Generic model based control of different specific rates in recombinant *E. coli* Fed-batch processes *

Julian Kager * Nora Horst *** Johanna Bartlechner ** Christoph Herwig *** Stefan Jakubek **

 * Competence Center CHASE GmbH, Linz, Austria (e-mail: julian.kager@chasecenter.at).
 ** Institute of Mechanics and Mechatronics, TU Wien, 1040 Vienna, Austria (e-mail: stefan.jakubek@tuwien.ac.at)
 *** Institute of Chemical, Environmental and Biological Engineering, TU Wien, Vienna, Austria (e-mail: christoph.herwig@tuwien.ac.at)

Abstract: In addition to often reported control of the specific growth rate we show in this contribution a novel method to control the biomass specific substrate uptake as well as the protein production rate during the production phase of microbial fed-batch cultures. The control laws base on a calibrated process model including experimental data sets with different constant glycerol feed rates. The model includes changing growth and production behaviour in function of the induced metabolic stress. Based on the method of non-linear feedback linearization, the control laws were derived from the model, using the glycerol feed rate as the manipulable variable.

Considering comparable setpoints at three different levels the effectiveness of the controller as well as their potential to improve the production phase was assessed by simulations. The direct control of product formation rates enables to ensure constant production performance and therefore to reach highest product levels in short time. As a drawback, the controller needs high feed adaption towards the end to counteract the advancing decay of the attainable production rates. This, to a lesser extend, also applies for the growth rate. A constant substrate uptake rate resulted in a valuable alternative to a uncontrolled, constant feed rate during the production phase of recombinant protein production processes.

Keywords: Bioprocess control, feedback linearization, generic model control, fed-batch, advanced control.

1. INTRODUCTION

Fed-batch processes using genetically modified $E.\ coli$ organisms are widely used for the production of high value recombinant proteins (Selas Castiñeiras et al., 2018). State of the art processes are divided into three main phases (Yee and Blanch, 1992). The batch phase, followed by an exponential feeding phase to ensure high amounts of metabolically active cells, which in the induction phase catalyze the transcription and translation of the target proteins.

The cell behavior during the growth phase is well understood and optimized to reach high biomass concentrations within short times (Lee, 1996), while avoiding the formation of unwanted side products (Abadli et al., 2021). Due to the metabolic load of the forced production of the target protein, cell metabolism changes in course of the induction phase (Neubauer et al., 2003). Therefore, control of the induction phase is a challenging task and often characterized by sub optimal, static parameters including constant feed rates (Wechselberger et al., 2012), leading to unfavorable changes during the production phase, without any possibility to intervene.

To improve these unfavorable operation, predetermined dynamic feed profiles or feedback control can be applied as reported by Mears et al. (2017). To predetermine optimal profiles as well as to enable effective feedback control mathematical models to describe the occurring non-linear process dynamics are required (Almquist et al., 2014) as normal PID control can fail without knowing the underlying dynamics (Kager et al., 2020; Zotică et al., 2020). In addition to that biological processes are very sensitive to deviations, so that little over shots can have irreversible effects. Therefore, different advanced and predictive control structures were developed and proposed for biotechnological processes (Mears et al., 2017). From these proposed structures especially generic model based control (Abadli et al., 2021) and model predictive control (Kager et al., 2020; Dewasme et al., 2015) gained special attention, as they can be based on simple and interpretable mechanistic models (Almquist et al., 2014). Abadli et al. (2021) for example derived a robust model based controller for optimal microbial growth, while at the same time controlling the

^{*} Funding was provided by the Competence Center CHASE GmbH, funded by the Austrian Research Promotion Agency (FFG) (No 868615).

formation of unwanted acetate concentration. Kager et al. (2020) verified a model predictive controller to foresee and therefore avoid unwanted byproduct formation in a penicillin production process. Model predictive control was also adopted to very sensitive mammalian cell cultivations to recursively optimize substrate addition while at the same avoiding to activate unwanted metabolic pathways (Dewasme et al., 2015).

Although adaptive and predictive model based control concepts were already introduced in biotechnological processes their application is often limited to describe the growth or the consumption of different substrates or by product formation. There are only a few works including product formation which are either simple products from primary metabolism or simple relations to other control variables (Mears et al., 2017).

Within this work we concentrate on the product formation phase of an *E. coli* fed-batch process and introduce and analyze the potential control of specific rates including the often discussed growth rate as well as the biomass substrate uptake rate and the novel and direct control of the biomass specific production rate during the production phase. After the description of the non-linear process model including the calibration on three experimental data sets, the establishment of the control laws by the method of non-linear feedback linearization is given. Besides showing the real and modelled process dynamics of the three experiments, simulation results of the derived control laws on comparable setpoints are given in the results section before concluding with potential usage of the control laws to improve the production phase of biotechnological processes.

2. MODEL AND CONTROL LAW ESTABLISHMENT

2.1 Fed-batch model including metabolic load and product formation dynamics

The system differential equations for the ideally stirred tank reactor in fed-batch mode are:

$$\frac{dV_R}{dt} = F_{in}$$

$$\frac{dc_X}{dt} = \mu c_X - \frac{F_{in}}{V_R} c_X$$

$$\frac{dc_S}{dt} = -q_S c_X + \frac{F_{in}}{V_R} (c_{S,in} - c_S) \quad (1)$$

$$\frac{dc_P}{dt} = q_P c_X - \frac{F_{in}}{V_R} c_P$$

$$\frac{dS_{met}}{dt} = q_S$$

describing the concentration changes over time $\left(\frac{dx}{dt}\right)$ of the four macroscopic components: biomass c_X , glycerol c_S , product c_P , metabolized substrate S_{met} and the reactor volume V_R . Glycerol is added to the system by the substrate inflow F_{in} with a concentration $c_{S,in}$. The biological conversion rates, namely the growth rate μ , the substrate consumption rate q_S and the product formation rate q_P are described by the following reaction kinetics: The substrate uptake rate q_S described by a Monod kinetic with $q_{S,max}$ being the maximum uptake rate in function of the glycerol concentration c_S and the half saturation constant K_S .

$$q_S = q_{S,max} \frac{c_S}{c_S + K_S} \tag{2}$$

Biomass growth μ is derived from consumed substrate q_S by substrate conversion yield $Y_{X/S}$ reduced by the substrate needed for cell maintenance m_S . Whereas $Y_{X/S}$ is reduced by an asymptotic decay $(K_{Y_{X/S}})$ in function of the metabolized substrate (S_{met}) .

$$\mu = Y_{X/S} (q_S - m_S)$$

$$Y_{X/S} = Y_{X/S,max} \exp(-S_{met} K_{Y_{X/S}})$$
(3)

The product formation q_P is composed by a Monod term describing the dependency of $q_{P,max}$ on q_S with the half saturation constant K_{Sq_S} and a Haldane term to describe the underlying start up and decline phase with S_{met} as trigger, K_{Sq_P} as the delay coefficient, K_{Iq_P} as the decay coefficient and k as the Haldane exponent determining the shape of the decay.

$$q_{P} = q_{P,max} \frac{q_{S}}{q_{S} + K_{Sq_{S}}} \frac{S_{met}}{\frac{S_{met}^{k}}{K_{Iq_{P}}} + S_{met} + K_{Sq_{P}}}$$
(4)

The model was fitted to three experimental data sets, with three different constant feed rates, displayed in figure 1. The model parameters are given in table 1. As fitting criterion the weighted sum of squared errors between measurements of biomass, product and sugar concentrations and respective model simulations of all three experiments was minimized by a local optimizer (Fmincon:MATLAB).

Table 1. Model Parameter

Parameter	Value	Unit	Description
$c_{S,in}$	850	gL^{-1}	feed concentration
$q_{S,max}^{*}$	1.0	$gg^{-1}h^{-1}$	max. substrate uptake
K_S^*	0.0050	gL^{-1}	half saturation constant
m_S	0.02	h^{-1}	maintenance constant
$Y_{X/S,max}^*$	0.47	gg^{-1}	max. growth yield
$K_{Y_X/S}$	0.34	gg^{-1}	growth yield decay
$q_{P,max}$	0.0066	$gg^{-1}h^{-1}$	max. production rate
K_{Sq_S}	0.082	$gg^{-1}h^{-1}$	q_P affinity to q_S
k^{-}	4.6	—	Haldane exponent
K_{Iq_P}	3.8	$(gg^{-1})^{k-1}$	product decay
K_{Sq_P}	0.096	gg^{-1}	product formation delay

2.2 Experimental data

A modified K12 *E. coli* strain with a rhamnose-inducible expression system (rhaBAD promoter), producing a single chain antibody fragment with a transporter sequence to be transported into the periplasm was used. Fermentations were conducted in a DASGIP multibioreactor system (Eppendorf, Germany). Composition of the used minimal media was based on Wilms et al. (2001). Temperature was kept at 35°C, stirrer speed at 1400 rpm and aeration at 1.4 vvm for the whole process. The pH was controlled at 7.0 with addition of 12.5 % NH₄OH solution. The dissolved

oxygen (DO₂) was kept over 25% by supplementing pure oxygen to the air. A volume of 2.5% of the 1.0 L batch volume equivalent of an overnight pre-culture (30°C and 170 rpm) was used to initiate the batch phase of the reactors with 20 gL⁻¹ glycerol and a duration of approx. 12 h. For the pre-induction phase an exponential feed ramp with a μ of 0.14 according to the equation given in Lee (1996)) was used to reach high biomass concentrations of 45 gL⁻¹. Recombinant protein production was induced by a one point addition of sterile filtrated rhamnose solution (1.5 g L-rhamnose) and the glycerol feed rate was set to different constant values (*high*= 14.5 mLh⁻¹; *medium* = 8.5 mLh⁻¹; *low* = 2.2 mLh⁻¹).

Periodical offline measurements were taken and analyzed. Biomass dry content was determined by centrifugation at 4500 g, 10 min, 4°C including one washing step before drying at 105°C for min. 72h. For intra- and extracellular product content (homogenized and gel filtrated (PD MiniTrap g-25) cellsupension and supernatant) was applied on a protein G affinity column (HiTrap ProtG (GE Healthcare; USA) with a flow rate of 2 mlmin⁻¹ (20 mM phosphate buffer) at 25°C and elution with a change of the pH from 7.4 to 2.5). Acetate and glycerol concentrations were quantified from the supernatant by enzymatic, photometric principle in a robotic system (BioHT, Roche, Germany) and were under the detection limit during induction. Measured reaction rates μ , q_S and q_P , displayed in figure 1, were calculated based on the local difference between the measurement points.

2.3 Non-linear control by feedback linearization

To derive different control laws to keep the reaction rates constant, feedback linearization was applied in order to include the non-linearities into a standard feedback control law as described by Abadli et al. (2021), which defined the procedure as generic model based control. The system of differential equations given in Eq. 1 represents a nonlinear, input-affine system

$$\frac{d\boldsymbol{x}}{dt} = \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})u
y = h(\boldsymbol{x})$$
(5)

with the non-linear vector fields \boldsymbol{f} and \boldsymbol{g} of the state vector \boldsymbol{x} , the manipulable variable $u(F_{in})$ and the output variable y as function $h(\boldsymbol{x})$. By finding a suitable nonlinear feedback control law the non-linearity of the plant can be compensated exactly. The derivative of y becomes:

$$\dot{y} = \frac{\mathrm{d}h(\boldsymbol{x})}{\mathrm{d}t} = \frac{\partial h(\boldsymbol{x})}{\partial \boldsymbol{x}} \dot{\boldsymbol{x}} = \frac{\partial h(\boldsymbol{x})}{\partial \boldsymbol{x}} \boldsymbol{f}(\boldsymbol{x}) + \frac{\partial h(\boldsymbol{x})}{\partial \boldsymbol{x}} \boldsymbol{g}(\boldsymbol{x}) \boldsymbol{u} \quad (6)$$

The resulting partial derivatives from the time derivative of y can be defined as Lie derivatives $L_f h(x)$ or respectively $L_g h(x)$ resulting in:

$$\dot{y} = L_f h(\boldsymbol{x}) + L_g h(\boldsymbol{x}) u \tag{7}$$

with \dot{y} being a linear function of the manipulated input u. The transformed system can than be expressed as a simple proportional control law

$$\dot{y} = \alpha(y^* - y) \tag{8}$$

with the time constant $1/\alpha$ to track a setpoint y^* , which yields

$$u = \frac{\alpha(y^* - y) - L_f h(\boldsymbol{x})}{L_g h(\boldsymbol{x})}.$$
(9)

Deviations of the output variable y from the desired setpoint y^* are accounted for by the driving term $\alpha(y^* - y)/L_gh(x)$. Changes of the system due to nonlinear dynamics are considered by the maintenance term $L_fh(x)/L_gh(x)$. The method of feedback linearisation, briefly described here with an extended description in Cheng et al. (1984), was applied to obtain control laws for the system described in Eq. 1 by choosing the output y as the specific rates (μ , $q_S \& q_P$) and the input u as feed rate (F_{in}). It is noted that for all specific rates the relative degree was one as $L_gh(x) \neq 0$ and the internal dynamics therefore resulted in the order of four. The resulting proportional control laws for μ , q_S and q_P are:

Growth rate μ

$$\dot{F}_{in} = \frac{V_R}{c_{S,in} - c_S} \frac{\alpha(\mu^* - \mu) + q_S \left(K_{Y_{X/S}} \mu + Y_{X/S} c_X b_1 \right)}{Y_{X/S} b_1}$$
(10)

Substrate uptake q_S

$$\dot{F}_{in} = \frac{V_R}{c_{S,in} - c_S} \frac{\alpha(q_S^* - q_S) + c_X q_S b_1}{b_1}$$
(11)

Productivity q_P

$$\dot{F}_{in} = \frac{V_R}{c_{S,in} - c_S} \frac{\alpha(q_P^* - q_P) + q_S q_P (b_2 c_X + b_3)}{b_2 q_P} \quad (12)$$

with b_1 , b_2 and b_3

$$b_1 = \frac{q_{S,max} - q_S}{K_S + c_S} \tag{13}$$

$$b_2 = \frac{1}{K_S + c_S} \left(\frac{K_S}{c_S} - \frac{q_{S,max} - q_S}{K_{Sq_S} + q_S} \right), \qquad (14)$$

$$b_{3} = \frac{\frac{(K-1)S_{met}}{K_{Iq_{P}}} - K_{Sq_{P}}}{S_{met} \left(K_{Sq_{P}} + S_{met} + \frac{S_{met}^{k}}{K_{Iq_{P}}}\right)}$$
(15)

where α is the proportional factor and μ^* , q_S^* and q_P^* the aimed setpoints.

3. PROCESS DYNAMICS AND SETPOINT DEFINITION

In figure 1 the measured and modelled rate dynamics of the three experiments are given. As already discussed in the introduction the constant feed rates lead to highly dynamic and potentially unfavorable changing reaction rates throughout the production phase. No matter if high or low feed rates are applied reaction rates strongly change over time. Especially the q_P trajectory seems to be highly influenced by the feed level with higher maximum values at



Fig. 1. Model identification experiments with three (low, mid, high) constant glycerol feedrates F_{in} and corresponding specific rates (μ , $q_S \& q_P$) obtained from measurements (light colors) and calibrated model (dark colors)

high feeds but due to the lower metabolic stress prolonged productivities at low feed rates.

Within table 2 the root mean square errors normalized by the overall observed range (NRMSE) are displayed. Overall the identified kinetics of the process models with errors < 15 % describe well the rate trajectories of the three experiments with slightly higher errors for the mid and high feed rates. Therefore, the model can be regarded as suitable and trustworthy predictive within the calibrated region.

Table 2. normalized root mean square error (NRMSE) between modelled and measured rates $(\mu, q_S \& q_P)$ of different constant feeding levels

level	μ	q_S	q_P	Unit
low	2.7	3.2	5.5	%
mid	11.8	14.8	10.4	%
high	9.8	12.0	13.5	%

To deduce controller setpoints, which are comparable to the experiments with high, mid and low feedrates the average measured rates were calculated from the three experiments. The resulting, average rates are displayed in table 3. It can be seen, that with constant feeds, different average rate levels were achieved. Additionally, it can be seen that every specific rate is effected to another extent by the feed level. For example a low feed rate leads to a ten fold reduction of overall growth compared to a high feed rate but only a reduction of factor of four for the average product formation rates. This indicates that, although maximum production rates can be increased by higher feed rates, lower feed rates can lead to a more stable production at lower levels. The average rates, given in table 3 were used as setpoints for the subsequent controller simulation study.

> Table 3. Average measured rates of μ , $q_S \& q_P$ based on different constant feeding levels with corresponding feed rate F_{in}

feeding	F_{in}	μ	q_S	q_P
level	mLh^{-1}	h^{-1}	$gg^{-1}h^{-1}$	$gg^{-1}h^{-1}$
low	2.2	0.0059	-0.0294	0.00063
mid	8.5	0.0230	-0.1191	0.00126
high	14.5	0.0675	-0.2580	0.00204

4. CONTROLLER SIMULATION RESULTS

Based on the defined and comparable setpoints, controller simulations to keep either the growth rate μ , the substrate uptake rate q_S or the specific production rate q_P were performed. For each controller three different setpoint levels were simulated as displayed in table 3. All simulations run with a proportional factor of $\alpha = 1000$ without any modelplant-mismatch or other additional disturbances. Based on the feedback linearized control laws, perfect control behaviours can be therefore expected. The resulting feed rates F_{in} and reaction rates are displayed in figure 2.



Fig. 2. Simulation results of developed control laws to control: a) the growth rate μ , b) the substrate uptake rate q_S as well as c) the biomass specific productivity q_P during the production phase of the analyzed *E. coli* fed-batch process at three different setpoints (*low, mid, high*)

As the productivity is of central interest during the production phase it is shown for every controller simulation. Overall, all three control laws worked for the three selected setpoint levels and resulted in constant μ (figure 2 a), constant q_S (figure 2 b) as well as constant q_P (figure 2 c).

The manipulable variable F_{in} was hereby adapted in function of the feed-back linearized control laws, with the general form given in 9. Both the control of μ as well as q_P needed strong feed adaptions especially at higher setpoints. This is due to the observed and modelled decrease of growth and production capacities in function of the metabolic load. To counteract this the controller needs higher feed amounts over time resulting in a overexponential trajectory of the manipulated variable. Especially towards the end of the q_P controlled processes this effect is strongly pronounced and high feed rates are reached. Although, simulations didn't show any substrate accumulation in this region, this can be problematic in a real process.

The controller output of the substrate uptake behaves differently. Due to highly reduced growth at a low setpoint only little adaption is necessary to keep biomass specific substrate uptake constant. For mid and high setpoints strong adaptions are needed at the beginning, whereas towards the end, adaption is flattened in function of the decreasing growth capacity. As growth and production efficiency decreases over time overall substrate demand stagnates, which is visible in the resulting, more linearly increasing feed profiles of the q_S controller.

The trajectories of the productivity reveal that a control of μ or q_S yield in similar output behaviour, which is similar to the resulting q_P trajectory of constant feed rates as displayed in figure 1. Also the yielding maximum product

concentrations summarized in table 4 reach very similar levels. In all cases, with lowest setpoints highest product concentrations can be reached with highest overall values for q_S and μ controlled simulations. The simulation of a controlled q_P at a low level show that, although reaching a slightly lower maximum, it is reached 10 h earlier than the others.

Although, being of central interest in pre-production phases, the control of the specific growth rate μ does not show any clear advantages compared to constant feed rates but bears the risk of overfeeding due to strong feed rate adaptions to counteract the decreasing growth capacity. The simulations revealed, that the biomass specific uptake rate is unaffected by this decrease as feed rate adaption stagnates in course of time. This control concept could be adopted to real processes with the potential to stabilize production by keeping the substrate availability constant. Also, from a measurement point of view the specific substrate uptake rate is easier accessible in real-time as the growth or the production rate. An interesting but still very challenging concept from a measurement point of view, would be the direct control of the cell specific productivity. In addition to reach maximum product concentration faster in time a constant product formation rate could also have beneficial effects in product quality as subsequent steps of protein refolding and/or secretion can occur at a constant and optimal rate.

Table 4. Maximum overall product concentration and needed time for the three experiments with constant feeds (F_{in}) and the corresponding controller simulations with constant growth (μ) , substrate uptake (q_S) and production rate (q_P) .

level	constant F_{in}	constant μ	constant q_S	constant q_P
-	gL^{-1} (h)	gL^{-1} (h)	${\rm g}{\rm L}^{-1}$ (h)	gL^{-1} (h)
low	2.73(101)	2.83(100)	2.83(100)	2.67(90)
mid	2.37(44)	2.20(34)	2.29(33)	2.28(36)
high	1.21(21)	1.49(16)	1.45(12)	1.75(17)

5. CONCLUSION

Within this contribution novel possibilities to control the production phase of biotechnological fed-batch processes are shown. Based on a rather simple kinetic model functional control laws could be deduced by using the method of non-linear feedback linearization. Realistic model parameters and setpoints could be deduced from three different experiments. The controller simulations, revealed that due to the changing growth and production dynamics along the process different controller actions are needed. In the analyzed process a low constant feed and low growth. substrate uptake and production rate (q_P) setpoints lead to highest product concentration. Although the results does not indicate an increase in product under the usage of constant rates, such as q_P but it revealed the potential to shorten the needed time as well as the insurance of a constant reaction rate could lead to a better time space yield while ensuring a constant product quality. Besides future experimental verification of the presented findings in future work optimal setpoint trajectories could reveal potential to increase product amounts and to ensure product quality by the presented control approach.

ACKNOWLEDGEMENTS

Funding was provided by the Competence Center CHASE GmbH, funded by the Austrian Research Promotion Agency (FFG) (No 868615).

REFERENCES

- Abadli, M., Dewasme, L., Tebbani, S., Dumur, D., and Wouwer, A.V. (2021). An experimental assessment of robust control and estimation of acetate concentration in escherichia coli bl21 (de3) fed-batch cultures. *Biochemical Engineering Journal*, 108103.
- Almquist, J., Cvijovic, M., Hatzimanikatis, V., Nielsen, J., and Jirstrand, M. (2014). Kinetic models in industrial biotechnology-improving cell factory performance. *Metabolic engineering*, 24, 38–60.
- Cheng, D., Tarn, T.j., and Isidori, A. (1984). Global feedback linearization of nonlinear systems. In *The 23rd IEEE Conference on Decision and Control*, 74–83. doi: 10.1109/CDC.1984.272276.
- Dewasme, L., Fernandes, S., Amribt, Z., Santos, L., Bogaerts, P., and Wouwer, A.V. (2015). State estimation and predictive control of fed-batch cultures of hybridoma cells. *Journal of process control*, 30, 50–57.
- Kager, J., Tuveri, A., Ulonska, S., Kroll, P., and Herwig, C. (2020). Experimental verification and comparison of model predictive, pid and model inversion control in a penicillium chrysogenum fed-batch process. *Process Biochemistry*, 90, 1–11.
- Lee, S.Y. (1996). High cell-density culture of escherichia coli. Trends in biotechnology, 14(3), 98–105.
- Mears, L., Stocks, S.M., Sin, G., and Gernaey, K.V. (2017). A review of control strategies for manipulating the feed rate in fed-batch fermentation processes. *Journal of biotechnology*, 245, 34–46.
- Neubauer, P., Lin, H., and Mathiszik, B. (2003). Metabolic load of recombinant protein production: inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene in escherichia coli. *Biotechnology and bioengineering*, 83(1), 53–64.
- Selas Castiñeiras, T., Williams, S.G., Hitchcock, A.G., and Smith, D.C. (2018). E. coli strain engineering for the production of advanced biopharmaceutical products. *FEMS microbiology letters*, 365(15), fny162.
- Wechselberger, P., Sagmeister, P., Engelking, H., Schmidt, T., Wenger, J., and Herwig, C. (2012). Efficient feeding profile optimization for recombinant protein production using physiological information. *Bioprocess and biosys*tems engineering, 35(9), 1637–1649.
- Wilms, B., Hauck, A., Reuss, M., Syldatk, C., Mattes, R., Siemann, M., and Altenbuchner, J. (2001). High-celldensity fermentation for production of l-n-carbamoylase using an expression system based on the escherichia coli rhabad promoter. *Biotechnology and bioengineering*, 73(2), 95–103.
- Yee, L. and Blanch, H. (1992). Recombinant protein expression in high cell density fed-batch cultures of escherichia coli. *Bio/technology*, 10(12), 1550–1556.
- Zotică, C., Alsop, N., and Skogestad, S. (2020). Transformed manipulated variables for linearization, decoupling and perfect disturbance rejection. *IFAC-PapersOnLine*, 53(2), 4052–4057.