2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

Latent Variable MPC for Trajectory Tracking in Batch Processes: Role of the model structure

Masoud Golshan, John F. MacGregor, Mark-John Bruwer, Prashant Mhaskar

Abstract— **The Multiphase Latent Variable Model Predictive control (MLV-MPC) is developed based on the Principal component analysis (PCA) model. The proposed control methodology is capable of trajectory tracking as well as disturbance rejection. The model that is used in the course of MPC is a multiphase PCA model that is constructed based on the available data from the measurements on the process. Different data arrangements are studied and their effects on the performance of the control algorithm are evaluated.**

I. Introduction

The operation of processes in batch mood brings forth several distinguishing characteristics as opposed to continuous processes. Specifically, batch processes are characterized by a specific beginning and end time over the course of which the process variables often change by significant amounts. Additionally, the process variables at the end of the batch need not be at equilibrium, but rather correspond to desirable product properties. Both of these features have significant impact on the way batch processes are modeled, monitored and controlled. Another difficulty with the control of batch processes is that the nonlinearity and time varying characteristics cannot be ignored as the operating point is changing continuously. On the other hand the acceptable product quality is a tight range that makes the application of a high performance control methodology necessary. Predictive controllers are capable of doing this important job as long as a reliable model of the process is available. For nonlinear batch processes this has usually meant the use of fundamental nonlinear models embedded within an optimization. The problem with this approach is that the modeling effort is large, the computation time is unacceptable, and the solvers are complex. In this paper we present an approach using empirical Latent Variable models that capture the benefits of these nonlinear MPC's without implementation difficulties mentioned above.

Selection of the type of empirical model to be used is important as some of them are nonlinear without an important benefit. Choice of nonlinear empirical models brings forth several concerns such as computation time, convergence uncertainty and nonconvex optimization problem. On the other hand some linear models have a number of shortcomings. They are valid locally and include oversimplified assumptions. The aim is to come up with a model that can have the benefits of a linear model and can avoid the aforementioned drawbacks.

In dealing with empirical models, the only information one can count on is the measurements on the process variables which are collected repeatedly for a number of batches. However, most batches just have historical data on the process operated in normal way and may not have enough information about causal relationships among the variables. Usually, one has the opportunity to generate few batches according to designed experiments. Hence, a small number of tools are in hand to extract the most information out of the process.

The paper is organized as follows: in the next section the new control methodology is briefly explained. Next, the role of different data arrangements and considering multiphase models, in system identification and control is explored. Results from control simulations are then provided to illustrate the main issues involved with different approaches. Finally, the conclusion and recommendations are drawn.

II. Latent Variable (LV) Modeling approach

The data set for a batch process is a cube as the information is distributed in three dimensions (No. of Batches × No. of Variables× Time duration of Batches). In order to make the PCA model a data rearrangement is necessary. Several data arrangement approaches are studied in the literature [4]. A well known approach is batch-wise unfolding as shown in fig.1 where variables of different sample times are placed beside each other. An alternative unfolding approach is Variable-wise unfolding approach in which variables of different sample times are placed underneath each other. The batch-wise unfolding approach can model the nonlinearity and time varying properties of the process, but it requires many observations to build a PCA model. On the other hand, in the Variable-wise unfolding approach one can get the great advantage of building an LV model using a small number of batch runs by considering each time step during a batch as an

Manuscript received September 22, 2008. (Write the date on which you submitted your paper for review.)

(corresponding author to provide phone: 303-555-5555; fax: 303-555-5555; e-mail: author@ boulder.nist.gov).

M. Golshan is a PhD student in the Dept. of Chemical Engineering, Mcmaster University, Hamilton, ON, Canada L8S 4L7 (e-mail: golsham@mcmaster.ca).

J.F. MacGregor was with the Dept of Chemical Engineering, McMaster University, Hamilton, ON, Canada L8S 4L7. He is now President of ProSensus, Inc., Hamilton, ON, Canada L8P 0A1 (email: john.macgregor@prosensus.ca).

M.J. Bruwer is with ProSensus, Inc., Hamilton, ON, Canada L8P 0A1 (email: mj.bruwer@prosensus.ca)

P. Mhaskar is with the Dept. of Chemical Engineering, McMaster University, Hamilton, ON, Canada L8S 4L7. (e-mail: mhaskar@mcmaster.ca).

observation. However, the underlying assumption is that the correlation structure among the dataset does not vary with time and a static average model is enough to explain the process. However, time varying property is quiet common in batch processes. As a result a combination of the variable-wise and batch-wise unfolding approach includes time-lag in the variable-wise unfolded dataset [6]. This approach is called Variable-wise with Time-lag unfolding [1]. Hence, an average dynamic PCA model is obtained for the batch process. During this study two different approaches are analyzed. The first approach is the batch-wise unfolding approach as shown in fig. 1.

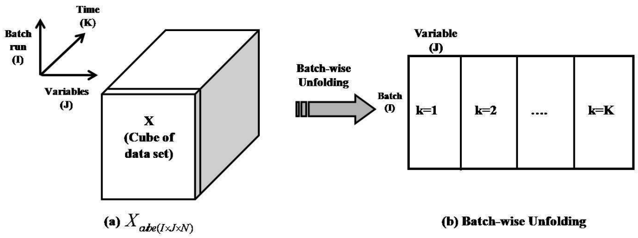

Figure 1- batch process dataset and batch-wise unfolding

The other approach is Variable-wise with Time-lag unfolding approach as shown in fig. 2.

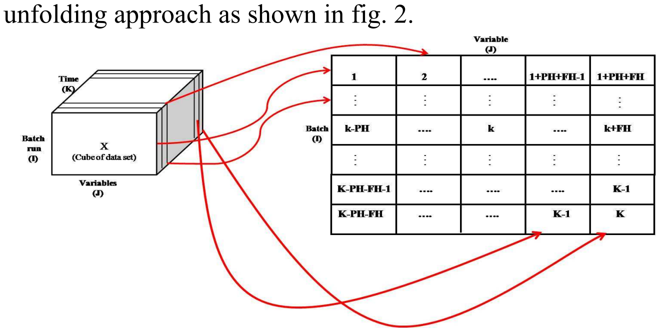

Figure 2- Variable-wise with time-lag Unfolding

If the unfolded matrices of Figs. 1 and 2 are considered as matrix $X_{(a\times b)}$, the PCA model is of the form:

$$X = TP^T + E \tag{1}$$

$$T = X P \tag{2}$$

Where T is a $(a\times A)$ matrix ($A\leq b$) of latent variable scores that summarizes the major differences among the batch trajectories, and P is a $(b\times A)$ matrix of loadings that show how the latent variable scores are related to the trajectory data (X). The score values of the A latent variables for each batch summarize the time varying behavior of its trajectories relative to all the other batches.

III. MULTIPHASE LV MODELING

a. Multiphase Batch-Wise Unfolding

In the batch-wise unfolding algorithm, data on each variable at all time intervals are included in a row. Thus, with mean centering a PCA model [1] is capable of explaining the time varying and nonlinearity characteristics of the batch. However, the resulting unfolded matrix contains many variables (JK variables). One such very large global LV model is less desirable as it requires more latent variables (which implies more batches may be needed in the training set, it leads to more ill-conditioned matrices in the control computations and does not focus as well on the local behavior of the trajectories.

In order to have more resolution on the local variations, the LV-MPC is conducted using a multiphase modeling approach. In this approach some phases are identified along the batch and the single batch-wise unfolded dataset is partitioned in many phases. Then for each phase the LV-MPC algorithm is implemented. This algorithm is regarded as Multiphase LV-MPC (MLV-MPC).

The phases should be selected in a way that the correlation structure among the data of the same phase has the minimum variation and the number of variables in each phase is in balance with the number of observations. For example, variables in the preheating step in a reactor and reaction step are likely to have different behaviors and can be considered as different phases. The better the phases are selected, the fewer principal components are needed for the PCA model.

b. Multiphase Variable-wise with Time-Lag Unfolding

By including time-lags in the variable-wise unfolding, dynamics are introduced into the PCA model. However, the model is still an average model over the range of sample times included in each column and the order of the dynamics increases by increasing the number of lagged data.

The main problem with the time-lagged unfolding approach is that it provides an averaged dynamic model that may decrease the model prediction ability which is crucial in the course of MPC. However, the Multiphase approach can alleviate this problem as well. The phase construction is to find phases along the batch in a way that minimum change of correlation structure is included in each phase. Then, the resulting phases are rearranged according to Variable-wise with Time-lag unfolding approach. Figure 3 shows the schematic of the phase selection.

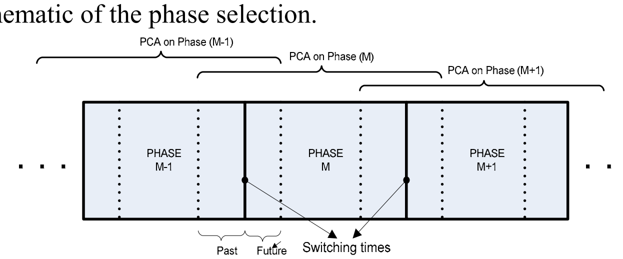

Figure 3- Phase construction and overlapping on a batch dataset

In both unfolding approaches an overlap is considered between two adjacent phases for smooth (bumpless) switching between phases. After determining the borders of the phases, one should use data over as many sample times as the selected model future horizon (fh) from the next phase and the data of as many sample times as the selected model past horizon (ph) from the previous phase as shown in figure

3. Then the current phase must be augmented with these two wings and the PCA model must be build based on the augmented phase. Now the algorithm can switch between phases as soon as the batch reaches the sample time corresponding to the border of the original phases. As a result the algorithm will never face the expanding past horizon and shrinking future horizon except at the beginning and end of the batch, respectively. The values of the *fh* and *ph* depend on the type of the process. The range 10-30 sample times is typical.

IV. Methodology

A. Identification

The training data can be the data from the previous batches run in normal conditions augmented with additional batches executed according to identification experiments to provide information on the causal relationship between the manipulated variables and the controlled variables at every time interval throughout each phase. The direct identification approach based on closed loop data is used in this study. Closed loop identification is preferred over open loop identification for batch processes in order to maintain the process close to its desired trajectories and to minimize the final product quality variations. A PRBS signal was added on top of the manipulated variable trajectories coming from an existing controller (PID) to provide some additional excitation of the process.

B. Prediction

For most linear and nonlinear dynamic models used for MPC, the future prediction is calculated using integration of the dynamic model over the prediction horizon (*fh*) and adapting it assuming a simple random walk type disturbance model is affecting the controlled variable (CV). The prediction step for the PCA latent variable model is accomplished via statistical missing data imputation methods. These methods use all past data up to the current point in time and the time varying batch model to impute the future (missing data) in any batch phase. Several missing data imputation methods have been proposed for latent variable models in the literature [2,3]. The method used in this study is the Trimmed Score Regression (TSR) method. It can be briefly stated as follows:

Any new observation (z) can be divided in two parts as shown in figure 4:

$$z^T = [z^{*T}\ z^{\#T}] \tag{3}$$

Where, z^* is the known data and $z^\#$ is the missing data. For the batch process analysis z^* corresponds to the past data and future setpoints and $z^\#$ corresponds to the future data. The loading matrix can also be divided into two parts in the same way as z.

$$P^T = [P^{*T}\ P^{\#T}] \tag{4}$$

The PCA score is computed based on the assumption that the known part of the data is the complete data in the observation. Thus,

$$\tau^* = P^T z^* \tag{5}$$

The final scores are then estimated by regressing the real scores (τ) from the training data on the fake scores (τ^*). Finally, the score estimation formula is [3]:

$$\tau = \Theta_{1:A} P^{*T}_{1:A} P^{*}_{1:A} (P^{*T}_{1:A} P^{*}_{1:Q} \Theta_{1:Q} P^{*T}_{1:Q} P^{*}_{1:A})^{-1} P^{*T}_{1:A} z^* \tag{6}$$

Where, Θ is the covariance matrix of the scores ($\Theta = (T^T \times T)/I$) in the PCA model, where "I" is the total number of batches in the dataset. The number of scores considered in Θ (Q) can be more than or equal to A.

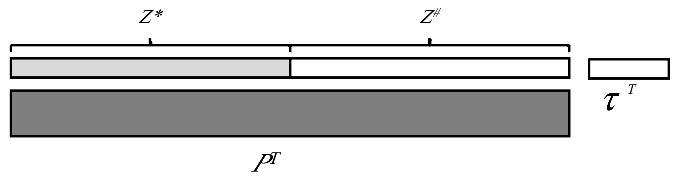

Figure 4- An observation containing missing data and its corresponding PCA model

C. Control

During this study, two control formulations are selected that are elaborated in [1]. They are briefly presented here to maintain continuity.

C.1. Control in Latent Variable Space

The objective of the control is to run a new batch to follow desired trajectories on the CV's. Assume the control algorithm is in the middle (sample time k) of a new batch and ζ_k is defined by equation (7).

$$\zeta_k^T = [x^T_{me,k}, y^T_{cv,k}, u^T_{c,k}, y^T_{sp,k}] \tag{7}$$

Where x_{me}, y_{cv}, u_c, and y_{sp} are measured variables, controlled variables, manipulated variables, and set point variables respectively. The existing information in the current batch can be rearranged as follows:

$$\begin{aligned} x^T &= [\zeta_1^T, \zeta_2^T, ..., \zeta_k^T, ..., \zeta_K^T] = \\ &[\zeta_j^T |_{j=1:k-1}, x^T_{me,k}, y^T_{cv,k}, y^T_{sp,k}\ \ y^T_{sp,j} |_{j=k+1,...,k+PH} \\ &u^T_{c,k}, u^T_{c,j} |_{j=k+1,...,k+CH}, x^T_{me,j} |_{j=k+1,...,k+PH}\ \ y^T_{cv,j} |_{j=k+1,...,k+PH}] \\ &= [x^T_{P1}, x^T_{P2}; x^T_{f1}, x^T_{f2}] \end{aligned} \tag{8}$$

Where, *CH* and *PH* are Control and Prediction (Model) horizons respectively and may be determined according to the general guidelines. The corresponding loadings, *P* matrix in the PCA model, can also be separated in the same way. The past data can be used to estimate the score of the current batch, τ_k, which summarizes the current position of the batch using missing data imputation methods discussed earlier. Then a correction to the score, $\Delta\hat{\tau}_k$ can be computed to bring the batch trajectories closer to their desired values by optimizing the following quadratic objective:

$$\begin{aligned} \min_{\Delta\hat{\tau}_k} & \tfrac{1}{2}(y_{cv} - y_{sp})^T V_1 (y_{cv} - y_{sp}) + u_f^T V_2 u_f \\ &= \tfrac{1}{2}(x_{f2} - x_{P2})^T V_1 (x_{f2} - x_{P2}) + u_f^T V_2 u_f \end{aligned} \tag{9}$$

Comparing equations (9) and (7), y_{cv} and y_{sp} correspond to

x_{22} and x_{12} respectively. Using the PCA model it can be shown that x_{22} and u_f can be written as a function of the decision variable, $\Delta\hat{\tau}_k$.

$$y_{cv} = P_{22}\left(P_f^T P_f\right)^{-1}\left(\hat{t}_i + \Delta\hat{t}_i - P_p^T x_{p,i}\right) \qquad (10)$$

$$u_f = P_{uf}\left(P_f^T P_f\right)^{-1}\left(\hat{t}_i + \Delta\hat{t}_i - P_p^T x_{p,i}\right) \qquad (11)$$

Combining equations (9), (10), and (11) and following optimization procedures, one can obtain the optimum $\Delta\hat{\tau}_k$. $\Delta\hat{\tau}_k$ contains information on the adjustments to all future inputs till the end of the batch ("infinite" horizon control [1]). The corresponding u_f can be computed using PCA model. Then according to MPC algorithm its first element is implemented to the process. At the next sample time the same procedure will be repeated.

C.2. Control in Space of MV's

The data for the current batch can be partitioned in a more explicit way with respect to the manipulated variable:

$$x^T = [\zeta_1^T, \zeta_2^T, \ldots, \zeta_k^T, \ldots, \zeta_K^T] =$$
$$[\zeta_j^T |_{j=1:k-1}, x_{me,k}^T, y_{cv,k}^T, y_{sp,k}^T \quad y_{sp,j}^T |_{j=k+1,\ldots,k+PH}$$
$$x_{me,j}^T |_{j=k+1,\ldots,k+PH} \quad y_{cv,j}^T |_{j=k+1,\ldots,k+PH} \quad u_{c,j}^T |_{j=k,\ldots,k+CH-1}] \qquad (12)$$
$$= [x_{P1}^T, x_{P2}^T, x_{f1}^T, x_{f2}^T, u_f]$$

The main point of this method is to formulate the problem directly in terms of manipulated variables, u_f. In this approach the future manipulated variable will be considered as known (past) information which will be later determined during the optimization process. The scores are again estimated using the trimmed score regression of equation (6). Once the score is estimated, the future controlled variables ($y_{cv} = x_{f2}$) can be estimated from the PCA model as in equation (10)

It is evident that x_{f2} is a function of past data which includes x_{P1}, x_{P2}, and u_f. Thus,

$$x_{f2} = C_{P1}x_{P1} + C_{P2}x_{P2} + C_{uf}u_f \qquad (13)$$

Where C_{P1}, C_{P2}, and C_{uf} are corresponding coefficients that come from combining equations (12) and (13). If the objective function (9) is modified using the new formulation for x_{f2} and replacing $\Delta\hat{\tau}_k$ by u_f as the decision variable, the following objective function is obtained:

$$\min_{u_f} \tfrac{1}{2}\left(y_{cv} - y_{sp}\right)^T V_1 \left(y_{cv} - y_{sp}\right) + \tfrac{1}{2} u_f^T V_2 u_f \qquad (14)$$

By solving the above LQ problem, the optimal u_f over the horizon *CH* will be obtained directly and its first element should be implemented to the process. At the next sample time, the same procedure can be repeated.

It should be noted that in both of the aforementioned methods, it is possible to either solve the optimization problem analytically, if there is no hard constraint, or solve it by numerical optimization methods, in case of existence of hard constraints such as saturation elements using equation (11) in the first control formulation or direct u_f in the second control approach as the hard constraint. However, constraints are generally much less of a problem in batch processes except in the initial start-up phase.

In the above control formulations the LQ matrices (V_1 and V_2) should be chosen carefully. V_1 is a diagonal matrix that can be exponentially weighted to put stress on the early future values rather than the far values. However, V_2 matrix should be a derivative matrix to penalize the changes in the MV's.

V. CASE STUDY

A nonlinear model of a batch reactor is presented in the literature [7],[8]. This is a case study for temperature control problem of a batch reactor. The schematic figure of the reactor is shown in fig. 3.

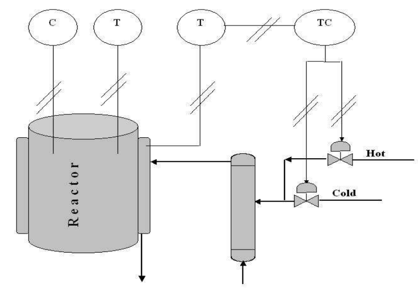

Figure 3- Schematic of the Reactor

The objective is to control the reactor temperature. The manipulated variable is the set point of the jacket temperature as shown in fig. 3. Once the set point is calculated by the controller, by combination of hot and cold water, the desired jacket temperature is generated immediately. However, it takes time for the jacket temperature (T_j) to achieve the T_{sp} (input). It is assumed to be a linear dynamic.

VI. RESULTS AND DISCUSSIONS

In this section the proposed algorithms are implemented on a case study. In this study, PCA modeling approaches based on different data arrangements discussed in section II are analyzed in order to find the methods with the best performance. MLV-MPC on both batch-wise unfolding and variable-wise unfolding with time-lagging are presented.

a. MLV-MPC on Batch-wise Unfolded Dataset

Figs. 4 and 5 show the results of implementation of the proposed algorithm on a batch-wise unfolding dataset.

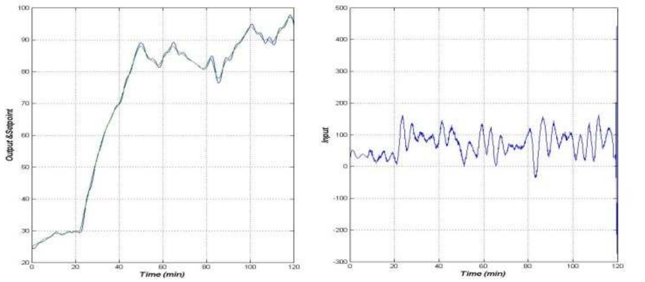

Figure 4- Control based on scores, MLV-MPC on batch-wise unfolded dataset, 6 phases (SP is dashed line)

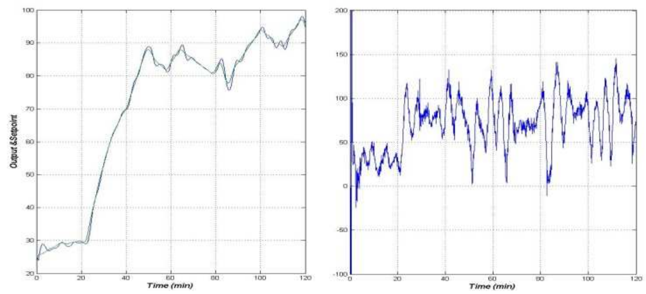

Figure 5- Control based on U, MLV-MPC on batch-wise unfolded dataset, 6 phases

It is of great importance to check the disturbance rejection power of the proposed algorithms. To provide a severe test of the disturbance rejection ability of the batch LV-MPC, a very large additional random walk disturbance was superimposed upon the measured temperature. This study was not intended to represent reality since such a noisy, nonstationary disturbance, unfiltered by passage through any part of the system, and appearing in no other measured variable would probably never occur in practice. The study is intended only as a severe test of the ability of the LV-MPC to eliminate offset. Figs. 6a and 6b show the ability of the LV-MPC algorithm to reject a random walk disturbance superimposed on the temperature.

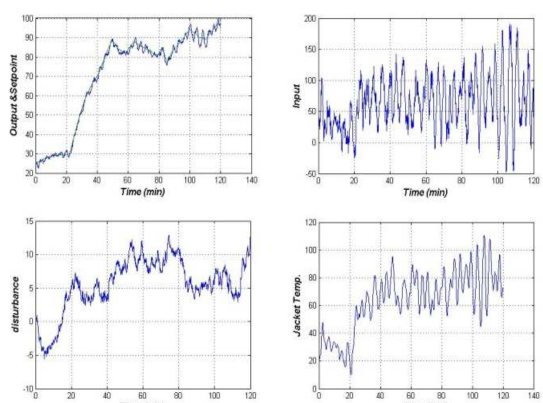

Figure 6a- Control based on scores, MLV-MPC on batch-wise unfolded dataset, 6 phases and disturbance during control (SP is dashed line)

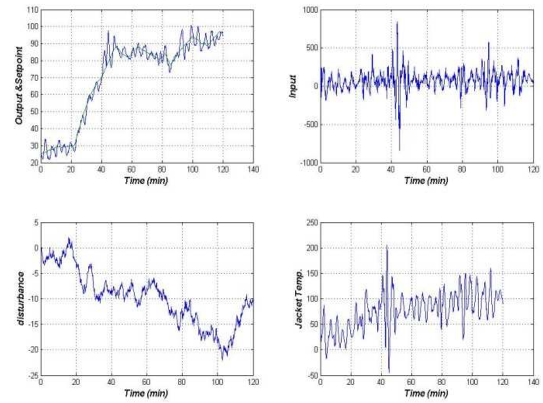

Figure 6b- Control based on U, MLV-MPC on batch-wise unfolded dataset, 6 phases and disturbance during control

It is observed that the LV-MPC both algorithms are able to track the set point trajectory without offset even in case of a nonstationary disturbance. It should be mentioned that the manipulated variable is the set point of the jacket temperature, but the actual controlling variable is the jacket temperature that is in direct contact with the plant. Thus, the jacket temperature is also shown in the above figures.

b. MLV-MPC on Time-Lag Unfolded Dataset

The results of the MLV-MPC on time-lag unfolded dataset are shown in figs. 7 and 8.

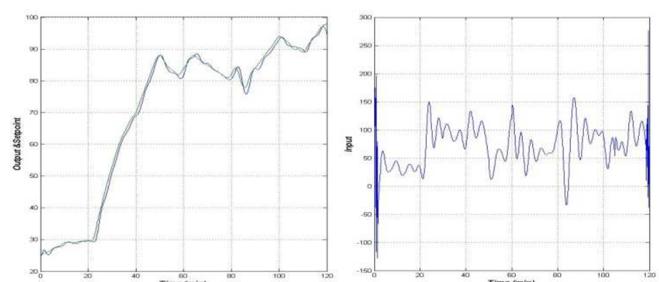

Figure 7- MLV-MPC based on on scores and time-lag unfolded dataset, 5 phases

It is seen that this modeling approach using the MLV-MPC is capable of tracking the set point, with approximately the same performance and even with a smoother manipulated and controlled variables. It is promising because this approach also needs the fewest number of batch identification runs. It should be noted that the multiphase algorithm plays a very critical role in the variable-wise time lagged approach. The reason is that this approach builds a time invariant dynamic model that is assumed to hold over the entire phase. As an example, fig. 9 shows the single phase LV-MPC on the variable-wise time-lag unfolded PCA model with a fast time varying heat transfer coefficient.

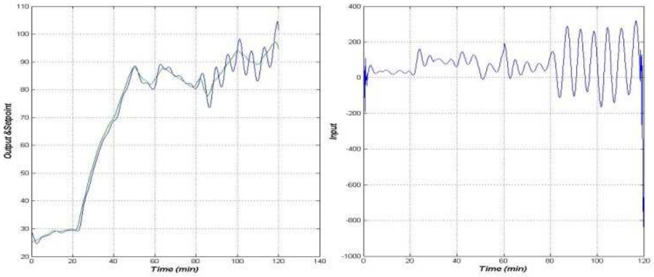

Figure 9- Single phase LV-MPC on the time-lag unfolded dataset

It is evident that the algorithm deteriorates considerably towards the end of the batch where the average model is poor. When a multiphase approach is used this poor late behavior disappears since different models are used for different phases and the time invariant phase models will stay reasonably valid during each phase.

The following figure show the random-walk disturbance rejection for the MLV-MPC on variable-wise with time-lag unfolding approach.

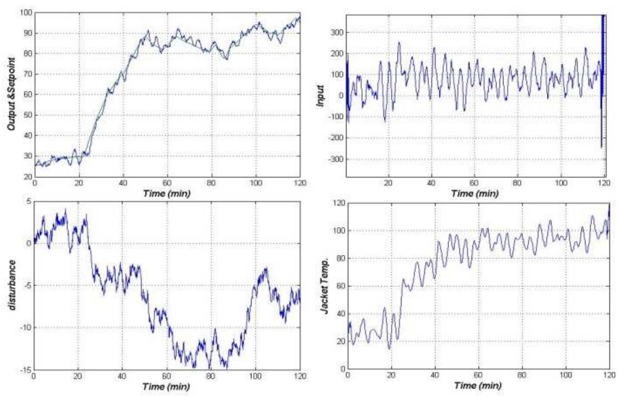

Figure 10- Control based on scores, Multiphase time-lag unfolded dataset, 5 phases, and random walk disturbance in control

VII. CONCLUSION

Several approaches to Multiphase Latent Variable MPC (MLV-MPC) for batch processes have been presented in this paper. In particular, two different approaches for data arrangement are explored in terms of the ability of the resulting PCA model to explain the real underlying process and its prediction power. Each approach has its own benefits and drawbacks. The batch-wise unfolding approach directly accommodates the time varying non-linearities within each phase, but has more stringent requirements for identification. The variable-wise unfolding with time lagging requires less data for identification, but assumes a time invariant linear model within each phase. However, with well chosen phases, both approaches appear to perform well in the MLV-MPC simulations. Two different control methodologies (control in the latent variable space and control in the manipulated variable (real) space) are presented. Both are shown to perform well. Closed-loop system identification issues for developing these batch PCA models are under investigation by the authors.

REFERENCES

[1] M. Golshan, J. F. MacGregor, M. J. Bruwer, "Latent Variable Model Predictive Control", Submitted for publication, Journal of Process control.

[2] P. P. C. Nelson, P. A. Taylor, J. F. MacGregor, "Missing data methods in PCA and PLS: Score calculations with incomplete observations", Chemometrics Intell. Lab. Syst., vol. 35, pp.45-65, 1996.

[3] F. Arteaga, A. Ferrer ," Dealing with missing data in MSPC: several methods, different interpretations, some examples", J. Chemometrics, vol. 16, pp.408-418, 2002.

[4] D. J. Louwerse, A. K. Smilde, "Multivariate statistical process control of batch processes based on three-way models", Chemical Engineering Science, vol. 55, pp.1225-1235, 2000.

[5] P. Nomikos, J. F. MacGregor , "Multiway partial least squares in monitoring batch processes", Chemometrics and intelligent laboratory systems, vol. 30, pp. 97-108, 1995.

[6] J. Flores-Cerrillo, J. F. MacGregor, "Latent Variable MPC for trajectory tracking in batch processes", J. Process Cont., vol. 15, pp. 651-663, 2005.

[7] N. Aziz, M. A. Hussain, I. M. Mujtaba, "Performance of different types of controllers in tracking optimal temperature profiles in batch reactors", Comput. Chem. Eng., vol. 24, pp. 1069–1075, 2000.

[8] B. J. Cott, S. Macchietto, "Temperature control of exothermic batch reactors using generic model control", Ind. Eng. Chem. Res., vol. 28, no.8, pp. 1177–1184, 1989.