# Systematic Methodology for Reproducible Optimizing Batch Operation

S. Bay Jørgensen[a*] and D. Bonné[a]

[a]CAPEC, Department of Chemical Engineering, Technical University of Denmark, Søltofts Plads, 2800 Kgs. Lyngby, Denmark

This contribution presents a systematic methodology for rapid acquirement of discrete-time state space model representations of batch processes based on their historical operation data. These state space models are parsimoniously parameterized as a set of local, interdependent models. The present contribution furthermore presents how the asymptotic convergence of Iterative Learning Control is combined with the closed-loop performance of Model Predictive Control to form a robust and asymptotically stable optimal controller for ensuring reliable and reproducible operation of batch processes. This controller may also be used for Optimizing control. The modeling and control performance is demonstrated on a fed-batch protein cultivation example. The presented methodologies lend themselves directly for application as Process Analytical Technologies (PAT).

## 1. Introduction

The batch processing types covered in this paper includes Batch, Fed-batch and periodic operation which all have the common traits of a repeated operation which start from nearly the same initial conditions. Thus the time within the bacth and the batch number are the two characteristic independent variables. Batch processing is subject to variations in raw material properties, in start-up initialization and other disturbances during execution. These different disturbances introduce variations in the final product quality. Compensating for these disturbances have been difficult in the past due to the nonlinear and time-varying behavior of batch processing and to the fact that reliable on- or in-line sensors for monitoring final product quality rarely are available. Consequently development of a systematic methodology which can ensure reliable reproducible operation may provide significant bennefits for batch processing.

Each batch operation may be defined as a series of operational tasks, i.e. mixing, reaction and separation. Within each task a set of subtasks, e.g. heating/cooling, (dis-)charging is handled. There may be more than one feasible set of operational tasks that can produce the specified product(-s). Consequently an optimal sequence of tasks and subtasks with respect to. a defined objective needs to be identified. This set of operational tasks is labeled the *optimal batch operations model*. Thus the Batch Operations Model combines the batch processing tasks normally specified in a generic recipe with the batch equipment under availability and other resource constraints.

*sbj@kt.dtu.dk

Several research groups have used general empirical model knowledge to develop methods for control of batch processes including an experimental adaptation (optimization) of the Batch Operations Model, labelled "the solution model" through tracking the Necessary Conditions of Optimality (NCO) Srinivasan and Bonvin (2004). Another approach Åkesson *et al.* (2001) exploits the knowledge that optimal batch operation consists of a sequence of operations to determine the presently (most) constrained variable. A data driven approach develops prediction of end of run properties (Flores-Cerrillo and MacGregor, 2003), while asimmilar inspiration lead to reconstruction of approximate time series models from available process data (Gregersen and Jørgensen, 2001; Bonné and Jørgensen, 2003).

This paper presents methodologies based upon a model for batch or periodic operation which can ensure reliable reproducible operation and which may enable optimizing operation. The contribution comprises data driven time series modeling of batch processes and a learning model predictive control methodology. The modelling methodology produces both a Linear Time-Invariant state space model representation for inter-batch prediction and a Linear Time-Varying state space model representation for intra-batch prediction. The modelling approximates the non-stationary and nonlinear behavior of batch processes with a set of local but interdependent linear regression models parameterized as AutoRegressive Moving Average models with eXogenous inputs (ARMAX). Tikhonov Regularization is applied to estimate the parameters of this model set. Learning Model Predictive Control is presented for control of repeated operation of stochastic Linear Time-Varying systems with finite time horizons together with tuning requirements for ensuring guaranteed convergence and hence closed-loop stability. The methodologies have been implemented as a Matlab toolbox Grid of Linear Models (GoLM).

## 2. Methods

Batch processes are modeled with the toolbox GoLM as a sets of $N$ LTI models. Such a set of LTI models could also be referred to as one LTV batch model. These LTI models can be parameterized in a number of ways, but in the present contribution an ARMAX parameterization was chosen. This choice of parameterization offers a simple multivariable system description with a moderate number of model parameters. The objective of the model set is to quantify the causal correlations between the process outputs $y_{k,i} \in \mathbb{R}^{n_y}$, inputs $u_{k,i-1} \in \mathbb{R}^{n_u}$, and distrubances $v_i \in \mathbb{R}^{n_y}$, for $i = 1, \ldots, t$, at times $t = 1, \ldots, N$ in batch $k$. To simplify notation, define the input $\boldsymbol{u}_k$, output $\boldsymbol{y}_k$, shifted output $\boldsymbol{y}_k^0$, and disturbance $\boldsymbol{v}_k$ profiles in batch $k$ as

$$
\begin{aligned}
\boldsymbol{u}_k &= \begin{bmatrix} u'_{k,0} & u'_{k,1} & \ldots & u'_{k,N-1} \end{bmatrix}' \\
\boldsymbol{y}_k &= \begin{bmatrix} y'_{k,1} & y'_{k,2} & \ldots & y'_{k,N} \end{bmatrix}' \\
\boldsymbol{y}_k^0 &= \begin{bmatrix} y'_{k,0} & y'_{k,1} & \ldots & y'_{k,N-1} \end{bmatrix}' \\
\boldsymbol{v}_k &= \begin{bmatrix} v'_{k,1} & v'_{k,2} & \ldots & v'_{k,N} \end{bmatrix}'
\end{aligned}
\tag{1}
$$

The GoLM toolbox models the differences between two successive batches with the ARMAX model

$$
\Delta \boldsymbol{y}_k = \boldsymbol{y}_k - \boldsymbol{y}_{k-1} = -\boldsymbol{A}\Delta \boldsymbol{y}_k^0 + \boldsymbol{B}\Delta \boldsymbol{u}_k - \boldsymbol{C}\boldsymbol{v}_k \tag{2}
$$

where $\Delta \boldsymbol{u}_k = \boldsymbol{u}_k - \boldsymbol{u}_{k-1}$ and $\Delta \boldsymbol{y}_k^0 = \boldsymbol{y}_k^0 - \boldsymbol{y}_{k-1}^0$. The batch ARMAX model (2) may be converted into different representations dependent on the particular application task. If the task at hand is to predict (or simulate) the behavior of a batch before it is started, the following form is convenient

$$\Delta \boldsymbol{y}_k = \boldsymbol{H} \Delta y_{k,0} - \boldsymbol{G} \Delta \boldsymbol{u}_k + \boldsymbol{F} \boldsymbol{v}_k \tag{3}$$

where $\Delta y_{k,0} = y_{k,0} - y_{k-1,0}$. The change in the initial conditions $\Delta y_{k,0}$ can be considered as either an input/control variable or a disturbance. Form (3) above is also convenient for the task of classification (e.g. normal or not) of a batch after it has been completed. Furthermore, the form (3) can be used to determine open-loop optimal recipes in the sense of optimizing an objective for the batch. If the objective is to minimize the deviations $\boldsymbol{e}_k = \left[e'_{k,1}, e'_{k,2}, \ldots, e'_{k,N}\right]'$, $e_{k,t} \in \mathbb{R}^{n_y}$, from a desired Batch Operations Model $\bar{\boldsymbol{y}}$, then (3) can be modified into

$$\boldsymbol{e}_k = \bar{\boldsymbol{y}} - \boldsymbol{y}_k = \boldsymbol{e}_{k-1} - \boldsymbol{H} \Delta y_{k,0} + \boldsymbol{G} \Delta \boldsymbol{u}_k - \boldsymbol{F} \boldsymbol{v}_k \tag{4}$$

The two forms (3) and (4) of the batch ARMAX model above are applicable to off-line or inter-batch type applications. For on-line estimation, monitoring, feedback control, and optimization however, it is convenient to use a state space realization of the batch ARMAX model. In an observer canonical form, which is structurally a minimal realization, the state space realization is given as

$$\begin{aligned} x_{k,t} &= \mathcal{A}_t x_{k,t-1} + \mathcal{B}_t \Delta u_{k,t-1} + \mathcal{E}_t v_{k,t} \\ \Delta y_{k,t} &= y_{k,t} - y_{k-1,t} = \mathcal{C} x_{k,t} \end{aligned} \tag{5}$$

with $\Delta u_{k,t-1} = u_{k,t-1} - u_{k-1,t-1}$ and the initial condition $x_{k,0} = \mathcal{C}' \Delta y_{k,0}$. Just as (3), the state space model form (5) is convenient for prediction, monitoring, and optimization type applications, and it facilitates on-line implementation of such applications. Furthermore, the state space model form (5) is particularly well suited for closed-loop or feedback control applications. For tracking control applications the state space model form (5) can be modified to

$$\begin{aligned} x_{k,t} &= \mathcal{A}_t x_{k,t-1} + \mathcal{B}_t \Delta u_{k,t-1} + \mathcal{E}_t v_{k,t} \\ e_{k,t} &= \bar{y}_t - y_{k,t} = e_{k-1,t} - \mathcal{C} x_{k,t} \end{aligned} \tag{6}$$

In order to use (6) for tracking control, it is necessary to estimate the states based on noisy observations of the outputs. Assume that during a batch ($k$), observations $z_{k,t}$ of the outputs $y_{k,t}$ are collected at times $t = 0, 1, \ldots, N$ and let the optimal estimate of the state $x_{k,t_1}$ in batch $k$ at time $t_1$ given data up to and including time $t_2$ be given as the conditional mean (e.g. obtainable with a Kalman filter) $\hat{x}_{k,t_1|t_2} = E\{x_{k,t_1} \mid \mathcal{I}_{k,t_2}\}$ where the information $\mathcal{I}_{k,t}$:

$$\mathcal{I}_{k,t} = \{z_{k,t}, \Delta u_{k,t}, \mathcal{I}_{k,t-1}\}, \quad \mathcal{I}_{k,-1} = \mathcal{I}_{k-1,N}, \quad \mathcal{I}_{0,-1} = \{\boldsymbol{y}_{-1}, \boldsymbol{u}_{-1}, \boldsymbol{z}_{-1}\} \tag{7}$$

Then the tracking error $e_{k,t_1}$ in batch $k$ at time $t_1$ given data up to and including time $t_2$ is estimated as $\hat{e}_{k,t_1|t_2} = \hat{e}_{k-1,t_1|N} - \mathcal{C} \hat{x}_{k,t_1|t_2}$ where the smoothened estimate of the error profile in batch $k-1$ is given as

$$\hat{\boldsymbol{e}}_{k-1|k-1} = \left[ \begin{array}{cccc} \hat{e}'_{k-1,1|N} & \hat{e}'_{k-1,2|N} & \cdots & \hat{e}'_{k-1,N|N} \end{array} \right]' = \bar{\boldsymbol{y}} - \hat{\boldsymbol{y}}_{k-1|k-1} \tag{8}$$

and $\hat{\boldsymbol{y}}_{k-1|k-1}$ is the smoothened output profile estimate from batch $k-1$ (e.g. obtainable with a Kernel Smoother). To utilize available information during a batch to obtain the best possible tracking performance, the following learning Model Predictive Control formulation

$$\begin{aligned}\{\Delta u_{k,l,t}\}_{l=t}^{N-1} = \arg \min_{\{\Delta u_{k,i}\}_{i=t}^{N-1}} & \left[ \sum_{i=t+1}^{N} \hat{e}'_{k,i|t} Q_i \hat{e}_{k,i|t} + \Delta u'_{k,i-1} R_i \Delta u_{k,i-1} \right] \\ s.t. \quad & \hat{x}_{k,i|t} = \mathcal{A}_i \hat{x}_{k,i-1|t} + \mathcal{B}_i \Delta u_{k,i-1} \\ & \hat{e}_{k,i|t} = \hat{e}_{k-1,i|N} - \mathcal{C}\hat{x}_{k,i|t} \\ & u_{\min i-1} \leq \Delta u_{k,i-1} + u_{k-1,i-1} \leq u_{\max i-1} \\ & y_{\min i} \leq \bar{y}_i - \hat{e}_{k,i|t} \leq y_{\max i}\end{aligned} \tag{9}$$

is solved at times $t = 0, 1, \ldots, N-1$ in batch $k$ and the control sequence

$$\Delta \boldsymbol{u}_k^\star = \left[ \begin{array}{cccc} \Delta u'_{k,0,0} & \Delta u'_{k,1,1} & \ldots & \Delta u'_{k,N-1,N-1} \end{array} \right]' \tag{10}$$

then approximates a closed-loop optimal control sequence. I.e., at time $t$ in batch $k$, (9) is solved based on updated state estimates, and the input $u_{k,t} = \Delta u_{k,t} + u_{k-1,t}$ is implemented the process.

## 3. Case Study: Protein production

The case study considers modeling, control and optimization of the production of a secreted protein in a fed-batch reactor (Banga *et al.*, 1998) as a simple illustrative example of a nonlinear batch process. The cultivation is subject to changing initial conditions, hence the specification and yield of the final product varies significantly, which clearly is undesirable a.o. for downstream processing.

**Process Identification:** To serve as an efficient alternative to the costly development of phenomena-based models of batch processes, this modeling methodology is data-based and a model can thus be relatively readily obtained from (designed) process operation data with the use of the modeling software GoLM. Once the modeling purpose has been defined, the process of identifying a batch process model is fully autonomous. Figures 1 and 2 illustrate the variablity resulting from random variations in initial conditions versus those obtained from PRBS variations in feeding profiles. Both data sets illustrate a large but realistic variability. The nominal model set used for control design is developed from the identification data.

**Control for Reproducibility:** Reproducible operation aims at realizing the same product quality, batch time and/or production yield from batch to batch. Reproducible operation is attractive because it simplifies forecasting final batch time, scheduling of batch units, forecasting load on up- and down-stream processes and/or down-stream processing. The reproducibility of a batch process can be expressed in terms of the deviations between the desired (optimal) operation trajectories and the trajectories realized during operation of the batch process. Optimal reproducibility can thus be formulated as a problem of tracking the optimal batch operations model. The learning MPC algorithm significantly improves the reproducibility the batch process illustrated in figure 3, since it utilizes process information as it becomes available during the batch runs.

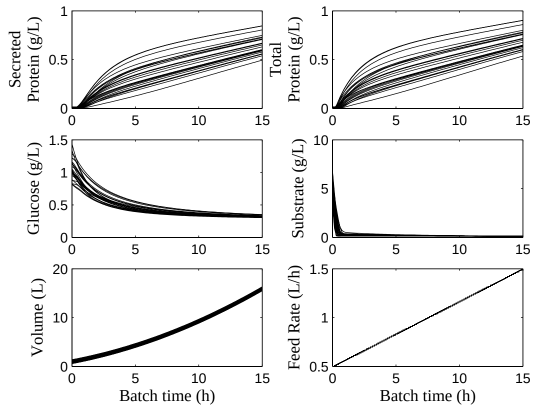


Figure 1. This figure shows the evolution of 25 batch runs on a cultivation reactor producing secreted protein. The cultivation reactor is operated in open-loop and according to the nominal batch operations model. Due to stochastic initial conditions, an undesired variability in the final secreted protein concentration is experienced.

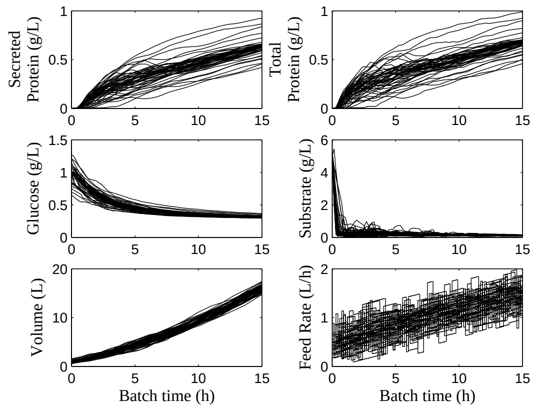


Figure 2. This figure shows the outcome of 50 designed experiments run on a cultivation reactor. The process excitation is generated by a Pseudo Random Binary Sequence (PRBS) dither signal. The cost in terms of lost productivity associated with these experiments is a modest 3% drop in the mean final secreted protein mass.

**Optimizing Control:** Optimizing control is realizing the best possible product quality, batch time and/or production yield in every single batch. Optimizing control is attractive because it autonomously optimizes the product yield and/or production rate and thus the profitability of every single batch unit. Since optimizing control and reproducible operation are implemented with the same process control software, a tractable combination of the two approaches is easily implemented. The obtained result in Figure 4 achieves an increased productivity, however the variabilty is increased compared to the reproducible operation. This increase is mainly due to only using one actuator variable, i.e. substrate feed flow rate. An extra degree of freedom available is substrate feed concentration which would enable reducing the variability further.

## 4. Conclusions

A systematic procedure for development of pieceewise linear model sets for batch operation from operating data have been presented together with a learning model predictive control methodology which may be tuned to achieve reproducible or optimizing operation or some combined objective. The methodlogies have been implemented in a Matlab toolbox labelled GoLM. The presented methodologies hold significant promise to enable data driven predictive control to harvest even more significant benefits in batch processing compared to the success already achieved on many continuously operating processes.

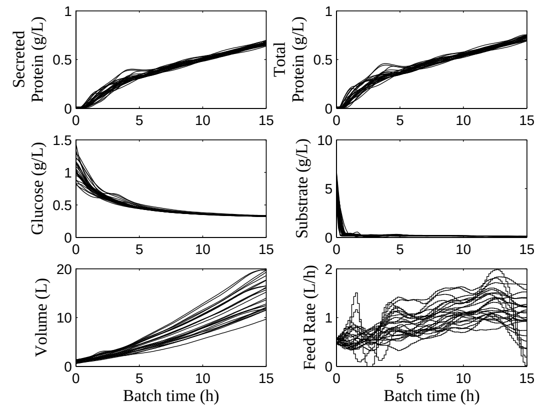


Figure 3. Closed-loop performance of a learning MPC algorithm implemented on a nonlinear batch process with stochastic initial conditions. The control objective is improved reproducibility. The mean productivity achieved is reduced by 4% compared to the nominal case.

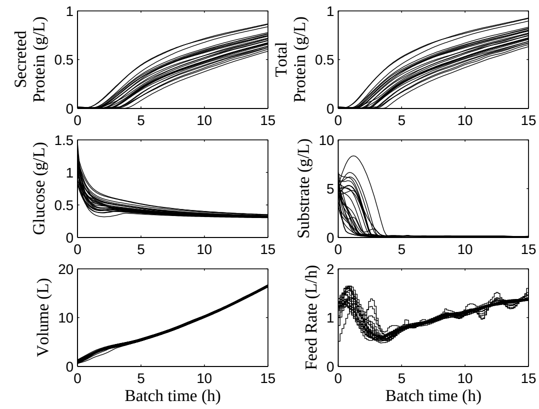


Figure 4. Closed-loop performance of a learning MPC algorithm implemented on a nonlinear batch process with stochastic initial conditions. The control objective is maximizing the final secreted protein mass. The mean protein productivity is increased by 12% over the nominal case.

## References

Åkesson, M.; Hagander, P. and Axelsson, J. (2001). Probing Control of Fed-Batch Cultures: Analysis and Tuning. *Control engng. practice*, **9**, 709–723.

Banga, J. R.; Irizarry-Rivera, R. and Seider, W. D. (1998). Stochastic Optimization for Optimal and Model-Predictive Control. *Computers and Chemical Engineering*, **22**(4/5), 603–612.

Bonné, D. and Jørgensen, S. B. (2003). Data-Driven Modeling of Nonlinear and Time-Varying Processes. In P. V. Hoff, editor, *SYSID*, pages 1655–1660. IFAC, Pergamon press.

Flores-Cerrillo, J. and MacGregor, J. (2003). Within-Batch and Batch- to-Batch Inferential-Adaptive Control of Batch Reactors: A PLS Approach. *Ind. Engng. Chem. Res.*, **42**, 3334–3345.

Gregersen, L. and Jørgensen, S. B. (2001). Identification of Linear Models for Batch Control and Optimisation. In G. Stephanopoulos; J. H. Lee and E. S. Yoon, editors, *6$^{TH}$ IFAC Symposium on Dynamics and Control of Process Systems*, pages 269–274.

Srinivasan, B. and Bonvin, D. (2004). Dynamic Optimization under Uncertainty Via NCO Tracking: A Solution Model Approach. In C. Kipparisidis, editor, *BatchPro - Symposium on Knowledge Driven Batch Processes*, pages 17–35. CEPERI, Thessaloniki.