Model Predictive Feeding Rate Control in Conventional and Single-use Lab-scale Bioreactors: A Study on Practical Application

O. Grigs,[a,b,*] V. Galvanauskas,[c] K. Dubencovs,[a,b] J. Vanags,[d] A. Suleiko,[a,b] T. Berzins,[d] and L. Kunga[a]

[a]Latvian State Institute of Wood Chemistry, Laboratory of Bioprocess Engineering, LV-1006 Riga, Latvia

[b]Riga Technical University, Department of Chemical Engineering, Riga LV-1048, Latvia

[c]Kaunas University of Technology, Department of Automation, LT-51367 Kaunas, Lithuania

[d]Biotehniskais Centrs, JSC, LV-1006 Riga, Latvia

A developed solution for fed-batch process modeling and model predictive control (MPC), facilitating good manufacturing practice (GMP) based on process elaboration, control, and validation, is presented in the paper. The step-by-step evolution of the so-called "golden batch" optimal biomass growth profile and its control during the process is demonstrated. The case study of an advanced fed-batch control was performed on the recombinant *E. coli* BL21 lab-scale (5.4 L) biomass production process using the conventional stirred tank glass reactor. Additionally, a test experiment for control reproducibility and applicability assessment of the proposed approach was carried out in a single-use stirred tank reactor (5.7 L). Four sequentially performed experiments are demonstrated as an example for desirable feeding profile evolution for *E. coli* BL21 biomass production in a glucose-limited fed-batch process. Under different initial biomass and glucose conditions, as well as for different reference feeding profiles selected in the explorative experiments, good tracking quality of preset reference trajectories by the MPC system has been demonstrated. Estimated and experimentally measured biomass mean deviations from the preset reference value at the end of the processes were 4.6 and 3.8 %, respectively. Biomass concentration of 93.6 g L^{-1} (at 24 h) was reached in the most productive run. Better process controllability and safer process run, in terms of avoiding culture overfeeding but still maintaining a sufficiently high growth rate, was suggested for the process with biomass yield of 79.8 g L^{-1} (at 24 h). Practical recommendations on the approach application and adaptation for fed-batch cultures of interest are provided.

Key words:
fed-batch, process reproducibility, model predictive control, model adaptation, bioreactors, Matlab

Introduction

In biotechnology, fed-batch processes are important when microorganism cultivation productivity can be increased by controlled substrate feeding rate. Satisfactory realization and performance of such processes is usually a challenge for technologists, who are confronted with new processes, lack sufficient experience in fermentation process modeling and/or lack sufficiently advanced equipment. One of the problems is the selection of a feeding rate profile that could maintain a microbial growth trajectory yielding in high product titers. Secondly, due to process disturbances, an undesirable, "overfed" or too limited microbial growth may occur.

This may lead to costly batch discard if no adequate immediate feeding rate correction by the operator or automatic control system takes place.

These processes, if implemented in commercial production, especially those in the pharmaceutical industry, need to follow strict good manufacturing practice (GMP) requirements regarding the fermentation process reproducibility for consistent batch-to-batch product quality. Moreover, prior to product commercialization, validation and potential risk assessment should be conducted for production process control[1]. From these requirements, one could state that the process under development should be well studied and precisely controlled on the basis of available and reliable on-line and off-line data. In this case, process automation and implementation of advanced knowledge- and model-based ap-

*Corresponding author: Oskars Grigs, e-mail: oskars.grigs@edu.rtu.lv

proaches could not only facilitate more safe and robust process control but also more easy validation and frequent re-validation procedure, due to minimized dependency on qualified personnel availability and potential risks related to the human-factor. Therefore, fed-batch processes that are sensitive to limiting substrate level require enough sophisticated feeding control approaches. At the same time, the latter should be relatively easily implementable and not too complicated.

While the first digital control units (DCU) for fermentation process control were introduced in the 1960's, many of the established techniques for fed-batch fermentation process control that are used today, such as on-line respiratory quotient (*RQ*) calculation for glucose feed-rate control or on-line estimation of biomass and substrate concentrations using O_2 and CO_2 measurements in off-gas, were introduced already in the following 20 years[2,3]. Further development of DCU capabilities lead to the expansion and the practical implementation of advanced fed-batch control techniques[4,5] such as adaptive control, soft sensing (process models, artificial neural networks (ANN), principal component analysis (PCA)), intelligent control (fuzzy logic, ANN, hybrid models), and model predictive control (MPC)[6] introduced for pilot-scale and industrial applications[7,8,9]. On the other hand, these systems are not widely accessible due to their specialized and non-commercial use, whether developed for the research laboratory or the biotech company. While investigating recently available information from the companies producing/developing bioprocess equipment, besides the common advanced dissolved oxygen (*p*O_2) adaptive control by substrate feeding, some other advanced fed-batch control examples could be found. Infors HT (Switzerland) offers the software package Iris 6, capable of feeding rate control on the basis of *RQ* calculations from off-gas analysis. New Brunswick/Eppendorf (USA) presented feed control based on the on-line dissolved CO_2 sensor data[10]. The YEWMAC Line Computer System (Yokogawa Electric Corporation, Japan) at the beginning of the 1990's was used for model predictive feeding rate control in the glutathione fermentation to compensate deviation from preset (modeled) ethanol concentration[7].

MPC has an advantage of relatively easier implementation possibilities as compared to ANN or fuzzy logic-based control. In MPC, standard and relatively simple process models can be used for controlled parameter calculation in real time and in the feed forward manner. Model re-identification during the process is desirable for MPC, if varying process conditions or phases cause significant estimation errors. Re-identification of parameters, such as biomass yield (Y_{xs}) and maximal specific substrate consumption rate (σ_{max}), is suggested[6]. For model re-identification, reliable methods should be used for on-line or off-line biomass (*X*), substrate (*S*), and volume (*V*) estimation. Precision of *V* estimation has significant influence on biomass and substrate modeling results. Here, several methods for culture volume on-line estimation can be used, for example, reactor direct weighing, using level sensors, calculation of volume on the basis of available information of pumped volumes and modeled flows to and from reactor, etc.

For optimal process performance, a feeding profile is applied to control the substrate consumption rate near its critical value, thus maintaining maximal biomass growth. At the same time, it is taken into account that the substrate consumption rate should not exceed critical value, at which significant synthesis of an inhibitory by-product (e.g., acetate) starts. The accumulated concentration of acetate could inhibit biomass growth. Therefore, to obtain a more suitable process run for safer control, a slightly lowered predefined substrate consumption rate may be applied, which does not reduce process yield significantly[11]. Model-based process optimization is desirable for such cases[12].

In this paper, the authors demonstrate a developed tool for model-based fed-batch control and feeding profile selection, implemented in a laboratory bioreactor system that is commercially available for a wide range of users. The developed and implemented MPC system was used to track the selected reference trajectory of the biomass amount. It demonstrated good control performance and reduced process variability as compared to the system with an open-loop feeding profile control. Practical recommendations for desirable feeding profile selection and control utilizing this tool are provided. Example of *E. coli* biomass production process is demonstrated. For applicability demonstration, the proposed approach is also demonstrated on a laboratory bioreactor of different configuration, belonging to the class of single-use bioreactors.

Materials and methods

Setup overview

The proposed approach was implemented and tested in a lab-scale bioreactor system EDF-5.4/BIO-4 (Biotehniskais Centrs, Riga, Latvia). An additional test run using the proposed approach was performed on a system consisting of a single-use bioreactor (SUB) CellVessel 5.7 (CerCell, Holte, Denmark) connected to BIO-4 controller. EDF-5.4/BIO-4 consists of a 5.4 L glass reactor with working volume of 2–4 L, two standard Rushton turbines, outlet gas condenser, and bioprocess controller BIO-4. Be-

sides standard measurement and control capabilities of temperature, pH, pO_2, foam and level, there is a possibility to measure O_2 and CO_2 concentrations in off-gas and culture turbidity for process on-line analysis and control. The SUB polycarbonate bioreactor with a total volume of 5.7 L and working volume of 2–4.5 L has four standard Rushton turbines and an outlet gas condenser. Some specifications of both the reactors are summarized in Table 1.

Table 1 – *Specifications of EDF-5.4 and CellVessel 5.7 reactors*

Characteristic	Value 'conventional'	Value 'SUB'
Vessel material	Glass (borosilicate), stainless steel (316 L)	Poly-carbonate
Vessel inner height (m)	0.32	0.35
Vessel inner diameter (m)	0.14	0.15
Turbines (number)	2	4
vessel fill till 1st turbine (L)	0.7	0.8
vessel fill till 2nd turbine (L)	2.0	1.4
vessel fill till 3rd turbine (L)	–	2.1
vessel fill till 4th turbine (L)	–	2.9
Turbine blades (number)	6	6
Turbine blade area (cm²)	3.4	0.8
Baffles (number)	3	3
Baffle height (m)	0.28	0.34
Baffle width (m)	0.018	0.017

The BIO-4 DCU has 3 digital inputs/outputs, 4 analog inputs/outputs, and 1 relay input unit (Siemens AG, Germany). The implemented router (Tosibox, Oulu, Finland) supports RJ-45 10/100 M auto-negotiation connection. All Ethernet connections are based on OSI 7 layers. Wireless is protected with WEP, WPA and WPA2 protection keys. The router uses 1024 bit encryption key for connection authentication. The PC SCADA – Supervision Control and Data Acquisition system was connected to the DCU through the router via Ethernet link. The communication diagram of the control system is shown in Fig. 1. Programming in Matlab (Mathworks, Inc.) .m code was used for implementation of the proposed MPC control algorithm. The on-line data exchange of the process and control variables between the MPC control algorithm and SCADA, programmed in PcVue (PcVue Solutions, Ltd.), was implemented every second (s) through an OPC server. The interval of one second was also used for process and control variable data exchange between
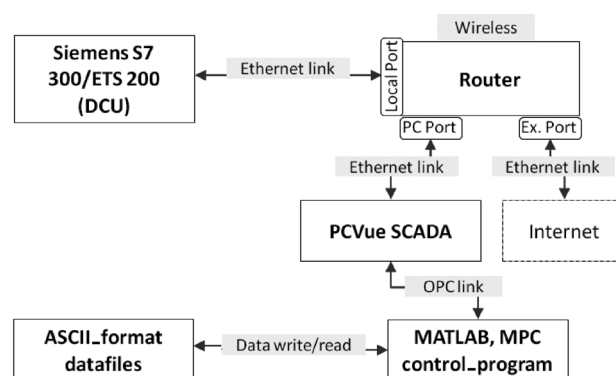

Fig. 1 – *Hardware implementation and communication diagram of the control system*

SCADA and DCU of the bioreactor. Additional data records from the MPC algorithm were written into ASCII format text files.

The process contains all the standard measurement/control equipment and controllers, which are correctly tuned using, e.g., internal model control (IMC) tuning rules, and do not negatively influence the performance of the MPC system discussed in this paper. The pump for glucose feeding (the feeding rate F_s is the main control variable of MPC) contains an internal PID controller for the substrate dosing that is set up at the factory, and needs no additional tuning, as it follows the set feeding rate profile with sufficient precision and quality.

Experimental conditions

The cultivation processes were carried out using recombinant *Escherichia coli* BL21. The batch cultivation and feeding media solutions used in this study were prepared according to R. Bajpai[13] where the batch solution was saturated by salts, glucose 5 g L^{-1}, and yeast extract 5 g L^{-1}. The feeding solution consisted of salts and glucose, where glucose concentration was 400 g L^{-1}. 100 mL of overnight shake flask pre-culture (14–16 h, optical density (*OD*) was 4.0–4.5) grown in batch media was used for inoculating the fermenter. Fermentations were started as batch cultures, and continued as fed-batch when MPC controller activated feeding to follow the predefined reference growth trajectory. The pH was controlled at 7.0 ± 0.2 using 30 % sodium hydroxide and 20 % sulfuric acid solutions. Temperature was kept at 37.0 ± 0.2 °C. The pO_2 was controlled by increasing stirrer rotation speed to the allowed maximum, and then enriching inlet air with pure oxygen. Constant air or air/oxygen mixture flow of 2.0 g min^{-1} (air at standard conditions) was maintained during all processes. The condenser was used for humidity condensing from outlet gases. The foam level was controlled by adding antifoam A (Sigma). Off-line sample measurements of biomass and glucose were made every 30–60 minutes.

Cell growth was observed by measurements of the *OD* at wavelength of 560 nm (Jenway, 6300, Essex, England). Biomass concentration was calculated by multiplying the measured *OD* by correlation coefficient of 0.4 determined experimentally in advance. Glucose was measured enzymatically (AccuChek ACTIVE, Roche, Basel, Switzerland).

Process modeling

Differential mass balance equations (eqs. 1–3) for biomass, substrate, and volume modeling in a fed-batch process were used. Specific substrate consumption rate (σ) (eq. 4) and specific biomass growth rate (μ) (eq. 5) are the state variables of the most notable influence on biomass growth modeling. For substrate consumption and biomass growth modeling, mechanistic terms for substrate (glucose) limitation as well as substrate and biomass inhibition were chosen. Substrate limitation and inhibition in batch/fed-batch *E. coli* cultivation processes are extensively investigated phenomena, described by many authors (e.g., Xu *et al.*[14]). The applied biomass inhibition term Kx_{max} comes from the logistic law[15]. In general, the model applied in this research may be considered a simplified version of the detailed model published by Xu *et al.*[14] augmented by the term that accounts for biomass inhibition.

$$\frac{dX}{dt} = \left(\mu_s - \frac{F}{V} \right) X \qquad (1)$$

$$\frac{dS}{dt} = -\sigma X - \frac{F}{V} S + \frac{F_S}{V} S_f \qquad (2)$$

$$\frac{dV}{dt} = F - F_{smp} \qquad (3)$$

$$\sigma = \sigma_{max} \frac{S}{K_s + S} \cdot \frac{K_{i,s}}{K_{i,s} + S} \cdot \left(1 - \frac{X}{Kx_{max}} \right) \qquad (4)$$

$$\mu = \sigma \cdot Y_{xs} \qquad (5)$$

where X – biomass concentration, g L^{-1}; S – glucose concentration, g L^{-1}; μ – biomass specific growth rate on substrate, h^{-1}; V – culture volume, L; F_s – feeding rate of substrate solution, L h^{-1}; F_{smp} – sampling rate, L h^{-1}; S_f – glucose concentration in feed, g L^{-1}; t – process time, h; Y_{xs} – biomass yield on substrate, g g^{-1}; K_s – Monod constant for substrate uptake, g L^{-1}; $K_{i,s}$ – substrate inhibition constant, g L^{-1}; Kx_{max} – biomass inhibition constant, g L^{-1}; σ – specific substrate consumption rate, g g^{-1} h^{-1}; σ_{max} – maximal specific substrate consumption rate, g g^{-1} h^{-1}. $F = F_s + F_b - F_c - F_e$, where F_c – carbon loss with the off-gas, L h^{-1}; F_b – alkali addition rate, L h^{-1}; F_e – evaporation rate, L h^{-1}.

The modeling procedure is a significant aspect for particular method application for fed-batch process control. First of all, the modeled trajectories of X, S, and V are selected as references. These trajectories are followed during MPC process control by appropriate correction of the feeding rate profile. Secondly, during the running process, the model is used for correction of the feeding rate, leading to achievement of a biomass growth profile as close as possible to the selected reference.

Prior to starting the modeling, some literature review should be made on significant model parameters Y_{xs} and σ_{max}. If information about the model parameters is inconsistent, average initial parameter values for $Y_{xs} = 0.5$ g g^{-1} and $\sigma_{max} = 1.0$ g g^{-1} h^{-1} may be used as an initial guess in the process under consideration. In the analyzed process, the substrate inhibition constant ($K_{i,s}$) and the biomass inhibition constant (Kx_{max}) are the parameters of minor priority. If the limiting substrate inhibition impact or biomass concentration influence on substrate consumption/biomass production is not observed or it is insignificant under current cultivation conditions, for model simplification, the terms $\frac{K_{i,s}}{K_{i,s} + s}$ and $\left(1 - \frac{X}{Kx_{max}} \right)$ may be excluded from the model eq. 4. It was shown that more simple models in many cases lead to comparable process optimization results[16].

In order to model culture volume, the flows into and out of the reactor should be taken into account. For typical processes, these flows consist of sampling, evaporation, and carbon loss in off-gas rates, as well as alkali, acid, antifoam agent, and feeding addition rates. A more detailed explanation on the equations for mass flow modeling for prediction and their parameters can be found in[12].

Process modeling was performed in Matlab environment. For the integration of differential equations, the standard Matlab ODE solver *ode15s* (a variable-order method for solving stiff differential equation systems) was used.

Reference profile selection

In the case of not yet experimentally evaluated/fitted fed-batch model and its parameters, it is advisable to initially select a reduced feeding profile, which, in case of insignificant substrate accumulation, could be increased step-by-step and modified for process productivity improvements. Model parameters during this procedure may be identified by model fitting to currently available experimental data. After model parameter identification is performed, the process operator may choose a desirable biomass growth profile or growth under

a defined specific growth rate, under which the model indicates no significant or potentially productivity-inhibiting substrate accumulation. The latter case is discussed in the results section (see section Evolution of desirable reference biomass and feeding profiles).

Biomass estimation and on-line feed-forward process control

For model re-identification and comparison of the estimated process biomass (XV_{est}) versus reference profile (XV_{ref}), and performing appropriate feeding rate control action, off-line biomass measurement and volume estimates were used. Volume estimation was made using calibrated peristaltic pump signals of alkali, antifoam and feeding pumps, carbon loss based on the off-gas analysis, empirically evaluated evaporation rate of 1.5 g h^{-1}, and sampling rate of 15 g h^{-1}. Acid addition was not used in this particular process.

After the entering of off-line biomass and sugar sample concentrations, and subsequent identification of Y_{xs} and σ_{max} values, the algorithm calculates 40 possible feeding rates around the pre-optimized profile, and selects the one that gives the closest XV_{est} fit to XV_{ref} in a one-step-ahead prediction. The control horizon interval corresponds to the time between the points in the feeding profile. In the explored experiments, this time was 60–120 minutes.

Process control

Optimal feeding rate evolution and control approach concept description

The implemented approach refers to the evolution and control of the so-called "golden-batch" biomass growth trajectory, desirable for a particular microbial cultivation process (see schematic overview in Fig. 2). For the first experiment, the operator needs to choose the biomass, substrate and volume profiles to be set as references. To calculate such reference profiles by means of a process model, it is necessary to know the numerical values of the model parameters, such as biomass yield from substrate (Y_{xs}), maximal substrate consumption rate (σ_{max}), (approximately) maximal achievable biomass (X_{max}). Usually, at the initial stage of fed-batch process development, this information is missing or it is of limited precision. This information about the model parameter values is continuously updated as the accumulated experimental database becomes more extensive, and the model parameter identification can be performed with higher precision.

The authors have developed a process-modeling tool (program), which allows modeling the process and selecting reference profiles. In the program, the reference profiles for modeled concentrations and initial feeding profile were selected. The process productivity (higher biomass amounts achieved) may be further increased within the model by increasing the values of particular feeding rate points until minimal substrate accumulation appears. And vice versa, if significant substrate accumulation is
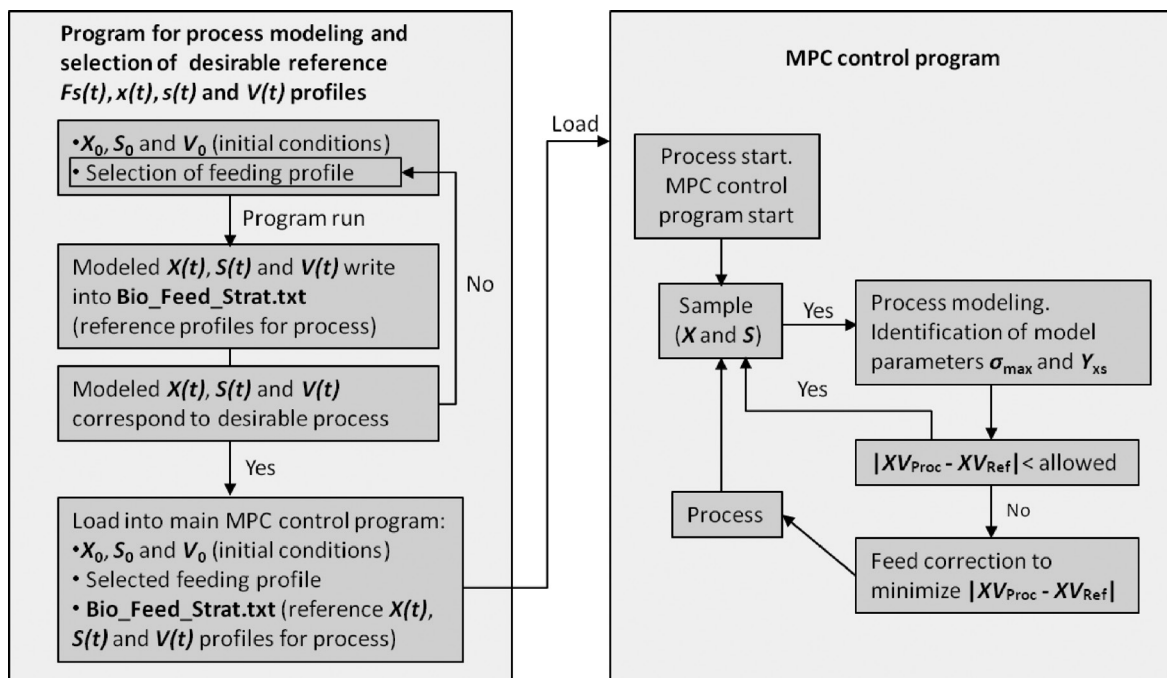

F i g . 2 – *Essential steps in preparation of the reference profiles and MPC algorithm*

observed in the modeling results, then the appropriate feeding profile points are decreased in order to avoid significant substrate accumulation in that particular phase of the modeled process. The latter action may be repeated from process to process until maximal biomass or product yields are reached. Later, the reference profiles from the most productive process may be used as "golden-batch" profiles for a desirable process run. After selection of the modeled reference trajectories and starting the real process with MPC (see section Process modelling), automatic correction of the selected reference feeding profile takes place in order to follow the process biomass reference trajectory XV_{ref}.

MPC controller and feeding rate control algorithm

The MPC principle[6] implemented in BIO-4 is based on feeding rate control to maintain the process biomass profile XV_{proc} as close as possible to the set reference biomass profile XV_{ref}. The reference profiles of $X(t)$, $S(t)$, and $V(t)$ for each particular process correspond to the modeled trajectories and are evaluated during modeling procedure (see section Process modeling). Later in the process, this model is fitted to the running experimental data to select a feeding rate for the achievement of preset reference biomass. The generalized form of the biomass dynamic sub-model directly used in the MPC algorithm is shown in eq. 6. The objective of the developed MPC system (its cost function) is defined as a tracking problem (see eq. 7) – i.e., minimization of the difference between the reference and the predicted biomass amount profiles by means of the main control variable (glucose feeding rate) variation within the constraints shown in eq. 8.

$$\frac{dX}{dt} = f(t, X(t), F_s(t), \Theta) \tag{6}$$

$$J(F_s(t), X(t_0)) =$$
$$= \sum_{t_i = t_1}^{t_H} [X_{pred}(t_i) V_{pred}(t_i) - X_{ref}(t_i) V_{ref}(t_i)]^2 \to \min \tag{7}$$

$$F_{s,\min} \leq F_s(t) \leq F_{s,\max} \tag{8}$$

where J – control cost function, g; X – biomass concentration, g L^{-1}; X_{pred} – predicted biomass concentration, g L^{-1}; X_{ref} – modeled reference biomass concentration, g L^{-1}; V_{pred} – predicted culture volume, L; V_{ref} – modeled reference culture volume, L; F_s – feeding rate of substrate solution (control variable), L h^{-1}; $F_{s,\min}$ – minimal feeding rate of substrate solution, L h^{-1}; $F_{s,\max}$ – maximal feeding rate of substrate solution, L h^{-1}; t – process time, h; $t_0 - 0$, h (start of the process); t_i – process time at the point i, h; t_1 –

initial time moment for MPC, h; t_H – control horizon time, h; Θ – vector of the tunable model parameters.

A model of the similar structure and MPC technique for Chinese hamster ovary (CHO) cell culture MPC were applied in research of Aehle et al.[17] Kuprijanov et al.[6] applied MPC for biomass concentration tracking. Further applications in chemical and refinery industrial processes were reported by Yüzgec et al.[18]

In the case investigated, the MPC non-convex problem is solved. In general, problems of this type are hard to solve and there are no guarantees on the exact optimum reached and solving time. Nevertheless, the solving time is not a critical issue in the case analyzed, as the process model parameter re-identification and calculation of the control variable values for the next optimal control horizon (that equals to approx. 0.5–1 hour) takes no more than 1–2 minutes in total, and does not cause a significant delay. The guarantee of reaching optimum within the practically required tolerances is limited only by the precision of the process model and the precision of biomass measurements. The actual deviation between the reference value and the measured value may also be influenced by the error produced in a particular measurement point.

Taking into account the measurement techniques used, the biomass and glucose measurement precision is approx. 5 %. After model parameter identification, the residuals of the modeled trajectories and measured data were uniformly distributed around zero-mean, and the modeling RMSD was varying at approx. 3–5 % depending on the process phase. This led to an unbiased estimation of the modeled trajectories with a precision that was comparable to the reference measurements and acceptable from the practical application point of view. As the initial reference profile of the control variable (glucose feeding rate) is provided as a result of the model-based off-line optimization of the process, the MPC calculations involve the correction of the control variable value around the reference point. This calculation is performed by means of a variation around the working point by using a scanning method, because the optimal control variable value for the control horizon is well predictable, and in practical applications usually varies within the range of 20–30 % of the initial reference value.

After model parameter identification (model fitting to experimental data), the further routine procedure for process modeling and reference profile selection is addressed to defining the initial process conditions (X_0, S_0 and V_0) and selecting a feeding profile that will lead to achievement of desirable biomass growth and limiting substrate levels. After running the modeling program, the modeled refer-

ence profiles for $X(t)$ and $V(t)$ are written in Matlab data file for reading/downloading into the MPC control algorithm responsible for feeding rate control/correction during the process.

In the analyzed case, the feeding profile consisted of 20 points, where linear approximation was used for feeding rate interpolation between these points. Time intervals between the points were 30, 60, or 120 minutes, depending on the process phase. In the initial process phase, shorter intervals were chosen, and at the later phase, the intervals were longer. Fed-batch mode start time was calculated automatically, on the basis of available off-line biomass measurements and *a priori* open-loop modeling estimates to activate fed-batch mode at the pre-defined biomass amount of 5 g. In this case, the predefined feeding start time of 210 minutes (process time) was applied or automatically shifted to an earlier or later time, if necessary, depending on the actual deviations of the process variables.

Investigation of model parameter sensitivity with respect to the optimization objective function (for biomass amount control problem) has shown[12] that the most sensitive parameters for this type of model are the conversion yields and maximal specific reaction rates of the main components involved. Therefore, the biomass/glucose conversion yield Y_{xs} and the maximal specific glucose consumption rate σ_{max} were selected as the main model parameters to be re-identified on-line every time the new sample for biomass and glucose measurement is evaluated. Every time a new biomass and glucose sample is measured and the results are entered into the SCADA system (see Fig. 2, 3), the numbers are passed over the OPC server into the on-line running MPC algorithm. After a defined number of entered samples (suggested are 3–5 measurement points), the re-identification of the process model parameters (Y_{xs} and σ_{max}) is started. After simulation of the current process model, the root-mean-square deviations (RMSD) between the model output and the measured biomass and glucose concentrations (obtained right from the beginning of the process until the current process time) are calculated. As a final modeling quality measure, weighted sum of biomass RMSD and glucose RMSD is calculated. During the parameter re-identification, the values of two model parameters – biomass/glucose conversion yield Y_{xs} and the maximal specific glucose consumption rate σ_{max} – are re-identified. For this purpose, a standard direct search algorithm (Nelder-Mead Simplex, Matlab *fminsearch* function) was applied. The optimization algorithm runs for 200 iterations (the number was selected empirically, in order to assure necessary precision of re-identified parameters). The calculation takes up to 1–2 minutes depending on the process phase. The re-identified parameter values are used for further process model simulations (including calculation of the optimal control action – feeding rate value for the time interval until the next sampling moment) until the next sample is taken. Then the re-identification procedure is repeated, and a new set of re-identified model parameters is obtained.

In the analyzed example, the set of Y_{xs} and σ_{max} model parameters was identified within the defined range (boundaries) in order to obtain the smallest possible deviations of the modeled biomass and substrate trajectories from the real values at the measurement points. The boundaries of their possible changes should be chosen based on *a priori* knowledge. The boundaries used by the authors were as follows: ± 30 % for Y_{xs}, and ± 30 % for σ_{max}. The diagram of the implemented MPC system is presented in Fig. 4. After a new sample is taken (approx. every 30–60 min.), the model parameters are re-identified on-line using the process model (eqs. 1–5). After the on-line model identification, a one-step-ahead prediction of the state variable trajectories is performed. In this step, a model-based optimization is executed. The aim of the optimization is to calculate/correct the control action (feeding rate) for the next control horizon in such a way that the deviations between the desirable (reference) biomass amount trajectory and the predicted one are minimized (eq. 7). The optimization is performed every time the off-line measurements of biomass and substrate are made and the model parameters have been re-identified. For control variable recalculation, a simple user-defined algorithm for scanning within the predefined range of F_s was applied. In total, every time 100 possible values of F_s within the defined control variable range were tested, the one leading to the lowest predicted tracking error (J) was chosen for implementation in the controller. The chosen iteration number assures the precision of the control and state variables that is sufficient from the practical application point of view.
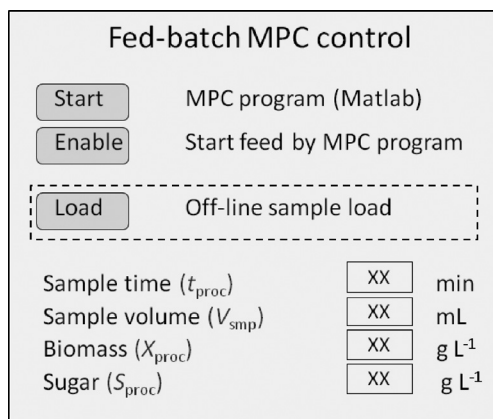

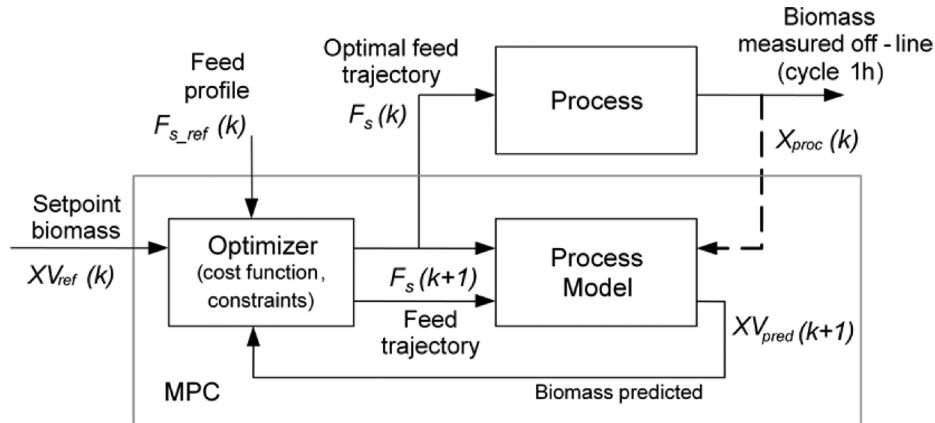F i g . 3 – *Screen-shot from SCADA fed-batch control window*

Fig. 4 – *Diagram of the implemented MPC system*

During optimization, the allowed feed correction span was ± 10 %. For the cases when predictions state that ± 10 % feed correction could be too small to avoid significant substrate accumulation, an appropriate feeding rate reduction exceeding 10 % took place (the latter case is herein after referred to as possible significant substrate accumulation). Off-line measurements as well as on-line modeling data were additionally stored in ASCII format files. In order to assure more reliable operation of the MPC controller, an intelligent filter was implemented to exclude faulty data reads from SCADA. The fault detection algorithm was based on the analysis of the current process variable values as well as their rates of change.

Results

Evolution of desirable reference biomass and feeding profiles

The procedure of the gradual evolution of the desirable reference feeding profile was demonstrated by means of 4 experiments performed one after another, where the initially reduced feeding profile was set. Later, it was specifically increased and modified to obtain a higher biomass yield (see Fig. 5, Fig. 7, and Fig. 8). The idea was to initially select feeding profile F_0, which would lead to biomass exponential growth phase, followed by a slowly decreasing biomass specific growth rate (μ). In case of possible biomass productivity increase, the feeding profile was gradually increased to the values that lead to maximum biomass amount, and at the same time significant substrate accumulation is avoided. By following this procedure in subsequent processes, it was possible to obtain increased μ in exponential growth phase and a very similar μ after exponential growth phase for all processes (except the one with significant substrate accumulation) (see Fig. 5).
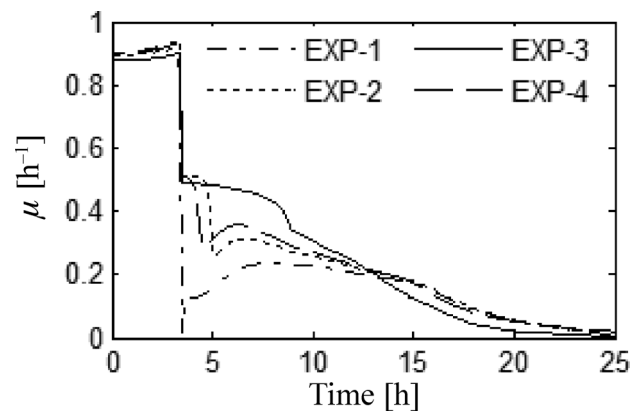

Fig. 5 – *Modeled specific growth rates resulting from the reference profiles for EXP-1 – EXP-4*

For selection of the reference profile as well as for on-line process modeling in the first experiment (EXP-1), the model parameters from *a priori* knowledge[14,15,19] were used. Initial values for the limitation and inhibition model parameters were taken from Xu *et al.*[14]; for biomass inhibition parameter *Kx*, the initial value was equal to the maximal biomass concentration achieved in *E. coli* B culture, 86 g L^{-1}, (Lee[20]). After parameter tuning (model fitting to experimental data) in subsequent experiments, the parameter values shown in Table 2 were used in further process modeling. Y_{xs} and σ_{max} were additionally re-identified on-line to fit the actually measured values of biomass and substrate.

Table 2 – *Model parameters (E. coli process)*

Parameter	Value
Y_{xs}	0.53 g g^{-1}
σ_{max}	1.05 g g^{-1} h^{-1}
K_s	0.05 g L^{-1}
$K_{s,i}$	30 g g^{-1}
Kx_{max}	97.5 g L^{-1}
S_f	400 g L^{-1}

EXP-1

For this experiment, the initial feeding profile $F_{ref} = F_0$ was chosen. In the first experiment, the feeding profile, selected in biomass exponential growth phase, resulted in biomass specific growth rate of 0.20 ± 0.05 h^{-1}. Biomass amount at the process end was 157 g.

EXP-2

For the second experiment, the points of initial feeding profile (F_0) were proportionally increased by 35 %, i.e., $F_{ref} = 1.35 \cdot F_0$, to obtain the biomass specific growth rate of 0.30 ± 0.05 h^{-1} in exponential growth phase. Biomass amount at the process end was 225 g.

EXP-3

For the third experiment, the points of the feeding profile from the previous run were increased by ~40 %, i.e., in total $F_{ref} = 1.75 \cdot F_0$, to obtain the biomass specific growth rate of 0.40 ± 0.10 h^{-1} in exponential growth phase. Although the model had predicted some sugar accumulation in the second part of the process, it was decided to test the chosen feeding strategy. In this case, one could test whether the MPC algorithm is able to correct the feeding profile to eliminate the critical substrate accumulation. As it was predicted by the model, at 16th process hour, glucose accumulation started and the feeding profile was significantly reduced. A significant mismatch in the biomass profile after 18th process hour was observed because significant glucose accumulation was avoided, which could potentially lead to more inhibited culture growth. At the very beginning of significant glucose accumulation (16 h, $X_{proc} = 80.6$ g L^{-1}, $S_{proc} = 4.6$ g L^{-1}), the process biomass had gained faster growth compared to the model predictions. This could indicate some model imprecision in this particular process phase when using relatively high feeding rates. Significantly increased X_{proc} values, as compared to $X_{ref,}$ at process time 9–11 h, are assumed as imprecise measurements caused by the operator. Biomass amount at the process end was 290 g.

EXP-4

Based on the EXP-3 results, the reference profile from EXP-3 was modified in order to test a slightly increased initial feeding profile, and to decrease the feeding rate points during the phase in which sugar accumulation was predicted by the process model. Finally, the feeding profile was selected, according to which no significant sugar accumulation was predicted during the modeled process. This feeding profile resulted in a similar growth compared to EXP-2, i.e., biomass specific growth rate of 0.30 ± 0.05 h^{-1} in exponential growth phase. Biomass amount at the process end was 227 g.

The theoretical estimation of the maximal productivity in particular cultivation process may be performed by means of a model-based design procedure similar to that described in the research of Galvanauskas *et al.*[12,16], i.e. by applying a realistic process model, taking into account the main limiting factors (constraints) of the real process (e.g., maximal oxygen transfer rate, maximal allowed culture volume, etc.), and performing a model-based optimization of the process. As a result, an optimal control strategy is obtained that leads to maximization of the numerically expressed productivity of the process. For the process investigated here, the authors have performed this study and have calculated that the optimal control strategy corresponds to the strategy experimentally achieved in EXP-2. Despite the higher final biomass amount achieved in EXP-3 and EXP-4, from the process controllability and product biosynthesis point of view, the observed significant glucose accumulation in EXP-3 and EXP-4 could potentially lead to significant acetate accumulation, resulting in inhibition of target product biosynthesis by *E. coli*.

Automatic fed-batch start

Due to slightly varying process initial biomass (X_0) and glucose (S_0) concentrations, and the resulting sub-optimal process control, the culture growth profile will vary as well. Therefore, the fed-batch mode initiation time potentially needs to be (automatically) shifted in one or another direction. Similar X_0 and S_0 parameters for EXP-1 and EXP-3 (see Fig. 6) resulted in the same fed-batch start at the process time of 210–212 minutes. Similar initial conditions were obtained for EXP-2 and EXP-4. In EXP-2 and EXP-4, initial X_0 was lower and S_0 was higher as compared to EXP-1 and EXP-3. Nevertheless, the faster biomass growth led to earlier fed-batch mode initiation at the process time of 180–186 minutes. The time points, when the process biomass reached the set reference values, are indicated with arrows in Fig. 6.

Model predictive feeding profile control

The experimental results are shown in Fig. 7 and Fig. 8. The estimated biomass amounts and concentrations at the process end for experiments EXP-1 – EXP-4 were 157 g (65.8 g L^{-1}), 225 g (79.8 g L^{-1}), 290 g (93.2 g L^{-1}), and 227 g (78.5 g L^{-1}), respectively. The estimated biomass ($XV_{end(est)}$) and experimentally measured ($XV_{end(exp)}$) differences ($d_{end(est)}$, $d_{end(exp)}$) from the reference profiles (XV_{ref}) at the process end (24 h) were 4.6 % and 3.8 %, respec-

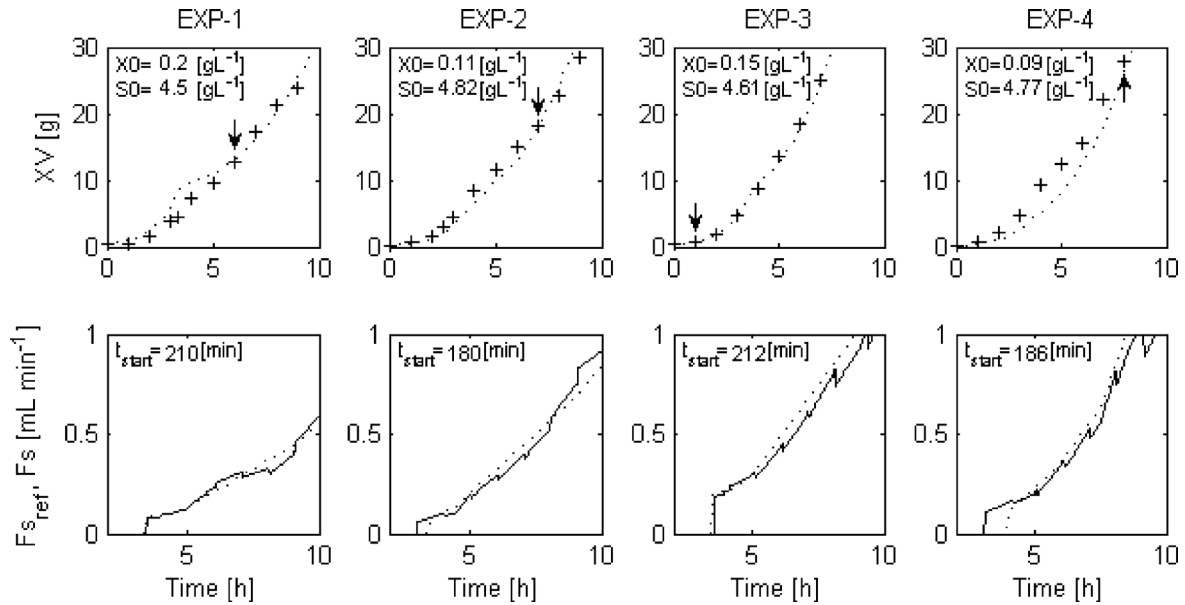Fig. 6 – *Process initial conditions (X_0 and S_0) and feeding start times (t_{start}) (below) for processes EXP-1 – EXP-4. (– –) set (modeled) reference profiles, (+) estimated, (–) applied feeding rate, (↓) indication of time moments when biomass reached set references.*

tively. Smaller XV_{est} deviations from the reference profiles were achieved at 20th process hour, where the mean deviation was 3.0 %. This could be explained by the fact that the model had not precisely enough modeled the late biomass stationary growth phases. To model this stage with higher precision, a more complex process model is required, which takes into account potential (by)product (acetate, protein, etc.[19,21]) formation, and biomass physiological states (cell starvation, ratio of viable cells, etc.[5]). Obviously, the precision of the selected (modeled) culture volume reference profile has an impact on $d_{end(est)}$. Higher numerical value of $d_{end(est)}$ as compared to $d_{end(exp)}$ indicates that biomass estimation
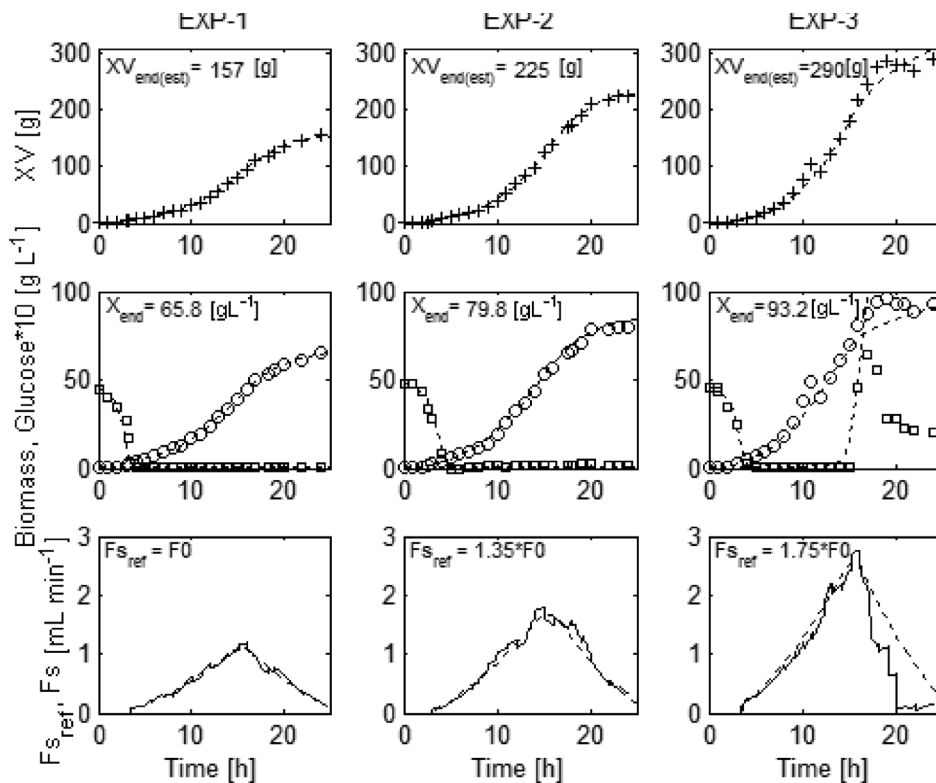

Fig. 7 – *Biomass amount control results in EXP-1 – EXP-3. (– –) Set (modeled) reference profiles, (+) estimated on-line, (□) off-line glucose, (○) off-line biomass, (–) applied feeding rate.*
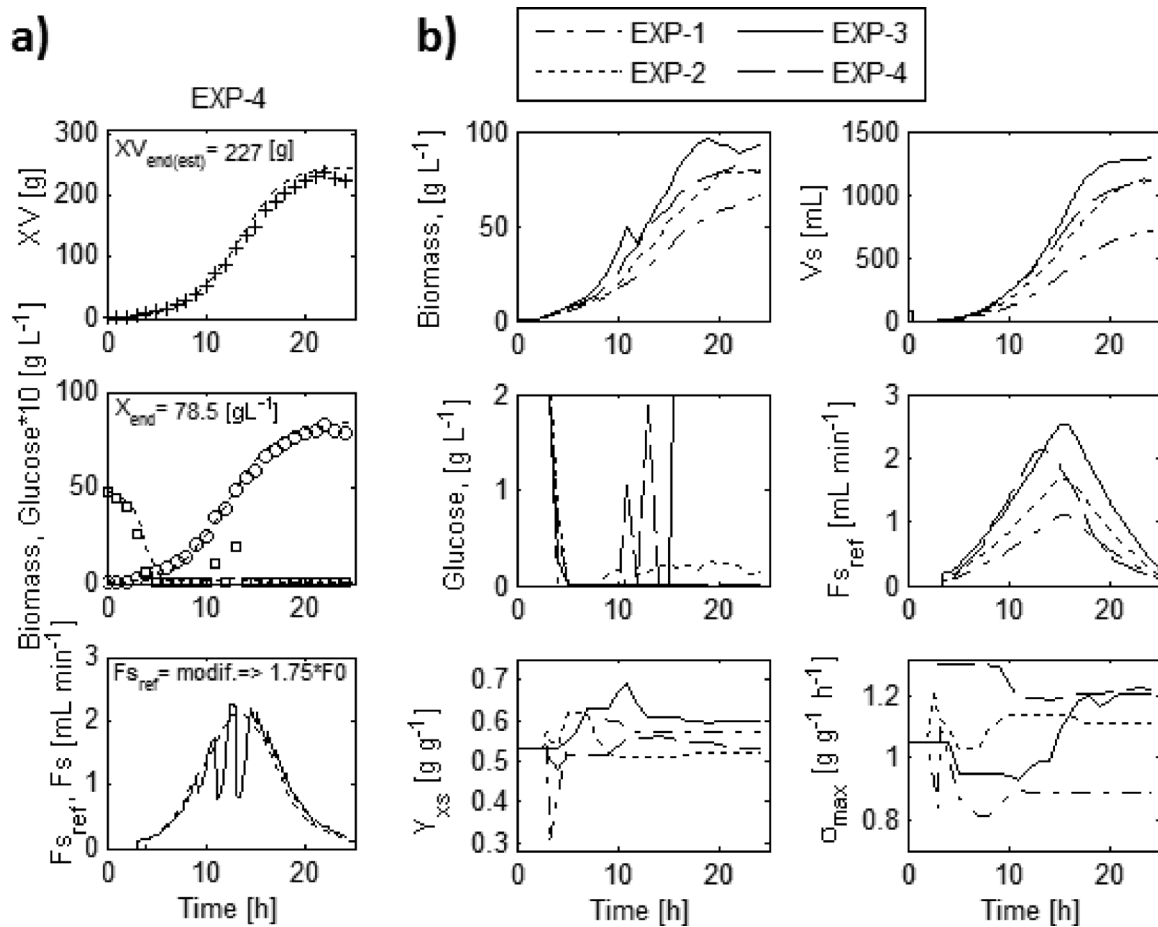
Fig. 8 – *Biomass amount control results (EXP-4) and summarized overview of parameter value evolution in experiments EXP-1 – EXP-4. a) (– –) Set (modeled) reference profiles, (+) estimated on-line, (▫) off-line glucose, (○) off-line biomass, (–) applied feeding rate b) Comparison of experiments EXP-1 – EXP-4. Lines for biomass, glucose, Y_{xs} and σ_{max} are drawn and interpolated between off-line analytical measurements made every 1–2 h.*

quality could be improved by more precise volume on-line estimation. Additionally, the results of EXP-4 suggest that the model parameters used for reference profile selection were not optimal for the case where an excessively increased feeding profile was chosen for the exponential growth phase. Taking into account the initial conditions (X_0 and S_0) of the EXP-4, the model does not reflect the observed significant glucose accumulation at 11th process hour. In this case, reduction of the values of model parameters Y_{xs} and σ_{max} by 5–10 % during the reference profile selection could lead to a more safe process in terms of avoiding significant glucose accumulation.

In EXP-1 and EXP-2, significant excess glucose accumulation during fed-batch phase did not occur. At the end phase of the process, significant glucose accumulation up to 6.39 g L⁻¹ was observed in EXP-3. It started at 16th process hour and continued until the end of the process: 4.61 g L⁻¹ (16 h), 6.39 g L⁻¹ (17 h), 5.58 g L⁻¹ (18 h), 2.77 g L⁻¹ (19 h), 2.79 g L⁻¹ (20 h), 2.23 g L⁻¹ (21 h), 2.12 g L⁻¹ (22 h), and 2.02 g L⁻¹ (24 h). This can be explained

by a too high reference feeding profile starting from the 15th process hour. Moreover, it is highly probable that the increased/incorrect biomass measurement results until the 12th process hour, when the process operator changed, were obtained in EXP-3 (significantly increased values at process time 9–11 h). Despite this fact, the biomass reference profile was reached in 2–3 h after more precise sampling events had occurred. For EXP-4, the feeding profile interval was reduced accordingly, and slightly increased feeding rates for the first part of the reference feeding profile werechosen, in order to test the faster growth ability. This resulted in significant glucose accumulation only in two sampling time moments in the first feeding profile part: 1.04 g L⁻¹ (11 h) and 1.87 g L⁻¹ (13 h). In the remaining part of the process fed-batch phase, the glucose concentration did not exceed 0.26 g L⁻¹.

From the Y_{xs} and σ_{max} adaptation results, additional useful information maybe extracted. If a brief overview of biomass production process is made, the biomass yield on substrate could become an indicator of how efficient the main carbon source is

converted into biomass. A higher yield potentially indicates that a smaller amount of feed is wasted for undesirable byproduct synthesis. Starting from process hour 8–9, some correlation between the estimated Y_{xs} trends for explorative experiments (see Fig. 8, b) could be observed. It is interesting that the highest yield at the end of the process was observed in the experiments with the slowest and the fastest feed rate profiles, in average 0.57 g g^{-1} for EXP-1 and 0.60 g g^{-1} for EXP-3. A relatively small difference in the yield between EXP-2 and EXP-4 was observed, 0.52 g g^{-1} and 0.54 g g^{-1}, respectively. This correlates well with the biomass at the end of the process (231 g and 247 g) taking into account that the substrate feed volume (V_s) for these experiments at the end of the process was practically the same. The estimation results for σ_{max} show that the values have fluctuated within the range of 1.05 ± 0.2 g g g L^{-1}. The variations may be related to different substrate limitation conditions within the performed experiments.

Test experiment in SUB

For the MPC control approach comparison in the reactors of different configuration, a test run in a single-use lab bioreactor (see configuration in Table 1) was carried out (see Fig. 9 (a) and (b)). For this experiment, the same reference profiles as in EXP-2 were chosen to maintain culture growth without critical glucose accumulation and, at the same time, to achieve comparatively high biomass yield. Initial biomass and sugar concentrations were similar to those in EXP-2. Nevertheless, possibly due to different temperature controller PID parameter tuning (until the 6th process hour) in order to compensate differences between two reactors (mainly because of different material and specific area of heat exchange element), frequent temperature oscillations of 36.8 ± 0.9 °C occurred in the initial phase of the process. This most probably led to longer (by 3–4 hours) batch lag phase and much lower process biomass at this stage. Regardless, the culture was able to recover and closely reached the reference biomass at 17 h (see Fig. 9, a). The feeding phase started at 210 minutes. Biomass of 212 g (78.1 g L^{-1}) was reached at the process end. Biomass estimate ($d_{end(est)}$) deviation from the pre-set reference at the end of the process was 4.3 %. Glucose concentration did not exceed 4.1 g L^{-1}.
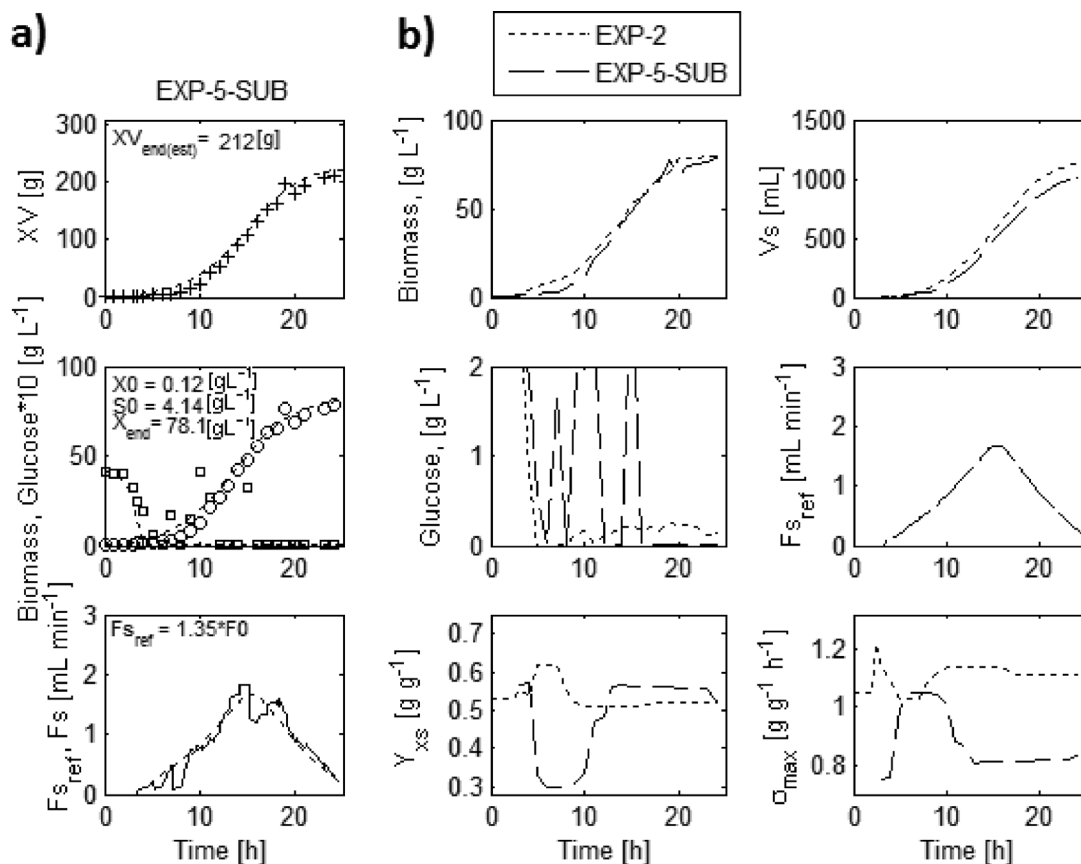

F i g . 9 – *Biomass amount control results (EXP-5-SUB) in single-use reactor and its comparison with EXP-2. a) (– –) Set (modeled) reference profiles, (+) estimated on-line, (◦) off-line glucose, (○) off-line biomass, (–) applied feeding rate b) Comparison of experiments EXP-5-SUB and EXP-2. Lines for biomass, glucose, Y_{xs} and σ_{max} are drawn and interpolated between off-line analytical measurements made every 1–2 h.*

Conclusions

Gradual model-based evolution of feeding profile and model-predictive control has been demonstrated on recombinant *E. coli* BL21 biomass cultivation processes in conventional and single-use bioreactors. The best match to the reference profiles was observed in the processes without significant glucose accumulation and in all processes until the beginning of biomass stationary growth phase, the 20th process hour. The reference biomass profile for the experiment without significant glucose accumulation, in which biomass of 79.8 g L^{-1} (225 g) was achieved at the end of the process, can be assumed as the optimal one for safe and high-yield process conduction with the particular equipment used. Estimated biomass mean deviation from the pre-set references at the end of the processes was 4.6 %. It was close to the deviations from the experimentally measured one (3.8 %). By means of sequentially performed experiments, it was possible to test the feeding profiles of different magnitude and shape, while maintaining exponential growth with specific growth rates of 0.2, 0.3, and 0.4 ± 0.05 h^{-1}. In this way, by performing the feeding profile modification, the so-called desirable "golden batch" feeding profile maybe achieved and selected as a reference for the particular process. A successful test experiment was conducted in the single-use laboratory bioreactor. Biomass estimate ($d_{end(est)}$) deviation from the pre-set reference at the end of the process for this experiment was 4.3 %.

The developed and implemented MPC system was used to track the selected reference trajectories of the biomass amount. The MPC system demonstrated good control performance and reduced process variability as compared to the system with an open-loop feeding profile control. The implemented process control and its performance may be additionally improved. A more specific process model adapted for *E. coli*, for example, taking into account by-product (protein, acetate, etc.) synthesis/consumption and information about biomass physiological state (cell starvation, ratio of viable cells, etc.), may enhance precision of the process modeling and biomass on-line estimation. Improved accuracy of volume estimation by means of direct weighing of solution bottles or reactor may also improve the precision of biomass estimation and its subsequent control.

The proposed approach has been implemented in a commercially available bioreactor system (Biotehniskais Centrs, JSC) and is dedicated for a wide range of potential users.

Nomenclature

ANN – Artificial Neural Network

DCU – Digital Control Unit

GMP – Good Manufacturing Process

IMC – Internal Model Control

MPC – Model Predictive Control

OD – Optical Density

OPC – Open Platform Communications

PCA – Principal Component Analysis

PID – Proportional, Integral, and Derivative control

RQ – respiratory quotient

SCADA – Supervisory Control and Data Acquisition

SUB – Single Use Bioreactor

$d_{end(est)}$ – estimated biomass ($XV_{end(est)}$) mean deviation from reference biomass ($XV_{end(ref)}$) at the end of the processes, %

$d_{end(exp)}$ – experimentally measured biomass ($XV_{end(exp)}$) mean deviation from reference biomass ($XV_{end(ref)}$) at the end of the processes, %

F_s – feeding rate of substrate solution, L h^{-1}

F_{s_ref} – reference feeding rate of substrate solution, L h^{-1}

F_{smp} – sampling rate, L h^{-1}

F_c – carbon loss via off-gass, L h^{-1}

F_b – alkali addition rate, L h^{-1}

F_e – evaporation rate, L h^{-1}

F_0 – initial feeding profile, L h^{-1}

K_s – Monod constant for substrate uptake, g L^{-1}

$K_{i,s}$ – substrate inhibition constant, g L^{-1}

Kx_{max} – biomass inhibition constant, g L^{-1}

S_0 – initial glucose concentration at $t_{proc} = 0$ h, g L^{-1}

S – glucose concentration, g L^{-1}

S_{proc} – process glucose at time t_{proc}, g L^{-1}

S_f – glucose concentration in feed, g L^{-1}

t, t_{proc} – process time, h

t_{start} – fed-batch start time, h, min

V – culture volume, L

V_s – fed substrate, L

X_0 – initial biomass concentration at $t_{proc} = 0$ h, g L^{-1}

X – biomass concentration, g L^{-1}

X_{proc} – process off-line biomass at time t_{proc}, g L^{-1}

$XV_{end(est)}$ – estimated biomass at the end of the process, g

$XV_{end(exp)}$ – experimentally measured biomass at the end of the process, g

$XV_{end(ref)}$ – reference biomass at the end of the process, g

XV_{proc} – biomass off-line estimate at time t_{proc}, g

XV_{ref} – set biomass reference profile, g

XV_{pred} – one step ahead predicted biomass, g

Y_{xs} – biomass yield, g g^{-1}

σ – specific substrate consumption rate, g g^{-1} h^{-1}

σ_{max} – maximal specific substrate consumption rate, g g^{-1} h^{-1}

μ_s – biomass specific growth rate, L h^{-1}

References

1. Directive 2001/83/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to medicinal products for human use: EU Legislation – Eudralex. **1** (2001) 1.

2. *Shimizu, K.*, An overview on the control system design of bioreactors, in *Fiechter, A.* (Ed.), Advances in Biochemical Engineering Biotechnology, Springer-Verlag, Berlin/Heidelberg, Germany, 1993, pp 65-84.

3. *Junker, B. H., Wang, H. Y.*, Bioprocess Monitoring and Computer Control: Key Roots of the Current PAT Initiative, Biotechnol. Bioeng. **95** (2006) 226.

4. *Lübbert, A., Jørgensen, B. S.*, Bioreactor performance: a more scientific approach for practice, J. Biotechnol. **85** (2001) 187.

5. *Pohlscheidt, M., Charaniya, S., Bork, C., Jenzsch, M., Notzel, T. L., Lübbert, A.,* Bioprocess and fermentation monitoring. in *Flickinger, M. C.* (Ed.), Upstream Industrial Biotechnology: Equipment, Process Design, Sensing, Control, and cGMP Operations, Vol. 2., 1st ed., John Wiley & Sons, Inc., New Jersey, USA, 2013, pp 1469-1491.

6. *Kuprijanov, A., Schaepe, S. V., Simutis, Lübbert, R. A.,* Model predictive control made accessible to professional automation,Biosystems and Information Technology **2** (2013) 26.

7. *Sakato, K., Tanaka, H.*, Advanced control of glutathione fermentation, process, Biotechnol. Bioeng. **40** (1992) 904.

8. *Hosobuchi, M., Kurosawa, K., Yoshikawa, H.*, Application of computer to monitoring and control of fermentation process: microbial conversion of ML-236B to pravastatin, Biotechnol. Bioeng. **42** (1993) 815.

9. *Weuster-Botz, D., Paschold, H., Striegel, B., Gieren, H. M., Kula, R., Wandrey, C.*, Continuous computer controlled production of formate dehydrogenase (FDH) and isolation on a pilot scale, Chem. Eng. Technol. **17** (1994) 131.

10. *Chen, Y., Krol, J., Huang, W. J., Cino P., Vyas R., Mirro R., Vaillancourt, B.*, DCO2 on-line measurement used in rapamycin fed-batch fermentation process, Process. Biochem. **43** (2008) 351.

11. *Schaepea, S., Kuprijanov, A., Simutis, R., Lübbert, A.*, Avoiding overfeeding in high cell density fed-batch cultures of *E. coli* during the production of heterologous proteins, J. Biotechnol. **192** (2014) 146.

12. *Galvanauskas, V., Simutis, R., Volk, N., Lübbert, A.*, Model based design of a biochemical cultivation process, Bioproc. Eng. **18** (1998) 227.

13. *Bajpai, R.*, Control of bacterial fermentations, Ann. N.Y. Acad. Sci. **506** (1987) 446.

14. *Xu, B., Jabic, M., Enfors, S. O.,* Modeling of overflow metabolism in batch and fed-batch cultures of *E. coli*, Biotechn. Progr. **15** (1999) 81.

15. *Nielsen, J.*, Microbial process kinetics, in *Ratledge, C., Kristiansen, B.* (Ed.), Basic biotechnology, Cambridge University Press, Cambridge, 2006, pp 155-180.

16. *Galvanauskas, V., Simutis, R., Lübbert, A.*, Model-based design of biochemical processes: simulation studies and experimental tests, Biotechnol. Lett. **19** (1997) 1043.

17. *Aehle, M., Bork, K., Schaepe, S., Kuprijanov, A., Horstkorte, R., Simutis, R., Lübbert, A.,* Increasing batch-to-batch reproducibility of CHO-cell cultures using a model predictive control approach, Cytotechnology **64** (2012) 623.

18. *Yüzgec, U., Palazoglu, A., Romagnoli, J. A.,* Refinery scheduling of crude oil unloading, storage and processing using a model predictive control strategy, Comput. Chem. Eng. **34** (2010) 1671.

19. *Galvanauskas, V., Grigs, O., Vanags, J., Dubencovs, K., Stepanova, V.*, Model-based optimization and pO_2 control of fed-batch *Escherichia coli* and *Saccharomyces cerevisiae* cultivation processes, Eng. Life Sci. **13** (2013) 172.

20. *Lee, S. Y.,* High cell-density culture of *Escherichia coli*, Trends Biotechnol. **14** (1996) 98.

21. *Gnoth, S., Jenzsch, M., Simutis, R., Lübbert, A.*, Control of cultivation processes for recombinant protein production: a review, Bioprocess. Biosyst. Eng. **31** (2008) 21.