

Real-time estimation of biomass and specific growth rate in physiologically variable recombinant fed-batch processes

Patrick Wechselberger · Patrick Sagmeister ·
Christoph Herwig

Received: 20 September 2012 / Accepted: 2 October 2012 / Published online: 23 November 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract The real-time measurement of biomass has been addressed since many years. The quantification of biomass in the induction phase of a recombinant bioprocess is not straight forward, since biological burden, caused by protein expression, can have a significant impact on the cell morphology and physiology. This variability potentially leads to poor generalization of the biomass estimation, hence is a very important issue in the dynamic field of process development with frequently changing processes and producer lines. We want to present a method to quantify “biomass” in real-time which avoids off-line sampling and the need for representative training data sets. This generally applicable soft-sensor, based on first principles, was used for the quantification of biomass in induced recombinant fed-batch processes. Results were compared with “state of the art” methods to estimate the biomass concentration and the specific growth rate μ . Gross errors such as wrong stoichiometric assumptions or sensor failure were detected automatically. This method allows for variable model coefficients such as yields in contrast to other process models, hence does not require prior experiments. It can be easily adapted to a different growth stoichiometry; hence the method provides good generalization, also for induced culture mode. This approach estimates the biomass (or anabolic bioconversion) in induced fed-batch cultures in real-time and provides this key variable for process development for control purposes.

Keywords Recombinant protein production · Process model · Soft sensor · Real-time biomass quantification · Process analytical technology (PAT)

List of symbols

t	Time (h)
S	Total amount of substrate in the cultivation broth (C-mol)
\dot{S}	Substrate feed rate (C-mol/h)
r	Conversion rate (C-mol/h)
q	Specific rate (g/g/h)
Y	Yield (C-mol/C-mol)
F	Flow/feed rate (g/h) for liquid and (nL/min) for gas
C	Concentration (C-mol/l)
X	Total amount of biomass in the cultivation broth (C-mol) or (g)
y	Mole fraction (–)
V_m	Molar volume of gas at norm condition (0 °C and 1 atm) (nl/mol)
Ra_{inert}	Inert gas ratio (–)
y_{wet}	O ₂ Conc. diluted by water content (without bioreaction) (–)
N	Total amount of ammonium in the cultivation broth (mol)
O_2	Total amount of oxygen in the cultivation broth (mol)
Z_i	Elemental composition of component i in biomass (–)
p_i	Elemental composition of component i in substrate (–)
V	Volume of the cultivation broth (l)
ex_{H_2O}	Water content in off-gas (–)
M	Molecular weight (g/c-mol)

P. Wechselberger · P. Sagmeister · C. Herwig (✉)
Research Area Biochemical Engineering, Institute of Chemical Engineering, Vienna University of Technology,
Gumpendorfer Straße 1a, 1060 Vienna, Austria
e-mail: christoph.herwig@tuwien.ac.at

P. Wechselberger
e-mail: pwechsel@mail.tuwien.ac.at

ρ_{feed}	Density of feed (g/L)
S_0	Feed concentration (g/L)
γ	Degree of reduction (–)
v	Coefficients e.g. v (–)
OD	Optical density 600 nm (–)
k	Specific growth rate if used for feed rate calculations (h^{-1})
μ	Measured specific growth rate (h^{-1})
CER	Carbon dioxide evolution rate (C-mol/h)
OUR	Oxygen uptake rate (C-mol/h)

Indices

in	Input
out	Output
conv.	Conversion
acc.	Accumulation
s	Substrate
f	Feed
x	Biomass
a	Air
CO_2	Carbon dioxide
O_2	Oxygen
N	Ammonium
b	Base
O	Oxygen
H	Hydrogen
n	Nitrogen
m	Measured
c	Estimated
i	Item number i
j	Item number j
t	Time point t
0	Initial

Introduction

Motivation

One of the key tasks in process development is maximization of space–time yield of the product while maintaining a previously defined product quality related attributes. Even though recombinant protein production is not strictly growth related as products of primary metabolism (e.g. ethanol), it is usually tied to the physiological state of the culture [1, 2]. Fed-batch process mode provides good metabolic control over the cell metabolism, since the availability of the limiting substrate is governed by the feeding profile. Hence, it is an important goal of process development to come up with a feeding strategy beneficial for both product yield and product quality. Here, a key variable is the biomass concentration which is required for further calculation of variables describing the metabolic

state of the culture such as specific rates (e.g. the specific growth rate) and yields. The reference method for biomass is gravimetric determination of biomass dry weight, which is typically defined as the insoluble fraction of the culture broth after removal of water (e.g. by drying at 105 °C). Conventionally, this is determined by off-line quantification, which is time-consuming and also comes with high operator dependent measurement error. Dry cell weight determination accounts for the insoluble fraction of mass in biomass, while it would physiologically more relevant to quantify the cells actually taking part in the bioreaction [3]. Especially in induced cultures there is variation in cell morphology [4], energy metabolism [5] and macroscopic composition of the cells; hence quantification of “biomass” or similar variables is not straight forward [6]. But most importantly there has to be a clear link to product quality or productivity, which is also process or product dependant. Furthermore, the variable should become available in real-time for control strategies and also to speed up process development using parallel processing [7] and automation of experiments [8, 9].

Differentiation from existing methods

Hard-type sensors

Various sensors for biomass using optical methods (turbidity, near infrared, NADH) or radio impedance (capacitance) are available for the in-line quantification of the biomass concentration [6, 10–12]. Since determination of biomass dry weight is not feasible directly (in situ) in the bioreactor for obvious practical reasons, this variable has to be quantified by relation to physical characteristics or morphological properties of the cells (e.g. electrical impedance or light scattering properties). However, changes in the physiological state of the culture are often accompanied by physiochemical or morphological changes in the cell population, potentially distorting established correlations. Different calibrations can be necessary for different process modes, media composition, and physiological states of the culture and also for different strains. Model maintenance or very good generalization (however, this often comes at the cost of estimation accuracy) is required to account for variability of the relation between the probe response and the actual dry cell weight, since the relation can vary, e.g., dependent on the media constituents and other process characteristics [13, 14], especially in process development with constantly changing process environments, culture conditions, strains, etc. it can be misleading to use these hard-type sensors as a sole quantification of biomass in process development. On the other hand this can be very useful to acquire interesting real-time information on the physiological state of the culture [15];

however, it is often not possible to determine the source of variability (e.g. morphological change or increase in biomass) from the real-time signal only. Thus, this should be seen as an auxiliary tool together with complementary methods (e.g. flow cytometry).

Soft-type sensors

Soft-type sensors make use of easily accessible items such as carbon dioxide and oxygen concentrations in the off-gas stream (using well established methods such as infrared for CO₂ or paramagnetic principle for O₂, respectively) to indirectly quantify less “easy” items such as the total biomass concentration [16]. There are two main competing approaches, empiric or data-driven methods and mechanistic models based on fundamental knowledge.

One school of thought came up with data-driven methods such as PCA, ANNs, PCR, PLSR [17–20]. No prior knowledge is required, but representative training data sets for the modeling problem to estimate model coefficients or to train weights (ANNs) are an imperative. Hence, these methods are useful if the process conditions do not vary too much (e.g. in production) and they come with constrained generic applicability as a soft sensor in process development [16].

A more direct approach to generate knowledge is using mechanistic models, which try to describe the system in question by fundamental principles (e.g. chemical or physical) on the interaction between process variables [21, 22]. The advantage of mechanistic models is also a drawback; detailed knowledge on the mechanistic of the process is not always available. In biological processes setting up mechanistic models is especially challenging due the great complexity of the living cell. Hence, one should avoid extensive use of prior knowledge or frequent re-fitting of parameters, as there would be no gain in the ease of use compared to hard-type sensors.

Soft-type sensors based on first principles

First principles are generally valid (e.g. the law of conservation, or equations of state for vapor–liquid equilibriums); hence, this is ideal for the dynamic field of process development, since these can be easily adapted to a new problem. For this reason first principles should be used whenever possible and calibration/training should be avoided. First principles require little prior knowledge, mostly items which are quantified anyways, e.g. growth stoichiometry and quantification of in-going and out-going mass streams (substrate, oxygen, carbon dioxide, etc.).

Here, a minimalistic mechanistic model relying on first principles rather than prior knowledge is suggested for extraction of information, consistency check and estimation of unknown items in physiologically variable cultures.

While the basic idea of this approach has been published decades ago [21, 23, 24], these ideas are interesting in the context of modern industrial recombinant bioprocesses, due to their ease of use and general applicability. The model approach is black box and unsegregated for sake of real-time capability. For general applicability the model is biochemically structured, based on macroscopic mass balances; hence, extensive use of model parameters which have to be experimentally determined was avoided. Ideally only natural constants, direct measurements and first principles should be used. Using enough constraints such as elemental balances, estimation of parameters such as yield coefficients from previous experiments can be avoided. Furthermore, redundancy should be applicable (the equation system is over determined), so that the consistency of the estimation can be verified. This is very useful to detect gross errors such as wrong stoichiometric assumptions or sensor failure.

Context overview

Induction of recombinant protein can have significant impact on the energy metabolism of cells, e.g. *E. coli* cultures this can result in partial cell lysis [25, 26] and unspecific release of proteins, carbohydrates and other building blocks to the supernatant. Alterations in the energy metabolism can also result in variations of yield coefficients. For this reason, models making use of fixed yields coefficients are wrong as soon as there is unaccounted variations in these coefficients. Some authors [27] came up with the idea of supplementing mechanistic models with data-driven methods such as neural networks to tackle that problem, but this does not eliminate the need for representative training data sets in the first place. Another approach to deal with model uncertainties such as poor knowledge of model coefficients is Kalman filters. Deviations of the process model can be mitigated by incorporation of off-line samples [28], however this does not eliminate off-line sampling and still requires prior knowledge on model coefficients. We want to avoid off-line sampling and the need for representative training data sets at all, using an elemental balancing approach, which relies on first principles only. While the basics for this approach were already published [16, 29], this contribution focuses on quantification of the biomass in the induction phase of fed-batch processes in red biotechnology, which is a key variable for process optimization as discussed above.

Figure 1 shows an overview on approaches to quantify biomass in real-time discussed in the previous chapters.

Goals

- We demonstrate an approach for the estimation of biomass concentration and specific growth rate for

	Hard-type Sensors	Soft-type Sensors data-driven	Soft-type Sensors mechanistic	Soft-type Sensors current contribution
Requirements	Probe available, sterilizable, cleanable, representative training data sets	PIMS capable of using soft-type sensors, representative training data sets	PIMS capable of using soft-type sensors, prior knowledge	PIMS capable of using soft-type sensors, little prior knowledge - mostly items which should be known anyway
Characteristics	In-situ probe, measurement of some physical biomass property which can be related to biomass	Non-invasive, statistical relation of process variables to biomass	Non-invasive, mechanistic relation of process variables to biomass	Non-invasive, mechanistic relation of process variables to biomass using only first principles
Field of application	Manufacturing, process development when generalization is not an issue			Manufacturing, process development
Pros	Simple installation, Linear regression very easy	Can be done with historical data, requires no prior knowledge	Can be done with historical data, can help with mechanistic understanding of process	Can be done with historical data, good generalization, can help with mechanistic understanding of process
Cons	Requires model maintenance, training data sets, possibly poor generalization	Requires model maintenance, training data sets, possibly poor generalization	Requires model maintenance, training data sets, possibly poor generalization	No model maintenance, growth stoichiometry must be correctly assumed (e.g. no metabolites, accumulation)

Fig. 1 Overview on approaches to quantify biomass in real-time

processes with variable cell metabolism and morphology and evaluate its applicability. For instance, this can be during induction phase of recombinant processes, where the quantification of biomass is a challenge. The method avoids off-line sampling and the need for representative training data sets.

- Different methods to calculate the specific growth rate in real-time, which is a key variable for process optimization and is typically calculated from off-line biomass concentrations, are compared, including a soft-sensor approach based on cumulative elemental balancing, a Luedeking–Piret-type (fixed yield) approach, based on off-gas rates and a hard-type capacitance probe.
- The approach should be useful for control and be available as a key variable for process development.
- Consistency check: gross errors such as wrong stoichiometric assumptions or sensor failure should be detected automatically.

Materials and methods

Culture

Pichia pastoris as eukaryotic microbial model system

The *Pichia pastoris* strain KM71H expresses the horseradish peroxidase isoenzyme C1A (HRP). The strain was

of MutS (methanol utilization slow) phenotype and HRP was secreted into the fermentation broth. Media were prepared according to [30]. After shaking flask preculture, a batch cultivation was initiated, followed by a fed-batch on glycerol to increase the biomass and induction phase on methanol employing a feeding strategy according to [28].

E. coli as prokaryotic microbial model system

A recombinant K12 *E. coli* strain with alkaline phosphatase on a rhamnose inducible promoter was used for the verification runs with stoichiometrically defined media [25]. A shaking flask preculture (100 ml for inoculation of 6 L batch medium, in 1 L shaking flask with baffles) was inoculated from frozen stocks and was used to inoculate the bioreactor. Culture conditions were pH = 7, temperature = 35 °C and DO₂ > 20 %. After a batch phase, which was detected by a drastic drop in the CO₂ off-gas signal and an increase in dissolved oxygen (DO₂), an exponential fed-batch with a specific growth rate of 0.15 (h⁻¹) was initiated. Equations (1) and (2) were used to calculate the feed profile for the exponential fed-batch. The specific growth rate before induction was set prior to the experiment, while constants such as the feed concentration (S_0), density (ρ_{feed}), the initial volume (V_0) and initial biomass concentration X_0 were measured. The biomass yield ($Y_{x/s}$) was determined in prior experiments. The molecular weight of substrate and biomass (M_S , M_X) can be found in

the literature or measured by a CHON analyzer in prior experiments. The exponential phase was followed by an induction phase with linear feeding, which was adjusted to an initial specific growth rate of different percentages of the exponential phase growth rate, 100, 70 and 40 % or 0.15, 0.1 or 0.06 (h⁻¹), respectively.

Feedrate in exponential fedbatch

$$F_{(t)} = F_0 * e^{k * t} \tag{1}$$

Initial feed rate in exponential fedbatch

$$F_0 = \frac{k * X_0 * M_s * \rho_{feed} * V_0}{S_0 * Y_{x/s} * M_X} \tag{2}$$

Biomass

Biomass concentrations were quantified by gravimetric measurement after drying for 72 h at 105 °C. Samples were centrifuged (5,000 rpm, 10 min) and the pellet was washed twice with distilled water to get rid of salts.

Substrate and small metabolites

Substrate and small metabolite concentrations were quantified using an HPLC method (Supelcogel C-610, Sigma Aldrich, flowrate: 0.5 ml/min, eluent: 0.1 % H₃PO₄/NaN₃, 30 °C, RI detector).

Protein determination: BCA

Extracellular protein concentrations were measured using the Bicinchoninic Acid Kit for Protein Determination (Sigma, BCA1-1KT). Bovine Serum Albumin (BSA) was used as a standard. The limit of quantification (blank + 9 standard deviations) was determined to be 0.151 (g/l) with a residual standard deviation of 0.008 (g/l).

Bioreactor setup and on-line analytics

Bioreactor

Two stainless steel bioreactors with working volumes of 10 and 20 L were used (Infors, Bottmingen, Switzerland). The systems come with a controller unit, which was used to adjust the process parameters: pH, temperature, aeration, reactor pressure and stirrer speed. Dissolved oxygen (DO₂) was controlled >20 % using a step controlled with reactor pressure, stirrer speed and air flow as manipulated variable. The pH was controlled using an integrated digital peristaltic pump and NH₄OH as a base. Air was filtered by a membrane-type filter and dispensed by a ring sparger. The culture vessel was sterilized at 121 °C for 20 min by in situ steam sterilization prior to inoculation.

Off-gas analysis

CO₂ and O₂ in the off-gas were quantified by a gas analyzer (Servomex, UK; M. Müller AG, Switzerland), using infrared and paramagnetic principle, respectively. Air flow was quantified by a mass flow controller (Vögtlin, Aesch, Switzerland).

Capacitance probe

An annular-type probe (Aber Instruments, Aberystwyth, Wales, UK) was used to measure capacitance during the fermentation. Capacitance values are calculated in real-time from the difference between two frequencies. At 1 MHz *E. coli* cells contribute to the capacitance while 10 MHz is the “background” depending on the medium, according to definitions of the supplier. The difference in capacitance relates to the viable cell concentration or more directly to intact biovolume, as only intact cells act as a capacitor [31].

Data management

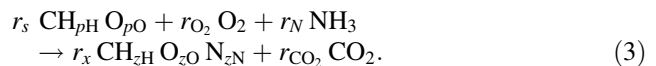
For recording of process data the process information management system Lucullus from Biospectra (Schlieren, Switzerland) was used. This system was also used for closed loop control (feed bottle on balance).

Setup of the soft sensor for estimation of biomass concentration and specific growth rate with variable growth stoichiometry

Conversion rates

Assuming oxidative metabolism, the bioreaction can be described according to Eq. 3. Although there are many different chemical reactions running in parallel in living cells, the conversion rates in Eq. 3 represent the overall summarized effect of all the different reactions.

Stoichiometric equations



General material balance

$$\text{Input} - \text{output} + \text{conversion} = \text{accumulation} \tag{4}$$

The conversion rates in Eq. (4) for the species substrate (S), biomass (X), carbon dioxide (CO₂), ammonia (N) as well as oxygen (O₂) can be derived from the general form of the material balance [Eq. (4)].

In fed-batch mode the conversion rates can be calculated as follows:

Conversion rate for substrate

$$r_s = \frac{d(s)}{dt} - \dot{S}_{in} + \dot{S}_{out} = -\frac{F_{f,in}}{\rho_{feed}} S_0. \quad (5)$$

In fed-batch mode, the outflow term \dot{S}_{out} is zero and the accumulation term $\frac{d(s)}{dt}$ can be neglected, as long $\mu < \mu_{max}$; hence, the conversion rate r_s is only dependent on the inflow term \dot{S}_{in} which is calculated from the feed rate.

Conversion rate for biomass

$$r_x = \frac{d(X)}{dt} - \dot{X}_{in} + \dot{X}_{out} = \frac{d(X)}{dt}. \quad (6)$$

Since there is no in- and outflow term, r_x is equal to the accumulation term $\frac{d(X)}{dt}$.

Conversion rate for carbon dioxide (=CER)

$$\begin{aligned} r_{CO_2} = CER &= \frac{d(CO_2)}{dt} - \dot{CO}_{2,in} + \dot{CO}_{2,out} \\ &= \frac{F_{a,in}}{V_m} (y_{CO_2,out} Ra_{inert} - y_{CO_2,in}) * 60. \end{aligned} \quad (7)$$

Conversion rate for oxygen (=OUR)

$$\begin{aligned} r_{O_2} = OUR &= \frac{d(O_2)}{dt} - \dot{O}_{2,in} + \dot{O}_{2,out} \\ &= \frac{F_{a,in}}{V_m} (y_{O_2,out} Ra_{inert} - y_{O_2,in}) * 60. \end{aligned} \quad (8)$$

Inert gas ratio

$$Ra_{inert} = \frac{1 - y_{O_2,in} - y_{CO_2,in}}{1 - y_{O_2,out} - y_{CO_2,out} - \frac{y_{wet}}{y_{O_2,in}}}. \quad (9)$$

Due to the low solubility of O_2 in the fermentation broth, $\frac{d(O_2)}{dt}$ can be neglected. The term $\frac{d(O_2)}{dt}$ can be also neglected, since the solubility of CO_2 in the fermentation broth is mainly a function of temperature and pH, which are typically kept constant. Hence, the rates r_{CO_2} and r_{O_2} are dependent on the in- and outflow terms (Eqs. 7 and 8). $F_{a,in}$, $y_{CO_2,out}$ and $y_{CO_2,in}$ are measured, while Ra_{inert} (Eq. 9) depends on the dilution by water stripping describes the ratio between the in- and outflow term. y_{wet} is the off-gas concentration of O_2 without bioreaction and relates to the dilution by water stripping [32].

Specific rates and yields

Conversion rates are the basis for the computation of yields (Eq. 10). Specific rates are calculated according to Eq. (11). Specific growth rates using off-line or soft-sensor biomass concentrations were calculated according to Eq. (12). The Luedeking–Piret-type Eq. (13) (as found in the literature [33], assuming constant maintenance) can be reformed to calculate specific growth rates from off-gas rates only, since the yield can be eliminated from the equation as long it is assumed to be constant (the

maintenance part does not vary). Specific growth rates based on total carbon or total oxygen were computed according to Eqs. (14) and (15). The total cumulated carbon dioxide $tCER$ is equal to the carbon dioxide rate CER integrated with Δt , plus an initial value $tCER_{initial}$. The specific growth rate is subsequently calculated by the fraction of CER by $tCER$. Similarly specific growth rates based on the capacitance signal were calculated according to Eq. (16) by the fraction of the capacitance rate by the total capacitance.

Calculation of yields

$$Y_j = \frac{r_i}{r_j}. \quad (10)$$

Calculation of specific rates

$$q_i = \frac{r_i}{X}. \quad (11)$$

Calculation of the specific growth rate from biomass conversion rates

$$\mu = \frac{r_x}{X}. \quad (12)$$

Luedeking–Piret-type equation, on the example of CO_2

$$CER = Y_{CO_2/X} * \mu * X. \quad (13)$$

Calculation of total cumulated off-gas rates, on the example of CO_2

$$tCER = CER * \Delta t + tCER_{initial}. \quad (14)$$

Calculation of the specific growth rate from off-gas rates on the example of CO_2

$$\mu_{tCER} = \frac{r_x * Y_{CO_2/X}}{X * Y_{CO_2/X}} = \frac{CER}{tCER}. \quad (15)$$

Calculation of the specific growth rate from the capacitance

$$\mu_{cap} = \frac{d(\text{Capacitance})}{dt} \frac{1}{\text{Capacitance}}. \quad (16)$$

Constraints

General form of constraints with k elemental balances

$$\sum_{i=1}^k r_i v_i = 0. \quad (17)$$

Using the law of conservation, elemental balances (=k) can be imposed on every element of the bioreaction as constraints (Eq. 17). In which r is the rate vector and v is the vector of coefficients for each element. This is useful as a consistency check of the data and to calculate non-measured items. In this contribution two balances were used, the carbon (C) balance and the degree of reduction (DoR) balance. The degree of reduction balance can be set

up with the following definition: $\gamma_N = -3$, $\gamma_C = C = 4$, $\gamma_{O_2} = O_2 = -4$, $\gamma_H = 1$. Other definitions would lead to the same results [34].

Sensitivities and error propagation

The random measurement error of off-gas analysis and gravimetric balances which are used as an input for this soft sensor are typically very low (relative error = 0.01–0.1 % propagated to the rates, using definitions according to supplier), hence can therefore be neglected. However, systematic errors such as miscalibration, sensor drift or the measurement error on the constants have to be considered. One percent deviation in feed concentration or density directly propagates as 1 % deviation of the substrate uptake rate. One percent deviation on the off-gas dilution by water ($y_{O_2_wet}$) propagates up to 3 % on the oxygen uptake rate, depending on the measured oxygen concentration. Furthermore, the off-gas rates can be prone to miscalibration and sensor. As the soft sensor is directly based on these rates, the error on the estimated accumulated biomass (the difference between the current value and the start value) is supposed to be in the magnitude of a linear combination of these effects. Gross errors such as wrong stoichiometric assumptions (e.g.: oxidoreductive instead of oxidative growth) also contribute to these effects.

Consistency check

To evaluate the residuals on the rates when applying the constraints in “Constraints” and compare them with the expected residuals due to measurement error (“Sensitivities and error propagation”), a statistical test adapted from the literature [35] was applied to get a quantitative measure on the validity of the observed system. Equation (17) can be written in matrix form (Eq. 18):

Matrix form of constraints

$$EW = 0. \tag{18}$$

W is the vector of the measured volumetric rates r .
 For noisy data a residue vector ε is added (Eq. 19):
 Matrix form of constraints with residue vector

$$E'W = \varepsilon. \tag{19}$$

For each rate an expected error (by default 3 % error on each rate) is specified in the variance–covariance matrix Ψ of the rates and is assumed to be non-correlated (square with the errors for each rate in the diagonal). The result of the statistical test value h is calculated with Φ as the variance–covariance matrix of the residuals Eqs. (20) and (21). The hypothesis of not having any errors exceeding the expected error specified in Ψ is rejected if h is greater than a certain threshold value. This threshold value can be read

from Chi-square distribution, which depends on the degree of redundancy of the equation system (or also the degree of freedom of the Chi-square distribution) and the significance level α (by default 0.95). The default α degree of redundancy of one (=estimation of one rate) results in a threshold of 3.84 for the statistical test value, which is exceeded if the current error is higher than the expected error (e.g. when gross errors such as wrong stoichiometric assumptions are present). The expected error was assumed to be 3 % error on each rate. An error of 3 % on each rate results in a deviation of about 10 % on the C- and DoR balance. The degree of redundancy of the equation system is equal to the rank of E if no conversion rates are estimated or to the rank of R if conversion rates are estimated.

Variance–covariance matrix

$$\Phi = E^T \Psi E. \tag{20}$$

Statistical test value

$$h = \varepsilon^T \Phi^{-1} \varepsilon. \tag{21}$$

Data reconciliation

A data reconciliation procedure according to [36] was applied. In addition to estimation of non-measured conversion rates, redundancy in the equation system can be also used to adjust the conversion rates to simultaneously close all elemental balances imposed in “Constraints”. The lumped residues of the equation system are distributed along the rates according the expected error for each rate. Using a least squares approach, the goal of reconciliation is to find a measurement error vector δ to calculate the reconciled vector W_b (Eq. 22), hence the vector of the best estimates of the volumetric reaction rates to fit all constraints. The solution (Eq. 23) to this problem is adapted from the literature [37].

Calculation of the reconciled vector Eq. (22)

$$W_b = W + \delta. \tag{22}$$

Calculation of the measurement error vector Eq. (23)

$$\delta = \Psi E^T \Phi^{-1} \varepsilon. \tag{23}$$

Estimation

For estimation of non-measured components the equation system from “Consistency check” is split into measured part ($E_m W_m$) and a calculated part [$E_c W_c$; Eq. (24)].

Split of equations system

$$EW = E_m W_m + E_c W_c = 0. \tag{24}$$

The law of conservation can be also applied on measured rates exclusively, but first the elemental matrix has to be stripped of relations with the calculated items, the

resulting matrix is called the redundancy matrix R (Eqs. 25 and 26). If R contains zero columns, there is no way to express the measured rates independent from the non-measured rates. The rank of R is equal to degree of redundancy of the equation system.

Introduction of the redundancy matrix R

$$RW_m = 0. \quad (25)$$

Calculation of the redundancy matrix R

$$R = E_m - E_c E_c^{-1} E_m. \quad (26)$$

For further calculation of the h value and the reconciliation procedure, R has to be stripped of rows with zero singular values, otherwise the inverse of Φ might not exist or be too close to singular, leading into unstable results. The singular values can be read from the matrix Σ acquired by singular value decomposition of R . The number of non-zero singular values in Σ determines the number of rows in conversion matrix Σ_{conv} ; which has 1 in the diagonal else 0 (Eqs. 27 and 28).

Singular value decomposition of R

$$R = U\Sigma V. \quad (27)$$

Calculation of the reduced redundancy matrix R

$$R_{\text{red}} = \Sigma_{\text{conv}} \Sigma \Sigma^T U^T. \quad (28)$$

Now the h value can be calculated and the rates reconciled as explained above by replacing E with R_{red} in Eq. (20). Finally the vector for estimated rates can be calculated according to Eq. (29), using the reconciled rates instead of raw measured rates. Furthermore, the estimated rates can be numerically integrated to estimate the cumulated biomass in the time window Δt according to Eq. (30), which is called cumulative elemental balancing further on.

Calculation of the estimated rate vector

$$W_c = E_c^{-1} E_m W_b. \quad (29)$$

Estimation of the biomass by numeric integration of the items in the estimated rate vector W_c (stoichiometric cumulation) Eq. (30)

$$X = r_{X_{\text{recon}}} * \Delta t + X_{\text{initial}}. \quad (30)$$

Results

Correct assumption of growth stoichiometry is an important prerequisite for the elemental balancing approach proposed in this contribution. By proper application of elemental balances products such as ethanol or acetate can be quantified [16]. However, this has to be addressed beforehand, by setting up the soft sensor accordingly. If the residuals on the elemental balances are higher than the defined error

according to Ψ , faulty definition of stoichiometric growth is detected. This is evaluated by statistical test, which takes measurement error according to the Chi-square distribution into account. Hence, Ψ has to be defined based on a realistic assumption of the measurement error, as failure to do so will result in potentially misleading results and/or statistical test values (“Consistency check”).

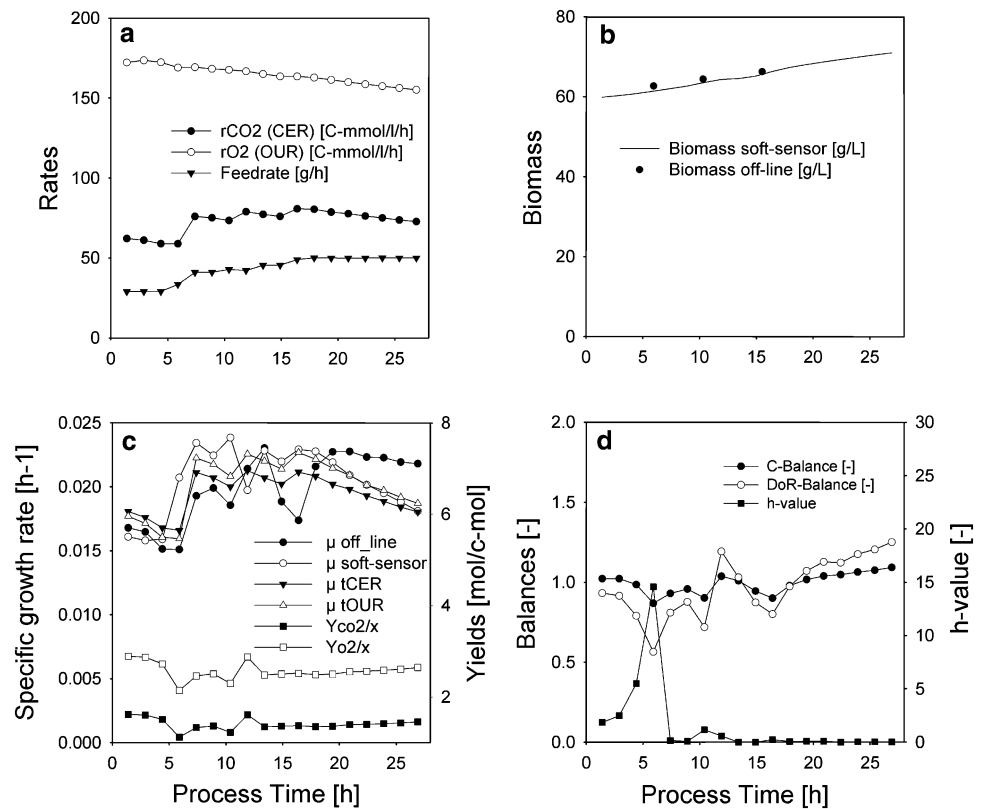
There are processes which generally follow balanced growth conditions, which means metabolic or morphological variations due stress and fermentation conditions are negligibly small and do not propagate to coefficients in data-driven or mechanistic models. In that case the estimation problem is much easier, since coefficients are constant throughout the experiment and also for follow-up experiments (see “Induced culture without significant extracellular product”). However, this is not true for any kind of process (see “Induced culture with significant extracellular product and variable yields”); hence, poor model generalization can become a major issue. Here, the elemental balancing approach poses a valuable alternative.

Induced culture without significant extracellular product

Data from an induced *P. pastoris* culture were used as an example for a culture without significant product (volumetric rate of substrate uptake is of magnitudes higher than the volumetric rate for extracellular product). Furthermore, this culture exhibits balanced growth conditions with regard to energy metabolism, since the yields are rather constant as also shown below (<10 % variation). This can probably be attributed to the careful adaption of the culture to the methanol feed and moderate expression rates due to the single-copy strain used in this process [30], which results in a constant maintenance part in the yields. A soft sensor for fed-batch with oxidative metabolism according to “Estimation” was implemented to estimate the biomass concentration in real-time from signals which are available online only. The input rates for this soft sensor were OUR, CER (measured by Off-gas analysis), the substrate feed rate and the reactor broth mass (by gravimetric principle, respectively) as shown in Fig. 2a. The output of the soft sensor was the biomass concentration (Fig. 2b) calculated using the cumulated biomass (Eq. 30) and current volume together with subsequently calculated variables such as the specific growth rate according to “Specific rates and yields”. Figure 2c shows the specific growth rate for calculated using different sources for the biomass quantification:

- *Off-line* Smoothed and interpolated off-line values according to Eq. (12) were used to calculate μ off-line (Fig. 2c). To check the quality of the off-line values, elemental balances such as the C and DoR

Fig. 2 Induced *P. Pastoris* fedbatch on methanol without significant extracellular product **a** inputs; **b** measured biomass concentration and estimated biomass concentration by soft sensor based on cumulative elemental balancing; **c** specific growth rate calculated from interpolated offline biomass samples (μ off-line), off-gas rates (μ tCER and μ tOUR, together with the respective yields: $Y_{\text{CO}_2/\text{x}}$ and $Y_{\text{O}_2/\text{x}}$) and based on cumulative elemental balancing (μ soft-sensor); **d** elemental balances for off-line biomass concentrations and h value for soft sensor; between process time 5 and 10 h the h value and the balances were off, probably due to consumption of substrate accumulated during the onset of the methanol feeding



balance can be imposed. If the law of conservation is satisfied the balance are close to 1, which is true here, if they deviate gross errors such as incorrectly assumed growth stoichiometry (e.g. oxidoreductive instead oxidative) or substrate accumulation are present (Fig. 2d).

- *tCER* and *tOUR* Estimation by off-gas rates only according to Eq. (15) was used to calculate μ tCER and μ tOUR. This method is only useful if the correlation of off-gas rates to biomass conversion rate is constant, or in other words the yields do not vary too much (only <10 %), which can be assumed for this culture (Fig. 2c). This can be probably attributed to the careful adaption of the culture to methanol and moderate foreign protein expression rates [28], which results in a constant maintenance part in the yields. This approach is comparable to similar methods without variable yields.
- *Soft sensor* Estimation of μ using cumulative elemental balancing as described in “Estimation” (Fig. 2c). This method allows flexible yields but the growth stoichiometry (e.g. oxidative metabolism) has to be correctly assumed beforehand. The h value is based on elemental balances (“Consistency check”) and is used as a real-time acceptance criterion. If the value is lower than the threshold value of 3.84, no gross errors are present and the growth stoichiometry was assumed correctly, which holds true for this culture (Fig. 2d).

Since the off-gas yields were rather constant (Fig. 2c) and the growth was purely oxidative, all methods give an estimate of the specific growth rate in good agreement with the values calculated from off-line biomass concentrations. Between process time 5 and 10 h the h value and the balances were off, probably due to consumption of substrate, which was accumulated during the onset of the methanol feeding.

Induced culture with significant extracellular product and variable yields

The applicability of the soft-type sensor based on cumulative elemental balancing was evaluated by estimation of biomass and the specific growth rate in a culture with significant extracellular product and variable yields. Results were compared with a Luedeking–Piret-type approach, a capacitance probe (“Capacitance probe”) and conventional off-line sampling. Two different fed-batch experiments were evaluated. In one experiment the initial specific growth rate was adjusted from $\mu = 0.15$ (h^{-1}) to a $\mu_{\text{initial}} = 0.10$ (h^{-1}) (Fig. 3), while in the other experiment the initial specific growth rate was adjusted from $\mu = 0.15$ (h^{-1}) to a $\mu_{\text{initial}} = 0.06$ (h^{-1}) (Fig. 4).

Signals which are available on-line were used as input signals for the soft sensor (Figs. 3, 4a). The estimated biomass concentration deviates from the off-line biomass

Fig. 3 Induction phase of *E. coli* culture with significant product with $\mu_{\text{initial}} = 0.10$ (h^{-1}); **a** inputs for the soft sensor based on elemental balancing; **b** measured biomass concentration and estimated biomass concentration by soft sensor based on elemental balancing; **c** specific growth rate calculated from interpolated offline biomass samples (μ off-line), off-gas rates (μ tCER and μ tOUR, together with the respective yields: $Y_{\text{CO}_2/\text{x}}$ and $Y_{\text{O}_2/\text{x}}$), capacitance probe (μ Cap) and based on cumulative elemental balancing (μ soft-sensor); **d** elemental balances for off-line biomass concentrations and h value for soft sensor

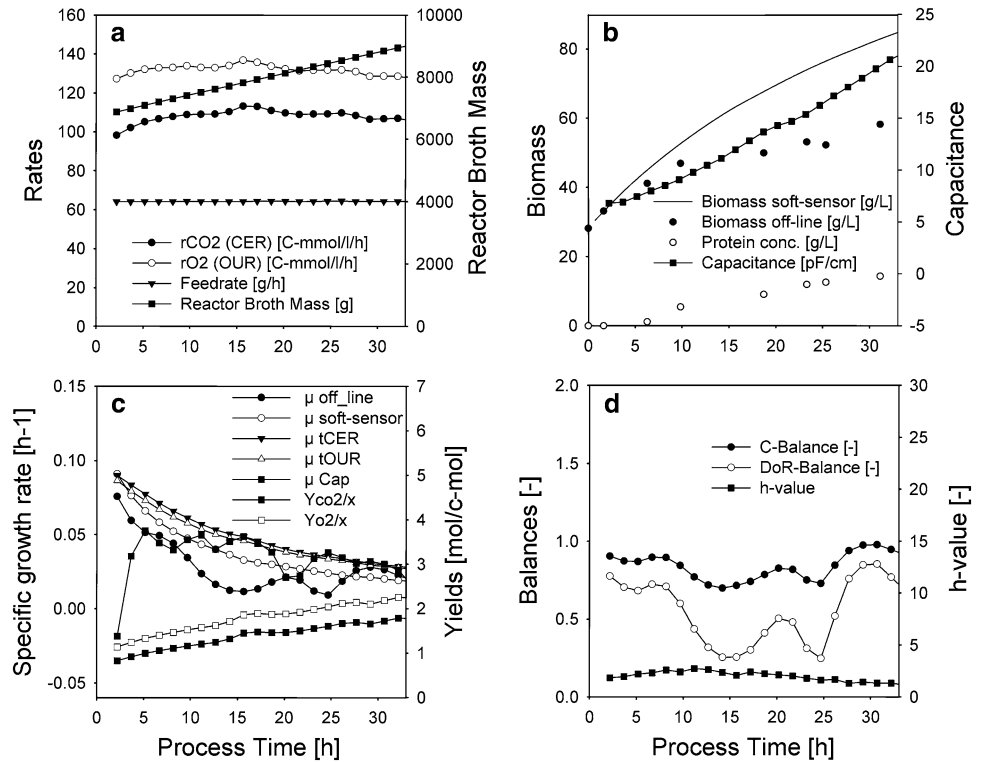
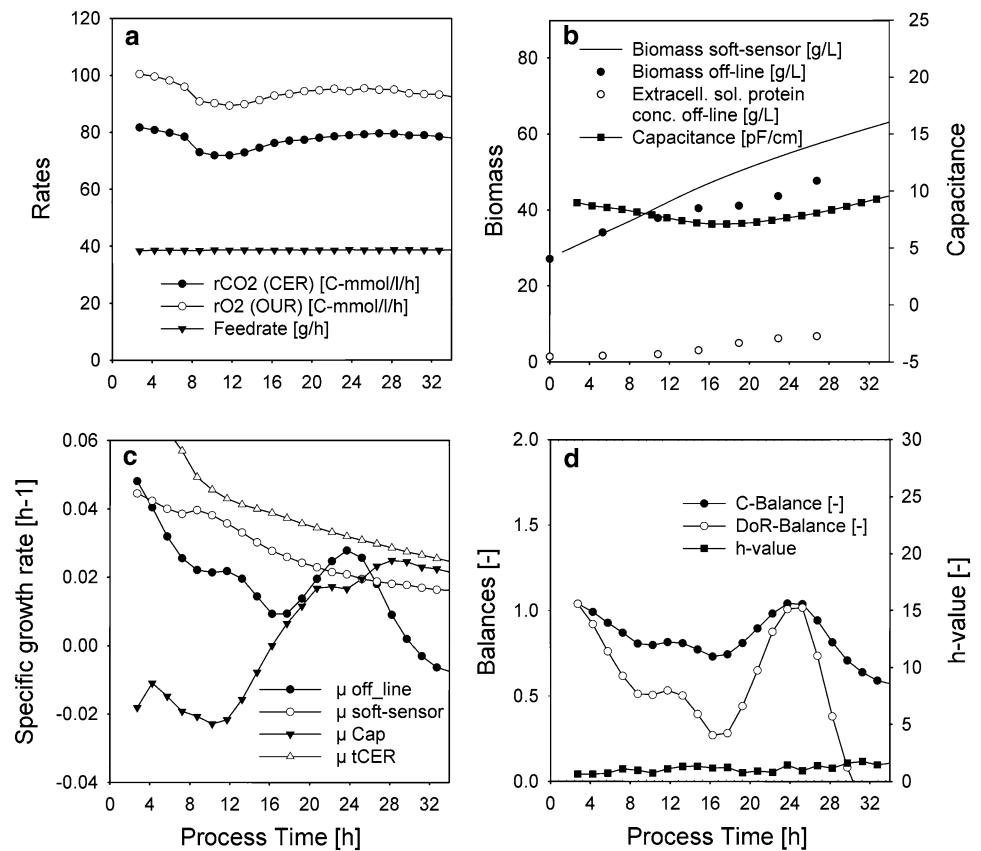


Fig. 4 Induction phase of *E. coli* culture with significant product with $\mu_{\text{initial}} = 0.06$ (h^{-1}); **a** inputs for the soft sensor based on elemental balancing; **b** measured biomass concentration and estimated biomass concentration by soft sensor based on elemental balancing; **c** specific growth rate calculated from interpolated offline biomass samples (μ off-line), off-gas rates (μ tCER), capacitance probe (μ Cap) and based on cumulative elemental balancing (μ soft-sensor); **d** elemental balances for off-line biomass concentrations and h value for soft sensor



concentration progressively with process time (Figs. 3, 4b). If the extracellular soluble protein concentration is added to the biomass, which is justifiable since the stoichiometry is very similar, this deviation is smaller and the remainder can probably be attributed to extracellular soluble non-protein content. This is due to the fact that the elemental balancing approach estimates overall bioconversion, biomass and any other soluble building block material. By this approach, it is not possible to distinguish between extracellular products such as soluble protein and cellular biomass since both have very similar stoichiometry (compared to classical oxidoreductive products such as ethanol).

Capacitance is described as a powerful tool for recombinant *E. coli* processes in literature [38]. The probe quantifies capacitance in (pF/cm), which typically has to be converted to more common units such as biomass (g/L) or optical density (–) for further use (by means of linear regression). An important thing to note is that capacitance relates to intact biovolume [31] and not necessarily to biomass. Schwan's model [31], which describes the relation between enclosed biovolume or capacitor volume and the dielectric increment, also suggests C_m [plasma membrane capacitance per unit of membrane area (F/m²)] as a parameter, which relates to the ability of the plasma membrane to store charge.

Schwan's model Eq. (31)

$$\Delta C * \frac{k}{\varepsilon_0} = \Delta \varepsilon' = \frac{9 * P * r * C_m}{4 * \varepsilon_0}. \quad (31)$$

Figures 3 and 4c show the specific growth rate calculated using different sources of the biomass concentrations:

- *Off-line* Smoothed and interpolated off-line values according to Eq. (12) were used to calculate μ off-line. To check the consistency of the off-line values, elemental balances, namely a C- and a DoR balance were imposed. Figures 3 and 4d show that both balances are <1, which can be attributed to the extracellular soluble protein unaccounted for in this balance.
- *tCER and tOUR* Estimation by off-gas rates only according to Eq. (15) was used to calculate μ tCER and tOUR. Figure 3 and 4c show that the yields are increasing for this culture (up to 70 % increase over process time), hence the estimated specific growth rate is artificially larger than the other growth rates. If this method is used to estimate the specific growth without verification by other methods and lacking a strategy to account for the variable maintenance part of the yield in real-time, results are misleading due to the variable yield.
- *Soft sensor* Estimation using elemental balancing and reconciliation as described in “Data reconciliation” and

“Estimation”. This method allows flexible yields but the growth stoichiometry (e.g. oxidative metabolism) has to be correctly assumed beforehand. The h value is based on elemental balances (“Consistency check”) and is used as a real-time acceptance criterion. If the value is lower than the threshold value of 3.84 no gross errors are present and the growth stoichiometry was assumed correctly. The value is below the threshold value (Figs. 3, 4d). The estimated specific growth rate is larger than the off-line specific growth rate due to the fact that this growth rate is based on the overall biomass conversion rate, including extracellular soluble components.

- *μ capacitance* The capacitance signal relates to biovolume and not to biomass, hence the linear regression model possibly also interferes with the parameter membrane capacitance as discussed above. There seems to be a dramatic decrease in biovolume due the reduction of specific growth rate from $\mu = 0.15$ (h⁻¹) to $\mu_{\text{initial}} = 0.1$ (h⁻¹) or 0.06 (h⁻¹) at the onset of induction (Figs. 3, 4c). Such a decrease in volume due to carbon depletion was also reported in literature [39]. The parameter C_m in Schwan's model (Eq. 31), which is in fact another morphological parameter, can also interfere with quantification. For this reason the apparent specific growth rate calculated from the capacitance signal is low or even negative during the initial hours of induction. This is even more pronounced in Fig. 4c, obviously since the drop in feed rate is larger here.

Discussion

Different methods to estimate the biomass concentration and also the specific growth rate μ in real-time were compared. These variables are typically determined by time-consuming off-line sampling, which is also prone to operator-dependent measurement error. The biomass concentration is estimated from signals which are available on-line only. A real-time biomass concentration is required for real-time calculation of specific rates and yields, which in turn provide valuable information on the bioprocess.

Specific growth rates calculated using total carbon or oxygen rates can only be used as long as the respective yields are constant. This was true for the induced *P. pastoris* culture (Sect. 3.1), but not for the *E. coli* culture (Sect. 3.2). Since the yields varied up to 70 %, estimated growth rates are artificially too large by this factor. Without verification by other means to estimate the biomass or a strategy to account for the variable maintenance part in the yields, the results are potentially misleading.

The soft sensor cannot differentiate between biomass and very similar soluble products such as protein in the

supernatant, hence estimates an overall bioconversion rate (anabolic conversion) of substrate, instead of conversion to cells or protein. If the off-line extracellular protein is added to the off-line biomass, the result is similar to the estimation of the soft sensor. Potentially extracellular protein can be quantified in real-time if required, e.g. by at-line spectroscopic or photometric methods.

Provided balanced growth conditions, the cell volume and therefore the capacitance signal linearly correlates to viable biomass dry weight. In induced systems the assumption of balanced growth does not necessarily hold true, as shown in this contribution. If this is not considered, it can be misleading to use this as a sole method to quantify biomass. Along with possible variability in the ratio biovolume per biomass, variations of the parameter C_m would also interfere with linear regression models for the estimation of biomass. While this can be interesting additional information on the bioprocess, further off-line measurements of morphological characteristics such as *E. coli* size distribution (e.g. by coulter principle) or radiofrequency scanning, as shown by other authors [40], are required to fully identify the source of variability. However, this was not within the focus of this contribution.

The quantification of biomass in the induction phase is not straight forward due to possibly variable cell metabolism, morphology and variables describing the former such as yields. There are multiple possible definitions for biomass. Conventionally biomass is defined as the non-soluble fraction of the culture broth after removal of water. It is not possible to apply this method in situ in real-time for obvious practical reasons. Furthermore, this definition might not be the best one from a physiological point of view, since there is more to biomass than just non-soluble mass. Elemental balancing for example allows quantifying an overall anabolic bioconversion, including soluble cell components secreted to the supernatant, such as carbohydrates and proteins. Capacitance provides information on the intact biovolume. If frequency scanning or additional off-line measurements are used, also other interesting morphological information can be revealed.

The biomass concentration, typically referred to as biomass dry cell weight concentration, is considered to be a key variable for the design of control strategies. This contribution outlined that there are multiple variables which describe (e.g.: biovolume, total bioconversion). Hence, the question arises on what measurement basis a control strategy, e.g. for the induction phase feed profile, should be used. Accordingly other authors suggest that the induced cell population is not homogeneous and segregated models need to be applied to account for different subpopulations [41]. In a control context, the question arises how to make different subpopulations quantitatively accessible by means of on-line and real-time sensors. This

is also a question of what is the link of the biomass definition to product quality or productivity. One sensor might not suffice for this task, but combinations of different analytical devices in combination with mechanistic models/soft sensors, might be able to unlock the current status of the cell population.

Conclusion

- A real-time capable soft sensor based on elemental balancing to quantify biomass concentration and specific growth rates was presented. Quantification of biomass in the induction phase is particularly challenging, due to the morphological, physiological and metabolic variations during induction. The presented method works with variable yields, which makes this approach especially interesting in the induced phase of recombinant bioprocesses, compared to other approaches with fixed yield or model coefficients as detailed in “Discussion”. As demonstrated, other approaches with fixed yield or model coefficients, such as the Lu-edecking–Piret-type approach and the linear regression model for the capacitance probe, have issues in the induction phase. Furthermore, it was shown that the reference method, gravimetric biomass dry weight determination, cannot quantitatively cover anabolic bioconversion, due to secretion of soluble contents to the supernatant.
- While the basic idea of elemental balancing, as a method to estimate unknown items such as the biomass concentration in a bioprocess, has been published decades earlier, the methodology is not used in modern bioprocess technology including recombinant bioprocesses so far. Furthermore, the approach is underrepresented in recent reviews [10, 42]. The ease of use and the low requirements on prior knowledge, as also requested by these reviews, make this method a particularly valuable tool for industrial bioprocesses.
- The results were compared and discussed with other state of the art approaches for real-time quantification of biomass. The optimum approach depends on the favored definition for “biomass” (conventional biomass, anabolic bioconversion, biovolume, etc.). Dependent on the process and/or product a definition with a clear link to product quality or productivity should be used.
- Since the results of this approach are obtained in real-time, this can be used to estimate process variables for control strategies. Specific rates such as the specific growth rate are especially important process variables, acting as a descriptor of the physiological state, with impact on biomass production, product quantity, and

product quality [43]. Furthermore, if a system is fully quantified, a great deal of understanding is achieved that can be easily communicated.

- This approach can be applied for platform approaches, new trends toward platform manufacturing and decreases time to market.

Acknowledgments The authors are very grateful to their colleague Christian Dietzsch for providing the data on the *P. pastoris* strain used in this study. The study was financially supported by the Austrian Science Fund (FWF Project P24154-N17).

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

- Looser V, Hammes F, Keller M, Berney M, Kovar K, Egli T (2005) Flow-cytometric detection of changes in the physiological state of *E. coli* expressing a heterologous membrane protein during carbon-limited fedbatch cultivation. *Biotechnol Bioeng* 92(1):69–78
- Christiansen T, Michaelsen S, Wuempelmann M, Nielsen J (2003) Production of savinase and population viability of *Bacillus clausii* during high-cell-density fed-batch cultivations. *Biotechnol Bioeng* 83(3):344–352
- Sonnleitner B, Locher G, Fiechter A (1992) Biomass determination. *J Biotechnol* 25(1–2):5–22
- Looser V, Hammes F, Keller M, Berney M, Kovar K, Egli T (2005) Flow-cytometric detection of changes in the physiological state of *E. coli* expressing a heterologous membrane protein during carbon-limited fed-batch cultivation. *Biotechnol Bioeng* 92(1):69–78
- Glick BR (1995) Metabolic load and heterologous gene expression. *Biotechnol Adv* 13(2):247–261
- Clements F, Bayer K (2006) Improvement of bioprocess monitoring: development of novel concepts. *Microb Cell Fact* 5:183–196
- Knorr B, Schlieker H, Hohmann H-P, Weuster-Botz D (2007) Scale-down and parallel operation of the riboflavin production process with *Bacillus subtilis*. *Biochem Eng J* 33(3):263–274
- Puskeiler R, Kaufmann K, Weuster-Botz D (2005) Development, parallelization, and automation of a gas-inducing milliliter-scale bioreactor for high-throughput bioprocess design (HTBD). *Biotechnol Bioeng* 89(5):512–523
- Harms P, Kostov Y, Rao G (2002) Bioprocess monitoring. *Curr Opin Biotechnol* 13(2):124–127
- Kiviharju K, Salonen K, Moilanen U, Eerikainen T (2008) Biomass measurement online: the performance of in situ measurements and software sensors. *J Ind Microbiol Biotechnol* 35(7):657–665
- Cervera AE, Petersen N, Lantz AE, Larsen A, Gernaey KV (2009) Application of near-infrared spectroscopy for monitoring and control of cell culture and fermentation. *Biotechnol Prog* 25(6):1561–1581
- Kaiser C, Carvell JP, Luttmann R (2007) A sensitive, compact, in situ biomass measurement system controlling and monitoring microbial fermentations using radio-frequency impedance. *Bio-Process Int* 5(Suppl. 1):52–55
- Kiviharju K, Salonen K, Moilanen U, Meskanen E, Leisola M, Eerikainen T (2007) On-line biomass measurements in bioreactor cultivations: comparison study of two on-line probes. *J Ind Microbiol Biotechnol* 34(8):561–566
- Knabben I, Regestein L, Schauf J, Steinbusch S, Buechs J (2011) Linear correlation between online capacitance and offline biomass measurement up to high cell densities in *Escherichia coli* fermentations in a pilot-scale pressurized bioreactor. *J Microbiol Biotechnol* 21:204–211
- Opel CF, Li J, Amanullah A (2010) Quantitative modeling of viable cell density, cell size, intracellular conductivity, and membrane capacitance in batch and fed-batch CHO processes using dielectric spectroscopy. *Biotechnol Prog* 26(4):1187–1199
- Wechselberger P, Seifert A, Herwig C (2010) PAT method to gather bioprocess parameters in real-time using simple input variables and first principle relationships. *Chem Eng Sci* 65(21):5734–5746
- Huang J, Kaul G, Cai C et al (2009) Quality by design case study: an integrated multivariate approach to drug product and process development. *Int J Pharm* 382(1–2):23–32
- Guebel DV, Canovas M, Torres NV (2009) Analysis of the *Escherichia coli* response to glycerol pulse in continuous, high-cell density culture using a multivariate approach. *Biotechnol Bioeng* 102(3):910–922
- Cimander C, Mandenius C-F (2002) Online monitoring of a bioprocess based on a multi-analyser system and multivariate statistical process modelling. *J Chem Technol Biotechnol* 77(10):1157–1168
- Kadlec P, Gabrys B, Strandt S (2009) Data-driven soft sensors in the process industry. *Comput Chem Eng* 33(4):795–814
- Bastin G, Dochain D (1990) Process measurement and control, 1: on-line estimation and adaptive control of bioreactors. Elsevier, Amsterdam, p 379
- Nielsen J, Villadsen J (1992) Modelling of microbial kinetics. *Chem Eng Sci* 47(17–18):4225–4270
- Stephanopoulos G, San K-Y (1984) Studies on on-line bioreactor identification. I. Theory. *Biotech Bioeng* 26(10):1176–1188
- van der Heijden RTJM et al (1994) Linear constraint relations in biochemical reaction systems: I. classification of the calculability and the balanceability of conversion rates. *Biotechnol Bioeng* 43:3–20
- Wilms B, Hauck A, Reuss M et al (2001) High-cell-density fermentation for production of L-N-carbamoylase using an expression system based on the *Escherichia coli* rhaBAD promoter. *Biotechnol Bioeng* 73(2):95–103
- Corchero JL, Cubarsi R, Vila P, Aris A, Villaverde A (2001) Cell lysis in *Escherichia coli* cultures stimulates growth and biosynthesis of recombinant proteins in surviving cells. *Microbiol Res* 156(1):13–18
- Geethalakshmi S, Narendran S, Pappa N, Ramalingam S (2012) Development of a hybrid neural network model to predict feeding method in fed-batch cultivation for enhanced recombinant streptokinase productivity in *Escherichia coli*. *J Chem Technol Biotechnol* 87(2):280–285
- Jenzsch M, Simutis R, Luebbert A (2006) Generic model control of the specific growth rate in recombinant *Escherichia coli* cultivations. *J Biotechnol* 122(4):483–493
- Jobe AM, Herwig C, Surzyn M, Walker B, Marison I, von Stockar U (2003) Generally applicable fed-batch culture concept based on the detection of metabolic state by on-line balancing. *Biotechnol Bioeng* 82(6):627–639
- Dietzsch C, Spadiut O, Herwig C (2011) A dynamic method based on the specific substrate uptake rate to set up a feeding strategy for *Pichia pastoris*. *Microb Cell Fact* 10(1):14
- Davey CL, Davey HM, Kell DB (1993) Introduction to the dielectric estimation of cellular biomass in real time, with special

- emphasis on measurements at high volume fractions. *Anal Chim Acta* 279:155–161
32. Heinzle E, Oeggerli A, Dettweiler B (1990) On-line fermentation gas analysis: error analysis and application of mass spectrometry. *Anal Chim Acta* 238:101–115
 33. Luedeking R, Piret EL (1959) Transient and steady state in continuous fermentation. Theory and experiment. *Journal Biochem Microbiol Technol Eng* 1(4):431–459
 34. Nielsen J, Villadsen J, Liden G (2002) *Bioreaction engineering principles*, 2nd edn. Kluwer Academic/Plenum Publishers, New York
 35. van der Heijden RTJM, Romein B, Heijnen JJ, Hellinga C, Luyben KCAM (1994) Linear constraint relations in biochemical reaction systems: II. diagnosis and estimation of gross errors. *Biotechnol Bioeng* 43:11–20
 36. van der Heijden RTJM, Romein B, Heijnen JJ, Hellinga C, Luyben KCAM (1994) Linear constraint relations in biochemical reaction systems: III. sequential application of data reconciliation for sensitive detection of systematic errors. *Biotechnol Bioeng* 44:781–791
 37. Madron F, Veverka V, Vanecek V (1977) Statistical analysis of material balance of a chemical reactor. *AIChE J* 23(4):482–486
 38. Kaiser C, Pototzki T, Ellert A, Luttmann R (2008) Applications of PAT-process analytical technology in recombinant protein processes with *Escherichia coli*. *Eng Life Sci* 8(2):132–138
 39. Shehata TE, Marr AG (1971) Effect of nutrient concentration on the growth of *Escherichia coli*. *J Bacteriol* 107(1):210–216
 40. Dabros M, Dennewald D, Currie DJ et al (2009) Cole–Cole, linear and multivariate modeling of capacitance data for on-line monitoring of biomass. *Bioprocess Biosyst Eng* 32(2):161–173
 41. Andersson L, Strandberg L, Enfors S-O (1996) Cell segregation and lysis have profound effects on the growth of *Escherichia coli* in high cell density fed-batch cultures. *Biotechnol Prog* 12(2):190–195
 42. Luttmann R et al (2012) Soft sensors in bioprocessing: a status report and recommendations. *Biotechnol J* 7(8):1040–1048
 43. Schuler M, Marison I (2012) Real-time monitoring and control of microbial bioprocesses with focus on the specific growth rate: current state and perspectives. *Appl Microbiol Biotechnol* 94(6):1469–1482