Computers and Chemical Engineering 173 (2023) 108203

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/cace

Simultaneous real-time estimation of maximum substrate uptake capacity and yield coefficient in induced microbial cultures

Don Fabian Müller [a], Daniel Wibbing [c], Christoph Herwig [a,b], Julian Kager [d,*]

[a] Institute of Chemical, Environmental and Bioscience Engineering, TU Wien, 1040 Vienna, Austria
[b] Competence Center CHASE GmbH, 4040 Linz, Austria
[c] Festo SE & Co. KG, 73770 Denkendorf, Germany
[d] Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 228A, 2800 Kgs. Lyngby, Denmark

ARTICLE INFO

Keywords:
Bioprocess model
Soft sensor
Maximum substrate uptake capacity
Real-time parameter estimation
Particle filter
Microbial cultivation
Bioprocess monitoring

ABSTRACT

In this work we present a soft sensor to accurately estimate the yield coefficient $Y_{X/S}$ and the substrate uptake capacity q_{Smax} in a cultivation process using offgas measurements and a nonlinear state observer with an underlying mechanistic model. The structural observability analysis of the mechanistic model showed that both parameters are observable given the available measurement information. In different simulation scenarios we analyzed under which conditions an accurate estimation is possible when measurements are uncertain. Testing the proposed soft sensor in-silico showed that q_{Smax} and $Y_{X/S}$ can be estimated with reasonable accuracy depending on the parameter sensitivity. Verification of the developed state and parameter estimation was carried out in induced *Escherichia coli* cultivations. Besides accurate prediction of living biomass, and substrate accumulation, decreasing $Y_{X/S}$ and q_{Smax} could be detected. Therefore, the developed soft sensor can be used to control induced cultures and to prevent overfeeding situations.

1. Introduction

The control of microbial fed-batch cultivation processes is predominantly based on the manipulation of the substrate feed rate (Mears et al., 2017), which is highly impacting the product formation rate of the culture (Kager et al., 2022). To avoid substrate accumulation, the controlled feed rate should not exceed a critical maximum substrate uptake capacity (q_{Smax}). Exceeding this biological limit can lead to overflow metabolism, toxic by-product formation and decrease of overall productivity (Santos et al., 2012; Neubauer et al., 2003; Pekarsky et al., 2019; Kager et al., 2020). The maximum substrate uptake capacity q_{Smax} is used as a parameter to describe the dependency of the cell specific substrate uptake rate (q_S) on the present substrate concentration in the medium (c_S) by different reaction kinetics e.g. Monod kinetics (Monod, 1949). In contrast to the traditional assumption of q_{Smax} being a constant parameter, previous studies revealed that q_{Smax} drastically varies during the process and often declines especially in the case of induced recombinant protein production (Lin et al., 2001; Neubauer et al., 2003; Reichelt et al., 2017). In addition, the activated recombinant genes decrease the biomass yield $Y_{X/S}$ as substantial amounts of substrate are needed for product formation. Therefore, monitoring of q_{Smax} and $Y_{X/S}$ in real-time would be highly beneficial in order to assure robust feed rate control without the risk of overfeeding.

There are several approaches presented in literature to estimate $Y_{X/S}$ or q_{Smax} dynamically in a bioprocess. Wechselberger et al. (2013) for example used elemental balancing in strictly substrate limited cultures to calculate the biomass yield during induced cultures. This approach could also be used to determine changing growth rates as a function of a dynamic temperature decrease (Sagmeister et al., 2013).

Other reported methodologies to determine cell characteristics are based on transient perturbations during the cultivation. Lin et al. (2001) found decreasing glucose uptake capacities in induced *E. coli* fed-batch cultures by frequently pulsing glucose just up to 0.5 to $2\,g\,L^{-1}$ to achieve short-term substrate unlimited conditions without significantly influencing overall cellular growth. The maximum glucose uptake capacity q_{Smax}, was then calculated based on the subsequent drop of the dissolved oxygen (pO_2). With this method decreasing maximum substrate uptake capacities after induction of the recombinant genes could be shown in *E. coli* cultivations (Neubauer et al., 2003). Although the analysis was applied offline and glucose was added manually with syringes, the authors pointed out the potential of automation and real-time applicability.

The automation and real-time application of this approach was realized within the so called probing control, first introduced by Akesson (1998) and later refined by de Maré (2016). Hereby the perturbation

* Corresponding author.
E-mail address: jukager@kt.dtu.dk (J. Kager).

https://doi.org/10.1016/j.compchemeng.2023.108203
Received 31 October 2022; Received in revised form 28 January 2023; Accepted 22 February 2023
Available online 25 February 2023

can be either an "up-pulse" meaning a short increase of the feed rate or a "down-pulse" meaning a short decrease of the feed rate. The response of the dissolved oxygen (pO_2) in function of the feed pulses was evaluated in real-time. In the case of substrate limited growth the perturbation of substrate availability will result in a peak of the pO_2 signal. In the case of substrate saturation however, the perturbation will not produce a pO_2 peak, since the culture is already operating at maximum capacity (q_{Smax}). In summary, the presence of a pO_2 peak will reveal if the culture is limited on substrate or saturated on substrate.

More generally, these approaches can be classified as extremum seeking control approaches. However, in contrast to the evaluation of the extremum condition via an objective function, here the control input perturbation is determined by a conditional statement through a peak detection algorithm (Dewasme and Wouwer, 2020).

By adopting these methods, the feed rate can be controlled to maintain the culture at q_{Smax} without the risk of overfeeding. However, the probing control method inherently requires to operate close to the maximum possible capacity q_{Smax} whereas methods based on elemental balancing only work on the other end, in strictly substrate limited conditions with neglectable substrate accumulation. Besides that pO_2 response based methods lead to deflections of the controlled dissolved oxygen. If deflections reach levels below 30% pO_2, unwanted shifts of the metabolism as acetate formation can occur. Therefore, these perturbations should be avoided as they can negatively influence process performance (Phue and Shiloach, 2005). Negative effects can be circumvented by building the probing control scheme exclusively on "down-pulsing" and feed supply interruption, which keeps the culture under strict glucose limitation and prevents the pO_2 from decreasing too rapidly (Whiffin et al., 2004). Although probing control and elemental balancing have shown their capability to determine changing cell characteristics such as q_{Smax} and $Y_{X/S}$ in real-time they are still limited in their applicability.

Kinetic models, which are more detailed and enable the description of cell internal behavior could be used to describe changing yields and uptake rates by algebraic relations. To address the above mentioned challenges, a kinetic overflow metabolism model was proposed by Pimentel et al. (2015), where a feedback linearizing control law was implemented to keep inhibitory by-product concentrations from overflow metabolism at low levels. Also other works successfully implement observers based on simple kinetic models for bioprocesses (Tuveri et al., 2022; Kager et al., 2018; Sinner et al., 2021) and specifically for overflow metabolism cultures (Bárzaga-Martell et al., 2021; Dewasme et al., 2013; Veloso et al., 2009). In this work however the focus was to detect shifts in the substrate uptake capacity and the biomass yield coefficient apart from the effects of the overflow metabolism bottleneck principle. Especially with combined state and parameter estimation and limited measurement information it is important to verify the parameter identifiability of the kinetic expressions (Margaria et al., 2004; Raue et al., 2009) and the system observability when the model is used for process monitoring (Lecca and Re, 2019). Further, a rather simple model reduces the need for prior knowledge on exact system kinetics and their parameters. Therefore, we took a basic Monod-based growth model and tried to estimate the kinetic parameters (q_{Smax} and $Y_{X/S}$) in real-time.

In addition to an adequate model and measurements, an observer algorithm is needed to accurately estimate the process states of interest. Today, several types of observers are available (Mohd Ali et al., 2015) and were also successfully used for bioprocesses (Goffaux and Wouwer, 2005). Considering bayesian type estimators, the traditional linear Kalman Filter is rarely applied, since the performance drops when dealing with nonlinear systems. The extended Kalman Filter (EKF) as a nonlinear variation of the Kalman Filter is more widely applied for bioprocesses (Yousefi-Darani et al., 2020). However, because of the linearization of the system through the Jacobians at each time step, the EKF lacks accuracy when dealing with highly nonlinear systems and can cause diverging predictions, for example when metabolic rates are changing rapidly between the end of the batch phase and start of a substrate feed (Wan and Van Der Merwe, 2000). To prevent those divergences, other nonlinear extensions of the Kalman Filter can be used such as the Unscented Kalman Filter (UKF) or the Particle Filter (PF), which are not based on a linearization, but on a direct propagation of numerous samples of the system (sigma points in case of the UKF and particles in case of the particle filter) through the nonlinear model. By that, highly nonlinear systems with non-Gaussian distributions can be estimated. Although the PF is computationally more demanding than the previously mentioned state estimators, but for additional parameter estimation it does not require to solve a nonlinear optimization problem over a time interval as for example required in a moving horizon estimator (MHE) (Rawlings and Bakshi, 2006).

In this contribution we aim to use a kinetic model combined with real-time offgas measurements to estimate the maximum substrate uptake rate q_{Smax} and the biomass yield coefficient $Y_{X/S}$ during induced fed-batch cultivations. To circumvent parameter identifiability and state observability issues the state estimator is based on a simple, unstructured model. By augmenting the model with the two parameters they can be additionally estimated without knowing their behavior beforehand. This novel approach enables to extend the validity range of these simple models to all phases of recombinant protein production processes. To do so we first set up an unstructured biomass growth model including Monod kinetics and elemental balances to link the measurable carbon evolution rate (CER) and oxygen uptake rate (OUR) to biomass growth. We considered two model parameters as variable over time, the yield coefficient $Y_{X/S}$ and the maximum substrate uptake capacity q_{Smax}. We analyzed the structural observability of the augmented model using Lie derivatives (Villaverde et al., 2016) and investigated parameter identifiability of $Y_{X/S}$ and q_{Smax} based on their sensitivity to CER and OUR measurements (Daume et al., 2019). After that we conducted simulation studies using a particle filter as a state observer to evaluate the feasibility of a simultaneous q_{Smax} and $Y_{X/S}$ estimation in different scenarios, including different feeding strategies and measurement accuracies. In the end the developed and parameterized state observer was verified on four different experiments to see overall estimation performance for biomass and substrate as well as the ability to indicate changing q_{Smax} and $Y_{X/S}$ during the production phases.

2. Material and methods

2.1. E. coli cultivations

In total five fed-batch cultivations were conducted, which are listed in Table 1. The first experiment P1 was solely used for the calibration of the observation model, described in Section 4.1. The other four experiments were used to verify the soft sensor algorithm containing the calibrated observation model.

The production host was a recombinant *Escherichia coli* BL21(DE3) strain producing the enzyme L-lactate dehydrogenase 1 (LDH), cultivated in 3.3 L Labfors 5 bioreactors (Infors, Bottmingen, Switzerland). A DeLisa minimal medium (DeLisa et al., 1999) was used containing 13.3 g L^{-1} KH_2PO_4, 4.0 g L^{-1} $(NH_4)2HPO_4$, 1.2 g L^{-1} $MgSO_4$, 1.7 g L^{-1} citric acid, 100 mg L^{-1} Fe(III) citrate, 13 mg L^{-1} $Zn(CH_3COO)_2$, 8.4 mg L^{-1} EDTA and 5 mL L^{-1} trace element solution containing 2.5 mg L^{-1} $CoCl_2.6H_2O$, 15 mg L^{-1} $MnCl_2.4H_2O$, 1.2 mg L^{-1} $CuCl_2.2H_2O$, 3.0 mg L^{-1} H_3BO_3 and 2.5 mg L^{-1} $Na_2MoO_4.2H_2O$. The medium was supplemented with Thiamine HCl to a final concentration of 4.5 mg L^{-1} and ampicillin to 50 mg L^{-1}. The pre-culture was cultivated in 1 L shake flasks for 16 h containing 100 mL DeLisa medium with 8 g L^{-1} glucose. 100 mL pre-culture was used to inoculate 1 L of batch medium. The batch phase was started with 15 g L^{-1} glucose as the carbon source. After the batch was finished a fed-batch phase with an exponential glucose feed profile was performed to keep the specific substrate uptake

Table 1
Overview of conducted experiments with *Escherichia coli*. The $q_{S,mean}$ values indicate the average q_S in the production phase.

Name	Purpose	Feeding	$q_{S,mean}$
P1	Model Calibration	Std. fed-batch	$0.25\,g\,g^{-1}\,h^{-1}$
P2	Verification	Std. fed-batch	$0.14\,g\,g^{-1}\,h^{-1}$
P3	Verification	Feed pulses	$0.35\,g\,g^{-1}\,h^{-1}$
P4	Verification	Feed pulses	$0.23\,g\,g^{-1}\,h^{-1}$
P5	Verification	Overfeeding	$0.31\,g\,g^{-1}\,h^{-1}$

rate q_S at a constant level. The feed rate was controlled according to Eq. (8) with a glucose concentration of $400\,g\,L^{-1}$.

The aimed q_s before induction was $0.25\,g\,g^{-1}\,h^{-1}$. After 24 h the transcription of the recombinant gene was induced by adding 1 mM isopropyl β-d-1-thiogalactopyranoside (IPTG) and adaption of the q_S setpoints to the values shown in Table 1. In addition to that, the temperature was lowered from 37 °C to 31.5 °C whereas the pH was controlled with NH_4^+ and H_3PO_4 to 6.75 throughout the whole process. The pO_2 was controlled to 30% by manipulating stirrer speed and oxygen content of the air inflow. The values for the induction phase were identified as beneficial in a previous study for recombinant production of inclusion bodies (IBs) in *E. coli* (Kopp et al., 2018).

2.2. Reference analytics

For reference analytics offline samples were taken every 2 h. The dry cell weight (DCW) and the optical density at 600 nm (OD_{600}) were measured for offline quantification of biomass. For DCW analysis 1 mL of culture broth was centrifuged at 20 000 g, the pellet was dried at 105 °C for 48 h and gravimetrically quantified in 1.5 mL reaction tubes. The glucose concentration was measured in an enzymatic analyzer (Cedex Bio HT, Roche, Basel, Switzerland) using the hexokinase based assay kit. The distinction between living cells and dead cells was made via flow cytometry according to Wurm et al. (2017), where the biomass sample was stained with 0.75 µM of the membrane potential sensitive DiBAC4(3) dye to stain dead cells and 3.0 µM of RH414 to stain all plasma membranes. A CyFlow® Cube 6 flow cytometer (Partec, Münster, Germany) was used for this purpose. For calculation of the dead cell mass X_d the ratio of dead cells to total cells was multiplied by the DCW. The remaining share was assumed to be viable cell mass.

2.3. Online analytics

Online measurements of the reactor setup included liquid temperature (ϑ_L), pH, dissolved oxygen tension (pO_2), the substrate feed rate F_R through a balance, airflow rate (F_{AIR}), oxygen flow rate (F_{O2}) and offgas analytics (BlueInOne, BlueSens, Herten, Germany) from which the oxygen uptake rate (OUR) and carbon evolution rate (CER) can be obtained. The calculation is based on F_{AIR} and F_{O2}, the molar gas volume under normal conditions V_M, the inert gas fraction $R_{a,inert}$ and the mole fractions of oxygen (O_2) and CO_2 in the offgas x_{O2} and x_{CO2} respectively (Aehle et al., 2012).

$$\mathrm{CER}_{meas} = \frac{F_{AIR} + F_{O2}}{V_M} \cdot (x_{CO2} \cdot R_{a,inert} - x_{CGin}) \tag{1}$$

$$\mathrm{OUR}_{meas} = \frac{F_{AIR} + F_{O2}}{V_M} \cdot (x_{O2} \cdot R_{a,inert} - x_{OGin}) \tag{2}$$

with

$$R_{a,inert} = \frac{1 - x_{OGin} - x_{CGin}}{1 - x_{O2} - x_{CO2} - \left(1 - \frac{Y_{wet}}{x_{OAIR}}\right)} \tag{3}$$

In order to obtain the measurement errors for CER and OUR the errors of the prime variables F_{AIR}, F_{O2}, x_{O2} and x_{CO2} were taken from the equipment documentation of the used mass flow controllers and the offgas analyzer and were propagated through the equation system (Eqs. (1) & (2)) using Gaussian error propagation (Müller et al., 2022).

3. Computational framework

3.1. Observer & plant model

Within this work we distinguish between two models: The observer model, which was used in the state observer (particle filter) to estimate the state and parameters based on incoming offgas measurements, and the plant model, which was used to generate the plant feedback in the simulation study. Besides the usage of different model parameters, the plant model includes an algebraic description for a decreasing q_{Smax} over time. The observation model contains q_{Smax} as an static parameter. An unstructured kinetic growth model, was used as the basis for both models. It is composed of:

- four states: Volume, viable and dead biomass, susbtrate $x = [V, X_v, X_d, S]^T$
- two measurements: CER and OUR $y = [\mathrm{CER}, \mathrm{OUR}]^T$
- and two inputs: Feed flow rate and sampling rate $u = [F_R, F_{out}]^T$.

The state equations (Eq. (4) – Eq. (7)) represent the dynamics of the system by mass balances. Although concentration balances are more commonly used in mechanistic bioprocess models to achieve a more concise description of the kinetics, we used mass balances to simplify the incorporation of first principle elemental balances.

$$\frac{dV}{dt} = F_R - F_{out} \tag{4}$$

$$\frac{dX_v}{dt} = Y_{X/S} \cdot q_S \cdot X_v - k_d \cdot X_v - F_{out} \cdot \frac{X_v}{V} \tag{5}$$

$$\frac{dX_d}{dt} = k_d \cdot X_v - F_{out} \cdot \frac{X_d}{V} \tag{6}$$

$$\frac{dS}{dt} = F_R \cdot c_{SR} - q_S \cdot X_v - F_{out} \cdot \frac{S}{V} \tag{7}$$

The state equations (Eq. (4)–Eq. (7)) denote the liquid volume, the living biomass, the dead biomass and the substrate mass, respectively. Rearranging the substrate mass balance (Eq. (7)) to the feed rate F_R and neglecting the outlet flow rate F_{out} yields the equation for open loop q_S control,

$$F_{Rw}(t) = q_{Sw}(t) \cdot \frac{X_v(t)}{c_{SR}} \tag{8}$$

with F_{Rw} being the calculated feed rate setpoint in $L\,h^{-1}$, q_{Sw} being the setpoint of q_S in $g\,g^{-1}\,h^{-1}$, X_v being the viable cell mass in g and c_{SR} being the glucose concentration in the feeding solution in $g\,L^{-1}$.

The CER and OUR model outputs can be obtained by solving the elemental balances (carbon balance and degree of reduction balance) (Roels, 1983) to calculate the yield coefficients for CO_2 production and O_2 uptake as a function of the biomass yield coefficient $Y_{X/S}$. Therefore, CER and OUR can be directly derived from the model outputs by Eqs. (9) and (10) with the C-molar masses of glucose ($M_S = 30\,g\,mol^{-1}$) and biomass ($M_X = 26.5\,g\,mol^{-1}$) as well as the degrees of reduction for glucose ($DOR_S = 4$), for oxygen ($DOR_{O2} = -4$) and for biomass ($DOR_X = 4.113$).

$$\mathrm{CER} = \left(\frac{1}{M_S} - \frac{Y_{X/S}}{M_X}\right) \cdot q_S \cdot X_v \tag{9}$$

$$\mathrm{OUR} = \left(\frac{DOR_S}{M_S} - \frac{Y_{X/S} \cdot DOR_X}{M_X}\right) \cdot q_S \cdot \frac{X_v}{DOR_{O2}} \tag{10}$$

The specific substrate uptake rate q_S was described using Monod kinetics, based on the model parameters q_{Smax} and k_S. For the calculation of kinetics the component masses have to be converted to concentrations ($c_S = \frac{S}{V}$).

$$q_S = q_{Smax} \cdot \frac{c_S}{k_S + c_S} \tag{11}$$

Table 2
Plant model parameters and model constants.

Parameter θ	Value	Unit
$Y_{X/S}$	0.45	$g\,g^{-1}$
k_d	0.01	$g\,g^{-1}\,h^{-1}$
c_{SR}	400	$g\,L^{-1}$
$q_{Smax,init}$	1.2	$g\,g^{-1}\,h^{-1}$
k_S	0.004	$g\,L^{-1}$
M_X	26.5	$g\,Cmol^{-1}$
M_S	30	$g\,Cmol^{-1}$
M_{O2}	32	$g\,mol^{-1}$
M_{CO2}	44	$g\,Cmol^{-1}$
DOR_X	4.113	$mol_{e^-}\,Cmol^{-1}$
DOR_S	4	$mol_{e^-}\,Cmol^{-1}$
DOR_{O2}	−4	$mol_{e^-}\,Cmol^{-1}$

In the plant model the q_{Smax} parameter was additionally decreased by Eq. (12) after a process time of 11 h to investigate if the state estimator, including the observation model is able to detect the changing maximum capacity.

$$q_{Smax}(t) = \begin{cases} q_{Smax,init} & \text{for } t < 11\,\text{h} \\ q_{Smax,init} \cdot e^{-0.09\,\text{h}^{-1}\cdot(t-11\,\text{h})} & \text{for } t \geq 11\,\text{h} \end{cases} \tag{12}$$

Besides of q_{Smax} all other model parameters were kept constant to study specifically the estimation of the unknown dynamics of this parameter. All parameter values of the plant model are given in Table 2.

The basic model structure of the observation model is the same as for the plant model. However, some adaptations were made. To enable model parameter estimation in real-time the state vector can be augmented with the target parameters, including zero dynamics in the respective differential equations (Kager et al., 2018; Patwardhan et al., 2012). We augmented the state vector with the parameters $Y_{X/S}$ and q_{Smax} yielding the new augmented state vector

$$\bar{x} = [V, X_v, X_d, S, Y_{X/S}, q_{Smax}]^T.$$

For the simulation study the remaining model parameters were assumed to be perfectly known and equal to the plant model. For the experimental validations $Y_{X/S}$, q_{Smax} and k_d were parameterized based on the experiment P1 (Fig. 1).

3.2. Model analysis and parameterization

For a successful parameter identification the model parameters must be identifiable from the measurements that are available. For model calibration (Section 4.1) the parameters $Y_{X/S}$, q_{Smax} and k_d are supposed to be identified using the offline measurements, which in this case are the states: viable biomass (X_v), dead biomass (X_d) and substrate mass (S). The structural identifiability can be evaluated by computing the rank of the sensitivity matrix. The local sensitivity of a model parameter θ_p on state x_i can be calculated by the partial derivative of x_i with respect to the parameter θ_p.

$$S_{i,p} = \frac{\partial x_i}{\partial \theta_p} \tag{13}$$

The sensitivity matrix is then calculated as the Jacobian matrix of the states with respect to the parameter vector.

$$S = \frac{\partial(x_1, \ldots, x_m)}{\partial(\theta_1, \ldots, \theta_n)} \tag{14}$$

If the sensitivity matrix S has full rank meaning the columns are linearly independent, the system is structurally identifiable.

The model parameters q_{Smax}, $Y_{X/S}$ and k_d were estimated with regard to the offline measured data $X_{v,\text{meas}}$, glucose mass (S_{meas}) and dead cell mass ($X_{d,\text{meas}}$) of process P1 (Fig. 1). The optimal parameter set was obtained by minimization of the squared error sum (J_{SSE}) between the offline reference measurements ($Y_{i,\text{meas}}$) and the simulated model states ($Y_{i,\text{model}}$).

$$J_{SSE} = \sum_{i=1}^{M} (Y_{i,\text{model}} - Y_{i,\text{meas}})^2 \tag{15}$$

As only approximate parameters from the calibration process P1 were needed, no weighting of the different measurements was applied. The estimated parameters were then used as initial values in the observation model of the experimental validation experiments in Section 4.3.

3.3. Structural generalized observability

To assess, whether the states and model parameters can be estimated in real-time, the structural observability of the system was analyzed. After augmentation of the two model parameters $Y_{X/S}$ and q_{Smax} into the state vector the generalized observability–identifiability matrix was constructed using Lie derivatives and the rank was determined (Villaverde et al., 2016). The first order Lie derivative of the output function $g(x)$ with respect to the state function $f(x,u)$ is given as

$$L_f g(x) = \frac{\partial g(x)}{\partial x} \cdot f(x,u). \tag{16}$$

The higher order Lie derivatives are calculated based on the previous Lie derivative by

$$L_f^i g(x) = \frac{\partial L_f^{i-1} g(x)}{\partial x} \cdot f(x,u) \tag{17}$$

and stacked over each other to derive the nonlinear observability matrix $\mathcal{O}(x)$.

$$\mathcal{O}(x) = \begin{pmatrix} \frac{\partial g(x)}{\partial x} \\ \frac{\partial L_f g(x)}{\partial x} \\ \frac{\partial L_f^2 g(x)}{\partial x} \\ \vdots \\ \frac{\partial L_f^{i-1} g(x)}{\partial x} \end{pmatrix} \tag{18}$$

If $\mathcal{O}(x)$ has full rank, the structural observability is given. The minimum number of Lie derivatives for which $\mathcal{O}(x)$ can be of full rank is given by $\frac{n+q}{m} - 1$ with n being the number of states and q being the number of augmented parameters (Villaverde et al., 2016). For the calculation of Lie derivatives and evaluation of structural identifiability the STRIKE-GOLDD Matlab® toolbox was used, which was published by Villaverde et al. (2016).

3.4. Particle filter

In this study we used a sequential importance resampling (SIR) particle filter algorithm to estimate the unknown model states and parameters from the OUR and CER measurements. The general algorithm of the particle filter is based on Simon (2006), Rawlings and Bakshi (2006). A particle filter contains multiple samples in the state space (particles) representing the distribution of the estimated state. Each particle is assigned to a corresponding weight reflecting the likelihood of the respective particle to fit the measurements. In contrast to the KF or EKF the distribution can be non-Gaussian and multi-modal. Usually, the more particles are sampled the better the distribution can be represented. Each particle is propagated simultaneously through the model in a prediction step. The process prediction was performed using the observation model (Section 3.1) with the augmented state vector $\bar{x}$, where the dynamics of the parameters $Y_{X/S}$ and q_{Smax} were set to 0 (exogenous form), with artificial noise on the parameters. This allows for simultaneous state- and parameter estimation. After that in the correction step, measurements are used to update the particle weights. Over time, more and more particles get very small weights causing a

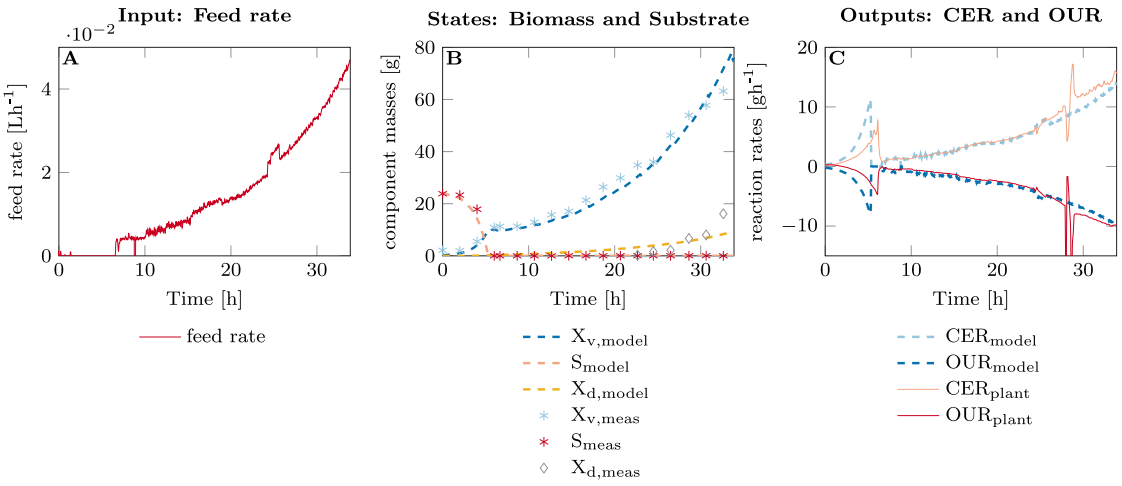

Fig. 1. Open-loop model simulation after parameter estimation based on the process P1 data. The estimated parameter set is: $q_{Smax} = 1.3395\,\mathrm{g\,g^{-1}\,h^{-1}}$, $Y_{X/S} = 0.4371\,\mathrm{g\,g^{-1}}$, $k_D = 0.0123\,\mathrm{g\,g^{-1}\,h^{-1}}$ with a relative error (NRMSE) of 5.12% for X_v, 7.85% for S and 47.4% for X_d. In C the corresponding reaction rates CER and OUR for the open loop model and the measurements are shown where negative reaction rates (OUR) refer to uptake of the component by the cell and positive reaction rates (CER) refer to an evolution of the component.

decrease of prediction accuracy (degeneration). In this case particles with higher weights are resampled if the number of effective particles drops under a certain threshold (Li et al., 2014). The integration time was defined to be 0.001 h or 3.6 s since the stiffness of the ODE system required a frequent solver step size. The update step (k) however was just performed when new measurements were available with 30 s as smallest interval.

Initialization. In the first time step ($k = 0$) a sample of $N = 100$ particles is randomly drawn respecting a known probability density function of the initial state ($x_0^i : i = 1, \ldots, N$). In this work, Gaussian distribution was assumed for the initial distribution described by the initial state covariance matrix

$$\Sigma_0 = diag[0, 0.05, 0, 0.05, 0.0004, 0.0004]. \tag{19}$$

Prediction step. Time propagation of the particles through the model state transition equation $f(x)$ (Eq. (5) – Eq. (7)) was calculated with additive process noise w_k^i.

$$\tilde{x}_k^i = f(x_{k-1}^i) + w_{k-1}^i (i = 1, \ldots, N) \tag{20}$$

Noise levels for particle propagation were assumed to be Gaussian with a covariance matrix

$$\Sigma_{wk} = diag[0, 0, 0, 0, 0.000025, 0.0003]. \tag{21}$$

It can be seen in Σ_{wk} that just the last two states $Y_{X/S}$ and q_{Smax} are assumed to be uncertain, which are the parameters to be estimated in real-time within this study. There is no direct state transition variance on the other states, but they are influenced by the parameters uncertainty through the model.

Correction (resampling) step. The measurement function was defined as $y(x) = [\mathrm{CER}, \mathrm{OUR}]$. The model outputs (Eqs. (1) and (2)) are compared in the correction step with the measured rates obtained from offgas-analysis (Eqs. (9) and (10)). The relative likelihood $\tilde{q}_k^i$ is computed for every particle analogous to Sinner et al. (2022). In the case of new measurements correction is performed by multiplication of the relative likelihood $\tilde{q}_k^i$ with the respective particle weight to update the particle weights. Systematic resampling is carried out when the effective particle ratio N_{eff} drops under the critical value $N_{\mathrm{eff,crit}} = 0.25$.

3.5. Simulation study and real-time implementation

Simulation studies and real-time calculations during the experiments were executed within MATLAB® and Simulink® (Version R20-22a, MathWorks, Natick, Massachusetts, USA). The same Simulink observer implementation was used for the simulation and the experimental validation experiments.

Real-time process control and communication to the equipment was realized through the Lucullus Process Information Management System PIMS (Securecell, Urdorf, Switzerland).

Lucullus was connected via OPC to a computer where in MATLAB®, the data preprocessing, the feed control and the state observer for the different experiments was running. The numerical derivative of the feed weight signal was computed and smoothed with a Savitzky–Golay-Filter approach (Savitzky and Golay, 1964) to obtain the feed rate F_R used as model input.

4. Results and discussion

4.1. Model analysis and calibration

Before applying the state observer on the plant model as well as on the different validation experiments (P2–P5) the observer model was structurally analyzed and parameterized using the model calibration experiment (P1).

4.1.1. Observer model calibration and observability

The parameters of the observer model were obtained from one single experiment (P1, Table 1). The purpose of the parameter estimation was hereby not to obtain a perfect simulation model but to get a reasonable set of initial parameter values that can be used to initialize the particle filter.

The model calibration resulted in the estimated parameter set of $q_{Smax} = 1.3395\,\mathrm{g\,g^{-1}\,h^{-1}}$, $Y_{X/S} = 0.4371\,\mathrm{g\,g^{-1}}$ and $k_D = 0.0123\,\mathrm{g\,g^{-1}\,h^{-1}}$. The open loop model simulation and its fit on the experiment P1 is shown in Fig. 1. In Fig. 1A the exponential feed rate profile of process P1 is shown, which was applied to control q_S to $0.25\,\mathrm{g\,g^{-1}\,h^{-1}}$ by open loop feed control according to Eq. (8). Fig. 1 shows the system states living biomass (X_v), dead biomass (X_d) and substrate mass (S) from offline measurements as well as the model simulation. The model fits the offline measured data well with a small normalized root mean square errors (NRMSE) of 5.12% for the living biomass fit and 7.85% for the substrate fit. However, the higher NRMSE of 47.4% for the dead biomass fit indicates that a constant k_d value does not describe the cell death accurately. Furthermore, in the first part of the fed-batch phase (7 h to 25 h) the model underestimates the measured biomass, whereas in the last few hours of the process the biomass is being slightly

overestimated. This indicates, that a constant biomass yield coefficient $Y_{X/S}$ does also not fully capture the dynamics of the biomass growth, even if $Y_{X/S}$ is fitted to the measured data. $Y_{X/S}$ must be decreasing throughout the process to counteract this missmatch.

Fig. 1C shows the measured and simulated CER and OUR. When comparing CER and OUR derived from the model outputs by elemental balances (Eqs. (9) and (10)) to the online measured CER and OUR (Eqs. (1) and (2)) it can be seen that the dervied rates are higher in the batch phase and indicate an earlier batch end. Also, in the late fed-batch phase (after 24 h) the derived rates start to deviate more from the online measurements. Overall, the model is able to describe the process dynamics of P1 and the estimated parameters can be used as the initial parameter values for the observer model.

To analyze if the parameters $Y_{X/S}$ and q_{Smax} can be estimated by CER and OUR the reduced sensitivity matrix S defined in Section 3.2 was computed. The matrix is included in the Appendix and has a full rank of 2. Therefore, structural parameter identifiability based on CER and OUR can be confirmed. For the assessment of structural observability the generalized observability–identifiability matrix was constructed using three Lie derivatives, which can be used to check for state observability as well as parameter identifiability that are augmented in the state vector. The generalized observability–identifiability matrix $\mathcal{O}(x)$ derived in Section 3.3 has a rank of 5, which is one less than the number of augmented states. By calculating the rank of different reductions of the matrix the dead biomass X_d was identified as non observable. To obtain a fully observable model either a direct measurement of dead biomass or measurements of total and viable biomass would be needed. However, for this application the dead biomass can be neglected to some extend as it is not crucial for the real-time estimation of the parameters $Y_{X/S}$ and q_{Smax}. Overall, the formal model analysis shows that it is possible to uniquely reconstruct the model parameters and states (except of dead biomass) from the measured CER and OUR signals.

4.1.2. Plant model dynamics and sensitivities

A simulation result of the plant model can be seen in Fig. 2A and B. In order to mimic a real process, initial states of $0.25\,\mathrm{g\,L^{-1}}$ viable biomass X_v, $0\,\mathrm{g\,L^{-1}}$ dead biomass X_d, $22.5\,\mathrm{g\,L^{-1}}$ glucose S and a starting volume of 1 L were defined. The feed rate was increased exponentially after 8 h to keep the specific substrate uptake rate q_S at a constant level of $0.4\,\mathrm{g\,g^{-1}\,h^{-1}}$. Eq. (8) was taken to calculate the feed. The biomass yield coefficient $Y_{X/S}$ was kept constant throughout the process. After a process time of 11 h, induction was assumed and the q_{Smax} was exponentially decreased according to Eq. (12). This decreasing profile was inspired by earlier studies, which revealed that q_{Smax} continuously decreases in induced recombinant *E. coli* processes (Neubauer et al., 2003).

To see the impact of the decrease of q_{Smax} on our measurement information the time resolved sensitivity for q_{Smax} and $Y_{X/S}$ on CER and OUR was computed and its normalized sum is displayed in Fig. 2C. Normalization occurred by a division of viable cell mass X_v. Although the resulting sensitivity matrix has full rank throughout the whole process their values change over time. These changing values affect the practical identifiability and subsequently the real-time estimation of the parameters. Both parameters are most sensitive in the batch phase, where due to unlimited growth the actual substrate uptake q_S is at maximum q_{Smax}. During the first fed-batch phase with q_S beeing 33 % of q_{Smax} the sensitivities of both parameters drop to lower levels. This means that changes in parameters have a lower effect on the measurements and the measurements are therefore less informative in this phase. Later in the fed-batch phase when q_{Smax} was exponentially decreased (11 h) the sensitivity for q_{Smax} rises again. This can be explained by the fact that if q_S and q_{Smax} are close to each other, small changes in q_{Smax} are influencing the model outputs CER and OUR more significantly. Therefore, the sensitivity is high, which leads to a better practical identifiability. On the other hand, if q_S is way below q_{Smax}, changes in q_{Smax} do not affect CER and OUR much. Therefore, the sensitivity is small with a poor practical identifiability.

In summary, the results show that especially in substrate saturated phases, the q_{Smax} parameter is highly sensitive, which usually means that it can be well identified and estimated. The sensitivity of $Y_{X/S}$ however seems to be more dependent on the overall q_S level and is not necessarily better in substrate saturated phases. Overall, we can expect good estimates of q_{Smax}, if actual substrate uptake q_S gets close to q_{Smax}. Although there might be phases with better and poorer estimation, the full rank of the time resolved sensitivity matrix indicates the possibility to simultaneously estimate both parameters together.

4.2. Simulation study

A simulation study was carried out to investigate different scenarios and to tune the state estimator to achieve satisfactory results. The particle filter described in Section 3.4 was therefore used to reconstruct the parameter trajectories shown in Fig. 2B from the measurable model outputs CER and OUR (Eqs. (9) and (10)). The particle filter contains the observation model, which in contrast to the plant model assumes a constant q_{Smax} during the whole process. The observer model was initialized with deflected parameters using a Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$ with a standard deviation (σ) of $0.15\,\mu$ with μ being the true mean value. Therefore, the true parameter trajectories and initial conditions for $Y_{X/S}$ and q_{Smax} are unknown to the particle filter. Fig. 2C shows the time-resolved sensitivity of both parameters with respect to CER and OUR since these are the online measurements that were obtained from the system. The parameter sensitivities will become important for the practical identifiability in the subsequent simulation study.

In the following we investigated different scenarios to reconstruct the ground truth from the plant model by the state estimator considering (i) perfect measurements with $\sigma = 0$, (ii) uncertain measurements with Gaussian noise $\mathcal{N}(\mu, \sigma^2)$ with $\sigma = 0.03\,\mu$, (iii) different feed profiles and (iv) the time-resolved sensitivity to tune the parameter perturbation during the state propagation step.

4.2.1. Measurement accuracy

To investigate the estimation accuracy for unknown q_{Smax} and $Y_{X/S}$ the simulations were carried out by deflecting both parameters using a Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$ with $\sigma = 0.15\,\mu$ with μ being the reference plant model values of $q_{Smax} = 1.2\,\mathrm{g\,g^{-1}\,h^{-1}}$ and $Y_{X/S} = 0.45\,\mathrm{g\,g^{-1}}$. The state propagation variance in the observation model was identical as in Eq. (21) (Section 3.4). The same state transition variance was used for all following simulation studies as well as for the experimental verification.

Fig. 3 shows the estimation results for a perfect measurement information assumption (Fig. 3A) and for uncertain measurement information (Fig. 3B) on the measurements CER and OUR. In case of perfect measurements (Fig. 3A), the q_{Smax} estimate jumps to the correct reference value of $1.2\,\mathrm{g\,g^{-1}\,h^{-1}}$ in the beginning of the process and then perfectly tracks the reference value. After 11 h when the reference value of the plant model starts to decrease, the estimation is lagging a little bit behind, but the estimate follows the reference value with a overall low error (RMSE = $0.0498\,\mathrm{g\,g^{-1}\,h^{-1}}$). The yield coefficient $Y_{X/S}$ is perfectly estimated over the whole process in case of perfect measurement information. Overall, the root mean square error (RMSE) between estimated parameters and the ground truth are rather low with $0.0498\,\mathrm{g\,g^{-1}\,h^{-1}}$ for q_{Smax} and $5.24 \times 10^{-4}\,\mathrm{g\,g^{-1}}$ for $Y_{X/S}$.

In the case of realistic measurement errors ($\sigma = 0.03\,\mu$) (Fig. 3B) the estimations are much worse for q_{Smax} (RMSE = $0.610\,\mathrm{g\,g^{-1}\,h^{-1}}$) as well as for $Y_{X/S}$ (RMSE = $0.0614\,\mathrm{g\,g^{-1}}$). Already at the beginning of the process the state estimator is not able to revert the parameter estimates from the deflected initial conditions to the reference values of the plant model. At the beginning of the fed-batch phase (8 h) the $Y_{X/S}$ estimations are approaching the reference values well,

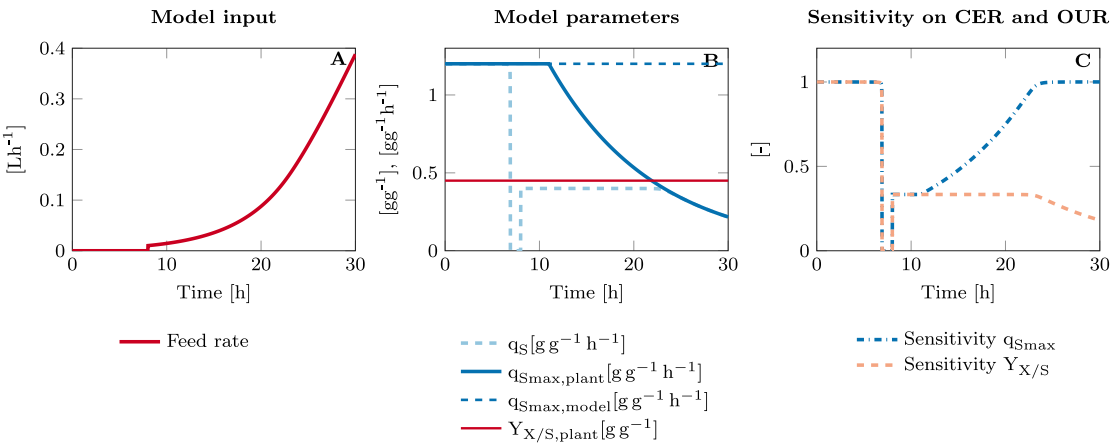

Fig. 2. Simulation study design with the feed rate trajectory as the control input (A), the model parameter trajectories of q_{Smax} and $Y_{X/S}$ (B) and normalized parametric sensitivities on CER and OUR (C).

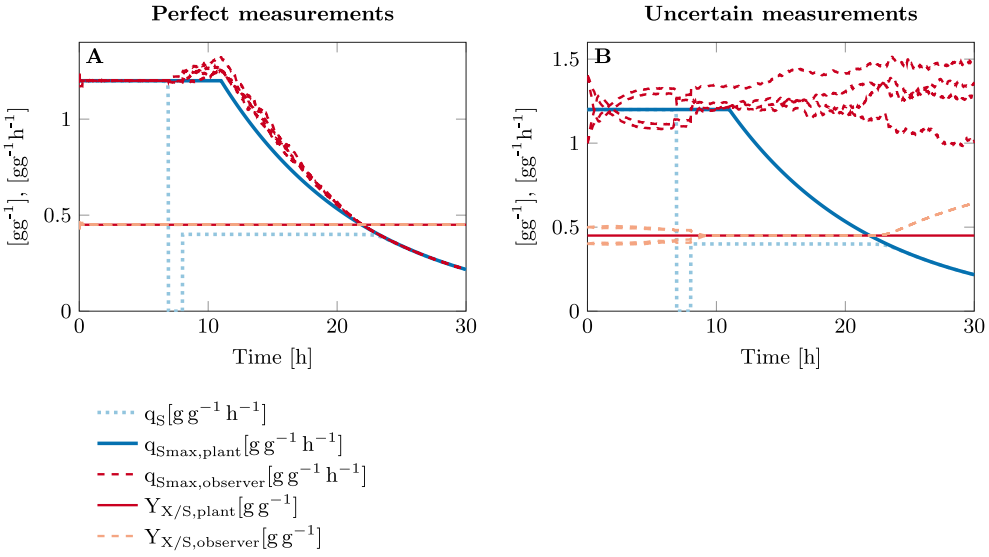

Fig. 3. Estimation of q_{Smax} and $Y_{X/S}$ based on: (A) perfect measurement information and (B) uncertain measurements with $\mathcal{N}(\mu, (0.03\,\mu)^2)$. The state observer was initialized with all four combinations (dotted lines) of deflected start parameters ($\mathcal{N}(\mu, (0.15\,\mu)^2)$).

whereas the q_{Smax} estimations remain inaccurate. Estimations remain also unaffected by the decline of the reference value (11 h onwards) and spread out in a random walk fashion. Since the sensitivity of q_{Smax} on CER and OUR is low, estimation during this phase seems very challenging with uncertain measurements. Although parametric sensitivity increases again towards the end, when actual q_S approaches q_{Smax} (see Fig. 2C), the estimations remain unaffected and does not reveal the decreasing q_{Smax} of the plant model.

The $Y_{X/S}$ estimations on the other hand are increasing towards the end of the process with values significantly above the reference value of $0.45\,\mathrm{g\,g^{-1}}$. The reason for the increasing $Y_{X/S}$ estimation at this time point is, that the substrate feed is now carried out above the maximum uptake capacity, which results in overfeeding. With a higher $Y_{X/S}$, CER and OUR balances are still closing. There are two options for the state estimator to match the model outputs. The first (and correct) option would be to decrease the q_{Smax} estimate to the reference value and thereby correctly detect the overfeeding regime. The other option is however to increase the $Y_{X/S}$ estimate to shift the elemental carbon balance more towards cell growth, which lowers CER. This however neglects the overfeeding and assumes that the added substrate is still completely converted to cell mass. This option is presumably taken by the particle filter because there are no particles available with a q_{Smax} estimate value near the true plant reference. The q_{Smax} estimates already completely separated from the reference at this point in contrast to the $Y_{X/S}$ estimates.

The results show that the overall parameter estimation is not satisfying under realistic CER and OUR measurement errors and needs to be further improved.

4.2.2. Feedback of parameter sensitivity

In this section we utilize the normalized time-resolved sensitivity of both parameters q_{Smax} and $Y_{X/S}$ shown in Fig. 2C for the particle filter to increase the accuracy of state estimation with uncertain CER and OUR measurements.

The basic idea hereby is to tune the state propagation variance of the observation model (Eq. (21)) based on the current parametric sensitivity. Hereby, the variance is increased if the parameter is highly sensitive and lowered if it is less sensitive. This ensures that the parameters with the highest impact on the measurements are adapted to a greater extent. To do so, the time-resolved normalized sensitivity is fed back into the particle filter and multiplied with the state transition variance (Σ_{wk}, Eq. (21)) of q_{Smax} and $Y_{X/S}$. Thereby, the parameter with very low sensitivities remain static and sensitive parameters are changed more drastically the higher the sensitivity is. This, on the one

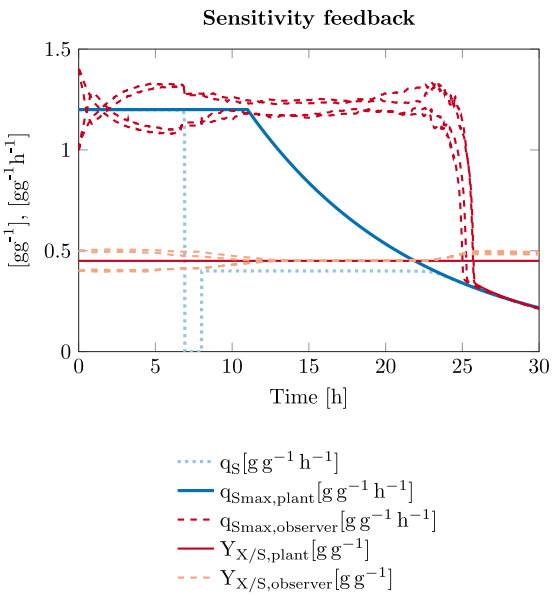

Fig. 4. Estimation of q_{Smax} and $Y_{X/S}$ based on uncertain measurements with $\mathcal{N}(\mu, (0.03\,\mu)^2)$ and feedback of parametric sensitivity for noise level tuning. The state observer was initialized with all four combinations (dotted lines) of deflected start parameters $\mathcal{N}(\mu, (0.15\,\mu)^2)$).

hand should prevent parameter estimates from randomly deviating if they are practically non identifiable (meaning the sensitivity is close to 0), and on the other hand should boost the estimation of parameters that are highly sensitive in that moment.

Fig. 4 shows the estimation results with the sensitivity feedback. Besides the sensitivity feedback and the dynamical adjustment of the parameter state transition variance, the simulation scenario is identical to the one used in Fig. 3B. It can be seen that in contrast to Fig. 3B the q_{Smax} estimates at the end of the process are now reverting to the reference value of the plant model. This effect is reached as the sensitivity for q_{Smax} in this phase is at the maximum (Fig. 2C) and therefore a strong adaption of the respective parameter is possible.

Compared to Fig. 3B, also the $Y_{X/S}$ estimates are deviating to a lesser extent from the reference value. Again, compared to Fig. 3B the RMSE was reduced from $0.610\,\mathrm{g\,g^{-1}\,h^{-1}}$ to $0.404\,\mathrm{g\,g^{-1}\,h^{-1}}$ in case of q_{Smax} (1.5-fold reduction) and from $0.0614\,\mathrm{g\,g^{-1}}$ to $0.0301\,\mathrm{g\,g^{-1}}$ in case of $Y_{X/S}$ (2.0-fold reduction). Although the effect was just analyzed for the particle filter estimator, this approach would be also applicable to other bayesian estimators such as the KF, EKF and UKF since all of them rely on the definition of a process noise covariance matrix that determines how much belief to give to the model. Making it variable as a function of parametric sensitivitities might also increase the performance of them. Here, a higher parametric sensitivity would mean a higher process noise covariance and therefore a bigger influence of the measurements than the model. Comparable effects to the one observed here with the particle filter can be therefore expected.

Despite the slight improvements of the estimation accuracy, especially the q_{Smax} estimation is still more of a qualitative information. In this configuration the q_{Smax} estimation could indicate whether overfeeding occurs. This is the case when the current uptake q_S is close or exceeds q_{Smax} as in this situation the q_{Smax} parameter is highly sensitive to CER and OUR. However, in normal operating regions (below overfeeding regimes) the sensitivity is still too low to make any distinctions between different levels of maximum uptake capacities. To get a real quantitative statement for q_{Smax} the parametric sensitivity during this operational space needs to be enhanced.

4.2.3. Superimposed feed pulses for enhanced sensitivity

In this work we adopted the principle of superimposed feed pulses also known as probing control (de Maré, 2016) to our model based state estimation approach. Therefore, we applied $10\,\mathrm{mg\,g^{-1}}$ (substrate/biomass) feed pulses every 30 min onto the underlying exponential feed profile. The effect on the parametric sensitivity of this short pulses can be seen in Fig. 5B, whereas the corresponding estimation results are displayed in Fig. 5A. It can be seen that every time a pulse is given, the sensitivity of q_{Smax} peaks. Therefore, the q_{Smax} estimation is more effective at those time points. For the simulation conducted in Fig. 5 we utilized the same sensitivity feedback procedure as introduced in Section 4.2.2 as well as measurement errors applied in Section Section 4.2.1. The feed pulses can also be seen in the specific substrate uptake rate q_S peaking to almost q_{Smax} for a short period of time.

The enhanced sensitivity has a major influence especially on the q_{Smax} estimation accuracy. Now, even in case of realistic measurement uncertainty, the estimation shows a reasonable tracking accuracy of the reference value with a RMSE of $0.0726\,\mathrm{g\,g^{-1}\,h^{-1}}$, which is a 8.4-fold reduction of RMSE compared to the estimation shown in Section 4.2.1 and a 5.6-fold reduction to the results obtained in Section 4.2.2. Also, the estimation accuracy of $Y_{X/S}$ is improved with a RMSE starting from $0.0614\,\mathrm{g\,g^{-1}}$ in Section 4.2.1, to $0.0301\,\mathrm{g\,g^{-1}}$ in Section 4.2.1 and now being at $0.0246\,\mathrm{g\,g^{-1}}$. All RMSE values of the different simulation scenarios are listed in Table 3.

The results of the simulation study indicate, that despite the low parametric sensitivity of q_{Smax} in a traditional substrate limited fed-batch process and the correlation between $Y_{X/S}$ and q_{Smax}, the simultaneous estimation of both parameters from CER and OUR measurements is possible. In comparison to the traditional probing control methodology the presented soft sensor is able to estimate $Y_{X/S}$ and q_{Smax} while overall maintaining physiologically controlled conditions.

4.3. Experimental validation

In order to verify the capability of the developed and tuned soft sensor to estimate the observable system states and the two parameters ($Y_{X/S}$ and q_{Smax}), four *E. coli* fed-batch cultivations were conducted. The cultivations listed in Table 1 were similar to the scenarios investigated in the simulation study. The cultivations were carried out according to the description in Section 2.1. The state transition covariance was identical to the simulation study (Section 4.2). The measurement covariance however was not assumed with a static relative measurement error like in the simulation study, but propagated from the errors of the measured prime measurement as described by Müller et al. (2022). The initial values of the parameters were taken from the model calibration in Section 4.1. The soft sensor configuration was identical for all four processes.

An overview of the different processes and their state estimation results are displayed in Fig. 6. Within Table 4 the state estimation errors are listed in comparison to a feed-forward model simulation using the calibrated model from Section 4.1.1. The parameter estimation results for the single experiments are presented in Figs. 7 to 9. As process P3 and P4 were replicates, results of P3 can be found in the appendix.

The processes mainly differ by their feed profiles. Process P2 has a similar feed profile as the calibration experiment P1. P3 and P4 included superimposed feed pulses to enhance the parametric sensitivity as shown in Section 4.2.3 of the simulation study. P5 had a sharply increasing exponential feed profile to provoke substrate accumulation in the later process phase. The process durations varied between 30 h and 80 h, which is mainly due to the different applied feeding rates.

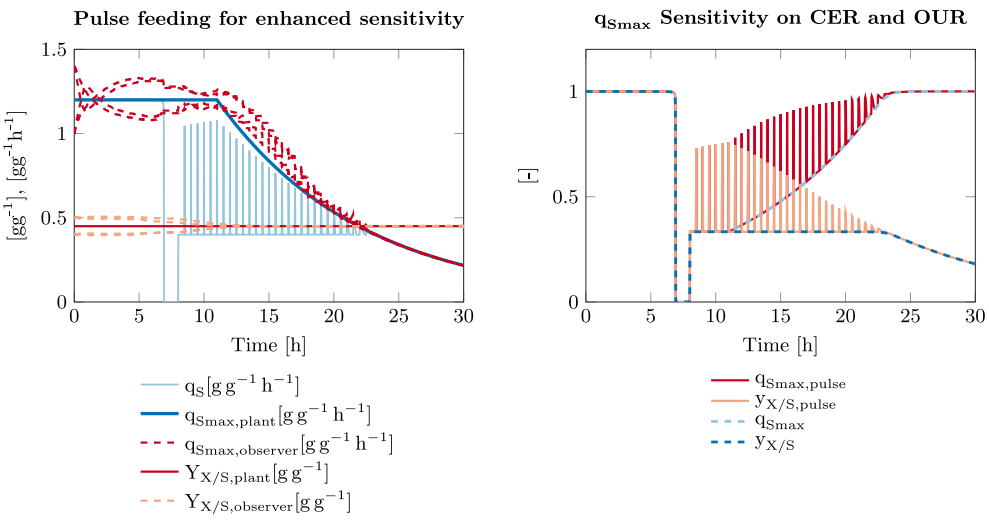

Fig. 5. Estimation of q_{Smax} and $Y_{X/S}$ based on uncertain measurements with $\mathcal{N}(\mu, (0.03\,\mu)^2)$ including superimposed feed pulses and feedback of parametric sensitivity for noise level tuning. The state observer was initialized with all four combinations of deflected start parameters ($\mathcal{N}(\mu, (0.15\,\mu)^2)$).

Table 3
Root mean square errors (RMSE) of the analyzed state estimation scenarios for $Y_{X/S}$ and q_{max} within the simulation study. Standard estimation A refers to perfect measurement assumption (Fig. 3A) and standard estimation B refers to uncertain measurement assumption (Fig. 3B).

	Std. estimation A (Section 4.2.1)	Std. estimation B (Section 4.2.1)	Sensitivity feedback (Section 4.2.2)	Superimposed pulsing (Section 4.2.3)	Final reduction
$Y_{X/S}$ RMSE [g g^{-1}]	5.24×10^{-4}	0.0614	0.0301	0.0246	2.49-fold
q_{Smax} RMSE [g g^{-1} h^{-1}]	0.0498	0.610	0.404	0.0726	8.40-fold

Table 4
Root mean square errors (RMSE) for biomass and substrate of the experimental validation experiments.

	State	P2	P3	P4	P5
Open-loop model RMSE [g]	X_v	15.86	8.03	11.44	13.53
	S	0.3764	2.276	1.338	30.91
State estimator RMSE [g]	X_v	6.202	7.922	7.666	9.632
	S	0.3480	1.230	0.7062	23.36
RMSE reduction by state estimator	X_v	**2.5x**	**1.01x**	**1.5x**	**1.4x**
	S	**1.1x**	**1.9x**	**1.9x**	**1.3x**

4.3.1. Biomass and substrate estimation results

The second row of Fig. 6 shows the estimation results of substrate and viable biomass. It can be seen that for each process the biomass and substrate estimates are in very good accordance to the offline determined reference measurements. Only a slight overestimation of biomass can be observed in P3 and P4. If we compare these estimation results with a feed-forward simulation of the calibrated model the effectiveness of the state estimation algorithm becomes evident. The root mean square errors (RMSE) given in Table 4 show that the state estimation errors are always lower than the open-loop simulation (with a reduction of up to 2.5 times), which corresponds to a reduction from 15.9 g L^{-1} (~15% error) down to 6.2 g L^{-1} (~6% error) for biomass in P2. Overall, the estimation error was below 10% for biomass and below 5% for glucose. This is in the range of other real-time measurement techniques that used elemental balancing based methods (Wechselberger et al., 2013), permittivity measurements (Reichelt et al., 2016), MIR spectroscopy (Siegl et al., 2022) or a hybridization of MIR spectroscopy and mass balancing methods (Siegl et al., 2022). This approach provides the advantage that only one calibration experiment was needed and the soft sensor was able to cover the batch, fed-batch and the production phase without any further adaptation.

4.3.2. Parameter estimation results

The improvement of the biomass estimation was reached by the real-time adaptation of the model parameters q_{Smax} and $Y_{X/S}$, which especially after induction can significantly change. This change over time is usually not reflected in simple unstructured models and therefore without any parameter adaptation their predictions are poor in the production phase.

Fig. 7 shows the real-time parameter estimates for $Y_{X/S}$ and q_{Smax} for process P2 with the standard exponential feeding profile. As expected from the simulation study q_{Smax} is only slightly adapted. The real q_S, which was calculated from the offline measurements through the mass balances, aligns very well with the predicted q_S from the state observer and is always clearly below the assumed q_{Smax}. This is also confirmed by the absence of substrate throughout the whole process. The $Y_{X/S}$ estimate however is significantly changed in Fig. 7 to match the model outputs with the CER and OUR measurements, shown in the last row of Fig. 6. The yield coefficients determined from offline measurements confirms this trend as the measured yield strongly decreased after induction. The estimated yield corresponds well to the offline measured yield with a RMSE of only 0.0366 g g^{-1}.

Fig. 8 and Fig. 11 in the appendix show the parameter estimation results for the two experiments with the superimposed feed pulses. Although the simulation study indicated a significant increase in the estimation accuracy of q_{Smax}, the trend is very similar to P2. Again, substrate concentration was limited all the time. The q_S, although frequently increased through the addition of the feed pulses, remained clearly below q_{Smax}. Under this condition, the parametric sensitivity remains low and no proper estimation can occur. It has also been noted, that the decreasing q_{Smax} of the plant model is purely assumption based and real experiments can behave differently. Similar to the results of P2 the yield coefficient decreased upon induction, although estimated and reference values have a clear offset. This offset is due to a wrong start parameter, which cannot be corrected by the state observer.

The last experiment P5 was designed to force overfeeding and subsequent substrate accumulation in order to test if the parameter estimation works under these extreme conditions. The parameter estimation results are displayed in Fig. 9. As soon as overfeeding took place, the q_{Smax} estimate was adjusted to the edge of q_S until 32 h, where the q_{Smax} started again to rise. During this time span

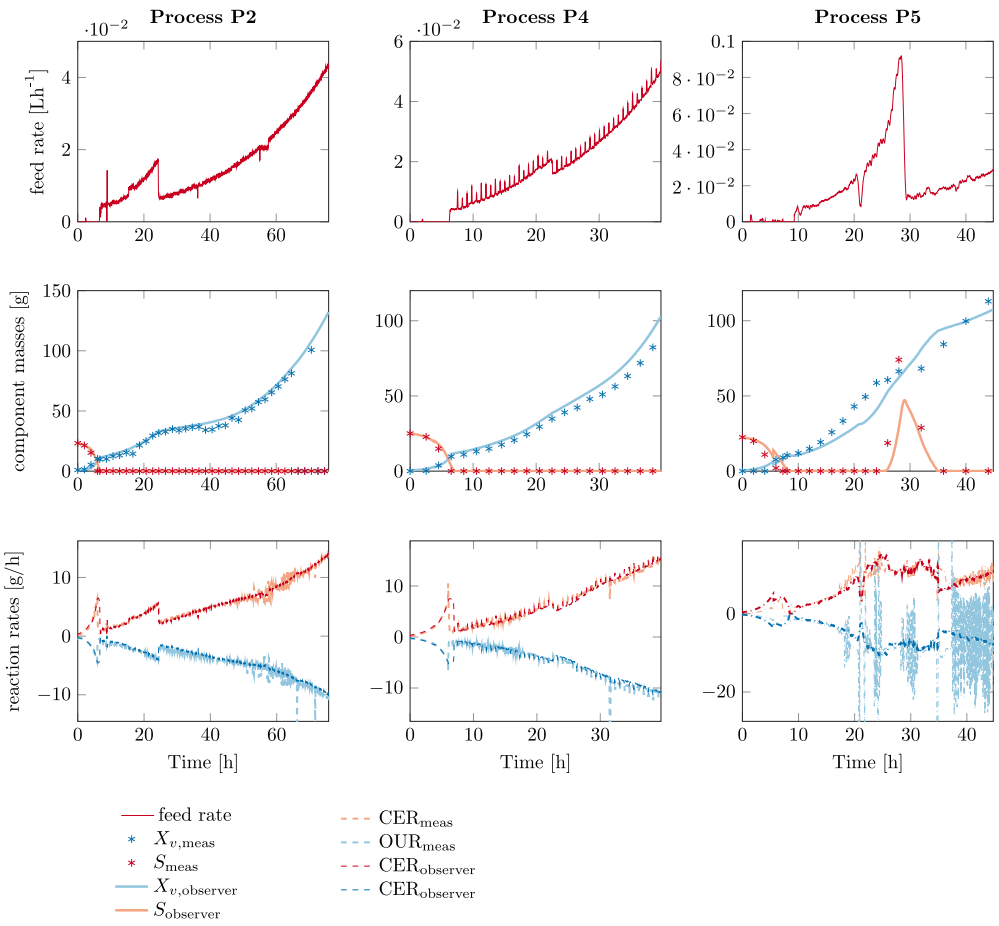

Fig. 6. Cultivation data and state estimation results of three *Escherichia coli* fed-batch experiments conducted in a 3.3 L lab-scale bioreactor.
row 1: Feeding profiles either standard exponential feed (P2, P5) or superimposed pulses (P2).
row 2: Stars represent offline substrate and biomass measurements and solid lines represent the respective soft sensor estimates.
row 3: Online CER and OUR measurements and the model outputs that are updated based on them. Negative reaction rates (OUR) refer to uptake of the component by the cell and positive reaction rates (CER) refer to an evolution of the component.

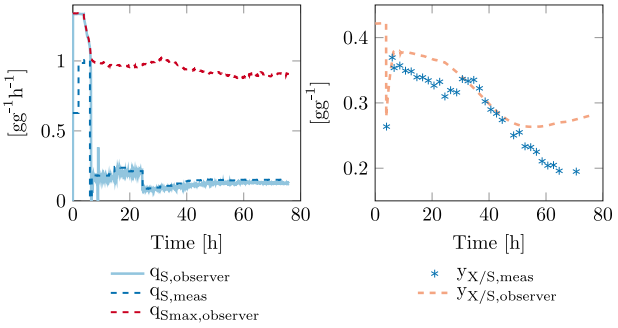

Fig. 7. Real-time q_{Smax} and $Y_{X/S}$ estimation of process P2 in comparison to offline determined yield and q_S.

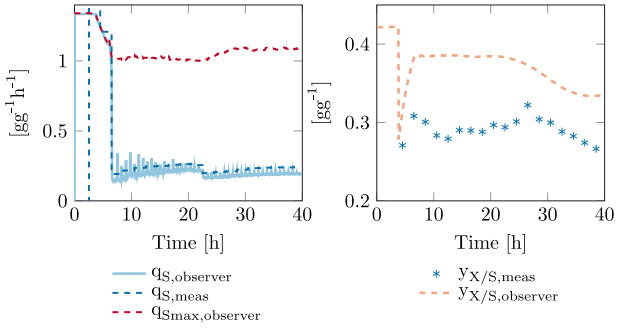

Fig. 8. Real-time q_{Smax} and $Y_{X/S}$ estimation of process P4 in comparison to offline determined yield and q_S.

of 5 h substrate accumulated in the culture medium. Because of the decreased q_{Smax} value excess substrate from the feed accumulated in the reactor. This can be also clearly seen by the offline measured substrate concentration of process P5 in Fig. 6. First, this experiment shows that the estimation of q_{Smax} works well when q_S approaches the maximum uptake capacity. Second, it confirms also the assumption that q_{Smax} decreases after the induction of the product formation. In this particular case q_{Smax} dropped from 1.35 g g^{-1} h^{-1} to 0.2 g g^{-1} h^{-1}. Maybe under less harsh feeding conditions the drop would have been not so pronounced. However, the result in this work is in line with some other studies (Neubauer et al., 2003; Reichelt et al., 2017) which observed similar drops in the substrate uptake capacities.

4.3.3. Parameter estimation discussion

When comparing this real-time parameter estimation approach of q_{Smax} and $Y_{X/S}$ to a state estimation approach of a more detailed model where the respiratory capacity is modeled by kinetics, there are some considerations to be made. On the one hand a parameter estimation approach of a simple model as used here reduces the need for prior knowledge about the system and therefore facilitates a generic

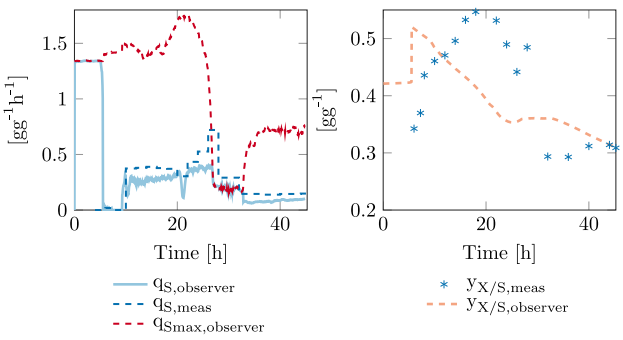

Fig. 9. Real-time q_{Smax} and $Y_{X/S}$ estimation of process P5 in comparison to offline determined yield and q_S.

applicability and transferability of the method. On the other hand the simple model can have poor convergence of the parameters due to low sensitivity and practical identifiability issues. In our specific case, the q_{Smax} parameter can be estimated properly under the condition of sufficient sensitivity as discussed in Section 4.2.3. However, when practical identifiability is getting poor the estimate is diverging as shown in Fig. 3. In those cases one could extend the model with suitable kinetics describing the behavior of the respiratory capacity which would increase robustness and accuracy. This, however necessitates a good knowledge about the system especially during the production phase where the metabolism is known to change over time. The parameters have to be well identified upfront which can be a time consuming task. Also, with increasing complexity the model can be harder to transfer to other host systems. Therefore, it depends on the case, which model complexity to choose.

The main challenge of a verification approach for q_{Smax} real-time estimation is, that there is no reference value available as long as q_{Smax} is not exceeded.

In this case study, we could show that particularly in an overfeeding situation in P5 the practical identifiability of q_{Smax} was high enough to properly estimate the reduced uptake capacity of the cells (Fig. 9) and to estimate the substrate accumulation consequently (Fig. 6). However, for all other processes (P2–P4) a dynamic trajectory of q_{Smax} was not detected, probably because of a too low sensitivity at uptake rates (q_S) wich are significantly below q_{Smax}. Also, the applied pulse feeding seemed to be too weak to get a significant increase of the parameter sensitivity. In following studies larger feed pulses could be applied in order to increase the sensitivity to a sufficient value to ensure real-time practical identifiability. This comes however with the cost that the influences of the feed deviations on the process are more pronounced and therefore potentially infeasible for production.

5. Conclusion

The goal of this contribution was to uniquely estimate the model parameters q_{Smax} and $Y_{X/S}$ in real-time from online CER and OUR measurements while keeping the culture under physiologically controlled conditions.

To achieve this we used a particle filter as a nonlinear state observer in combination with substrate perturbation techniques that have been developed in earlier studies (Akesson, 1998; de Maré, 2016; Whiffin et al., 2004). Our method however differs from the probing control methodology used in those earlier studies. The idea of probing control is, that the feed rate is getting increased as long as the culture is substrate limited and pO_2 peaks are detected after the feed pulses. Therefore, over time q_S is steered towards q_{Smax} and the feed rate is maximized without exceeding q_{Smax} and risking substrate accumulation. The problem is, that optimal productivity often takes place at a controlled q_S, which is below the maximum capacity q_{Smax}. The novelty of our contribution is the facilitation of simultaneous q_{Smax} and $Y_{X/S}$ estimation while maintaining controlled physiological rates below the maximum capacity. The method is based on a simple unstructured bioprocess model utilizing elemental balancing to derive CER and OUR from the model states. Thus, only one calibration experiment was needed to parameterize the soft sensor. Most strain specific parameters can be taken from literature or calculated in advance without much experimental effort. Although the assumption of those parameters as being constant and perfectly known in our simulation study is inaccurate to some extend, the observer still gave reasonable results in the experimental verification. The method serves as a monitoring strategy, which could also be applied as a PAT tool in industrial scale applications. It is completely automated since the feed pulses are realized by a transient increase of the pump setpoint rather than by manual additions of substrate.

The monitoring tool can also be seen as a basis for further improvements and extensions. Dead biomass could be included by either measuring the dead biomass directly (for example using a flow cytometer in an at-line fashion) or by measuring the total biomass (turbidity, spectroscopy) as well as the living biomass (e.g. conductivity).

In addition to that, the observer could be extended to describe acetate formation in function of the changing substrate uptake capacity and the conversion yield. An extension to include the product formation could be also of interest. Product formation however is very specific to certian strains, products and process conditions. In this study, we aimed on a more generic view on recombinant microorganisms, which we hope is generically applicable across different products and plasmid systems.

Furthermore, the state estimation algorithm provides the basis for more advanced bioprocess control applications that consider physiological rates and capacities. A feeding strategy controlling q_S to a specific fraction of q_{Smax} is imaginable, which was not possible with pure probing control. However, the biggest challenge for developing and improving such a method is, that there is no reliable reference value for q_{Smax} yet, which makes the estimator and controller difficult to validate.

CRediT authorship contribution statement

Don Fabian Müller: Conceptualization, Investigation, Formal analysis, Visualization, Writing – original draft. **Daniel Wibbing:** Software, Resources, Writing – review & editing. **Christoph Herwig:** Conceptualization, Supervision, Writing – review & editing, Funding acquisition. **Julian Kager:** Conceptualization, Methodology, Validation, Writing – review & editing, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors acknowledge financial support through the COMET Centre CHASE, funded within the COMET - Competence Centers for Excellent Technologies programme (No. 868615) by the BMK, the BMDW and the Federal Provinces of Upper Austria and Vienna. The COMET programme is managed by the Austrian Research Promotion Agency (FFG). Further financial support was provided by the Novo Nordisk Foundation (Start Package grant NNF220C0081250). The Authors thank DTU Library as member of the Royal Danish Library consortium for their open access funding program.

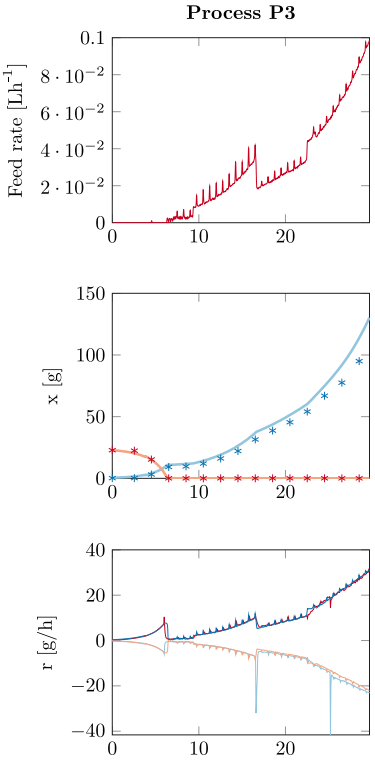

Fig. 10. Cultivation data of process P3 analogous to Fig. 6.

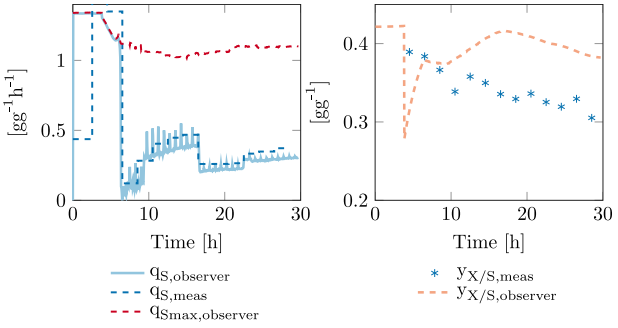

Fig. 11. Real-time q_{Smax} and $Y_{X/S}$ estimation of process P3 in comparison to offline determined yield and q_S.

Appendix. My appendix

Eq. (22) is the symbolic sensitivity matrix of CER and OUR related to the parameters q_{Smax} and $Y_{X/S}$.

$$S = \begin{bmatrix} -\frac{S \cdot X_v \cdot (\frac{Y_{X/S}}{M_S} - \frac{1}{M_S})}{k_S \cdot V_L + S} & -\frac{S \cdot X_v \cdot q_{Smax}}{M_X \cdot (k_S \cdot V_L + S)} \\ S \cdot X_v \cdot \frac{\frac{DOR_S}{M_S} - \frac{DOR_X \cdot Y_{X/S}}{M_X}}{DOR_{O2} \cdot (k_S \cdot V_L + S)} & -\frac{DOR_X \cdot S \cdot X_v \cdot q_{Smax}}{DOR_O \cdot M_X \cdot (k_S \cdot V_L + S)} \end{bmatrix} \tag{22}$$

See Figs. 10 and 11.

References

Aehle, M., Bork, K., Schaepe, S., Kuprijanov, A., Horstkorte, R., Simutis, R., Lübbert, A., 2012. Increasing batch-to-batch reproducibility of CHO-cell cultures using a model predictive control approach. Cytotechnology 64 (6), 623–634. http://dx.doi.org/10.1007/s10616-012-9438-1, URL: http://link.springer.com/10.1007/s10616-012-9438-1.

Akesson, M., 1998. A Probing Strategy for Substrate Feeding in Escherichia Coli Cultivations (Ph.D. thesis). Lund University of Technology, Lund.

Bárzaga-Martell, L., Duarte-Mermoud, M., Ibáñez Espinel, F., Gamboa-Labbé, B., Saa, P., Pérez-Correa, J., 2021. A robust hybrid observer for monitoring high-cell density cultures exhibiting overflow metabolism. J. Process Control 104, 112–125. http://dx.doi.org/10.1016/j.jprocont.2021.06.006, URL: https://www.sciencedirect.com/science/article/pii/S0959152421000974.

Daume, S., Kager, J., Herwig, C., 2019. Time resolved sensitivity & identifiability analysis for directed parametrization of highly dynamic models. Comput. Aided Chem. Eng. 46, 1111–1116. http://dx.doi.org/10.1016/B978-0-12-818634-3.50186-7, URL: https://linkinghub.elsevier.com/retrieve/pii/B9780128186343501867.

DeLisa, M.P., Li, J., Rao, G., Weigand, W.A., Bentley, W.E., 1999. Monitoring GFP-operon fusion protein expression during high cell density cultivation of Escherichia coli using an on-line optical sensor. Biotechnol. Bioeng. 65 (1), 54–64.

Dewasme, L., Goffaux, G., Hantson, A., Wouwer, A., 2013. Experimental validation of an Extended Kalman Filter estimating acetate concentration in E. coli cultures. J. Process Control 23, 148–157. http://dx.doi.org/10.1016/j.jprocont.2012.09.004, URL: https://www.sciencedirect.com/science/article/pii/S0959152412002235.

Dewasme, L., Wouwer, A., 2020. Model-free extremum seeking control of bioprocesses: A review with a worked example. Processes 8 (1209), http://dx.doi.org/10.3390/pr8101209, number: 10 Publisher: Multidisciplinary Digital Publishing Institute. URL: https://www.mdpi.com/2227-9717/8/10/1209.

Goffaux, G., Wouwer, A., 2005. Bioprocess state estimation: some classical and less classical approaches. In: Control and Observer Design for Nonlinear Finite and Infinite Dimensional Systems. pp. 111–128.

Kager, J., Bartlechner, J., Herwig, C., Jakubek, S., 2022. Direct control of recombinant protein production rates in E. coli fed-batch processes by nonlinear feedback linearization. Chem. Eng. Res. Des. 182, 290–304. http://dx.doi.org/10.1016/j.cherd.2022.03.043, URL: https://linkinghub.elsevier.com/retrieve/pii/S0263876222001460.

Kager, J., Herwig, C., Stelzer, I.V., 2018. State estimation for a penicillin fed-batch process combining particle filtering methods with online and time delayed offline measurements. Chem. Eng. Sci. 177, 234–244. http://dx.doi.org/10.1016/j.ces.2017.11.049, URL: https://linkinghub.elsevier.com/retrieve/pii/S0009250917307388.

Kager, J., Tuveri, A., Ulonska, S., Kroll, P., Herwig, C., 2020. Experimental verification and comparison of model predictive, PID and model inversion control in a Penicillium chrysogenum fed-batch process. Process Biochem. 90, 1–11. http://dx.doi.org/10.1016/j.procbio.2019.11.023, URL: http://www.sciencedirect.com/science/article/pii/S1359511319310335.

Kopp, J., Slouka, C., Strohmer, D., Kager, J., Spadiut, O., Herwig, C., 2018. Inclusion body bead size in E. coli controlled by physiological feeding. Microorganisms 6 (4), 116. http://dx.doi.org/10.3390/microorganisms6040116, Number: 4 Publisher: Multidisciplinary Digital Publishing Institute. URL: https://www.mdpi.com/2076-2607/6/4/116.

Lecca, P., Re, A., 2019. Identifying necessary and sufficient conditions for the observability of models of biochemical processes. Biophys. Chem. 254, 106257.

Li, T., Sun, S., Sattar, T., Corchado, J., 2014. Fight sample degeneracy and impoverishment in particle filters: A review of intelligent approaches. Expert Syst. Appl. 41, 3944–3954.

Lin, H.Y., Mathiszik, B., Xu, B., Enfors, S.-O., Neubauer, P., 2001. Determination of the maximum specific uptake capacities for glucose and oxygen in glucose-limited fed-batch cultivations ofEscherichia coli. Biotechnol. Bioeng. 73 (5), 347–357. http://dx.doi.org/10.1002/bit.1068, URL: https://onlinelibrary.wiley.com/doi/10.1002/bit.1068.

de Maré, L., 2016. Feeding Strategies Based on Probing Control (Ph.D. thesis).

Margaria, G., Riccomagno, E., White, L., 2004. Structural identifiability analysis of some highly structured families of statespace models using differential algebra. j. Math. Biol. 49, 433–454.

Mears, L., Stocks, S.M., Sin, G., Gernaey, K.V., 2017. A review of control strategies for manipulating the feed rate in fed-batch fermentation processes. J. Biotechnol. 245, 34–46. http://dx.doi.org/10.1016/j.jbiotec.2017.01.008, URL: https://linkinghub.elsevier.com/retrieve/pii/S0168165617300251.

Mohd Ali, J., Ha Hoang, N., Hussain, M., Dochain, D., 2015. Review and classification of recent observers applied in chemical process systems. Comput. Chem. Eng. 76, 27–41. http://dx.doi.org/10.1016/j.compchemeng.2015.01.019, URL: http://linkinghub.elsevier.com/retrieve/pii/S0098135415000216.

Monod, J., 1949. The growth of bacterial cultures. Annu. Rev. Microbiol. 3 (1), 371–394. http://dx.doi.org/10.1146/annurev.mi.03.100149.002103, URL: http://www.annualreviews.org/doi/10.1146/annurev.mi.03.100149.002103.

Müller, D.F., Lagoda, K., Wibbing, D., Herwig, C., Kager, J., 2022. Incorporation of error propagation into an elemental balancing based soft-sensor for improved online monitoring of microbial fed-batch processes. In: Montastruc, L., Negny, S. (Eds.), Computer Aided Chemical Engineering. In: 32 European Symposium on Computer Aided Process Engineering, vol. 51, Elsevier, pp. 1177–1182. http://dx.doi.org/10.1016/B978-0-323-95879-0.50197-1, URL: https://www.sciencedirect.com/science/article/pii/B9780323958790501971.

Neubauer, P., Lin, H.Y., Mathiszik, B., 2003. Metabolic load of recombinant protein production: Inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene inEscherichia coli. Biotechnol. Bioeng. 83 (1), 53–64. http://dx.doi.org/10.1002/bit.10645, URL: https://onlinelibrary.wiley.com/doi/10.1002/bit.10645.

Patwardhan, S., Narasimhan, S., Jagadeesan, P., Gopaluni, B., Shah, S., 2012. Nonlinear bayesian state estimation: A review of recent developments. Control Eng. Pract. 20, 933–953.

Pekarsky, A., Konopek, V., Spadiut, O., 2019. The impact of technical failures during cultivation of an inclusion body process. Bioprocess Biosyst. Eng. 42 (10), 1611–1624. http://dx.doi.org/10.1007/s00449-019-02158-x, URL: http://link.springer.com/10.1007/s00449-019-02158-x.

Phue, J.-N., Shiloach, J., 2005. Impact of dissolved oxygen concentration on acetate accumulation and physiology of E. coli BL21, evaluating transcription levels of key genes at different dissolved oxygen conditions. Metab. Eng. 7 (5), 353–363. http://dx.doi.org/10.1016/j.ymben.2005.06.003, URL: https://www.sciencedirect.com/science/article/pii/S1096717605000534.

Pimentel, G., Benavides, M., Dewasme, L., Coutinho, D., Wouwer, A., 2015. An observer-based robust control strategy for overflow metabolism cultures in fed-batch bioreactors. IFAC-PapersOnLine 48, 1081–1086. http://dx.doi.org/10.1016/j.ifacol.2015.09.112, URL: https://linkinghub.elsevier.com/retrieve/pii/S2405896315011933.

Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., Timmer, J., 2009. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics 25, 1923–1929.

Rawlings, J., Bakshi, B., 2006. Particle filtering and moving horizon estimation. Comput. Chem. Eng. 30, 1529–1541. http://dx.doi.org/10.1016/j.compchemeng.2006.05.031, URL: https://linkinghub.elsevier.com/retrieve/pii/S0098135406001566.

Reichelt, W.N., Brillmann, M., Thurrold, P., Keil, P., Fricke, J., Herwig, C., 2017. Physiological capacities decline during induced bioprocesses leading to substrate accumulation. Biotechnol. J. 12 (7), 1600547. http://dx.doi.org/10.1002/biot.201600547, _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/biot.201600547. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/biot.201600547.

Reichelt, W.N., Thurrold, P., Brillmann, M., Kager, J., Fricke, J., Herwig, C., 2016. Generic biomass estimation methods targeting physiologic process control in induced bacterial cultures. Eng. Life Sci. 16 (8), 720–730. http://dx.doi.org/10.1002/elsc.201500182, URL: http://doi.wiley.com/10.1002/elsc.201500182.

Roels, J.A., 1983. Energetics and Kinetics in Biotechnology. Elsevier Biomedical Press.

Sagmeister, P., Langemann, T., Wechselberger, P., Meitz, A., Herwig, C., 2013. A dynamic method for the investigation of induced state metabolic capacities as a function of temperature. Microb. Cell Factories 12, 1–11.

Santos, L., Dewasme, L., Coutinho, D., Wouwer, A.V., 2012. Nonlinear model predictive control of fed-batch cultures of micro-organisms exhibiting overflow metabolism: Assessment and robustness. Comput. Chem. Eng. 39, 143–151. http://dx.doi.org/10.1016/j.compchemeng.2011.12.010, URL: https://linkinghub.elsevier.com/retrieve/pii/S0098135411003474.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36 (8), 1627–1639. http://dx.doi.org/10.1021/ac60214a047, Publisher: American Chemical Society.

Siegl, M., Brunner, V., Geier, D., Becker, T., 2022. Ensemble-based adaptive soft sensor for fault-tolerant biomass monitoring. Eng. Life Sci. 22 (3–4), 229–241. http://dx.doi.org/10.1002/elsc.202100091, URL: https://onlinelibrary.wiley.com/doi/10.1002/elsc.202100091.

Simon, D., 2006. Optimal State Estimation: Kalman, H [Infinity] and Nonlinear Approaches. Wiley-Interscience, Hoboken, N.J, OCLC: ocm64084871.

Sinner, P., Stiegler, M., Goldbeck, O., Seibold, G.M., Herwig, C., Kager, J., 2022. Online estimation of changing metabolic capacities in continuous *Corynebacterium glutamicum* cultivations growing on a complex sugar mixture. Biotechnol. Bioeng. 119 (2), 575–590. http://dx.doi.org/10.1002/bit.28001, URL: https://onlinelibrary.wiley.com/doi/10.1002/bit.28001.

Sinner, P., Stiegler, M., Herwig, C., Kager, J., 2021. Noninvasive online monitoring of Corynebacterium glutamicum fed-batch bioprocesses subject to spent sulfite liquor raw material uncertainty. Bioresour. Technol. 321, 124395. http://dx.doi.org/10.1016/j.biortech.2020.124395, URL: https://linkinghub.elsevier.com/retrieve/pii/S0960852420316692.

Tuveri, A., Holck, H.E., Nakama, C.S., Matias, J., Jäschke, J., Imsland, L., Bar, N., 2022. Bioprocess monitoring: A moving horizon estimation experimental application. IFAC-PapersOnLine 55 (7), 222–227. http://dx.doi.org/10.1016/j.ifacol.2022.07.448, URL: https://linkinghub.elsevier.com/retrieve/pii/S2405896322008497.

Veloso, A., Rocha, I., Ferreira, E., 2009. Monitoring of fed-batch E. coli fermentations with software sensors. Bioprocess Biosyst. Eng. 32, 381–388. http://dx.doi.org/10.1007/s00449-008-0257-x.

Villaverde, A.F., Barreiro, A., Papachristodoulou, A., 2016. Structural identifiability of dynamic systems biology models. PLoS Comput. Biol. 22.

Wan, E., Van Der Merwe, R., 2000. The unscented Kalman filter for nonlinear estimation. In: Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium. Cat. No.00EX373, pp. 153–158. http://dx.doi.org/10.1109/ASSPCC.2000.882463.

Wechselberger, P., Sagmeister, P., Herwig, C., 2013. Real-time estimation of biomass and specific growth rate in physiologically variable recombinant fed-batch processes. Bioprocess Biosyst. Eng. 36 (9), 1205–1218. http://dx.doi.org/10.1007/s00449-012-0848-4, URL: http://link.springer.com/10.1007/s00449-012-0848-4.

Whiffin, V.S., Cooney, M.J., Cord-Ruwisch, R., 2004. Online detection of feed demand in high cell density cultures ofEscherichia coli by measurement of changes in dissolved oxygen transients in complex media. Biotechnol. Bioeng. 85 (4), 422–433. http://dx.doi.org/10.1002/bit.10802, URL: https://onlinelibrary.wiley.com/doi/10.1002/bit.10802.

Wurm, D.J., Marschall, L., Sagmeister, P., Herwig, C., Spadiut, O., 2017. Simple monitoring of cell leakiness and viability in Escherichia coli bioprocesses-A case study. Eng. Life Sci. 17 (6), 598–604. http://dx.doi.org/10.1002/elsc.201600204, URL: http://doi.wiley.com/10.1002/elsc.201600204.

Yousefi-Darani, A., Paquet-Durand, O., Hitzmann, B., 2020. The Kalman Filter for the Supervision of Cultivation Processes. Springer Berlin Heidelberg, Berlin, Heidelberg, imp. URL: http://link.springer.com/10.1007/10_2020_145.