



Process dynamics, control and monitoring

A Bayesian inference based two-stage support vector regression framework for soft sensor development in batch bioprocesses

Jie Yu

Department of Chemical Engineering, McMaster University, Hamilton, Ontario, Canada L8S 4L7

ARTICLE INFO

Article history:

Received 6 September 2011

Received in revised form 25 February 2012

Accepted 10 March 2012

Available online 28 March 2012

Keywords:

Soft sensor

Batch process

Support vector regression

Bayesian inference

Process and measurement uncertainty

Fed-batch penicillin cultivation process

ABSTRACT

Inherent process and measurement uncertainty has posed a challenging issue on soft sensor development of batch bioprocesses. In this paper, a new soft sensor modeling framework is proposed by integrating Bayesian inference strategy with two-stage support vector regression (SVR) method. The Bayesian inference procedure is first designed to identify measurement biases and misalignments via posterior probabilities. Then the biased input measurements are calibrated through Bayesian estimation and the first-stage SVR model is thus built for output measurement reconciliation. The inferentially calibrated input and output data can be further used to construct the second-stage SVR model, which serves as the main model of soft sensor to predict new output measurements. The Bayesian inference based two-stage support vector regression (BI-SVR) approach is applied to a fed-batch penicillin cultivation process and the obtained soft sensor performance is compared to that of the conventional SVR method. The results from two test cases with different levels of measurement uncertainty show significant improvement of the BI-SVR approach over the regular SVR method in predicting various output measurements.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Batch or semi-batch processes have been extensively used to produce low-volume but high value-added goods across various industries including special chemicals, polymers, biopharmaceuticals, semiconductors and materials. During batch operation, it is critically important to stabilize process, maintain safety and ensure product quality within specifications. Effective process control and monitoring have become essential to achieve such objectives in industrial plants (Chiang, Russell, & Braatz, 2001; Yu & Qin, 2009b). As prior requirements, the reliable sensor measurements and data collections play crucial roles in process automation, monitoring and optimization (Alford, 2006; Kadlec, Gabrys, & Strandt, 2009; Lee & Lee, 2007; Yu, 2011a). The lack of robust hardware for measuring key quality variables has posed a challenging issue for batch bioprocess control and monitoring (Clementschitsch & Bayer, 2006; Yu, 2011b).

Over the past two decades, soft sensors have attracted increasing attention in both academia and industry due to its inferential measurement capability based on process data and predictive models (Dochain & Perrier, 1997; Fortuna, Graziani, Rizzo, & Xibilia, 2007; Kano & Nakagawa, 2008). With soft sensors being developed, the variable measurements rely on the virtual computer programs instead of physical hardwares. The soft sensor methods in

literature can normally be attributed to two categories, one of which is mechanistic model based while the other is historical data driven (Kadlec et al., 2009). The former class of approaches require in-depth process knowledge to develop first-principle models, which are used to illustrate and characterize the complicated physical, chemical or biological relationships among the process and quality variables (Bastin & Dochain, 1990; Doyle, 1998; Pons, Rajab, Flaus, Engasser, & Cheruy, 1988). As a result, the effort of mechanistic model based soft sensors can be quite heavy and time consuming. Furthermore, the inadequate process understanding may make it impossible to build accurate first-principle models. In contrast, the data driven soft sensor techniques have been more appealing as they do not need a priori process knowledge but depend on historical process data only. The early attempt of utilizing data driven methods for soft sensor development has been focused on multivariate statistical techniques such as principal component analysis (PCA) and partial least squares (PLS) (Kano, Miyazaki, Hasebe, & Hashimoto, 2000; Kresta, Marlin, & MacGregor, 1994; Lin, Recke, Knudsen, & Jorgensen, 2007; Mejdell & Skogestad, 1991; Neogi & Schlags, 1998; Zamprogna, Barolo, & Seborg, 2005; Zhang & Lennox, 2004). This type of methods can deal with the variable co-linearity and identify the model within the projected lower-dimensional subspace that retains the covariance structure. However, the PCA and PLS methods are essentially linear modeling techniques and cannot handle process nonlinearity unless certain nonlinear variations like kernel functions are integrated (Zhang, Yan, & Shao, 2008). Alternately, artificial neural networks (ANNs) have been widely applied to build soft sensor models for

E-mail address: jieyu@mcmaster.ca

nonlinear processes (Assis & Filho, 2000; Dufour, Bhartiya, Dhurjati, & Doyle, 2005; Montague, Morris, & Tham, 1992). Moreover, some research effort has been explored to incorporate neural networks with multivariate statistical techniques like PCA and PLS for building nonlinear soft sensors (Qin & McAvoy, 1992; Qin, Yue, & Dunia, 1997). ANN is employed to model the nonlinear relationship in the processes while the PCA or PLS based latent variable procedure is adopted to handle the collinearity of measurement data as well as to replace the missing values. Though ANN technique has been proven capable of handling system nonlinearity and representing process knowledge (Hoskins & Himmelblau, 1988; Venkatasubramanian & Chan, 1989), its generalization ability cannot be guaranteed so that the well trained model may lead to poor predictions for new observations. In addition, the solution of neural networks often converges to local minima and thereby the obtained soft sensor measurements can be suboptimal.

More recently, some new machine learning techniques like independent component analysis (ICA) and support vector machine (SVM) have been applied to soft sensor development (Chen, Song, & Li, 2005; Desai, Badhe, Tambe, & Kulkarni, 2006; Kaneko, Arakawa, & Funatsu, 2009; Yan, Shao, & Wang, 2004; Yu, 2012). The ICA method is able to extract mutually independent components based on higher-order statistics and identify the non-Gaussian process features that cannot be dealt with by traditional PCA or PLS methods. SVM approach, on the other hand, overcomes the drawbacks of ANN and has the desirable properties such as globally optimal solution, strong generalization ability and self-learning capacity on network structure. Yan et al. suggested to use Bayesian evidence method to select the SVM model parameter values and optimize the soft sensor predictions. In spite of different methods for soft sensor modeling, they essentially lead to deterministic models for soft sensors and do not take into full consideration the process and measurement uncertainty due to various practical factors including missing values, unmeasured disturbances, measurement errors and drifts, outliers, sample delays and irregular data frequency. As a powerful probabilistic technique to handle system uncertainty, Bayesian analysis is gaining interest in various areas of process systems engineering including soft sensor estimation (Chen, Bakshi, & Goel, 2009; Khartibisepehr & Huang, 2008; Yu & Qin, 2008, 2009a).

In this study, a two-stage support vector regression framework is integrated with Bayesian analysis for soft sensor development in nonlinear fed-batch bioprocesses. Firstly, the multiway batch data matrix is unfolded and scaled in the preprocessing step. Then, the measurement data of process variables are used to estimate the posterior probabilities of all samples within the model input space through Bayesian inference. The subset of samples with the corresponding posterior probabilities less than the pre-defined confidence level are identified and calibrated with the Bayesian estimation technique. Further, the cleaned process measurements along with the output data of quality variables are employed to build the first-stage support vector regression model. The predicted values of training set are then used to estimate the posterior probabilities of output data. Likewise, the abnormal quality measurements with posterior probabilities outside of confidence bound can be detected and re-evaluated with Bayesian procedure. Thus the second-stage SVR model can be fit from the calibrated input and output data for the soft sensor predictions.

The remainder of the article is organized as follows. Section 2 reviews the support vector machine based model regression. The two-stage Bayesian inference based support vector regression (BI-SVR) approach is then described in Section 3. Section 4 shows the utility of the proposed soft sensor approach through the application of fed-batch penicillin fermentation process and its comparison to the regular SVR method. Finally, this work is concluded in Section 5.

2. Review of support vector regression

Support vector machine is a type of machine learning technique with attractive usages in classification, regression and probability density function estimation (Vapnik, 1995, 1998). In the support vector regression scheme, given a series of training data $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ with $x \in R^d$ being the d -dimensional input samples and $y \in R$ the output observations, the regression problem in the linear case is expressed as

$$f(x) = \omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n + b = \langle \omega, x \rangle + b \quad (1)$$

where $\omega = [\omega_1, \omega_2, \dots, \omega_n]^T$ represents the regression coefficients and b is the bias term. It can be solved from the following optimization problem (Smola & Schölkopf, 2004; Vapnik, 1995):

$$\min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n (\xi_i^- + \xi_i^+) \quad (2)$$

subject to the constraints

$$y_i - \langle \omega, x_i \rangle - b \leq \epsilon + \xi_i^- \quad (3)$$

and

$$\langle \omega, x_i \rangle + b - y_i \leq \epsilon + \xi_i^+ \quad (4)$$

where ϵ is the precision threshold, C denotes the regularization parameter and ξ^- and ξ^+ represent the slack variables with non-negative values to ensure feasible constraints. With ϵ -insensitive loss function, the optimization problem is reformulated as

$$\begin{aligned} \max_{\alpha, \alpha'} & -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \{(\alpha_i - \alpha'_i)(\alpha_j - \alpha'_j) \langle x_i, x_j \rangle\} \\ & + \sum_{i=1}^n \{\alpha_i (y_i - \epsilon) - \alpha'_i (y_i + \epsilon)\} \end{aligned} \quad (5)$$

where the Lagrange multipliers α_i and α'_i satisfy the conditions $0 \leq \alpha_i, \alpha'_i \leq C$ and $\sum_{i=1}^n (\alpha_i - \alpha'_i) = 0$.

In nonlinear case, the above objective function can be modified by substituting the dot product with a kernel function as follows (Gunn, 1998; Smola & Schölkopf, 2004; Vapnik, 1995):

$$\begin{aligned} \max_{\alpha, \alpha'} & -\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \{(\alpha_i - \alpha'_i)(\alpha_j - \alpha'_j) K(x_i, x_j)\} \\ & + \sum_{i=1}^n \{\alpha_i (y_i - \epsilon) - \alpha'_i (y_i + \epsilon)\} \end{aligned} \quad (6)$$

By selecting the support vectors whose Lagrange multipliers α_i and α'_i are greater than zero, the regression function may be rewritten as

$$f(x) = \sum_{i=1}^p (\alpha_i - \alpha'_i) K(x, x_i) + b \quad (7)$$

where p is the number of support vectors. Then the regression coefficients and bias can be computed as

$$\langle \omega, x \rangle = \sum_{i=1}^n (\alpha_i - \alpha'_i) K(x, x_i) \quad (8)$$

and

$$b = \begin{cases} y_i - f(x_i, b=0) - \epsilon & \text{if } 0 \leq \alpha_i \leq C \\ y_i - f(x_i, b=0) + \epsilon & \text{if } 0 \leq \alpha'_i \leq C \end{cases} \quad (9)$$

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات