

Original Research Paper

# Multi-Phase Support Vector Regression Soft Sensor for Online Product Quality Prediction in Glutamate Fermentation Process

<sup>1,2</sup>Rongjian Zheng and <sup>1</sup>Feng Pan

<sup>1</sup>Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi, China

<sup>2</sup>Faculty of Automation, Huaiyin Institute of Technology, Huaian, China

## Article history

Received: 28-02-2017

Revised: 12-05-2017

Accepted: 03-06-2017

Corresponding Author:

Rongjian Zheng

Key Laboratory of Advanced

Process Control for Light

Industry (Ministry of

Education), Jiangnan

University, Wuxi, China

E-mail: zhengrjian@163.com

**Abstract:** Glutamate fermentation is inherently nonlinear, multi-phase and an aerobic fermentation process. As long measurement delays and expensive apparatus cost, on-line measurement of the product concentration is not necessarily available. The present fermentation process monitoring and quality prediction involve manual interpretation of highly informative, however, the concentrations of substrates, biomass and products are only low frequency off-line measurements. In this paper, we propose a novel Multi-Phase Support Vector Regression (MPSVR) based soft sensor model for online quality prediction of glutamate concentration. The glutamate fermentation process can be divided into a sequence of five phases by detecting the trend variation events (also termed as singular points or inflection point) of online measured O<sub>2</sub> in the exhaust gas, the Inflection Point (IP) are easily identified through combining Moving Window (MW) with Pearson Correlation Coefficient (PCC). For each estimation phase, SVR soft sensor model are constructed and their performance is evaluated against fermentation data in a 5 L fermenter. The efficiency of the proposed soft sensor model for online product quality prediction has been demonstrated to be superior compared to that of reported techniques in a 5 L glutamate fermentation process.

**Keywords:** Multi-phase Quality Prediction, Glutamate Fermentation, Support Vector Regression, Soft Sensor

## Introduction

Glutamate is commercially one of the most important amino acids produced mainly by fermentation process, its fermentative production amount exceeds 2.2 million tons annually (Xiao *et al.*, 2006; Khan *et al.*, 2005). Glutamate are widely used for human and animal alimentation, as ingredients of pharmaceutical products, agrochemicals and some other industrial derivatives (Pal *et al.*, 2016).

Like other fed-batch processes, glutamate fermentation process requires a sophisticated operator involvement. In addition to run-to-run modifiability, this could result in abnormal situation, in which any deviation from desired operating regimes could lead to product quality change. This provides a strong incentive for automating operation supervision. During manual operation, a human operator is commonly responsible for

setting points to regulatory controllers, performing control actions, process supervising and taking remedial measures when an abnormal condition is detected (Muthuswamy and Srinivasan, 2003). The batch-to-batch change may root in the variation in the raw material quality or the seed culture variations. Typically, during manipulation, the product quality and batch performance are desired to be controlled and monitored by way of offline laboratory assays of concentrations of the product, substrates and biomass, which might take up to 2 h. These laboratory assays are high investment costs, manpower inputs and time consuming, are obtained at low frequencies and hence, may not necessarily acquire timely information about the fermentation status of the batch. Online measurements that are easily acquirable include temperature, pH, dissolved oxygen, agitation speed, exhaust CO<sub>2</sub> and O<sub>2</sub>, whereas these measurements do not show direct state of the process (Doan *et al.*,

2007). Consequently, this will lead to time delay to the quality control of product, since during this period of time, the fermentation process is without precise and continual information on the product quality (Ge *et al.*, 2011). Nevertheless, because of technical difficulty, high investment costs and large measurement delays, the laboratory assays apparatuses are limited use in practical plants. Furthermore, as we known, it requires a very significant effort to develop a first principles model for accurately depicting the fermentation process. Hence, in biochemical plants, soft sensors are used widely to estimate the primary quality variables that are difficult to measure online. An inferential model is constructed between objective variable which is difficult to measure online and process variables which are easy to measure online. (Facco *et al.*, 2009; Kadlec *et al.*, 2009; Kaneko *et al.*, 2009; 2011).

To date, many soft sensor methods have been presented for quality prediction objective, including Artificial Neural Network (ANN), Partial Least Squares (PLS) and Support Vector Machine (SVM) (Acuña *et al.*, 2014; Facco *et al.*, 2009; Wang *et al.*, 2014). Recently, SVR, an extension of SVM, has also been receiving increasing attention to solve nonlinear estimation problems. It has been successfully applied in different problems of time series prediction (Kavousi-Fard *et al.*, 2014; Lu *et al.*, 2009; Santamaria-Bonfil *et al.*, 2016; Were *et al.*, 2015).

Besides the inherently non-linear behavior, as we all known, biomass growth undergoes a series of phases in a fermentation process: Lag phase, exponential phase, stationary phase, decline phase (Khan *et al.*, 2005). The metabolism in each stage is different and each stage may have its special nature, using a single model will not be capable of entirely get the dynamic characteristic of the fermentation process. A straightforward method is to divide the fermentation process into different operation phases on the basis of the changes in variable cross-correlations and model each stage separately. Furthermore, minimize off-line sampling is desirable for the concomitant risk of contamination, at the same time, we need to obtain enough information on product formation and nurture uptake on-line. As a result, it is of critical importance to on-line identification of phases in fermentation process, phase partition is a crucial procedure before multi-phase modeling. The effectiveness of a multi-phase model is problematical without a proper phase division (Doan *et al.*, 2007; Sun *et al.*, 2011; Yao and Gao, 2009; Luo *et al.*, 2016).

In recent years, many phase identification methods have been developed by way of online analytical measurements of important bioprocess parameters such as the biomass concentration, or broth composition measurements, including use of ion chromatography, Near Infrared (NIR) and HPLC. These systems can

provide analysis of product concentration, nutrient compositions and other metabolites (Alford, 2006). However, these methods suffer from the aforementioned disadvantages. Another class of approach has focused on using the routinely available online data to qualitatively identify fermentation phases, this class most common methods include process knowledge, process analysis and the process data. A formal framework for inferring process trends from the online variables was exploited (Cheung and Stephanopoulos, 1990) and applied to fermentation process data (Stephanopoulos *et al.*, 1997; Doan *et al.*, 2007). Another method for detecting phase change uses singular points detection based only on online measurements (Maiti *et al.*, 2009; Régis *et al.*, 2008). The reader can find more multi-phase analysis methods in (Yao and Gao, 2009; Camacho *et al.*, 2008; Doan *et al.*, 2007; Luo *et al.*, 2016), which give different kinds of phase identification methods. Knowledge based phase identification fails when the process prior knowledge is not enough to divide processes into phases legitimately and difficult to customise for diverse fermentation processes. Process analysis based phase identification works well when certain required process features are known. Finally, process data based methods carry out phase partition by detecting variation in process data. Compared with the aforementioned two methods, data-driven methods are easier to perform because of their data-driven property. However, their phase partition results obtained by data-driven methods may or may not always consistent with actual operation phases (Sun *et al.*, 2011; Luo *et al.*, 2016).

In this paper, a novel phase partition method and a Multi-Phase SVR (MPSVR) modeling strategy are presented for online estimation and prediction of glutamate concentration. Glutamate fermentation process goes through a number of phases based on serial cell growth, substrate uptake and product formation. Besides, The production of glutamate is an aerobic process, the glutamate fermentation performance and the metabolic flux distribution are affected drastically by the concentrations of dissolved oxygen in the liquid phase in fermenter or oxygen concentrations in exhaust gas (Golobič and Gjerkeš, 1999; Xiao *et al.*, 2006), the fed-batch process can be divided into 5 phases based on the detection of Inflection Point (IP) by online measured O<sub>2</sub> in the exhaust gas, the IP are easily identified through combining Moving Window (MW) with Pearson Correlation Coefficient (PCC). The phase division result agrees well with actual fermentation process, it depends only on on-line measurements and fermentation processes can be easily automated to work. Then, for each estimation phase, SVR soft sensor models are designed for online prediction of glutamate concentration and their performance is evaluated against glutamate fermentation data in a 5 L fermenter. Also, a

comparison with Neural Networks (NN) based prediction approach in the literature is presented.

## Materials and Methods

### Experimental Methods

*C. glutamicum* S9114 was used in the present study, it was kept by the laboratory of industrial biotechnology, Jiangnan University. The fermentation conditions and the compositions of medium were the same as those ahead reported (Zhang *et al.*, 2005; Xiao *et al.*, 2006; Ding *et al.*, 2012; Cao *et al.*, 2013; Zheng and Pan, 2016).

*C. glutamicum* S9114 was cultured for glutamate production in a 5 L bioreactor. PH was controlled in 7.0-7.2 by feeding 25% (v/v) ammonia water. The O<sub>2</sub> and CO<sub>2</sub> concentrations in the exhaust gas were measured on-line by a gas analyzer (LKM2000, Lokas Co., Korea), Dissolved Oxygen (DO) was controlled at various levels by automatically or manually controlling the agitation speed, O<sub>2</sub> Uptake Rate (OUR) and CO<sub>2</sub> Evolution Rate (CER) were computed accordingly. Temperature was controlled at about 32°C. Electronic balances, which was connected to a PC via RS232, was used to compute the glucose and ammonia consumption rates (Zhang *et al.*, 2005; Xiao *et al.*, 2006; Ding *et al.*, 2012; Cao *et al.*, 2013; Zheng and Pan, 2016).

### Support Vector Regression

The SVR aims to provide a nonlinear mapping function to map the training data  $\{x_i, y_i; i = 1, \dots, n\}$  to a high dimensional feature space (Kavousi-Fard *et al.*, 2014). Then, the nonlinear relation can be represented as follows:

$$f(x) = w^T \phi(x) + b \quad (1)$$

where,  $w$  and  $b$  are the coefficients to be adjusted,  $\phi(x)$  denotes a mapping function of the feature space. The empirical risk can be defined as following:

$$R_{emp}(f) = \frac{1}{n} \sum_{i=1}^n \Theta_{\varepsilon}(y_i, f(x_i)) \quad (2)$$

where,  $\Theta_{\varepsilon}$  denote the  $\varepsilon$ -insensitive loss function and is described as follows:

$$\Theta_{\varepsilon}(y_i, f(x)) = \begin{cases} |f(x) - y| - \varepsilon, & \text{if } \Theta_{\varepsilon}(y_i, f(x)) \geq \varepsilon \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Then, an optimum hyper plane can be acquired by utilizing the function. With the help of hyper plane, the training data were divided into two linear separable

subsets with maximum separation distance. As it is, SVR is an optimizing problem with objective function is:

$$\text{Min}_{w, b, \xi, \xi^*} R_{\varepsilon}(w, \xi^*, \xi) = 0.5w^T w + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (4)$$

where,  $C$  is the regularization parameter. The constraint conditions of this optimization problem are as follows:

$$\begin{cases} y_i - w^T \phi(x_i) - b \leq \xi_i^* \varepsilon, & i = 1, \dots, n \\ -y_i + w^T \phi(x_i) + b \leq \xi_i \varepsilon, & i = 1, \dots, n \\ \xi_i, \xi_i^* \geq 0, & i = 1, \dots, n \end{cases} \quad (5)$$

By solving the above describing optimization problem, the coefficients of Equation (1) can be got as below:

$$w = \sum_{i=1}^n (\beta_i - \beta_i^*) \phi(x_i) \quad (6)$$

where,  $\beta_i$  is the Lagrangian coefficients. The SVR regression function can be described as below:

$$f(x) = \sum_{i=1}^n (\beta_i^* - \beta_i) K(x_i, x_j) + b \quad (7)$$

where,  $K(x_i, x)$  denotes the Kernel function, it can be described as follows in the feature space:

$$K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \quad (8)$$

In this study, the RBF Kernel function is utilized, it can be expressed as:

$$K(x_i, x_j) = e^{-\frac{\|x_i - x_j\|^2}{\sigma^2}} \quad (9)$$

where,  $\sigma$  denotes the width of the RBF.

### Phase Partition Technique

The statistical model used in this approach is

$$Y_t = a + bt + \varepsilon_t \quad (10)$$

where,  $1 \leq t \leq n$ , so  $t$  represents the time with the initial time taken as minute 1,  $b$  is the slope, it indicates the current variation trend of the O<sub>2</sub> with fermentation time  $t$ ,  $a$  is the intercept and  $\varepsilon$  are random errors. The errors are assumed to be identically distributed and independent. When working with the online measured O<sub>2</sub> data, this postulation may not be valid. Whereas, this assumption is more likely to be satisfied if time average is used as the response variable (Li *et al.*, 2013; Hess *et al.*, 2001).

After taking time averages, the regression coefficients are estimated by using the least squares method. Thus the estimates of the intercept and slope are given by:

$$\hat{b} = \frac{\sum (t - \bar{t})(Y - \bar{Y})}{\sum (t - \bar{t})^2} \quad (11)$$

$$\hat{a} = \bar{Y} - \hat{b}\bar{t} \quad (12)$$

where,  $\bar{t}$  and  $\bar{Y}$  are the arithmetic means for  $t_i$  and  $Y_i$ .

Since glutamate fermentation is a continuous process,  $b$  changes continuously and smoothly relative to  $Y$  ( $O_2$ ), we then employ Pearson Correlation Coefficient (PCC) to identify the Inflection Point (IP) of variation trend  $b$ . The PCC was developed by Karl Pearson from a related idea introduced by Galton in the late 19th century, it is a well-established measure of correlation and has range of -1 (perfect but negative correlation) to +1 (perfect correlation) with 0 denoting the short of a relationship (Adler and Parmryd, 2010).

PCC ( $r$ ) is given as follows:

$$r = \frac{\sum_{i=1}^m (b_i - \bar{b})(T_i - \bar{T})}{\sqrt{\sum_{i=1}^m (b_i - \bar{b})^2 \sum_{i=1}^m (T_i - \bar{T})^2}} \quad (13)$$

where,  $\bar{b}$  and  $\bar{T}$  is the means for  $b_i$  and  $T_i$ . Fig. 1 gives the phase partition procedure.

As the glutamate fermentation continues, the fermentation characteristics will change over time, the Moving Window (MW) technology can be used for tracking  $O_2$  changes. It is essential to discard the old data and add the newest data to the model. In fact, the most challenging thing lies in the selection of the window length, enough information can be included to detect the real change trend of parameter via setting suitable window size. If the window size is too small, the variation trend will be disturbed by the process noises and phases recognition may be lagged when the window size is too long (Yuan *et al.*, 2016). The window length of  $t$  is set to be 180 by trial-and-error, while the window length of  $b$  is set to be 60. Fig. 2 shows the recognition process of inflection point, the profile of the online measured  $O_2$  is shown in Fig. 2A. As shown in Fig. 2B, the  $r$  changes sharply at transition point of two different phases.

Glutamate fermentation experiences about five phases: Growth phase, transition phase, initial and middle production phases, late production stage, end of fermentation phase. Nevertheless, according to offline measurements and analysis, the growth of the cells passes through 5 phases: Lag phase, accelerate phase, decelerate phase, stationary phase and decline phase.

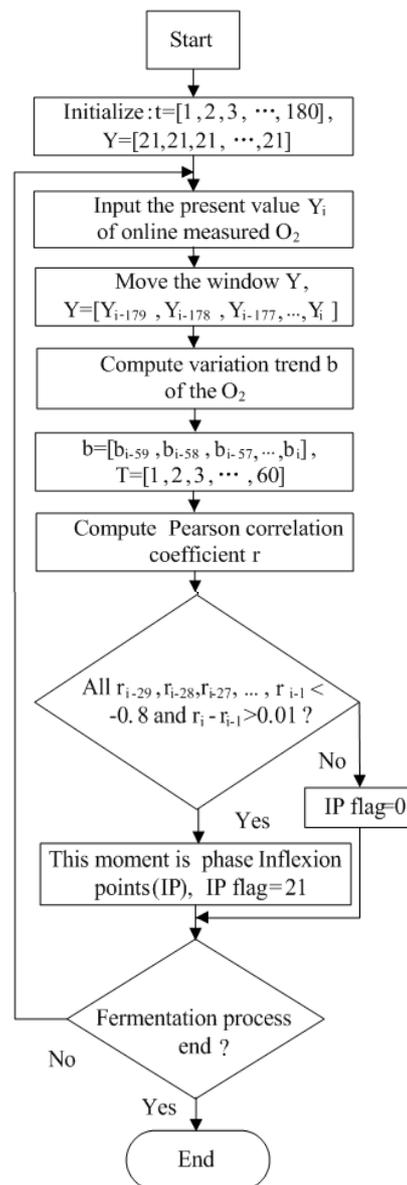


Fig. 1. Phase partition procedure

A key challenge for exploiting such a multi-phase partition is finding a appropriate  $b$  in normal fermentation process, analysis and experiments of the fermentation data for all the normal batches revealed that  $b_1$  is set to  $b_1 < 0$ , while  $(b_1/10) < b_2 < 0$ ,  $(b_1/3) < b_3 < b_2$ ,  $b_4 > 0$  and  $b_5 > b_4$  ( $b_1, b_2, b_3, b_4$  and  $b_5$  denotes phase 1, phase 2, phase 3, phase 4 and phase 5, respectively).

#### Quality Prediction Based on Multi-phase SVR

The prediction model based on Multi-Phase SVR (MPSVR) for glutamate concentration is shown in Fig. 3.

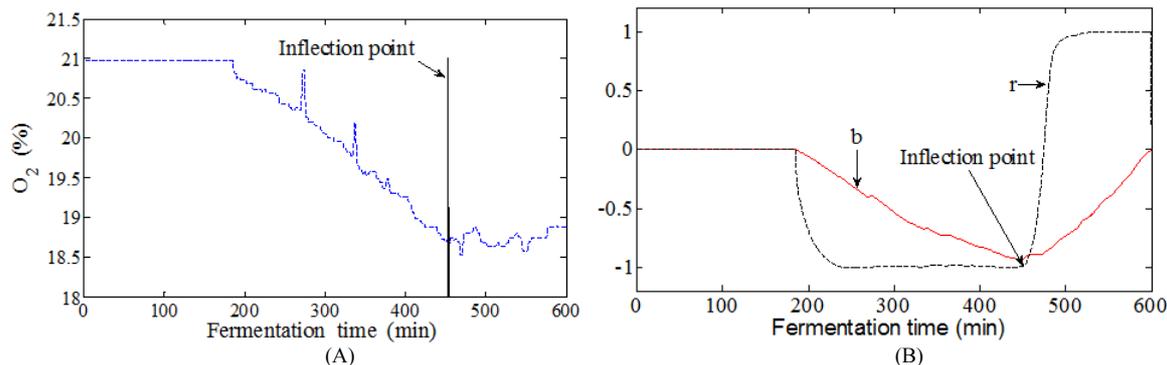


Fig. 2. Recognition process of inflection point

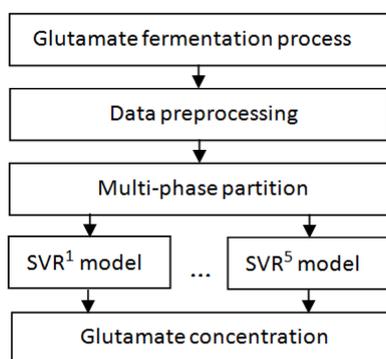


Fig. 3. Prediction model structure of multi-phase SVR

Table 1. Input and output variables of the prediction model

Input variables	Output variable
t	Glutamate concentration
T	
pH	
DO	
AG	
OUR	
CER	
AR	

In soft sensors, the secondary variables are used to act as the inputs of the soft sensor model and the primary variables such as glutamate concentration is employed to act as the output of soft sensor model.

The online measurement variables, such as fermentation time (t), fermentation Temperature (T), pH, Dissolved Oxygen Concentration (DO), Agitation Rate (AG), O<sub>2</sub> Uptake Rate (OUR), CO<sub>2</sub> Evolution Rate (CER) and ammonia water consumption Rate (AR) were chose as input variables. Table 1 shows the inputs and outputs configuration of the prediction model. There are 10 batches data for modeling and testing, 9 batch used for training and the other one for testing (Zheng and Pan, 2016).

In order to evaluate the prediction performance, Root Mean Square Error (RMSE) and coefficient of determination ( $R^2$ ) are computed as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (14)$$

$$R^2 = \frac{\left( n \sum_{i=1}^n \hat{y}_i y_i - \sum_{i=1}^n \hat{y}_i \sum_{i=1}^n y_i \right)^2}{\left( n \sum_{i=1}^n \hat{y}_i^2 - \left( \sum_{i=1}^n \hat{y}_i \right)^2 \right) \left( n \sum_{i=1}^n y_i^2 - \left( \sum_{i=1}^n y_i \right)^2 \right)} \quad (15)$$

where,  $y_i$  is the offline measured value;  $\hat{y}_i$  is the prediction value and  $n$  is the number of samples.

## Results and Discussion

The phase partition results of glutamate fermentation process are shown in Fig. 4, comparisons of multi-phase partition and offline measurements method are shown in Fig. 5. It can be seen that the phase partition results are well consistent with the physiological states of the real process by offline measurements and analysis method. The phase partition approach can be applied to other fermentation processes where minimal process information concerning phase shifts is available. Moreover, it can be used as a guide online in making a decision about the timing of off-line sampling. Note that in Fig. 4, range of variables has been stretched processing for clear show of each curve.

Once the phase recognition of fermentation process has been determined, the proposed MPSVR prediction models are therefore built for online product quality prediction of glutamate fermentation base on the phase partition results. For comparison, the SVR, NN and Multi-Phase NN (MPNN) models are equally constructed to predict production concentration with the same fermentation batches.

For the NN, Functions included to automatically train and test standard 1-layer neural networks using the MATLAB functions "train" and "sim". The number of hidden neurons is cross validated. For the SVR, standard support vector implementation for regression and function approximation using the Libsvm toolbox. Prediction values comparisons of different models are shown in Fig. 6.

Table 2. Quality prediction results of different models

Models	RMSE	R <sup>2</sup>
NN	7.2582	0.9694
SVR	3.7979	0.9942
MPNN	7.8957	0.9145
MPSVR	2.8653	0.9908

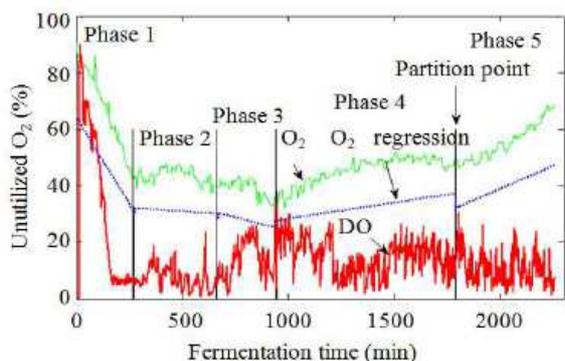


Fig. 4. Phase partition results of glutamate fermentation process

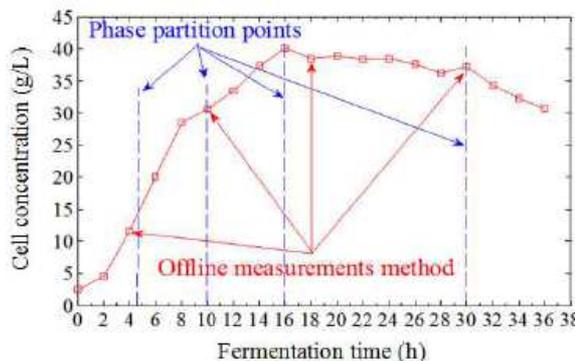


Fig. 5. Comparisons of multi-phase partition and offline measurements method

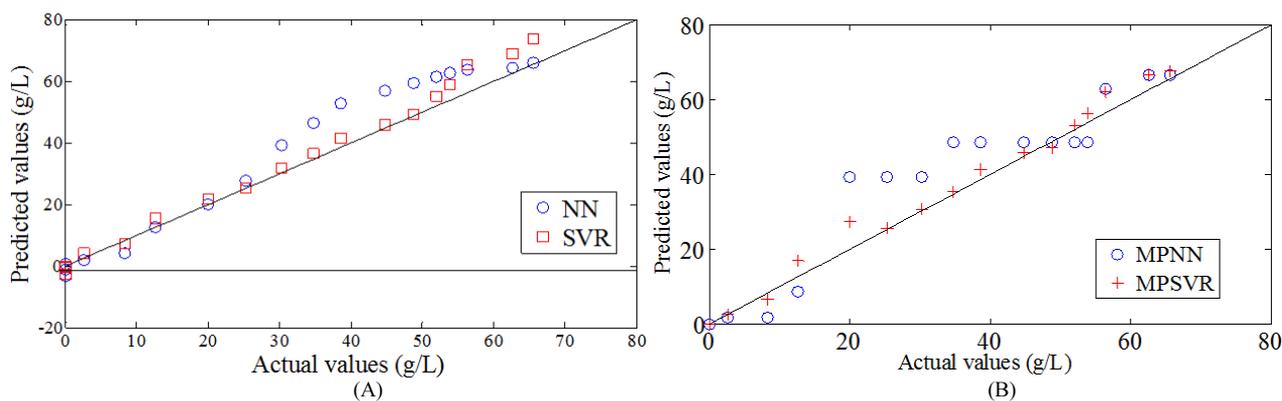


Fig. 6. Prediction values comparisons of different models. (A) Overall simple models, (B) Multi-phase models

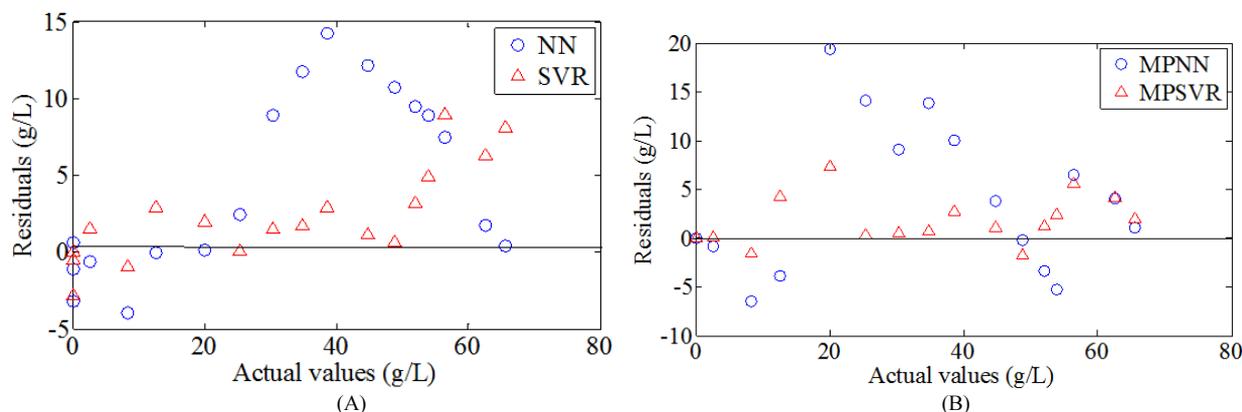


Fig. 7. Prediction residuals comparisons of different models. (A) Overall simple models, (B) Multi-phase models

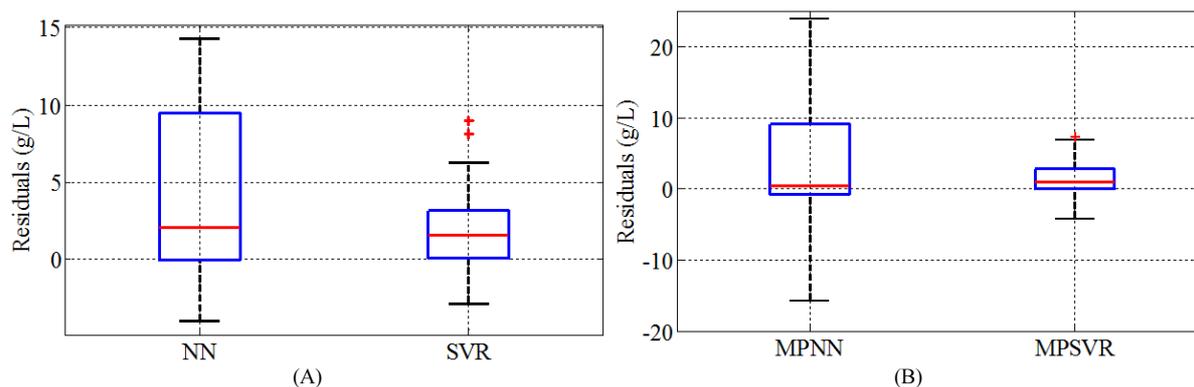


Fig. 8. Residuals boxplot comparisons of different models. (A) Overall simple models, (B) Multi-phase models

Prediction residuals and residuals boxplot comparisons of different models are shown in Fig. 7 and 8. It can be seen that the best prediction result has been obtained by the multi-phase SVR (MPSVR) model based soft sensor. The maximum residuals reaches 7.32 at fermentation time of 12 h. The reason is that it is decelerate phase, the bacteria activity becomes stronger at fermentation time of 12 h. Meanwhile, From Fig. 6-8, we can see that the MPSVR based prediction model exhibits better tracking performance than that by NN, SVR and MPNN models. Quality prediction results of different models are shown in Table 2, as can be seen that MPSVR has got the best prediction performance, since it has the minimum Root Mean Square Error (RMSE) 2.8653 among all prediction models, however, RMSE of SVR model is 3.7979. Furthermore,  $R^2$  of the MPSVR is 0.9908.

## Conclusion

In this paper, a soft sensor prediction model based on Multi-Phase SVR (MPSVR) was proposed for online quality prediction in glutamate fermentation process.

The states detection and partition is based on the detection of inflection points of the online measured  $O_2$  variations using Moving Window (MW) and Pearson Correlation Coefficient (PCC). The phase identification result agrees well with the physiologic phase changing of the glutamate fermentation process. Hence this method can help to automatically control and optimize the glutamate fermentation process. As the phase identification merely lies on the process data, it can be applied to other fermentation processes where minimal process knowledge regarding phase shifts is available. Besides, the phase partition approach can be used as a guide online in making decisions about the timing of off-line sampling.

The performance comparisons among SVR, NN, Multi-phase NN and Multi-phase SVR are implemented. Through glutamate fermentation process case study, the feasibility of the proposed MPSVR soft sensor has been conformed, the MPSVR model exhibits excellent

prediction performance, it can provide effective information for monitoring and operation of glutamate fermentation process.

## Acknowledgement

The Authors would like to appreciate the National Natural Science Foundation of China under Grant No.61273131 for financial support.

## Author's Contributions

**Rongjian Zheng:** Designed the research plan, participated in all experiments and writing the manuscript.

**Feng Pan:** Provided guidance, acquisition and analysis of data.

## Ethics

The authors declare that they have no conflict of interest.

## References

- Acuña, G., C. Ramirez and M. Curilem, 2014. Software sensors for biomass concentration in a SSC process using artificial neural networks and support vector machine. *Bioprocess Biosyst. Eng.*, 37: 27-36. DOI: 10.1007/s00449-013-0925-3
- Adler, J. and I. Parmryd, 2010. Quantifying colocalization by correlation: the Pearson correlation coefficient is superior to the Mander's overlap coefficient. *Cytometry Part A*, 77: 733-742. DOI: 10.1002/cyto.a.20896
- Alford, J.S., 2006. Bioprocess control: Advances and challenges. *Comput. Chem. Eng.*, 30: 1464-75. DOI: 10.1016/j.compchemeng.2006.05.039
- Camacho, J., J. Picó and A. Ferrer, 2008. Multi-phase analysis framework for handling batch process data. *J. Chemometrics*, 22: 632-643. DOI: 10.1002/cem.1151

- Cao, Y., E. Mpofu and Z.P. Shi, 2013. A novel metabolic model incorporating directed signal flow diagram with enzymatic activities data for evaluating glutamate yield in glutamate fermentation. *Biochem. Eng. J.*, 77: 136-146. DOI: 10.1016/j.bej.2013.05.015
- Cheung, J.T.Y. and G. Stephanopoulos, 1990. Representation of process trends-part I. A formal representation framework. *Comput. Chem. Eng.*, 14: 495-510. DOI: 10.1016/0098-1354(90)87023-I
- Ding, J., Y. Cao, E. Mpofu and Z.P. Shi, 2012. A hybrid support vector machine and fuzzy reasoning based fault diagnosis and rescue system for stable glutamate fermentation. *Chem. Eng. Res. Design*, 90: 1197-1207. DOI: 10.1016/j.cherd.2012.01.004
- Doan, X.T., R. Srinivasan, P.M. Bapat and P.P. Wangikar, 2007. Detection of phase shifts in batch fermentation via statistical analysis of the online measurements: A case study with rifamycin B fermentation. *J. Biotechnol.* 132: 156-166. DOI: 10.1016/j.jbiotec.2007.06.013
- Facco, P., F. Doplicher, F. Bezzo and M. Barolo, 2009. Moving average PLS soft sensor for online product quality estimation in an industrial batch polymerization process. *J. Process Control*, 19: 520-529. DOI: 10.1016/j.jprocont.2008.05.002
- Ge, Z.Q., T. Chen and Z.H. Song, 2011. Quality prediction for polypropylene production process based on CLGPR model. *Control Eng. Practice*, 19: 423-432. DOI: 10.1016/j.conengprac.2011.01.002
- Golobič, I., H. Gjerkeš and J. Malenšek, 1999. On-line estimation of the specific growth rate in the bacitracin fermentation process. *AIChE J.*, 45: 2550-2556. DOI: 10.1002/aic.690451211
- Hess, A., H. Lyer and W. Malm, 2001. Linear trend analysis: A comparison of methods. *Atmospheric Environ.*, 35: 5211-5222. DOI: 10.1016/S1352-2310(01)00342-9
- Kadlec, P., B. Gabrys and S. Strandt, 2009. Data-driven soft sensors in the process industry. *Comput. Chem. Eng.*, 33: 795-814. DOI: 10.1016/j.compchemeng.2008.12.012
- Kaneko, H., M. Arakawa and K. Funatsu, 2011. Novel soft sensor method for detecting completion of transition in industrial polymer processes. *Comput. Chem. Eng.*, 35: 1135-1142. DOI: 10.1016/j.compchemeng.2010.09.003
- Kaneko, H., M. Arakawa and K. Funatsu, 2009. Development of a new soft sensor method using independent component analysis and partial least squares. *AIChE J.*, 55: 87-98. DOI: 10.1002/aic.11648
- Kavousi-Fard, A., H. Samet and F. Marzbani, 2014. A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting. *Expert Syst. Applications*, 41: 6047-6056. DOI: 10.1016/j.eswa.2014.03.053
- Khan, N.S., I.M. Mishra, R.P. Singh and B. Prasad, 2005. Modeling the growth of corynebacterium glutamicum under product inhibition in L-glutamic acid fermentation. *Biochem. Eng. J.*, 25: 173-178. DOI: 10.1016/j.bej.2005.01.025
- Li, Y.P., J.D. Chen and F. Pan, 2013. Online monitoring of fermentation processes based on trend analysis. *J. Jiangnan University*, 12: 379-383.
- Lu, C.J., T.S., Lee and C.C., Chiu, 2009. Financial time series forecasting using independent component analysis and support vector regression. *Decision Support Syst.*, 47: 115-125. DOI: 10.1016/j.dss.2009.02.001
- Luo, L.J., S.Y. Bao, J.F. Mao, D. Tang and Z.L. Gao, 2016. Fuzzy phase partition and hybrid modeling based quality prediction and process monitoring methods for multiphase batch processes. *Industrial Eng. Chem. Res.*, 55: 4045-4058. DOI: 10.1021/acs.iecr.5b04252
- Maiti, S.K., R.K. Srivastava, M. Bhushan and P.P. Wangikar, 2009. Real time phase detection based online monitoring of batch fermentation processes. *Process Biochemistry*, 44: 799-811. DOI: 10.1016/j.procbio.2009.03.008
- Muthuswamy, K. and R. Srinivasan, 2003. Phase-based supervisory control for fermentation process development. *J. Process Control*, 13: 367-382. DOI: S0959-1524(02)00063-X
- Pal, P., R. Kumar, D. Vikramachakravarthi and S. Chakraborty, 2016. Modeling and simulation of continuous production of L (+) glutamic acid in a membrane-integrated bioreactor. *Biochemical Eng. J.*, 106: 68-86. DOI: 10.1016/j.bej.2015.11.008
- Régis, S., A. Doncescu and J. Desachy, 2008. Detection and characterization of physiological states in bioprocesses based on Hölder exponent. *Knowledge-Based Syst.*, 21: 70-79. DOI: 10.1016/j.knosys.2007.01.001
- Santamaria-Bonfil, G., A. Reyes-Ballesteros and C. Gershenson, 2016. Wind speed forecasting for wind farms: A method based on support vector regression. *Renewable Energy*, 85: 790-809. DOI: 10.1016/j.renene.2015.07.004
- Stephanopoulos, G., G. Locher, M. Duff and R. Kamimura, 1997. Fermentation database mining by pattern recognition. *Biotechnol. Bioeng.*, 53: 443-452. DOI: 10.1002/(SICI)1097-0290(19970305)53:5
- Sun, W., Y. Meng, A. Palazolu, J.S. Zhao and H. Zhang *et al.*, 2011. A method for multiphase batch process monitoring based on auto phase identification. *J. Process Control*, 21: 627-638. DOI: 10.1016/j.jprocont.2010.12.003
- Wang, T., J.W. Sun, W.D. Zhang and J.Q. Yuan, 2014. Prediction of product formation in 2-keto-l-gulononic acid fermentation through Bayesian combination of multiple neural networks. *Process Biochem.*, 49: 188-194. DOI: 10.1016/j.procbio.2013.11.003

- Were, K., D.T. Bui, Ø.B. Dick and B.R. Singh, 2015. A comparative assessment of support vector regression, artificial neural networks and random forests for predicting and mapping soil organic carbon stocks across an Afrotropical landscape. *Ecological Indicators*, 52: 394-403.  
DOI: 10.1016/j.ecolind.2014.12.028
- Xiao, J., Z.P. Shi, P. Gao, H.J. Feng and Z.Y. Duan *et al.*, 2006. On-line optimization of glutamate production based on balanced metabolic control by RQ. *Bioprocess Biosyst. Eng.*, 29: 109-117.  
DOI: 10.1007/s00449-006-0059-y
- Yao, Y. and F.R. Gao, 2009. A survey on multistage/multiphase statistical modeling methods for batch processes. *Annu. Rev. Control*, 33: 172-183.  
DOI: 10.1016/j.arcontrol.2009.08.001
- Yuan, X.F., Z.Q. Ge and Z.H. Song, 2016. Spatio-temporal adaptive soft sensor for nonlinear time varying and variable drifting processes based on moving window LWPLS and time difference model. *Asia-Pac. J. Chem. Eng.*, 11: 209-219.  
DOI: 10.1002/apj.1957
- Zhang, C.Y., Z.P. Shi, P. Gao, Z.Y. Duan and Z.G. Mao, 2005. On-line prediction of products concentrations in glutamate fermentation using metabolic network model and linear programming. *Biochem. Eng. J.*, 25: 99-108. DOI: 10.1016/j.bej.2005.03.012
- Zheng, R.J. and F. Pan, 2016. Soft sensor modeling of product concentration in glutamate fermentation using gaussian process regression. *Am. J. Biochemistry Biotechnol.*, 12: 179-187.  
DOI: 10.3844/ajbbsp.2016.179.187