Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

Soft Sensor for Multiphase and Multimode Processes Based on Gaussian Mixture Regression

Xiaofeng Yuan† , Zhiqiang Ge† *, Hongwei Zhang† , Zhihuan Song† *, Peiliang Wang‡

† State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China
(e-mail: yxf80@zju.edu.cn;gezhiqiang@zju.edu.cn;zhanghongwei@zju.edu.cn; zhsong@iipc.zju.edu.cn).
‡ School of Information and Engineering, Huzhou Teachers College, Huzhou, Zhejiang, People's Republic of China
(e-mail: wpl@hutc.zj.cn)
**To whom all correspondence should be addressed*

Abstract: For complex industrial plants with multiphase and multimode characteristic, traditional multivariate statistical soft sensor methods are not applicable as Gaussian distribution assumption of data is not met. Thus Gaussian mixture model (GMM) is used to approximate data distribution. In the previous GMM-based soft sensor modeling researches, GMM is only used to identify operating mode, then other regression algorithms like PLS are used for quality prediction in different localized modes. In this article, an existing method—Gaussian mixture regression (GMR) is introduced for soft sensor modeling, which has been already successfully applied in robot programming by demonstration (PbD). Different from past GMM-based soft sensors, GMR is directly used for regression. In GMR, data mode identification and regression are incorporated into one model, thus there is no need to switch prediction model when data mode has changed. Feasibility and efficiency of GMR based soft sensor are validated in the fermentation process and TE process.

Keywords: Soft sensor, Gaussian mixture regression (GMR), multiphase, multimode, quality prediction

1. INTRODUCTION

To improve process efficiency and product quality, key product variables are essential for process operations like process control and optimization in many industrial plants. In some cases, the processes often encounter the great challenge of lacking accurate real-time measurements of these key process variables. Over the past decades, soft sensors have been widely used to predict key process variables through those that are easy to measure online (Ge, Huang et al. 2013; Kaneko and Funatsu 2013; Kim, Okajima et al. 2013). The most commonly used soft sensors are the data-driven modeling methods, which construct inferential models using abundant process measurement data. As a category of data-driven soft sensors, the multivariate statistical techniques such as principal component analysis (PCA)(Jolliffe 2005), partial least squares (PLS)(Kim, Okajima et al. 2013), and independent component analysis (ICA)(Kaneko, Arakawa et al. 2009) are extensively researched and applied to diverse processes. This class of methods often extract the variable features by projecting the original measurement data onto a linear subspace. In addition, machine learning based methods like support vector machine(SVM)(Ge and Song 2010) have also been successfully applied to soft sensor modeling.

Though different types of soft sensor modeling techniques have been applied for quality prediction, most of them are based on the assumption the process data come from a single operating region and follow a unimodal Gaussian distribution. For complex multiphase and multimode processes that are running at multiple operating conditions, the basic assumption of multivariate Gaussian distribution is not met because of the mean shifts or covariance changes. Consequently, data distribution may be complicated with arbitrary non-Gaussian patterns. As mixture models can represent arbitrarily complex probability density functions, they are ideal tools for modeling complex multi-class data set.

The most commonly used mixture model is the Gaussian mixture model (GMM). By taking sufficient linear combinations of basis single multivariate Gaussian distributions, GMM can smoothly approximate almost any continuous density to arbitrary accuracy. Thus the state-of-the-art GMM technique has shown strong ability in dealing with non-Gaussian data and been successfully applied in fields like speech recognition(Lu, Ghoshal et al. 2013), image segmentation(Greggio, Bernardino et al. 2012), and robotic learning(Núñez, Rocha et al. 2013) in past researches. With respect to the process industrial, GMM is widely utilized for multiphase and multimode process monitoring and soft sensor. For example, a novel multimode process monitoring approach based on finite Gaussian mixture models and Bayesian inference strategy is proposed in(Yu and Qin 2008). Another application example in the soft sensor area is the integration of multiway Gaussian mixture model and adaptive kernel partial least squares for online quality prediction in(Yu 2012).

Despite a plenty of applications for GMM, most of them only use GMM for classification problems, not for the regression ones. Even for soft sensor regression in (Yu 2012), GMM is exploited in order to identify and isolate different operating phases, then the KPLS models are built for regression in different identified phases. However, the GMM can also be extended to solve regression problems, which is denoted as Gaussian mixture regression (GMR). A relation model is directly constructed between the input and output variables in GMR. The detailed implementation of GMR algorithm is described in(Sylvain 2009). As mentioned in the paper, GMR is easy to implement, and satisfies robustness and smoothness criterions that are common to a variety of fields. Although the theoretical considerations of GMR were presented two decades ago, it has come out with only few applications. This is surprising since GMM algorithm has been put into practice in various fields. By far, GMR has mainly been utilized in area of robot programming by demonstration (PbD) for imitation learning of multiple tasks(Calinon, Guenter et al. 2007; Cederborg, Li et al. 2010). To our best knowledge, this method has not been applied in other fields yet. Especially, no literature has been found for soft sensor using GMR model in chemical process up to now. Therefore, the use of GMR in soft sensor framework has yet to be explored. In this study, GMR based soft sensor model will be established for quality prediction of key variables in chemical process.

The remaining of this paper is organized as follows. In Section 2, the definition of GMM and the EM algorithm for parameter estimation are briefly revisited. Then the regression derivation the GMM (GMR) is introduced for soft sensor in Section 3. In Section 4, the algorithm is validated on two application examples: the fermentation process for penicillin production and the Tennessee Eastman (TE) process. Finally, the conclusions and future directions are presented in Section 5.

2. GAUSSIAN MIXTURE MODEL

2.1 Introduction to GMM

Assuming $x \in R^d$ to be a d-dimensional data point from a single d-dimensional multivariate Gaussian model with parameters $\theta = \{\mu, \Sigma\}$ (μ is the mean vector and Σ is the covariance matrix), then the probability density function of the single Gaussian component is given by

$$f(x \mid \theta) = \frac{1}{\sqrt{(2\pi)^d \left|\Sigma\right|}} \exp\left\{-1/2(x-\mu)^T \Sigma^{-1}(x-\mu)\right\} \tag{1}$$

If x is from a mixture Gaussian model, then its probability density function can be formulated as following(Yu 2012)

$$p(x \mid \theta^T) = \sum_{k=1}^{K} \omega_k f(x \mid \theta_k) \tag{2}$$

where K means the number of Gaussian components included in GMM, ω_k is the probabilistic weight of the kth Gaussian component and subjects to the equation $\sum_{k=1}^{K} \omega_k = 1$, $\theta_k = \{\mu_k, \Sigma_k\}$ represent the parameters of the kth Gaussian model. $\theta^T = \{\theta_1, \theta_2, \cdots \theta_K\} = \{\mu_1, \Sigma_1, \mu_2, \Sigma_2, \cdots \mu_K, \Sigma_K\}$ denote the total parameters of all the Gaussian components, respectively.

The derivation of the probability density function of GMM can also be explained like the following:

Define the hidden state $z_k \in \{0,1\}$, where $z_k = 1$ denotes that x is in the kth Gaussian component and vice versa. Hence we have

$$\begin{aligned} &\sum\nolimits_k z_k = 1 \\ &p(z_k = 1) = \omega_k \end{aligned} \tag{3}$$

Let $Z = [z_1, z_2, \ldots z_K]$, the joint probability distribution of Z and conditional probability distribution of x on Z can formulated as

$$\begin{aligned} p(Z) &= \prod_{k=1}^{K} \omega_k^{z_k} \\ p(x \mid Z) &= \prod_{k=1}^{K} f(x \mid \theta_k)^{z_k} \end{aligned} \tag{4}$$

The total probability distribution of x is

$$p(x \mid \Theta) = p(Z)p(x \mid Z) = \prod_{k=1}^{K} \omega_k^{z_k} f(x \mid \theta_k)^{z_k} \tag{5}$$

where $\Theta = \{\{\omega_1, \mu_1, \Sigma_1\}, \{\omega_2, \mu_2, \Sigma_2\}, \cdots \{\omega_K, \mu_K, \Sigma_K\}\}$ .It is easy to see that formula (5) is equivalent to formula (2). Meanwhile, formula (5) is more conducive to the derivation of EM algorithm for parameter estimation.

2.2 EM algorithm for GMM

To use a GMM, the unknown parameter set Θ of probabilistic weights and model parameters of each Gaussian component should be estimated first. Common methods for this problem include the maximum likelihood estimation (MLE) and expectation maximization (EM). Given a set of training samples $X = \{x_1, x_2, \cdots, x_N\}$ and their corresponding hidden states $Z^T = \{Z^1, Z^2, \cdots, Z^N\}$ with its nth entry equal to $Z^n = \{z_1^n, z_2^n, \cdots, z_K^n\}$, the likelihood function is

$$L(X, \Theta) = \prod_{n=1}^{N} \prod_{k=1}^{K} \omega_k^{z_k^n} f(x_n \mid \theta_k)^{z_k^n} \tag{6}$$

The corresponding log-likelihood function is

$$\log L(X, \Theta) = \sum_{n=1}^{N} \sum_{k=1}^{K} z_k^n \{\log(f(x_n \mid \theta_k)) + \log \omega_k\} \tag{7}$$

The parameters are estimated by maximizing the log-likelihood function. In practice, the EM algorithm is used to estimate parameters of the maximum likelihood function. By giving an initial parameter set $\Theta^{(1)}$ and successively applying the E step and M step, the EM algorithm can produce a sequence of parameters $\{\Theta^{(1)}, \Theta^{(2)}, \cdots, \Theta^{(m)}, \cdots\}$. The details of these two steps are performed as follows(Yu 2012):

E-step:

$$\gamma^{(m)}(z_k^j \mid x_j, \Theta^{(m)}) = \frac{\omega_k^{(m)} f(x_j \mid \mu_k^{(m)}, \Sigma_k^{(m)})}{\sum_{i=1}^{K} \omega_i^{(m)} f(x_j \mid \mu_i^{(m)}, \Sigma_i^{(m)})} \quad (8)$$

where $\gamma^{(m)}(z_k^j \mid x_j, \Theta^{(m)})$ denotes the posterior probability of the jth training sample with the kth Gaussian component at the mth iteration.

M-step:

$$\begin{aligned}
\omega_k^{(m+1)} &= \frac{\sum_{j=1}^{N} \gamma^{(m)}(z_k^j \mid x_j, \Theta^{(m)})}{N} \\
\mu_k^{(m+1)} &= \frac{\sum_{j=1}^{N} \gamma^{(m)}(z_k^j \mid x_j, \Theta^{(m)}) x_j}{\sum_{j=1}^{N} \gamma^{(m)}(z_k^j \mid x_j, \Theta^{(m)})} \\
\Sigma_k^{(m+1)} &= \frac{\sum_{j=1}^{N} \gamma^{(m)}(z_k^j \mid x_j, \Theta^{(m)})(x_j - \mu_k^{(m+1)})(x_j - \mu_k^{(m+1)})^{\mathrm{T}}}{\sum_{j=1}^{N} \gamma^{(m)}(z_k^j \mid x_j, \Theta^{(m)})}
\end{aligned} \quad (9)$$

where $\omega_k^{(m+1)}, \mu_k^{(m+1)}, \Sigma_k^{(m+1)}$ are the probabilistic weight, mean and covariance of the kth Gaussian component at the (m+1)th iteration, respectively.

3. GAUSSIAN MIXTURE REGRESSION AND GMR-BASED SOFT SENSOR

3.1 Gaussian mixture regression

This sub-section introduces the derivation of GMR. Assume the data sample x in GMM consists of two parts: the input x^I and output x^O. Subsequently, the mean and covariance can be divided into the input and output parts like the following

$$\mu_k = \begin{bmatrix} \mu_k^I \\ \mu_k^O \end{bmatrix}, \Sigma_k = \begin{bmatrix} \Sigma_k^I & \Sigma_k^{IO} \\ \Sigma_k^{OI} & \Sigma_k^O \end{bmatrix} \quad (10)$$

For a given input variable x^I and its corresponding Gaussian component k, x^O obeys the Gaussian distribution. The expected distribution of x^O can be defined by

$$p(x^O \mid x^I, k) \sim f(x^O \mid \hat{x}_k, \hat{\Sigma}_k) \quad (11)$$

where the mean $\hat{x}_k$ and covariance $\hat{\Sigma}_k$ are calculated by

$$\begin{aligned}
\hat{x}_k &= \mu_k^O + \Sigma_k^{OI}(\Sigma_k^I)^{-1}(x^I - \mu_k^I) \\
\hat{\Sigma}_k &= \Sigma_k^O - \Sigma_k^{OI}(\Sigma_k^I)^{-1}\Sigma_k^{IO}
\end{aligned} \quad (12)$$

In fact, we only know the input x^I while the component label k is unknown. Considering the complete GMM, the expected distribution of output x^O consists of K parts as

$$p(x^O \mid x^I) \sim \sum_{k=1}^{K} h_k f(x^O \mid \hat{x}_k, \hat{\Sigma}_k) \quad (13)$$

where h_k corresponds to the probability that k is responsible for x^I. By the Bayes formula, we can get h_k as

$$h_k = p(k \mid x^I) = \frac{p(k)p(x^I \mid k)}{\sum_{i=1}^{K} p(i)p(x^I \mid i)} = \frac{\omega_k f(x^I \mid \mu_k^I, \Sigma_k^I)}{\sum_{i=1}^{K} \omega_i f(x^I \mid \mu_i^I, \Sigma_i^I)} \quad (14)$$

Finally, given x^I, the expectation and covariance of output can be approximated as

$$\begin{aligned}
\hat{x}^O &= \sum_{k=1}^{K} h_k \hat{x}_k \\
\hat{\Sigma} &= \sum_{k=1}^{K} h_k^2 \hat{\Sigma}_k
\end{aligned} \quad (15)$$

The detailed derivation of GMR can be found in (Sylvain 2009).

3.2 Soft sensor development based on GMR

The purpose of soft sensor is to build relation models between key process variables and easy-to-measure variables in industrial process. Obviously, GMR can be used for soft sensor modeling in multiphase and multimode processes. We assume the process input variables are x, and the output variables are y. There are K different operation modes in the process. The procedure of GMR for soft sensor can be explained as follows:

Training step: firstly, we combine input variables x and its corresponding output variables y into a new variable vector $v = [x^T \; y^T]^{\mathrm{T}}$. Assume the probability density distribution of v follows the GMM distribution

$$p(v) = \sum_{k=1}^{K} p(k)p(v \mid k) \quad (16)$$

Then by updating and recalculating the E-step and M-step of EM algorithm, the optimal probabilistic weights and Gaussian model parameters (described in (10)) can be obtained using the training data set. Thus a corresponding soft sensor can be constructed.

Testing step: when a new data sample x_{new} comes in, its posterior probability and the conditional probability density of y on x_{new} in the kth mode can be obtained by

$$p(k \mid x_{new}) = \frac{p(x_{new} \mid \theta_k)p(k)}{p(x_{new} \mid \theta_k)} \quad (17)$$

$$p(y_{new,k} \mid x_{new}, k) \sim f(y_{new,k} \mid \hat{y}_{new,k}, \hat{\Sigma}_{new,k}) \quad (18)$$

where $\hat{y}_{new,k}$ and $\hat{\Sigma}_{new,k}$ can be calculated using (12). Therefore, the combined prediction output is given as

$$\hat{y}_{new} = \sum_{k=1}^{K} p(k \mid x_{new}) \hat{y}_{new,k} \quad (19)$$

4. CASE STUDIES

In this part, two case studies on a simulated fermentation process and TE process will be carried out to verify the validity and effectiveness of GMR.

4.1 Fermentation process for penicillin production

The fermentation process for penicillin production is a biochemical batch benchmark extensively used for soft sensor and fault diagnosis algorithm. An online simulation tool of this process can be found at the website of http://simulator.iit.edu/web/pensim/simul.html.The simulator contains a fermenter where the biological reaction takes place. Fig. 1 shows a flow sheet of the process, and a detailed description of the process can be found in reference (Birol, Ündey et al. 2002). For different demands, the simulator provides several settings including the controller, process duration, sampling rates, etc. The settings in this study are given as default in the webpage except that the sampling interval is 1 hour and the simulation time is 1000 hours.

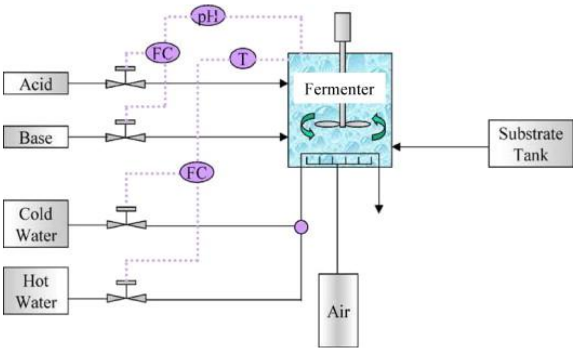

Fig. 1. The flow sheet of the fermentation process

Table 1. Input variables in the fermentation process

No.	Variable description	Units
1	Aeration rate	L/h
2	Agitator power	W
3	Substrate feed rate	L/h
4	Substrate feed temperature	K
5	Substrate concentration	g/L
6	Dissolved oxygen concentration	g/L
7	Biomass concentration	g/L
8	Culture volume	L
9	Carbon dioxide concentration	g/L
10	pH	—
11	Fermentor temperature	K

There are totally 16 measured variables in the simulation plant. For soft sensor model construction, the penicillin concentration is chosen as the output variable. 11 variables, which are highly related to it, are selected as the input variables. The description of each variable is listed in Table 1. In this study, a total of 1000 samples are generated. 500 of them are used for training set and the remaining for the testing set.

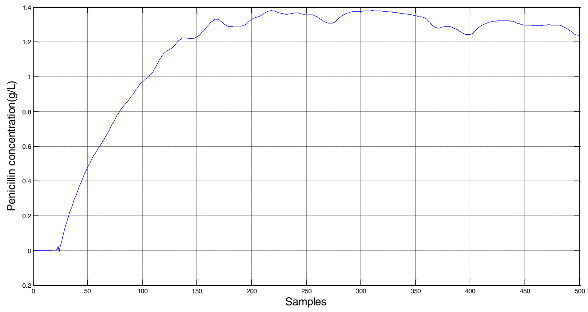

Fig. 2. Data characteristic of output variable

In order to examine the multiphase behavior of the process data, the values of output variable are shown in Fig. 2. As can be seen from the figure, the entire penicillin fermentation process consists of three distinct phases: the lag phase, exponential phase and stationary phase. As a single global model cannot meet the demand of prediction accuracy, multi-model method of GMR is necessary. For performance comparison, the GMM-based PLS (GMM-PLS) model has also been used for quality prediction. In GMM-PLS algorithm, the input and output data are firstly divided into different specific blocks using GMM, then different localized PLS models are constructed in these blocks, respectively. For new testing sample, it is categorized into an individual phase through posterior probability calculation based on Bayesian inference. After that, output of quality variable is estimated using the corresponding localized KPLS model. For performance comparison, the root-mean-square error (RMSE) criterion is adopted in the following.

Table 2. RMSE values of different models

Soft sensors	GMR	GMM-PLS
RMSE values	0.0203	0.0257

To be fair, numbers of Gaussian components in GMM-PLS and GMR are both selected as 3. The component number of PLS is chosen as 6, which can explain most of the process data information. The RMSE values of GMR and GMM-PLS based quality predictions are shown in Table 2. It can be seen that GMR is superior to GMM-PLS as the RMSE value of GMR is smaller than that of GMM-PLS. Furthermore, the detailed trend plots of quality variable prediction results in the testing data are displayed in Fig. 3. The red and blue lines represent the trends of real and predicted values of penicillin concentration, respectively. Fig. 3(a) shows the results of GMR based soft sensor model while the results of GMM-PLS based models are exhibited in Fig. 3(b). Though the predictive superiority is not obvious here, GMR based soft sensor model has more advantages than the GMM-PLS model. Firstly, the GMR simultaneously integrates classification procedure and regression step into a single model, while GMM-PLS takes two steps for soft sensor modeling. Secondly, GMR does not need to switch the prediction model if the running phase has been changed, which means its automation level is higher than that of GMM-PLS.

a

b

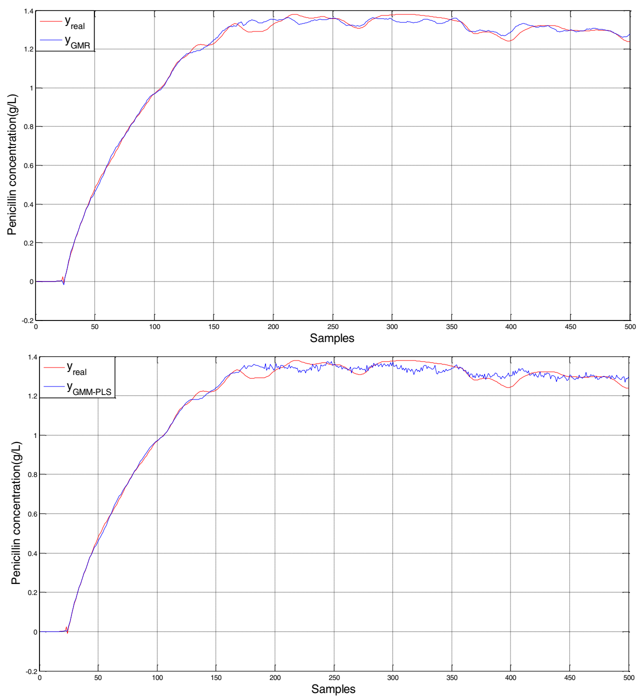

Fig. 3 Prediction results of the testing data on the fermentation process, (a) GMR, (b) GMM-PLS

4.2 TE benchmark process

TE process is a benchmark simulation process, which has been extensively used for algorithm testing and performance evaluation(Downs and Vogel 1993). It consists of five operating units: a reactor, a condenser, a compressor, a separator and a stripper. Fig. 4 presents the flowchart of TE process.

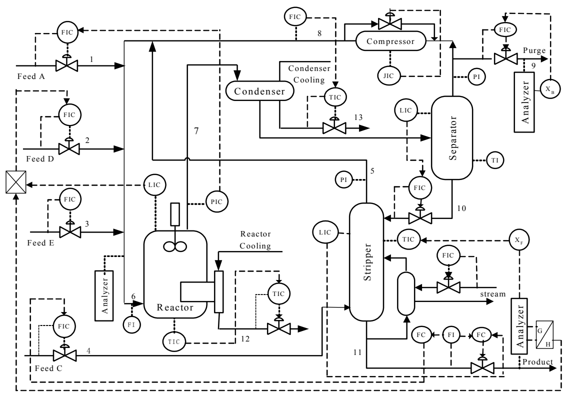

Fig. 4 The flowchart of TE process

There are 41 measured variables and 12 manipulated variables in this process. Among the 41 measured variables, there are 19 component variables that are difficult to be measured online. For soft sensor purpose, C component in the product stream is selected as the output variable. Meanwhile, 16 easy to be measured variables are chosen as inputs, which are tabulated in Table 3.

For soft sensor of GMR, six different operation modes are simulated. In each operation mode, a total of 2000 data samples are collected and equally partitioned into two parts: training data and testing data. Then the training data and testing data of different operation modes are incorporated, respectively.

Table 3 Input variables in TE process

No.	Input variables	No.	Input variables
1	A feed	9	Separator temperature
2	D feed	10	Separator pressure
3	E feed	11	Separator underflow
4	A and C feed	12	Stripper pressure
5	Recycle flow	13	Stripper temperature
6	Reactor feed rate	14	Stripper steam flow
7	Reactor temperature	15	Reactor cooling water outlet temperature
8	Purge rate	16	Separator cooling water outlet temperature

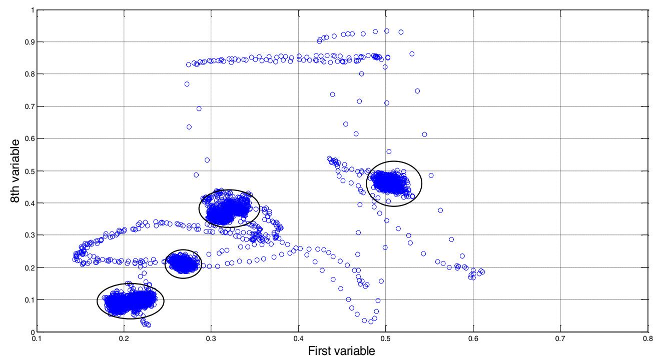

Fig. 5 Data characteristic of 2 input variables

To examine the multimode behavior of the data, the scatter plot of two representative input variables (the first and 8^{th} input variables) is shown in Fig. 5. It can be seen that the data do not obey a single Gaussian distribution. Several typical different modes are highlighted in ellipses in the figure. In addition, different modes overlap with each other. Therefore, mixture models are necessary to approximate the data distribution.

Again, GMR and GMM-PLS based soft sensor models are constructed for quality variable prediction in the TE process. Here, K is set to six. Number of latent variables is determined as 8. Performance comparisons of the two methods are tabulated in Table 4. Similarly, the GMR based soft sensor model performs better than the GMM-PLS method as the RMSE value of the former is smaller than that of the latter. In detail, the prediction results of the two different soft sensor models are given in Fig. 6(a) and Fig. 6(b), respectively. From the figure, we can see that the predictive results of GMR based soft sensor are better than that of GMM-PLS based method. The predicted curve matches better with the real curve of C component. Especially, at the switch points between different operating modes, the actual values of C component are better tracked by GMR based method than that of GMM-PLS. Comparison results show that GMR is very effective for soft sensor modeling.

Table 4 RMSE values of different models

Soft sensors	GMR	GMM-PLS
RMSE values	0.3470	0.5322

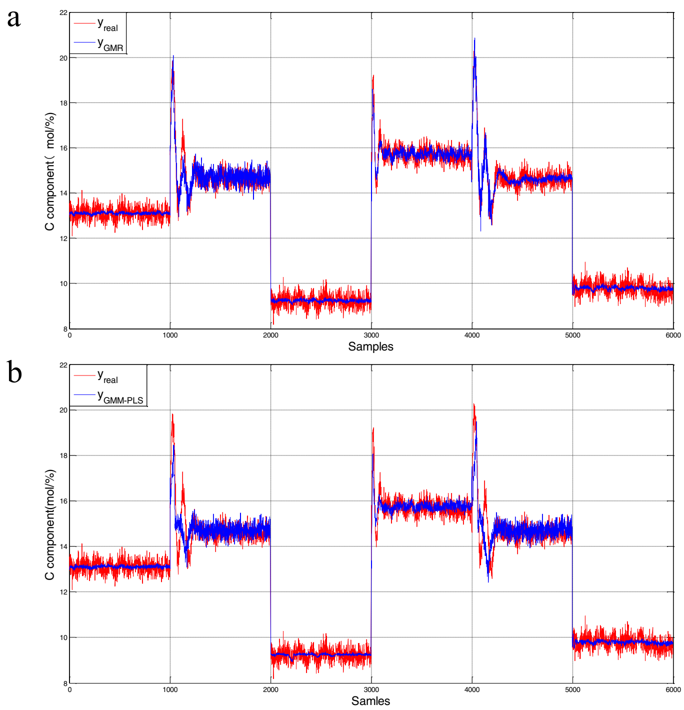

Fig. 6 Prediction results of the testing data on TE process, (a) GMR, (b) GMM-PLS

5. CONCLUSION

In the present paper, a new method of GMR based soft sensor model is introduced for quality prediction of multiphase and multimode processes. Different from other soft sensors for these processes, this method can simultaneously address cluster identification and regression in one model. Beside, the new soft sensor method can provide a predictive distribution for the output variables, which can give the uncertainty information of soft sensor. Through the fermentation process and TE process, the feasibility and efficiency of GMR based soft sensor are demonstrated. This is yet a fringe attempt of GMR application in soft sensor, further researches on this method will be carried out in detail.

ACKNOWLEDGEMENT

This work was supported in part by the National Natural Science Foundation of China (NSFC) (61370029), the National Project 973 (2012CB720500).

REFERENCES

Birol, G., C. Ündey, et al. (2002). "A modular simulation package for fed-batch fermentation: penicillin production." Computers & Chemical Engineering 26(11): 1553-1565.

Calinon, S., F. Guenter, et al. (2007). "On learning, representing, and generalizing a task in a humanoid robot." Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on 37(2): 286-298.

Cederborg, T., M. Li, et al. (2010). Incremental local online Gaussian mixture regression for imitation learning of multiple tasks. Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on, IEEE.

Downs, J. J. and E. F. Vogel (1993). "A plant-wide industrial process control problem." Computers & Chemical Engineering 17(3): 245-255.

Ge, Z., B. Huang, et al. (2013). "Mixture semi-supervised principal component regression model and soft sensor application." Aiche Journal: n/a-n/a.

Ge, Z. and Z. Song (2010). "A comparative study of just-in-time-learning based methods for online soft sensor modeling." Chemometrics and Intelligent Laboratory Systems 104(2): 306-317.

Greggio, N., A. Bernardino, et al. (2012). "Fast estimation of Gaussian mixture models for image segmentation." Machine Vision and Applications 23(4): 773-789.

Jolliffe, I. (2005). Principal component analysis, Wiley Online Library.

Kaneko, H., M. Arakawa, et al. (2009). "Development of a new soft sensor method using independent component analysis and partial least squares." Aiche Journal 55(1): 87-98.

Kaneko, H. and K. Funatsu (2013). "Adaptive Soft Sensor Model Using Online Support Vector Regression with Time Variable and Discussion of Appropriate Parameter Settings." Procedia Computer Science 22: 580-589.

Kim, S., R. Okajima, et al. (2013). "Development of soft-sensor using locally weighted PLS with adaptive similarity measure." Chemometrics and Intelligent Laboratory Systems 124: 43-49.

Lu, L., A. Ghoshal, et al. (2013). "Cross-lingual Subspace Gaussian Mixture Models for Low-resource Speech Recognition."

Núñez, P., R. P. Rocha, et al. (2013). "Novelty detection and segmentation based on Gaussian mixture models: A case study in 3D robotic laser mapping." Robotics and Autonomous Systems 61(12): 1696-1709.

Sylvain, C. (2009). Robot programming by demonstration: A probabilistic approach, EPFL Press.

Yu, J. (2012). "Multiway Gaussian Mixture Model Based Adaptive Kernel Partial Least Squares Regression Method for Soft Sensor Estimation and Reliable Quality Prediction of Nonlinear Multiphase Batch Processes." Industrial & Engineering Chemistry Research 51(40): 13227-13237.

Yu, J. and S. J. Qin (2008). "Multimode process monitoring with Bayesian inference - based finite Gaussian mixture models." Aiche Journal 54(7): 1811-1829.