Journal of Mechanical Science and Technology 28 (1) (2014) 317~322
www.springerlink.com/content/1738-494x
DOI 10.1007/s12206-013-0982-1

Tool wear prediction considering uncovered data based on partial least square regression †

Guofeng Wang*, Zhiwei Guo and Lei Qian
Key Laboratory of Mechanism Theory and Equipment Design of Ministry of Education, Tianjin University, Tianjin, China

(Manuscript Received March 25, 2013; Revised August 20, 2013; Accepted August 28, 2013)

Abstract

Tool wear prediction plays an important role in guaranteeing the workpiece quality and improving the production efficiency. However, because of the uncertainty and complexity of tool wear process, it is hard to ensure that the samples related to all tool wear values can be collected during the training stage. Therefore, the accuracy of tool wear prediction for these uncovered data will deteriorate severely. In this paper, partial least square regression is presented to realize the tool wear prediction based on force signal. The main characteristic of this method is that the regression analysis is in the principal component space so that the multicollinearity between explanatory variables can be avoided effectively. Side milling experiment was carried out to validate the effectiveness of the proposed model. The analysis and comparison under different number of uncovered data show that the partial least square regression based tool wear prediction is more accurate.

Keywords: Tool wear prediction; Partial least square regression; Multiple linear regression; Neural network; Uncovered data

1. Introduction

During the machining process, the cutting tool will gradually wear out and become blunt due to the friction, collision, thermal fracture and chemical reaction. The blunt tool will inevitably affect the surface quality of the workpiece. Therefore, it is essential to develop a monitoring system based on sensory signal to predict the tool wear status so as to avoid undesirable consequences. Indirect monitoring method which correlates the sensory signal with the tool wear process is preferred because the information can be obtained more easily and economically in comparison with the direct method [1].

Currently, there are two main methods to realize the prediction of the tool wear. One commonly used is artificial neural networks based method. Ghosh et al. used the back propagation (BP) neural network to realize the estimation of the average flank wear of the main cutting edge by using cutting force, vibration and spindle current signal [2]. Using the features extracted from average milling force as input, Dong et al. realized the estimation of tool wear by the combination of the Bayesian inference and neural network [3]. Srinivasa et al. realized the estimation of the flank wear in face milling by using the fixed randomly hidden layer with batch fuzzy C-means algorithm [4]. They claimed that the proposed model can be used to realize more accurate tool wear estimation. Čuš et al. utilized a multilayer perceptron neural network to detect the tool wear [5]. It can be proved that BP neural network is capable of predicting tool wear with high accuracy [5]. Kuo developed a tool wear estimation system through the integration of radial basis function (RBF) neural network with fuzzy neural network [6]. Ozel et al. realized the flank wear prediction of the cutting tools under various conditions by utilizing back propagation neural network model [7]. The main advantage of the neural network method is that there is no need to build the analytic model to describe the complex internal mechanism of the tool wear because the information of the tool wear status can be memorized by the weight value of the neural network. However, this characteristic also results in that the prediction accuracy depends on the training samples greatly. If the new input data is not covered by the training samples, which is called “uncovered data”, the prediction accuracy will deteriorate greatly. Another approach which has been adopted by some researchers is the statistics based method. This method is realized by firstly supposing a statistical model and then calculating the model coefficients based on the training samples. Chen [8] and Zeng [9] adopted the multiple linear regression (MLR) model to predict the tool wear based on features extracted from sensory signals and the results showed that this method can get acceptable prognostic results. Bhattacharyya et al. proposed a new method by integrating simple time domain features from the cutting force

*Corresponding author. Tel.: +86 138 2016 2837, Fax.: +86 22 27406260
E-mail address: gfwangmail@tju.edu.cn
† Recommended by Associate Editor Jihong Hwang
© KSME & Springer 2014

signal and MLR model [10]. Kaya et al. also built a MLR model in which the cutting force and torque were used as the explanatory variables and maximum flank wear as response variable [11]. The experimental analysis showed that the MLR model can reflect the relationship between the tool wear status and sensory signal accurately.

In comparison with neural network based method, the MLR analysis above can get stronger generalization ability because it exerts an analytic model on the training data in advance. However, it is hard to get rid of the collinearity between these variables because the variation of the selected feature variables is commonly influenced by the tool wear. Therefore, if this model is used for the uncovered data, the variance of the response variables will increase and the tool wear prediction accuracy will deteriorate obviously [12, 13].

In this paper, a partial least square regression (PLSR) model is presented. The main characteristic of this method is that the regression coefficients are calculated within the principal component space so that the collinearity phenomenon between the explanatory variables can be eliminated. Therefore, the generalization ability can be enhanced and the prediction accuracy can be improved greatly even some data are missing in the case of real industrial application. In order to validate the effectiveness of the presented method, the side milling experiment of Titanium alloy was carried out and six features extracted from the cutting force were utilized as explanatory variables to realize tool wear prediction. To verify the effectiveness of the PLSR model for the uncovered data, the prediction is carried out under different number of the missing data. At the same time, BP, RBF neural network and MLR are adopted simultaneously to make comparison. The analysis and comparison show that the PLSR method can get higher prediction accuracy with the increase of uncovered data.

This paper is organized as follows: in section 2, the principle of the PLSR is introduced and a performance criterion is presented to analyze and compare the prediction accuracy. In Section 3, the experiment setup is described in details and six features are extracted from the force signals of different directions to depict the relationship between the sensory information and the tool wear value. In section 4, the comparison of the PLSR with BP, RBF and MLR under different number of the uncovered data is realized and the results show that the PLSR outperforms the other methods obviously with the increase of the uncovered data. Some useful conclusions are presented in section 5.

2. Principle of PLSR modeling and model validation

2.1 Principle of PLSR based prediction

PLSR is a multivariate calibration technique which is proposed to find the relationship between a set of explanatory variables X(m×N) (m is the dimension of the input data and N is the sample number) and a set of response variable Y(1×N). The main difference between PLSR and MLR is that the PLSR model is built in the principal component space so that it can give stable predictions even when X contains highly correlated variables. The PLSR model is obtained by a pre-specified maximum number of iterations [14] and the procedure can be described in the following four steps for the h_{th} iteration.

(1) The input matrix $E^h(m\times N)$ is decomposed into score vector t^h and a loading vector p^h

$$E^h = t^h p^h + E_0^h \tag{1}$$

where the superscript h represents the transpose and E_0^h (m×N) denotes the matrix of the residual data. The output matrix $F^h(1\times N)$ is decomposed in a similar manner

$$F^h = u^h (q^h)^T + F_0^h \,. \tag{2}$$

In Eq. (2), u^h and q^h are the score vector and the loading vector for the hth iteration, F_0^h (1×N) is the residual matrix.

Where t^h and u^h are (m×k) matrices of the k extracted score vectors; p^h and q^h are (N×k) matrices that represent matrices of loadings (k denotes the number of extracted score vectors). E^h and F^h are the standardized input and output matrices whose initial values are calculated from X and Y, respectively.

$$E^1 = \frac{X - \bar{X}}{S_X} \tag{3}$$

$$F^1 = \frac{Y - \bar{Y}}{S_Y} \,. \tag{4}$$

$\bar{X},\bar{Y}$ are the mean value of X and Y, respectively. S_X and S_Y correspond to their variance.

(2) A linear relationship is established between the E^h score vector t^h and F^h score vector u^h:

$$u^h = b^h t^h + d^h \tag{5}$$

where b^h is the (k×k) regression diagnose coefficients, d^h is the regression residual.

(3) A new model of F^h is built on T^h, Q^h, and B^h to obtain $(F^h)^*$

$$F^h = T^h (Q^h)^T B^h + (F^h)^* \tag{6}$$

where $T^h = [t^1,...,t^h]$, $Q^h = [q^1,...,q^h]$, and $B^h = [b^1,...,b^h]$ are the regression coefficients.

(4) E^h and F^h in Eqs. (1) and (2) are replaced by E_0^h (Eq. (1)) and $(F^h)^*$ (Eq. (6)) respectively and another iteration starts from the first step.

After all iterations are completed, the regression coefficient vector B in the principle component space is obtained and coefficients in the original data are calculated correspondingly by linear transformation.

$$\alpha = \sum_{i=1}^{h} b^i W^i \tag{7}$$

where, W^i is the eigenvector calculated from score matrices P^h and Q^h. The prediction equation based on input variable $(x_1,x_2,\ldots x_m)$ is given as.

$$y=[\bar{y}-\sum_{i=1}^{m}\alpha_i\frac{S_y}{S_{x_i}}\bar{x}_i]+\alpha_1\frac{S_y}{S_{x_1}}x_1+\ldots+\alpha_m\frac{S_y}{S_{x_m}}x_m\,. \tag{8}$$

2.2 Evaluation criteria of the prediction accuracy

The performance of the PLSR model can be evaluated using the root mean square error of prediction (RMSEP). The RMSEP represents the error associated with the model and can be computed by [15]

$$RMSEP=\sqrt{\frac{\sum_{i=1}^{N}(y_i^{'}-y_i)^2}{N}} \tag{9}$$

where, $y_i^{'}$ and y_i represent the model computed and measured values of the variable, and N represents the number of observations. As a measure of the goodness of fit, it can best describe an average measure of the error in predicting the response variable.

3. Experiment setup and feature extraction

To validate the necessity and effectiveness of the proposed method, a series of milling experiments were conducted in a Makino vertical machining center. The schematic diagram of the experiment setup is illustrated in Fig. 1. The workpiece was Titanium alloy Ti-6Al-4V which was clamped on a three-axis piezoelectric dynamometer (Kistler, type 9257A) and a Mitsubishi cutter with VP15TF coating was used for side-milling operation. The geometric parameters and type of the cutter is listed in Table 1. Cutting speed (v_c), axial depth of cut (a_p) and radial depth of cut (a_e) and feed per tooth (f_z) were kept constant at 40 m/min, 0.4 mm, 6 mm and 0.1 mm/tooth respectively. The length of each cutting pass was 150 mm. The main wear [16] appeared around the flank face (as shown in Fig. 2) and maximal length of this zone was measured after every cutting pass by an optical microscope. The cutting force signal was initially pre-amplified by a multi-channel charge amplifier (Kistler 5070), and then directly collected by a data acquisition card with the sampling frequency of 10 kHz. The experiment was performed until the value of the flank wear (VB) exceeded 0.3 mm. Because when the wear value exceeds 0.3 mm, the cutting tool usually has to be replaced [17]. The tool wear values under thirty-eight cutting passes were measured during the milling process and given in Fig. 2. To show the proceeding progress of the tool wear morphology, the pictures corresponding to some cutting passes (painted in blue dot) are also selected and demonstrated.

To demonstrate the force variation trend with the change of the tool wear status, the force signals along the direction of transverse feed x and longitudinal feed y under different tool wear values are illustrated in Figs. 3 and 4. The peak value P_k is also listed simutaneously. It can be seen that the peak value of the dynamic force signal in both directions increase with the growth of VB value. To depict the relationship between cutting force and tool wear value effectively and completely, root mean square (RMS) and standard deviation [5, 18] are also extracted from the cutting force in both x and y direction and combined with the peak value feature to construct the explanatory variables. The mathematical expressions of these indicators are listed in Table 2. For each tool wear state, the cutting force signal is first divided into two datasets. One is used for training and another is for test. Each dataset includes 20 segments with the length of 2000 and six time domain based features are extracted correspondingly. The mean values of these features are finally used as the explanatory variables to characterize the current wear status.

Table 1. Geometric parameters and type of the cutter.

Tool nose radius	Cutter diameter	Clearance angle	Inclination angle	Insert type	Tool holder type
0.8 mm	12 mm	11º	90º	APMT1135 PDER-H2 VP15TF	DEREK 300RC

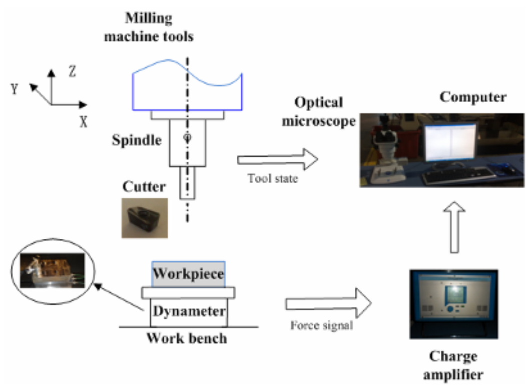

Fig. 1. Schematic diagram of the experiment setup.

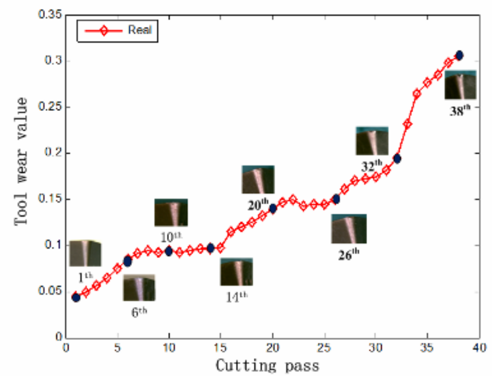

Fig. 2. Tool wear curve and morphology (v_c = 40 m/min, a_p = 0.4 mm, a_e = 6 mm and f_z = 0.1 mm/tooth).

Table 2. Mathematical expression of the extracted time domain indicators.

Feature	Function equation
RMS	$X_{RMS} = \sqrt{\frac{1}{m}\sum_{i=1}^{m} x_i^2}$
Standard deviation	$\sigma = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(x_i - \mu)}$
Peak value	$P_k = \max(\lvert x \rvert)$

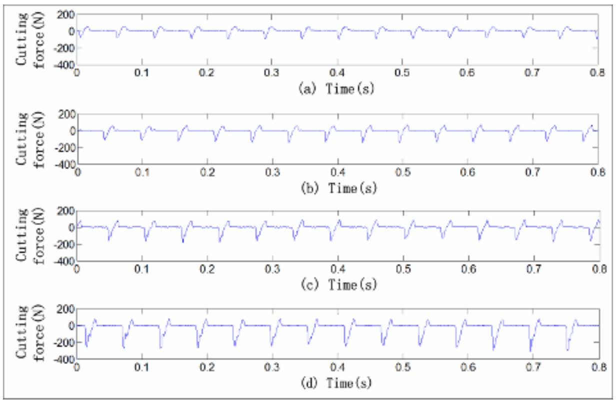

Fig. 3. Waveform of the cutting force under different tool wear value in x direction: (a) VB = 0 mm, P_k = 151 N; (b) VB = 0.09 mm, P_k = 217 N; (c) VB = 0.16 mm, P_k = 269 N; (d) VB = 0.31 mm, P_k = 317 N.

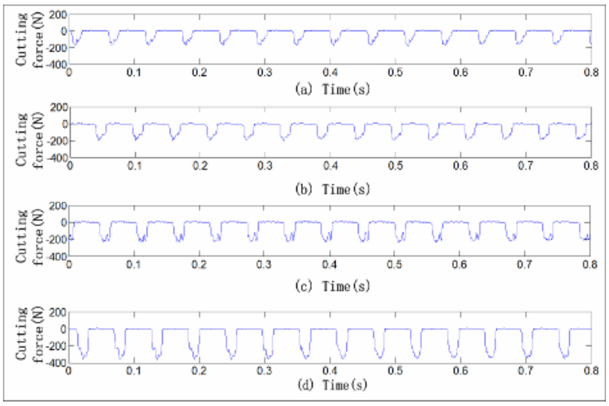

Fig. 4. Waveform of the cutting force under different tool wear value in y direction: (a) VB = 0 mm, P_k = 195 N; (b) VB = 0.09 mm, P_k = 215 N; (c) VB = 0.16 mm, P_k = 255 N; (d) VB = 0.31 mm, P_k = 376 N.

4. Tool wear prediction based on the PLSR

4.1 Simply description of BP, RBF and MLR model

Based on the features extracted above, PLSR is utilized for tool wear prediction in this section. At the same time, MLR, BP [7] and RBF [4] network are adopted simutaneously to compare the prediction accuracy with PLSR. MLR is a statistical model in which the relationship between the input features and the tool wear value is described by a linear equation. The estimation of regression coefficients can be obtained by least square fitting method. BP and RBF neural network are feedforward models which usually comprise three layers: input layer, hidden layer and output layer [4, 7]. The input layer takes feature vectors as the input and the output layer calculates the tool wear value accordingly. In this paper, the structure of the BP network is selected 6-10-1. The mapping function between the input layer and hidden layer is Sigmoid type and the linear transfer function is adopted in the output layer to predict the tool wear value. The Levenberg-Marquardt algorithm is used to realize the training of the network [19]. For RBF neural network, Gaussian function is used to realize nonlinear mapping in the hidden layer and the linear transfer function is used to map the output of the hidden layer into the tool wear value. To enhance the flexibility of the network, the hidden neuron of RBF network is added automatically during the training process. Therefore, the final number of the hidden layer depends on the training samples.

4.2 Comparison of PLSR with other methods

To show the influence of the uncovered data on the prediction accuracy, the data corresponding to some tool wear values are missed deliberately in the training dataset. In this paper, the number of the training data is selected as two to thirty-eight respectively. Based on each training data, the prediction model can be built using these four methods correspondingly. Then, thirty-eight test data are inputted into each model to predict the tool wear value. The prediction curves of these four methods under different number of uncovered data can be drawn out and compared with the measured value. Some of the results are demonstrated in Fig. 5. It can be seen that the missing of the training data results in severe deterioration of the prediction accuracy. Moreover, the error becomes larger with the increase of the missing data. Among these models, RBF gets the worst results and PLSR can achieve the best results. Because different position of the uncovered data usually results in different prediction results, it is hard to get a robust conclusion only by observing the prediction curve intuitively. Therefore, the RMSEP for each kind of position under different number of uncovered data are calculated and utilized as an index to reflect the prediction accuracy. The variation of RMSEP under different number of the uncovered data are illustrated in Fig. 6. It can be seen that the prediction error of RBF is the largest among these models for any number of the uncovered data. When the number of the uncovered data is small, the prediction accuracy of MLR, BP and PLSR is the same more or less. However, the prediction errors of BP and MLR tend to be larger than PLSR with the increase of the uncovered data, which testify that the PLSR model has stronger stability. By combining with Fig. 5, it can be concluded that PLSR outperforms the other methods if the uncovered data appears during the monitoring process.

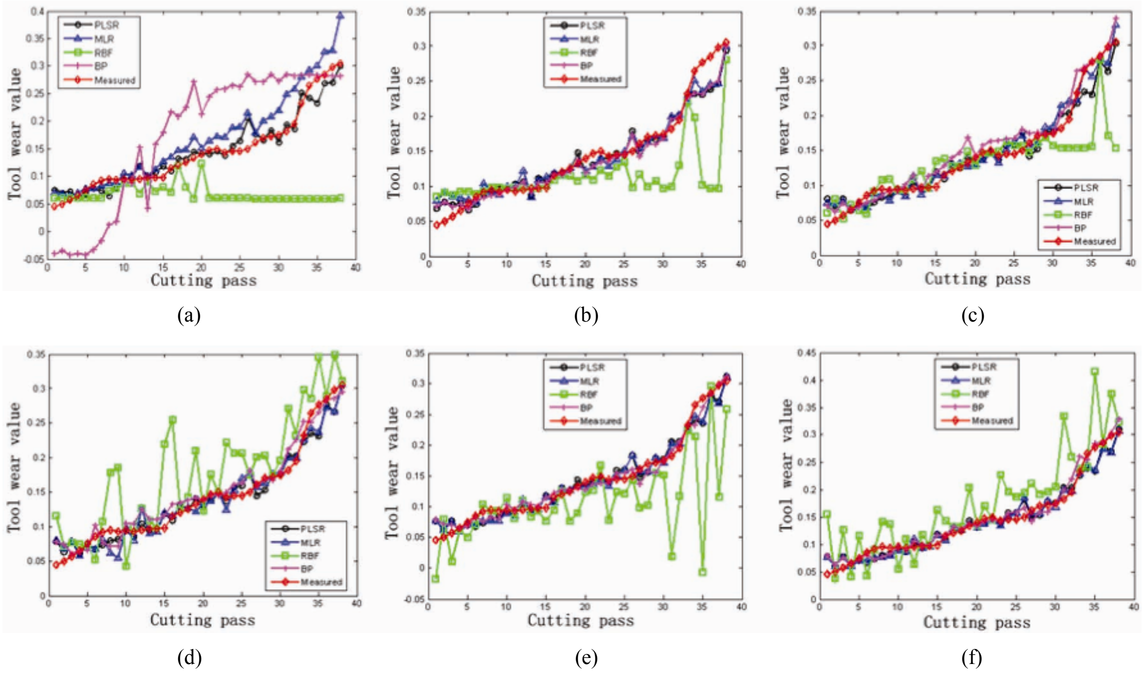

Fig. 5. Prediction curve of different models under different number of uncovered data: (a) 35 uncovered data; (b) 30 uncovered data; (c) 23 uncovered data; (d) 16 uncovered data; (e) 8 uncovered data; (f) 0 uncovered data.

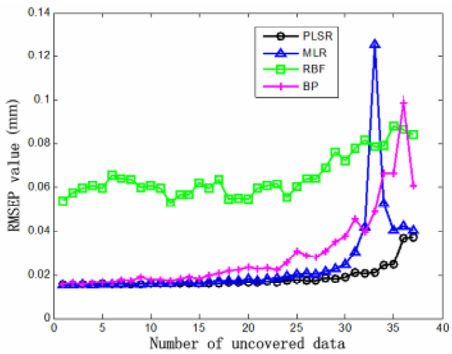

Fig. 6. Comparison of prediction error under different number of uncovered data.

5. Conclusions

In this paper, the PLSR model is presented to realize the accurate tool wear prediction. The main characteristic is that the multicollinearity between the explanatory variables can be avoided because the regression coefficients are not calculated in the original variable space but the principal component space. Therefore, the prediction accuracy for the uncovered data can be improved greatly. The analysis of the tool wear prediction for the milling process shows that, with the increase of the number of the uncovered data, the PLSR model can achieve more accurate results in comparison with the neural network and MLR based models. This method casts a new light on the prediction of the tool wear in real industrial environment. However, it should be noted that the current model can only be used to predict the tool wear value for the predefined cutting parameters. If the cutting parameters are changed, the prediction model should be rebuilt according to the new collected data correspondingly. In the future work, the factorial design of the experiments should be carried out to reflect the tool wear variation under different levels of the cutting parameters.

Acknowledgment

This project is supported by National Natural Science Foundation of China (51175371).

Nomenclature

$X\,(m\times N)$	: Input variables, m is the dimension of the input data and N is the sample number
$Y(1\times N)$	: Response variable
h	: The h_{th} iteration
E^h (m×N)	: Input matrix
$F^h\,(1\times N)$	: Output matrix
t^h, u^h	: Score vector
p^h, q^h	: Loading vector
E_0^h, F_0^h	: Residual matrix
$\overline{X}, \overline{Y}$	: Mean value
S_X, S_Y	: Variance
b^h	: Regression diagnose coffecient
d^h	: Regression residual
T^h, Q^h, B^h	: Regression coefficients
P^h, Q^h	: Score matrices

W^i : Eigenvector

α : Regression coefficients

References

[1] J. V. Abellan-Nebot and F. R. Subirón, A review of machining monitoring systems based on artificial intelligence process models, *Int. J. Adv. Manuf. Tech.*, 47 (2010) 237-257.

[2] N. Ghosha, Y. B. Ravib, A. Patrac, S. Mukhopadhyayc, S. Pauld, A. R. Mohantyd and A. B. Chattopadhyayd, Estimation of tool wear during CNC milling using neural network-based sensor fusion, *Mech. Syst. Signal Pr.*, 21 (2007) 466-479.

[3] J. F. Dong, K. V. R. Subrahmanyam, Y. S. Wong, G. S. Hong and A. R. Mohanty,. Bayesian-inference-based neural networks for tool wear estimation, *Int. J. Adv. Manuf. Tech.*, 30 (2006) 797-807.

[4] P. Srinivasa Pai, T. N. Nagabhushana and P. K. Ramakrishna Rao,. Flank wear estimation in face milling based on radial basis function neural networks, *Int. J. Adv. Manuf. Tech.*, 20 (2002) 241-247.

[5] F. Čuš, U. Župerl, Real-time cutting tool condition monitoring in milling, *Stroj. Vestn. -J. Mech. E.*, 57 (2) (2011) 142-150.

[6] R. J. Kuo and P. H. Cohen, Multi-sensor integration for on-line tool wear estimation through radial basis function networks and fuzzy neural network, *Neural Networks*, 12 (1999) 355-370.

[7] T. Ozel and A. Nadgir, Prediction of flank wear by using back propagation neural network modeling when cutting hardened H-13 steel with chamfered and honed CBN tools, *Int. J. Mach. Tool Manu.*, 42 (2002) 287-297.

[8] J. C. Chen, A multiple-regression model for monitoring tool wear with a dynamometer in milling operations, *The Journal of Technology Studies*, 30 (4) (2005) 71-77.

[9] X. Li, H. Zeng, J. H. Zhou, S. Huang, T. B. Thoe, K. C. Shaw and B. S. Lim, Multi-modal sensing and correlation modelling for condition-based monitoring in milling machine, *SIMTech technical reports*, 8 (1) (2007) 50-56.

[10] P. Bhattacharyyaa, D. Senguptaa and S. Mukhopadhyay, Cutting force-based real-time estimation of tool wear in face milling using a combination of signal processing techniques, *Mech. Syst. Signal Pr.*, 21 (2007) 2665-2683.

[11] B. Kaya, C. Oysu and M. H. Ertunc, Force-torque based on-line tool wear estimation system for CNC milling of Inconel 718 using neural networks, *Adv. Eng. Softw.*, 42 (2011) 76-84.

[12] S. Wold, M. Sjostrom and L. Eriksson, PLS-regression: a basic tool of chemometrics, *Chemometr. Intell. Lab.*, 58 (2001) 109-130.

[13] H. A. Farahani, A. Rahiminezhad, L. Same and K. immannezhad, A comparison of partial least squares (PLS) and ordinary least squares (OLS) regressions in predicting of couples mental health based on their communicational patterns, *Procedia Social and Behavioral Sciences*, 5 (2010) 1459-1463.

[14] H. Yang, R. P. Griffiths and J. D. Tate, Comparison of partial least squares regression and multi-layer neural networks for quantification of nonlinear systems and application to gas phase Fourier transform infrared spectra, *Anal. Chim. Acta*, 489 (2003) 125-136.

[15] J. M. Poveda, A. Garcia, P. J. Martin-Alvarez and L. Cabezas, Application of partial least squares (PLS) regression to predict the ripening time of Manchego cheese, *Food Chem.*, 84 (2004) 29-33.

[16] S. Khamel, N. Ouelaa and K. Bouacha, Analysis and prediction of tool wear, surface roughness and cutting forces in hard turning with CBN tool, *J. Mech. Sci. Technol.*, 26 (11) (2012) 3605-3616.

[17] K. Venkatesh, M. Zhou and R. J. Caudill, Design of artificial neural networks for tool wear monitoring, *J. Intell. Manuf.*, 8 (1997) 125-226.

[18] R. Teti, K. Jemielniak, G. O'Donnell and D. Dornfeld, Advanced monitoring of machining operations, *CIRP Annals - Maunf. Techn.*, 59 (2010) 717-739.

[19] G. Wang and Y. Cui, On line tool wear monitoring based on auto associative neural network. *J. Intell. Manuf.* DOI: 10.1007/s10845-012-0636-7.

Guofeng Wang is currently an associate professor in School of Mechanical Engineering, Tianjin University, China. He received his Ph.D. degree from Tianjin University, China, in March 2002. His research interests include dynamic modeling and condition monitoring of machining process.