2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Discriminatory Learning based Performance Monitoring of Batch Processes

Shailesh Patel, Ramprasad Yelchuru*, Srikanth Ryali, and Ravindra Gudi

Abstract— **This paper proposes a novel approach towards performance monitoring of batch processes that is oriented towards the requirements of real time assessment of batch health and online batch qualification. The proposed approach is based on the use of discriminant analysis and exploits class information that is generally known (but ignored) from the archive of historical batches. Wavelet approximations are shown to provide for a parsimonious representation of the batch profiles. A framework for batch classification that is based on the above discrimnatory learning is proposed to facilitate the task of performance monitoring. The developed methods are evaluated on a Penicillin fermentation process for their ability to monitor and to detect the faults both for real time batch qualification as well as for batch release procedures.**

I. Introduction

BATCH manufacturing is fraught with critical problems that stem from significant batch-to-batch variation in end quality indices. While this variation may appear relatively less important from the perspective of a single stage batch, the impact of these variations get significantly amplified when one considers manufacturing in multi-stage batch processes. Tight control of quality at each of the individual stages would help maximize the overall quality/ productivity of the manufacturing, and will help reduce quality variations and hence batch rejections and product recalls. With the recent PAT initiative by the US-FDA to bring in advances in life sciences manufacturing that adheres to regulatory guidance, as well as continuing interest in enhancing performances in other batch manufacturing areas such as polymers and fine chemicals, batch performance control has been receiving significant academic and industrial interest.

As is well known, by virtue of their finite (but variable) duration batch processes pose additional complexities towards the task of process monitoring, when compared with their continuous counterparts. The multi-way directions of variation in a typical batch archive and the approaches to mine the information present in them using appropriate multivariate statistical tools and matrix unfolding methods have been well documented in literature [7][9]. Furthermore, approaches to accommodate varying batch durations in the above multivariate framework (such as functional space methods [6] and dynamic time warping (DTW) [8], have also been popular. Batch processes also exhibit behavior such as time varying correlation and these has been addressed in literature using staged or multi-regime approaches [6]. The task of online monitoring of batch processes also requires a completed batch record, so that the projections about the batch state can be meaningfully done. Various approaches to addressing this problem are discussed in [9][15]. Yelchuru et al. (2008) have also recently proposed an approach based on similarity assessment of the current batch with those in the archive followed by suitable weighting based on a Euclidean measure of the similarities.

The notion of process monitoring for batch processes has traditionally looked at differentiating an abnormal batch from a normal one by the use of multivariate statistical models. From an operating perspective, the need for continuous and real-time batch qualification, even for a normal batch, is also additionally important. Accurate prediction of end quality indices of the batch at every time step (or at least at key logical time points) during the batch evolution would be expected to help in batch health assessment and also shorten batch release times. Such an assessment, when done during the early time steps of the batch evolution, would also help to initiate corrective control or remedial measures to improve batch productivity. Often times, it is also instructive to compare the relative health of an evolving batch, *i.e.* the batch health relative to the best known or golden batch. As well, from a learning perspective it would be useful to mine the data in the archived database of batch measurements and identify factors that contribute to batches evolving closer to the golden batch.

In this paper, we propose a novel approach based on discriminatory learning of archived batch information. Firstly, we extend the notion of process monitoring, which is relatively restrictive in batch qualification, and propose a new approach based on performance monitoring that explicitly quantifies the relative health of normal batches on a continuum scale, in addition to the abnormal batch flagging. To achieve this task, we propose the use of supervised learning / classification and discrimination analysis that are known to help in simplifying nonlinearities, as well as in detection and resolution of incipient parametric signatures that could be of help in longer term predictions. The task of batch labeling and the classification for the discrimination based learning is based on the end quality indices. Parametric signatures associated with different batches in the archive are then mapped onto these classes to model the overall nonlinear relationships as well as to facilitate the classification of new online batches. The above mapping of the parametric signatures needs the (i) accommodation of unequal batch durations as well as

Manuscript received Sep 27, 2010
Shailesh Patel is with Department of Chemical Engineering, McMaster University, Hamilton, ON, Canada, (E-mail: patelsr@mcmaster.ca).
*Ramprasad Yelchuru, is with the Department of Chemical Engineering, Norwegian Science and Technological University, Trondheim, Norway (E-mail: ramprasad.yelchuru@chemeng.ntnu.no) (Author for correspondence)
Dr. Srikanth Ryali is with the Department of Medical Institute, Stanford University, California, USA. (E-mail: srikanthryali@yahoo.co.uk).
Dr. Ravindra Gudi is with Department of Chemical Engineering, Indian Institute of Technology, Bombay, India (E-mail: ravigudi@iitb.ac.in).

compaction of the information present in the trajectories, and (ii) a reliable approach for completing the batch record online during the batch evolution. To achieve the former, we propose and evaluate a wavelet based signal approximation approach that can accommodate varying time scales or dynamics associated with these trajectories. To address the latter problem, we propose the use of the Gustafson-Kessel fuzzy classification approach to select batches from the archive that are similar to the current batch, and use them for batch record completion. The proposed approach based on discriminatory learning has been validated using simulations involving a nonlinear batch penicillin fermentation.

The remainder of the paper is organized as follows: In the next section, wavelet based functional approximation is presented to address the problem of signal representation. The batch classification framework including batch labeling, representative cluster prototype generation and their use in a supervised fisher discriminant analysis algorithm is presented in section 3. The results with combined approach of wavelet based functional approximation and FDA and the prediction of the batch duration and yield for both offline/online situations are presented in section IV.

II. WAVELET BASED FUNCTIONAL APPROXIMATION

A batch process data record consists of time series variables with nonstationary dynamic trends and sharp discontinuities. For dimensionality reduction, classification and reliable fault detection, the features in these time trajectories need to be parsimoniously represented. Commonly used functional approximation methods such as Legendre polynomials require more number of bases functions to model the non-stationarities and sharp discontinuities in the data accurately. This result in an increase in the dimensionality of feature space which in turn affects the performance of models built for fault detection and classification. In this scenario, data approximation using wavelet bases provides an attractive alternative. Wavelets provides a multi resolution representation of data using which the important features such as sharp discontinuities in the data can be represented using relatively few number of bases functions. Wavelet based methods have been successfully used for applications such as data compression and de-noising [1, 2].

A given time series *x(t)* can be represented in terms of wavelet bases $\varnothing_{\ln}(t)$ as follows

$$x(t) = \sum_{l=1}^{L}\sum_{n=1}^{n_l} w_{\ln}\varnothing_{\ln}(t) \tag{1}$$

where, w_{ln} are wavelet coefficients, L is the number of resolutions and n_l is the number of bases at signal resolution l. The number of resolutions is generally chosen as $L = \lfloor \log_2(T) \rfloor$, where T is the length of the time series and $\lfloor . \rfloor$ is the rounding operator. The wavelet coefficients can be efficiently computed using the two-channel filter bank algorithm [3].

This commonly used wavelet based data approximation method involves representing the coarsest-scale signal first (l=1) and then adding increasingly finer levels of resolution l >1. The approximation error reduces with the increase in the number of levels and the characteristics of the data are better represented. However, to represent sharp discontinuities in the data, which are important features in batch process data, more number of resolution levels need to be added. This results in an increase in the number of wavelet coefficients for data approximation and hence greater feature dimension. Therefore, we need a procedure for parsimonious representation of data which also preserves important characteristics of the data.

Low frequency components of the data are adequately represented by coarser scales while the sharp jumps in the data are manifested in coarser to finer scales. Owing to better time-frequency resolution of wavelet bases, the magnitude of only few wavelet coefficients in finer scales within the neighborhood of this discontinuity will be significant when compared to the rest of the coefficients. Therefore, by retaining only these few significant wavelet coefficients at finer scales along with the coefficients at coarser scales adequately represents the characteristics of the data. We use the procedure suggested in [4] to select wavelet coefficients which parsimoniously represent data under consideration.

The optimum number of wavelet coefficients w_{ln} is to be selected in such a way that both the reconstruction error and the number of wavelet coefficients should be small. To achieve a compromise between these conflicting objectives, we define the following cost function which we refer to here as Relative Reconstruction Error (RRE). RRE is a weighted average of normalized mean square error (nMSE) and the number of coefficients (C) used for reconstruction.

$$RRE(C) = w \times nMSE + (1-w) \times \frac{C}{N} \tag{2}$$

here, w is a weight ($0<=w<=1$) and N is the total number of coefficients. We select C for which RRE is minimum. We select the required wavelet bases for each variable time series of batch process using a training data set of normal operation following the procedure described below.

1. Apply L - level wavelet decomposition [3] of time series in the training data set.
2. For each batch in the training set, select C number of absolute significant coefficients and compute RRE. Repeat this for different values of C. Select that C for which RRE is minimum.
3. The minimum of all the selected optimum C's obtained for all the training data is the optimum C, say C*, for this variable.
4. For each batch, select C* absolute significant coefficients and note their corresponding bases.
5. The union of all the selected bases is the optimum bases set for batch data.

The optimum bases set and the corresponding wavelet coefficients for each variable are then used for building models for fault detection and classification.

III. Batch Classification Framework

The task of batch performance monitoring involves a characterization of the batch health defined in the terms of key quality indicators and placed relative to the health of the best known batch. Towards this objective, we propose a batch classification framework that broadly consists of (i) Information representation, (ii) Discriminatory learning, and (iii) Batch qualification. In the sequel, we briefly describe the individual aspects of this framework.

A. Information representation via Classification and clustering

The information present in the archived database of batch measurements can be represented and exploited in different ways depending on the objective sought. In the current work, we propose to represent the information in a way that facilitates the tasks of classification and discriminatory learning. Figure 1 represents the classification of batches with respect to their end quality indices. For instance, time duration and yield are two important quality indicators for a batch. Using these indicators, batches that belong to Cluster 1 achieve a high yield in relative shorter times and therefore could be considered as the best batches. A similar qualification could be attributed to each of the batches in the archive based on these quality indicators so that the relative health of these batches can be established with respect to the best batch.

In addition to representing the known batches into each of these clusters via batch labeling, it is also important to define the class prototypes or representative batches that best define the behavior within each class/cluster. Generally, such supervised classification approaches recommend using the class mean as a representative prototype. In our work, we propose to additionally use the class members on the boundaries of each class (see Figure 2) to better represent the behavior of each class

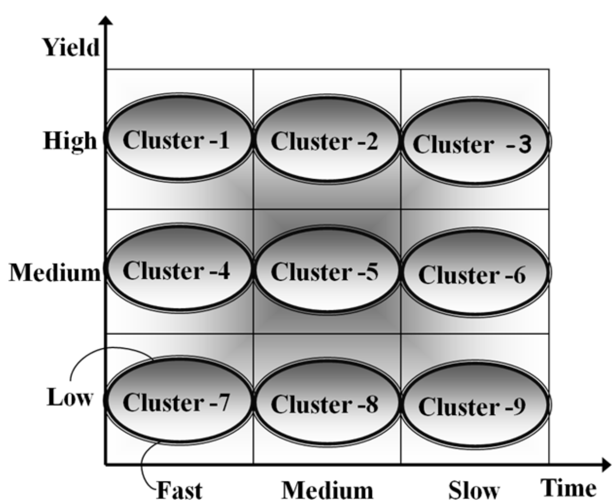


Fig 1. Classification of a batch into clusters based on batch yield and duration

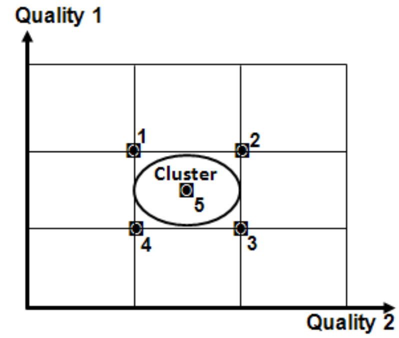


Fig. 2. Cluster prototype definition for the batch classification framework

B. Learning via discriminant analysis

Having represented the performance of a batch in terms of its end quality indicators, the next step is to establish the discriminants or classifiers that help to map the (suitably wavelet transformed) variable trajectories to each of these classes. Towards this objective we propose to use the linear discriminant analysis to build these classifiers as follows:

Considering that (i) the three way matrix of variation is appropriately unfolded along the batch direction [7], and (ii) the time trajectories of the variables are approximated by wavelet coefficients (see section II), we now have the matrix X of these coefficients as row vectors for each batch. Furthermore, assuming that there are c classes of batches in this matrix X, and that a subset of samples in X, which belong to class j is denoted as $X(j)$, the discriminant analysis begins by calculating the. total mean vector $\overline{\mathbf{x}}$ and the mean vector for class j($\overline{\mathbf{x}}\langle j\rangle$) which are given by

$$\overline{\mathbf{x}} = \frac{\sum_{i=1}^{m}\mathbf{x}_i}{m} \text{ and } \overline{\mathbf{x}}\langle j\rangle = \frac{\sum_{\mathbf{x}_i\in\mathbf{X}\langle j\rangle}\mathbf{x}_i}{m\langle j\rangle}$$

where $m(j)$ is the number of batches in class j. Next, The total scatter matrix $\mathbf{S}_t$, the within-class scatter matrix for class j , $\mathbf{S}\langle j\rangle$, the within-class scatter matrix $\mathbf{S}_w$ and the between-class scatter matrix $\mathbf{S}_b$ are defined as

$$\mathbf{S}_t = \sum_{i=1}^{m}\left(\mathbf{x}_i-\overline{\mathbf{x}}\right)\left(\mathbf{x}_i-\overline{\mathbf{x}}\right)^T \tag{3}$$

$$\mathbf{S}\langle j\rangle = \sum_{\mathbf{x}_i\in\mathbf{X}\langle j\rangle}\left(\mathbf{x}_i-\overline{\mathbf{x}}\langle j\rangle\right)\left(\mathbf{x}_i-\overline{\mathbf{x}}\langle j\rangle\right)^T \tag{4}$$

$$\mathbf{S}_w = \sum_{j=1}^{h}\mathbf{S}\langle j\rangle \tag{5}$$

$$\mathbf{S}_b = \sum_{j=1}^{h} m\langle j\rangle\left(\overline{\mathbf{x}}\langle j\rangle-\overline{\mathbf{x}}\right)\left(\overline{\mathbf{x}}\langle j\rangle-\overline{\mathbf{x}}\right)^T \tag{6}$$

Fisher discriminant analysis then seeks to find the discriminants $\mathbf{v}$ that maximize the between class scatter while minimizing the within class scatter using a suitably formulated objective function as,

$$J\left(\mathbf{v}\right) = \max_{\mathbf{v}\neq\mathbf{0}}\frac{\mathbf{v}^T\mathbf{S}_b\mathbf{v}}{\mathbf{v}^T\mathbf{S}_w\mathbf{v}} \tag{7}$$

which results in the solution of the generalized eigenvalue problem given by

$$\mathbf{S}_b \mathbf{v} = \lambda \mathbf{S}_w \mathbf{v} \tag{8}$$

C. Batch classification using probabilistic clustering

During online operation, performance monitoring and prediction in real time requires that the information about the remainder of the batch be completed to form a batch record. In our approach, we propose to use the Gustafson-Kessel fuzzy-probabilistic classification method to find the membership value of an ongoing batch to pre-defined cluster prototypes, which could be then used to construct the future batch evolution as a weighted average of these cluster prototype profiles. We outline the batch classification steps as follows:

Future Data filling and Quality Prediction Method :

Step 1: The membership value of an ongoing batch with respect to its similarity to the cluster prototypes is first evaluated. We consider cluster prototypes as cluster centroid for the Fuzzy GK algorithm to estimate the membership of the ongoing batch in terms of its similarity to the cluster prototype as

$$\mu_{ij} = \frac{1}{\sum_{k=1}^{c} \left(D_{ij,A} / D_{kj,A} \right)^{2/(m-1)}} \qquad (m \neq 1) \tag{9}$$

where, μ_{ij} is the membership value of x_i to cluster prototype j, $D_{ij,A}$ is weighted Euclidean distance from sample x_j to cluster prototype CP_j., m is fuzzy exponent parameter and A is semi positive definite weighting matrix

Step 2: Next, the end product quality indices of the ongoing batch is calculated as a weighted average of the quality indices of the cluster prototype as,.

$$\begin{aligned} E_{Quality1} &= \sum \mu_{ij} BQ_{1j} \\ E_{Quality2} &= \sum \mu_{ij} BQ_{2j} \end{aligned} \tag{10}$$

where BQ_{1j} is the quality 1(e.g. batch duration) of cluster prototype j, BQ_{2j} is the quality 2 (e.g batch yield) of cluster prototype j, and $E_{Quality1}$ is the estimated value of quality 1.

Step 3: The future record of current batch from the current time $t=k$ to predicted batch duration ($E_{Quality}$) is calculated as weighted average of wavelet coefficients of all cluster prototypes. The weighted average wavelet coefficients are inverted to get the time domain profiles of the ongoing batch as,

$$WC_{NewBatch} = \sum \mu_{ij} WCC_j \tag{11}$$

where WC_j is the wavelet coefficients of batch j

$$x_j \big|_{i=k+1:E_{Quality1}} = inv(WC_{NewBatch}) \big|_{i=k+1:E_{Quality1}} \tag{12}$$

where $x_j \big|_{i=k+1:E_{Quality1}}$ is the prediction of variable j from current instant to predicted batch duration($E_{Qualtiy}$)

The next steps in the quality prediction would then involve a wavelet based approximation of the complete batch records as above (Equation 10 -12) to arrive at the feature vector of the evolving batch. This feature vector is then projected onto the FDA directions $\boldsymbol{v}$ (see Equation (8)) to identify the evolving batch cluster.

Figure 3 below shows the schematic of the proposed performance monitoring algorithm.

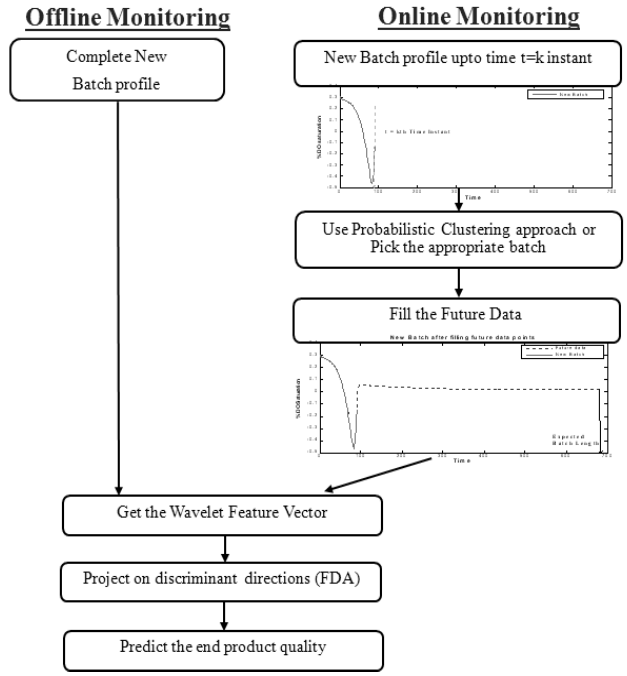


Fig. 3. Batch FDA Algorithm

In the next section, we discuss the results obtained for validation of the proposed algorithm using simulations involving a nonlinear antibiotic fermentation

IV. RESULTS

The proposed method is evaluated on data obtained from simulations involving penicillin fermentation. The basic model equations used for this simulator are taken from Birol et. al. [11] and were simulated in Matlab. The simulations were performed for total of 210 batches and the data was generated by varying the initial conditions assuming a sampling time of 0.5 hrs. The data consisted of time profiles of 8 variables (such as temperature, aeration rate, pH etc.) varying over batch durations of approximately 180 to 325 hours and batch yields of 0.6 g/l to 1.1 g/l, across the batches. In addition to this, 6 faulty batches were also simulated with steps and ramp disturbances of different magnitudes in aeration rate and agitation power.

Out of the 210 batches, 180 batches were randomly selected as training batches and the rest 30 batches were used for algorithm validation. The time domain data of all the variables in all the batches are approximated by wavelet approximation method as discussed in section 3. The data was classified into 5 different clusters based on the end product quality indices (batch duration and yield). Each cluster was defined by cluster prototypes as discussed in previous section. The training batch data were used to build the Batch FDA model and define the control limits for normal batch operation. Figure 4 shows the FDA scores plot for the training batch data. It can be seen that the algorithm is able to successfully segregate the data into 5 distinct clusters.

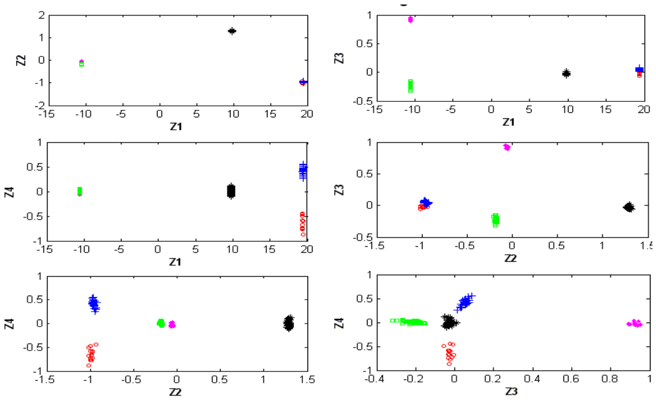
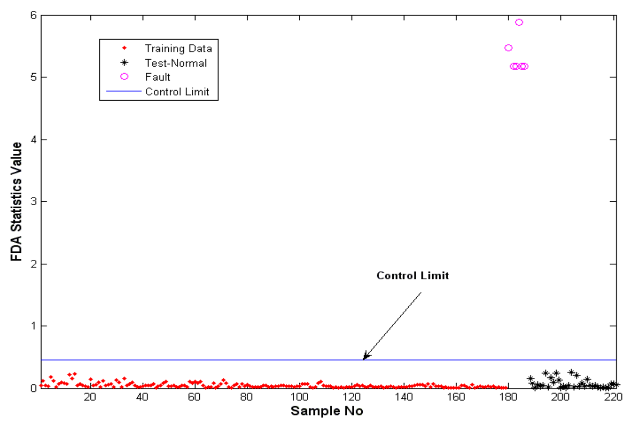


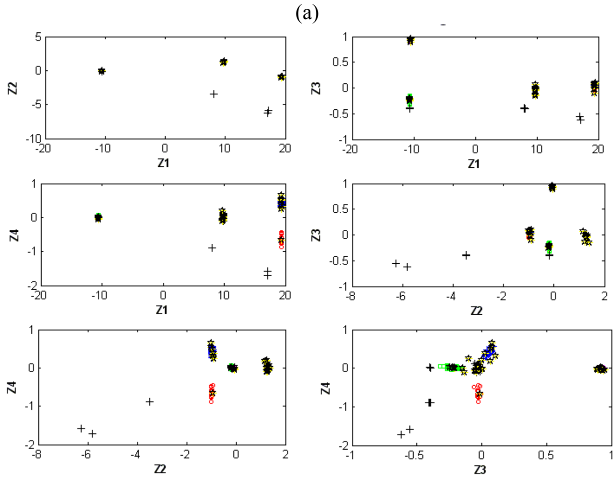
Fig. 4 Scores plot - Training data set

A. *Offline batch process and performance monitoring*

The representative feature vector (wavelet coefficients) for the test batch data are calculated and projected on to the FDA directions to predict the batch health and its quality. As per the expectation, the normal test data set are projected well below the FDA statistical control limit while faulty batches are seen to violate the control limit (Figure 5).

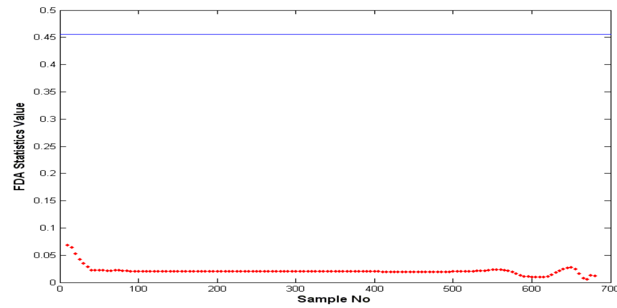


(a)

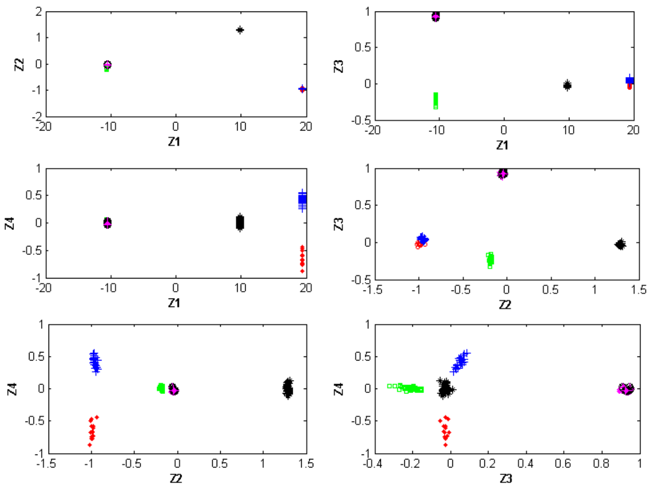


(b)

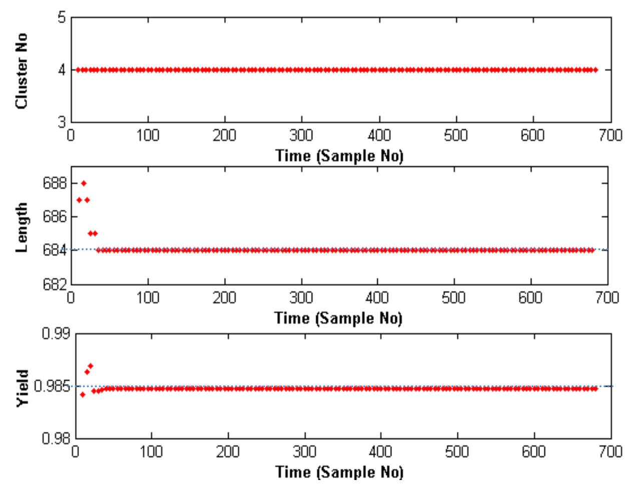
Fig. 5 Offline-monitoring Result (a) FDA Control Chart (b) Scores plot



(a)



(b)



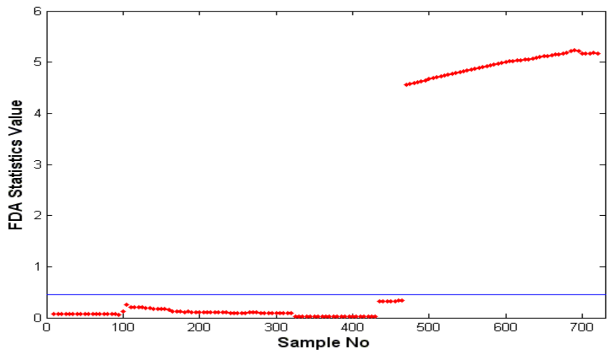
(c)

Fig. 6. The online performance monitoring result (a) FDA control chart (b) FDA scores plot (c) Quality prediction chart

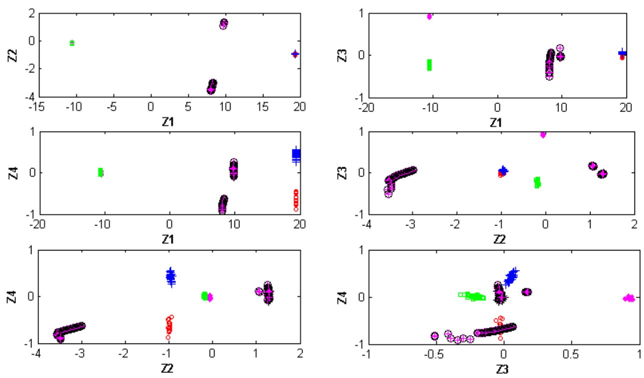
B. *Online batch process and performance monitoring*

The online batch monitoring requires a complete batch record. As discussed in section 3, the partial batch record is filled with future data using cluster prototypes. The time trajectories of variables of the ongoing batch are approximated by wavelet analysis and the data are projected on the FDA discriminant directions.

Figure 6 represents the result of online monitoring and quality prediction of normal test batch. The projection of the ongoing batch at each time instant falls into the cluster 4 which qualifies the ongoing batch to be a normal batch having similar dynamics as batches in cluster 4. The dotted line shows the actual quality value (duration: 342 hrs (684 samples), Yield 0.985 g/l) of test batch It is seen that this end quality prediction are made with high accuracy early during the evolution of the batch.



(a)



(b)

Fig. 7 The online performance monitoring result (a) FDA control chart (b) FDA scores plot

Figure 7 represents the results of online monitoring of a faulty batch. The fault is introduced in aeration rate at 200th sampling instant (100 hrs). The projection of the ongoing batch started to violate the control limit after 460th sampling instant and remains outside the control limit which confirms that the batch is not normal. The FDA scores plot also indicates that the projection of ongoing batch does not fall on any of the normal clusters.

V. Conclusions

A novel approach for online assessment of batch health and batch classification is proposed. Issues associated with time varying correlations and unequal batch records are addressed using an improved wavelet based functional approximation method. The batch data are classified based on the quality indices and discriminant analysis is performed to predict the batch quality. The proposed online performance monitoring approach is validated for both offline and online scenario using simulation involving a nonlinear fermentation case study.

References

[1] Process Analytical Technology Initiative. Available : http://www.niroinc.com/pharma_systems/process_analytical_technology_PAT.asp

[2] Said, A. and W.A. Pearlman, *An image multiresolution representation for lossless and lossy compression.* IEEE Trans Image Process, 1996. 5(9), 1303-10.

[3] Donoho, D.L. and I.M. Johnstone, *Ideal Spatial Adaptation by Wavelet Shrinkage.* Biometrika, 1994. 81(3), 425-455.

[4] Mallat, S.G., *A wavelet tour of signal processing*. 2nd ed. 1999, San Diego: Academic Press. xxiv, 637 p.

[5] Lada, E.K., J.C. Lu, and J.R. Wilson, *A wavelet-based procedure for process fault detection.* IEEE Transactions on Semiconductor Manufacturing, 2002. 15(1), 79-90.

[6] Chen, J., and Liu, J., “Derivation of Function Space Analysis Based PCA Control Charts for Batch ProcessMonitoring”, Chemical Engineering Science, 56, 2001, 3289-3304.

[7] Nomikos, P., and MacGregor, J.F., “Monitoring Batch Processes using Multiway Principal Component Analysis” AIChE Journal, 40(8), 1994, 1361-1375.

[8] Kassidas, A., Macgregor, J.F., and Taylor, P. A., “Synchronization of Batch Trajectories using Dynamic Time Warping”, AIChE Journal, 44(4), 1998, 864-875.

[9] Nomikos, P., and MacGregor, J.F., “Multivariate SPC Charts for Monitoring Batch Processes”, Technometrics, 37, 1995(a), 41-59.

[10] Nomikos, P., and MacGregor, J.F., “Multiway Partial Least Squares in Monitoring Batch Processes”, Chemometrics Intell. Lab. Syst., 30, 1995(b), 97-108.

[11] Birol G., Undey C., Ali Cinar, “A Modular Simulation Package for Fed-batch Fermentation: Penicillin productions”, Computers and Chemical Engineering, 26, 2002, 1553-1565.

[12] Marjanovic, O., Lennox, B., Sandoz, D., Smith, K., Crofts, M., Real-time monitoring of an industrial batch process, Computers and Chemical Engineering, 30, 2006, 1476-1481.

[13] Guidance for Industry PA - A Framework for Innovative Pharmaceutical Development,4527, Manufacturing, and Quality Assurance, US-FDA report 2004.

[14] Yelchuru, R., Patel, S., Ryali, S. and Gudi, R.D., Prediction of batch quality indices using functional space approximation and partial least squares, American Control Conference, Seattle, USA, June 11-13, 2008.

[15] Eric N. M. van Sprang, Henk-Jan Ramaker, Johan A. Westerhuis, Stephen P. Gurden, Age K. Smilde, Critical evaluation of approaches for on-line batch process monitoring, Chemical Engineering Science 57, 2002, 3979-3991