Computers and Chemical Engineering 33 (2009) 992–1003

Contents lists available at ScienceDirect

Computers and Chemical Engineering

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

Multi-model based real-time final product quality control strategy for batch processes

D. Wang[a], Rajagopalan Srinivasan[a,b,*]

[a] Institute of Chemical Engineering and Sciences, 1 Pesek Road, Jurong Island, Singapore 627833, Singapore
[b] Department of Chemical and Biomolecular Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 117576, Singapore

ARTICLE INFO

Article history:
Received 11 March 2008
Received in revised form 2 October 2008
Accepted 26 October 2008
Available online 13 November 2008

Keywords:
Quality control
PLS
Optimization
Model predictive control
Design of experiment
Batch operations

ABSTRACT

A novel real-time final product quality control strategy for batch operations is presented. Quality control is achieved by periodically predicting the final product quality and adjusting process variables at pre-specified decision points. This data-driven methodology employs multiple models, one for each decision point, to capture the time-varying relationships. These models combine real-time batch information from process variables and initial conditions with information from prior batches. Design of experiments is performed to generate informative data that reveal the relationship between process conditions and the final product quality at various times. Control action is also taken at pre-specified decision points; at these times, the manipulated variable values are calculated by solving an optimal control problem similar to model predictive control. A key benefit of this strategy is that missing data imputation is obviated. The proposed modeling and quality control strategy is illustrated using a batch reaction case study.

1. Introduction

The batch mode of production is common for manufacturing many value-added products such as pharmaceuticals and agro-chemicals. For a specific batch production, the possible synthesis routes have already been investigated by the chemists and a recipe and operation mode selected. During each batch, one usually needs to just follow the pre-specified procedures and establish the pre-scribed process conditions; these will be repeated batch after batch. Batch processes usually suffer a lack of reproducibility from batch to batch due to changes in raw material purities, variations in initial conditions, and disturbances. These changes are inherent in the processes and may be difficult for operators to discern a priori, but could have an adverse effect on the final product quality. Online monitoring of process variables for rapidly detecting abnormalities and taking remedial actions is therefore essential to ameliorate the effects of such changes and to produce on-spec products from each batch.

Statistical process control (SPC) has been used to ensure batch product quality. Numerous SPC approaches have been reported. Of these, the use of principal component analysis (PCA) and partial least squares (PLS) for batch process monitoring has been extensively investigated. In such approaches, the behavior of the process is characterized using a statistical model derived through multi-way analysis of online measurements obtained when the process is in a state of statistical control. Subsequently, future unusual events are detected by projecting the process measurements against this "in-control" model (Doan & Srinivasan, 2007; Kourti, Nomikos, & MacGregor, 1995; Nomikos & MacGregor, 1994; Wise, Gallagher, Butler, White, & Barna, 1999; Wold & Sjostrom, 1998). Through such monitoring, an abnormal batch can be detected online, without waiting for the final quality to be measured at the end of the batch.

Even though online process monitoring can detect abnormality promptly, the separation between normal and abnormal batches in terms of product quality is often ambiguous. Quality variations even among normal batches can be quite significant; some abnormal batches can be rectified by appropriate remedial actions during the batch. This motivates the development of with-in batch recovery schemes for final product quality control.

A number of approaches have been developed to reduce the variation in product quality. One of the most popular approaches adjusts the operating condition of a new batch based on data collected from previous batches in an attempt to bring the new batch's final quality close to the desired target. In such batch-to-batch control strategies, if the end-of-batch quality measurements consistently show a statistically significant deviation from the nominal

* Corresponding author at: Department of Chemical and Biomolecular Engineering, National University of Singapore, 10 Kent Ridge Crescent, Singapore 117576, Singapore. Tel.: +65 65168041; fax: +65 67791936.
E-mail address: chergs@nus.edu.sg (R. Srinivasan).

0098-1354/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2008.10.022

Nomenclature

C_A	concentration of reactant A
C_B	concentration of reactant B
CE	cooling effect
$e_j(\Delta X_i^n)$	error between the current prediction and the target: $e_j(\Delta X_i^n) = \hat{y}_j^i - y_j$; $(j = 1, 2, \ldots, M)$
F_B	flow-rate of reactant B
I	number of batch datasets
J	number of process variables
K	number of samples in a batch
k_i	decision point $(i = 1, 2\ldots)$
P	weight matrix for setpoint tracking, $P = \mathrm{diag}(p_1, p_2, \ldots, p_M)$
Q	weight matrix put for control penalty, $Q = \mathrm{diag}(q_1, q_2, \ldots, q_K)$
Q_C	final (end-of-batch) quality of product C
Q_D	final (end-of-batch) quality of by-product D
SPE	squared prediction error
T^2	Hotelling statistics which is a scaled squared 2-norm of an observation vector from its mean
T_A	temperature of reactant A
T_B	temperature of reactant B
T_{sp}	setpoint of reactor temperature
X	a three-way measurement array of size $(I \times J \times K)$
X^n	partition of X representing the manipulated process inputs
X^m	partition of X representing the other process measurements
X_{k_i}	process measurements collected from start of batch up to decision point k_i
X_i	two-way matrix formed by batch-wise unfolding of X_{k_i}
ΔX_i^n	control adjustment of input variables at decision point k_i, $\Delta X_i^n = X_i^n - X_i^{n,\mathrm{present}}$
Y	a $(I \times M)$ matrix representing end-point product quality, response variable
$\hat{Y}_i$	final (end-of-batch) quality prediction at decision point k_i, $\hat{Y}_i = [\hat{y}_{i,1}, \hat{y}_{i,2}, \ldots, \hat{y}_{i,M}]^T$
Z	initial condition $(I \times L)$ matrix

Greek letters

θ	regression coefficient
θ_i	model parameters at decision point k_i
ξ	optimal experimental settings
Ξ	data matrix of predictor variable
Ξ_i	data available up to decision point k_i
$\Psi(V_\theta)$	measure of the variance-covariance matrix V_θ of parameters to be estimated

case, the operator would identify the cause and adjust the operating conditions for subsequent batches (Crowley, Harrison, & Doyle, 2001; Dong, McAvoy, & Zafiriou, 1996; Edgar et al., 2000; Flores-Cerrillo & MacGregor, 2003; Ott & Schilling, 1990). The common characteristic of these approaches is that the correction is made for a new batch as a whole, thus it is an *offline* quality control strategy.

In another approach, a reference trajectory is recommended for all batches. The online control objective in this scheme is to maintain the operating conditions as per the reference, even in face of disturbances. The reference trajectory is determined *a priori* by optimizing an off-line process model, derived either from first principles or from mining historical data (Clarke-Pringle & MacGregor, 1998; Srinivasan, Bonvin, Visser, & Palanki, 2003; Srinivasan, Palanki, & Bonvin, 2003; Vander Wiel, Tuker, Faltin, & Doganaksoy, 1992). Even though this approach is effective in rejecting some disturbances, it may not always produce on-spec products at the end of the batch, even with perfect tracking control of process variables (Russell, Kesavan, Lee, & Ogunnaike, 1998; Russell, Robertson, Lee, & Ogunnaike, 1998). This is because significant batch-to-batch variations could arise in raw material impurity profiles or process parameters (kinetics, heat transfer, etc.) whose effects are not included in the nominal model. In many cases, the raw material has large stochastic variations originating from prior processing steps and the regulatory model cannot sufficiently capture their subtle effects on final product quality. Hence, the implementation of an off-line calculated trajectory does not guarantee optimal batch performance.

It is possible to express batch quality control as an optimization problem, where uncertainty in parameters and disturbances are considered (Srinivasan et al., 2003). If an accurate process model is not available, a robust optimization strategy can be used to derive process inputs, which once implemented would drive the final quality within specs. This approach usually produces a conservative solution. A measurement-based optimization scheme that tracks the necessary conditions for optimization can both cope with uncertainty and lead to a less conservative optima. However, this relies on appropriate parameterization of input profiles to satisfy the necessary condition of optimality. The central assumptions here are that the set of active constraints is known a priori and the set does not change due to the process uncertainties, i.e., the structure of the optimal solution of the true system is known *a priori*.

Mid-course correction (MCC) strategies are used during a batch's evolution in order to reduce variations in final quality. These strategies recognize that process conditions during the batch will tend to dominate systematic batch-to-batch variations (Flores-Cerrillo & MacGregor, 2003; Kesavan, Lee, Saucedo, & Kishnagopalan, 2000; Russell and Kesavan et al., 1998; Russell and Robertson et al., 1998; Yabuki & MacGregor, 1997). In this approach, online measurements at some midcourse points are used to predict the final product quality. If the predicted quality deviates beyond a statistically defined "in control" zone, a model is used to calculate the control move which would bring the batch back to statistical control. The success of such schemes is mainly dependent on the quality of the model. Usually an inferential model for quality prediction is developed using historical data. For process recovery purpose, the training data needs to contain sufficient input variability and disturbance information to allow proper model identification, i.e., the model identification requires persistency of excitation. Although historical batch information can be used for model development, the control inputs in such data often do not have sufficient excitation since they were selected based on optimality for a specific batch.

Another significant issue while using models for quality prediction and control in MCC approaches is that the online measurements that form the basis for quality prediction are incomplete, i.e., all the data necessary for predicting the end of batch quality becomes available only when the batch has finished. The conundrum from the absence of future data is usually solved in an *ad hoc* fashion, e.g., by using data imputation methods, or assuming a known correlation between the available measurements and future ones. Such imputation inevitably results in additional uncertainty during the prediction and control task, especially at early stages of the batch when limited data are available (Nelson, MacGregor, & Taylor, 2006).

In this article, a data-driven real-time batch quality control strategy, similar to model predictive control, is developed. With a