EXPERIMENTAL INVESTIGATION OF DATADRIVEN MODELLING FOR CONTROL OF FED-BATCH CULTIVATION

Jan Kamyno Rasmussen [*,1] **Sten Bay Jørgensen** [*]

[*] *CAPEC, Department of Chemical Engineering, Technical University of Denmark, Building 229, DK-2800 Lyngby, Denmark*

Abstract: Application of the data driven modelling framework entitled Grid of Linear Models to derive a predictive model for a fed-batch fermentation process is presented. The modelling framework utilises a large number of local linear models to approximate the behaviour of the non-linear process. Industrial data from a fed-batch fermentation has been supplied by Novozymes A/S. A model for the process has been identified based on these data. The model possesses good predictive capabilities and it is intended to be implemented in a model predictive control framework.

Keywords: Datadriven modelling, industrial fermentation, modelling for control

1. INTRODUCTION

Fed-batch processes play a very important role in chemical and biochemical industry. Fermentations are widely used in biochemical industry and these are most often carried out as fed-batch processes. Present control schemes do not utilize the full potential of the production facilities and may often fail to achieve uniform product quality and optimal productivity. Application of advanced multivariable control schemes can help solve this problem. However the introduction of model based control strategies is considered difficult because suitable models are not readily available and require a significant investment in experimental work to develop. First principles engineering models can be used but the usually limited knowledge of the regulatory network in the micro-organism makes them very time consuming to develop. Another strategy is to use a purely data-driven approach where only limited prior knowledge of the process is required. A new methodology for generation of such black-box models has recently been developed (Bonné and Jørgensen, 2003). This method is called *Grid of Linear Models* (GoLM) and it is developed for estimation of models for batch and other periodic operations. The resulting models are large Linear Time Invariant (LTI) models which capture the time varying dynamics of fed-batch processes rather well.

In this paper the GoLM method is used to generate a model for a fed-batch fermentation process for later use in a Model Predictive Control (MPC) framework.

The paper is structured as follows: Section 2 gives an introduction to the industrial process studied. Section 3 gives an overview of the model approach used. Section 4 describes the selection of variables used in the model. Section 5 shows how the modelling identification has been carried out.

[1] Corresponding author.

A validation of the identified model is given in 6. Finally a discussion is given in section 7.

2. PROCESS DESCRIPTION

The process studied is fermentation of the filamentous fungi *Aspergillus oryzae* for production of the enzyme amylase. The cultivation is initiated by inoculation of a seed tank with the desired strain. When a certain criterion has been satisfied, as specified in the recipe, the contents are transferred to the main fermentation tank. The main fermentation tank contains an initial amount of substrate and the main fermentation process starts as soon as the inoculation has occurred. Only the main fermentation will be considered in this paper and will be referred to as the fermentation. The starting time of data sets used in this work correspond to the transfer to the main tank. The fermentation is carried out in two phases, initially as a batch phase and later as a fed-batch phase. Aeration and addition of ammonia takes place during the batch phase and the pH is kept at a constant level. When the initial substrate has been consumed by the microorganisms the fed-batch phase is initiated. Feed dosing is started at a low level and increased to its final value within a certain amount of time. Aeration and ammonia addition continue throughout the entire fermentation. The main objective of the batch phase is to produce bio mass which acts as a catalyst for the production of enzyme which takes place in the fed-batch phase. The process is run a predefined manner according to the given recipe. The fermentors are equipped with sensors for online measurements of different properties and these measurements are constantly monitored by process operators. If one of the monitored variables is outside the operating region specified in the recipe action is taken by the operators and one or more of the inputs to the process are manipulated. Samples are taken from the fermentor at regular intervals for laboratory analysis. These off-line measurements are not used for control purposes but for later evaluation of the batch. The quality of the batch is evaluated as the activity of the produced enzyme in the end of the batch. Successful operation of the fermentation requires that certain conditions are fulfilled during the process. Substrate feeding is necessary for product formation, but a high feeding rate can lead to production of excess biomass which increases the viscosity and decreases the oxygen transfer rate. It is known that production of the enzyme is repressed if the level of dissolved oxygen becomes too low. One of the control problems is thus to maintain the substrate feeding rate within these limits.

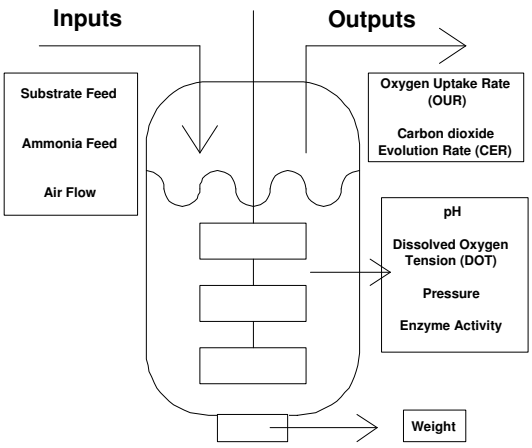

Fig. 1. Sketch of the fermentor. Inputs to process are shown on the left side and outputs are shown on the right side.

2.1 Available data

Production data from an industrial fermentation plant has been made available by Novozymes A/S for use in the BatchPro project. The data set contains data from batches which have been run with the same strain under similar conditions. The basic recipe used is the same for all batches but in some cases operator intervention has been necessary because certain process variables have exceeded the limits specified in the recipe. The data set contains data from a total of 54 batches leading to both satisfactory and unsatisfactory final enzyme activity. 17 of the batches supplied do not contain all the desired variable trajectories and are discarded immediately. Thus only 37 batches have undergone further treatment. By closer inspection 13 of the 37 remaining batches have had problems with one or more of the measured variables. Among these erroneous measurements are: DOT values, gas analysis measurements (CER and OUR) and unreliable enzyme activity measurements. These 13 batches have been discarded to avoid modelling of corrupt data, and only 24 batches are used for actual modelling. The variables in these data sets are given in table 1.

There is a total of 21 variables, where of 17 are available as on-line measurements whereas 4 are only available as off-line measurements. Three of these are sampled every 12 hours but the enzyme activity, which can be considered the quality variable for the process, is only measured once every 24 hours. The on-line variables are sampled every 10 minutes and they all contain 1153 samples, corresponding to a batch length of 192 hours. The Oxygen Uptake Rate (OUR) and Carbon Dioxide evolution Rate (CER) are calculated by the control system installed at the fermentation plant. The system also calculates the

Table 1. Variables available from data set supplied by Novozymes A/S

Variable	Type	Sampling rate
Time	On-line	$6h^{-1}$
Air flow	On-line	$6h^{-1}$
NH_3 flow	On-line	$6h^{-1}$
Accumulated NH_3 flow	On-line	$6h^{-1}$
Dissolved O_2 tension	On-line	$6h^{-1}$
pH	On-line	$6h^{-1}$
Feed flow measured	On-line	$6h^{-1}$
Accumulated feed flow	On-line	$6h^{-1}$
Feed flow set point	On-line	$6h^{-1}$
Back pressure	On-line	$6h^{-1}$
Bottom pressure	On-line	$6h^{-1}$
O_2 uptake rate (OUR)	On-line	$6h^{-1}$
Accumulated O_2 uptake	On-line	$6h^{-1}$
CO_2 evolution rate (CER)	On-line	$6h^{-1}$
Accumulated CO_2 evolution	On-line	$6h^{-1}$
Respirative quotient (RQ)	On-line	$6h^{-1}$
Weight	On-line	$6h^{-1}$
Laboratory pH	Off-line	$2day^{-1}$
Refractive index	Off-line	$2day^{-1}$
Volume percent of mycelia	Off-line	$2day^{-1}$
Enzyme activity	Off-line	$1day^{-1}$

Respirative Quotient (RQ) as the ratio between CER and OUR.

3. MODELLING FRAMEWORK

The framework used for modelling the fermentation data is termed "Grid of Linear Models" (GoLM) (Bonné and Jørgensen, 2003). The approach is purely data-driven and only limited prior knowledge of the process is required. The time span of the entire process is subdivided into grid points, each containing a Linear Time Invariant (LTI) model. This framework has been implemented as a toolbox for Matlab. The methodology is suited for modelling repeated batch processes or other periodic processes. The distance between the grid points can either be constant or dependent on the operating region of the process, ie. reflecting the process dynamics. The current version of the toolbox supports three different linear time series models: Finite Impulse Response (FIR), Auto Regressive with eXogenuos inputs (ARX) and Auto Regressive Moving Average with eXogenuos inputs (ARMAX). The time series models are represented as state space (SS) models as well. In this study only the ARX structure has been applied.

The following time series have been defined:

- Input variables $u_t \in \mathbb{R}^{n_u(t)}$
- Output variables $y_t \in \mathbb{R}^{n_y(t)}$
- Disturbance variables $w_t \in \mathbb{R}^{n_y(t)}$

and corresponding to the input and output variables are their reference trajectories, $\bar{u}_t \in \mathbb{R}^{n_u(t)}$ and $\bar{y}_t \in \mathbb{R}^{n_y(t)}$ respectively. The ARX model parameterisation is applied to describe the output deviation $\bar{y}_t - y_t$ at sample time t as a weighted sum of the past $n_b(t)$ input and $n_a(t)$ output deviations formulated as:

$$\begin{aligned}\bar{y}_t - y_t = &- a_{t,t-1}\big(\bar{y}_{t-1} - y_{t-1}\big) - \dots \\ &- a_{t,t-n_a(t)}\big(\bar{y}_{t-n_a(t)} - y_{t-n_a(t)}\big) \\ &+ b_{t,t-1}\big(\bar{u}_{t-1} - u_{t-1}\big) + \dots \\ &+ b_{t,t-n_b(t)}\big(\bar{u}_{t-n_b(t)} - u_{t-n_b(t)}\big) \\ &+ w_t\end{aligned} \tag{1}$$

Here $n_a(t)$ and $n_b(t) \in [1, \dots, t]$ define the model orders of the local ARX model at each grid point, while $a_{i,j} \in \mathbb{R}^{n_y(i),n_y(j)}$ and $b_{i,j} \in \mathbb{R}^{n_y(i),n_u(j)}$ are the parameters of the local ARX model.

Having an operation with N sampling points, the input $\boldsymbol{u}$, output $\boldsymbol{y}$, shifted output $\boldsymbol{y}^0$, and disturbance $\boldsymbol{w}$ profiles are defined as:

$$\begin{aligned}\boldsymbol{u} &= [u_0^T u_1^T \dots u_{N-1}^T]^T \\ \boldsymbol{y} &= [y_1^T y_2^T \dots y_N^T]^T \\ \boldsymbol{y}^0 &= [y_0^T y_1^T \dots y_{N-1}^T]^T \\ \boldsymbol{w} &= [w_1^T w_2^T \dots w_N^T]^T\end{aligned} \tag{2}$$

The ARX model can then be formulated in matrix form:

$$\bar{\boldsymbol{y}} - \boldsymbol{y} = -\boldsymbol{A}(\bar{\boldsymbol{y}}^0 - \boldsymbol{y}^0) + \boldsymbol{B}(\bar{\boldsymbol{u}} - \ \boldsymbol{u}) + \boldsymbol{w} \tag{3}$$

where $\boldsymbol{A}$,$\boldsymbol{B}$ are structured lower block triangular matrices.

3.1 Batch to Batch Modeling

The disturbance profile $\boldsymbol{w}$ is composed of contributions from a number of sources which can be subdivided into repeated disturbances *e.g.* recipe/input bias, model bias and erroneous readings, and random disturbances *e.g.* process upsets with no correlation between subsequent batches.

By adopting the batch-to-batch approach from Iterative Learning Control (Bonné and Jørgensen, 2003) showed how it is possible to model the disturbance profile as a random walk with respect to the batch index k:

$$\boldsymbol{w}_k = \boldsymbol{w}_{k-1} + \boldsymbol{v}_k \tag{4}$$

where $\boldsymbol{v} = [v_1^T v_2^T \dots v_N^T]^T \in \mathbb{R}^{n_y(t)\cdot N}$ represents the part of the disturbance sequence which is not repeated; assumed to be zero-mean, independent and identically distributed.

The difference between two successive operations then becomes:

$$\begin{aligned}\Delta\boldsymbol{y}_k &= \boldsymbol{y}_k - \boldsymbol{y}_{k-1} \\ &= \boldsymbol{A}(\boldsymbol{y}_k^0 - \boldsymbol{y}_{k-1}^0) - \boldsymbol{B}(\boldsymbol{u}_k - \boldsymbol{u}_{k-1}) \\ &\quad + \boldsymbol{w}_k - \boldsymbol{w}_{k-1} \\ &= \boldsymbol{A}\Delta\boldsymbol{y}_k^0 - \boldsymbol{B}\Delta\boldsymbol{u}_k + \boldsymbol{v}_k\end{aligned} \tag{5}$$

This is a model with an ARX structure, independent of the reference profiles ($\bar{\boldsymbol{y}}$,$\bar{\boldsymbol{u}}$) and of repeated disturbances.

3.2 Model Identification

For the estimation of the model parameters and model orders, a Least Squares (LS) methodology can not be used directly due to the many local models. However an interrelation between models in neighbouring grid points can be expected, especially when time difference between the grid points are shorter than the time constants of the relevant process dynamics. Using Tikhonov regularisation the model parameters can be estimated by solving the extended LS problem:

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{\boldsymbol{\Lambda}} &= \arg\min_{\boldsymbol{\theta}} \left[(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\theta})^T(\boldsymbol{Y}-\boldsymbol{X}\boldsymbol{\theta})\right. \\ &\quad \left. + (\boldsymbol{\Lambda}\boldsymbol{L}\boldsymbol{\theta})^T(\boldsymbol{\Lambda}\boldsymbol{L}\boldsymbol{\theta})\right] \\ &= (\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{X}+\boldsymbol{L}^{\boldsymbol{T}}\boldsymbol{\Lambda}^{\boldsymbol{2}}\boldsymbol{L})^{-1}\boldsymbol{X}^{\boldsymbol{T}}\boldsymbol{Y}\end{aligned} \tag{6}$$

where the penalty $\boldsymbol{\Lambda}\boldsymbol{L}\boldsymbol{\theta}$ is a column vector of weighted differences between parameters in neighbouring grid point models. $\boldsymbol{L}$ being a structured penalty matrix, maps the parameter vector $\boldsymbol{\theta}$ into the appropriate parameter difference where the diagonal regularisation matrix $\boldsymbol{\Lambda}$ finally weights these differences. The interdependency between the grid point models is determined by the choice of structure of the penalty matrix $\boldsymbol{L}$ and the values in the regularisation matrix $\boldsymbol{\Lambda}$ (Bonné and Jørgensen, 2003).

The model orders and regularisation weights can be estimated based on minimisation of the mean square prediction error obtained from using the proposed model on a validation data set. Different model properties can be obtained depending on the choice of prediction horizon for the prediction error *e.g.* one step ahead or pure simulation for the full batch length.

4. VARIABLE SELECTION

The supplied data set contains a total of 21 variables (including time) but not all of them contain useful information for the model estimation, eg. some of the variables may be highly correlated. To facilitate the variable selection Multi-way Principal Component Analysis (MPCA) has been applied (Kosanovich *et al.*, 1994; Gregersen, 1999). 9 of the 24 batches have been selected for the analysis. All available variables except time have been used in the analysis. Common for all batches is that they have no operator interference. The feed flow rate has been maintained at the predefined level and the batches have been run exactly according to the specified recipe, ie. the substrate feed rate is the same for all batches. The purpose is not to model the influence of the feed rate but to investigate which process outputs are connected to the enzyme formation and to derive a control strategy which can control the outputs and indirectly force the process in the desired direction. As the batches are carried out using a predetermined batch operations model, it is expected that the trajectories of the measurements are very similar for all batches. The aim of the analysis is to identify which variables are responsible for the variations in product quality.

The first step in the analysis is to calculate mean trajectories for all the variables for the 9 batches and subtracting this mean trajectory from the actual trajectory. These deviation trajectories thus represent the difference from the mean batch, see figure 2.

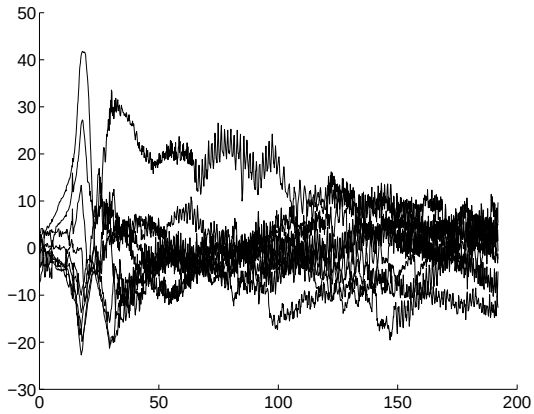

Fig. 2. Difference trajectories for the 9 batches used for the MPCA analysis. Here DOT is shown, the same procedure has been used for all the variables.

The deviation trajectories have been stored in a three-way matrix X ($J \times K \times I$), where J is the number of variables, K is the number of samples from each batch, and I is the number of batches. In this case the dimensions are $20 \times 1153 \times 9$, as time is not included in the analysis. The matrix X is unfolded to a two-way matrix as seen in figure 3. This two-way matrix is called X ($J \times IK$). Each column in X represents a certain variable for all batches and all points in time. Each column is mean centred and scaled to unit variance (auto scaling).

The variance captured by each principal component (PC) is shown in figure 4. The two first PC's capture a relatively large part of the variance and the variance captured by PC 3 and 4 is significantly lower. Another drop appears from the 4th to the 5th PC and from here the decrease is almost linear. The cumulated variance captured by the 4 first PC's is 54.8 %.

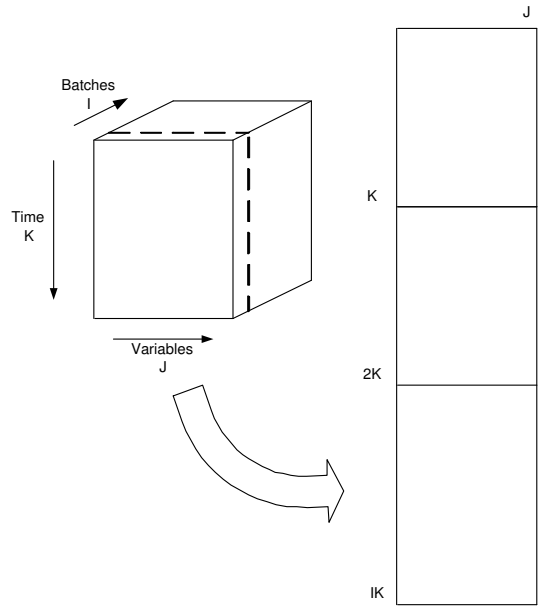

Fig. 3. The difference trajectories have been organised into a 2-dimensional structure and are unfolded by stacking them underneath each other.

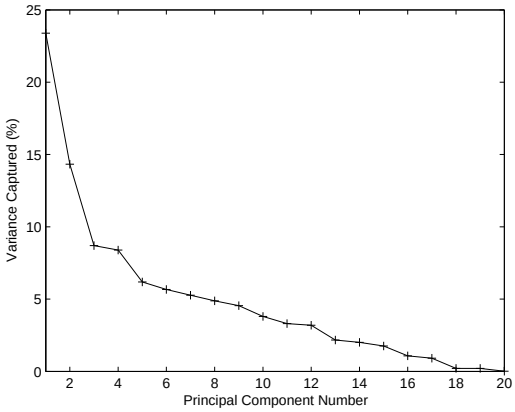

Fig. 4. Variance captured by each principal component.

Figure 5 shows the loadings of each variable on the two first principal components. It is seen that the enzyme activity has a relatively small loading on both PC 1 and PC 2, indicating that it is only responsible for a small explanation of the entire variation of the data. The variables having the highest loadings on PC 1 are accumulated ammonia flow (negative loading), weight, accumulated feed flow and accumulated CER. This shows that a large part of the deviation from the mean batch originates from accumulated variables. This is not surprising as a higher total substrate feed results in a higher weight of the fermentor, a higher conversion of substrate and consequently a higher oxygen demand. The lack of correlation between the accumulated variables and the enzyme activity shows that the product

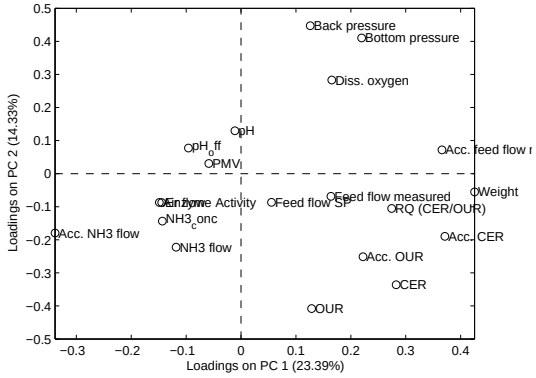

Fig. 5. Loadings of all variables on the two first principal components.

yield, ie. the ratio between enzyme produced and substrate consumed, is not constant.

An interesting observation is that there is a negative correlation between enzyme activity and DOT. This suggests that it is beneficial for the product formation to follow a relatively low DOT trajectory. This hypothesis is supported by investigating PC 3 and 4 (figure 6).

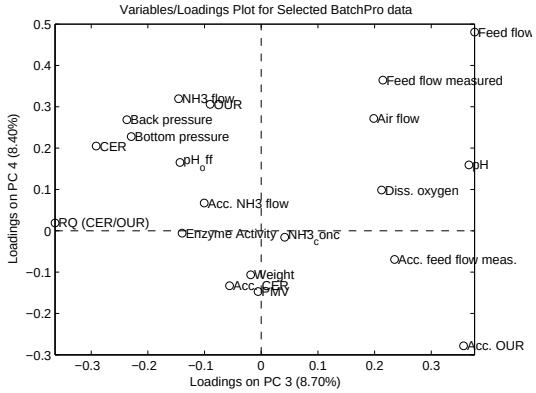

Fig. 6. Loadings of all variables on the third and fourth principal component.

Here it is seen that the enzyme activity has a negative loading on PC 3 and that DOT has a positive loading on the same PC.

The selection of variables for modelling is based on investigation of the loadings of each variable on the first four PC's and on knowledge of the process. The goal has been to keep the number of variables as low as possible but still be able to get good predictive capabilities and capture as much of the variation from the mean batch as possible. To keep the model as simple as possible accumulated gas analysis measurements have not been included. The actual CER and OUR measurements are used instead in order to be able to capture some of the more dynamic phenomena in the system, eg. conversion of a suddenly changing feed flow rate. The only input used in the model is the feed flow set point. This is because it is

the controlled variable having the largest impact on the process and it is intended to use this as the actuator (manipulated variable) in the MPC controller. The variables chosen for the model are given in table 2.

Table 2. Variables used for modelling

Variable	Type
Feed rate	Input
DOT	Output
EnzA	Output
Weight	Output
OUR	Output
CER	Output
Ammonia flow	Output
Air flow	Output

5. MODELLING

The raw batch data has been filtered using a Butterworth filter of 4th order. Different cut-off frequencies have been applied for each variable. The last few hours of the data set have been truncated because the emptying of the fermentor has started at this time for some of the batches. It is not desired to model these phenomena because the actual fermentation process is over at this point. The data has been truncated to a duration of 190 hours. The cut-off frequencies have been tuned carefully for each variable in order to filter away as much noise as possible and still not lose too much information about the process dynamics. The number of variables has been kept at a minimum in order not to make the model too complicated and only one input (to be used as the manipulated variable in the controller) has been chosen.

The 24 batches have been divided into 16 modelling and 8 validation batches. Both modelling and validation batches have had operator intervention on the feed flow set point.

The maximum order for each output/output and input/output relationship has been set to 4.

112 grid points have been used to model the process from t=0 hours to t=190 hours. For the first 60 hours the grid point spacing has been set to 1 hour in order to cover the very dynamic behaviour of the batch phase. For t=61 hours to t=190 hours the grid point spacing has been increased to 2.5 hours as the process is less dynamic in the fed-batch phase.

6. MODEL VALIDATION

The orders of the individual ARX models representing each relationship between inputs and outputs are given in table 3. The model orders do not have a direct physical interpretation but it can be noted that the relation between the feed flow set point (input) and all the outputs have high model orders, supporting the fact that variations in feed flow play a fundamental role for the process.

The resulting GoLM model can be represented as a State Space (SS) model. This representation has been used for the simulations. State estimation is carried out at each grid point by using the actual measurements obtained at that point in a Kalman filter. An initial estimate of the system state is obtained by estimating the initial state (t=0) using the initial measurements available at this time point. The pure simulation (PS) is made by using only this initial estimated state and all future states are calculated based on this. The one step ahead (OSA) simulation is obtained by using the state estimated at each grid point to predict the outputs at the next grid point.

The 8 validation batches have been simulated using the model developed from the 16 modelling batches. Results for one of these simulations is given in figure 7 to 14. Figure 7 shows the input signal used for the simulation and the input signal which has been applied to the reference batch. It is seen that the reference input reaches a constant level after around 40 hours. The input for the simulated system (actual input) also reaches this level but is manipulated further by intervention from process operators after around 70 hours.

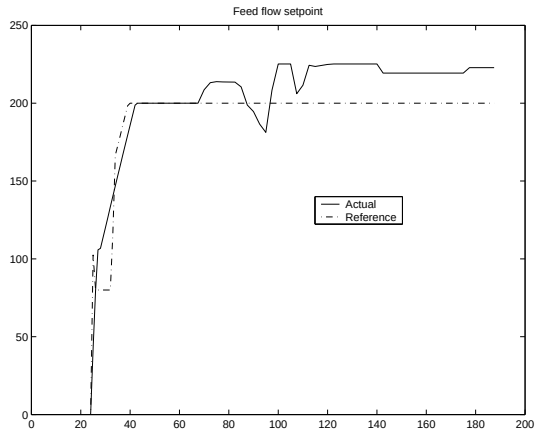

Fig. 7. Feed flow set point. Actual: Simulated trajectory. Reference: Trajectory used in the reference batch.

The enzyme activity (figure 8) is simulated rather well by the model. The reference is significantly lower than the actual and it is seen that the model is able to capture the change in process conditions and predict a higher enzyme activity than the one for the reference batch. The pure simulation follows the reference trajectory closely up to around 35 hours. It then increases and follows the measured trajectory. The pure simulation of the enzyme activity is especially important for control

Table 3. Model orders for the identified ARX models. Combination of outputs and inputs (rows) onto outputs (columns).

	DOT	EnzA	Weight	OUR	CER	NH_3 flow	Air flow
DOT	1	1	2	1	1	0	4
EnzA	2	1	0	0	4	0	0
Weight	0	0	4	4	4	1	1
OUR	0	1	4	1	1	4	0
CER	1	4	4	1	1	2	0
NH_3 flow	1	0	3	4	1	4	4
Air flow	3	0	0	1	0	4	3
Feed flow set point	4	4	4	4	3	2	4

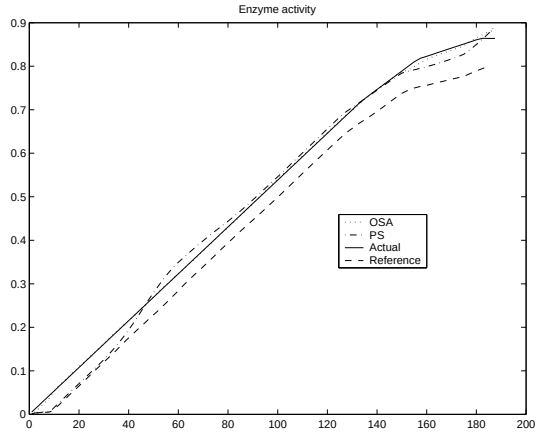

Fig. 8. Enzyme activity. OSA: One step ahead prediction. PS: Pure simulation. Actual: Simulated batch. Reference: Reference batch.

purposes because no measurements are available during the process to give information on the actual activity.

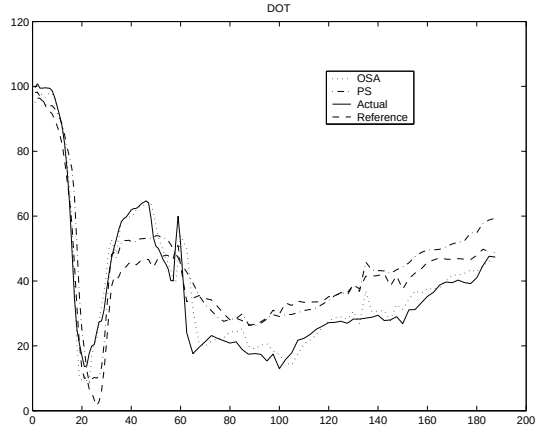

Fig. 9. Dissolved oxygen tension. OSA: One step ahead prediction. PS: Pure simulation. Actual: Simulated batch. Reference: Reference batch.

Figure 9 shows the prediction of the dissolved oxygen tension. The actual trajectory deviates somewhat from the reference and the pure simulation predicts something in between for most of the simulation. The one step ahead prediction follows the actual trajectory very well. The pure simulation is less important in this case than for the enzyme activity because the measurement is available on-line but the MPC controller still needs a reliable model for prediction of the future behaviour of the process.

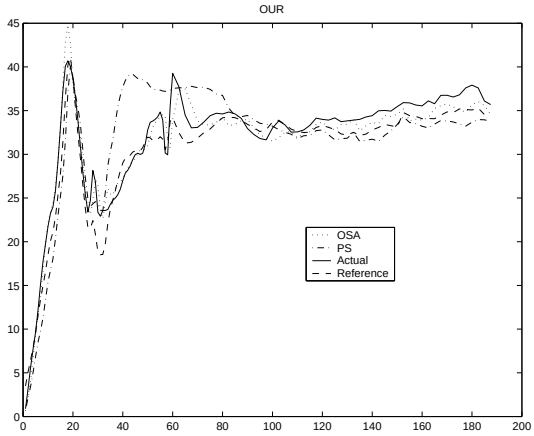

Fig. 10. Oxygen uptake rate. OSA: One step ahead prediction. PS: Pure simulation. Actual: Simulated batch. Reference: Reference batch.

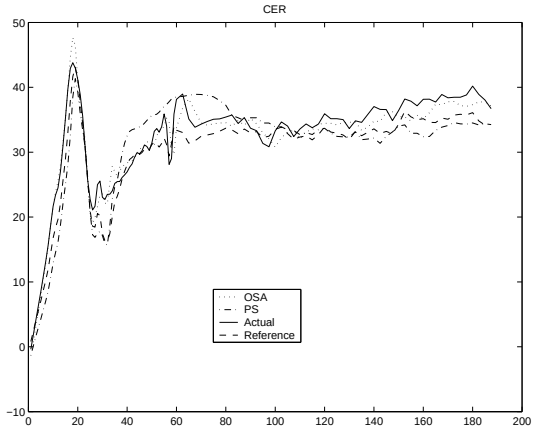

Fig. 11. Carbon dioxide evolution rate. OSA: One step ahead prediction. PS: Pure simulation. Actual: Simulated batch. Reference: Reference batch.

Figure 10 and 11 show the predicted OUR and CER respectively. The simulations are quite good but some problems occur for the pure simulation from 30 to 80 hours for both variables.

The ammonia flow rate is shown in figure 12. Here the model is able to predict the trajectory very well, both for one step ahead and pure simulation.

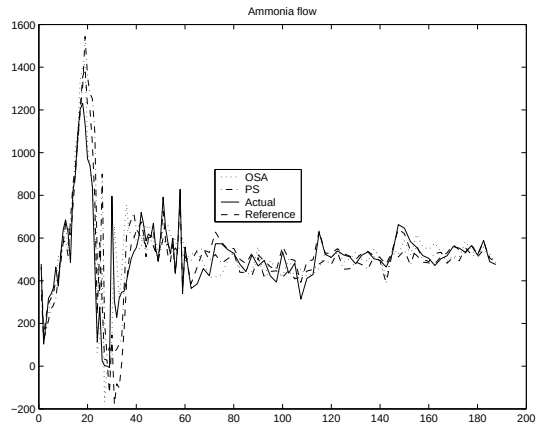

Fig. 12. Ammonia flow rate. OSA: One step ahead prediction. PS: Pure simulation. Actual: Simulated batch. Reference: Reference batch.

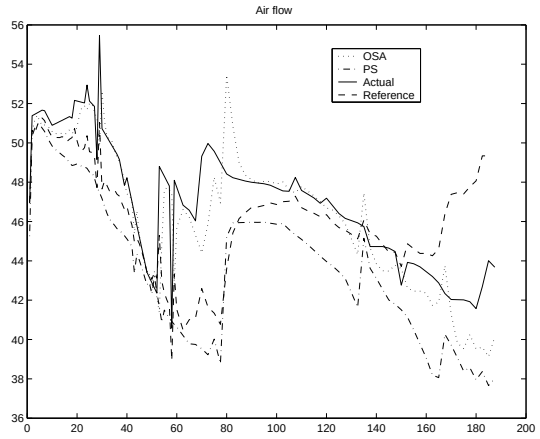

Fig. 13. Air flow rate. OSA: One step ahead prediction. PS: Pure simulation. Actual: Simulated batch. Reference: Reference batch.

The prediction of the air flow rate (figure 13) is quite poor, mainly because it is caused by disturbances of the pressure in the air supply line which can not be modelled. The reason for modelling this variable is to incorporate information on the amount of air supplied to the system in the model.

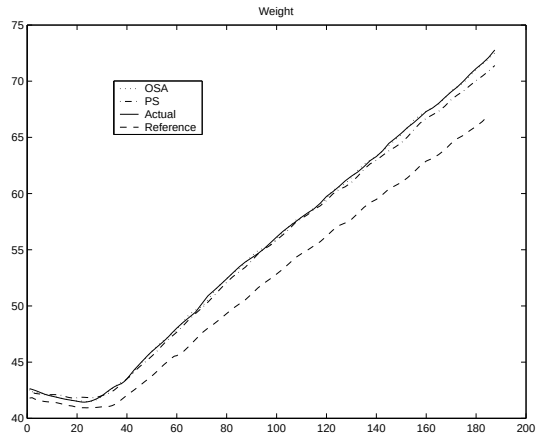

Fig. 14. Weight. OSA: One step ahead prediction. PS: Pure simulation. Actual: Simulated batch. Reference: Reference batch.

The weight of the fermentor (figure 14) is predicted very well which is due to the fact that this variable is easy to model. The variation in weight is to a large extent due to the accumulated feed flow but substrate conversion and evaporation of water also plays a role.

7. DISCUSSION

The GoLM modelling framework has been applied to model industrial production data from a fed-batch fermentation process. The modelling framework divides the entire batch duration into a large number of grid points and local linear models are fitted to each of those points. The combination of all these models is able to approximate the behaviour of the entire process. The variable selection has been facilitated by application of multi-way principal component analysis on the data, identifying the parameters having the largest impact on the process dynamics. The developed model generally possesses good predictive capabilities and the model structure makes it well suited for implementation in a MPC framework.

8. ACKNOWLEDGEMENT

This work has been carried out as a part of BatchPro, an European Research and Training Network. Contract No. HPRN-CT-2000-0039.

REFERENCES

Bonné, Dennis and Sten Bay Jørgensen (2003). Data-driven modeling of nonlinear and time-varying processes. In: *Proceedings to IFAC Symposium on System Identification*. pp. 1655–1660.

Gregersen, Lars (1999). Supervision of fed-batch fermentations. *Chemical Engineering Journal* **75**(1), 69–76.

Kosanovich, K.A., M.J. Piovoso, K.S. Dahl, J.F. MacGregor and P. Nomikos (1994). Multi-way pca applied to an industrial batch process. *American Control Conference, 1994* **2**, 1294–1298.