Journal of Process Control 21 (2011) 817–829

Contents lists available at ScienceDirect

Journal of Process Control

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

# Statistical analysis and online monitoring for handling multiphase batch processes with varying durations

Chunhui Zhao, Shengyong Mo, Furong Gao*, Ningyun Lu[1], Yuan Yao

*Department of Chemical and Biomolecular Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong*

ARTICLE INFO

*Article history:*
Received 15 September 2010
Received in revised form 10 February 2011
Accepted 15 April 2011
Available online 6 May 2011

*Keywords:*
Unequal multiphase batches
Uneven-length group
Subspace separation
Common and specific correlations
Variable-unfolding

ABSTRACT

In the present work, statistical analysis and online monitoring is presented for handling uneven-length multiphase batch processes. Firstly, the irregular batches are classified into different uneven-length groups according to the changes of underlying characteristics. Then multi-source measurement data can be dealt with, each corresponding to one operation mode. The basic principle is that over different uneven-length groups, despite the uneven-length operation patterns, there are both similarity and dissimilarity to a certain extent among their underlying correlations. By an adequate decomposition, two different subspaces are separated, modeling the group-common and specific information respectively. Their corresponding confidence regions are constructed by searching similar patterns respectively. Accordingly, the online monitoring system is set up, which can track different types of variations closely. This analysis adds a detailed insight into the inherent nature of uneven-length multiphase batch processes. Its feasibility and performance are illustrated by a typical practical case with uneven cycles.



## 1. Introduction

Batch processes are common in chemical, pharmaceutical, and food industries. Monitoring these batch processes is needed for various reasons such as safety, waste-stream reduction, consistency and quality improvement. Multivariate statistical techniques, such as multiway principal component analysis (MPCA) [1,2] and multiway partial least squares (MPLS) [3] were introduced by Nomikos and Macgregor for batch process modeling and monitoring. Since then, many applications and extensions of statistical batch process monitoring have been reported [4–12]. Generally, a vast amount of historical database on the measurement profiles is needed with completed batch runs that produced on-spec products. Subsequent to data acquisition, multivariate statistical analysis methods can thus be used to empirically model the successful historical operation batches. The variation within this data serves as reference distribution, against which the performance of independent new batches can then be compared. To apply the conventional statistical analysis methods, an implicit assumption is that the batch profiles are already time-scaled so that the operation events are synchronized over batches, resulting in equal durations. However, very often this assumption does not hold. The main reason for this wide range in batch time is related to the changes in operation conditions or control objectives. Take for instance the case when the end time depends on the amount of product produced or a quality constraint which has to be met. Commonly in these cases, data have to be properly aligned or equalized prior. Therefore, the data synchronization itself is a very crucial preprocessing step in the successful application of these multivariate methods.

Work has been carried out into various methods for equalizing batch lengths. In the simplest case, two batches have different lengths but the trajectories overlap in the common time part [13]. If this is the case, the problem can be very easily solved. Provided there are enough long batches, a model can be derived using the information from the long batches, while the absent part of trajectory of the shorter batches is treated as missing data [13]. The opposite is cutting the batches to a minimum length [14], and modeling focusing on this part, in which, however, the important information towards the end of longer batches would have been lost. Therefore, they are suitable only when the uneven-length problem is not serious and the main events have occurred in the common time part. Unfortunately, this is seldom satisfied since in the general case, the trajectories of the variables may have different shapes for most of the common time, showing different operation patterns. Regarding varying batch-to-batch process time, other handling methods [2,9,15] were presented including using rescaled batch time as a maturity index, tracking the batch progress with an indicator variable or using local batch time as the response vector in a PLS model. Nomikos and MacGregor [2] have suggested the use of another measured variable in place of

* Corresponding author. Tel.: +852 23587136; fax: +852 23580054.
*E-mail address:* kefgao@ust.hk (F. Gao).
[1] College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu Province, PR China.

0959-1524/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2011.04.005