G Model
JJPC-1756; No. of Pages 15
ARTICLE IN PRESS
Journal of Process Control xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Journal of Process Control

journal homepage: www.elsevier.com/locate/jprocont

Phase analysis and statistical modeling with limited batches for multimode and multiphase process monitoring

Chunhui Zhao *

State Key Laboratory of Industrial Control Technology, Department of Control Science and Engineering, Zhejiang University, Hangzhou 310007, China

ARTICLE INFO

Article history:
Received 27 August 2013
Received in revised form 18 March 2014
Accepted 6 April 2014
Available online xxx

Keywords:
Limited batches
Multimode
Between-mode relative analysis
Concurrent phase partition
Batch process monitoring

ABSTRACT

For multimode batch processes, the conventional modeling methods in general require that sufficient batches should be available for every mode, which, however, cannot be guaranteed in practice. It may be impractical to conduct enough trial runs and wait until sufficient batches are available before development of monitoring models for each mode. Starting from limited batches, how to derive reliable process information and develop monitoring models has been an important question for successful online multimode batch process monitoring. To address this problem, this article proposes a phase analysis and statistical modeling strategy with limited batches. One mode which has obtained sufficient batches is chosen as the reference mode while the other modes which can only get limited batches work as alternative modes. Starting from limited batches, the proposed algorithm addresses two issues, concurrent phase partition and analysis of between-mode relative changes. First, for each alternative mode, generalized time-slices are constructed by combining several consecutive time-slices within a short time region to explore local process correlations. The time-varying characteristics are then concurrently analyzed across modes so that multiple sequential phases are identified simultaneously for all modes. Then phase-representative data units are arranged by variable-unfolding the conventional time-slices for the reference mode and the generalized time-slices for each alternative mode respectively. Between-mode statistical analysis is performed within each phase where the relative changes from the reference mode to each alternative mode are analyzed. From the between-mode perspective, different types of relative variations in each alternative mode are separated and modeled for online monitoring. Starting from limited batches, online batch process monitoring can be conducted, providing reliable fault detection performance. The proposed algorithm is illustrated with a typical multiphase batch process with multiple modes.

1. Introduction

Batch and semi-batch processes play an important role in the processing of specialty chemical, semiconductor, food and biology industries for producing high-value-added products to meet today's rapidly changing market. Characterized by finite duration, batch process operation is in general carried out in different steps to produce products of desired quality at the lowest possible cost. Process disturbances, which may vary with the development of time and from batch to batch, affect both process and product reproducibility. Proper process monitoring and diagnosis is important to not only process safety but also quality improvement [1–7]. Multivariate statistical methods such as principal component analysis (PCA) [8] and partial least square (PLS) [9,10] have been successfully used in modeling multivariate continuous processes. Several extensions of the conventional PCA/PLS to batch processes have also been reported, among which most batch process monitoring methods are based on multi-way principal component analysis (MPCA) and partial least squares (MPLS) [11–16]. Variable-unfolding modeling methods [17] also have been developed which can be put into online application without fulfilling missing data. However, they may not well reflect multiplicity of process characteristics across different time regions since they use a single monitoring model to describe the entire process. Considering that the multiplicity of operation phases is an inherent nature of many batch processes and each phase exhibits significantly different underlying behaviors, it is desirable to develop multiple phase-based models [6]. Then each model represents a specific phase and explains local process behaviors, which can effectively enhance process understanding and improve monitoring reliability. Phase-specific nature [18–20]

* Tel.: +86 571 87951879; fax: +86 571 87951879.
E-mail addresses: chhzhao@zju.edu.cn, huihuizh@gmail.com

http://dx.doi.org/10.1016/j.jprocont.2014.04.001
0959-1524/

Please cite this article in press as: C. Zhao, Phase analysis and statistical modeling with limited batches for multimode and multiphase process monitoring, J. Process Control (2014), http://dx.doi.org/10.1016/j.jprocont.2014.04.001

of batch processes has drawn increasing attentions and is handled with different methods. Automatic phase division algorithms and sub-phase modeling methods [20–26] have been developed and widely used. They do not require fulfilling missing process observations and well preserve dynamic process relationships. Also they can help to understand the operation process, revealing similar process characteristics within a certain time region which can be described by a unified model structure. Since then, phase-based sub-PCA/PLS modeling methods have been well developed and widely used to handle different problems in batch process monitoring with multiphase nature.

The wide and successful report of phase-based modeling and monitoring strategies, however, is in general limited to single-mode batch processes. In practice, manufacturing processes frequently go through mode changes due to various factors, such as alterations of feedstock and compositions, fluctuations in the external environment, and various product specifications. Especially, to satisfy fast changing market demands, the manufacturing strategies and operation conditions have to be adjusted frequently which also makes multimode a popular and important problem. When the process behaviors to be evaluated are operating under a normal condition other than the one the reference model is built upon, false alarms will be issued, indicating that the current process is abnormal. The statistical modeling and monitoring of multimode processes is a challenging problem, which has drawn increasing attention recently [27–29] and most of which are for continuous processes. However, few research work regarding statistical analysis and monitoring of multimode and multiphase batch processes have been reported. Moreover, for conventional modeling methods, it is in general required that sufficient batches should be available for every mode to cover statistically normal batch-to-batch variation, which, however, cannot be guaranteed in practice. Sometimes, only limited batches are available for some modes. It is impractical to conduct enough trial runs and wait until sufficient batches are available before development of monitoring models for each mode. It is hoped that initial monitoring models can be developed only based on limited batches and used for online application before sufficient batches are available. Therefore, multiphase and multimode analysis with limited batches is of significance.

In the present work, the problem of multimode and multiphase modeling and monitoring with limited batches is addressed. One mode which has collected sufficient batches is chosen as reference mode while the other modes which only have limited batches work as alternative modes. Using conventional methods, the monitoring models for each alternative mode cannot be obtained only based on limited batches. The objective of the current work is to reveal the changes of process characteristic of each alternative mode starting from limited batches and then develop the initial monitoring models for online process monitoring. While the conventional time-slices are directly used for the reference mode, generalized time-slices are organized for each alternative mode each of which is composed of several consecutive time-slices. Concurrent phase partition is then performed by simultaneously checking the time-varying process characteristics of the reference mode with sufficient batches and alternative modes with limited batches. After that, to capture the changes from the reference mode to each alternative mode, phase-based between-mode relative analysis is performed. The underlying process behaviors in each alternative mode are decomposed based on limited batches, revealing different types of relative variations in comparison with the reference mode. They are then modeled differently and online monitored separately. Although starting from limited batches, the proposed method can efficiently identify mode affiliation and detect faults for each mode during online application. Its feasibility and performance are illustrated with a typical multiphase batch process with multiple modes. Considering that it is more common that sufficient batches cannot be guaranteed for every mode, the proposed modeling and analysis algorithm with limited batches is significantly meaningful for multimode and multiphase process monitoring. From another viewpoint, the multimode case with sufficient batches for every mode can be regarded as one extreme case of the concerned problem.

The rest of this paper is organized as follows. The specific of the proposed method is addressed in Section 2. It includes preparation of analysis units for different modes, concurrent phase partition, phase-based between-mode relative analysis for model development and online multimode monitoring. In Illustrations section, its applications to a typical multiphase batch process, injection molding, are presented where four modes are considered, one as the reference mode with sufficient batches while the others as alternative modes with limited batches. The illustration results suggest feasibility and efficacy of the proposed algorithm for online multimode and multiphase monitoring. Finally, conclusions are drawn in the last section.

2. Methodology

As mentioned before, sufficient modeling batches cannot be guaranteed for all modes. Here, except for the reference mode, only limited batches can be collected for each alternative mode. Clearly, statistical models can be readily developed based on sufficient batches for the reference mode and then used for online monitoring of the same mode. However, since these alternative modes have different characteristics from the reference mode, they cannot be successfully described and supervised using the monitoring models developed from the reference mode. Because of model mismatch, alarms will be falsely issued to indicate the normal alternative modes as some faults. Therefore, new monitoring models should be developed for each alternative mode to reflect their different characteristics. However, using conventional methods, the monitoring models cannot be readily developed for each alternative mode only based on limited batches. The proposed methodology tries to analyze the phase nature and derive the monitoring models for different alternative modes starting from insufficient modeling batches. Three important analysis steps are included, preparation of new modeling units, multimode phase analysis, and phase-based model development. They are described in the following subsections.

2.1. Modeling data with limited batches

In each batch run (batch index $i = 1, 2, \ldots, I$), assume that J process variables are measured online at $k = 1, 2, \ldots, K$ time instances throughout the operation cycle, forming each regular batch data set, denoted as $\mathbf{X}(K \times J)$. In the present work, batches are of equal length without special declaration so that the specific process time can be used as an indicator to data preprocessing. There are M modes in all. Here only one mode has sufficient batches, which is chosen as the reference mode, covering statistically enough batch-to-batch variations. A three-way data array is available by collecting data from I batches, $\underline{\mathbf{X}}_r(I_r \times J \times K)$ (subscript r indicates the reference mode). The other modes only have limited batches. They work as the alternative modes and a three-way data array, $\underline{\mathbf{X}}_a(I_a \times J \times K)$ (subscript a indicates alternative mode), is obtained for each of them. As shown in Fig. 1, for the reference mode, time-slice $\mathbf{X}_{r,k}(I_r \times J)$ can well reflect batch-wise variations and process characteristics at each time. In contrast, for alternative modes, the "short" time-slice $\mathbf{X}_{a,k}(I_a \times J)$ fails to reveal batch-wise variations at each time resulting from insufficient batches. Therefore, before statistical analysis, a new data unit should be organized for each alternative mode.