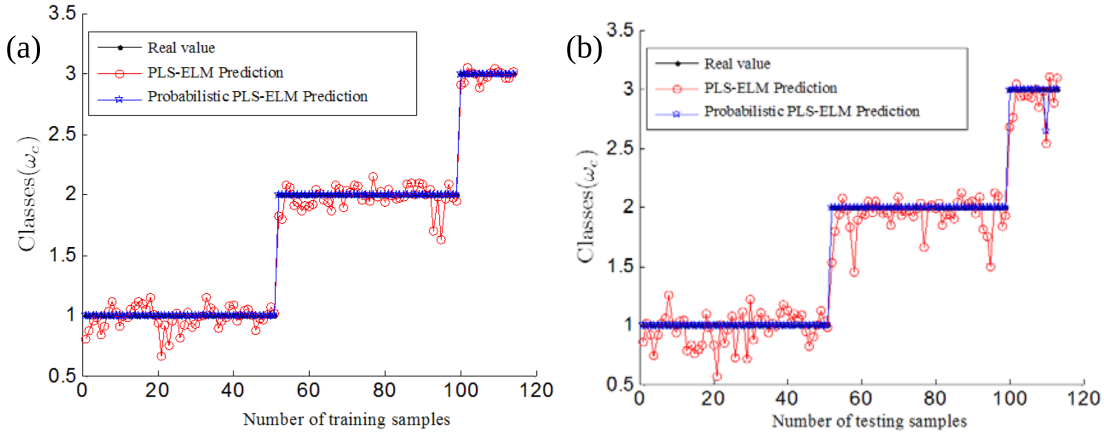


*Figure 2: Enhanced Prediction Of Probabilistic KPCA_PLS-ELM Model (A) Training Model;(B) Testing Model*

In order to evaluate the performance of multi-classification models, direct multi-classification and OAO multi-classification are compared for the training and testing data. Detail results of the multi-classification are shown in Table II. It can be seen that (1) model accuracy using KPCA feature extraction are better than the model based on the full frequency spectrum; (2) classification accuracy based on PLS-ELM model outperforms ELM; (3) direct multi-classification model has higher accuracy than OAO multi-classification model based on P_KPCA_PLS-ELM.

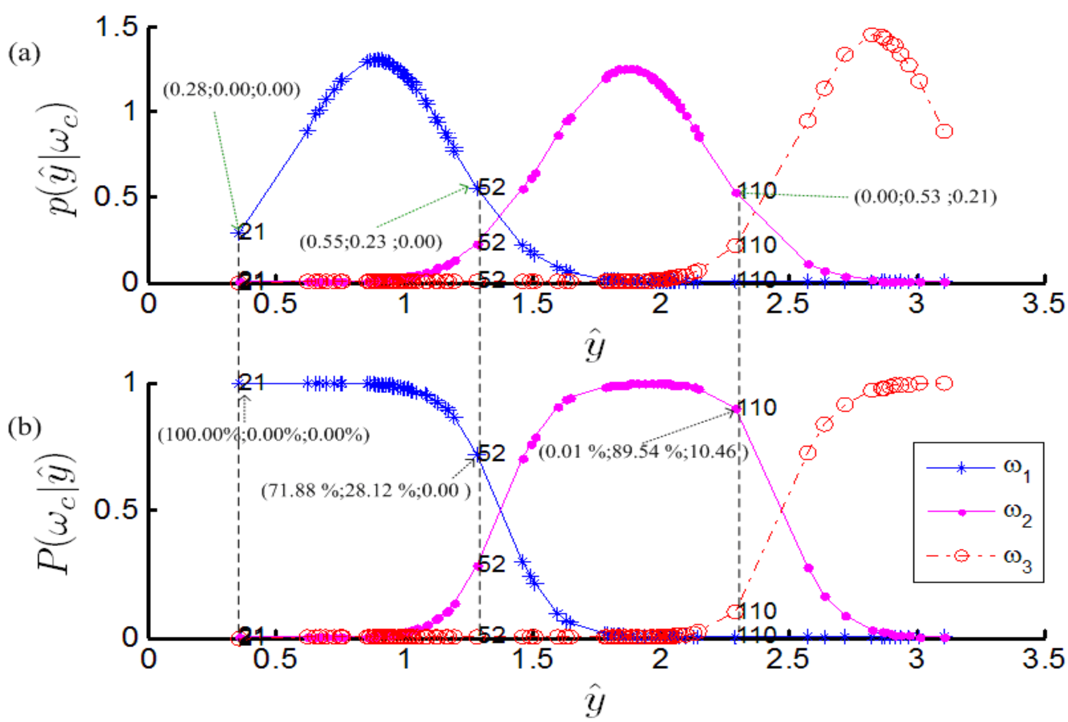


*Figure 3: Direct Multi-Classification Results For Three Testing Samples Based On Probabilistic KPCA_PLS-ELM Model*

*(A) Probability Density Function; (B) Posterior Probabilities*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

*Table II: Performance Comparisons Of Multi-Class Classification Model*

| Method | Model | Training accuracy | Right assigned samples /Total samples | Testing accuracy | Right assigned samples /Total samples |
|---|---|---|---|---|---|
| OAO Multi-classification | KPCA_ELM | 100% | 114/114 | 90.27% | 102/113 |
| | KPCA_PLS-ELM | 94.74% | 108/114 | 92.92% | 105/113 |
| Probabilistic OAO Multi-classification | P_KPCA_ELM | 100% | 114/114 | 95.57% | 108/113 |
| | P_ KPCA_PLS-ELM | 98.25% | 112/114 | 97.35% | 110/113 |
| Direct Multi-classification | KPCA_ELM | 100% | 114/114 | 92.04% | 104/113 |
| | KPCA_PLS-ELM | 100% | 114/114 | 96.46% | 109/113 |
| Probabilistic Direct Multi-classification | P_ KPCA_ELM | 100% | 114/114 | 97.35% | 110/113 |
| | P_ KPCA_PLS-ELM | 100% | 114/114 | 100% | 113/113 |

## 6. CONCLUSION

Based on shell vibration signal with the high sensitivity and less disturbance, a multi-classification model of mill load status based on probabilistic PLS-ELM is proposed to identify the operating status of ball mill load. It is simple and easy to implement. The following conclusions are drawn from the investigations: (1) The multi-classification accuracy using KPCA feature extraction are better than the model based on the full frequency spectrum; (2) The multi-classification accuracy based on PLS-ELM model outperforms ELM; (3) The direct multi-classification model has higher accuracy than OAO multi-classification model based on P_KPCA_PLS-ELM; (4) The direct multi-classification model has less computational cost than OAO multi-classification model.

## ACKNOWLEDGEMENTS

This work was supported by National Natural Science Foundation of China (61203102 and 60874057 ) and National Science Foundation for Post-doctoral Scientists of China (No. 20100471464).

## REFRENCES:

[1]. N. Xu, J.W. Zhou, Q. Wang, "Discussion on intelligent controlling of mineral processing", Copper Engineering, Vol. 107, No. 1, 2011, pp. 54-60.(In Chinese).
[2]. J. Tang, L.J. Zhao, J.W. Zhou, H. Yue, T.Y. Chai, "Experimental analysis of wet mill load based on vibration signals of laboratory-scale ball mill shell", Minerals Engineering, Vol. 23, No. 9, 2010, pp. 720-730.
[3]. P. Zhou, T.Y Chai, "Intelligent optimal-setting control for grinding circuits of mineral processing", IEEE Transactions on Automation Science and Engineering, Vol. 6, No. 4, 2009, pp. 730-743.
[4]. Néstor F. Pérez, Joan Ferré, Ricard Boqué, "Calculation of the reliability of classification in discriminant partial least-squares binary classification", Chemometrics and Intelligent Laboratory Systems, Vol. 95, No. 2, 2009, pp.122-128.
[5]. L.J. Zhao, X.K. Diao, D.C. Yuan, W. Tang, "Enhanced classification based on probabilistic extreme learning machine in wastewater treatment process", Procedia Engineering, Vol. 15, No. 1, 2011, pp.5563-5567.
[6]. G.B. Huang, Q.Y. Zhu, Chee khong Siew, "Extreme Learning Machine: Theory and Applications", Neurocomputing, Vol. 70, No.1-3, 2006, pp. 489-501.
[7]. G.B. Huang, D.H. Wang, Yuan Lan, "Extreme Learning Machines: A Survey", International Journal of Machine Learning and Cybernetics, Vol. 2, No.2, 2011, pp.107-122
[8]. Bons Igelnik, Yoh-Han Pao, "Stochastic Choice of Basis Functions in Adaptive Function Approximation and the Functional-Link Net", IEEE Transactions On Neural Networks, Vol. 6, No. 6, 1995, pp. 1320 -1329.
[9]. Yoh-Han Pao, Gwang-Hoon Park, Dejan J. Sobajic, "Learning and generalization characteristics of the random vector Functional-link net", Neurocomputing, Vol. 6, 1994, pp.163-180.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[10]. L.P. Wang, C.R Wan, “Comments on "the extreme learning machine"”, IEEE Transactions on Neural Networks, Vol.19, No.8, 2008, pp. 1494-1495.

[11]. Néstor F. Pérez, Joan Ferré, Ricard Boqué, “Multi-class classification with probabilistic discriminant partial least squares (p-DPLS)”, Chemometrics and Intelligent Laboratory Systems, Vol. 95, No. 2, 2009, pp.122-128.

[12]. L.J. Zhao, T.Y. Chai, X.K. Diao, D.C. Yuan. “Multi-class classification with one-against-one using probabilistic extreme learning machine”. Lecture Notes in Computer Science, Vol. 7368, 2012, pp.10-19.

[13]. W.W. Zong, G.B. Huang. Face recognition based on extreme learning machine. Neurocomputing, in press, 2012.

[14]. L.J. Zhao, D.H. Wang, T.Y. Chai, “Estimation of effluent quality using PLS based extreme learning machines”, Neural Computing and Applications, Vol. 21, No. 1, 2012, pp. 1-11.

[15]. L.J. Zhao, X. Feng, D.C. Yuan, H. Xiao, “Feature selection of frequency spectrum for the ball mill load based on interval partial least squares”, Journal of Theoretical and Applied Information Technology, Vol. 43, No. 1, 2012, pp.119-126.

[16]. B. Scholkopf, A. Smola, K. R. Muller, “Nonlinear component analysis as a kernel eigenvalue problem”, Neural Computation, Vol. 10, 1998, pp.1299-1319.