Journal of Process Control 21 (2011) 1370–1377

Contents lists available at ScienceDirect

Journal of Process Control

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

Experiences in batch trajectory alignment for pharmaceutical process improvement through multivariate latent variable modelling

Salvador García-Muñoz[a,*], Mark Polizzi[a], Andrew Prpich[a], Cathal Strain[b], Adam Lalonde[b], Vilmary Negron[c]

[a] Pfizer Global Research & Development, Groton, CT 06340, USA
[b] Pfizer Global Manufacturing, Groton, CT 06340, USA
[c] Pfizer Global Manufacturing, Barceloneta, PR, United States

ARTICLE INFO

Article history:
Received 1 December 2010
Received in revised form 26 July 2011
Accepted 26 July 2011
Available online 15 September 2011

Keywords:
Batch process
Multivariate monitoring
Multi-way PCA
Multi-way PLS
Alignment

ABSTRACT

The primary objective of batch data as trajectory alignment (or synchronization) is to standardize the data sampling per batch according to the evolution of the process, and secondarily to homogenize the samples per run. The use of an indicator variable performs both objectives well. Two examples from the pharmaceutical sector are discussed to illustrate the different ways to deal with uneven samples across batches and across variables in the same batch. Since trajectory alignment requires large time investment, a simple triage approach is proposed to assess the need to analyze the dynamics of a given process and hence perform alignment. The presented examples are representative of a broad variety of batch processes that are operated by recipe in the pharmaceutical sector. In our experience, the variables associated with the automation triggers in these recipes are the best indicator variables to use for alignment. This is due to (i) the fact that the trigger variables are easy to identify from the automation of the recipe, (ii) operators are familiar with these, (iii) the target values for triggers are known a priori and hence the resulting alignment scheme can be performed in real-time for monitoring applications and (iv) it makes the monitoring scheme easy to understand and justify around the design-space since the design-space may originally be defined in terms of the trigger variables for each phase of the batch.

1. Introduction

The application of multivariate latent variable models to analyze batch processes has been widely studied and discussed in literature in applications ranging from the analysis and troubleshooting of the process using historical data from the network of sensors installed in the process; process monitoring and fault detection for continuous quality assurance; process operation design and optimization; and control. These techniques have been successfully implemented in industrial settings, with some applications available in the public literature and excellent reviews written on the topic [1,2].

A batch recipe will commonly be executed by a series of instructions that trigger the available actuators based on process targets (temperature, weights, pressured, etc.) and will often have uneven time-length. Due to this uneven time duration across batches a key exercise in performing any statistical analysis of batch data is the need to align (or synchronize) the samples taken throughout the run, for all batches available. This is necessary so that sample *i* (across batches) correspond to the same state of evolution for a given process variable (i.e., temperature from the heating phase for batch *A* should not be contrasted with the temperature during the reactive phase for batch *B*). Nomikos and MacGregor [3] identified the indicator variable approach to synchronize the data. Kassidas et al. [4] later proposed to use Dynamic Time Warping when there was no other observation of the evolution of the batch.

This work presents our experience in dealing with these situations with two examples representative of those in the pharmaceutical sector. We also comment on the expectations of a batch alignment exercise from a practical perspective and finally present a triage method to assess the potential impact of the dynamics of a process onto the final product quality; and hence determine the need to invest the necessary time and effort to align the data, or not.

2. 2D Multi-way methods and process dynamics

Data sampled from a dynamic system will contain samples of variables as they change with time. The data can then be arranged in any number of ways, depending on the model structure to be used (data is only a set of numbers with some contextual

* Corresponding author. Tel.: +1 860 715 05 78.
E-mail address: sal.garcia@pfizer.com (S. García-Muñoz).

0959-1524/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2011.07.013