

Document downloaded from:

<http://hdl.handle.net/10251/53073>

This paper must be cited as:

González Martínez, JM.; Ferrer Riquelme, AJ.; Westerhuis, JA. (2011). Real-time synchronization of batch trajectories for on-line multivariate statistical process control using Dynamic Time Warping. *Chemometrics and Intelligent Laboratory Systems*. 105(2):195-206. doi:10.1016/j.chemolab.2011.01.003.



The final publication is available at

<http://dx.doi.org/10.1016/j.chemolab.2011.01.003>

Copyright Elsevier

Additional Information

Real-time Synchronization of Batch Trajectories for On-line Multivariate Statistical Process Control using Dynamic Time Warping

J.M. González-Martínez^{a,*}, A. Ferrer^a, J. A. Westerhuis^b

^a*Departamento de Estadística e Investigación Operativa Aplicadas y Calidad, Universidad Politécnica de Valencia, Camino de Vera s/n, 46022, Valencia, Spain*

^b*Biosystems Data Analysis, Swammerdam Institute for Life Sciences, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands*

Abstract

This paper addresses the real-time monitoring of batch processes with multiple different local time trajectories of variables measured during the process run. For Unfold Principal Component Analysis (U-PCA)- or Unfold Partial Least Squares (U-PLS)-based on-line monitoring of batch processes, batch runs need to be synchronized, not only to have the same time length, but also such that key events happen at the same time. An adaptation from Kassidas *et al.*'s approach [1] will be introduced to achieve the on-line synchronization of batch trajectories using the Dynamic Time Warping (DTW) algorithm. In the proposed adaptation, a new boundaries definition is presented for accurate on-line synchronization of an ongoing batch, together with a way to adapt mapping boundaries to batch length. A relaxed greedy strategy is introduced to avoid assessing the optimal path each time a new sample is available. The key advantages of the proposed strategy are its computational speed and accuracy for the batch process context. Data from realistic simulations of a fermentation process of the *Saccharomyces cerevisiae* cultivation are used to illustrate the performance of the proposed strategy.

Keywords: Dynamic Time Warping, On-line Synchronization, Batch Processes, Multivariate Statistical Process Control.

1. Introduction

Batch processes play an important role in several industries such as pharmaceutical, biochemical, food or biological, among others. The development of analysis and monitoring schemes for this kind of processes is an important task from an economical point of view. It is desirable when monitoring schemes are able to detect and diagnose disturbances efficiently in order to take actions on time to obtain consistent and good quality products in a consistent and profitable way. For this reason, the development of new methods based on the projection onto latent structures have experimented a great interest in the last years. These statistical methods are key tools for process understanding, multivariate statistical process control, abnormal situation detection, fault diagnosis and process control. All of these goals are included in modern improvement technologies such as Process Analytical Technology (PAT), fostered by United States Food and Drugs Administration (FDA) for the pharmaceutical industry.

Nomikos and MacGregor [2, 3], in their pioneering work, introduced Batch Multivariate Statistical Process Control (BMSPC) based on Unfold Principal Component Analysis (U-PCA) and Unfold Partial Least Squares (U-PLS). Afterwards, other multiway methods were proposed to build empirical models for its subsequent application in monitoring batch process, such as Tucker models [4] and Parallel Factor Analysis (PARAFAC) [5]. One of the strong assumptions of most of these models is that all batch trajectories have

*Corresponding author

Email address: jogonmar@doctor.upv.es (J.M. González-Martínez)

equal duration and are synchronized (i.e., similar events happen at the same time points). Usually, data collected from batch processes have different time length because the time is not a deciding factor for the completion of the batch, but another process variable such as, e.g., the biomass amount produced in a batch run. Kourti [6] discussed different patterns of varying-time batch trajectories and their possible solutions to be synchronized. In case batches have different duration but trajectories overlap in the common time part, an U-PCA model can be fitted using the information from the long batches while the absent part of the trajectory for the short batches is treated as missing data [7]. For some batches data collection starts earlier than in other batches, so a shift can be applied to align/synchronize the batches. In more complex cases, trajectories of the variables have different shapes, even when batch durations are the same, indicating that the timing for key events during each batch is different.

Some approaches for batch synchronization have been proposed in literature. One of the most successful procedures, called the Indicator Variable approach, was suggested by Nomikos and MacGregor [2]. Their idea was to replace time by another measured variable that progresses monotonically in time and has the same starting and ending value for each batch, and to perform the on-line monitoring relative to its progress. Several applications of this approach can be found in [8, 9, 10, 11, 12, 13, 14]. However, such an indicator variable is not always available.

Other authors dealt with the batch alignment issue based on trivial ideas, such as truncating the trajectories of all batches to the shortest batch length, or compressing/expanding the trajectories using linear time adjustments by dividing each time point along the trajectory by the time at the 100 % completion point [15, 14, 16]. These ideas, although simple, are often inadequate for aligning batch trajectories [17]. Lakshminarayanan [15] proposed to extend each batch trajectories up to the longest batch duration by adding artificial values, in this case, the last measured value. This implicitly assumes that all batches evolve at the same pace and all the timing differences between trajectories appear at the last stage of the batch process. This is clearly a strong assumption which is not true in most cases. In addition to this, it is not correct to fill in the batch data by repeating the last observation values, because it may break the correlation and make the latent model detect this batch as abnormal [6].

Wold *et al.* [18] suggested to use their observation-wise unfolding (OWU) approach as a tool to predict the batch 'maturity' and align. Their purpose was based on unfolding the three-way batch data matrix to preserve the variable direction instead of preserving the batch direction as in Nomikos and MacGregor's approach. A PLS analysis is achieved between the unfolded matrix and the 'local batch time' or the 'batch maturity index'. Wold *et al.* proposed to analyse differences among batches by batch-wise unfolding the score matrix \mathbf{T} from the PLS-OWU model. Westerhuis *et al.* [17] showed that the Wold *et al.*'s approach can provide poor predictions of time for the analyzed batches, due to non-linear relationships between the latent variables for the process and time variable spaces. Another drawback is that the OWU scores, \mathbf{T} , need to be aligned before the batch modeling of the unfolded scores is achieved.

Other methodologies based on stretching, compressing and translating signals were proposed in literature, such as Correlation Optimized Warping (COW) and Dynamic Time Warping (DTW). The COW approach was originally designed to correct peak shifts in chromatographic profiles [19, 20, 21]. This method consists of stretching or expanding pieces of sample signal (a series of points) to synchronize it against a reference signal using two input parameters, segment length (m , length of pieces) and slack (t , maximum allowed adjustment for the length of pieces). The optimal warping path is assessed by maximizing correlation factors between different pieces of signals. The necessity to optimize input parameters and its high computational cost are two drawbacks to bear in mind. One application in on-line batch process monitoring can be found in Fransson and Folestad [22].

Regarding Dynamic Time Warping (DTW), it is a dynamic-programming-based technique originated from speech recognition to match similar events between signals. DTW has the ability to synchronize two trajectories by appropriately translating, stretching and shrinking sample points to match with reference points in order to minimize the distance between both trajectories. In a process chemometrics context, DTW was first used as a supervision and fault detection tool, and also as a method to recognize different phases in bioprocesses [23]. Kassidas *et al.* [1] adapted the DTW algorithm to batch statistical process control as a tool to synchronize batches with different lengths and/or different local times. Later on, a modification of the Kassidas *et al.*'s DTW algorithm was proposed [24].

Some authors emphasize the importance of using the warping information that comes out of the synchronization. Once raw data with unequal batch duration is aligned, the original time profile is removed from the data. Recovery of this primary information is essential to improve quality product prediction and also fault detection during a batch run. Westerhuis *et al.* [25] used the cumulative warp information as an extra variable in the U-PCA model in order to ease the fault detection. García-Muñoz *et al.* [8] used the cumulative time deviation from the average time as a new variable to reach a better fault detection, after synchronizing the batch data by the Indicator Variable approach.

Most of the research done on batch synchronization has been focused on off-line batch process monitoring. Unfortunately, there has not been a high interest to face the issue of batch synchronization in multivariate statistical process control for real-time applications. Only Kassidas *et al.* [1] and Ramaker *et al.* [24] comment on this topic. Nevertheless, there are several open issues deserving research. This is the main goal of this paper.

Section 2 introduces a brief review of the Dynamic Time Warping algorithm. A modification from Kassidas *et al.*'s approach [1], based on introducing a semi-greedy search strategy, will be proposed in Section 3 to improve the alignment fit. Issues as the use of the warping information in the latent model, the strategies to give more importance to certain variables in the alignment procedure, some details about the inner DTW algorithm and the convenience of using warping information to improve the detection power of the model will also be discussed. In addition, some guidelines will be presented to carry out real-time synchronization for on-line multivariate statistical batch process control. Section 4 illustrates the performance of the proposed approach for real time alignment using simulated data from a *Saccharomyces Cerevisiae* fermentation process. Finally, conclusions will be drawn in Section 5.

2. Off-line DTW synchronization for Batch Monitoring

Let us assume a set of batch profiles are collected under normal operating conditions (NOC). Let $\mathbf{B}_n (K_n \times J)$ be the matrix containing the trajectories of J variables measured in K_n time intervals for n -th batch, $n = 1, \dots, N$. Each trajectory set over all batches may have time-varying runs, and/or may be unsynchronized (i.e. non-overlapping or different local time). As commented before, the DTW algorithm allows a non-linear warping of batch profiles in such a way that similar events which appear in different time instants will be synchronized by means of several shifts, compression and/or expansion of trajectory features, so that a minimum distance between them is achieved. DTW works with pairs of patterns, thus, alignment must be reached batch to batch and in a multivariate approach. Let us assume there is a batch $\mathbf{B}_{ref} (K_{ref} \times J)$ that can be used as the reference to achieve synchronization. In this paper, the batch whose length is closest to the median duration of collected NOC batches will be used as reference batch, \mathbf{B}_{ref} . The median is less sensitive to outliers and thus it is preferred to other choices such as the average.

Before applying DTW, the batches usually need to be scaled properly to account for the differences in the units used to record the trajectories of the J variables. Kassidas *et al.* [1] suggested that each variable in all batches is divided by its average range as

$$\tilde{\mathbf{b}}_{n,j}^T = \frac{\mathbf{b}_{n,j}^T \cdot N}{\sum_{n=1}^N \max(\mathbf{b}_{n,j}^T) - \min(\mathbf{b}_{n,j}^T)} \quad (1)$$

where $\mathbf{b}_{n,j}^T (1 \times K_n)$ is the row vector of the \mathbf{B}_n trajectory matrix.

Other choices such as the average standard deviation or the average interquartile range could also be used. The latter has the advantage of being less sensitive to outliers. Note that working out scaling weights across batches for a variable at a particular time point in the trajectory is not sensible because batches are not synchronized.

If batch data has the same measurement units of similar magnitude, as it is usually the case in spectroscopy [24] or with data registered from artificial tongues and noses, the scaling step for DTW can be omitted.

Following Kassidas *et al.*'s approach [1], a symmetric DTW algorithm combined with an asymmetric synchronization is used to synchronize each \mathbf{B}_n with \mathbf{B}_{ref} .

After setting the reference batch and scaling all batches properly, the symmetric DTW algorithm is applied. A $K_n \times K_{ref}$ grid is displayed for all sampling times of \mathbf{B}_n and \mathbf{B}_{ref} . Let us define the warping function or path as

$$\begin{aligned} \mathbf{f}_n^T &= \{w(1), w(k), \dots, w(K_{w_n})\} \\ \max(K_{ref}, K_n) &\leq K_{w_n} \leq K_{ref} + K_n \end{aligned} \quad (2)$$

where each $w(k)$ is a position in the grid indicating the elements of the multivariate trajectory of both batches that are matched:

$$w(k) = [i(k), j(k)], \quad k = 1, \dots, K_{w_n} \quad (3)$$

At each grid point, a weighted local distance measure between the $i(k)$ -th row vector of the \mathbf{B}_n trajectory matrix, $\tilde{\mathbf{b}}_{n,i(k)}^T$, and the $j(k)$ -th row vector of the \mathbf{B}_{ref} trajectory matrix, $\tilde{\mathbf{b}}_{ref,j(k)}^T$, which reflects the dissimilarity between them, is assessed as:

$$d(i(k), j(k)) = \left(\tilde{\mathbf{b}}_{n,i(k)} - \tilde{\mathbf{b}}_{ref,j(k)} \right)^T \cdot \mathbf{W} \cdot \left(\tilde{\mathbf{b}}_{n,i(k)} - \tilde{\mathbf{b}}_{ref,j(k)} \right) \quad (4)$$

where \mathbf{W} is a diagonal nonnegative matrix used to give more weight to certain variables. Let us define the cumulative weighted local distance between two batches along the path \mathbf{f}_n as:

$$D(f_n) = D(i(K_{w_n}), j(K_{w_n})) = \sum_{k=1}^{K_{w_n}} d(i(k), j(k)) \quad (5)$$

Some constraints must be defined to reduce the number of possible paths as the search space is very large. The solution search is computationally expensive, in terms of both memory and space. Hence, the optimal path \mathbf{f}_n^* is given by the solution to the following optimization problem:

$$D^*((i(K_{w_n}), j(K_{w_n}))) = \min_f [D((i(K_{w_n}), j(K_{w_n}))) \quad (6)$$

Dynamic programming is used to find the optimal solution by searching on the constrained grid under continuity, local and global constraints. The continuity constraint avoids the optimal path to go back in time. Local constraints define a set of predecessors for each point (i, j) in the grid. Global constraints define a subset of the $K_n \times K_{ref}$ grid points to be the actual search space for the optimal path. Usually, Sakoe-Chiba local and global constraints [26] are used. This global constraint does not allow the path to deviate $\pm M$ grid points from the diagonal path defining the upper u_t and lower l_t bounds at each sampling time. The local constraint allows a point (i, j) in the grid to be reached only either through $(i-1, j)$, $(i, j-1)$ or $(i-1, j-1)$ point. So, the optimization problem is solved by the next dynamic programming recursive equation:

$$D^*((i(k), j(k))) = d(i, j) + \min \left\{ \begin{array}{l} D(i-1, j) \\ D(i-1, j-1) \\ D(i, j-1) \end{array} \right\} \quad (7)$$

Later on, an asymmetric synchronization step is achieved by replacing the points of \mathbf{B}_n that are aligned with the same point of \mathbf{B}_{ref} with the mean value of these original values of \mathbf{B}_n . After this step, all trajectories will have the same duration and will be synchronized with each other, not only with the reference batch. For more details about the DTW algorithm, readers are referred to [1].

As previously mentioned, the weight matrix \mathbf{W} is used to give more importance to certain variables based on a particular criterium. Several choices are found in the literature. Kassidas [1] proposed an iterative procedure to give more weight to variables that are consistent from batch to batch. One drawback of Kassidas *et al.*'s weighting criteria is that consistent horizontal trajectories, with no warping information,

could have a high weight in defining the optimal path. Ramaker *et al.* [24] provided a new definition of the weight matrix to give more importance to variables containing more warping information such as curved-like process variables. Important warping information is associated to the optimal path. Ramaker *et al.* proposed a procedure to estimate which variables are more important from a warping perspective.

A combination of both methods would provide high weights to variables that are more consistent from batch to batch and containing warping information. It is desirable that variables with these features have more importance in the synchronization. Hence, let us define a new weight definition as

$$w_{j,j} = \sqrt{w_{j,j|Ram} \times w_{j,j|Kass}} \quad (8)$$

the geometric mean of the weights estimated following the Ramaker *et al.*'s approach ($w_{j,j|Ram}$) and Kassidas *et al.*'s approach ($w_{j,j|Kass}$) for each variable.

Note that in all the above approaches, \mathbf{W} is normalized in such a way that the sum of the diagonal elements equals the number of process variables J .

A paramount issue to be taken into account is the use of warping information as a tool to improve detection of disturbances and wrong alignments. When DTW algorithm is applied to a set of batch runs, all of them are transformed from original to a time scale of reference batch. If the original time scale is erased from the data set, we lose a potential way to detect errors of different type which might have an effect on the normal process performance or the product quality. Hence, it is advisable to use warping information over all batch runs by introducing the warping information as a new variable in the latent model, as Westerhuis *et al.* [25] pointed out. However, the correlation of this variable with the rest of the process measurements is an important issue to be taken into account. The most likely, this variable is not correlated with any process variable in the U-PCA model, thus this effect would be hidden by the model. One solution may be to give more weight to this variable in the model. Another choice is to analyse the warping information separately. From the warping information of the NOC synchronized batches, a univariate control chart is built to monitor the performance of the alignment.

3. On-line DTW synchronization for Batch Monitoring

The Multivariate Statistical Process Control (SPC) monitoring scheme provides an on-line diagnostic tool to interrogate the latent model about possible reasons of faults in non-finished batches [2]. Later on, corrective actions can be applied to avoid future batches that are affected by the same faults. Kassidas *et al.* [1] proposed an approach to carry out real-time alignment for on-line batch process monitoring. In this paper, several modifications from the previous approach are developed.

3.1. Kassidas *et al.*'s approach

Let \mathbf{B}_{ref} be the trajectory matrix of the reference batch and \mathbf{W} the weight matrix used to carry out the off-line synchronization of the NOC batch data. When new measurements are available at t sampling time, each variable is scaled by the average range estimated from the trajectories of NOC unsynchronized batches. Let \mathbf{B}_{new} ($t \times J$) be the trajectories of J variables measured from time zero until the t current time for an evolving new batch.

The use of the off-line DTW algorithm for on-line synchronization is inadequate because the endpoint is unknown at each sampling time. Therefore, the endpoint constraint must be relaxed in order to obtain the optimal path that provides the best matching between \mathbf{B}_{ref} and the uncompleted batch \mathbf{B}_{new} .

Following the above idea, a set of cumulative weighted distances $D_{t,j}$ at t sampling time is calculated, where $j = l_t, \dots, u_t$ are the allowed points of \mathbf{B}_{ref} imposed by the lower l_t and upper u_t bound at t sampling time. Let us define e_t^* as the time point of the batch reference \mathbf{B}_{ref} where the minimum cumulative weighted distance $D_{t,j}$ occurs at t sampling time. Hence, the optimal path with the minimal total distance up to the e_t^* point can be determined. After applying the previous procedure, an asymmetric synchronization is required to have a duration of e_t^* sampling points.

Once the ongoing batch is synchronized, it can be monitored on-line using an U-PCA or U-PLS model built previously. However, the prediction of the future behaviour of the evolving batch is necessary for the

monitoring. In the literature, several methods to achieve the prediction of absent part of batches can be found [2, 27, 28].

With every new measurement coming in, the above procedure has to be repeated.

3.2. Relaxed Greedy approach

Band global constraint is a paramount issue that has to be taken into account for on-line applications. The use of this constraint avoids large deviations from the diagonal path, preventing similar features in different time periods are not matched. For a feasible search, the band size M has to be at least equal or greater than the absolute value between the reference and test batch duration. As the duration of a new batch is unknown beforehand, this parameter cannot be fitted correctly, thus, it should be fixed by a defined criterium, such as taking the maximum, median or average batch duration from NOC batches. Consequently, a large search space is defined. It just allows us to search along the grid without taking into account the variability of the process. It may happen that the variability is different at different stages of batches, therefore the search space should be defined based on the consistency from batch-to-batch, i.e. the bands ought to be narrowed at the time periods with less deviation from the diagonal path.

In this paper, a band boundary more suitable for batch process monitoring is defined to provide greater accuracy in the synchronization. The warping window is optimized based on batch process understanding and off-line alignment of NOC batch data. Once a set of synchronized batches is available after aligning without global constraints, boundaries from historical NOC batch data are obtained. Note the off-line synchronization is performed without boundaries in this case because the main goal is to find the optimal path across all batches. In case global constraints were used, there would not be any guarantee the optimal path can be obtained.

Let \mathbf{F} be a matrix ($N \times K_{w_n}$) with all optimal paths \mathbf{f}^* obtained after applying the DTW algorithm, where N is the number of NOC batches and K_{w_n} is the number of warping indexes for the n -th batch. From all optimal paths, new boundaries are defined as:

$$\begin{aligned} \hat{u}_t &= \max(\mathbf{F}_{1,w(t,j(k_1))}, \dots, \mathbf{F}_{N,w(t,j(k_N))}) \\ \hat{l}_t &= \min(\mathbf{F}_{1,w(t,j(k_1))}, \dots, \mathbf{F}_{N,w(t,j(k_N))}) \\ t &= 1, \dots, \max(K_n) \text{ and } n = 1, \dots, N \end{aligned} \quad (9)$$

where \hat{u}_t and \hat{l}_t are the upper and lower boundaries at the t sampling time, respectively. All of them make up an envelope that includes all possible ways that the optimal path can take to synchronize multivariate batch trajectories based on the alignment of NOC batches (see Figure 1).

A crucial issue that was not considered by Kassidas *et al.* [1] was the updating of the upper and lower boundaries. This is necessary because the duration of a new batch is unknown and it may also occur that the batch length is greater or even smaller than those from NOC batches. For instance, in a fermentation process for *Saccharomyces Cerevisae* cultivation, when environmental and culture of bacterial conditions are appropriate and process is in control, the biomass concentration target can be reached sooner or later than in off-line NOC batches registered. In order to solve this situation, a boundary adaptation is required to achieve real-time alignment. In case global constraints are not updated, wrong alignment may be produced and the use of this warping information in the monitoring scheme would have no sense and it may introduce uncertain predictions and false alarms. An issue to keep in mind is that the updating of the boundaries will not affect the fault detection because the warping information is used in the monitoring scheme. In case an evolving batch length is longer or shorter than off-line NOC batches, this will be captured by the warping information and detected by the latent model.

One important issue in the online DTW alignment in U-PCA-based monitoring is the optimal path search at each sampling time. Once a new sample is available, achieving a backward search to find the best path from the current time point t up to the start point (1,1) would provide a different optimal path with regard to the optimal one obtained at the previous sampling time $t-1$. The immediate consequence is the values of the samples synchronized at the previous time point $t-1$ may change with the synchronization at the time point t (procedure followed by Kassidas *et al.* [1]). Then, those samples must be monitored again,

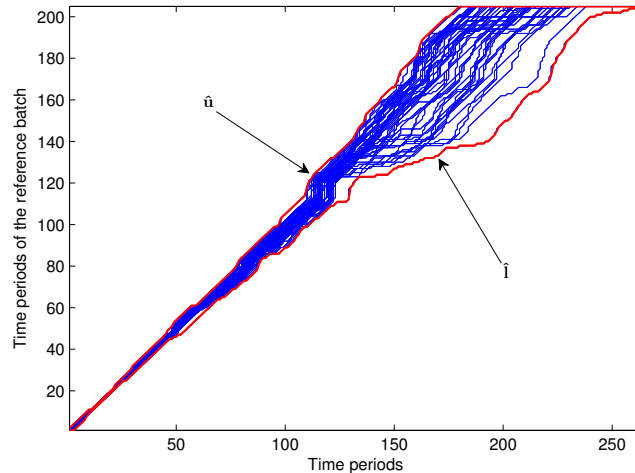


Figure 1: Set of optimal paths found by applying the off-line DTW algorithm without global constraints. The red lines are the minimum (\hat{l}_t) and maximum (\hat{u}_t) values at each time point, defining an envelope.

giving different values on T^2 and SPE statistics. Therefore, it might cause an uncertainty in the monitoring scheme and false alarms. It is necessary to use a strategy to overcome the above issue. One option would be to change the search procedure to find the optimal path. Instead of using a backward search to obtain the optimal path, a DTW variant based on a greedy search may be employed to synchronize the test batch locally. In other words, this variant does not assess the optimal path every time a new sample is coming in, but it estimates what matching is the best choice between points enclosed by a suitable window.

A time warping algorithm based on a relaxed greedy strategy, called Relaxed-Greedy Time Warping (RGTW), is proposed in this paper. Frequently, greedy algorithms fail to find the optimal solution because not all possible points are taken into account in the search procedure. The RGTW algorithm builds up a piecewise solution; always taking the next piece providing the most efficient and a quick benefit to a global solution. It is extremely useful when the problem is NP-complex, that is, problems where one solution can be found and verified quickly, being a solution closed to the optimal path. A dynamic programming-based search guaranteed to find an optimal solution if the search is achieved on the whole search region and makes decisions based on the decision from previous stages. On the other hand, greedy search makes one decision after another without having the chance to go back to obtain one better solution.

In our context, the real-time U-PCA-based monitoring scheme projects a new available measurement onto the NOC latent model. If one pure greedy search is chosen to find a matching path, a suboptimal path will be found. It is desirable to carry out the alignment providing a solution close to the optimal one. For that, we propose to use a window $\zeta_{t-\gamma+1,t}$ with a width γ , which constrains greedy choices to minimize dissimilarity between optimal and suboptimal path found (see Figure 2).

RGTW algorithm

Let us assume a new vector of measurements of the evolving batch $\tilde{\mathbf{b}}_{new,t}^T$ ($1 \times J$) is available at t sampling time and scaled by the average range estimated from the NOC batch trajectories. By using Eq. 4, local distances are assessed between $\tilde{\mathbf{b}}_{new,t}^T$ and the row vectors of the \mathbf{B}_{ref} trajectory matrix corresponding to the batch reference time points belonging to the interval defined from the boundaries shown in Fig. 1, i.e., from $\tilde{\mathbf{b}}_{ref,\hat{l}_t}^T$ to $\tilde{\mathbf{b}}_{ref,\hat{u}_t}^T$. The weights of the \mathbf{W} matrix follow Eq. 8.

Afterwards, a set of cumulative weighted distances $D_{t,j}$ at t sampling time is estimated, where $j = \hat{l}_t, \dots, \hat{u}_t$. Let e_t^* be the time point of the batch reference \mathbf{B}_{ref} where the minimum cumulative weighted

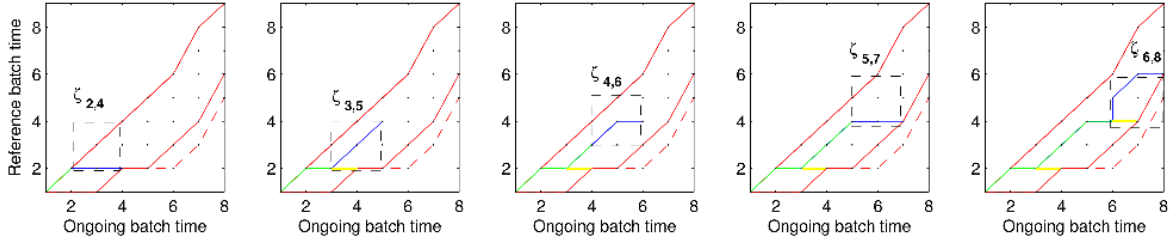


Figure 2: Performance of the on-line synchronization of an ongoing batch from the 4th up to the 8th sampling time, using the Relaxed-Greedy Time Warping with a window width γ equal to 3 units. Red solid lines represent the boundaries that restrict the optimal path search whereas the red dashed line shows the lower boundary update. The final optimal path $\hat{\mathbf{f}}_t^T$ and the optimal path $\hat{\mathbf{f}}_{t-\gamma+1}$ found within the window $\zeta_{t-\gamma+1}$ are plotted by green and blue solid lines, respectively. Yellow solid lines are path segments that are not taken into account on the final solution.

distance $D_{i,j}$ occurs at t sampling time. The optimal path at first γ sampling times of the evolving batch is derived following Kassidas *et al.*'s approach [1], from the start point $(1, 1)$ to the end point e_t^* .

Let us define the optimal path found from the window $\zeta_{t-\gamma+1,t}$ at the interval $[t - \gamma + 1, t]$ as

$$\mathbf{f}_{t-\gamma+1,t}^T = \{w(1), \dots, w(K_{t-\gamma+1,t})\} \quad (10)$$

where $K_{t-\gamma+1,t}$ is the number of points of the optimal path within the window $\zeta_{t-\gamma+1}$.

In case $t < \gamma$, any pairwise points of the optimal path from the window $\zeta_{1,t}$ is saved to the final solution. Otherwise, coordinates from the optimal path within the window $\zeta_{t-\gamma+1}$ for the $(t - \gamma + 2)$ -th sampling time will be saved to the final solution $\hat{\mathbf{f}}^T$ as

$$\hat{\mathbf{f}}_t^T = \hat{\mathbf{f}}_{t-1}^T \cup \hat{\mathbf{f}}_{t-\gamma+1|w(t-\gamma+2,j(k))}^T \quad (11)$$

being $\hat{\mathbf{f}}_{t-\gamma+1|w(t-\gamma+2,j(k))}^T$ every coordinates of the optimal path $\hat{\mathbf{f}}_{t-\gamma+1}$ found within the window ζ belonging to the reference batch points that match with the $(t - \gamma + 2)$ -th evolving batch point. $\hat{\mathbf{f}}_{t-1}^T$ and $\hat{\mathbf{f}}_t^T$ are the final optimal path at the previous and current sampling time, respectively.

Later on, the sliding window will move one position forward to go on with the alignment task with the $\zeta_{t-\gamma+1}$ window. It has a direct implication which is the imposition of a start point for the next sliding window. During the first γ points, the start point will be $(1, 1)$ but after the first movement of the window, successive start points will be the last pairwise point belonging to $w(t - \gamma + 2, j(k))$.

In case the endpoint e_t^* at the current time is lying on one extreme of the band, an updating of the upper or lower boundary, depending on where e^* lies on, must be carried out as

$$\begin{aligned} u_i &= \min(u_i + 2, K_{ref}) \quad i = t + 1, \dots, \max(K_n) \\ l_i &= \begin{cases} l_t & i = \{t + 1, t + 2\} \\ l_i - (l_{t+2} - l_t) & i = t + 3, \dots, \max(K_n) \end{cases} \\ & \quad n = 1, \dots, N \end{aligned} \quad (12)$$

When the duration of the ongoing batch is longer than the longest duration of NOC batches, with every new measurement coming in, the boundaries will be extended by taking the value of the latter one belonging to both the upper and lower boundaries.

After handling the alignment step, synchronized samples for the $(t - \gamma + 1)$ -th batch time are available. Depending on whether the algorithm compressed or expanded the samples, an update of the last monitored sample (the asymmetric DTW algorithm compressed, so an average with previous values that match with the same reference point is achieved, producing a new value) or some new values (stretching step) will be available, respectively. As a consequence of the previous step, an uncompleted batch in the reference time scale is available to be monitored. Hence, the prediction of the unknown batch part becomes a paramount

issue to achieve monitoring. Nomikos and MacGregor [2] proposed three approaches to predict unknown samples up to the end of the batch. In this work, Trimmed Score Regression method (TSR) [28] is used to impute unknown samples as missing data. TSR is a simple and efficient method to estimate the missing data. For more details about the algorithm and comparison with other approaches, readers are referred to [28]. Finally, once the whole batch is available with the unknown part estimated, it can be monitored in real-time by following Nomikos and MacGregor’s scheme [3, 2].

Cross-validation for the global constraint and warping window

Regarding the choice of the window width γ and the boundaries, in this paper a procedure to maximize similarity, minimizing the width, is proposed. For that, it is essential to know and understand the process dynamics. For instance, wastewater treatment or fermentation processes are featured by a slow dynamics. Thus, in these cases, a greater window width would improve the alignment quality without losing detection capability of disturbances on the process.

Then, we propose an off-line cross-validation for estimating the performance of the RGTW algorithm proposed using different window widths γ . Firstly, the matrix with N optimal paths \mathbf{F}_{DTW} is achieved after applying the off-line DTW algorithm without global constraints to the raw batch data. Later on, the raw batch trajectories are split into a training set (train) and test set (test) N times in the loop, being N as the number of batches. This step is carried out following the leave-one-out procedure. Every number of window widths γ that is considered is used to synchronize the train set in each iteration. As a result, a matrix with $N - 1$ suboptimal paths and window width γ $\mathbf{F}_{RGTW,\gamma}$ is obtained. Once the upper and lower boundaries are defined from those $N - 1$ paths, the test batch is synchronized using the on-line approach, leading to a suboptimal path $\mathbf{f}_{n,RGTW}^T$, being $n = 1, \dots, N$. This procedure is repeated until all N batches are synchronized over all window widths. The Pearson’s correlation coefficient ($r(\mathbf{f}_{n,DTW}, \mathbf{f}_{n,RGTW})$) between the suboptimal paths found in the loop for each window width γ and the optimal path from the off-line synchronization for all NOC batches is assessed. So, the similarity index for the different window width is obtained as

$$\delta_\gamma = \prod_{n=1}^N r(\mathbf{f}_{n,DTW}, \mathbf{f}_{n,RGTW}) \quad (13)$$

where $\mathbf{f}_{n,DTW}$ and $\mathbf{f}_{n,RGTW}$ are the optimal and suboptimal paths found by DTW and RGTW, respectively, for a certain window width equal to γ , for the n -th batch. It is expected that the similarity index increases when width is greater, yielding a high Pearson’s correlation r . In case the RGTW algorithm is applied with window width $\gamma = 1$, a pure greedy strategy is carried out. Wider widths will provide a strategy less greedy and more accurate.

4. Results

Saccharomyces cerevisiae is a yeast widely used in biotechnical and pharmaceutical industries for the production of proteins, so this yeast was chosen to simulate a fermentation process to illustrate the performance of the approach proposed in this paper. The model of the aerobic growth of *S. cerevisiae* on glucose limited medium introduced by Lei [29] is used as a case study.

Fermentation is performed in batch mode where glucose is in excess in the medium, cells are not able to consume the whole amount of glucose, hence ethanol is produced together with the excretion of pyruvate and acetate. Later on, the initial amount of glucose is consumed by the growing cells. Before ethanol is consumed, the accumulated amount of pyruvate and acetate are consumed. During growth on ethanol, acetate is produced again. For more details about theoretical assumptions and additional information concerning the first principles model, readers are referred to [29].

Process data was generated using Simulink for Matlab release 2007b[®] (©The MathWorks, Inc). In order to simulate the physical uncertainty caused by the biological variability, slightly modified values of constants of the first principles model were introduced into the parametric space that defines the Simulink

scheme for simulation. Also, Gaussian noise of low magnitude in the initial conditions (10%) and measurements (5%) to simulate the typical errors produced by sensors were added. Furthermore, the simulation achieved here have also taken into account the nature of this kind of processes. In fermentation processes, characterized by a duration of several days, some microorganisms may have different generation times, having a significant influence on biomass growth and quality features. Consequently, each batch performed has different durations.

A set of 50 unsynchronized batches was simulated under normal conditions and under the assumption commented previously. Ten variables were measured every sampling time over all batches: concentrations (glucose, pyruvate, acetaldehyde, acetate, ethanol and biomass), active cell material, acetaldehyde dehydrogenase (proportional to the measured activity), specific oxygen uptake rate and specific carbon dioxide evolution rate. Figure 3(a) shows 4 variables (selected from the 10 process variables) for 50 good quality batches from the simulated process. The variables shown in the Figure 3(a) illustrates the lack of synchronization among the trajectories and also the dissimilarity in the batch length. The total length of the batches varied from 180 to 264 data points. Before carrying out the statistic modeling, the raw data need to be synchronized. In the selection of the batch which represents overall typical batch to batch variations, the closest batch to median length was taken as reference batch to carry out the synchronization by DTW. The median length was taken and not the average as statistical index to determine the reference batch because this statistic is less sensitive to outliers. In this case, the median length is 207 sampling times and the batch #4, whose batch length is 205, is the closest. All conditions and constraints described in section 2, except for the Sakoe and Chiba band constraint, were fixed. The reason why this global constraint was not established is because it is desirable to find optimal paths over all batches in order to get an overall envelope defined for all optimal pathways. As it is a preprocessing step characterized by alignment and process understanding, DTW algorithm is not totally constrained.

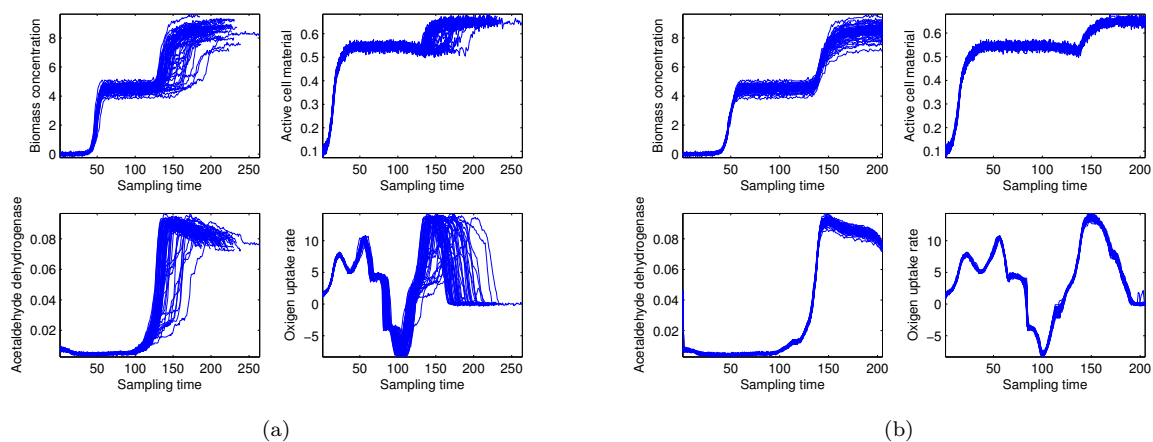


Figure 3: Trajectories belonging to four fermentation variables: before synchronization (a) and after synchronization using DTW based on the weights calculated as the geometric mean of Ramaker *et al.*'s and Kassidas *et al.*'s weights (b)

An issue to be taken into account is which variables should be introduced into the DTW algorithm. It may happen that there are on/off variables providing information about different episodes or phases in the process that would be extremely useful in the synchronization step to improve the alignment. In contrast, these variables would not likely provide important information in MPCA modeling because they are not correlated with other variables. In this case, all variables measured were used to carry out the alignment. So Kassidas *et al.*'s and Ramaker *et al.*'s approaches were applied to calculate the weight matrix \mathbf{W} , giving as a result (at 4th and 6th iteration, respectively) the weights shown in Figure 4. Later on, the final weights used by the DTW algorithm were calculated as the geometric mean of the previous weights (see Figure 4). After the DTW algorithm was applied to 50 NOC batches, a set of synchronized batches against batch #4

was reached (see Figure 3(b)).

As explained in section 2, Kassidas *et al.*'s approach gives more weight to those variables more consistent from batch to batch, i.e. those variables whose trajectories are not notably deviating from the average trajectory. These variables are characterized by a high signal to noise ratio, hence high weights are given. In this case, the iterative procedure gave less weight to certain variables, such as ethanol and biomass concentration, active cell material and acetaldehyde dehydrogenase (from 5th to 8th variable marked as asterisks in Figure 4), whereas the remaining variables were given more weight, being the 9th variable as the most weighted. This can be understood from Figure 3(b) where from the 6th to the 9th variables (biomass concentration, active cell material, acetaldehyde dehydrogenase and oxygen uptake rate) are plotted after synchronization. The first three plotted variables have a low signal to noise ratio in certain batch stages. In the 7th variable (active cell material), an important variation through the whole batch duration can be seen, therefore a low weight is expected. On the contrary, the 9th variable (oxygen uptake rate) has small differences between batch trajectories and the average trajectory, so a high weight is given. In an intermediate situation are the 6th and 8th variable, where larger differences are appreciated. In the case of the 6th variable (biomass concentration), from the beginning up to the 50th sampling time there is not much variation between the average trajectory and the test batches whereas from the 50th sampling time onwards, there is a considerable variation. Regarding the 8th variable, it seems to be consistent from the beginning to the 150th sampling time but after this point, a clear variation is detected. Hence, these two last variables are given an intermediate weight, being lower on the 6th variable than the 8th variable, as can be seen in Figure 4.

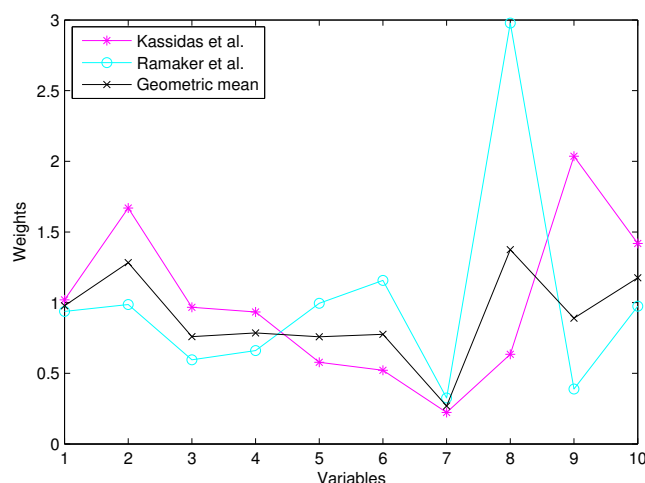


Figure 4: Values of the weight matrix \mathbf{W} assessed by following the Ramaker et al.'s and Kassidas *et al.*'s approaches together with the geometric mean of both weights

Since Ramaker *et al.*'s viewpoint, the weights should be given according to how much warping information is provided by a variable. The weights calculated by following this idea are drawn by circles in Figure 4. This plot shows that variables with less warping information are the active cell material (#7) and the oxygen uptake rate (#9), whereas the variable acetaldehyde dehydrogenase (#8) has a high warping information content. Furthermore, it's worth emphasizing the fact that the 8th variable receives more weight following Ramaker *et al.*'s approach than Kassidas *et al.*'s approach. The reason for that is that in spite of its variability around average trajectory, the shape of the 8th variable's trajectory explains the main occurrences in the process represented by the majority of the fermentation variables, providing this way a better fit on the alignment. Our proposal, introduced in Section 2, is to average out geometrically the weights of both approaches in order to give weights more suitably. The idea is to give more weight to those variables more

consistent from batch to batch and also providing important warping information. As can be seen in Figure 4, those variables that contain more warping information but are not consistent from batch to batch or vice versa, are weighted to an intermediate value between Ramaker *et al.*'s and Kassidas *et al.*'s weights. In case variables are weighted with similar values following Ramaker *et al.*'s and Kassidas *et al.*'s approaches, approximate weights are given by the geometric mean.

Before building the PCA model and carrying out the on-line monitoring of several batches using the RGTW algorithm, the width of the sliding window must be set and validated by the iterative procedure explained in Section 3. For that, from raw and synchronized batch data of the fermentation process, Pearson's correlation coefficients of warping information over all batches for a defined width were iteratively assessed, following the leave-one-out strategy. This allows us to determine the best window width that reduces the lack of fit caused by the greedy strategy. An Analysis of Variance (ANOVA) was performed on the Fisher Z-transformed correlation coefficients for the different window widths γ , showing statistically significant differences (p -value < 0.05). The Least Significant Difference (LSD) intervals displayed in Figure 5 show that correlation is statistically higher for window width $\gamma = 4$ (average $r = 0.9999$), whereas no statistically significant differences between window widths $\gamma = 1$ and $\gamma = 2$ were found. As a compromise between computational complexity and performance is desired in this case, a window width $\gamma = 3$ (average $r = 0.9999$) was selected.

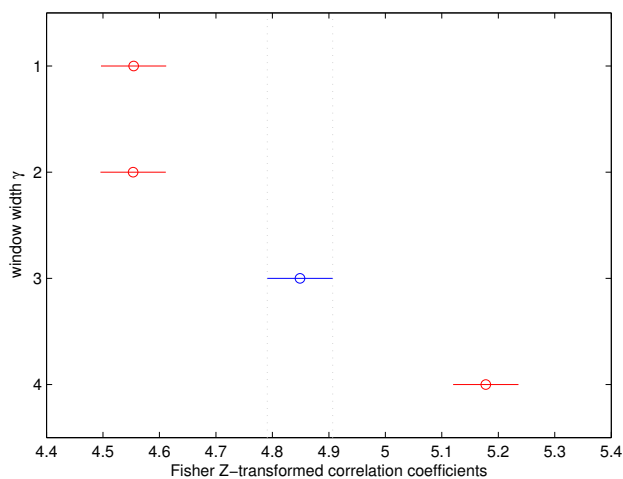


Figure 5: LSD intervals for the Fisher Z-transformed correlation coefficients for the different window widths γ .

Figure 6 shows the warping information from all suboptimal paths found in the validation procedure drawn in Section 3. From the beginning up to the 50th sampling time, the warping information is smooth for all batches, meaning nearly no compressions or expansions were necessary to synchronize all batches. However, from the 50th to 120th sampling time, a greater number of vertical and horizontal transitions were required to align the batches. Specifically, a series of expansions (horizontal paths) can be observed between the 105th to 120th sampling time. One reason for this fact could be the phase change during the process. From this point there were an important number of compressions (vertical paths) because many batches have longer durations than the reference one.

All this warping information can be used to improve the detection of problems that may be masked by the synchronization. Nevertheless, special care must be taken when the warping information is taken into account in latent-based models. In processes characterized by large amount of variables, the inclusion of warping information as a new variable may cause this one to have no influence in the latent model. It is due to the fact that the warp may not be correlated with any of the process variables, thus this new variable would have little weight. In such cases, this variable should be considerably weighted in order to avoid this

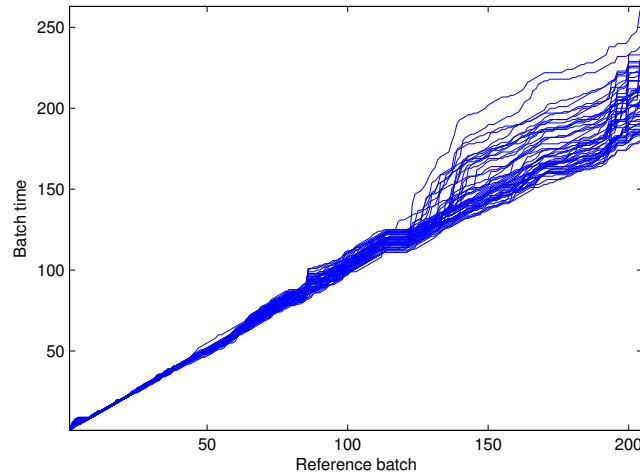


Figure 6: Distorsion of the batch time over all batches (warping information) after applying the RGTW algorithm

problem. In this work, the warping information will be introduced as a new variable into the model U-PCA without applying any special weight because of the features of this process. However, the capability of the model U-PCA to detect disturbances in the synchronization was checked as commented below.

Once all batches were synchronized, they were arranged together with the warping information in a $50 \times (11 \times 205)$ [batches \times (variables \times time)] two-way matrix. Multivariate analysis based on U-PCA did not show any outlier. Three principal components were selected based on cross-validation results explaining 57% of total variation. Two monitoring charts, T^2 -Hotelling and SPE, were developed from the validated model, and also confidence limits were initially estimated from theoretical results [30, 31]. In order to achieve a suitable performance of the monitoring charts, control limits for the SPE chart are readjusted using data in a leave-one-out basis. Limits are raised or lowered so that the Overall Type I risk is equivalent to the imposed significance level (in this case at 95% and 99% confidence level for both statistics).

In order to illustrate the approach proposed in his paper, three simulated batches with abnormalities and a batch with good quality were simulated. The out-of-control batches were simulated by modifying the initial values of the constants k_6 , k_{10} and k_{11} . Changing these constants, an abnormal behaviour on the formation of ethanol from acetaldehyde, the degradation rate of the active compartment (dependent on the glucose and ethanol concentration) and the degradation of acetaldehyde dehydrogenase compartment is expected, respectively.

The on-line synchronization approaches of Kassidas *et al.* and the one proposed in this paper are then used to achieve the real-time alignment of these simulated batches from the fermentation process. Both approaches are able to synchronize unequal batches but our method has some improvements over Kassidas *et al.*'s on-line implementation that will be shown by this example. Having built and validated both the PCA model and the parameters of the RGTW algorithm, the next step is to evaluate its ability to detect the out-of-control batches and time points and diagnose the cause of the out-of-control signal, i.e. the process variable(s) responsible for the abnormality. It is also worth knowing which approach produces more false alarms. For that, the NOC batches and three abnormal batches are synchronized and monitored in real-time. The results of the on-line synchronization from the start until the end of the batch run are summarised in Figure 7. Suboptimal paths from the NOC batch and three abnormal batches assessed from the on-line synchronization, applying the RGTW algorithm (see Figure 7(a)) and Kassidas *et al.*'s on-line approach (see Figure 7(b)), are shown. The batches that present an abnormality on ethanol, active cell material and acetaldehyde are plotted with a yellow (bk6), pink (bk10) and red line (bk11), respectively. Kassidas *et al.*'s on-line DTW implementation needs to estimate the optimal path for each sampling time

to achieve the synchronization. This causes the search algorithm of the optimal path to give different synchronization sequences as time goes by correcting decisions taken in the past. This can be observed from Figure 7(a), where many branches come out from the root path. An immediate consequence will appear on the monitoring charts. The monitoring scheme will monitor the whole ongoing batch, from the start point up to the current sampling time, calculating the T^2 -Hotelling and SPE statistics every sampling time. Therefore, it is expected that the monitoring charts produce many false alarms (see Figure 8). This behaviour is not seen in the RGTW algorithm (see Figure 7(b)) due to the fact that this algorithm uses a greedy strategy that provides a synchronized sample after the first γ steps. All optimal paths calculated for all synchronized batches are smooth, unlike the Kassidas *et al.*'s on-line implementation (see Figure 7(b)). Hence, the most important advantages of using the RGTW algorithm are the computational time and memory save and the noteworthy reduction of false alarms.

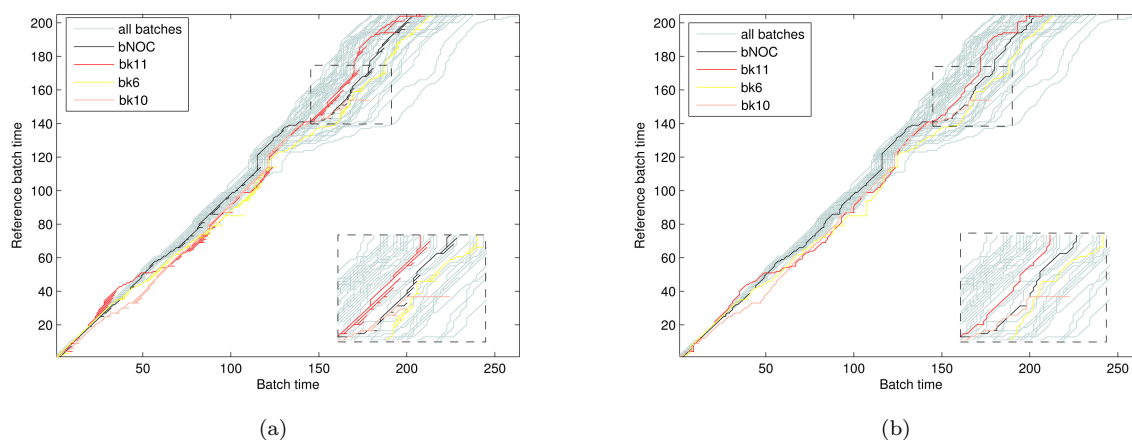


Figure 7: Suboptimal paths from a test batch (black path) and three abnormal batches (red, yellow and pink path) calculated by applying the Kassidas *et al.*'s on-line DTW implementation (a) and applying the RGTW algorithm with a window width γ equal to 3 units (b)

Another issue to discuss is the role of the warping information in the monitoring scheme. The result of the synchronization on the abnormal batches (see Figure 7) is different from the one on NOC batches (lightgreen paths). From the 25th up to the 60th and from the 70th up to 90th sampling time a different behaviour is detected in the warping information. During this time interval, there were more expansions (vertical paths) registered in the batch which presents an abnormality caused by a modification on the constant k_{11} (red path) than in NOC batches. Nevertheless, a different behaviour is presented in the abnormal batch which represents the degradation rate of the active cell material (the pink path). Unlike the abnormal batch (bk11), more compressions (horizontal paths) are shown from the 25th to the 60th sampling time than in the NOC batches. It seems the warping information may provide an extra information to detect the abnormality.

Once the result of the synchronization is obtained, Figure 9 shows the result of the on-line monitoring of the NOC batch and three abnormal batches through the whole batch run. SPE and T^2 -Hotelling charts are shown for all simulated batches with an abnormal behaviour and under NOC. It can be seen that the SPE control charts correctly signals the abnormal batches from the very beginning, since each simulated fault breaks down the correlation structure from the start point.

In Figure 10, the SPE contribution plots at the time point 80 for the three faulty batches are shown. These charts allow diagnosing the particular fault. For the simulated batch with an abnormal behaviour on the formation of ethanol from acetaldehyde, the SPE contribution highlights the variables acetaldehyde (#3), acetate (#4), ethanol concentration (#5) and biomass concentration (#6) as responsible for the abnormality. Figure 10(b) represents the variables responsible for the degradation rate of the active compartment, mainly, the active cell material (#7) and CO_2 evolution rate (#10). Finally, the monitoring scheme detected the

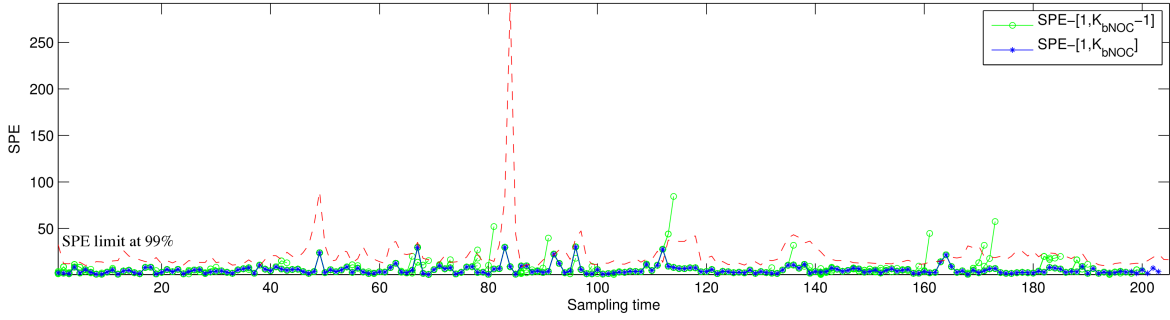


Figure 8: SPE chart for a NOC batch after applying the Kassidas *et al.* on-line DTW implementation at each sampling time. Circles represent SPE statistic obtained at the sampling time t whereas, the asterisks represent SPE statistics assessed at the end of the batch after the synchronization was carried out.

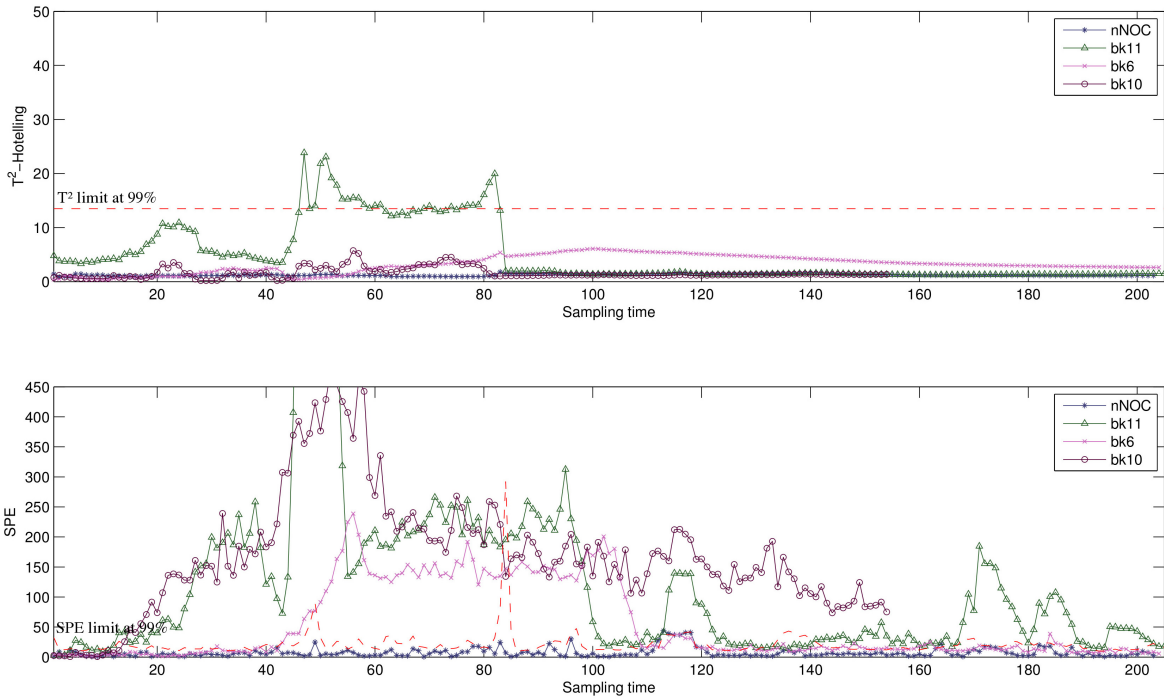


Figure 9: Multivariate control charts monitoring a synchronized batch and three abnormal batches from the start point up to the end after applying the RGTW algorithm in each sampling time.

acetaldehyde (#3), acetate (#4) and biomass concentration (#6) to be the main variables responsible for the abnormality of the third faulty batch for the SPE statistic (Figure 10(c)), while the ethanol concentration and the warping information received less weight. Another important issue is that for the three faulty batches the contributions plot gave at least some weight to the warping information variable. This shows that the monitoring scheme is able to detect an abnormal behaviour during the synchronization, as commented above. The proposed RGTW algorithm, as well as reducing false alarms produced by the synchronization, provides computation benefits compared to Kassidas *et al.*'s on-line implementation. In order to demonstrate

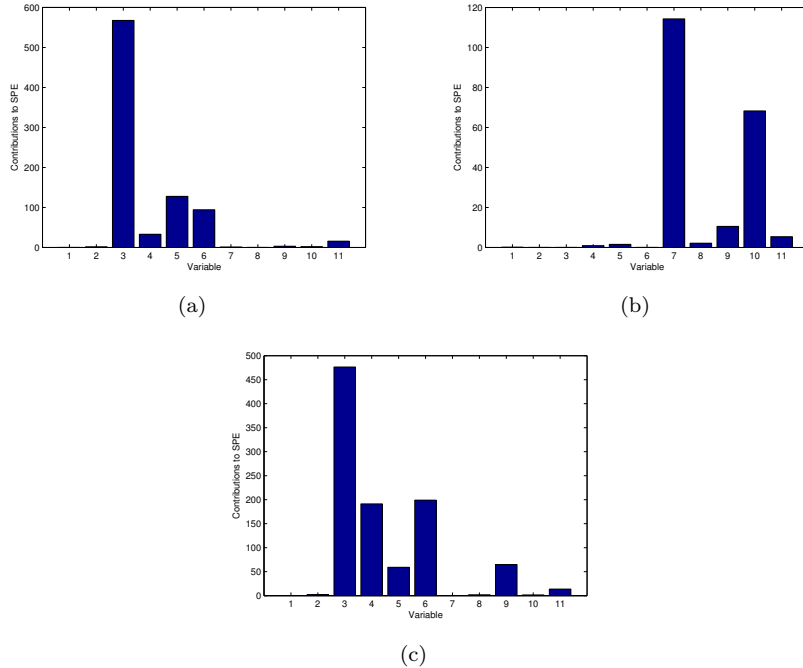


Figure 10: Variable contribution to SPE statistic for fault diagnostics of the abnormal batches at the 80th sampling time: an abnormal behaviour on the formation of ethanol from acetaldehyde (a), degradation rate of the active compartment (dependent on the glucose and ethanol concentration) (b) and degradation of acetaldehyde dehydrogenase compartment (c).

the efficiency of the new approach, both algorithms were applied 1000 times to the NOC batch simulated, measuring the runtime each time. For that, the system clock on an Intel Core Duo with 4GB of RAM memory with minimal background processes running was used. At the end of the computation experiment, the average of runtimes measured was assessed. The RGTW algorithm was run by using three different window widths ($\gamma=1,3,5$) together with the bands defined by the cross-validation procedure. In the case of the Kassidas *et al.*'s implementation, Sakoe-Shiba global constraint with a width equal to $M=59$ (the difference between the largest length of NOC batches and the reference batch length) was set. The evaluation gave as a result that the RGTW algorithm was significantly faster than Kassidas *et al.*'s on-line DTW implementation for all window widths tested. Particularly, our approach required on average 0.241, 0.251 and 0.262 seconds to synchronize the NOC batch with a window width γ equal to 1, 3 and 5 units, respectively. In contrast, Kassidas *et al.*'s on-line implementation took 1.935 seconds on average. So, the maximum speed-up factor for this batch length is 8. An ANOVA was also performed on the run times obtained for each approach, showing statistical significant differences (p -value < 0.05) between RGTW and Kassidas *et al.*'s approach. When data are larger in terms of time points, differences on execution time will be greater. This is caused by the modifications proposed in this paper. Firstly, using the new global band constraint definition, the synchronization procedure is speeded up by a constant factor because less cells of the weight matrix \mathbf{W} must be evaluated. Secondly, the relaxed greedy strategy avoids calculating the suboptimal path for each sampling time, thus the optimization problem is reduced to the window ζ instead of the whole batch length. Thirdly, the relaxed-greedy synchronization does not require the complete cumulative weighted local distance \mathbf{D} be kept in the memory, as in the case of Kassidas *et al.*'s approach. It only needs to keep the current and $\gamma - 1$ previous columns in memory because the RGTW algorithm finds the suboptimal path within the window ζ instead of through the ongoing batch time.

5. Conclusions

This paper addresses the problem of unsynchronized batch trajectories for multivariate statistical process monitoring. As U-PCA- and U-PLS-based monitoring schemes assume that all batch trajectories have equal length, all process variables are measured at the same sampling frequency and key events happen at the same time, synchronization becomes a primary pre-processing step both in the modeling and in the exploitation phase.

Based on the original dynamic time warping (DTW) Kassidas *et al.*'s approach, a new time warping algorithm for on-line batch synchronization based on relaxed greedy strategy, called Relaxed-Greedy Time Warping (RGTW), is proposed in this paper. The new proposal avoids assessing the optimal path each time a new sample is available, increasing the computational speed. Furthermore, it reduces the uncertainty in the monitoring statistics and predictions in such a way that false alarms are notably reduced. The main contributions of the RGTW approach are: (1) a new global band constraint definition that takes into account the variability across all batches in each time period; (2) a definition of a cross-validated monitoring window in order to use an optimized window size yielding the minimum local optimal paths (a relaxed greedy strategy); and (3) a way to adapt mapping boundaries to the batch length. This last point is highly critical because in the on-line approach the endpoint of ongoing batch is unknown, hence, normal boundaries must be adapted by the current batch run.

Data from realistic simulations of a fermentation process of the *Saccharomyces cerevisiae* cultivation are used to illustrate the good performance of the proposed strategy.

Acknowledgements

This research work was supported by the Spanish government under the project (DPI2008-06880-C03-03). We also gratefully acknowledge José Camacho PhD. for providing simulated data from a fermentation process of *Saccharomyces cerevisiae*. The authors would also like to acknowledge the valuable suggestions made by Prof. Paul Taylor.

References

- [1] A. Kassidas, J. F. MacGregor, P. Taylor, Synchronization of batch trajectories using dynamic time warping, *AIChE Journal* 44 (1998) 864–875.
- [2] P. Nomikos, J. MacGregor, Multivariate spc charts for monitoring batch processes, *Technometrics* 37 (1995) 41–59.
- [3] P. Nomikos, J. MacGregor, Monitoring batch processes using multiway principal component analysis, *AIChE Journal* 40 (1994) 1361–1375.
- [4] A. K. Smilde, S. van der Wiel, Application of tucker models for batch monitoring, *AIChE Journal* 62 (1999) 1320.
- [5] X. Meng, A. Morris, E. Martin, On-line monitoring of batch processes using a parafac representation, *Journal of Chemometrics* 17 (2003) 65–81.
- [6] T. Kourti, Abnormal situation detection, three-way data and projection methods; robust data archiving and modeling for industrial applications, *Annual Reviews in control* 27 (2003) 131–139.
- [7] T. Kourti, Multivariate dynamic data modeling for analysis and statistical process control of batch processes, start-ups and grade transitions, *Journal of Chemometrics* 17 (2003) 93–109.
- [8] S. García-Muñoz, T. Kourti, J. MacGregor, Troubleshooting of an industrial batch process using multivariate methods, *Industrial and Engineering Chemistry Research* 42 (2003) 3592–3601.
- [9] C. Undey, S. Ertunç, A. Çinar, Online batch/fed-batch process performance monitoring, quality, prediction, and variable-contribution analysis for diagnosis, *Industrial & Engineering Chemistry Research* 42 (2003) 4645–4658.
- [10] D. Neogi, C. Schlags, Multivariate statistical analysis of an emulsion batch process, *Industrial and Engineering Chemistry Research* 37 (1998) 3971–3979.
- [11] T. Kourti, J. Lee, J. MacGregor, Experiences with industrial applications of projection methods for multivariate statistical process control, *Computers & Chemical Engineering* 20 (1996) S745–S750.
- [12] C. Duchesne, T. Kourti, J. MacGregor, Multivariate spc for startups and grade transitions, *AIChE Journal* 48 (2002) 2890–2901.
- [13] Y. Zhang, M. Dudzic, V. Vaculik, Integrated monitoring solution to start-up and run-time operations for continuous casting, *Annual Reviews in Control* 27 (2003) 141 – 149.
- [14] M. Zarzo, A. Ferrer, Batch process diagnosis: Pls with variable selection versus block-wise pcr, *Chemometrics and Intelligent Laboratory Systems* 73 (2004) 15–27.
- [15] S. Lakshminarayanan, R. Gudi, S. Shah, Monitoring batch processes using multivariate statistical tools: extensions and practical issues, in: *Proceedings of IFAC Worm Congress*, pp. 241–246.

- [16] D. J. Louwerse, A. A. Tates, A. K. Smilde, G. L. M. Koot, H. Berndt, Pls discriminant analysis with contribution plots to determine differences between parallel batch reactors in the process industry, *Chemometrics and Intelligent Laboratory Systems* 46 (1999) 197 – 206.
- [17] J. Westerhuis, T. Kourti, J. MacGregor, Comparing alternative approaches for multivariate statistical analysis of batch process data, *Journal of Chemometrics* 13 (1999) 397 – 413.
- [18] S. Wold, N. Kettaneh, H. Fridén, A. Holmberg, Modelling and diagnostics of batch processes and analogous kinetic experiments, *Chemometrics and Intelligent Laboratory Systems* 44 (1998) 331 – 340.
- [19] N. Nielsen, J. Carstensen, J. Smedsgaard, Aligning of single and multiple wavelength chromatographic profiles for chemometrics data analysis using correlation optimised warping, *Journal of Chromatography* 805 (1998) 17–35.
- [20] V. Pravdova, B. Walczak, D. Massart, A comparison of two algorithms for warping of analytical signals, *Analytica chimica acta* 456 (2002) 77–92.
- [21] G. Tomasi, F. van den Berg, Correlation optimized warping and dynamic time warping as preprocessing methods for chromatographic data, *Journal of Chemometrics* 18 (2004) 231–241.
- [22] M. Fransson, S. Folestad, Real-time alignment of batch process data using cow for on-line process monitoring, *Chemometrics and Intelligent Laboratory Systems* 84 (2006) 56–61.
- [23] K. Gollmer, C. Posten, Supervision of bioprocesses using a dynamic time warping algorithm, *Control Engineering Practice* 4 (1996) 1287–1295.
- [24] H. Ramaker, E. van Sprang, J. Westerhuis, A. K. Smilde, Dynamic time warping of spectroscopic batch data, *Analytica Chimica Acta* 498 (2003) 133–155.
- [25] J. A. Westerhuis, A. Kassidas, T. Kourti, P. A. Taylor, J. F. MacGregor, On-line synchronization of the trajectories of process variables for monitoring batch processes with varying duration, in: *Proceedings of SSC6: 6th Scandinavian Symposium on Chemometrics*, Porsgrunn Norway, August 15–20.
- [26] H. Sakoe, S. Chiba, Dynamic programming algorithm optimization for spoken word recognition, *IEEE Transactions on acoustics, speech, and signal processing* 26 (1978) 43–49.
- [27] S. García-Muñoz, T. Kourti, J. F. MacGregor, Model predictive monitoring for batch processes, *Industrial and Engineering Chemistry Research* 43 (2004) 5929–5941.
- [28] F. Arteaga, A. Ferrer, Dealing with missing data in mspc: several methods, different interpretations, some examples, *Journal of Chemometrics* 16 (2002) 408–418.
- [29] F. Lei, M. Rotbøll, S. Jørgensen, A biochemically structured model for *saccharomyces cerevisiae*, *Journal of biotechnology* 88 (2001) 205–221.
- [30] N. Tracy, R. M. J.C. Young, Multivariate control charts for individual observations, *Journal of Quality Technology* 24 (1992) 88.
- [31] J. Jackson, G. Mudholkar, Control procedures for residuals associated with principal component analysis, *Technometrics* 21 (1979) 341.