



Adaptive soft sensor modeling framework based on just-in-time learning and kernel partial least squares regression for nonlinear multiphase batch processes

Huaiping Jin, Xiangguang Chen*, Jianwen Yang, Lei Wu

Department of Chemical Engineering, Beijing Institute of Technology, Beijing 100081, People's Republic of China

ARTICLE INFO

Article history:

Received 20 December 2013
Received in revised form 3 April 2014
Accepted 19 July 2014
Available online 28 July 2014

Keywords:

Adaptive soft sensor
Batch process
Kernel partial least squares
Just-in-time learning
Partial mutual information
Chlortetracycline fermentation process

ABSTRACT

Batch processes are characterized by inherent nonlinearity, multiple phases and time-varying behavior that pose great challenges for accurate state estimation. A multiphase just-in-time (MJIT) learning based kernel partial least squares (KPLS) method is proposed for multiphase batch processes. Gaussian mixture model is estimated to identify different operating phases where various JIT-KPLS frameworks are built. By applying Bayesian inference strategy, the query data is classified into a particular phase with the maximal posterior probability, and thus the corresponding JIT-KPLS framework is chosen for online prediction. To further improve the predictive accuracy of the MJIT-KPLS algorithm, a hybrid similarity measure and an adaptive selection strategy are proposed for selecting local modeling samples. Moreover, maximal similarity replacement rule is proposed to update database. A procedure of input variable selection based on partial mutual information is also presented. The effectiveness of the MJIT-KPLS algorithm is demonstrated through application to industrial fed-batch chlortetracycline fermentation process.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

Batch or semibatch processes have been widely used to produce special chemicals, materials for microelectronics, pharmaceutical and agricultural products (Cinar et al., 2003). As prior requirements, the reliable real-time measurements play a crucial role in process automation, monitoring and optimization (Alford, 2006). However, the lack of reliable online sensors, which can accurately detect the important state variables, has become one of the major challenges of controlling batch processes accurately, automatically and optimally (Chen et al., 2006; Dochain, 2008; Nicoletti and Jain, 2009).

Over the past two decades, soft sensors have received increasing attention in both academia and industry due to their inferential estimation capability. Although soft sensors have been applied in broad fields, the online prediction remains the dominant application (Kadlec et al., 2009). When online analyzers are not available, soft sensing technology aims to provide estimations of difficult-to-measure variables based on some easy-to-measure variables. Generally, soft sensors can be classified into two groups: first-

principle models and data-driven models (Kadlec et al., 2009). The focus of this work is on the data-driven soft sensor modeling. Reviews of this type of soft sensor application have been published in references Fortuna et al. (2007), Haimi et al. (2013), Kadlec et al. (2009), Kano and Nakagawa (2008), Kano and Fujiwara (2013), Slišković et al. (2011).

The most common data-driven modeling techniques for developing soft sensors are multivariate statistical techniques such as multivariate linear regression (MLR) (Kano and Ogawa, 2010), principle component regression (PCR) (Jolliffe, 2002) and partial least squares (PLS) (Lin et al., 2007; Kano and Fujiwara, 2013; Sharmin et al., 2006). These linear modeling methods account for 90% of the soft sensors used in industry (Kano and Ogawa, 2010) due to their statistical background, ease of interpretability, and because they can deal efficiently with data collinearity, which is common among industrial datasets (Kadlec and Gabrys, 2011). Nevertheless, these linear methods cannot function well when applied to highly nonlinear processes like batch processes. Thus many efforts have been attempted to nonlinear approaches, such as artificial neural networks (ANN) (Cui et al., 2012; Gonzaga et al., 2009), KPLS (Jia et al., 2013; Yu, 2012a), support vector regression (Yu, 2012b), neuro-fuzzy system (Jang, 1993; Jassar et al., 2009) and Gaussian process regression (GPR) (Grbić et al., 2013). However, there exist some issues remained to be solved to develop soft sensors for batch processes.

* Corresponding author. Tel.: +86 13601333018; fax: +86 010 68914662.

E-mail addresses: jinhuaiping@gmail.com (H. Jin), xgc1@bit.edu.cn (X. Chen), yangjianwen@bit.edu.cn (J. Yang), wulei@bit.edu.cn (L. Wu).

One particular drawback of many current soft sensors is their non-adaptive nature. Traditionally, the predictive models are not adaptive, and once deployed into the real-life operation, the models will not change, whereas the operation environment is often changing. To cope with changes in process characteristics, model maintenance is essential to maintain high estimation performance for a long time. Without process or expert knowledge, the soft sensor model has to be updated automatically. Thus many kinds of recursive modeling methods, which update models by prioritizing newer samples, have been proposed (Kadlec et al., 2011). Although these methods can adapt the soft sensor to a new operation condition recursively, they cannot cope with abrupt changes in process characteristics, which are caused by replacement of a catalyst, cleaning of equipment, etc., because a query sample just after an abrupt change becomes significantly different from the prioritized samples. It is also a common practice to enhance the model adaptation by the ensemble learning framework (Grbić et al., 2013; Kadlec et al., 2011; Kadlec and Gabrys, 2011; Yu, 2012c). In the ensemble framework, data are divided into different sub-domains and local sub-models are constructed over each domain. In this way, instead of using a single global model, multiple simpler models are developed and then combined to obtain the final prediction. Some other adaptive soft sensors were developed by partitioning the process data into multiple clusters corresponding to different operating phases where local predictive model is built for each phase (Yu, 2012a; Yu and Qin, 2008, 2009). When implemented online, the query sample, for which an output estimation is required, is firstly classified into a particular phase, and then the corresponding local model representing the identical phase is adaptively chosen for prediction. Although such phase based multi-model methods outperforms the single models, they are unable to effectively capture the between-phase transient dynamics. Thus, a Bayesian model averaging (BMA) based multi-model method was proposed to tackle this issue (Yu et al., 2013). In practice, however, it is difficult to determine the partition number due to the lack of the quantitative and precise information of phase divisions. More importantly, the local models used in the BMA based multi-model methods are built offline and not updated once deployed online, thus the changes in process characteristics cannot be well dealt with due to the time-varying nature of the real-life processes.

Recently, just-in-time (JIT) learning has attracted increasing attention in process modeling and soft sensor development (Cheng and Chiu, 2004; Fujiwara et al., 2009). By applying JIT learning, a local model is constructed from the samples similar to the query sample. Thus, on one hand, JIT based model can cope with abrupt changes as well as gradual ones. On the other hand, it can deal with nonlinearity since it builds a local model repeatedly. Compared to the traditional modeling methods which can be considered as global modeling, JIT based method exhibits a local model structure where a local model is built from the historical dataset selected by some similarity measure to the query data when the estimation is required. Once the estimated output is obtained, the built local model is discarded. Nevertheless, the estimation accuracy of JIT based models is expected to be further improved by selecting the optimal combination of local regression function, input variable selection, similarity measure, and database updating scheme, etc., simultaneously.

To enhance the predictive performance and adaptability of JIT based models, there are some problems remain to be addressed. Although linear local model based JIT methods (Cheng and Chiu, 2004; Kim et al., 2013) can successfully address the process nonlinearity, high nonlinearity variable relationships are widely existed in batch processes where a local linear model may not always function well. Thus a nonlinear modeling technique with high computational efficiency is preferable. In addition, the commonly used similarity measures rarely take into account process

characteristics. In real applications, the most frequently used similarity measures are based on the Euclidean distance or the Mahalanobis distance (Cheng and Chiu, 2004; Ge and Song, 2010; Schaal et al., 2002), most of which are defined only from the perspective of sample algebraic space irrespective of specific process knowledge. Moreover, the optimal local modeling size is usually determined offline as reported in (Ge and Song, 2010). An adaptive strategy is required to choose local modeling samples adaptively for each query data. Besides, a reliable database updating scheme is also crucial for JIT based models to tackle changes in process characteristics. Usually, the database is updated only by simply removing the oldest samples and adding the new samples (Shigemori et al., 2011).

The second problem encountered in soft sensor modeling is the lack of a systematic guideline for input variable selection. As reported in literature work (Cui et al., 2012; Kadlec and Gabrys, 2011; Kim et al., 2013; Pani et al., 2013), input variables are often selected based on engineers' personal experience and prior process knowledge. However, it is time-consuming for the engineers to select the input variables since trial and error is inevitable. Additionally, the selected variables may not be optimal. Also, it becomes very difficult even for experienced engineers to properly select input variables when a large number of variables are measured and physical and chemical phenomena are not sufficiently understood. Consequently, various data-based methods have been proposed for selecting proper input variables.

One popular approach for reducing the input variable dimension is by projecting the original input space into an adequate lower dimensional space. The most popular methods for achieving such projection task are based on linear projection of input space, such as the widely known principle component analysis (PCA) and PLS. However, the new variables resulting from such methods are difficult to interpret in terms of actual process variables (Delgado et al., 2009). More importantly, the underlying assumption of linearly structured dependence contradicts to the development of statistical model of nonlinear processes.

Another approach for reducing input space dimension is based on selecting the most important variables from all potential variables according to some criteria. As a nonparametric and nonlinear measure of relevance derived from information theory, mutual information (MI) has recently been applied to nonlinear processes modeling (May et al., 2008a), process monitoring (Chen et al., 2013; Rashid and Yu, 2012a, 2012b) and soft sensor design (Grbić et al., 2013). Unlike those linear methods that only consider linear relationships between variables, MI is theoretically able to identify relations of any type. It furthermore makes no assumption about the distribution of the data. However, several issues have arisen in the formulation of MI-based selection algorithms, which are: the ability of handling the inter-dependencies between candidates and the lack of an appropriate principle for determining when to halt the selection procedure (Chow and Huang, 2005). To tackle this issue, the partial mutual information (PMI) criterion has been developed by considering the effect of the already selected input variables when evaluating the relevance between one plausible input variable and the output variable (May et al., 2008a, 2008b; Sharma, 2000).

Apart from the model adaptation and input variable selection, during data-driven soft sensor modeling, much attention is only paid to the plant data, whereas the process characteristics are usually ignored. In practice, batch processes are often characterized by their multiphase characteristics where multiple operating phases are involved. Thus the accuracy and reliability of quality variable prediction can heavily degrade as the operating phase and process dynamics change. Multiphase modeling strategy with phase identification has been reported more efficient than the conventional single-model based methods (Yu, 2012a; Yu and Qin, 2008,

Download English Version:

<https://daneshyari.com/en/article/6595508>

Download Persian Version:

<https://daneshyari.com/article/6595508>

[Daneshyari.com](https://daneshyari.com)