Journal of Process Control 19 (2009) 520–529

Contents lists available at ScienceDirect

Journal of Process Control

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

Moving average PLS soft sensor for online product quality estimation in an industrial batch polymerization process

Pierantonio Facco [a], Franco Doplicher [b], Fabrizio Bezzo [a], Massimiliano Barolo [a,*]

[a] DIPIC-Dipartimento di Principi e Impianti di Ingegneria Chimica, Università di Padova, Via Marzolo 9, 35131 Padova PD, Italy
[b] Sirca S.p.A. Resins and Coatings, Viale Roma 85, 35010 San Dono di Massanzago PD, Italy

ARTICLE INFO

Article history:
Received 10 July 2007
Received in revised form 28 April 2008
Accepted 10 May 2008

Keywords:
Statistical process control
Partial least squares
Soft sensing
Multivariate quality control

ABSTRACT

This paper considers the development of multivariate statistical soft sensors for the online estimation of product quality in a real-world industrial batch polymerization process. The batches are characterized by uneven length, non-reproducible sequence of processing steps, and scarce number of measurements for the quality indicators with uneven sampling of (and lag on) these variables. It is shown that, for the purpose of quality estimation, the complex series of operating steps characterizing a batch can be simplified to a sequence of three estimation phases. The switching from one phase to the other one can be triggered by easily detectable events occurring in the batch. For each estimation phase, PLS software sensors are designed, and their performance is evaluated against plant data. The estimation accuracy can be substantially improved if some form of dynamic information is included into the models, either by augmenting the process data matrix with lagged measurements, or by averaging the process measurements values on a moving window of fixed length. In particular, the moving average three-phase PLS estimator shows the best overall performance, providing accurate estimations also during estimation Phase 2, which is characterized by a very large variability between batches.

1. Introduction

Batch processing is used to manufacture high added-value goods, such as specialty chemicals and biochemicals, materials for microelectronics, and pharmaceuticals. With respect to their continuous counterparts, batch processes are easier to set up and require only limited fundamental knowledge of the underlying process mechanisms. One key feature of batch processes is that, by properly adjusting the operating recipe, they can achieve a consistently high and reproducible quality of the product, in spite of changes in the raw materials and in the state of the equipment or of the utilities.

In principle, the operation of a batch process is easy, because the processing usually evolves through a "recipe", i.e. a series of elementary steps (e.g.: charge; mix; heat-up/cool; react; separate; discharge) that can be easily carried out even without supervision, if the production facility is outfitted with a fairly large degree of automation. However, it is often the case that batch plants are poorly automated, and may require intervention by the operating personnel to provide online adjustments of the operating recipe in order to avoid the production of off-spec products. In fact, with respect to product quality control, most batch processes are run in an open-loop fashion, because information about product quality is not available online, but is obtained offline from laboratory assays of few product samples. Because of the lack of real time information on the product quality, it may be difficult to promptly detect quality shifts and to counteract them by adjusting the operating recipe accordingly. Therefore, a quality control strategy for a batch process often reduces to the online control of the trajectories of some key process variables (which can be measured online), and possibly to some midcourse intervention on the operating recipe to compensate for the shifts detected in the product quality measured offline.

The performance of a batch process could be improved if accurate and frequent information on the product quality were available. Software sensors (also called virtual sensors or inferential estimators) are powerful tools for this task. They are able to reconstruct online the estimate of "primary" quality variables from the measurements of some "secondary" process variables (typically, temperatures, flow rates, pressures, valve openings), by using a model to relate the secondary variables to the primary ones.

This work considers the design a soft sensor for the online estimation of the quality of a resin produced in a real-world industrial batch polymerization process. The resin quality is determined by the values of two chemical properties, i.e. the resin acidity number (N_A) and the resin viscosity (μ). The process is monitored online through a fairly large number of process measurements. However, these measurements are noisy, auto-correlated and cross-correlated. Quality measurements are only available offline, and they

* Corresponding author. Tel.: +39 049 827 5473; fax: +39 049 827 5461.
E-mail address: max.barolo@unipd.it (M. Barolo).

0959-1524/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jprocont.2008.05.002

Nomenclature

A	total number of latent variables
ΔK	lag on the process variables in the TP–PLS model
$\Delta K'$	length of the moving window in the MATP–PLS model
h	sampling instant of the quality variables
H_i	total number of quality samples for the batch i
i	generic batch of the reference dataset
I	total number of batches in the reference dataset
j	generic process variable
J	total number of process variables
k	sampling instant for the process variables
K_i	total number of process variables samples of batch i
LTP–PLS	lagged three-phase PLS
LV	latent variable
LV1	first latent variable
LV2	second latent variable
m	quality variable
MATP–PLS	moving average three-phase PLS
$MRPE_{i,m}$	mean relative prediction error for quality variable m in batch i during a single estimation phase
MVS	multivariate statistical methods
M	total number of quality variables
N	number of samples of the matrix X
N_A	acidity number
PLS	projection on latent structure method or partial least squares
Q	squared prediction error
T^2	Hotelling statistic
TP–PLS	three-phase PLS method
VIP	variable importance in the projection methods
$\mathbf{X}$	input matrix of the process variables
$\underline{\mathbf{X}}$	three-way matrix of the process variables
$\mathbf{X}_i$	matrix of the time trajectories of all the process variables for batch i
$\bar{\mathbf{X}}_i$	process variables' moving averages matrix for batch i
$\mathbf{X}^{\text{delayed}}$	matrix of delayed variables
$\mathbf{X}^{\text{lagged}}$	input matrix for the LTP–PLS model
$\mathbf{X}_i^{\text{delayed}}$	matrix of the delayed variables in batch i
$\mathbf{X}^{\text{unfolded}}$	bi-dimensional matrix obtained by variable-wise unfolding the $\underline{\mathbf{X}}$ matrix
$\bar{X}^{\text{unfolded}}$	bi-dimensional matrix of the process variables' moving averages for the MATPPLS model
$\mathbf{x}_{i,j}$	jth variable time profile in batch i in form of column array of the $\mathbf{X}_i$ matrix
$X_{i,j}^{\Delta K}$	time profile of the variable j in batch i delayed of ΔK samples, column vector of the $\mathbf{X}_{\text{delayed}}$ matrix
$\mathbf{x}_{j,j,k}$	element of the $\underline{\mathbf{X}}$ matrix
$\bar{\mathbf{x}}_{i,j,k}$	moving average of the variable j on batch i in the kth time instant, element of the $\bar{\mathbf{X}}_i$ matrix
$\mathbf{Y}$	matrix of the quality variables
$\underline{\mathbf{Y}}$	three-dimensional matrix of the quality variables
$y_{\text{i,m,h}}$	element of the $\underline{\mathbf{Y}}$ matrix
$\hat{y}_{i,m,h}$	estimated value of $y_{i,m,h}$
$\mathbf{Y}^{\text{unfolded}}$	bi-dimensional matrix obtained by variable-wise unfolding the $\underline{\mathbf{Y}}$ matrix

Greek symbol

μ	viscosity

are scarce, delayed and unevenly spaced in time. The operating procedure for a batch evolves through a nominal recipe, which is subject to several online adjustments made by the plant personnel depending on the actual evolution of the batch, as it is monitored by the quality measurements. The batch length exhibits a large variability. All of these features make each batch hardly reproducible, and the online quality estimation a challenge.

It is well known that developing a first-principles model to accurately describe the chemistry, mixing and heat-transfer phenomena occurring in a polymerization process requires a very significant effort. Several designed experiments may be needed to identify the most representative set of equations and all the related parameters. Furthermore, if the plant is a multi-purpose one (as is often the case in batch polymer processing), this effort must be replicated for all the products obtained in the same facility. Finally, the resulting first-principles soft sensor may be computationally very demanding for online use.

Data-driven soft sensors overcome these difficulties, and will be considered in this paper. This class of inferential estimators does not require to develop extra information on the process in terms of mechanistic equations or values assigned to physical parameters. Rather, it extracts and exploits the information already embedded in the process data as these data become available in real time from the measurement sensors. In the last two decades, multivariate statistical (MVS) techniques have proved to be excellent tools for the analysis and monitoring of processes where lots of process data are available [12]. These techniques are able to compress the information contained in the available data down to a low-dimensional space that retains almost all of the information originally embedded in the data. Within this space, it is possible to obtain a relationship between the (transformed) process data and the (transformed) quality data so as to design computationally inexpensive online estimators.

In this paper we exploit one of these projection methods (namely, partial least squares, PLS, regression; [8]) to design a soft sensor for the online estimation of the resin quality properties. Several studies about the online estimation of product quality through MVS techniques are available for continuous polymerization processes (for example [22,15,23]. On the other hand, examples for batch processes are mostly related to the use of these techniques for batch classification, or for the prediction of the end-point product quality only, or are limited to simulation studies [19,18,21,4,24,14,3]. Very few papers present industrial applications of MVS software sensors for the real time prediction of the product quality for industrial batch polymerization processes [16]. The additional complexity of the case discussed in this paper is given by the fact that there is no easily detectable "standard" behavior in the batches, and that most operations are performed manually by the operators and therefore are hardly reproducible.

In the following, after a description of the process, we consider the design of three alternative software sensors, with the aim to show how the performance of the estimators can be successively improved by separately taking into account the nonlinear nature of the process and its dynamic behavior.

2. The industrial process

The industrial process under study is the production of a polyester resin used in the manufacturing of coatings via batch polycondensation between a diol and a long-chain dicarboxylic acid. The reaction is carried out in a stirred tank reactor having the nominal capacity of 12 m^3, heated up through a dowtherm oil through an external coil. Several other resins are produced in the same reactor in different production campaigns.

Besides the desired product, the poly-condensation reaction leads to the formation of water, which must be removed from the reaction environment to promote the forward reaction. To allow for the removal of water, the plant is equipped with a packed distillation column (which is run in dry mode for the production of the resin under study), an external water-cooled condenser, and a scrubber. A vacuum pump allows to operate the plant under vacuum when needed.

The plant is equipped with several online measurement sensors. The values of thirty-four variables are routinely collected online and recorded by a process computer every 30 s. Typically, these variables include process measurements (temperatures, pressures, valve openings) and controller setpoints (which are adjusted manually by the operators); 4500–7500 recordings are collected for each process measurement during a batch. For the purpose of this study, the online recordings of twenty-three process variables were considered. These variables were chosen on the basis of a preliminary engineering analysis.

Product quality measurements (N_A and μ) are not available online. Product samples are taken manually, quite infrequently and unevenly (i.e. one sample every 1.5–2 h, depending on the operators' availability and on the actual evolution of the batch), and are sent to the laboratory for analysis. The analysis takes ~20 min to be completed, and the accuracy of the laboratory assay is ~10% of the reading. The quality measurements are not available for the entire duration of the batch; in fact, the product sampling begins 8–10 h after the batch starts (i.e. after at least 1000 time instants have elapsed). As a result, only 15–20 measurements for each quality index are typically available during a batch.

Each batch is run through a sequence of operating steps, most of which are triggered manually by the operators. The switching from one operating step to the subsequent one is determined by the current values of the resin viscosity and acidity number. A typical sequence of the operating steps runs as follows. Cleaning of the equipment and lines is done when a different resin has been produced in the preceding batch. Then, the reactants, additives and catalyst are loaded into the reactor. The charge of liquid diol is automated, while the acid is charged manually as a solid. Being the dicarboxylic acid a product of fermentation, its quality may vary markedly from batch to batch; minor changes may be experienced in the quality of fresh diol. However, the feed quality cannot be measured before the batch starts. During the reactor loading, the mixing and heating systems are switched on, and heat-up continues until the reactor reaches the setpoint temperature (202 °C).

The dowtherm oil temperature may vary from batch to batch, because this oil serves as a hot utility for other reactors in the same production facility and the duty of the heating furnace is limited; this results in different durations of the heat-up period from one batch to another one.

The raising temperature in the reactor activates the poly-condensation reaction; hence, water is produced and must be removed to improve the yield of resin. Water is generated as a vapor phase that leaves the reactor. In the early stages of the batch, this vapor phase contains significant amounts of diol, which must therefore be recovered and recycled for further processing. Therefore, the vapor phase leaving the reactor is sequentially processed in the following ways: (i) by differential condensation in the packed column, in such a way as to recover liquid diol and recycle it back to the reactor; (ii) by total condensation in the condenser; (iii) by washing and contact condensation in the scrubber.

Vacuum needs to be applied during the course of a batch to adjust the viscosity and the molecular weight distribution of the resin. Furthermore, to ensure that the final product quality is on specification, the operating recipe always requires at least two additions of fresh raw materials and catalyst during the course of a batch. The first addition is made before vacuum is applied for the first time. Therefore, when fresh materials and catalyst are charged into the reactor again, vacuum must be broken and then resumed.

When N_A and μ fail to approach the target values in the expected amount of time, further amounts of fresh material are charged. Following the operators' jargon, these supplementary additions are known as "corrections" to a batch. Corrections are the way the operators act online to compensate for any unmeasured disturbance affecting a batch, and more than one third of the batches undergoes to corrections.

When the end of the batch is approaching, the reactor temperature is increased to 220–230 °C. The batch is stopped and the product is discharged when the resin reaches the desired quality targets in terms of both N_A and μ.

The net result of this quite complex (and mostly manually driven) operating recipe is that, although the end-point quality of the resin usually falls within a very narrow range, the "internal" variability of the batches is very large. Indeed, there are several sources of variability within a batch, most of which cannot be eliminated:

- At the batch start, the equipment may be "hot" or "cold".
- The amount and the quality of the raw materials and additives may vary from batch to batch (because of errors in weighting, different level of impurities in the raw materials and additives, partial activity loss of the catalyst).
- The heating loop performance depends on the numbers of the batches being run simultaneously in the plant.
- The number of the midcourse corrections is subject to the available number of (and delay on) the quality measurements and to the operators' judgment and availability, all of which are not predictable.
- The set points are manipulated manually by the operators, and are therefore subject to the experience and confidence of each operator.

Most of this variability reflects itself in the trajectories of the process measurements, and eventually in the total batch duration [7]. As an example, the total batch duration in the set of batches considered in this study ranges between 40 and 70 h.

As was noted, the switching from one operating step to the subsequent one is triggered by the measured value of the resin viscosity and acidity number. However, because quality measurements are available quite infrequently, the switching may be substantially delayed, with the result of poor monitoring of the product quality and increase of the duration of a batch. Therefore, as a first approach to the design of a system for the online monitoring of the whole production process, the design of a soft sensor for the estimation of μ and N_A is considered, with the objective to make available online frequent and accurate estimations of the product quality indicators.

3. Quality estimation using PLS regression

The quality monitoring approach we have developed relies on the PLS regression technique [8,25]. For the batch process under study, the available dataset includes measurements of the process variables and of the quality variables from 33 batches (16 months of operating effort). This dataset was split into two subsets: 27 batches constitute the reference (i.e. calibration) dataset, while the remaining 6 batches represent the validation dataset.[1] The reference process data are collected into a three-way matrix $\mathbf{X}$ ($I \times J \times K_i$). Each of the $j = 23$ columns of this matrix contains one

[1] The whole dataset can be made available to the interested reader by requesting it to the corresponding author.

measured process variable, while each row corresponds to one of the $I = 27$ reference batches; time occupies the third dimension, and K_i is the total number of recordings taken for each of the J process measurements during batch i. As was already mentioned, the duration of the generic batch i is not fixed, and this makes K_i change from batch to batch. The generic element of matrix $\mathbf{X}$ is denoted with the symbol $x_{i,j,k}$.

The arrangement of the three-way $\mathbf{Y}$ ($I \times M \times H_i$) matrix is similar; however, only $M = 2$ columns are present, which correspond to the two quality variables to be estimated (μ and N_A). The third dimension of $\mathbf{Y}$ is scanned unevenly and with a much lower frequency than the one of the $\mathbf{X}$ matrix (i.e. $H_i << K_i$, where H_i corresponds the total number of midcourse quality measurements in batch i). To remove the time delay due to sample transportation and analysis, "time" in the $\mathbf{Y}$ matrix is set to correspond to the time when a sample is taken from the reactor, not to the time when the quality analysis is made available from the lab.

Since the sequence of operating steps is quite complex and almost non-reproducible, no attempts were made to synchronize the batch time evolution. The three-way matrices were therefore unfolded according to the variable-wise technique [27]. The adopted technique transforms the three dimensional process data matrix into a bi-dimensional array:

$$\overline{\mathbf{X}}_{\text{unfolded}} = \begin{bmatrix} \overline{\mathbf{X}}_1 \\ \overline{\mathbf{X}}_2 \\ \vdots \\ \overline{\mathbf{X}}_i \\ \vdots \\ \overline{\mathbf{X}}_1 \end{bmatrix} \tag{1}$$

where

$$\mathbf{X}_i = \begin{bmatrix} x_{i,1,1000} & x_{i,2,1000} & \cdots & x_{i,J,1000} \\ x_{i,1,1000} & x_{i,2,1001} & \cdots & x_{i,J,1001} \\ \vdots & \vdots & & \vdots \\ x_{i,1,k_i} & x_{i,2,k} & \ddots & x_{i,J,k} \\ \vdots & \vdots & & \vdots \\ x_{i,1,K_i} & x_{i,2,K_i} & \cdots & x_{i,J,K_i} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{i,1} & \mathbf{X}_{i,2} & \cdots & \mathbf{X}_{i,J} \end{bmatrix}. \tag{2}$$

Note that, since the quality data are available only 8–10 h after the beginning of the batch, the first 999 process measurements had to be discarded when a static estimator was designed. Therefore, for such an estimator, $\mathbf{x}_{i,j}$ in (2) represents the vector of measurements of process variable j in batch i, starting from time instant #1000. On the other hand, as will be clarified later, dynamic estimators can make use of "early" process measurements, taken when quality assessment has not been started yet.

Because, for a given batch i, $K_i >> H_i$ results, when static estimators were designed the $\mathbf{X}_i$ matrices were pruned in such a way as to eliminate all the rows that do not correspond to a time instant where a quality measurement is available. Note, however, that this pruning is needed only during the PLS model calibration phase. When the model is built, it can be interrogated any time a process measurement is available, regardless of the fact that a quality measurement is available or not, thus obtaining N_A and μ estimates at the same frequency as the process measurements.

3.1. Single-phase PLS model

As discussed previously, the operating recipe for the production of the resin results in a complex series of operations, most of which are subject to the operators' manual intervention. Therefore, also owing to the intrinsic nonlinear nature of the process, it is quite unlikely that the correlation structure between the variables remains the same during the whole duration of a batch. In turn, this means that a single linear PLS model may not be able to provide an accurate prediction of the quality variables along the whole duration of a batch. To check this conjecture and provide a term for comparison, a single PLS model on 5 LVs for the estimation of μ and N_A was built from the reference dataset as a first attempt. The number of latent variables was chosen in such a way as to minimize the estimation error in the validation dataset. Typical validation results are reported in Fig. 1, where the acidity number and the viscosity predicted by this model are compared to the actual values.

Although the frequency at which the quality estimations are made available is much higher than the frequency of the lab measurements (which can improve the monitoring of the process), the estimation accuracy is not satisfactory. The use of nonlinear transformations on the process variables or of a nonlinear inner relationship in the PLS algorithm did not improve the results significantly. More sophisticated nonlinear methods (like those proposed, for example, by Kosanovic and Piovoso [10], and by Maulud et al. [17]) were not explored.

3.2. Multi-phase PLS model

An approach to overcome the nonlinearity problems (i.e. a changing correlation structure among the variables) is to divide a batch in different phases, and to develop a linear PLS submodel for each of these phases [11,28]. In this case, a criterion also needs to be found to detect online a phase change so as to dictate the switching between one submodel and the subsequent one.

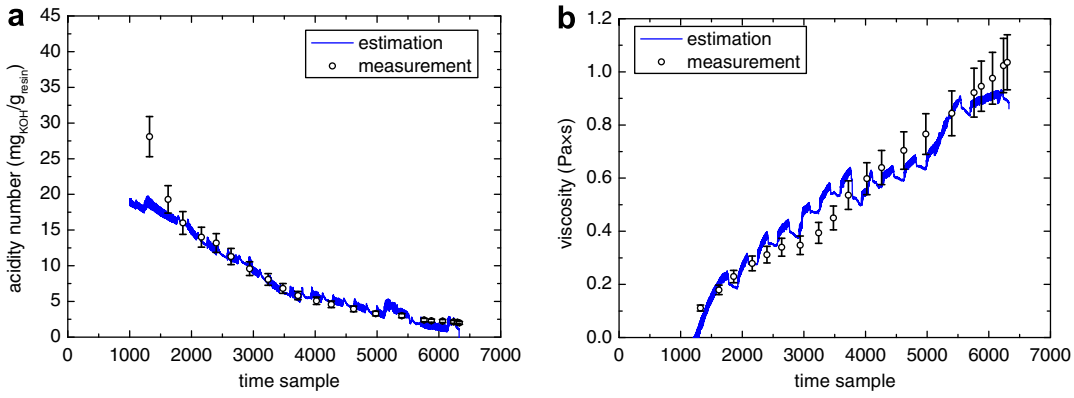

Fig. 1. Prediction of (a) the acidity number and (b) the viscosity for validation batch #4 using a single PLS model. The vertical bars represent the laboratory assay accuracy.

In the process under study, building a different PLS model for each operating step is not a viable solution, not only because the number of operating steps is large, but also because too few quality measurements are usually available in a single operating step to build the relevant PLS submodel. Furthermore, the actual number of operating steps in a batch is not known *a priori*, being dependent on the number of corrections that the batch will be subject to. An alternative approach is to check whether different operating steps in a batch share the same correlation structure among the measurements. If this is the case, the same PLS submodel can be used to represent these operating steps. These "shared" operating steps constitute the same estimation phase. Following this approach, we may end up with a number of estimation phases that is (possibly much) lower than the number of the operating steps.

To detect how many estimation phases can be recognized in the reference batches, the scores on the first two LVs can be plotted one against the other when a single PLS model is built from the reference dataset. In Fig. 2a, each point represents the batch state in a certain instant of time for each of the reference batches.

It can be seen that the score points are mainly clustered into three distinct regions of the score plane. A closer inspection of the score points related each single batch (Fig. 2b) revealed that all batches are characterized by a similar pattern in the "movement" of the score points: at the beginning of the operation, a score point is located at the left of the score plane ("Phase 1" cluster), then it moves to the center of the plane ("Phase 2" cluster) as time progresses, and finally it shifts to the plane right ("Phase 3" cluster) towards the end of the batch. The correlation structure between variables is more similar for points within a cluster than for points between clusters. Otherwise stated, each cluster represents an estimation phase, and can be envisioned as a series of operating steps that maintain the same correlation structure among the variables. Therefore, one distinct PLS submodel can be developed for each estimation phase to predict the quality variables from the process ones within that phase. The resulting quality estimator is called a three-phase PLS (TP–PLS) estimator. Note that clusters could also be identified without using process knowledge. To this purpose, *k*-means clustering based on PCA and PLS can be an effective way to obtain automatic cluster detection [14]; in the presence of auto-correlated and cyclic process data, like the ones encountered in the process under study, the clustering algorithm proposed by Beaver et al. [2] can also prove useful. Note, however, that the number of clusters must be kept as small as possible because if too few quality measurements are available within a cluster, it may be impossible to design the relevant PLS submodel.

A key issue in the development of such a multi-phase estimator is finding a proper criterion to switch from one PLS submodel to the subsequent one [3,14]. Switching from one submodel to the subsequent one means being able to recognize in real time that the correlation structure of the data is changing. It was observed that, due to the large inter-batch variability, "time" is not a good indicator to assess phase switching in the process under study. Therefore, submodel switching was linked not to time, but to events: there are certain events that do occur in all batches and mark a change in the correlation structure, although they occur at a different time from batch to batch. Analysis of the process and quality data for all the reference batches revealed that the switching from Phase 1 to Phase 2 occurs the first time vacuum is applied to the reactor, while Phase 3 begins as soon as the final rise of temperature takes place. Following this approach, not only does the number of submodels to be developed remain sufficiently low, but clearly detectable events can also be recognized during a batch to trigger the switching between submodels.

It should be stressed that each submodel is representative only of the phase it refers to. Fig. 3 clarifies this issue with respect to Phase 2 submodel. The scores plot of Fig. 3a refers to a typical validation batch, and shows the similarity of each sample to the Phase 2 samples of the reference set of batches. Only during Phase 2 do the validation scores fall within the 95% confidence ellipse of the Phase 2 submodel. When the process measurements start to be recorded (Phase 1), the score points all lie well outside the confidence ellipse; they enter into, and stay within, the ellipse during Phase 2, while during Phase 3 they tend to move again outside the ellipse to a different region of the score plot. The squared prediction error plot in Fig. 3b shows that indeed Phase 2 submodel is not reliable as a quality estimator during Phase 1 or Phase 3. Therefore, care must be taken to identify the proper switching criterion and detect it online; anticipated or delayed detection may lead the soft sensor to provide unreliable quality estimates.

The typical estimation performance of the TP–PLS soft sensor on 5 LVs for every Phase in a validation batch is shown in Fig. 4. It can be seen that the performance is greatly improved with respect to that of the single PLS model. The estimation accuracy is generally within the accuracy of the laboratory analysis. Yet, some noise in the estimation is present (for example, in the estimation of the acidity number during Phase 1, and in the estimation of viscosity during Phase 3). Furthermore, the viscosity estimation displays a somewhat erratic behavior during Phase 2. In the next section, two different approaches will be considered to further improve the estimation performance.

4. Including time information to improve the estimation performance

The variable-wise unfolding of the three-way **X** and **Y** matrices [27] has the advantage of being very simple to carry out, because it

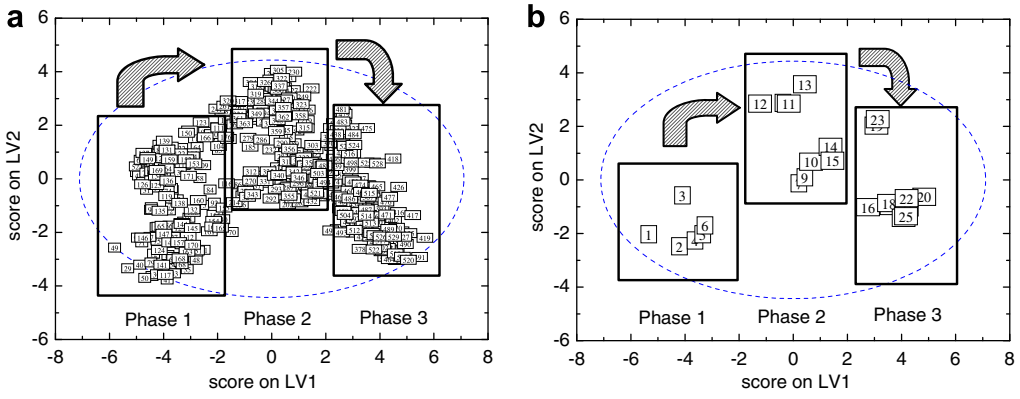

Fig. 2. Scores plot on the first and second latent variables for a single PLS model: (a) whole reference dataset and (b) validation batch #3. The 95% confidence limits are indicated with dashed lines. The rectangular boxes indicate the approximate boundaries of the estimation phase regions.

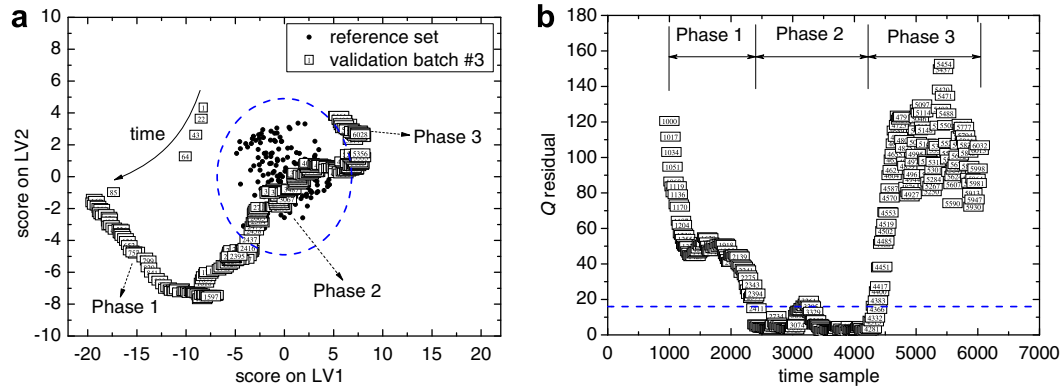

Fig. 3. TP–PLS estimator: reliability of the Phase 2 submodel in the estimation of viscosity for validation batch #3. (a) Scores plot for the first two latent variables, (b) squared prediction error plot. The dashed lines indicate 95% confidence limits. To improve the readability, most of the samples have not been plotted.

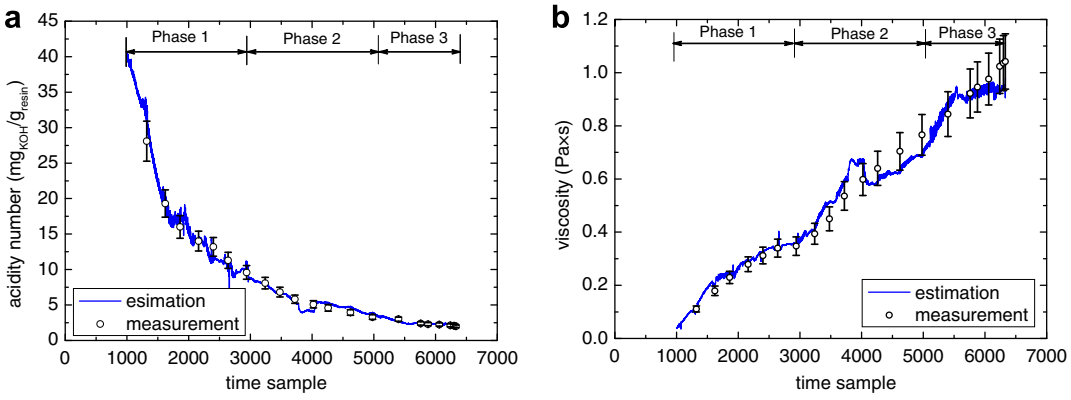

Fig. 4. Prediction of (a) the acidity number and (b) the viscosity for validation batch #4 using the TP-PLS model. The vertical bars represent the laboratory assay accuracy.

can be applied in a straightforward way to sets of batches which have different time duration, without the need of synchronizing the batch length. The price to pay for this simplicity is that the "time footprint" of the data is lost, because the order in which the rows of the two-way **X** and **Y** matrices are assembled following a variable-wise unfolding is unimportant for the design of a PLS estimator. Otherwise stated, the TP–PLS model is inherently static, and this may affect the estimation performance given the fact that a batch process in inherently dynamic. To account for the process dynamics, two different techniques were evaluated, namely the augmentation of the process data matrix with lagged values, and the use of averaged values instead of point values in the **X** matrix.

4.1. Process data matrix augmentation with lagged measurements

A PLS model on variable-wise unfolded data is inherently static. To take into account the dynamic behavior of a batch process, the use of dynamic PLS models has been suggested [13,4,23]. By following this approach, the process data matrix is augmented with lagged values of the process variables at the past sampling instants. To keep reasonably small the column dimension of the $\mathbf{X}_{\text{unfolded}}$ process data matrix, lagged values of only the three most important variables, as identified by the VIP method [5], were considered in three past time instants. These variables are the column top temperature, the column bottom temperature, and the reactor temperature (i.e. variables #7, 10, and 21). By trial and error, it was found that a good performance of the estimator could be obtained by considering the current measurement value plus the values lagged by 1 h (120 time instants), 3 h (360 time instants) and 5 h (600 time instants) for the selected process variables. Information on the most proper values for the lags can be obtained also by studying the autocorrelation and cross-correlation structure of the process and quality measurements. By using this approach, it was found that, depending on the process variable and on the estimation phase, the most appropriate lags range from 300 to 900 time instants, which is consistent with the values we considered in our simplified approach. Note that the variety of lags existing for different process variables is an indication of the variety of time scales that may exist in the process dynamics. In order to account for the effect of different time scales, a more rigorous approach could be taken using a multi-resolution PLS approach [1]. However, this was found to be unnecessary in the present application.

The reference process data matrix was therefore augmented by including $3 \times 3 = 9$ additional columns:

$$\mathbf{X}^{\text{lagged}} = \left[\mathbf{X}^{\text{unfolded}} \quad \mathbf{X}^{\text{unfolded}}\right], \tag{3}$$

where

$$\mathbf{X}^{\text{delayed}} = \begin{bmatrix} \mathbf{X}_1^{\text{delayed}} \\ \mathbf{X}_2^{\text{delayed}} \\ \vdots \\ \mathbf{X}_i^{\text{delayed}} \\ \vdots \\ \mathbf{X}_I^{\text{delayed}} \end{bmatrix} \tag{4}$$

and

$$\mathbf{X}_i^{\text{delayed}} = \left[\mathbf{X}_{i,7}^{-120} \;\; \mathbf{X}_{i,7}^{-360} \;\; \mathbf{X}_{i,7}^{-600} \;\; \mathbf{X}_{i,10}^{-120} \;\; \mathbf{X}_{i,10}^{-360} \;\; \mathbf{X}_{i,10}^{-600} \;\; \mathbf{X}_{i,21}^{-120} \;\; \mathbf{X}_{i,21}^{-360} \;\; \mathbf{X}_{i,21}^{-600}\right] \tag{5}$$

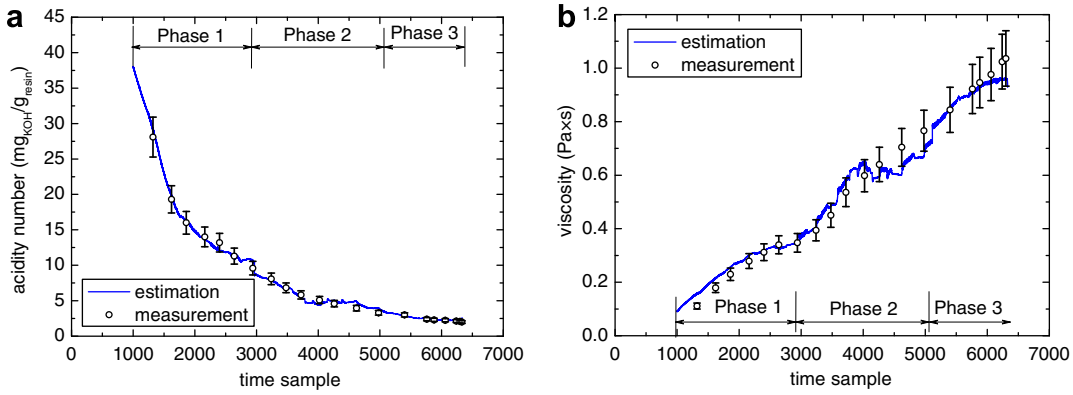

Fig. 5. Prediction of (a) the acidity number and (b) the viscosity for validation batch #4 using the LTP–PLS model. The vertical bars represent the laboratory assay accuracy.

This approach introduces process variables that are more collinear with the quality ones. As a result, the variability in $\mathbf{X}^{\text{lagged}}$ is more representative of the variability in $\mathbf{Y}^{\text{unfolded}}$, and the estimation capability is improved.

It was verified that, also in this case, each batch can be segmented into three estimation phases. The resulting soft sensor is called a lagged three-phase PLS (LTP–PLS) estimator, in which 5, 4 and 3 LVs were chosen for Phase 1, 2 and 3, respectively. The reduction of the number of LVs in Phase 2 and 3 prevents overfitting problems, particularly when the signal-to-noise ratio is low.

Fig. 5 shows the estimation results for this model on a typical validation batch. It can be seen that including information about the batch dynamics, through the use of lagged measurements, suppresses most of the noise that was apparent in the TP–PLS estimations. A slight improvement is also obtained in the accuracy of the viscosity estimation during Phase 2, although the estimated values of this quality indicator in this phase still seem to suffer from some inaccuracy. It should be noted, however, that quality estimation during Phase 2 is inherently difficult because all the corrections to the operating recipe take place during Phase 2, which is therefore subject to a much larger inter-batch variability than the other phases.

4.2. Use of averaged data in the process data matrix

An alternative way to account, although indirectly, for dynamics in the process data is to build the process data matrix with averaged values of the measurements, instead than with current process measurement values. Modifications to the standard PLS algorithm that consider moving windows on weighted past process measurement values have already been proposed for use within recursive and adaptive process control strategies [26,6,21,20]. However, we take a different approach here. The PLS algorithm itself is not altered; what is altered instead is the process measurement matrix: each entry in a column of the process data matrix $\bar{\mathbf{X}}^{\text{unfolded}}$ represents the average value $\bar{x}_{i,j,k}$ of the relevant process measurements, in a certain batch, within a window including the previous $\Delta K'$ time samples. Namely, the value included in the process data matrix at any time instant is the average of the last $\Delta K'$ samples:

$$\bar{\mathbf{X}}^{\text{unfolded}} = \begin{bmatrix} \bar{\mathbf{X}}_1 \\ \bar{\mathbf{X}}_2 \\ \vdots \\ \bar{\mathbf{X}}_i \\ \vdots \\ \bar{\mathbf{X}}_1 \end{bmatrix} \tag{6}$$

where

$$\bar{\mathbf{X}}_i = \begin{bmatrix} \bar{x}_{i,1,1000} & \bar{x}_{i,2,1000} & \cdots & \bar{x}_{i,J,1000} \\ \bar{x}_{i,1,1001} & \bar{x}_{i,2,1001} & \cdots & \bar{x}_{i,J,1001} \\ \vdots & \vdots & \ddots & \vdots \\ \bar{x}_{i,1,K_i} & \bar{x}_{i,2,K_i} & \cdots & \bar{x}_{i,J,K_i} \end{bmatrix} \tag{7}$$

and

$$\bar{X}_{i,j,k} = \frac{\sum_{r=k}^{k-\Delta K'} x_{i,j,r}}{\Delta K'}. \tag{8}$$

The moving averages are used not only to dampen measurement noise (as done for example by Kamohara et al. [9]), but also to smooth out short-term fluctuations (process noise) while, at the same time, preserving the capability to highlight long-term trends. Smoothing process noise was necessary because when a correction takes place in a batch most process variables change abruptly (e.g. when vacuum is broken, most process variables undergo a step-wise change), while bulk properties (like μ and N_A) are practically insensitive to such abrupt changes. Therefore, the performance of a purely static estimator may be disrupted when these events occur. A similar effect is found when, due to poor controller tuning, some process variables tend to cycle (typically, the reactor temperature), while at the same time this cycling does not affect the product quality properties.

The length of the moving window was set by trial-and-error to 900 time samples (7.5 h). The wide extension of the time window also allows to incorporate the variability within most of the first part of the batch, when no quality measurements are taken. The resulting estimator is a moving average three-phase PLS (MATP–PLS) soft sensor built on 5 LVs for every Phase, and its performance is illustrated in Fig. 6. As expected, the estimated profiles of the quality variables are smoother than with the other models. An improvement is apparent in the accuracy of the estimated viscosity profile during Phase 2.

Table 1 allows for a quantitative comparison of the three three-phase estimators that were designed: the static one, the "lagged" dynamic one, and the "averaged" dynamic one.

It is clear that including some form of time information into the **X** matrix greatly increases the amount of variance that can be explained on the quality data, and using averaged measurements (MATP–PLS estimator) appears to be better than using lagged measurements (LTP–PLS estimator). Note that the amount of variance captured in the **Y** matrix during Phase 3 is relatively small for all estimators and both quality variables. This is due to the fact that, when the end of the batch is approaching, the process measurements profiles flatten considerably and the signal-to-noise ratio decreases, making the process variables much less effective predictors of the quality variables. It is also interesting to note that in the

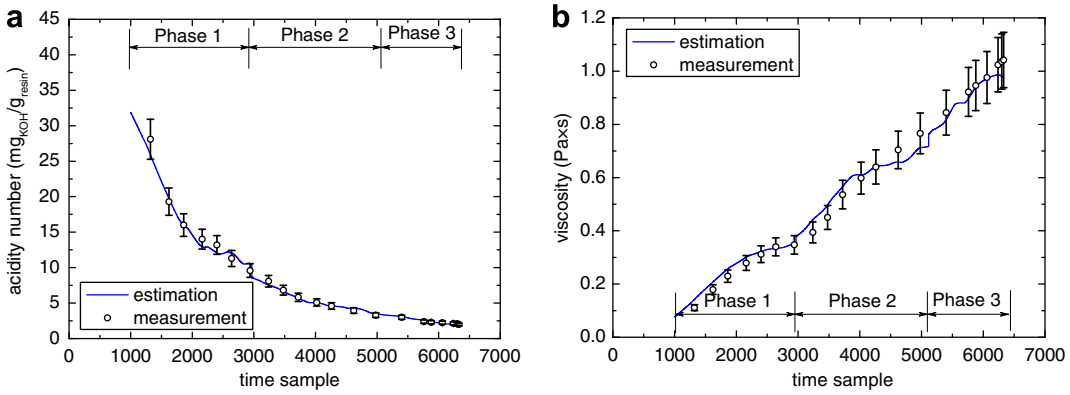

Fig. 6. Prediction of (a) the acidity number and (b) the viscosity for validation batch #4 using the MATP–PLS model. The vertical bars represent the laboratory assay accuracy.

Table 1
Explained variance of the TP–PLS, LTP–PLS and MATP–PLS estimators on the process and quality variables for both acidity number and viscosity (calibration dataset)

Phase	TP–PLS mode				LTP–PLS model				MATP–PLS model			
	N_A estimation		μ estimation		N_A estimation		μ estimation		N_A estimation		μ estimation	
	on X (%)	on Y (%)	on X (%)	on Y (%)	on X (%)	on Y (%)	on X (%)	on Y (%)	on X (%)	on Y (%)	on X (%)	on Y (%)
1	62.00	88.50	63.74	86.57	66.12	96.47	67.47	93.14	70.77	95.56	71.74	94.52
2	67.38	82.97	67.47	78.46	57.11	88.92	57.14	83.66	67.42	91.13	68.63	85.04
3	73.78	59.84	72.56	52.93	61.68	67.47	61.01	55.59	74.21	72.26	75.57	61.90

LTP–PLS estimator the variance captured in the **Y** matrix increases considerably in Phases 2 and 3 with respect to the TP–PLS model, despite a smaller number of retained LVs and a larger number of process variables that causes the captured variance of the **X** matrix to decrease. This indicates that the lagged measurements do bring "new" valuable information for the prediction of quality, and this information was not present in the original **X** matrix; although this new information contained in the "lagged" **X** matrix cannot be captured to a large extent, it is nevertheless much more predictive of the quality matrix.

During each phase of the generic validation batch i, the estimation accuracy on the quality variable m can be evaluated in terms of mean relative prediction error $\mathrm{MRPE}_{i,m}$:

$$\mathrm{MRPE}_{i,m} = \frac{\sum_{h=1}^{n_{\mathrm{sample}}} \left[\frac{\sqrt{\left(y_{i,m,h} - \hat{y}_{i,m,h}\right)^2}}{y_{i,m,h}} \right]}{n_{\mathrm{sample}}} \times 100, \quad (9)$$

where $y_{i,m,h}$ is the (measured) value of quality variable m at the hth sampling instant of that phase, ^ indicates an estimated value, and n_{sample} is the total number of quality samples in the phase. This error can be averaged on all the validation batches, to get an MRPE value for each estimated quality variables during any of the estimation phases. In Fig. 7, the MRPE on the acidity number and on the viscosity for the three soft sensors is shown.

Although the explained variance on the viscosity during Phase 3 is lower then the other phases, the predictions are very accurate. It can be seen that, although all three soft sensors provide an estimation accuracy generally within the one of the laboratory analysis, the MATP–PLS model shows a superior overall performance.

It should also be noted that, using diagrams similar to those reported in Fig. 3, the reliability of an estimate provided by any of the estimators can be assessed online during each estimation phase. For example, the estimated viscosity profile during validation batch #5 is shown in Fig. 8a. While the batch is being run, the quality measurements of course are not available to assess the reliability of the estimation. However, complementing the estimation results (Fig. 8a, full line) with a Q-residuals plot (Fig. 8b; a T^2 plot can be used additionally), one can detect online when the estimated quality values are not reliable.

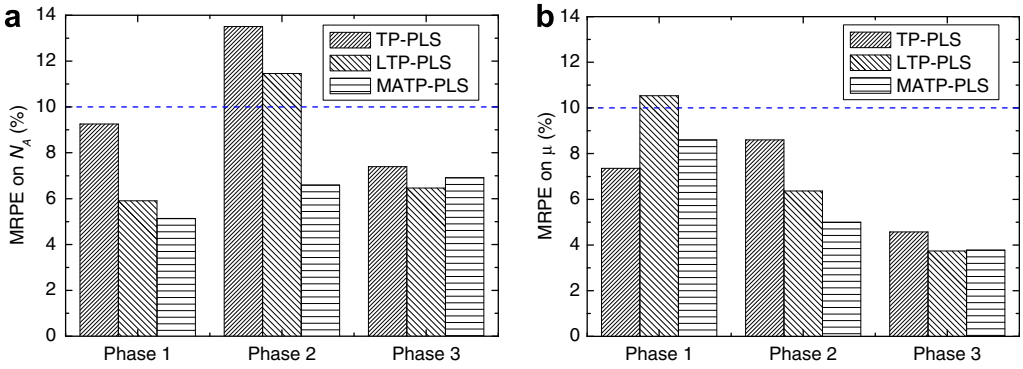

Fig. 7. Comparison between the estimation accuracy of the three-phase PLS models in terms of the average mean relative prediction errors (MRPE) on the validation datasets for (a) the acidity number and (b) the viscosity. The dashed lines represent the laboratory analysis accuracy.

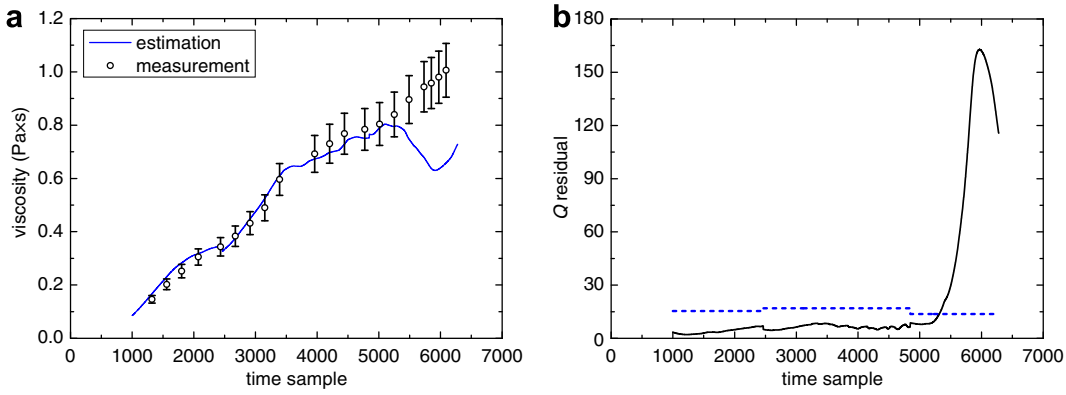

Fig. 8. Online assessment of the reliability of viscosity prediction using the MATP–PLS estimator in validation batch #5: (a) viscosity profile and (b) Q-residuals control chart with 95% confidence limits.

5. Conclusions

Partial least squares regression has proved to be a reliable tool for the online estimation of the product quality properties in an industrial batch polymerization process for the manufacturing of a resin. The process under study was characterized by a large number of available process measurement samples, uneven batch duration, scarce number of quality variable measurements with uneven sampling of (and lag on) these variables, complex and almost entirely non-reproducible sequence of processing steps.

To compensate for the nonlinear nature of the input/output mapping, a segmentation of the batches into a limited number of estimation phases was carried out by highlighting three clusters of score points in the scores plot of the reference dataset. Within each of these phases, linear PLS submodels were shown to provide quality estimations with an accuracy comparable to the one of the lab measurements, but with no delay and at a much higher frequency. Switching between one submodel to another one was triggered by clearly detectable landmark events occurring in the process.

Inclusion of time information into the process data matrix was shown to substantially improve the estimation accuracy. Namely, augmenting the process measurement matrix with lagged measurements dampened most of the noise on the estimated values of the quality variables. Averaging the process measurement values on a moving window of fixed length provided a sort of "memory" of the batch evolution that proved to be useful both to suppress measurement noise and to attenuate process noise, especially during Phase 2, which is characterized by a high degree of inter-batch variability.

One of the advantages of the resulting moving average three-phase PLS estimator is that it is very easy to implement, because it does not require to modify the structure of the PLS algorithm. Furthermore, using averaged measurement represents an easy way to handle noise spikes or temporarily missing values of the process measurements. However, care must be taken in selecting the length of the moving window, because too wide a window may delay the appearance of out-of-threshold values in the T^2 or Q-residuals control charts.

Acknowledgements

Partial financial support to this work was provided by the University of Padova within Progetto di Ateneo 2005 "Image analysis and advanced modelling techniques for product quality control in the process industry".

PF, FB and MB would like to thank Sirca S.p.A. for allowing them to publish their industrial data.

References

[1] B.R. Bakshi, Multiscale PCA with application to multivariate statistical process monitoring, AIChE J. 44 (1998) 1596–1610.
[2] S. Beaver, A. Palazoglu, J.A. Romagnoli, Cluster analysis for autocorrelated and cyclic chemical process data, Ind. Eng. Chem. Res. 46 (2007) 3610–3622.
[3] J. Camacho, J. Picò, Online monitoring of batch processes using multi-phase principal component analysis, J. Process Control 16 (2006) 1021–1035.
[4] J. Chen, K. Liu, Online batch process monitoring using dynamic PCA and dynamic PLS models, Chem. Eng. Sci. 57 (2002) 63–75.
[5] I.G. Chong, C.H. Jun, Performance of some variable selection methods when collinearity is present, Chemom. Intell. Lab. Syst. 78 (2005) 103–112.
[6] B.S. Dayal, J.F. MacGregor, Recursive exponentially weighted PLS and its applications to adaptive control and prediction, J. Process Control 7 (1997) 169–179.
[7] P. Facco, M. Olivi, C. Rebuscini, F. Bezzo, M. Barolo. Multivariate statistical estimation of product quality in the industrial batch production of a resin, in: B. Foss, J. Alvarez (Eds.), Proceedings of DYCOPS 2007 – 8th IFAC Symposium on Dynamics and Control of Process Systems, Cancun, Mexico, June 6–8, 2 (2007) 93–98.
[8] P. Geladi, R. Kowalski, Partial least squares regression: a tutorial, Anal. Chim. Acta 185 (1986) 1–17.
[9] H. Kamohara, A. Takinami, M. Takeda, M. Kano, S. Hasebe, I. Hasimoto, Product quality estimation and operating condition monitoring for industrial ethylene fractionator, J. Chem. Eng. Japan 37 (2004) 422–428.
[10] K.A. Kosanovich, M.J. Piovoso, PCA of wavelet transformed process data for monitoring, Intell. Data Anal. 1 (1997) 85–99.
[11] T. Kourti, Multivariate dynamic data modeling for analysis and statistical process control of batch processes, start-ups and grade transitions, J. Chemom. 17 (2003) 93–109.
[12] T. Kourti, J.F. MacGregor, Process analysis, monitoring and diagnosis, using multivariate projection methods, Chemom. Intell. Lab. Syst. 28 (1995) 3–21.
[13] W. Ku, R.H. Storer, C. Georgakis, Disturbance detection and isolation by dynamic principal component analysis, Chemom. Intell. Lab. Syst. 30 (1995) 179–196.
[14] N. Lu, F. Gao, Stage-based process analysis and quality prediction for batch processes, Ind. Eng. Chem. Res. 44 (2005) 3547–3555.
[15] N. Lu, Y. Yang, F. Gaom, F. Wang multivariate dynamic inferential modeling for multivariable processes, Chem. Eng. Sci. 59 (2004) 855–864.
[16] O. Marjanovic, B. Lennox, D. Sandoz, K. Smith, M. Crofts, Real-time monitoring of an industrial batch process, Comput. Chem. Eng. 30 (2006) 1476–1481.
[17] A. Maulud, D. Wang, J.A. Romagnoli, A multi-scale orthogonal nonlinear strategy for multi-variate statistical process monitoring, J. Process Control 16 (2006) 671–683.
[18] D. Neogi, C.E. Schlags, Multivariate statistical analysis of an emulsion batch process, Ind. Eng. Chem. Res. 37 (1998) 3971–3979.
[19] P. Nomikos, J.F. MacGregor, Multi-way partial least squares in monitoring batch processes, Chemom. Intell. Lab. Syst. 30 (1995) 97–108.
[20] S.G. Qin, Recursive PLS algorithm for adaptive data modeling, Comput. Chem. Eng. 22 (1998) 503–514.
[21] S. Rännar, J.F. MacGregor, S. Wold, Adaptive batch monitoring using hierarchical PCA, Chemom. Intell. Lab. Syst. 41 (1998) 73–81.
[22] S.A. Russell, P. Kesavan, J.H. Lee, B.A. Ogunnaike, Recursive data-base prediction and control of batch product quality, AIChE J. 44 (11) (1998) 2442–2458.
[23] R. Sharmin, U. Sundararaj, S. Shah, L.V. Griend, Y.J. Sun, Inferential sensor for estimation of polymer quality parameter: industrial application of a PLS-based soft sensor for a LDPE plant, Chem. Eng. Sci. 61 (2006) 6372–6384.

[24] C. Ündey, S. Ertunç, A. Çinar, Online batch/fed-batch Process performance monitoring, quality prediction, and variable-contribution analysis for diagnosis, Ind. Eng. Chem. Res. 42 (2003) 4645–4658.
[25] B.M. Wise, N.B. Gallagher, The process chemometrics approach to process monitoring and fault detection, J. Process Control 6 (1996) 329–348.
[26] S. Wold, Exponentially weighted moving principal components analysis and projection on latent structures, Chemom. Intell. Lab. Syst. 23 (1994) 149–161.
[27] S. Wold, N. Kettaneh, H. Fridèn, A. Holmberg, Modeling and diagnostics of batch processes and analogous kinetics experiments, Chemom. Intell. Lab. Syst. 44 (1998) 331–340.
[28] S.J. Zhao, J. Zhang, Y.M. Xu, Performance monitoring of process with multiple operating modes through multiple PLS models, J. Process Control 16 (2006) 763–772.