Process model adaption for the autotrophic cultivation of *Ralstonia eutropha* mutant strains

vorgelegt von
Dipl.-Ing.
Flavia Neddermeyer

von der Fakultät III - Prozesswissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
-Dr.-Ing.-

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr.-Ing. Matthias Kraume
Gutachterin: Prof. Dr.-Ing. Anja Drews
Gutachter: Prof. Dr.-Ing. Rudibert King

Tag der wissenschaftlichen Aussprache: 14. August 2020

Berlin 2020

Eidesstattliche Erklärung

von Flavia Neddermeyer, geboren am 11. April 1986 in Peine

Ich versichere an Eides statt, dass ich vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Berlin, den 1. September 2020,

Statutory Declaration

for Flavia Neddermeyer, born on April 11th 1986 in Peine

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Berlin, September 1, 2020,

Acknowledgement

While working on this thesis I was supported in many ways. At this point I would like to express my thanks.

My sincere gratitude goes first and foremost to the head of the department Prof. Rudibert King, who instructed me and generously and understandably shared his experience with me. In important situations he stood by me both professionally and mentally.

Prof. Anja Drews reviewed and evaluated my dissertation. I would like to thank her for taking on this time-consuming task with interest, thoroughness and for sharing her valuable expertise with me.

I would like to express my appreciation to my colleagues for the experimental support, the scientific exchange, the inviting working atmosphere and the cooperation.

My deepest gratitude goes to my family and my friends for their support, understanding and lots of childcare hours.

Furthermore, I would like to thank all the above mentioned for the numerous helpful comments, which resulted from proofreading and test presentations, which have increased the quality of this work.

Abstract

Bioprocesses, in which gases serve as substrates, have become popular in the past years, as they are used, for example, to upgrade industrial exhaust gases. In this thesis, first a regulatory concept for the autotrophic cultivation of the bacterium *Ralstonia eutropha* is presented. The focus lies on a model-based gas phase controller that works independently of the strain. It adjusts the gas composition in the reactor headspace and maintains the desired excess pressure. The gas phase controller is framed by the general, model-based control system for process optimization.

The second part deals with rapid model adaption for genetically modified strains. For process optimization, strain-dependent process models are used. Thus, the general process model has to be modified and reformulated for each strain before it can be utilized for model-based optimization. In order to shorten the labour-intensive, iterative step of modeling, different adaption methods were developed and applied.

The combination of a strain-independent gas phase control and a toolbox for fast adaption of the process model of genetically modified strains, enables a fast and resource-saving process adaption.

Kurzfassung

Bioprozesse, in denen Gase als Substrate dienen, haben in den vergangenen Jahren an Bedeutung gewonnen, da mit ihnen beispielsweise industrielle Abgase aufgewertet werden können. Diese Arbeit stellt zunächst ein Regelungskonzept zur autotrophen Kultivierung des Knallgasbakteriums *Ralstonia eutropha* vor. Dabei steht ein Regler zur Einstellung der Gasanteile sowie des Drucks im Gasraum im Vordergrund, der zwar modellbasiert, aber stammunabhängig arbeitet. Der Gasphasenregler ist ein Bestandteil der allgemeinen, modellbasierten Regelung zur Prozessoptimierung.

Im zweiten Teil geht es um die schnelle Modelladaption für genetisch modifizierte Stämme. Zur Prozessoptimierung werden stammabhängige Prozessmodelle verwendet, die für jeden genetisch modifizierten Stamm neu formuliert werden müssen, bevor die Kultivierung geregelt werden kann. Um den arbeitsintensiven, iterativen Schritt der Modellierung zu verkürzen, wurden verschiedene Methoden entwickelt und angewandt.

Die Kombination einer stammunabhängigen Gasphasenregelung und einem Methodenspektrum zur schnellen Anpassung des Prozessmodells an genetisch veränderte Stämme ermöglicht schließlich eine schnelle und ressourcenschonende Prozessanpassung.

Contents

List of Figures

List of Tables

General symbols

a	relative enzyme activity
A	total enzyme activity
c	concentration
$\mathbf{CV}_y$	covariance matrix of the measurements
d	diameter
e	control error, measurement error
ϵ	extinction coefficient
Fr	Froude number (dimensionless)
g	the outcome/value of a kinetic function, gravitational constant
γ	system dependent constant for $\mathrm{k_L a}$ calculation
H	Henry coefficient
i, l, m	running index
I	inhibition state, current
k	kinetic parameter
K, F	constant factor, controller gain K
$\mathrm{k_L a}$	volumetric gas transfer coefficient
L	length
m	mass
M	molar weight
μ	reaction rate for cell components
n	amount of material, stirring speed
N	number
ν	molar gas consumption rate, kinematic viscosity
Ne	Newton number (dimensionless)
p	partial pressure
P	pressure
ΔP	excess pressure
P_0	ambient pressure
$\mathbf{P}_0$	initial covariance of the states
ϕ	cost function
q	gaseous volume flow
Q	gas load
$\mathbf{Q}$	spectral density matrix of the systems noise
r	reaction rate, reference
R	universal gas constant, $R = 8.314\ \mathrm{J\,mol^{-1}\,K^{-1}}$

ρ	density
$\mathbf{R}$	covariance matrix of measurement noise
Re	Reynolds number (dimensionless)
σ	sample variance
t	time
T	temperature
T_1	first order delay
θ	parameter
$\underline{u}$	vector of manipulated variables
U	voltage
$U_{i,j}$	specific substrate uptake coefficient: substrate i per product j
v	empirical variation coefficient
V	volume
w	gas empty tube speed
$\mathbf{W}$	measurement weighing matrix
x	molar fraction
$\underline{x}$	vector of state variables
y	measurement
$\underline{y}$	vector of measurement variables

Modifiers of variables

$\bar{a}$	mean value a
$\underline{a}$	a has vector size
$\hat{a}$	simulated value of a
$\dot{a}$	time derivative of a

Superscripts

app	apparent
in	inlet
n	molar
prior	of the previous cycle
T	transposed
red	reduced

Subscripts

0	quantity at time t_0, ambient
30	30 °C
a	enzymatic activity
con	microbially consumed
cy	cyanophycin
CO_2	carbon dioxide
deg	degradation
est	estimated
exp	experiments
feed	inflow
Fe	iron and trace elements
ff	feedforward
FFDR	feedforward disturbance rejection
gas	index for H_2, CO_2, O_2
gas model	gas phase model
head	headspace of the reactor
int	integrated
In	long-term inhibition caused by oxygen exposure
H_2	hydrogen
i	component i
I	integral
k	sampling point k
kla model	gas transfer model
l	liquid
leak	leakage
max	maximal
meas	measured
min	minimal
MBH	membrane-bound hydrogenase
model I	general process model (I)
model II	process model without gas transport (II)
MP	membrane protein
N	ammonium
OD	optical density
O_2	oxygen

P	phosphate
P	proportional
ΔP	excess pressure
PHA	polyhydroxyalkanoat
PHB	polyhydroxybutyrate
Pr	protein
rel	relative
sat	saturation concentration in equilibrium
rest	combined residual gases
SH	soluble hydrogenase
t	total
trans	transferred via the liquid-gas interface
v	vapour
X	biomass or active biomass

Abbreviations

ABC	Advanced Batch Control software
Ai	Aiba kinetic function
CL	control loop
DCU	Digital Control Unit of the bioreactor
DEKF	Dynamic Extended Kalman Filter
DGL	differential equations
EKF	Extended Kalman Filter
FFDR	feedforward disturbance rejection
HB	hydroxybutyrate
lb	lower bound
min	minimum
max	maximum
MBH	membrane-bound hydrogenase
MFCS	Multiple Fermenter Control System
Mo	Moser kinetic function
MiMe	Michaelis–Menten kinetic function
NAD^+	oxidized nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
OD	optical density
OED	Optimal Experimental Design
PHB	polyhydroxybutyrate
PI	parameter identification/parameter estimation
PI	proportional-integral
REatc	*Ralstonia eutropha* autotrophic and gas phase controlled
Ro_1	one-parametric specific kinetic function
Ro_2	two-parametric specific kinetic function
rpm	rounds per minute
TP	trajectory planning
R. e.	*Ralstonia eutropha*
SH	soluble hydrogenase
SI	international system of units
Spec	Specific kinetic function
SPKF	Sigma Point Kalman Filter
ub	upper bound

Chapter 1

Introduction

One of today's great challenges is to reduce global warming. Sustainable and clean alternatives to petroleum-based fuels are sought after. Nowadays, sugar-based plants like corn are used to produce ethanol, respective biodiesel. However, using plant biomass as industrial fuel leads to ethical conflicts because it could as well be used to nourish people. Hence, the sugar-based bioindustry must adapt by using alternative and ethically correct substrates like industrial exhaust gases. Bioprocesses fed with exhaust gases have already been investigated in the past on a laboratory scale for various products. Humphreys and Minton (2018) also mention an industrial application example, in which such a gas fermentation was coupled to a steel mill. Liew et al. (2016) stated that the advantages of gas fermentation over classical chemical processes are product specificity and the reduction of the need for fossil fuels. An interesting organism that is able to grow on the exhaust gas component CO_2 is *Ralstonia eutropha*, also named *Cupriavidus necator*. This bacterium is of industrial relevance because it grows very fast with CO_2 as substrate and naturally produces the bioplastic polyhydroxybuturate (PHB). Modifying the PHB-pathway on a genetic level, enables to produce other industrial relevant compounds such as alkenes, which was investigated by Crépin et al. (2016) or terpenes (Krieg et al. (2018)). In order to convert CO_2 into industrially relevant products by cultivation, a process has to be setup first, later optimized and the strain must get metabolically engineered as pointed out by Takors et al. (2018). How the process is optimally designed depends on the metabolic properties of the final organism. This means that process optimization is theoretically only possible after metabolic engineering. It is desirable, however, to simultaneously optimize both the genetics of the strain and the process in order to save time.

1.1 Motivation and outline of this thesis

In a bioprocess, the cultivation conditions, e.g., temperature, nutrients, pH value, shear stress, determine growth and product formation of the selected organism. If the effects of these conditions on the metabolism of the organism are known, the bioprocess can be modeled mathematically. This model makes it possible to predict the effects of changing cultivation conditions. And vice versa, if a certain product is to be maximized, the model can be used to calculate the optimal trajectories for the cultivation conditions, i.e., to design the process. Since the conditions or states, e.g., nutrient concentrations, depend among other things on the system inputs, e.g., feeding rates, it is possible to calculate the optimal system inputs with the process model. During cultivation, the process model combined with observation techniques can also be used to track the prior optimized state trajectories, which is especially important when the system is disturbed and the model no longer provides an accurate description. In summary, a process model allows model-based control that is important to run a cultivation successfully, e.g., to maximize the target compound.
But the development of mathematical models for bioprocesses is very time-consuming and requires a lot of metabolic information about the strains and a lot of cultivation data. To generate data, cultivations must be performed before a process model is developed. These first cultivations are usually designed by experts who postulate appropriate input trajectories, e.g., feeding profiles, based on their experience. During cultivation, the experts react manually to disturbances, e.g. adjusting the feeding in case the cells of the innoculum are less vital than expected. Successful first cultivations would provide dynamic data of the states, which would enable metabolic conclusions to be recognized and thus modeled.
Parallel to process design, the organism is often genetically modified to improve the expression of a particular product. Genetic engineering alters the metabolism of the organism and therefore the process model originally developed for the original strain no longer provides an exact description of the process. It must be adapted to the metabolic behavior of the mutant strain designed.
For cultivation with autotrophic organisms growing on gases, such as *Ralstonia eutropha*, nutrients are of liquid and gaseous nature. The organisms consume the dissolved gases and to maintain the gas solubility equilibrium, the gas is physically transported from the gas phase to the liquid phase. As a result, the pressure in the gas phase of the culture vessel decreases. In order to counteract the pressure losses, the cultivation vessel needs a gas supply. The required gas flow consisting of hydrogen, carbon dioxide and oxygen is unknown and changes during cultivation because it depends on the metabolism of the organisms and their total number that drastically increases in a fed-batch cultivation. As a fed-batch is a nonlinear process, the individual gas flows cannot be calculated with linear approaches. Thus, it seems that the cultivation of autrotrophic organisms requires a model not only for maximizing the target substance. The process model is also needed for the gas flow calcula-

tion for cultivation. But to develop a model, data from cultivations must already exist. In order to overcome this vicious circle, the required gas flows must be calculated in cultivation without using metabolic information of the strain. Hence, to carry out the first cultivations, a process model-free gas control is required, which calculates the gas flows. Later, when the process model exists, it still is favorable to use a gas control that is not based on the process model for two reasons: First, the gas control works although the process model may be wrong. And second, if the strain is genetically modified and changes its metabolic behavior, the gas control does not fail. Hence, strain-independent gas control implies that unknown organisms can be cultivated autotrophically.

In this thesis, the framework for the parallel engineering of genetics and the process for the gas cultivation of *Ralstonia eutropha* (*R. e.*) was created. For this purpose, a strain-independent gas control system was developed. It enables to run first autotrophic cultivations and thus the generation of data, which is necessary for modeling the process. Using the cultivation data, a process model for the *R. e.* wild-type H16 was developed, which forms the basis for model adaptions. Cultivations with mutated strains were also carried out to generate data by using the strain-independent gas phase control. Obtained cultivation data was then utilized to adapt the original process model. Adaptions were made with the help of special methods to save time and not have to perform the entire labor-intensive modeling for each genetically modified strain.

Particular metabolic properties of the modified organisms and those of *R. e.* H16 (wild-type) are described in Chapter 2. These metabolic properties were taken into account when the process model for *R. e.* H16 was developed and later adapted to mutant strains. To develop process models, quantities such as nutrient or biomass concentration must be measured. The measurement methods used are explained in Chapter 3. All *R. e.* strains were cultivated autotrophically, i.e., gases instead of sugars served as substrates. A suitable bioreactor design and the control architecture are explained in Chapter 4. This model-based control architecture enables process optimization, e.g., the maximization of a certain product. Due to the autotrophic nature of the process, the gas phase must be controlled. A part of the control architecture is the gas phase controller, which operates independently of the metabolic properties of the organism but is based on a physical gas phase model. Gas phase control is presented in the Section 4.4 together with the gas phase model. As mentioned above, the gas phase controller works independently of the strain and thus enables the cultivation of mutant strains with unknown metabolism and obtained cultivation data can later be used for model adaption. Basis for the model adaption was a process model of *R. e.* H16, which is presented and discussed together with its cross-validations in Chapter 5. The process model for H16 contains, in contrast to the gas phase model, descriptions of the metabolism and therefore only describes the cultivation of this specific strain. For the genetically modified strains, it must be adapted according to the properties of the mutant strains in order to enable process optimization, e.g., product maximization. A methodology for a rather fast

model adaption of the general H16 process model for mutant strains is explained in Chapter 6. Resulting process models are presented in Chapter 7 and the success of the presented model adaption framework is discussed in Chapter 8.

Chapter 2

Cultivated strains

The present thesis deals with model adaption for mutant strains. Starting point for adaption is the model for cultivating the wild-type (H16) of *Ralstonia eutropha*.
Ralstonia eutropha (R. e.) is a bacterium that can live on both sugar and carbon dioxide, which corresponds to the heterotrophic or lithoautotrophic metabolism shown in Figure 2.1.

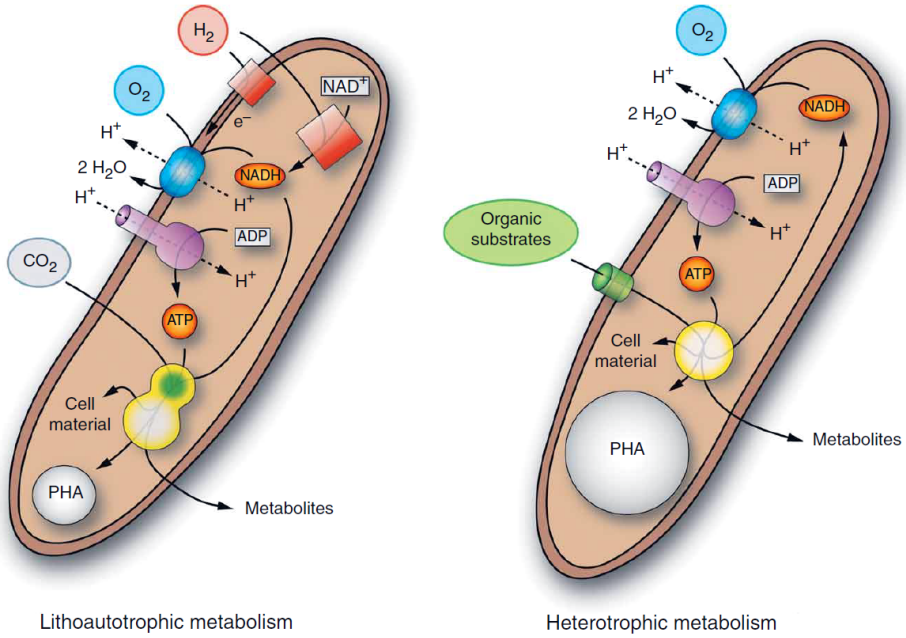


Figure 2.1: The main growth modes of *R. eutropha* published by Pohlmann et al. (2006). Schematic representation illustrating the key aspects of lithoautotrophic and heterotrophic metabolism. The yellow discs represent the process of the central metabolism whereas the green disc is the Calvin-Benson-Bassham cycle. The red squares symbolize the two energy-conserving hydrogenases. The gray round symbols indicate polyhydroxyalkanoate (PHA) storage granules.

In all cultivations conducted for the present thesis, a gas mix containing carbon dioxide, hydrogen and oxygen was fed. Carbon dioxide served as carbon source. Hydrogen together

with oxygen were absorbed to generate energy, namely ATP and other high-energy reduction equivalents such as NADH. *R. e.* consumes energy to assimilate carbon dioxide, ammonium, phosphate, iron and trace elements, which are utilized by the cells as building blocks for macromolecules. In case of ammonium or phosphate insufficiency, *R. e.* is well equipped and converts the present gaseous substrates into the storage polymer polyhydroxybutyrate (PHB) that belongs to the group of polyhydroxyalkanoates (PHAs), which were investigated to be used in tissue engineering (Sultana (2012)). Once all required nutrients are present, PHB can be transformed into active biomass. The following metabolic effects resulting from a deficient gas compound can be derived from Figure 2.1:

- Hydrogen deficiency:
  Smaller NADH/NAD^+-ratio or decreased oxygen uptake, smaller H^+-gradient and consequently less ATP, which leads to decreased oxygen and carbon dioxide uptake.
- Oxygen deficiency:
  Higher NADH/NAD^+-ratio or decreased hydrogen uptake, smaller H^+-gradient and thus less ATP, which leads to decreased carbon dioxide uptake.
- Carbon dioxide deficiency:
  Either energy is used for the conversion of PHB to biomass that comes along with a slightly decreased uptake of hydrogen and oxygen, or the gas uptake is decreased drastically because the internal PHB storage is empty, and therefore a conversion impossible.

In addition to H16, genetically modified strains were also cultivated for this thesis. The metabolic properties of the modified strains are explained in the following sections.

2.1 *Ralstonia eutropha* HF805

The strain HF805 was developed by Goris et al. (2011) to allow simpler purification of the enzyme membrane-bound hydrogenase (MBH), which is an important enzyme in hydrogen assimilation and thus energy metabolism. It catalyzes the reversible cleavage of H_2 and, moreover, is oxygen tolerant (Saggu et al. (2009)) and thus investigated for the biologic fuel production of H_2 (see Goldet et al. (2008)). In HF805, the MBH is expressed with a Strep-tag® bound to its terminal. Strep is a synthetic peptide that shows an affinity to Strep-Tactin® that can be utilized for a chromatographic purification after cell disruption. Additionally, this strain cannot produce the soluble hydrogenase (SH) since relevant genes were removed from its megaplasmid. The enzyme SH reduces co-factors such as NAD(P), as reported by Lauterbach et al. (2013). Since the regeneration of co-factors is relevant for biotransformation processes, *R. e.* and SH were investigated in this context by Ratzka (2011), for instance.

2.2 *Ralstonia eutropha* HF951

In HF951, gene sequences for the H_2-metabolism were removed from the native megaplasmid DNA and transferred to an additional plasmid. It features the genetic code for the synthesis of cyanophycin (cy) and resistance towards tetracycline, which avoids a loss of the extra plasmid. As *R. e.* can be grown autotrophically as well on SI-labelled CO_2, cyanophycin is an interesting candidate for a SI-labelled oligomere. The strain HF951 is unable to produce the carbon storage compound PHB.

2.3 *Ralstonia eutropha* H798

Lauterbach (2013) developed the strain H798 that expresses the soluble hydrogenase (SH) but no MBH. All genes for SH are coded on an extra plasmid together with the genes to express a fluorescent protein named Frex, which was originally created by Zhao et al. (2011), and a tetracycline resistance to avoid losing inserted genes. An expression of Frex is coupled to the SH promoter meaning that SH and Frex are translated simultaneously. Consequently, the amount of Frex proteins in a cell is related to the amount of SH. On a metabolic level, the SH regenerates cofactors by reducing NAD^+ to NADH. Once NADH binds to the Frex protein, it fluoresces at 515 nm while being stimulated at 490 nm. Reason for the fluorescence are conformational changes. For process development towards an improved SH production, this modification is of interest with respect to an easy online but indirect measurement of SH.

In order to evaluate cultivations and describe the growth and production behavior of these strains with mathematical models, liquid and gaseous substrates, intra- and extracellular products, biomass and intracellular components such as PHB have to be measured. For the variables important for this work, the next chapter summarizes the measurement methods and their uncertainties, whereby the measurement uncertainties of automated sensors are discussed in detail in Rossner (2014).

Chapter 3

Measuring methods

Regardless of the measured compound (i), at least two manually-drawn samples were analyzed in parallel and their mean value served for parameter estimation. A relationship between mean values and their corresponding sample variances ($\sigma_{i,\mathrm{meas}}$) was observed. The measurement variance of an N_l-multiple analysis (e.g., $N_l = 2$ for duplicates, $N_l = 3$ for triplicates,...) is defined as

$$\sigma_{i,\mathrm{meas}} = \sqrt{\frac{\sum_{l=1}^{N_l}(y_{i,l,\mathrm{meas}} - \bar{y}_i)^2}{N_l - 1}} \tag{3.1}$$

$$\bar{y}_i = \frac{\sum_{l=1}^{N_l} y_{i,l,\mathrm{meas}}}{N_l}, \tag{3.2}$$

with $y_{i,l,\mathrm{meas}}$ being the l-th measurement of analyte i and $\bar{y}_i$ the corresponding mean value. The more often one sample is analyzed, the higher N_l and the more accurate $\sigma_{i,\mathrm{meas}}$. However, analysis is costly, and thus samples were mostly evaluated in duplicates. An alternate way to eq. (3.1)–(3.2) for estimation of sample variance was sought for. The aim was that only for a reduced sample number N_m a multiple analysis with $N_l > 2$ was required and not for all drawn samples.
It was observed already by Heine (2004) that for higher mean values the sample variance increased as well. Therefore, a linear relationship was postulated using

$$\hat{\sigma}_i = \sigma_{0,i} + \theta_{\sigma,i} \cdot \bar{y}_i. \tag{3.3}$$

Employing a maximum likelihood estimation with unknown variance of $\sigma_{m,\mathrm{meas}}$ as described in Söderström (1989), the parameters $\sigma_{0,i}$ and $\theta_{\sigma,i}$ were identified. Because the variance needs to be estimated for each compound i as well, the likelihood description reduced to

the cost function

$$\phi = \ln \frac{1}{N_m - N_\theta} \sum_{m=1}^{N_m} \left(\hat{\sigma}_m - \sigma_{m,\text{meas}}\right)^2, \tag{3.4}$$

that was minimized, with N_m being the number of multiple-analyzed samples, N_θ the number of parameters to be identified in the approximation function eq. (3.3), herein assumed to be linear and therefore N_θ =2. The variables $\sigma_{m,\text{meas}}$ and $\hat{\sigma}_m$ are the measurement variance of analysis according to eq. (3.1) and variance calculations as in eq. (3.3) for each measurement sample m, respectively. According to the cost function, the minimized cost value equals the logarithmized variance of the approximation, and therefore the approximation error is defined as

$$\text{error}_\phi = \sqrt{e^\phi}, \tag{3.5}$$

i.e., the sigma confidence interval which includes about 68 % of all values that were used in this linear regression. The variance approximation methods described above have been applied to the manual analysis methods that will be introduced in the following sections. Parameter values obtained by the approximation and errors will be shown in tables and graphs. The following sections deal with the employed methods of analysis to gather measurements that are needed to identify model parameters. For parameter estimation, the agreement of model prediction $\underline{\hat{y}}$ and experimental data $\underline{y}$ was measured by weighted squares formula employing the weighing matrix $\mathbf{W}$

$$\phi(\underline{\theta}) = \sum_{m=1}^{N_\text{exp}} \sum_{l=1}^{N_y} \left((\underline{y} - \underline{\hat{y}})^T \cdot \mathbf{W} \cdot (\underline{y} - \underline{\hat{y}}) \right), \tag{3.6}$$

which served as cost function ϕ to be minimized by optimizing the parameter vector $\underline{\theta}$

$$\underline{\theta}_\text{opt} = \arg \min_{\underline{\theta}} \left[\phi \left(\underline{y}, \underline{\hat{y}}(\underline{x}, \underline{u}, \underline{\theta}), \mathbf{W} \right) \right]. \tag{3.7}$$

To this end, the weighted deviations were summed up over the number of measuring points (N_y) and the number of experiments (N_exp) used.
To analyze the parameter uncertainties, the Fisher information matrix for one sampling time $\mathbf{F}$ was calculated according to Walter and Pronzato (1997) by

$$\mathbf{F} = \left(\frac{\partial \underline{y}}{\partial \underline{\theta}} \right)^T \cdot \mathbf{CV}_y^{-1} \cdot \left(\frac{\partial \underline{y}}{\partial \underline{\theta}} \right), \tag{3.8}$$

with $\mathbf{CV}_y$ representing the measurement noise matrix that is multiplied with the sensitivities. Off-elements of $\mathbf{CV}_y$ were set to zero, i.e., covariances of measurement errors neglected, and the diagonal entries were the squared approximated measurement errors

$$\mathbf{CV}_{y,ii} = \hat{\sigma}_i^2. \tag{3.9}$$

When **F** is summed over all sampling times and then inverted, the resulting matrix gives lower bounds for the covariances of the estimated parameters, where diagonal values represent variance values.
When modeling and estimating parameters using the cost function definition of eq. (3.6), it is more important that simulations cover measurement points with low variance than those showing high noise. Thus, sample variances need to be considered in the weighing matrix **W**. For in-situ measurements, such as dissolved gas concentrations, sample variance equals the probe's accuracy, which is usually indicated by the manufacturer or discussed in Rossner (2014). For all analytical methods applied to manually-drawn samples, the measurement variance has been estimated as described above.
Analytical procedures for the quantification of phosphate, ammonium and biomass concentrations, including an estimation of the sample variances, are presented in Section 3.1. The hydrogenase activity essay and its accuracy is described in Section 3.2. Details for PHB analysis are given in Section 3.3. Moreover, for selected cultivations, extracellular metabolites were analyzed according to the description given in Section 3.4 in order to identify carbon-based secreted metabolites.

3.1 Analytic methods for salts and biomass

Phosphate was quantified with the Phosphate FS* test kit from DiaSys (Germany) and the Berthelot reaction described in Rhine et al. (1998) served for ammonium determination. Both analytic methods rely on extinction and were performed in microplates using plate reader Sunrise from TECAN (Austria). After diluting, the reaction was arranged in duplicates that were measured photometrically twice each.
Optical density (OD) was measured twice after dilution in the absoption rage of 0.2–0.3 with Pharo 300 from Spectroquant (Germany) at 436 nm using 5 mL cuvettes.
To measure biomass concentrations, a defined volume of cultivation broth was sucked through preweighed membrane filters of 2 µm pore size, the filters were dried at 95 °C and reweighed afterwards. The biomass concentration was measured in duplicates.
As in Section 3, the sample variances according to eq. (3.1) were approximated in a linear way according to eq. (3.3). To this end, selected samples were analyzed in triplicates in order to obtain more reliable mean values and variances. These analysis and approximation results are visualized in Figure 3.1. Identified sample variance parameters and remaining approximation errors are listed in Table 3.1. The presented evaluation of measurement noise does not only imply inaccuracies of the method itself that is caused by technical devices and inaccurate handling. It also takes into account errors caused by the different types of sampling, i e., automatic sampling by an autosampler and manual sampling by one person, as well as deviations by the different laboratory staff performing the procedures. In addi-

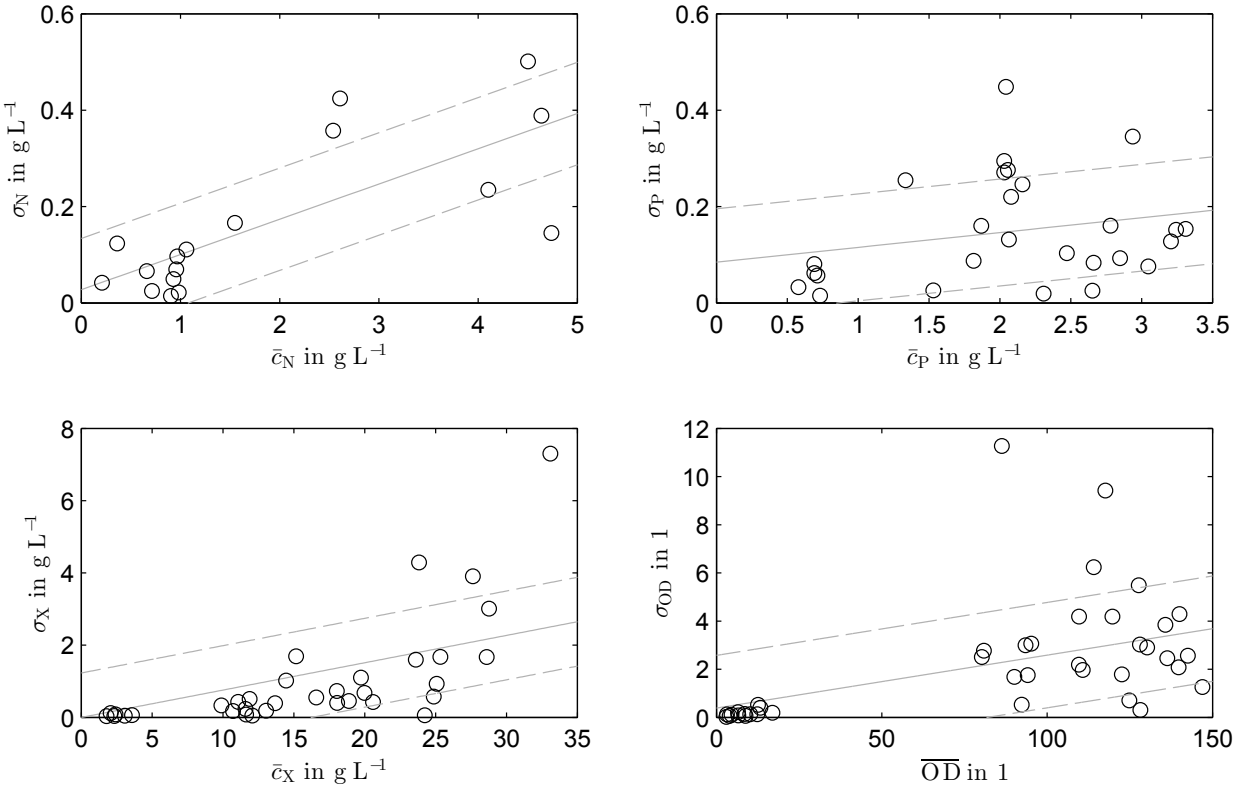


Figure 3.1: Sample variances plotted over the mean ammonium, phosphate, biomass concentration or OD (black) and linear approximation (solid, gray) with confidence intervals (dashed, gray).

Parameter	Value	Unit	Parameter	Value	Unit
$\sigma_{0,\mathrm{N}}$	0.03	g L^{-1}	$\sigma_{0,\mathrm{P}}$	0.09	g L^{-1}
$\theta_{\sigma,\mathrm{N}}$	0.07	-	$\theta_{\sigma,\mathrm{P}}$	0.03	-
error$_{\phi,\mathrm{N}}$	0.11	g L^{-1}	error$_{\phi,\mathrm{P}}$	0.11	g L^{-1}
$\sigma_{0,\mathrm{X}}$	0	g L^{-1}	$\sigma_{0,\mathrm{OD}}$	0.39	-
$\theta_{\sigma,\mathrm{X}}$	0.08	-	$\theta_{\sigma,\mathrm{OD}}$	0.02	-
error$_{\phi,\mathrm{X}}$	1.29	g L^{-1}	error$_{\phi,\mathrm{OD}}$	2.18	-

Table 3.1: Maximum likelihood estimation results for linear measurement noise approximations of salts and biomass visualized in Figure 3.1

tion, measurement deviations are included that can be justified with a storage time of up to 8 months and error-prone sample dilutions before analysis.

3.2 Activity analysis of hydrogenases

To estimate the model parameters belonging to hydrogenase formation and degradation, measurements of the internal membrane-bound and soluble hydrogenases were required. In all experiments of the wild-type strain H16, the enzyme was not tagged and thus a pu-

rification would have been complicated. Instead, an activity assay was performed and the measured activity was related to the amount of membrane-bound protein and soluble protein, respectively. The measurement of membrane-bound hydrogenase (MBH) and soluble hydrogenase (SH) of four sampling times takes roughly two days. In the following description of the analysis method, the duration of time-consuming steps is specified in parentheses. If not indicated, a duplicate sample was taken from the reactor at each sampling time, processed and measured twice each, leading to four measurements per sampling. MBH and SH were measured using a two-step method consisting of cell disruption and activity measurements. All required liquids as well as the cells and the lysate were stored on ice during the assay and technical equipment was precooled at 4 °C. After the description of the analysis stages, sample variance approximations are presented, which were calculated as described in the Section 3.

Stage 1: Cell disruption

The cell disruption stage was limited to 8 samples (4 duplicates) due to technical equipment. The cultivation pellets were defrosted and dissolved in potassium phosphate buffer by vortexing after adding approximately 1 mg of DNAse (0.5 h). Then, the cells were disrupted (2.5 h) by employing a french press HTU-DIGI-Press (G. Heinemann Ultraschall und Labortechnik). To remove cell debris, the samples were centrifuged at 4000 rpm (0.5 h). Membrane compounds were separated from liquid cell content by ultracentrifuging with Optima XE 90 (Beckman Coulter, USA) at 36000 rpm (1 h). After centrifuging, the supernatant contained soluble proteins and was transferred in Eppendorf tubes (0.25 h) for the SH activity assay. The pellet included membrane-bound proteins and was resuspendized in phosphate buffer for washing (0.5 h). A second ultracentrifugation step again separated the membrane compounds from the buffer (1 h). The liquor fractions were discarded and the pellets resuspendized in phosphate buffer, transferred in Eppendorf tubes and stored in the freezer at -80 °C until the MBH activity assay (Stage 2) was performed.

Stage 2: SH and MBH activity assay

During the second ultracentrifugation run, the activity of SH was measured by an H_2-driven NAD^+ reduction (3 h) as described in Lauterbach and Lenz (2013). Here, for activity analysis, the stable soluble extract was used instead of purified protein, and thus it was not necessary to add reducing agents. The total soluble protein amount was measured (1.5 h), using a BCA protein assay kit from Thermo Fisher (USA), so that measured activities could be related to the amount of protein in the soluble fraction.
Before measuring the activity of MBH, the resuspendized pellets of Stage 1 were defrosted. MBH catalytically changes optical properties of methylene blue. Using this effect, the activity was measured photometrically (3 h) according to Schlink and Schlegel (1979), but

without a glucose-based chemical oxygen trap. Instead, rubber-stoppered cuvettes were used. To obtain relative activity values, the total membrane-bound protein amount of the pellets was measured, again using the BCA protein test kit (1.5 h).

Measurement variance of MBH and SH

After the samples of several cultivations had been analyzed for MBH and SH, variations were observed between the four activity measurements (Stage 2) of one sample (each duplicate was measured twice). Based on the sample variance $\sigma_{\mathrm{MBH,meas}}$ and $\sigma_{\mathrm{SH,meas}}$ assessed with eq. (3.1) and the corresponding mean values $\bar{y}_{\mathrm{MBH}}$ and $\bar{y}_{\mathrm{SH}}$, the empiric variation coefficient was calculated by

$$v_{i,\mathrm{meas}} = \frac{\sigma_{i,\mathrm{meas}}}{\bar{y}_i} \cdot 100, \tag{3.10}$$

leading to values as high as 50 %. In order to find the cause of these large variances, various hypotheses were tested. Finally, this test led to a change in the procedure of the analysis of MBH and SH. Possible origins of these large variations are discussed in the following before presenting the final measurement variance approximation.

Enzymes, such as hydrogenases, are temperature sensitive. In this case, the temperature had only a small influence on the analysis variance because most rooms and equipment, which were used for analyzing and sampling, were temperature controlled.

Moreover, the cell disruption step has an impact on the purity of the processed samples and its total amount of hydrogenases. In this analysis of MBH and SH, variations in the efficiency of a cell disruption step did not have a major impact, since the relative amount of activity was measured. Hence, the activity per amount of membrane-bound or soluble protein was recorded and it was assumed that the cell disruption quality affected both, measured activity and amount of protein.

The duration of each analysis step given above is only an approximate value and depends strongly on the member of staff as well as on the total amount of sample load. Varying times of an analysis step lead to variations in air exposure, and hence to different degrees in protein inactivation. Air exposure is critical, when the enzymes are active, i.e., defrosted. There is only one step in the whole procedure where the different samples were stored for significantly different periods of time in the defrosted state, namely photometric measurement. In total, this measurement took 3 h for MBH or SH. However, only one sample after the other could be measured and so some samples were stored almost 3 h before the measurement, while others were measured immediately. The activity loss over time during analysis was determined for SH and MBH by measuring the same sample every 30 to 60 min for three hours in total. For each sample, averaged activity values and the measurement noise resulting from the four activity measurements were calculated. In Figure 3.2 the measured activities including errors for MBH at different times of analysis (circles) as well as the linear regression

are plotted. Averaging all six slopes results in an activity loss over time of -0.4 U mg_{MP}^{-1} h^{-1}

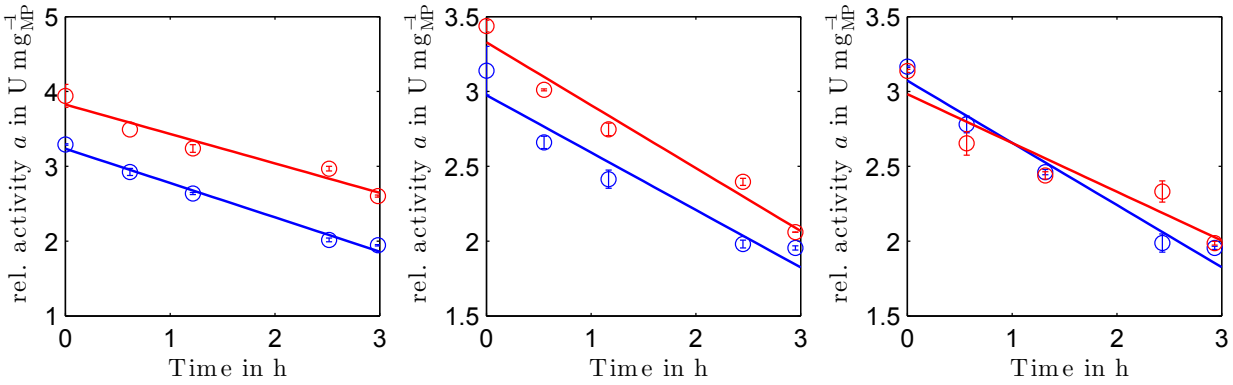


Figure 3.2: MBH activity during the analysis period (circles) and linear regression. Each figure corresponds to one sampling time in a cultivation where nitrogen gassed (red) and normal samples (blue) were taken in duplicates.

with a relative error of 25 %. For all prior MBH measurements, the approximate exposure time was known because the photometric measurements were recorded with time stamps. Thus, all previous MBH measurements were corrected for that activity loss over time. For all future MBH analysis, the procedure was modified such that one sample after another passed the processing steps of defrosting, dispersing and measuring. Thus, the storage time on ice during photomtetric analysis was decreased and a potential activity loss diminished. For SH, the activity loss over time was determined in a similar way using all six regressions resulting in an average slope of -0.13 U mg_{Pr}^{-1} h^{-1} with a relative error of 70 %, see Figure 3.3. Due to the large error, the linear approximation of activity loss over time seems to apply less to SH than to MBH. Therefore, all previous SH measurements were not corrected.
In the PhD thesis of Fritsch (2011), the influence of different oxygen concentrations on the

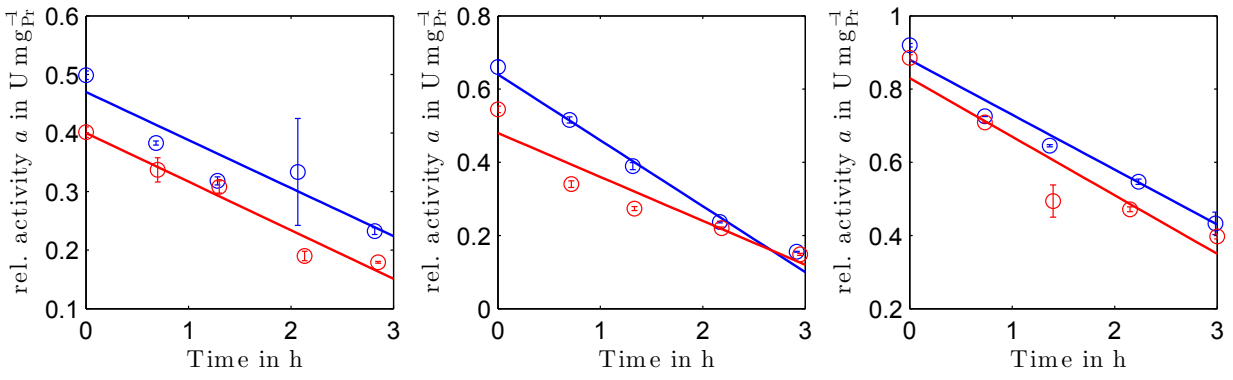


Figure 3.3: SH activity during the analysis period (circles) and linear regression. Each figure corresponds to one sampling time in a cultivation where nitrogen gassed (red) and normal samples (blue) were taken in duplicates.

activity of MBH was investigated. Accordingly, high oxygen leads to decreased activities,

which explains the above mentioned activity loss by air exposure during photometric analysis. Not only during analysis, but also during sampling, the hydrogenase in its active state was in contact with environmental oxygen.
To analyze the effect of variations in duration of oxygen exposure while sampling in this thesis, quadruplicate samples were taken. One sample (duplicate) was gassed with nitrogen whenever possible during the sampling procedure and the other one (duplicate) was taken according to the standard procedure, exposed to 20 % atmospheric oxygen. When comparing the measured activities as presented in Figure 3.3 and 3.2, respectively, it cannot be observed that the nitrogen gassed samples show systematically higher or lower relative activities neither for SH nor MBH. Nonetheless, between some of the two duplicates, variation coefficients up to 50 % were calculated. To cope with these large sample variances for MBH and SH, measurement variance approximations as described in Section 3 were performed and identified parameters are listed in Table 3.2. The approximations are intended to com-

Parameter	Value	Unit	Parameter	Value	Unit
$\sigma_{0,\mathrm{MBH}}$	0.0523	$\mathrm{U\,mg_{MP}^{-1}}$	$\sigma_{0,\mathrm{SH}}$	0.0287	$\mathrm{U\,mg_{Pr}^{-1}}$
$\theta_{\sigma,\mathrm{MBH}}$	0.0699	-	$\theta_{\sigma,\mathrm{SH}}$	0.1256	-
$\mathrm{error}_{\phi,\mathrm{MBH}}$	0.28	$\mathrm{U\,mg_{MP}^{-1}}$	$\mathrm{error}_{\phi,\mathrm{SH}}$	0.14	$\mathrm{U\,mg_{Pr}^{-1}}$

Table 3.2: Maximum likelihood estimation results for a linear measurement noise approximation visualized in Figure 3.4

pensate for errors that include the time-induced loss of activity for SH that, unlike MBH, was not quantifiable. For MBH, all measurements were first corrected and then evaluated. Results of variance approximation are plotted in Figure 3.4. Mainly duplicates (i.e., four activity measurements) but also quadruplicate samples (i.e., eight activity measurements) were employed to calculate mean values and sample variances. To get a more accurate idea

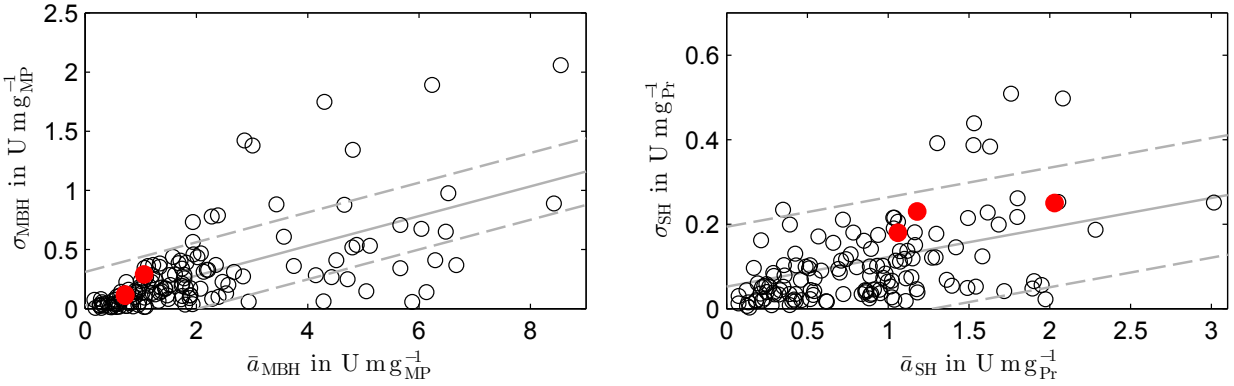


Figure 3.4: Measurement errors of MBH and SH duplicates/quadruplicates (black circles) and octuplicates (filled red circles) plotted against the mean activity values $\bar{a}_i$ and the linear regression (solid line) with confidence intervals (broken lines).

of the measurement noise by increasing sample size, an eightfold analysis was performed

for three selected samples. Eight samples were taken at the same time and then processed further. Six of the eight were centrifuged immediately, whereas two samples were stored in the fridge at 4 °C for 15 min before centrifuging. For each of the eight samples, the absolute activity of SH and MBH was determined photometrically twice. Then, to obtain relative values the amounts of membrane-bound and soluble proteins were also determined in duplicates.
Important results were that the refrigerating step after sampling seemed to have no influence on the measured activity. The sample noises determined from the eightfold analysis yielded values between 0.1 $\mathrm{U\,mg_{MP}^{-1}}$ – 0.3 $\mathrm{U\,mg_{MP}^{-1}}$ for MBH and between 0.18 $\mathrm{U\,mg_{Pr}^{-1}}$ – 0.25 $\mathrm{U\,mg_{Pr}^{-1}}$ for SH. All octuplicate sample variance values lie within the sigma confidence range of the identified linear regression as shown in Figure 3.4.

3.3 Quantification of polyhydroxybutyrate

In some cultivations, the manual samples were also analyzed for polyhydroxybutyrate (PHB). This analysis method used is based on the two-stage PHB analysis of Jacquel et al. (2008). Because PHB is an intracellular compound, at first the cells were disrupted and the non-soluble fraction was isolated (Stage 1). Then, the PHB content of the non-soluble fraction was determined enzymatically (Stage 2). Both stages are described in this section followed by an estimation of the sample variance for this method.

Stage 1: Isolation of non-soluble molecules

During sampling, 40 mL cell broth were filled into Falcon tubes and centrifuged for 20 min at 6000 rpm (4 °C) to remove supernatant. The duplicate samples were stored in a freezer at -80 °C until analyzing.
After defrosting and for cell disruption, 15–20 mL sodium-hypochlorite (14 %) were added into the Falcons. For a better mixing, the suspension was transferred into beakers and stirred for 3.5–4 h until the disruption was complete. Then, to dilute the sodium-hypochlorite, 60 mL of distilled water were added and the cell debris suspensions were centrifuged for 15 min at 7500 rpm and 4 °C in fresh Falcon tubes. After discarding the supernatant and another washing step with water, the white to yellowish pellets were resuspended in 5–20 mL destilled water and then transferred into preweighed Petri dishes. When all water had evaporated, the Petri dishes were reweighed to determine the mass of the non-soluble fraction. To quantify its PHB percentage, an enzymatic analysis followed in Stage 2.

Stage 2: PHB detection in the non-soluble fraction

In this stage, PHB was hydrolytically cleaved into hydroxybutyric acid monomers and then quantified by an enzymatic reaction. For this sake, 5–10 mg of the dry pellet were transferred into Eppendorf tubes. To cleave the polymer bonds, 150 µL NaOH (1 molar) and 150 µL H_2O were added. To speed up the cleavage reaction, it was incubated at 85 °C and shaken from time to time until the suspension had become totally clear. Then, it was diluted with 100 µL water in order to partly neutralize the pH. A total neutralization using acid was not favored, because it would have led to flocculation of PHB and interfere with the following analysis steps. The cleavage mix was diluted tenfold before starting the enzymatic test. A photometric test cell was filled with 880 µL Tris buffer of pH 8, 100 µL NAD^+ solution of 33 mM and 10 µL of the diluted cleavage mix. A blank extinction was measured at 340 nm before starting the enzymatic reaction. The enzyme β-hydroxybutyrate-dehydrogenase oxidizes mono-hydroxybuturate to ketobutyrate by reducing the co-factor NAD^+ to photometrically detectable NADH. After initiation by adding 10 µL of enzyme solution, the test cell was closed, shaken and incubated at 30 °C for 2–3 h until no more change in extinction was notable. Employing the law of Lambert-Beer,

$$E = c^{\mathrm{n}}_{\mathrm{NADH}} \cdot \epsilon \cdot d, \tag{3.11}$$

with E being the difference of measured blank and final extinctions, d being a test cell diameter of 1 cm and ϵ of 6200 $\mathrm{L\,mol^{-1}\,cm^{-1}}$ the extinction coefficient for NADH, the molar concentration of NADH was calculated. In this thesis, molar and mass concentrations are both symbolized by c, but the molar ones are tagged with the superscript "n". Per molecule of mono-hydroxybutyrate, one molecule of ketobutyrate and of NADH were produced. Thus, the calculated molar concentration of NADH directly equaled that one of mono-hydroxybuturate. Multiplication with the molecular weight of mono-hydroxybuturate (HB) resulted in a mass concentration

$$c_{\mathrm{HB}} = \frac{E}{\epsilon \cdot d} \cdot M_{\mathrm{HB}}. \tag{3.12}$$

Keeping track of weights and volumes used for hydrolysis and enzymatic reaction, the concentration of non-soluble cell fraction in the test cell was known. Employing eq. (3.12), the percentage of HB in the non-soluble fraction of the test cell could be calculated. By relating it to the total measured mass of non-soluble compounds (Stage 1), the amount of PHB in the drawn samples was finally calculated.

Measurement variance of PHB

As described in the previous paragraphs, PHB was usually analyzed in duplicates. Since the PHB analysis is costly and very time consuming, a sextuplicate analysis was done only for selected samples. For each duplicate and sextuplicate the mean value and sample variance were calculated and related to each other according to the procedure described in Section 3. Approximation results are visualized in Figure 3.5 and identified parameters listed in Table 3.3.

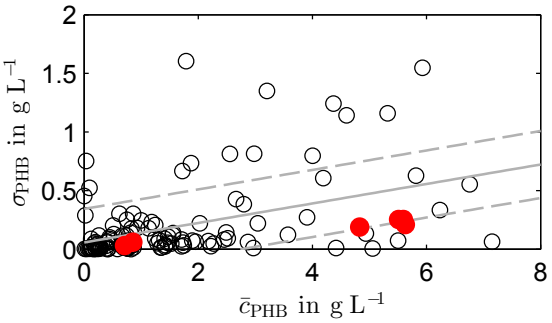


Figure 3.5: Measurement errors of PHB plotted over the mean values (duplicates in black, six-fold analysis in red) and the linear regression (solid line) with confidence intervals (broken lines)

Parameter	Value	Unit
$\sigma_{0,\mathrm{PHB}}$	0.0575	g L^{-1}
$\theta_{\sigma,\mathrm{PHB}}$	0.0832	-
error$_{\phi,\mathrm{PHB}}$	0.2862	g L^{-1}

Table 3.3: Maximum likelihood estimation results for a linear measurement noise approximation of PHB visualized in Figure 3.5

3.4 Quantification of extracellular metabolites

Analysis of some cultivations led to the hypothesis that extracellular metabolites were produced in significant amounts. In order to investigate whether extracellular metabolites were formed, some sample filtrates were analyzed for proteins and polysaccharides. Proteins were quantified with the BCA protein test kit in microtiter plates. Polysaccharides were analyzed according to the method described in Dubois et al. (1956). This analysis was carried out in glass tubes (10 mL) under the fume hood. Glucose solutions with concentrations up to 200 mg L^{-1} served as standards for calibration. During the development phase of the method in the laboratory, it turned out that the test result strongly depends on the total volume applied. The bigger the total volume, the more accurate the results. The analysis included the following steps:

1. A sample volume of 400 µL was pipetted into a test tube.

2. Then, 400 µL 5 % (w/w) phenol were added.
3. After sealing the test tube to prevent the evaporation of phenol, the content was mixed.
4. To start the reaction, 2 mL of concentrated sulphuric acid were added and the tube was sealed again before vortexing.
5. After 10 minutes, the tube content was mixed again.
6. The reaction was complete after 30 minutes. Then, the extinctions were measured photometrically in micro-cuvettes at 490 nm.

Each type of polysaccharide requires its own calibration because the type of molecule determines the number of reduction groups that bind to phenol and sulfuric acid and thus become photometrically active. As it was not known what kind of extracellular polysaccharides were built in the course of an autotrophic *R. e.* cultivation, the calibration was done with glucose. Hence, the measured concentrations did not correspond to the actual polysaccharid values. However, they give qualitative indications as needed for testing the hypothesis. The extracellular metabolites (sum of proteins and polysaccharides) measured at random in the cultivations corresponded to values of less than 1 % of the biomass. Due to the small values they seem to have little influence on the model and cannot be the main reason for estimated $k_L a_{CO_2}$ values that disagree the predictions of the film theory, see Section 5.1. Hence, according to these measurements the stated hypothesis could be ruled out.

3.5 Additional measurements

In addition to the manual measurement methods presented in this chapter, certain quantities were measured online. The supplied gas flows, dissolved carbon dioxide and dissolved oxygen, the fed volume of base and acid were recorded with appropriate sensors. Further information on the online measurement methods can be found in Rossner (2014).

During cultivation, it is useful to calculate the necessary and constantly changing gas flows with a controller. For this purpose, the next chapter presents a closed-loop process control setup that allows gas controlled cultivation with new strains without prior biological information in non-full control mode. But even if a process model is available for the new strain and the process runs in full-control mode, strain-independent gas control is preferable to gas control based on process models because mechanistic biological models tend to be inaccurate as they only provide very rough approximations of the metabolism.

Chapter 4

Cultivation system

Autotrophic cultivation requires a special bioreactor setup described in Section 4.1. An overview of medium composition and liquid feedings is given in Section 4.2. When cultivating, several controllers enable the tracking of setpoints and trajectories. An appropriate control architecture suitable for optimizations of the autotrophic process is explained in Section 4.3. An integral part of closed-loop process control is the so-called gas phase controller, which is introduced in Section 4.4. This controller does not rely on metabolic information, and thus works for all autotrophic strains. The gas phase controller can be run as a part of the overall closed-loop process control system but also independently once the user specifies reference trajectories.

4.1 Process setup

R. e. was cultivated autotrophically. Autotrophic growth means that gaseous feedings are required besides liquid salt feed flows. A gas phase controller to be developed in Section 4.4 calculated the required volume flows of H_2, CO_2 and O_2 to track the desired gas phase composition trajectories and an excess pressure in the headspace of the reactor, which was set to 40 mbar. In this study, an excess pressure was required to prevent ambient air from entering the vessel. Nitrogen as a component of the air would contaminate the system. Moreover, the excess pressure reduced the risk of surrounding organisms entering the system and contaminating the cultivation. The gaseous feed flows were realized by mass flow controllers FCR13-15 type El-flow F201CV from Bronkhorst (Netherlands), see Figure 4.1. Although the gases entered the headspace through an overpressure, the transported volumes were calculated by the mass flow controllers directly for ambient pressure, which will be important when reformulating the volume flows into mass flow quantities. Since the organisms grew inside the liquid phase, the gas mix of the gaseous phase had to be transferred into the medium. To this end, a compressor type N726FT.29E manufactured by KNF Neuberger

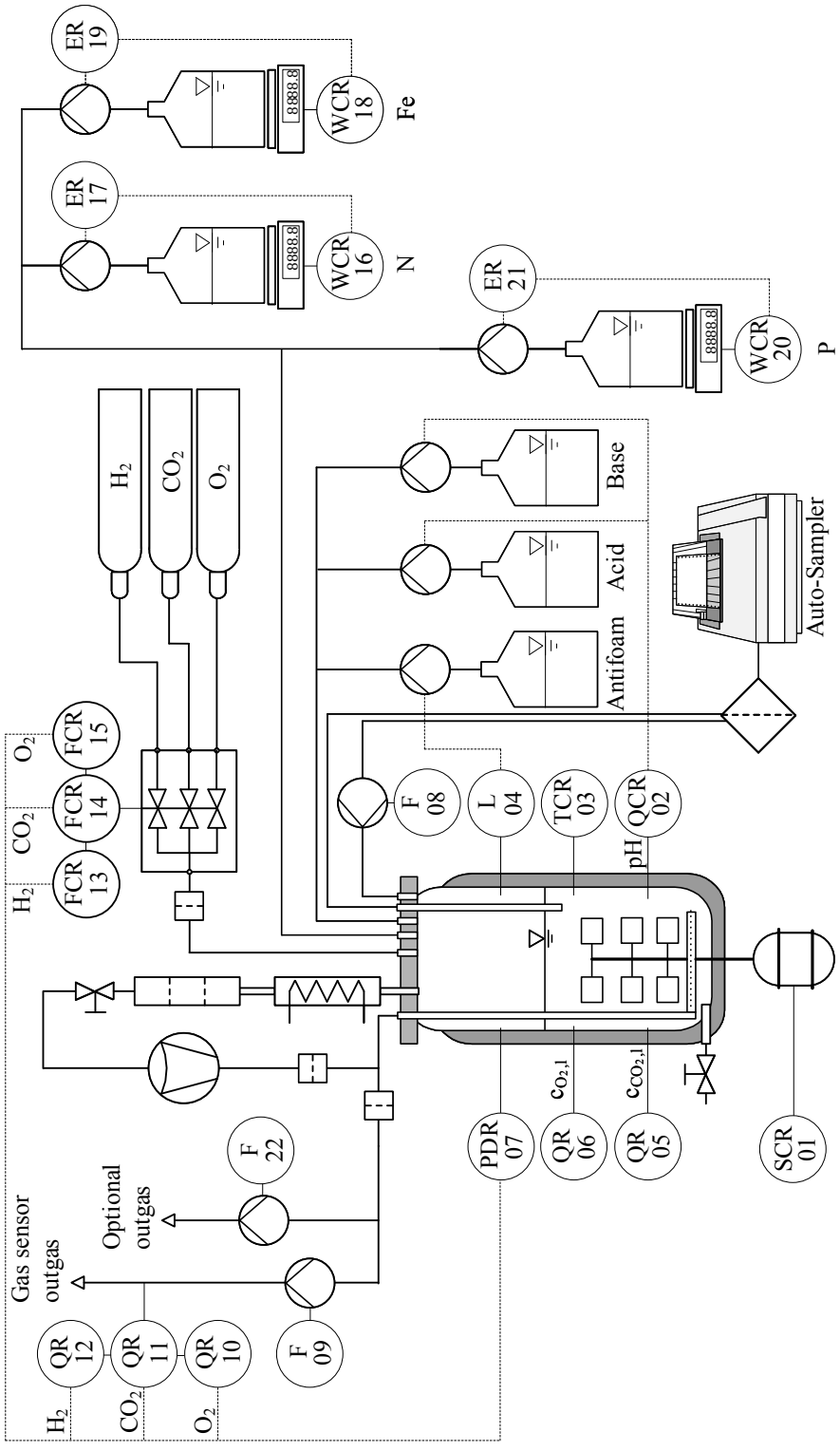


Figure 4.1: Setup scheme for the autotrophic cultivation of *R. e.*

(Germany) sucked the cooled and filtered gas from the headspace and blew it through a

microfilter and a sparger into the liquid phase at a rate of about 7 $\mathrm{L\,min^{-1}}$. Additional stirring (SCR01) by three Rushton turbines at 500 rpm dispersed the gas bubbles and led to better gas transfer. A hose pump F09 removed a defined outgas flow of 6.5 $\mathrm{L\,h^{-1}}$ from the gas cycle and led it to the gas sensors QR10-12 for analyzing the gas composition with a BlueInOne Cell and BCP-H2 from BlueSens (Germany). Measured gas data together with the pressure values PDR07 monitored by El-Press P502C from Bronkhorst were passed to the gas phase controller as in Section 4.4. In selected experiments, an additional variable outgas flow was pumped (F22) out of the gas cycle for an improved adjustment of the gas phase at low microbial gas consumption rates.

Besides gaseous feedings, *R. e.* consumes ammonium (N), phosphate (P), iron and trace elements, summarized as Fe in Figure 4.1, which are dosed as liquids by weight controlled hose pumps WCR/ER16-21 connected to the Multiple Fermenter Control Software (MFCS) that communicates with the Digital Control Unit (DCU) type Biostat® from B.Braun (Germany). Employed balances were BP6100 from Sartorius (Germany) connected to the pumps 101U/R from Watson Marlow (Great Britain) for weight controlled feeding. Depending on the purpose of each cultivation, the set trajectories for ammonium and phosphate feed flows were varied to study metabolic responses. Since iron and trace elements were not measured, a sufficient supply had to be ensured by coupling the feedrate to the one of ammonium in most cultivations.

When the cells assimilate ammonium, a proton is set free. To avoid pH changes due to ammonium assimilation or due to dissolved CO_2 fluctuations, the pH was controlled (QCR02) at a value of 6.8 being the setpoint. A PI control algorithm of the DCU type Biostat® from B.Braun (Germany) was connected to an inline pH probe InPro3100 from Mettler Toledo (Germany) and realized the desired setpoint. The pH control calculated required volumes of alkaline 3 normal sodium hydroxide and 3 normal sulfuric acid. Meaning that in each liter 3 moles of the active groups, i.e., proton for acid and hydroxide ion for base, were dissolved. Base and acid were fed via the hose pumps of the DCU and the fed amounts were recorded. A level control L04 fed an antifoam agent P2000 to prevent foam bubbles entering the compressor while the antifoam sensor IFA antifoam D19 from BiOENGiNEERiNG (Switzerland) sent a signal.

Besides recording volumes of fed acid, base and antifoam, the excess pressure and gas composition, as well as dissolved O_2 and dissolved CO_2 were registered every 5 seconds and used as measurements in the modeled system. The sensors InPro6820 and InPro5000i from Mettler Toledo measured dissolved CO_2 ($c_{\mathrm{CO_2,l}}$) and O_2 ($c_{\mathrm{CO_2,l}}$) that was displayed as dissolved oxygen partial pressure $p_{\mathrm{O_2,l}}$, respectively. The temperature setpoint of 30 °C was maintained by a PI controller TCR03. Together with pH and stirring speed the temperature was recorded but these values only served for monitoring not for modeling.

An automated sampling unit, which sampled cultivation filtrate at most every 15 min, was connected to a pumped (F08) bypass. It was analyzed manually for the concentrations of

phosphate, ammonium and extracellular metabolites. The same compounds were quantified in the manual samples taken from a valve located at the bottom of the 16 L vessel from BiOENGiNEERiNG. Additionally, the manual samples were investigated for biomass concentration, the amount of PHB, MBH and SH activities. Further information about the reactor and its peripheral setup can be found in Rossner (2014).

4.2 Feedings and medium

Compositions of liquid feedings are given in Table 4.1. The initial volume of the cultivation was 10 L of defined medium including inoculum unless otherwise stated. The initial medium

Table 4.1: Feedings composition for *R. e.* cultivations

Feeding	Chemical	Concentration [g·L^{-1}]
Phosphate (P)	Na_2HPO_4 ($2H_2O$)	39
	KH_2PO_4	13
Ammonium (N)	NH_4Cl	148
Iron and trace elements (Fe)	$MgSO_4$ ($7H_2O$)	20
	$CaCl_2$ ($2H_2O$)	1
	$FeCl_3$ ($6H_2O$)	1
	$NiCl_2$ ($6H_2O$)	0.024
	$ZnSO_4$ ($7H_2O$)	0.01
	$MnCl_2$ ($2H_2O$)	0.003
	H_3BO_3	0.03
	$CuCl_2$ ($2H_2O$)	0.001
	Na_2MoO_4 ($2H_2O$)	0.003
	$CoCl_2$ ($6H_2O$)	0.02

concentrations of iron and trace elements were 1/100 of those in the feeding. The ammonium concentration was 1/50 of the feeding concentration and that of phosphate was 1/10 or 1/30 (depending on the cultivation). For H798, 0.01 g L^{-1} and 0.09 g L^{-1} tetracycline was added to the initial medium and ammonium feeding, respectively.

4.3 Closed-loop control scheme

The mathematical model to be introduced in Section 4.4.1 for the gas phase controller like all mathematical models in Chapters 5–7 relate state variables $\underline{x}(t)$ with inputs $\underline{u}(t)$ and

measurements $\underline{y}(t)$. In the general case, a dynamic model

$$\underline{\dot{x}}(t) = f\left(\underline{x}(t), \underline{u}(t), \underline{\theta}\right) \quad , \underline{x}(0) = \underline{x}_0 \tag{4.1}$$

$$\underline{y}(t) = \underline{h}\left(\underline{x}(t), \underline{u}(t), \underline{\theta}\right) \tag{4.2}$$

results, where $\underline{x}(t)$ represents the system and model parameters are given by $\underline{\theta}$. Different state vectors will be considered in this thesis for different models, e.g., $\underline{x}_{\text{gas model}}$ for the gas phase model given below. The control inputs $\underline{u}(t)$ were calculated by different controllers for cultivations that ran in full-control process mode which will be explained here. The proposed, overall closed-loop control scheme was already published in Neddermeyer and King (2019), but to keep the thesis self-contained it is given here again.
Since *R. e.* was grown autotrophically, the system's input vector $\underline{u}$ in eq. (4.3) encompasses flow rates for liquid substrate feedings and trace elements (u_{N}, u_{Fe}, u_{P}), feed flows for the gaseous substrates ($q_{\text{H}_2,\text{v}}$, $q_{\text{CO}_2,\text{v}}$, $q_{\text{O}_2,\text{v}}$), summarized outgas flows ($q_{\text{leak,v}}$) and correction fluid flows for pH and antifoam control (u_{base}, u_{acid}, u_{antifoam}).

$$\underline{u}^T_{\text{system}} = (u_{\text{N}} \quad u_{\text{Fe}} \quad u_{\text{P}} \quad q_{\text{H}_2,\text{v}} \quad q_{\text{CO}_2,\text{v}} \quad q_{\text{O}_2,\text{v}} \quad q_{\text{leak,v}} \quad u_{\text{base}} \quad u_{\text{acid}} \quad u_{\text{antifoam}}) \tag{4.3}$$

Instead of a monolithic controller, sub-controllers were used to perform specific tasks. Except for $q_{\text{leak,v}}$, in full-control process mode, all entries of the input vector $\underline{u}_{\text{system}}$ were calculated by appropriate controllers united in the overall process control scheme that is depicted in Figure 4.2. Detailed information about signal adjustments among the different controllers and process units can be found in the Appendix B. The correction fluid flows were calculated by standard PI controllers located in the DCU of the bioreactor and the liquid reference feed flows for the substrates based on given references were realized by PI controllers of the Multiple Fermenter Control System (MFCS) via the pump/balance systems. MFCS is a basic control software that comes together with the biorector. Fed volumes were quantified and considered in the employed models. All volumetric gaseous feed flows ($q_{\text{H}_2,\text{v}}$, $q_{\text{CO}_2,\text{v}}$, $q_{\text{O}_2,\text{v}}$) were calculated by the PI-MIMO controller and a feedforward disturbance rejection (FFDR) as in Section 4.4, in order to maintain the desired excess pressure $r_{\Delta\text{P}}$ and the desired gas fractions ($r_{\text{H}_2,\text{v}}$, $r_{\text{CO}_2,\text{v}}$, $r_{\text{O}_2,\text{v}}$). Trajectories of the latter and liquid reference feed flows (r_{N}, r_{P}, r_{Fe}) were either designed by the operator when no general process model (I), which will be introduced in Section 5.2, was available. Or, the trajectories resulted from model-based optimizations done beforehand, and from closed-loop control active during the cultivation using a strain-dependent general process model. Employing the same model, the state vector ($\underline{x}_{\text{model I}}$) was estimated (subscript "est") by a Sigma Point Kalman filter (SPKF) or an Extended Kalman filter (EKF) with a dynamic system noise matrix (DEKF) as suggested by Schneider and Georgakis (2013). To this end, optical density (OD), dis-

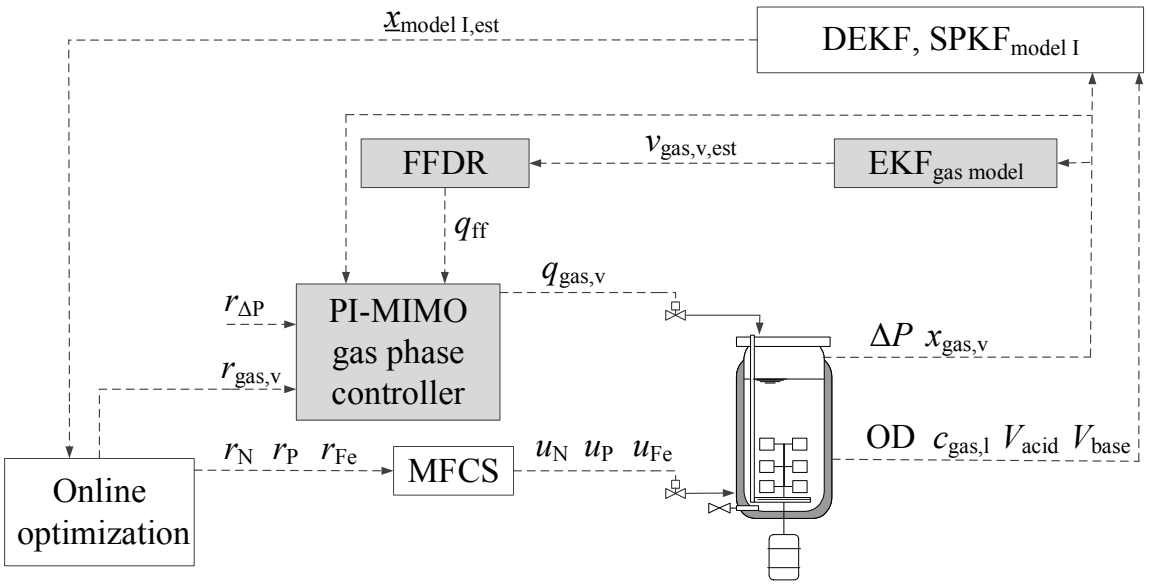


Figure 4.2: Overall control scheme for the autotrophic cultivation of *R. e.* The measurements $p_{\text{O}_2,\text{l}}$ and $c_{\text{CO}_2,\text{l}}$ are summarized as $c_{\text{gas,l}}$. Basic closed-loop controller for pH, foam level and temperature are independent and not depicted.

solved gas concentrations ($c_{\text{gas,l}}$), fed amounts of base and acid (V_{base}, V_{acid}), gas fractions in the headspace $x_{\text{gas,v}}$ and excess pressure ΔP served as online or atline measurements for state estimation. For the closed-loop control of the gaseous substrates different setups were proposed and tested in experiments from this thesis. Here, only the most successful setup of the gas phase control layer (gray) is given in the next Section 4.4, which works independent of biologic information and therefore serves the purpose of model adaption. Additionally, results utilizing this controller are given as well. The main idea proposed here is to estimate the gas transfer rates from the gaseous into the liquid phase and use these estimations for control. Estimations will be done as well with an EKF.

4.4 Model-based FFDR PI-MIMO gas phase control

For the sake of completeness, the gas phase control including a nonlinear headspace model, gas phase control laws, Extended Kalman filter parameter and gas control performance, which we already published in Neddermeyer and King (2019), are repeated here.
Measuring and controlling all dissolved gases H_2, CO_2 and O_2 directly is impossible since inline sensors for hydrogen are unavailable as discussed below in Section 5.3. Even if dissolved hydrogen were measurable, planned trajectories for the three gas compounds often would not be realizable due to an initially unknown gas consumption by the bacterium.
Instead of controlling the dissolved gases, a gas phase control is suggested that realizes the

desired gas composition as well as a constant excess pressure in the headspace. Ideally, four variables of the gaseous phase should be controlled, but only three manipulating variables are available as gaseous feed flows. The controller to be developed has to compensate for harsh external disturbances such as pressure drops due to sampling, internal disturbances like nonlinear drifts in microbial gas consumption rates and quick metabolic changes due to nutrient limitations. Because of the system's nonlinearity and its different time constants, controller concepts for linear systems fail and robust approaches are too slow for reference tracking when fast disturbances occur. Instead, a newly developed smart feedforward disturbance rejection (FFDR), which is independent of the metabolic behavior of the cells, and hence can be applied for all kind of autotrophic gas cultivations, is proposed here combined with a double loop PI controller.
In the present system, gas fractions of the headspace together with pressure are adjusted in two serial PI control loops, $\mathrm{CL}_{\Delta\mathrm{P}}$ and $\mathrm{CL}_{\mathrm{gas}}$, as originally suggested by Rossner (2014). Note, that the subscript "gas" in $\mathrm{CL}_{\mathrm{gas}}$ represents hydrogen, carbon dioxide and oxygen in what follows. Each individual feed flow can be expressed as

$$q_{\mathrm{gas,v}} = x_{\mathrm{gas,v,feed}} \cdot q_{\mathrm{t,v}}, \tag{4.4}$$

where $x_{\mathrm{gas,v,feed}}$ represents the gas fraction of a specific component in the total feed flow $q_{\mathrm{t,v}}$, the calculation of which is shown in Section 4.4.2. The total feed flow is used for maintaining a constant excess pressure and is manipulated by $\mathrm{CL}_{\Delta\mathrm{P}}$. The former is obtained via $\mathrm{CL}_{\mathrm{gas}}$ for the specific component "gas". As the time constant of the headspace changes by a factor of almost 50 during the course of cultivation, this is taken into account by the feedforward disturbance rejection. To calculate the FFDR, the consumption rates ν of the gases have to be known, which are estimated with a gas phase model described in Section 4.4.1, before Section 4.4.2 outlines the controllers. Tuning parameters of the EKF to estimate the consumption rates are provided in Section 4.4.3 and performance results are presented in Section 4.4.4.

4.4.1 Nonlinear gas phase model

Balancing the reactor headspace, a nonlinear model was formulated with the aim to estimate the gas consumption rates and forward them in terms of a disturbance rejection. Figure 4.3 illustrates major state and manipulating variables of the gas phase model. It consists of nine state variables: system pressure P and amount of material in the headspace of the three gas compounds $n_{\mathrm{gas,v}}$. A fourth gas amount, $n_{\mathrm{rest,v}}$, compensates for the remaining nitrogen and evaporated water in the gas phase. The amount of nitrogen might increase due to incoming air in periods of underpressure. Three gas consumption rates, $\nu_{\mathrm{gas,v}}$, mainly describe gas flows across the liquid-gas interface, which is indicated by an arrow in Figure 4.3, but also

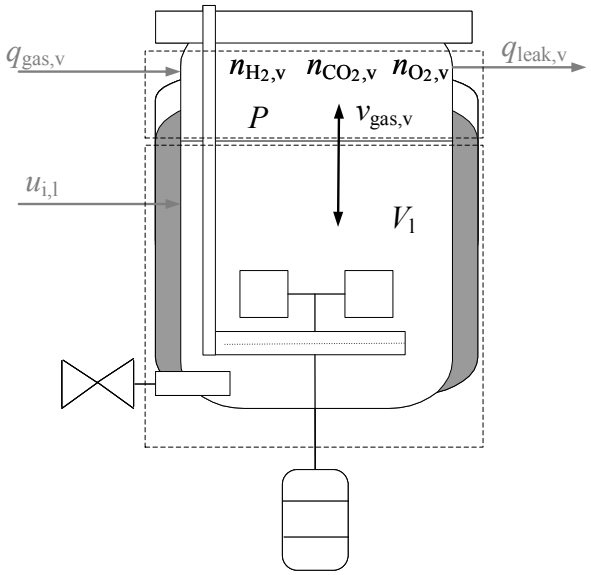


Figure 4.3: Scheme of the gas phase model derived from two connected systems. Both system boundaries are indicated with broken lines. States are shown in black and manipulating variables in gray.

compensate for sudden pressure drops due to sampling. These consumption rates were included as states as well. Instead of balancing, it was assumed that these states change slowly over time, i.e., $\dot{\nu}_{\mathrm{gas}}$ will be equated to zero below in the so-called undisturbed model. Since the headspace volume (V_{head}) changes due to liquid feedings $u_{i,\mathrm{l}}$, the liquid volume (V_{l}) has to be considered as well. This results in the state vector

$$\underline{x}^{T}_{\mathrm{gas\ model}} = (P \quad n_{\mathrm{H_2,v}} \quad n_{\mathrm{CO_2,v}} \quad n_{\mathrm{O_2,v}} \quad n_{\mathrm{rest,v}} \quad \nu_{\mathrm{H_2,v}} \; \nu_{\mathrm{CO_2,v}} \quad \nu_{\mathrm{O_2,v}} \quad V_{\mathrm{l}}), \tag{4.5}$$

with V_{l} being the only liquid state. The system's manipulating vector of eq. (4.3), which combines liquid flows $u_{i,\mathrm{l}}$, gaseous inflows and outflows $q_{\mathrm{gas,v}}$ and $q_{\mathrm{leak,v}}$, equals the input vector of this gas phase model. All volume flows are given in $\mathrm{L\,h^{-1}}$ so they must be multiplied by 10^{-3} to comply with SI units. When reformulating $q_{\mathrm{gas,v}}$ as molar flows applying the ideal gas law

$$\dot{n}_{\mathrm{gas,v,feed}} = P \cdot \frac{q_{\mathrm{gas,v}} \cdot 10^{-3}}{R \cdot T} \tag{4.6}$$

and considering the leakage gas flow ($q_{\mathrm{leak,v}}$)

$$\dot{n}_{\mathrm{leak,v}} = \frac{P}{R \cdot T} \cdot q_{\mathrm{leak,v}} \cdot 10^{-3} \tag{4.7}$$

the undisturbed dynamic state system given in SI units, with the exception V_l expressed in L, reads

$$\dot{P} = \frac{\sum \dot{n}_\mathrm{gas,v} \cdot R \cdot T}{V_\mathrm{head} \cdot 10^{-3}} \tag{4.8}$$

$$\dot{n}_\mathrm{H_2,v} = \dot{n}_\mathrm{H_2,v,feed} - \nu_\mathrm{H_2,v} - \dot{n}_\mathrm{leak,v} \cdot \frac{n_\mathrm{H_2,v}}{P \cdot V_\mathrm{head} \cdot 10^{-3}} \cdot R \cdot T \tag{4.9}$$

$$\dot{n}_\mathrm{CO_2,v} = \dot{n}_\mathrm{CO_2,v,feed} - \nu_\mathrm{CO_2,v} - \dot{n}_\mathrm{leak,v} \cdot \left(1 - \frac{(n_\mathrm{rest,v} + n_\mathrm{H_2,v} + n_\mathrm{O_2,v}) \cdot R \cdot T}{P \cdot V_\mathrm{head} \cdot 10^{-3}}\right) \tag{4.10}$$

$$\dot{n}_\mathrm{O_2,v} = \dot{n}_\mathrm{O_2,v,feed} - \nu_\mathrm{O_2,v} - \dot{n}_\mathrm{leak,v} \cdot \frac{n_\mathrm{O_2,v}}{P \cdot V_\mathrm{head} \cdot 10^{-3}} \cdot R \cdot T \tag{4.11}$$

$$\dot{n}_\mathrm{rest,v} = 0 \tag{4.12}$$

$$\dot{\nu}_\mathrm{H_2,v} = 0 \tag{4.13}$$

$$\dot{\nu}_\mathrm{CO_2,v} = 0 \tag{4.14}$$

$$\dot{\nu}_\mathrm{O_2,v} = 0 \tag{4.15}$$

$$\dot{V}_\mathrm{l} = u_\mathrm{N} + u_\mathrm{Fe} + u_\mathrm{P} + u_\mathrm{antifoam} + u_\mathrm{acid} + u_\mathrm{base} - u_\mathrm{sampling}. \tag{4.16}$$

The variable V_head is

$$V_\mathrm{head} = V_\mathrm{reactor} - V_\mathrm{l}, \tag{4.17}$$

with a total reactor volume of $V_\mathrm{reactor} = 16$ L, including vessel and tubing.
Since all gas phase compounds are measured as gas fractions ($x_\mathrm{gas,v}$) and $\Delta P = P - P_0$ is monitored as excess pressure (mbar), the measurement vector of the gas phase model

$$\underline{y}^T_\mathrm{gas\ model} = (y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5) \tag{4.18}$$

as a function of states yields

$$y_1 = y_{\Delta P} = \frac{P - P_0}{100} \tag{4.19}$$

$$y_2 = x_\mathrm{H_2,v} = n_\mathrm{H_2,v} \cdot \frac{R \cdot T}{P \cdot V_\mathrm{head} \cdot 10^{-3}} \tag{4.20}$$

$$y_3 = x_\mathrm{CO_2,v} = n_\mathrm{CO_2,v} \cdot \frac{R \cdot T}{P \cdot V_\mathrm{head} \cdot 10^{-3}} \tag{4.21}$$

$$y_4 = x_\mathrm{O_2,v} = n_\mathrm{O_2,v} \cdot \frac{R \cdot T}{P \cdot V_\mathrm{head} \cdot 10^{-3}} \tag{4.22}$$

$$y_5 = V_\mathrm{l}. \tag{4.23}$$

It is assumed that in the process major changes of the liquid volume are caused by correction fluids and feedings. Thus, an online calculated volume serves as synthetic measurement y_5 whereas y_1 to y_4 are detected every five seconds by the sensors introduced in Section 4.1. To apply a Kalman filter, all equations are assumed to be extended by appropriate noise terms resulting in the disturbed system model. By this, the gas consumption rates $\nu_{\mathrm{gas,v}}$, which change drastically over the course of cultivation, can be estimated (subscript "est"). This information will be used to determine the FFDR. Reusing the ideal gas law leads to the desired feedforward part of each individual gas component.

$$q_{\mathrm{gas,ff}} = q_{\mathrm{gas,v,est}} + q_{\mathrm{gas,v,leak}} \tag{4.24}$$

$$= \nu_{\mathrm{gas,v,est}} \cdot \frac{R \cdot T}{10^{-3} \cdot P} + x_{\mathrm{gas,v,est}} \cdot q_{\mathrm{leak,v}}. \tag{4.25}$$

The overall feedforward gas flow is

$$q_{\mathrm{ff}} = \frac{R \cdot T}{10^{-3} \cdot P} \cdot \sum \nu_{\mathrm{gas,v,est}} + q_{\mathrm{leak,v}}. \tag{4.26}$$

4.4.2 Gas phase control laws

Ideally, in the present cultivation system, three manipulating variables $q_{\mathrm{gas,v}}$, see eq. (4.3), serve to adjust four control variables, which are the gas fractions $x_{\mathrm{gas,v}}$ and the excess pressure ΔP. This is impossible in practice because the problem is underactuated. Hence, one gas compound has to remain uncontrolled, which can be selected by choosing an appropriate mode. For this work, due to the large solubility of CO_2, the gas fractions of hydrogen and oxygen ($x_{\mathrm{H_2,v}}$ and $x_{\mathrm{O_2,v}}$) were controlled in parallel PI loops ($\mathrm{CL_{H_2}}$ and $\mathrm{CL_{O_2}}$) that together represent the first part of the two-stage gas phase controller. In Figure 4.4, both control loops (gas content in white and pressure in gray) and their interface are shown schematically. Manipulating variables $q_{\mathrm{gas,v}}$ are calculated in two stages. In stage two, the required total gas flow ($q_{\mathrm{t,v}}$) needed to maintain the desired excess pressure is calculated by a PI controller ($\mathrm{CL_{\Delta P}}$) with FFDR-modulated gain scheduling and feedforward disturbance rejection. The total gas flow $q_{\mathrm{t,v}}$ thus consists of the feedforward $\mathrm{q_{ff}}$ and a closed-loop part q_{c}.

$$q_{\mathrm{t,v}} = \underbrace{q_{\mathrm{ff}}}_{\mathrm{FFDR}} + q_{\mathrm{c}}. \tag{4.27}$$

Combining the equations of the headspace model with eq. (4.27) leads to the dynamic evolution of the excess pressure

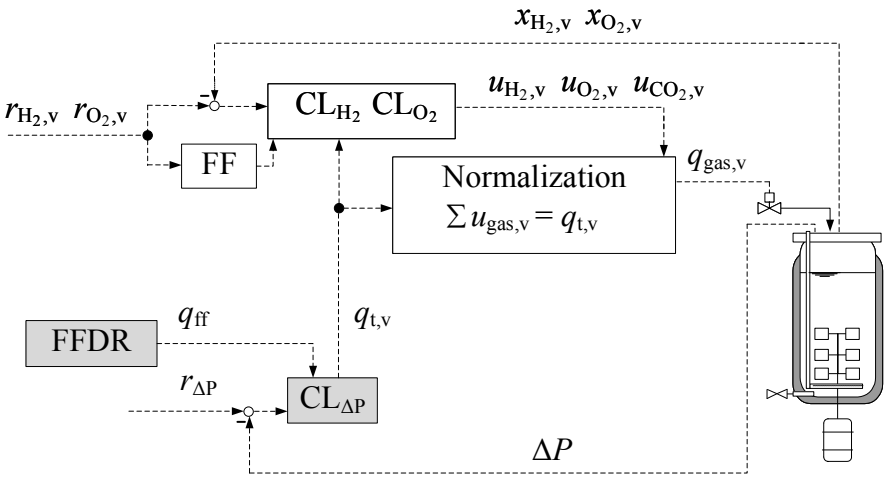


Figure 4.4: Block diagram of a serial PI-MIMO gas phase controller for the adjustment of pressure, hydrogen and oxygen in the gas phase. Broken and solid lines represent data and mass transport, respectively.

$$\Delta\dot{P} = \frac{q_c}{V_{head}} \cdot \Delta P + \frac{q_c}{V_{head}} \cdot P_0 + \underbrace{\frac{R \cdot T}{10^{-3} \cdot V_{head}} \left(\sum \nu_{gas,v,est} - \sum \nu_{gas,v} \right)}_{z}, \tag{4.28}$$

where P_0, $\nu_{gas,v,est}$ and z represent the ambient pressure, the estimated rates of gas transfer, which will be obtained below, and a remaining term that is interpreted as a disturbance, respectively. In a more standard-control oriented notation with ΔP, q_c, V_{head}^{-1} and $V_{head}^{-1} P_0$ being x, u, a and b, respectively,

$$\dot{x}(t) = a \cdot u(t) \cdot x(t) + b \cdot u(t) + z(t), \tag{4.29}$$

the nonlinearity of the process is obvious. Now, instead of a more complicated nonlinear approach, a simple control law is suggested. To guarantee exact setpoint tracking for constant reference values of the excess pressure $r_{\Delta P}$, given in mbar, integral action will be necessary and thus a PI controller is selected. For fixed proportional and integral gains, two fixed poles of the linearized system in the left half plane result. The dynamic behavior of the plant can be characterized by a time constant

$$T_{\Delta P} = \frac{V_{head}}{q_{ff}} \tag{4.30}$$

that approximates the residence time in the headspace. As this time constant significantly decreases in the course of cultivation, it turned out to be beneficial when the closed-loop bandwidth was increased accordingly. This can be achieved by multiplying the integral part with q_{ff}. As a result, the pole of the linearized closed-loop system nearer to the imaginary axis is shifted to the left. Therefore, finally, the control law reads

$$q_{\mathrm{c}} = K_{\Delta\mathrm{P},P} \cdot e_{\Delta\mathrm{P}} + K_{\Delta\mathrm{P},I} \cdot q_{\mathrm{ff}} \cdot e_{\Delta\mathrm{P,int}}, \tag{4.31}$$

with the pressure related control error

$$e_{\Delta\mathrm{P}} = r_{\Delta\mathrm{P}} - \Delta P. \tag{4.32}$$

Using the trapezoidal rule for integration yields the integrated error according to

$$e_{\Delta\mathrm{P,int}} = e_{\Delta\mathrm{P,int}} + \frac{(e_{\Delta\mathrm{P}} + e_{\Delta\mathrm{P,prior}})}{2} dt, \tag{4.33}$$

with dt being the sampling time and $e_{\Delta\mathrm{P,prior}}$ the control error of the previous time step. The total feed flow $q_{\mathrm{t,v}}$ has to be realized by the individual feeds

$$q_{\mathrm{t,v}} = q_{\mathrm{H_2,v}} + q_{\mathrm{CO_2,v}} + q_{\mathrm{O_2,v}}. \tag{4.34}$$

With the $\mathrm{CL_{gas}}$ controller only two of them are determined. An obvious solution such as

$$q_{\mathrm{CO_2,v}} = q_{\mathrm{t,v}} - q_{\mathrm{H_2,v}} - q_{\mathrm{O_2,v}} \tag{4.35}$$

will be ruled out below to allow for softer changes of the gas fractions, and thus prevent overshootings of $x_{\mathrm{CO_2,v}}$.
Starting points are reference values $r_{\mathrm{gas,v}}$ for the individual component fractions, which cannot be chosen independently but must obey

$$r_{\mathrm{H_2,v}} + r_{\mathrm{CO_2,v}} + r_{\mathrm{O_2,v}} + x_{\mathrm{H_2O,v}} = 100\ \%, \tag{4.36}$$

with $x_{\mathrm{H_2O,v}}$ being the constant gas fraction of evaporated water and assuming that nitrogen is absent. Simple PI controllers based on control errors

$$e_{\mathrm{gas}} = r_{\mathrm{gas,v}} - x_{\mathrm{gas,v}}, \qquad \mathrm{gas}\ \epsilon\ \{\mathrm{O_2}, \mathrm{H_2}\} \tag{4.37}$$

are used with tuned proportional and integral gains $K_{\mathrm{gas},P}$ and $K_{\mathrm{gas},I}$, respectively. The output $\Delta u_{\mathrm{gas,v}} = K_{\mathrm{gas},P} \cdot e_{\mathrm{gas}} + K_{\mathrm{gas},I} \cdot e_{\mathrm{gas,int}}$ of a controller multiplied by $q_{\mathrm{t,v}}$ represents a

designed feed flow that can be realized by a mass flow controller. An open-loop term based on the desired reference value is added to unburden the controller; i. e., the controller output reads

$$u_{\mathrm{H_2,v}} = \underbrace{r_{\mathrm{H_2,v}} \cdot q_{\mathrm{t,v}}}_{\text{open loop}} + \Delta u_{\mathrm{H_2,v}} \cdot q_{\mathrm{t,v}}, \tag{4.38}$$

$$u_{\mathrm{O_2,v}} = (r_{\mathrm{O_2,v}} + \Delta u_{\mathrm{O_2,v}}) \cdot q_{\mathrm{t,v}}. \tag{4.39}$$

The desired feed flow of CO_2 is written down in a similar fashion

$$u_{\mathrm{CO_2,v}} = (r_{\mathrm{CO_2,v}} + \Delta u_{\mathrm{CO_2,v}}) \cdot q_{\mathrm{t,v}}. \tag{4.40}$$

However, as pointed out above, the unknown part calculated according to

$$\Delta u_{\mathrm{CO_2,v}} = -(\Delta u_{\mathrm{H_2,v}} + \Delta u_{\mathrm{O_2,v}}) \tag{4.41}$$

led to unsatisfactory results in experiments performed because in the beginning of the cultivation large time constants of the process make control slow. As soon as one dissolved gas component is insufficient, growth is restricted, which leads to reduced substrate uptake and slows down the process respectively control even more. Since the gas components add up to 1, often one gas is insufficient and another is in excess. In order to accelerate the control again, the excess gas must first be consumed, which takes the longest for CO_2, as it is assimilated least by the organism. An excess of CO_2 therefore leads to slower control over a longer period of time and must be avoided. Therefore, the unknown part is reduced according to the fed CO_2 fraction of the prior sampling instant

$$\Delta u_{\mathrm{CO_2,v}} = -x^{\mathrm{prior}}_{\mathrm{CO_2,v,feed}} \cdot (\Delta u_{\mathrm{H_2,v}} + \Delta u_{\mathrm{O_2,v}}). \tag{4.42}$$

All feed rates $u_{\mathrm{gas,v}}$ are constrained from below zero, as negative flow rates cannot be realized. As with eq. (4.38)–(4.40) the desired overall flow $q_{\mathrm{t,v}}$ is not met, the normalized and actually applied feed streams finally read

$$q_{\mathrm{H_2,v}} = \frac{q_{\mathrm{t,v}}}{u_{\mathrm{H_2,v}} + u_{\mathrm{CO_2,v}} + u_{\mathrm{O_2,v}}} \cdot u_{\mathrm{H_2,v}} = x_{\mathrm{H_2,v,feed}} \cdot q_{\mathrm{t,v}} \tag{4.43}$$

$$q_{\mathrm{CO_2,v}} = \frac{q_{\mathrm{t,v}}}{u_{\mathrm{H_2,v}} + u_{\mathrm{CO_2,v}} + u_{\mathrm{O_2,v}}} \cdot u_{\mathrm{CO_2,v}} = x_{\mathrm{CO_2,v,feed}} \cdot q_{\mathrm{t,v}} \tag{4.44}$$

$$q_{\mathrm{O_2,v}} = \frac{q_{\mathrm{t,v}}}{u_{\mathrm{H_2,v}} + u_{\mathrm{CO_2,v}} + u_{\mathrm{O_2,v}}} \cdot u_{\mathrm{O_2,v}} = x_{\mathrm{O_2,v,feed}} \cdot q_{\mathrm{t,v}}. \tag{4.45}$$

4.4.3 Extended Kalman filter parameter

Here, it is assumed that the equations for an Extended Kalman Filter (EKF) are known. For further information, the reader is referred to, e.g., Gelb et al. (1974). In Section 4.4.1, a headspace gas model was introduced that enables FFDR by estimating the needed summarized gas flow ($\sum \nu_{\mathrm{gas,v}}$) in each timestep during the cultivation. To this end, the estimated gas amount rates of eq. (4.13) to (4.15) are transformed to volume flows

$$q_{\mathrm{gas,v,est}} = \nu_{\mathrm{gas,v,est}} \frac{R \cdot T}{P} \cdot \frac{1000\ \mathrm{L}}{\mathrm{m^3}} \tag{4.46}$$

and later inserted in eq. (4.24) on page 46 to calculate the FFDR. When using an EKF to estimate the gas consumption rates of eq. (4.46), three matrices have to be designed. Matrix $\mathbf{Q}$ describes the system's noise, matrix $\mathbf{P}_0$ gives the variance of the initial values and matrix $\mathbf{R}$ is the covariance matrix of the measurements. In the following, the matrix designs are explained. Off-diagonal entries of $\mathbf{Q}$ were set to zero and the diagonal elements ($\mathbf{Q}_{ii}$) of the EKF $\mathbf{Q}$-matrix were calculated empirically. To obtain the maximal slopes of the differential equations $\dot{x}_{i,\max}$, prior experiments were evaluated. It is assumed that a dynamical inaccuracy of max 10 % in the right hand side is described by the system noise. This leads to the $\mathbf{Q}$-matrix with squared values on the diagonal

$$\mathbf{Q}_{ii} = \left(\frac{\dot{x}_{i,\max}}{10}\right)^2 = \begin{pmatrix} 4.42 \cdot 10^{4}\ \mathrm{Pa^2} \\ 8.52 \cdot 10^{-6}\ \mathrm{mol^2} \\ 1.1 \cdot 10^{-5}\ \mathrm{mol^2} \\ 1.1 \cdot 10^{-5}\ \mathrm{mol^2} \\ 2.13 \cdot 10^{-5}\ \mathrm{mol^2} \\ 5.63 \cdot 10^{-2}\ \mathrm{m^6\,h^{-2}} \\ 7.7 \cdot 10^{-9}\ \mathrm{m^6\,h^{-2}} \\ 7.41 \cdot 10^{-3}\ \mathrm{m^6\,h^{-2}} \\ 1.69 \cdot 10^{-3}\ \mathrm{L^2} \end{pmatrix}. \tag{4.47}$$

Also for the initial states covariance matrix $\mathbf{P}_0$ all off-diagonal entries were set to zero. To calculate $\mathbf{P}_{0,ii}$, a relative error (e_i) of each initial state (subscript "0") was assumed, multiplied by the initial value and squared

$$\mathbf{P}_{0,ii} = (e_i \cdot x_{i,0})^2 = \begin{pmatrix} (0.1 \cdot P_0)^2 \\ (0.08 \cdot n_{0,\mathrm{H_2,v}})^2 \\ (0.08 \cdot n_{0,\mathrm{H_2,v}})^2 \\ (0.08 \cdot n_{0,\mathrm{O_2,v}})^2 \\ (0.05 \cdot n_{\mathrm{rest,v}})^2 \\ (0.3 \cdot \nu_{\mathrm{H_2,v}})^2 \\ (0.3 \cdot \nu_{\mathrm{H_2,v}})^2 \\ (0.3 \cdot \nu_{\mathrm{O_2,v}})^2 \\ (0.02 \cdot V_{0,\mathrm{l}})^2 \end{pmatrix}. \tag{4.48}$$

Known initial values were inserted, and for all unknown initial states, namely $n_{\mathrm{rest,v}}$ and $\nu_{\mathrm{gas,v}}$, a set of estimates derived from a prior experiment were taken:

$$n_{\mathrm{rest,v}} = 0.02 \text{ mol} \tag{4.49}$$

$$\nu_{\mathrm{H_2,v}} = 2.13 \text{ mol h}^{-1} \tag{4.50}$$

$$\nu_{\mathrm{CO_2,v}} = 0.28 \text{ mol h}^{-1} \tag{4.51}$$

$$\nu_{\mathrm{O_2,v}} = 0.71 \text{ mol h}^{-1}. \tag{4.52}$$

To build the covariance matrix of the measurement noise $\mathbf{R}$, relative (subscript "rel") sample errors were postulated since the linear gas sensors were calibrated by a one-point calibration. Slightly wrong calibrated slopes lead to smaller errors at lower concentrations than at higher ones. Due to the detection limit at very low concentrations a minimum (min) error is suggested (see Table 4.2) leading to

$$\mathbf{R}_{ii} = \left[\max(e_{\mathrm{min}}, e_{\mathrm{rel}})\right]^2. \tag{4.53}$$

The relative error of V_l in the measurement noise matrix $\mathbf{R}$ (see Table 4.2) is set to a higher value than for the initial system's noise $\mathbf{P}_0$ in eq. (4.48) to account for the unmodelled microbial water production during the cultivation. As for $\mathbf{Q}$ and $\mathbf{P}_0$, all off-diagonal elements in $\mathbf{R}$ were set to zero.

For the gas phase model, global observability was proven analytically by Lie derivatives using the toolbox STRIKE-GOLDD2 v2.01 developed by Villaverde et al. (2019).

Table 4.2: Postulated measurements errors of the gas phase model

Measurement	Symbol	Unit	Absolute tolerance $e_{\min}$	Relative tolerance e_{rel}
Excess pressure	ΔP	mbar	0.5	$0.01 \cdot y_1$
Gaseous fraction of H_2	$x_{H_2,v}$	%	0.5	$0.01 \cdot y_2$
Gaseous fraction of CO_2	$x_{CO_2,v}$	%	0.3	$0.02 \cdot y_3$
Gaseous fraction of O_2	$x_{O_2,v}$	%	0.3	$0.02 \cdot y_4$
Volume	V_l	L	0.05	$0.05 \cdot y_5$

4.4.4 Gas phase control performance

For all cultivations of H16, HF805 and H798 carried out within the scope of this thesis, the gas phase was controlled. The same applies for the strain HF951, which was, however, not cultivated within the context of this thesis. But since the gas phase controller was constantly evolving, only a few cultivations were carried out with the final strain-independent controller described in the previous sections. Two of these experiments (REatc33a and REatc34) are presented here, because they demonstrate the performance of the controller well. In these cultivations the controller had to cope with common situations such as sampling, abrupt gas transfer changes caused by antifoam dosage, metabolic changes of the organisms and a temporary server disconnection. Temporary missing data exchange is not common, but the controller should still be able to restart control after the connection has been restored. In both cultivations, the employed controller kept the reference trajectories for pressure and gas composition to a satisfactory degree as presented in Figures 4.5 and 4.6. In REatc34 (see Figure 4.5), when large volumes of cultivation broth were sampled, causing a sudden pressure drop as in the case between batch age 22–26 h, the excess pressure was adjusted immediately. Due to an employed gas model estimation sampling rate of 0.1 h, these pressure drops were not compensated for by a correction of the estimates of $\nu_{\text{gas,v}}$. Instead, the $CL_{\Delta P}$ unit of the double-layer PI controller in Figure 4.4 adjusted fast pressure changes with its lower sampling time of 0.0042 h. Internal disturbances resulting in quickly changing gas uptake rates were also compensated for by the controller. Here, between batch age 8–17 h indicated by vertical lines, the gas uptake rates decreased abruptly due to the punctual feeding of an antifoam agent, which inhibits gas transfer; consequently, all dissolved gas concentrations (not shown) decreased. When gases are missing in the liquid phase, the organisms starve, and therefore their gas uptake diminishes as well. It can be observed that a gas uptake drop led to a short overshoot in pressure until the gas flow rates are adjusted by the $CL_{\Delta P}$ control unit.

Microorganisms change their metabolism, and thus gas uptake composition during a cultivation. At batch age 10 h in REatc34, 5.7 $\text{L}\,\text{h}^{-1}$ of the entire gas flow (34.1 $\text{L}\,\text{h}^{-1}$) was oxygen, representing 16.7 % of $q_{\text{t,v}}$. Nonetheless, only the wanted 7 % oxygen was measured in the gas phase. Apparently, the controller needed to feed relatively more oxygen to maintain this

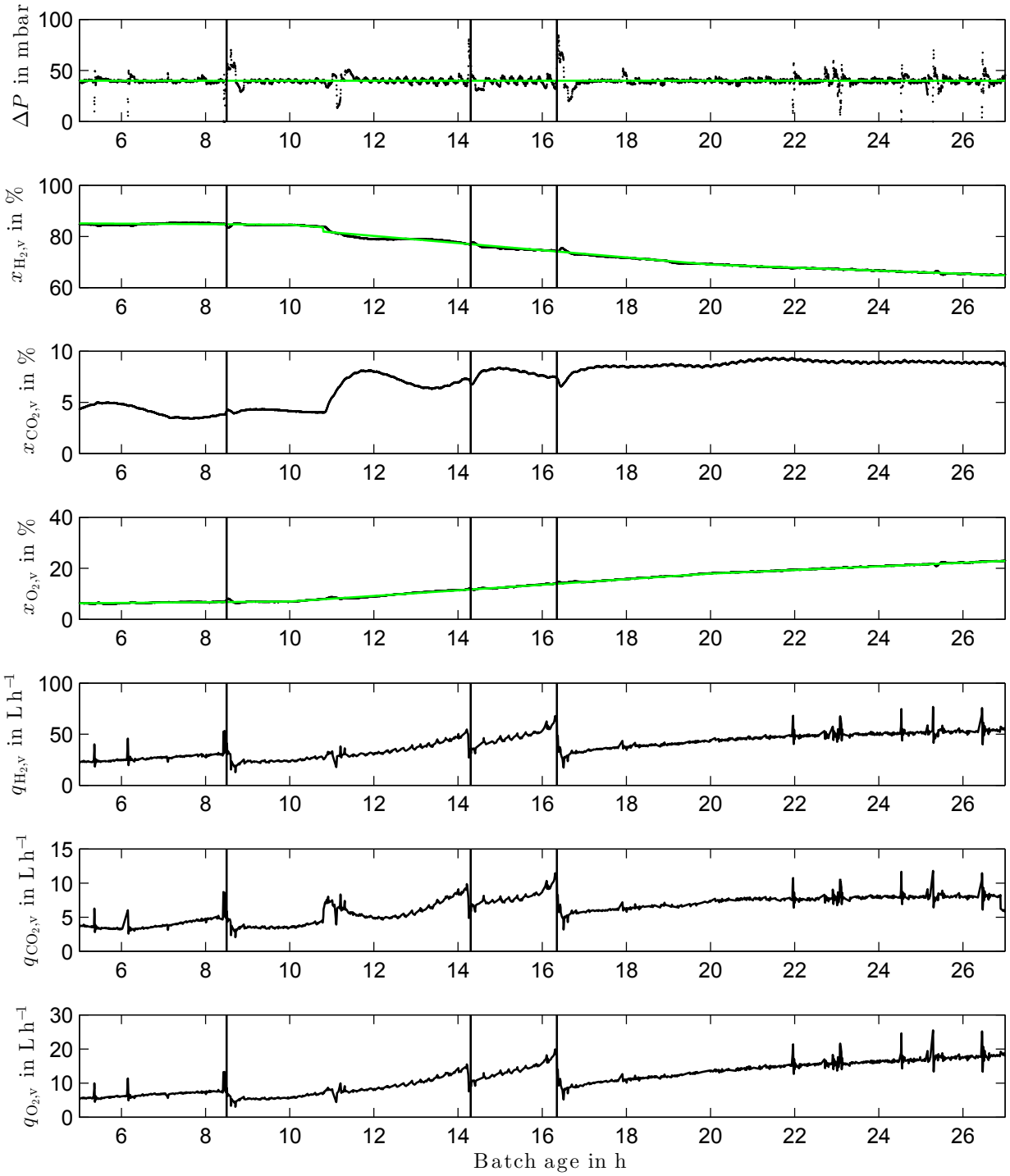


Figure 4.5: Gas phase variables ΔP and $x_{\mathrm{gas,v}}$ of REatc34 with abrupt changes in the fed gas flows ($q_{\mathrm{gas,v}}$ in black lines) due to the addition of antifoam or sampling of large volumes. Control setpoints are given in green and measured values in black dots.

low reference at this time in the cultivation. In the course of the experiment, this relation shifted towards the opposite, indicating different gas consumption coefficients at different stages of cultivation, which was also reported by Morinaga et al. (1978). These rather slow changes were tracked and compensated for by the integral part of $\mathrm{CL_{gas}}$.

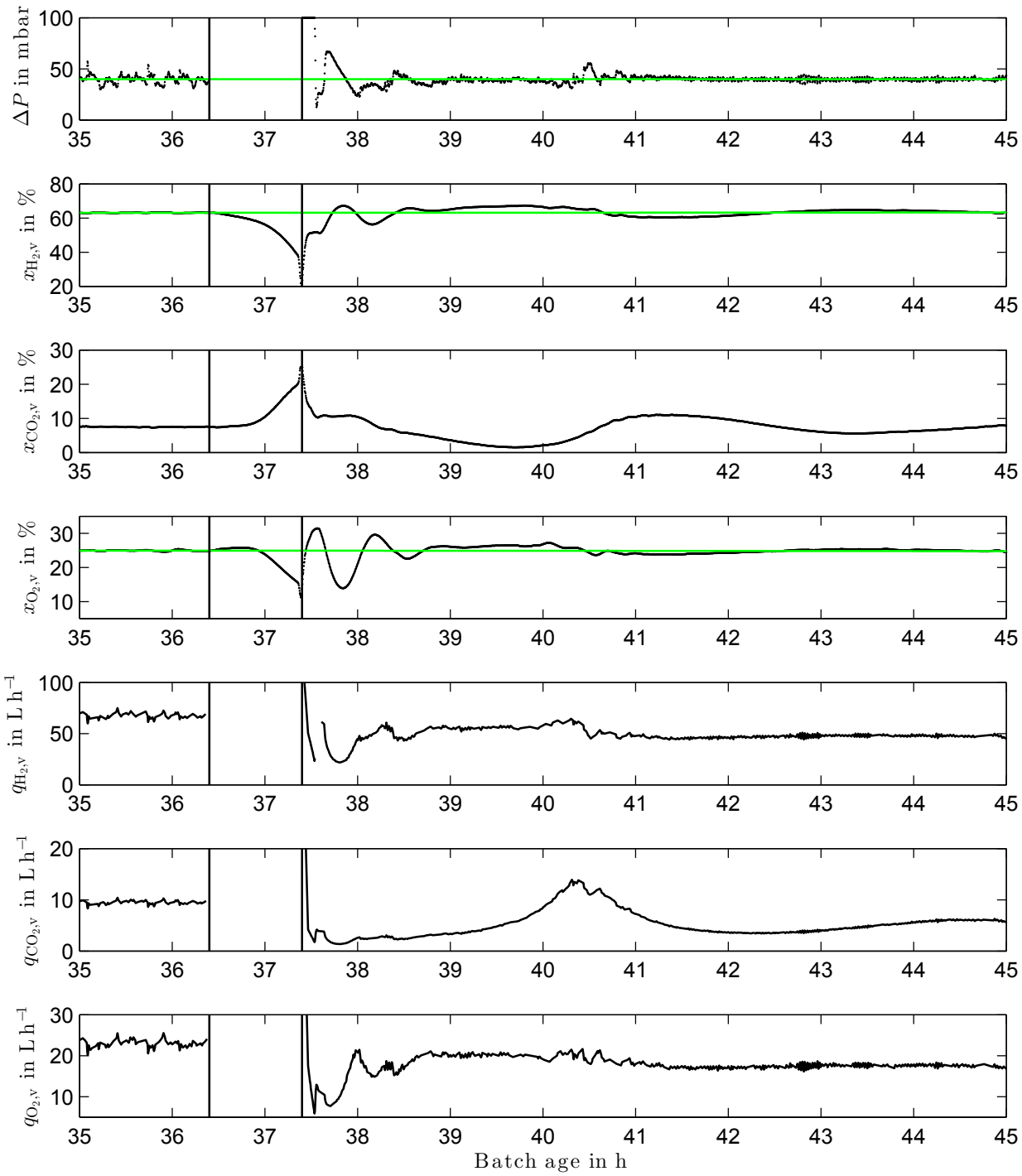


Figure 4.6: In cultivation REatc33a a phase with a strong disturbance occured. Due to a one-hour database disconnection (vertical lines), the gas phase was uncontrolled and environmental nitrogen entered the system as the result of negative excess pressure. The gas phase variables ΔP and $x_{\mathrm{gas,v}}$ were maintained by calculating the fed gas flows ($q_{\mathrm{gas,v}}$ in black lines). Control setpoints are given in green and measured values in black dots.

The developed gas controller also worked when atmospheric air entered the system allowing nitrogen to pollute the gas phase. In Figure 4.6, the gas phase variables of a period

of REatc33a are presented, where the data exchange was interrupted for one hour (vertical lines) due to a server breakdown. Accordingly, no gas was fed, the excess pressure dropped drastically below zero, and the tightness limit was passed, resulting in an inflow of air. In such an event, summarizing the three measured gas fractions plus evaporated water yields a value below 100 % and indirectly confirms the presence of nitrogen in the system. Restarting the gas phase controller after fixing the server communication problem first led to oscillations and after an hour the setpoints were well tracked again. Oscillations after the communication gap resulted from the large discrepancies between pressure and gas composition towards their corresponding set values. These initially large control errors were integrated in both layers $CL_{\Delta P}$ and CL_{gas} and slowly reduced in an oscillating manner. Large control errors result from software or hardware interruptions, but also from stepwise setpoint changes. To avoid oscillations in standard operation, setpoint trajectories should be ramped.

The gas controller described above does not require any biological information about the organism. It can be used both in fully controlled cultivations and in experiments in which the operator specifies the trajectories of the gas components. The latter are of particular interest when the process model is still unknown, e.g., because a new strain is cultivated to collect data for subsequent model adaption. This is based on an adaption of the process model for the wild type H16, which is introduced in the following chapter. In addition, cultivations that served as cross-validations are presented and discussed.

Chapter 5

Process models

In the previous chapter, a strain-independent gas phase controller was introduced, which relies on a model of the gas phase. This model does not involve any metabolic information of the organism and serves to estimate gas transfer or gas consumption in cultivations. In the present chapter, so-called process models (I) and (II) are proposed. They contain strain-specific biological information and describe growth in autotrophic cultivations of *R. e.* H16. Process model (I) includes the gas phase, and thus models the gas transport between liquid and gaseous phase and its fundamentals are presented in Section 5.1 before the model itself will be given in Section 5.2. All dissolved gas concentrations and gas consumption rates are calculated, in contrast to the gas phase model of the previous chapter, and can be compared to the measurements, except for dissolved hydrogen. Process model (I) serves for estimations of nutrients, biomass and products in full-control mode as explained in Section 4.3 and represents a basis for model adaption of mutant strains. Model (I) is very complex and in order to simplify steps in model adaption, e.g., estimation of parameters, it was reduced to a second model.

Process model (II) is a more simple model as it only considers the liquid phase and thus does not model gas transport. Hence, dissolved gas substrates cannot be calculated but they are needed as they are used directly as model inputs. For oxygen and carbon dioxide, measurements ($p_{O_2,l}$ and $c_{CO_2,l}$) were available. Although a measuring probe for dissolved hydrogen was developed as in Section 5.3, it could not be used for this thesis. Nevertheless, results of the probe in cultivation are shown and briefly discussed. A data-driven calculation of the model input $c_{H_2,l}$ is proposed instead and the resulting values compared to the simulations of process model (I) in Section 5.4 before process model (II) is introduced in Section 5.5. Its limitations are pointed out in Section 5.6.

5.1 Fundamentals of the process model (I)

Each model is valid for a certain subspace of the state-space. This limits model-based design of cultivation factors, such as environmental conditions, possible feeding and state trajectories of cultivations: Whenever a single factor is beyond the scope, the applicability of the model is questionable. Some elements, which define the scope of process model (I), will be discussed below.

The process model (I) of this chapter, does not include the dynamics of the gas phase control explained above, although an integration of these dynamics was tested, as described in the Appendix B. It turned out, though, that the resulting superior model could rarely be utilized, because solving of the system was computationally too demanding to be employed in closed-loop process control.

Parts of process model (I) were taken from Rossner (2014), including the description for gas transport. Gas transport via the liquid-gas interface was modeled by using gas transfer coefficients k_La for each gaseous compound. Rossner (2014) reported that the $k_La_{CO_2}$ was identified to a value of 41.7 h^{-1} that is about ten times smaller than predicted by the film theory explained in Garcia-Ochoa and Gomez (2009). Assuming that the value of 41.7 h^{-1} is wrong and the true value is higher would mean that CO_2 gets dissolved faster. Evaluating the well matching dissolved CO_2 trajectories of measurements and simulations, a faster transport is only possible, when CO_2 is consumed at a higher rate than explained by the model. If there was an additional and unmodelled cell compound consisting of carbon mostly, a discrepancy between measured and simulated cell dry weight (c_X) would result. As this has not been observed, extracellular carbon based metabolites were suspected and an appropriate analysis method of the filtrated medium was established (see Section 3.4). As no carbon could be detected, the idea of an unmodelled and secreted metabolite was dismissed and step response experiments to determine the $k_La_{CO_2}$ in medium without organisms were performed. Results and details are given in Section 5.1.1.

In order to extend the scope of the model with regard to limitations in gas supply, an increased gas transport would be beneficial. This could be achieved by varying the stirring speed to overcome growth limitations due to the absence of dissolved gases. Therefore, the Section 5.1.2 proposes a description of the gas transfer coefficient k_La as a function of the stirring speed.

In the previous chapter, the cultivation REatc34 was presented, where the controller had to compensate for sudden drops of gas consumption caused by pulsed antifoam dosages. The physical and chemical relationships leading to this occurrence will be explained in Section 5.1.3.

5.1.1 Gas transfer rate of carbon dioxide

The identified $k_L a_{CO_2}$ value of 41.7 h^{-1} was significantly smaller than reported in the literature and this discrepancy could not be explained by an unmodelled carbon dioxide assimilation. Experiments, so-called step responses, without organisms were performed to determine the gas transfer rate.
A simple model that describes the system of the experiments was developed that allows $k_L a_{CO_2}$ identification. In the model, gaseous and dissolved CO_2 ($x_{CO_2,v}$, $c_{CO_2,l}$) are regarded as input and output, respectively. Dissolved CO_2 in the liquid phase measured by an inline probe serves as output signal of the system whereas the fraction of CO_2 in the headspace serves as one of the inputs. In order to be able to evaluate the experiments, part of the setup of Section 4.1 is particularly relevant. This part is shown in Figure 5.1 and it had to be modified before carrying out the experiments for the determination of the $k_L a_{CO_2}$. These modifications are motivated and explained below before introducing the experimental procedure for step responses and describing the gas transport model on which this study is based.
In the system, the gas composition in the headspace is not measured by an inline probe but by peripherally located gas sensors (QR10-12) instead, and thus a delay of model input $x_{CO_2,v}$ has to be considered. Peripherally means that a defined volume flow of the circulating headspace gas is pumped (F09) with a flowrate $q_{leak,v} = 6.5\ L\,h^{-1}$ through tubes and a water lock system to the gas sensors. A tubing with the length $L_1 = 0.8$ m and the inner diameter $d_1 = 6$ mm connected to a longer tube with the length $L_2 = 2.7$ m and the inner diameter $d_2 = 3$ mm is linked to the water lock bottle of headspace volume = 210 mL before the bottle is joined with the gas sensors passing another tube that is $L_3 = 0.5$ m long and has the inner diameter $d_3 = 3$ mm. In all three tubes, a plug flow is postulated that leads to a dead time T_0 calculated according to

$$T_0 = \frac{V_{tube}}{q_{leak,v}} = \sum_{i=1}^{3} \frac{\pi \cdot d_i^2 \cdot L_i}{4 \cdot q_{leak,v}}. \tag{5.1}$$

Inserting the given volumes and lengths for the present system into eq. (5.1) leads to a total dead time of approximately 25 seconds, which is referred to as "tubing delay" below.
In contrast to the tubing, the gas concentration in the headspace of the water lock bottle is not approximated as a plug flow. Gas approaching the bottle is mixed with previous gas compositions leading to a time delay of first order. Assuming a homogeneous gas mix in the bottle, the change of $x_{CO_2,v}$ over time can be derived by a material balance of inlet, subscript "in", and outlet of CO_2

$$\frac{dn_{CO_2,v}}{dt} = \dot{n}^{in}_{CO_2,v} - \dot{n}_{CO_2,v}. \tag{5.2}$$

Assuming the ideal gas law, material quantities and material flows of eq. (5.2) are replaced by volumes and volume flows, yielding

$$V_{\text{bottle}} \cdot \frac{dx_{\text{CO}_2,\text{v}}}{dt} = q_{\text{leak,v}} \cdot x^{\text{in}}_{\text{CO}_2,\text{v}} - q_{\text{leak,v}} \cdot x_{\text{CO}_2,\text{v}}. \tag{5.3}$$

Separating variables and integrating from time $t_0 = 0$ and initial gas concentration in the bottle $x^0_{\text{CO}_2,\text{v}}$ to the time of interest t_1 and gas concentration of interest $x^1_{\text{CO}_2,\text{v}}$ results in

$$\ln \frac{x^{\text{in}}_{\text{CO}_2,\text{v}} - x^1_{\text{CO}_2,\text{v}}}{x^{\text{in}}_{\text{CO}_2,\text{v}} - x^0_{\text{CO}_2,\text{v}}} = -\frac{q_{\text{leak,v}}}{V_{\text{bottle}}} \cdot t_1, \tag{5.4}$$

with $x^1_{\text{CO}_2,\text{v}}$ and $x^0_{\text{CO}_2,\text{v}}$ being both either smaller or larger than $x^{\text{in}}_{\text{CO}_2,\text{v}}$. Solving eq. (5.4) for the target concentration $x^1_{\text{CO}_2,\text{v}}$ gives

$$\frac{x^{\text{in}}_{\text{CO}_2,\text{v}}}{x^{\text{in}}_{\text{CO}_2,\text{v}} - x^0_{\text{CO}_2,\text{v}}} - \frac{x^1_{\text{CO}_2,\text{v}}}{x^{\text{in}}_{\text{CO}_2,\text{v}} - x^0_{\text{CO}_2,\text{v}}} = e^{-\frac{q_{\text{leak,v}}}{V_{\text{bottle}}} \cdot t_1} \tag{5.5}$$

and finally

$$x^1_{\text{CO}_2,\text{v}} = x^{\text{in}}_{\text{CO}_2,\text{v}} - (x^{\text{in}}_{\text{CO}_2,\text{v}} - x^0_{\text{CO}_2,\text{v}}) \cdot e^{-\frac{q_{\text{leak,v}}}{V_{\text{bottle}}} \cdot t_1}. \tag{5.6}$$

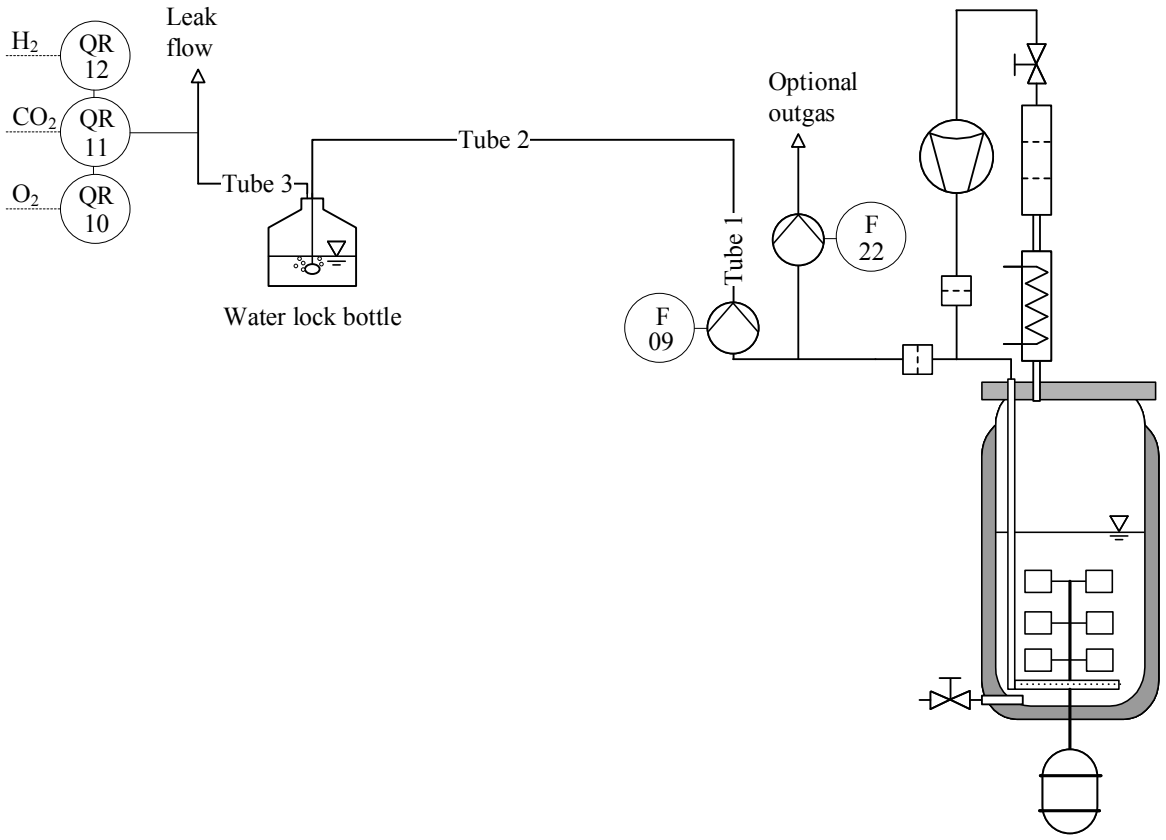


Figure 5.1: Detailed vaporous gas measurement setup, zoomed in from Figure 4.1 on page 38

Simplifying equation (5.6) by assuming an initial gas concentration $x^0_{CO_2,v} = 0$ yields

$$x^1_{CO_2,v} = x^{in}_{CO_2,v} \cdot \left(1 - e^{-\frac{q_{leak,v}}{V_{bottle}} \cdot t_1}\right). \tag{5.7}$$

For exponents in eq. (5.6) and eq. (5.7) greater than 4.7, more than 99 % of the final gas concentration are reached. With $V_{bottle} = 210$ mL and $q_{leak,v} = 6.5\ L\,h^{-1}$ it takes about 9.4 minutes until the gas composition in the bottle equals the influent to more than 99 %, which is referred to as "bottle delay" below.
Summing up bottle and tubing delay, it takes about 10 minutes until the gas composition of the headspace is detected by the gas sensors. To reduce the input time delay for $k_La_{CO_2}$ experiments, the water lock bottle was removed. A constant supervision of the experiment ensured that even without the water lock bottle no liquid of the reactor would reach the gas sensors in the event of pressure fluctuations. But for these experiments that are described below, another bottle was connected to the headspace of the reactor, see point 1 of the procedure description, to allow an additional outflow (besides the leakage flow). This was necessary because sufficient gas had to flow out so that sufficient gas could be pumped in at constant pressure during the step responses.

Experimental procedure

To record step responses of dissolved CO_2, a certain procedure needed to be completed. It included conditioning, initiating the step and detection of the steady state.

1. The headspace of the reactor was connected via a silicone tube to an explosion protected bottle of 1 L filled with water of volume ≈ 0.7 L. The connection tube ending was located at the bottom of the bottle in order to maintain a hydrostatic excess pressure. To allow a gaseous flow through the bottle, there was a gas outlet in the bottle's lid.

2. The reactor was loaded with CO_2-free medium tempered at 30 °C.

3. To condition the headspace of the reactor, it was filled with gas that contained 10 % of CO_2 and 90 % of H_2. Stirrer and circulation pump were switched off to prevent a significant CO_2 transport into the liquid phase before initiating the step.

4. The stirrer speed was then increased from zero to 500 rpm and the circulation pump, which pumped the gas from the headspace through the sparger into the liquid phase, was switched on. Fresh gas containing 10 % CO_2 and 90 % H_2 was pumped into the headspace with an average cultivation condition flow rate of 50 $L\,h^{-1}$. Due to the bottle from point 1, inflow rates higher than the leakage flow were possible. The reactor was tempered at 30 °C.

5. During the entire experiment, $c_{CO_2,l}$, $x_{CO_2,v}$ and the excess pressure ΔP were monitored.

6. The step response was complete when no change of gas concentration neither in the gaseous nor in the liquid phase was observed.

7. The thereby measured data was utilized for estimating and validating the parameter $k_L a_{CO_2}$ of the gas transport model.

Four experiments were carried out according to this procedure to collect data and to model the gas transport. The model used is explained below.

Gas transport model

The experimental data of the step responses served for the determination of the $k_L a_{CO_2}$. They were used for simulation with a model that only describes the transport of carbon dioxide and the dynamics of the $c_{CO_2,l}$ sensor. This model consists of two states and two inputs. Dissolved carbon dioxide $c_{CO_2,l}$ in $g\,L^{-1}$ is the first state, derived from a mass balance and assuming a constant volume, and it changes with the transfer of CO_2 into the medium

$$V_l \cdot \dot{c}_{CO_2,l} = V_l \cdot \dot{c}_{trans,CO_2} \tag{5.8}$$

$$\dot{c}_{CO_2,l} = k_L a_{CO_2} \cdot (c_{CO_2,sat} - c_{CO_2,l}). \tag{5.9}$$

The saturation concentration $c_{CO_2,sat}$ is calculated with Henry's law

$$c_{CO_2,sat} = H_{CO_2,30} \cdot p_{CO_2,v} \cdot M_{gas}, \tag{5.10}$$

with $p_{CO_2,v} = x_{CO_2,v} \cdot (P_0 + \Delta P)$. The fraction of carbon dioxide $x_{CO_2,v}$ and the excess pressure ΔP are both measured in the headspace and these measurements serve as inputs for the gas transport model.
To describe the delay time of the $c_{CO_2,l}$ sensor, which was approximated with a first order system as in Rossner (2014), a second state eq. (5.11) was introduced. The apparent dissolved carbon dioxide concentration changes with the reciprocal time constant T_{1,CO_2} and with the difference of dissolved carbon dioxide $c_{CO_2,l}$ and apparent, superscript "app", dissolved CO_2. It was calculated in $mg\,L^{-1}$ so that it could be compared directly to measurement values of the $c_{CO_2,l}$ sensor

$$\dot{c}^{app}_{CO_2,l} = \frac{1}{T_{1,CO_2}} \cdot (c_{CO_2,l} \cdot 1000\frac{mg}{g} - c^{app}_{CO_2,l}). \tag{5.11}$$

Rossner (2014) performed experiments to evaluate the $c_{CO_2,l}$ sensor behavior and identified T_{1,CO_2} to be 0.15 h. For parameter estimation, this value was used together with three out of four obtained step response data sets, yielding in an estimated value for $k_L a_{CO_2}$ of 31 h^{-1}. Data of the three identification experiments and cross-validation are shown in Figure 5.2. As soon as the step was initiated by switching on stirring, gas flow and circulation pump,

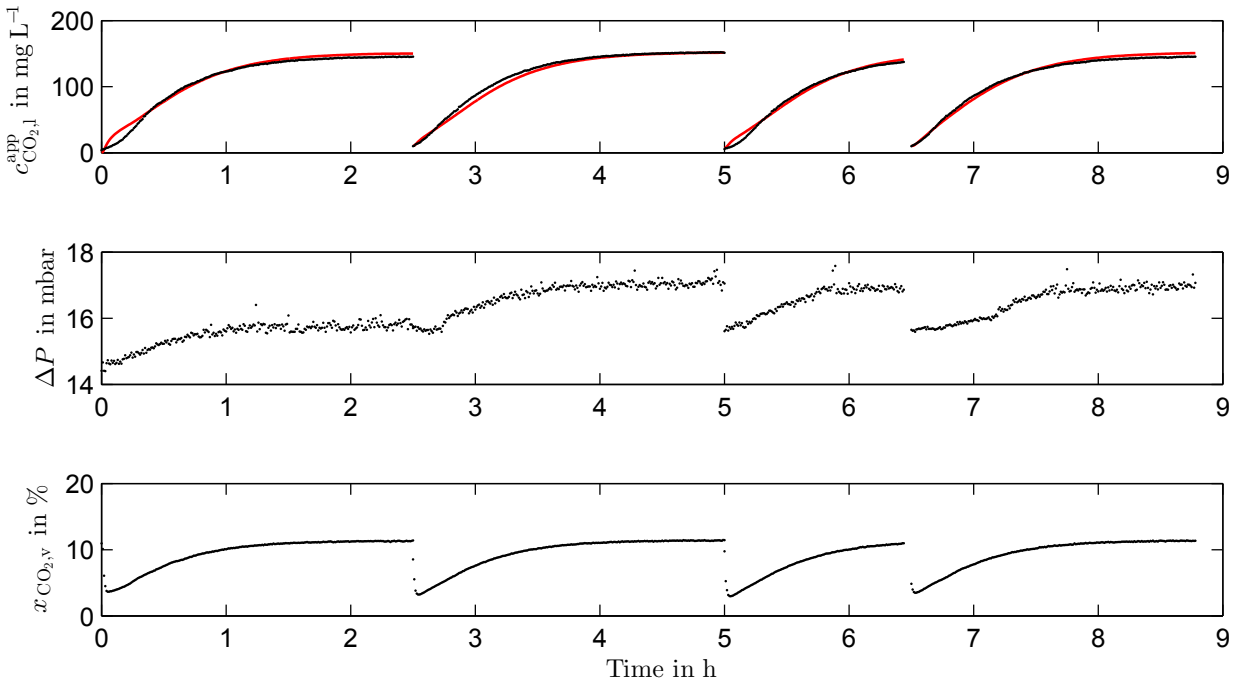


Figure 5.2: Upper plot: Simulations (red) versus experimental data (black) of four concatenated step responses, the fourth is a cross-validation. Lower plots: Corresponding inputs of the gas transport model.

the gaseous fraction of CO_2 decreased as it was transferred into the liquid phase to replace the dissolved air, which in turn outgassed into the headspace and diluted CO_2. The excess pressure increased over time because the bottle outlet was equipped with a microporous filter that got blocked by evaporating water in the course of the experiment.
All in all, the gas transport model gives a satisfactory approximation of $c_{\mathrm{CO_2,l}}$, when $\mathrm{k_L a_{CO_2}}$ adopts a value of 31 $\mathrm{h^{-1}}$, which is fairly close to the value identified by Rossner (2014), but again contradicts the value that is predicted by the film theory. Apparently, in the present system, dissolved CO_2 got consumed that is why $\mathrm{k_L a_{CO_2}}$ adopted a lower value than reported by the literature. An unmodelled biological consumption of CO_2 was neglected because the experiments were carried out without organisms. Moreover, the analysis results for extracellular carbon-based metabolites (see Section 3.4) supported this assumption. Hence a chemical consumption is suggested. Dissolving CO_2 reacts with water to carbonic acid and its residue anions,

$$\mathrm{H_2O + CO_2 \rightleftharpoons H_2CO_3 \rightleftharpoons HCO_3^- + H^+ \rightleftharpoons CO_3^{2-} + 2\,H^+}\,. \tag{5.12}$$

The cultivation medium is based on a sodium-potassium buffer as listed in Section 4.2, and hence the following equilibrium reactions occur:

$$\mathrm{Na_2HPO_4} \rightleftharpoons \mathrm{NaHPO_4^-} + \mathrm{Na^+} \rightleftharpoons \mathrm{HPO_4^{2-}} + 2\,\mathrm{Na^+} \tag{5.13}$$

$$\mathrm{KH_2PO_4} \rightleftharpoons \mathrm{HPO_4^-} + \mathrm{K^+} \rightleftharpoons \mathrm{PO_4^{2-}} + \mathrm{K^+} + \mathrm{H^+}. \tag{5.14}$$

Both cations of eq. (5.13)–(5.14) react by binding with the residue anions of eq. (5.12) to

$$\mathrm{K^+} + \mathrm{HCO_3^-} \rightleftharpoons \mathrm{KHCO_3} \tag{5.15}$$

$$\mathrm{Na^+} + \mathrm{HCO_3^-} \rightleftharpoons \mathrm{NaHCO_3} \tag{5.16}$$

$$2\,\mathrm{K^+} + \mathrm{CO_3^{2-}} \rightleftharpoons \mathrm{K_2CO_3} \tag{5.17}$$

$$2\,\mathrm{Na^+} + \mathrm{CO_3^{2-}} \rightleftharpoons \mathrm{Na_2CO_3}\,. \tag{5.18}$$

The removal of residue anions by reactions (5.15)–(5.18) is compensated for by transforming more dissolved CO_2 into carbonic acid and its residue anions as given in eq. (5.12). Until all reactions are in equilibrium, less physically dissolved CO_2 is present in the liquid phase than it would be without the chemical reactions listed above. As the $k_L a_{CO_2}$ seems to be lower than reported by literature, gas transport takes longer to reach equilibrium. Hence, the $k_L a_{CO_2}$ that was experimentally determined by Rossner (2014) contains both, the delays caused by chemical reactions and the actual gas transport. In this thesis, the process model (I) will not include the above listed chemical reactions but the step responses carried out give an order of magnitude for the expected parameter value for $k_L a_{CO_2}$ of the process model (I).

5.1.2 Impact of stirring on gas transfer

A constant value for $k_L a_{gas}$ can only be assumed for experiments with constant stirring speed. Since dissolved gases can become growth limiting in high-density cultivations, an optionally faster gas transport was desired. In this thesis, the influence of the parameter stirring speed on gas transport was investigated. Hence, a gas transport model that depends on stirring was formulated. Experiments were conducted and their data used for parameter estimation. Similar to the experimental setup for $k_L a_{CO_2}$ determination, these tests were performed without organisms and without the water lock bottle of Figure 5.1. Also for these tests, similar to the step response experiments, an explosion protected bottle was connected to the headspace as gas outlet to allow the feeding of gas flows larger than the leakage flow. For these preliminary investigations, pH control was switched off and the reactor was filled with tap water instead of medium. A constant total inlet volume gas flow of 50 $\mathrm{L\,h^{-1}}$ was

blown into the reactor.
Henzler (1982) suggested a dimensionless formulation of k_La for each gas component, which depends on the gas empty tube speed w, the kinematic viscosity ν of the fluid, here water, the gravitational constant g, the power required to rotate the stirrer, the liquid phase volume of the reactor V_l, density ρ, the stirring speed n, the stirrer diameter d and two system dependent constants γ_1, γ_2, which were identified in a parameter estimation

$$\frac{k_L a_{gas}}{w_{gas}} \left(\frac{\nu_{H_2O}^2}{g} \right)^{\frac{1}{3}} = \gamma_1 \left(\frac{\text{power}}{V_l \cdot w_{gas} \cdot \rho_{H_2O} \cdot g} \right)^{\gamma_2}. \tag{5.19}$$

The power input is unknown, but the bracketed term on the right side of eq. (5.19) can be replaced by

$$\frac{\text{power}}{V_l \cdot w_{gas} \cdot \rho_{H_2O} \cdot g} = Ne_{gas} \frac{n^3 \cdot d^5}{V_l \cdot w_{gas} \cdot g}, \tag{5.20}$$

as suggested by Kraume (2012), where the dimensionless Newton number Ne_{gas} is calculated with the empirical correlation

$$Ne_{gas} = 1.5 + \frac{1}{0.5 \cdot Q_{gas}^{0.075} + 1600 \cdot Q_{gas}^{2.6}}. \tag{5.21}$$

These correlations are only valid for a Reynolds number $Re \geq 10^4$ and Froude numbers $Fr \geq 0.65$. Moreover, the gas load Q_{gas} must be in the interval $[1.8 \cdot 10^{-4}; 0.5]$. Q_{gas} is defined as

$$Q_{gas} = \frac{q_{gas}}{n \cdot d^3} \tag{5.22}$$

with the volumetric gas flows q_{gas}. Equation (5.19) is simplified by substituting

$$B_{gas} = w_{gas} \cdot \left(\frac{\nu_{H_2O}^2}{g} \right)^{-\frac{1}{3}}. \tag{5.23}$$

The fraction of the right hand side in eq. (5.20) is replaced by

$$Z_{gas} = \frac{n^3 \cdot d^5}{V_l \cdot w_{gas} \cdot g} \tag{5.24}$$

before the expression is inserted in the term on the right side of eq. (5.19). The final equation for the k_La that depends on stirring speed reads

$$k_L a_{gas} = B_{gas} \cdot \gamma_1 \cdot (Ne_{gas} \cdot Z_{gas})^{\gamma_2}. \tag{5.25}$$

The k_La_{gas} calculation is part of a model that in parts is similar to the headspace model from Section 4.4.1. It encompasses the amounts of gases in the headspace and the molar

concentrations for CO_2 and O_2 and its state vector reads as

$$\underline{x}^T_{\text{kla model}} = (n_{\text{H}_2,\text{v}} \quad n_{\text{CO}_2,\text{v}} \quad n_{\text{O}_2,\text{v}} \quad c^{\text{n}}_{\text{CO}_2,\text{l}} \quad c^{\text{n}}_{\text{O}_2,\text{l}} \quad c^{\text{n,app}}_{\text{CO}_2,\text{l}}). \tag{5.26}$$

It is assumed, though, that the amount of evaporated water is constant and gas phase contamination by nitrogen neglectable. The associated system of differential equations obtained from material balances of the headspace is listed below:

$$\begin{aligned}\dot{n}_{\text{H}_2,\text{v}} =& \frac{P}{R \cdot T}\left(q_{\text{H}_2,\text{v}} \cdot 10^{-3} - q_{\text{leak,v}} \cdot 10^{-3} \cdot x_{\text{H}_2,\text{v}}\right) \\ & - \text{k}_\text{L}\text{a}_{\text{H}_2} \cdot (c^{\text{n}}_{\text{H}_2,\text{sat}} - c^{\text{n}}_{\text{H}_2,\text{l}}) \cdot V_\text{l} \end{aligned} \tag{5.27}$$

$$\begin{aligned}\dot{n}_{\text{CO}_2,\text{v}} =& \frac{P}{R \cdot T}\left(q_{\text{CO}_2,\text{v}} \cdot 10^{-3} - q_{\text{leak,v}} \cdot 10^{-3} \cdot x_{\text{CO}_2,\text{v}}\right) \\ & - \text{k}_\text{L}\text{a}_{\text{CO}_2} \cdot (c^{\text{n}}_{\text{CO}_2,\text{sat}} - c^{\text{n}}_{\text{CO}_2,\text{l}}) \cdot V_\text{l} \end{aligned} \tag{5.28}$$

$$\begin{aligned}\dot{n}_{\text{O}_2,\text{v}} =& \frac{P}{R \cdot T}\left(q_{\text{O}_2,\text{v}} \cdot 10^{-3} - q_{\text{leak,v}} \cdot 10^{-3} \cdot x_{\text{O}_2,\text{v}}\right) \\ & - \text{k}_\text{L}\text{a}_{\text{O}_2} \cdot (c^{\text{n}}_{\text{O}_2,\text{sat}} - c^{\text{n}}_{\text{O}_2,\text{l}}) \cdot V_\text{l} \end{aligned} \tag{5.29}$$

$$\dot{c}^{\text{n}}_{\text{CO}_2,\text{l}} = \text{k}_\text{L}\text{a}_{\text{CO}_2} \cdot (c^{\text{n}}_{\text{CO}_2,\text{sat}} - c^{\text{n}}_{\text{CO}_2,\text{l}}) \tag{5.30}$$

$$\dot{c}^{\text{n}}_{\text{O}_2,\text{l}} = \text{k}_\text{L}\text{a}_{\text{O}_2} \cdot (c^{\text{n}}_{\text{O}_2,\text{sat}} - c^{\text{n}}_{\text{O}_2,\text{l}}) \tag{5.31}$$

$$\dot{c}^{\text{n,app}}_{\text{CO}_2,\text{l}} = \frac{1}{T_{1,\text{CO}_2}}(c^{\text{n}}_{\text{CO}_2,\text{l}} - c^{\text{n,app}}_{\text{CO}_2,\text{l}}), \tag{5.32}$$

and with the ideal gas law the gas fractions read

$$x_{\text{gas,v}} = n_{\text{gas,v}} \cdot \frac{R \cdot T \cdot 10^3}{P \cdot V_\text{head}}, \tag{5.33}$$

with the pressure defined as

$$P = P_0 + \Delta P. \tag{5.34}$$

The saturation concentrations of dissolved gases are calculated as in

$$c^{\text{n}}_{\text{gas,sat}} = H_{\text{gas,30}} \cdot p_{\text{gas,v}}. \tag{5.35}$$

Since no liquids were added during the experiments and the trials were of short duration, compared to a cultivation, the assumption of a constant volume is justified. Thus, dissolved gas amount balances $\dot{n}_{\text{gas,l}}$ have been converted into molar concentrations $\dot{c}^{\text{n}}_{\text{gas,l}}$, see eq. (5.30)–(5.31). Moreover, for dissolved CO_2 the sensor delay was taken into account as

in eq. (5.32). The input vector contains stirring speed n and reads

$$\underline{u}^T_{\text{kla model}} = (\Delta P \quad n \quad q_{\text{H}_2,\text{v}} \quad q_{\text{CO}_2,\text{v}} \quad q_{\text{O}_2,\text{v}}). \tag{5.36}$$

To compare the simulations with the measurements, $c^{\text{n}}_{\text{O}_2,\text{l}}$ and $c^{\text{n,app}}_{\text{CO}_2,\text{l}}$ of the state vector are converted into dissolved partial pressure and mass concentration, respectively. Both quantities are multiplied by the molecular weights M_{gas} and the concentration of dissolved oxygen is then related to the saturation concentration with atmospheric oxygen, yielding $p_{\text{O}_2,\text{l}}$, which is given in %. Finally, the measurement vector is defined as

$$\underline{x}^T_{\text{kla model}} = (x_{\text{H}_2,\text{v}} \quad x_{\text{CO}_2,\text{v}} \quad x_{\text{O}_2,\text{v}} \quad p_{\text{O}_2,\text{l}} \quad c^{\text{app}}_{\text{CO}_2,\text{l}}). \tag{5.37}$$

The system dependent parameters γ_1, γ_2 of the gas transport description above were fitted by utilizing experimental data of initial experiments without organisms, which are not shown. In order to decrease the 1-σ-uncertainties of both γ, model-based Optimal Experimental Design was employed and the optimized experiment carried out. As shown in Figure 5.3, the optimizer changed the gas composition together with the stirring speed to allow for a more accurate subsequent parameter estimation leading to the values in Table 5.1. The relative standard deviations (rel. std. dev.) were calculated with the Fisher information matrix, as in eq. (3.8) on page 26 and represent only a lower limit of uncertainties. Additionally to the measurements and inputs, the simulated $\text{k}_\text{L}\text{a}_{\text{gas}}$ values are shown in Figure 5.3. For both parameters γ_1 and γ_2, which were identified with data of the OED experiment, the calculated standard deviation adopted small values meaning the estimated parameters are rather certain. Moreover, the identified values are quite close to those for water listed in Kraume (2012). Model-based OED seems an appropriate method to identify reliable values for γ. However, some cultivations with variations in stirring speed gave evidence that this may lead to unpredictable metabolic side effects, such as an enhanced PHB production. Consequently, to model the effect of stirring speed variations on gas transport does not suffice. Further studies on the effects on growth are needed and must be included in the process model (I). But since that would have gone beyond the scope of this thesis, all cultivations discussed here were run with constant stirring speed.

Table 5.1: Estimated gas transport parameters that were fitted to the data of an optimally planned experiment (model-based OED)

Parameter	Value (-)	Literature value (-)	rel. std. dev (%)
γ_1	$3.04{\cdot}10^{-5}$	$7.5{\cdot}10^{-5}$	7.18
γ_2	0.442	0.43	3.54

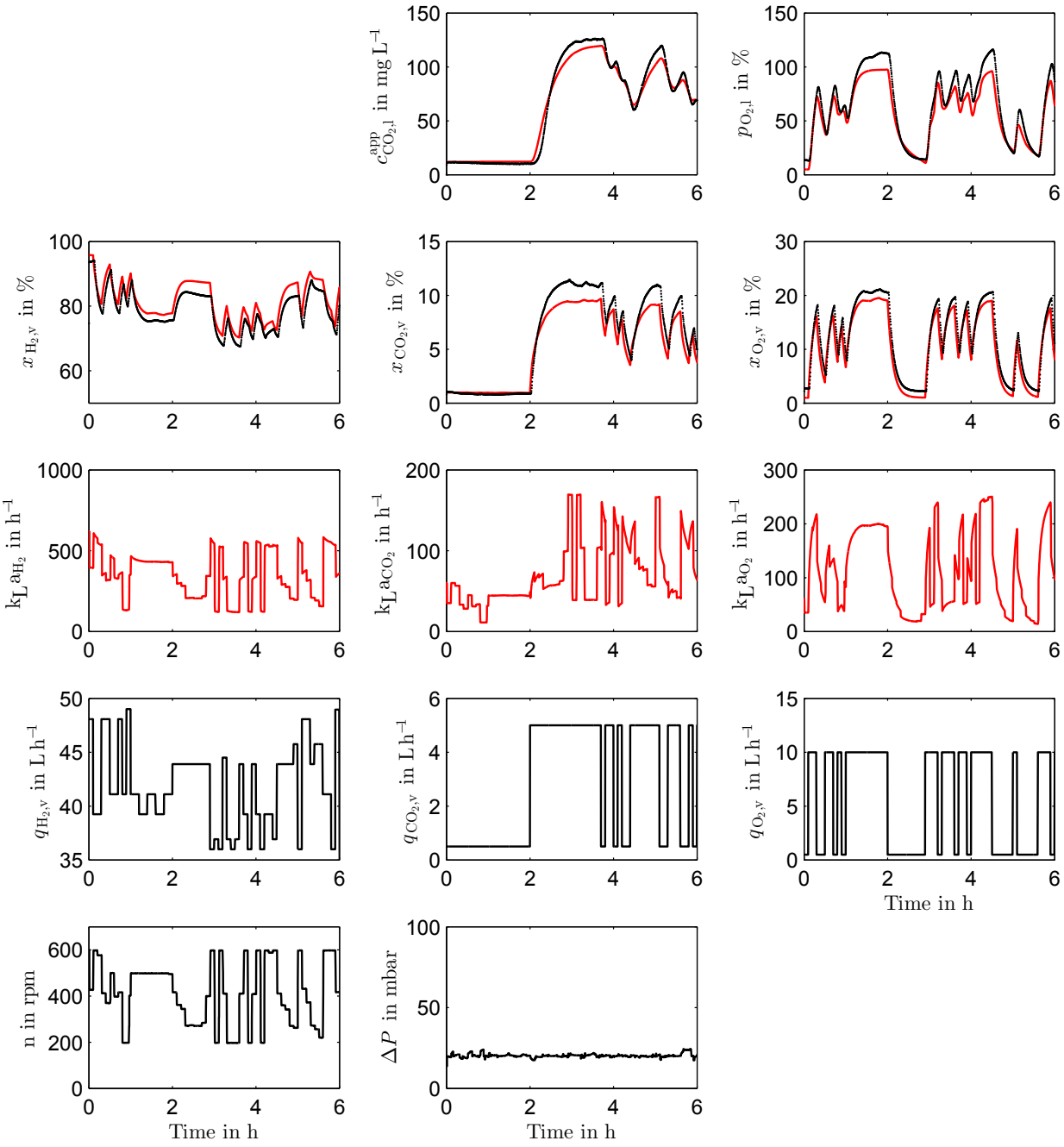


Figure 5.3: Simulations (red) and measurements (black dots) of the OED planned experiment. The underlying model relates stirrer speed and $k_L a$ to each other and its inputs are shown in the graphs of the two lower rows.

5.1.3 Influence of antifoam agent on gas transfer

During cultivations of *R. e.* it was observed that adding antifoam agent (PG 2000) led to sudden pressure peaks, which were compensated for by the gas phase controller by calculating reduced gas feed flows as was shown in Figure 4.5. A similar phenomenon was described by Seletzky (2009). He detected a relation between the alkoxylated fatty alcohol

antifoam agent Plurafac LF 1300 and decreased gas transfer coefficients k_La in cultivations of *Corynebacterium glutamicum*. Also Arjunwadkar et al. (1998) using silicon antifoam A, Calik et al. (2005) utilizing reagents fluorocarbon-hydrocarbon unsymmetrical bolaform and antifoam A, and Morão et al. (1999) employing polypropylene glycol, soybean oil and silicone oil observed similar effects.

Before we found this explanation for the reduced gas volume flows, we examined other hypotheses, which are briefly introduced below. Since *R. e.* was cultivated autotrophically, we hypothesized decreased gas feed flows resulting from diminished gas consumption by switching to heterotrophic growth. A sudden decrease of gas consumption is possible, when the organisms obtain energy and carbon from an additional source, e.g., the metabolization of antifoam agent. This hypothesis was rejected for two reasons. First, a stoichiometric analysis stated that the amount of antifoam agent mixture, containing PG 2000, ethanol and water, added at certain instances, cannot cause the notable drops in gaseous feed flows. Second, it was observed that after adding the antifoam agent mix, the $p_{O_2,l}$ level decreased. Again, one might conclude that the cells consume antifoam agent aerobically, and therefore require more oxygen causing the $p_{O_2,l}$ to decrease. In this case, gas transport would be accelerated because the driving difference in dissolved gas concentration would increase, leading to increased gas feed flows. Instead, the opposite, i.e., reduced feed flows, were observed in several cultivations, including a period of REatc33a shown in Figure 5.4. In order to prove that in *R. e.* cultivations the k_La value was changed by antifoam agents, affecting the gas transport, some calculations are carried out below. In the headspace of the reactor, the quantity of a gas compound changes with

$$\frac{d\dot{n}_{gas,v}}{dt} = \dot{n}_{gas,v,feed} - \dot{n}_{gas,leak} - \underbrace{k_La \cdot (c^n_{gas,sat} - c^n_{gas,l}) \cdot V_l}_{\text{gas transfer}}, \tag{5.38}$$

with $\dot{n}_{gas,v,feed}$, $\dot{n}_{gas,leak}$ and $c^n_{gas,l}$ being the gaseous feed flow, molar leakage flow and molar dissolved gas concentration, respectively. Note, the superscript "n" indicates that these are molar concentrations. Applying the ideal gas equation and reformulating the leakage flow gives

$$\frac{d\dot{n}_{gas,v}}{dt} = q_{gas,v} \cdot 10^{-3} \cdot \frac{P_0}{R \cdot T} - q_{leak,v} \cdot 10^{-3} \cdot x_{gas,v} \cdot \frac{P}{R \cdot T} - k_La \cdot (c^n_{gas,sat} - c^n_{gas,l}) \cdot V_l, \tag{5.39}$$

with P_0 being the ambient pressure, as the mass flow controllers record values for the gas flows that have been converted to ambient pressure. Employing a gas phase control of the gas fractions and pressure, and thus assuming pseudo-steady state with $\dot{n}_{gas,v} = 0$, results in

$$q_{gas,v} = \frac{R \cdot T}{P_0} \left(q_{leak,v} \cdot 10^{-3} \cdot x_{gas,v} \cdot \frac{P}{R \cdot T} + k_La \cdot \underbrace{(c^n_{gas,sat} - c^n_{gas,l})}_{\Delta c^n_{gas,l}} \cdot V_l \right). \tag{5.40}$$

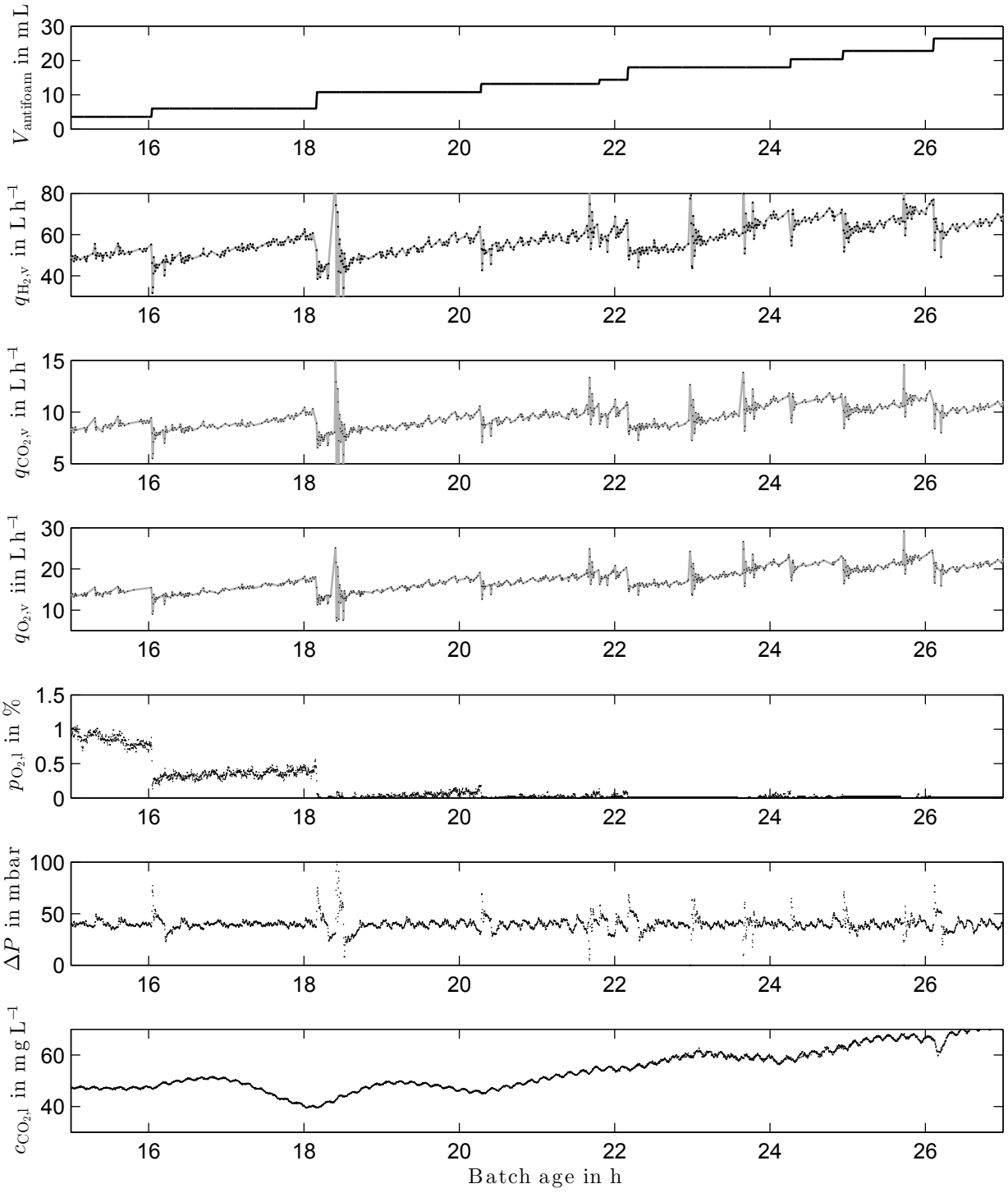


Figure 5.4: System response towards stepwise antifoam feeding in REatc33a. For better visualization the volume flows $q_{\text{gas,v}}$ are interpolated with gray lines.

As the system pressure P, $q_{\text{leak,v}}$, $x_{\text{gas,v}}$ as well as the reactor liquid volume V_{l} are constant in pseudo-steady state and R, T, P_0 are constant per definition, only the gas transfer can be held responsible for sudden drops of the gaseous feed flow rates as a response to adding antifoam. A decreased gas transfer might result either by a drop of the $k_{\text{L}}a$ or a reduced

driving molar concentration difference $\Delta c^{\mathrm{n}}_{\mathrm{gas,l}}$. Evaluating dissolved oxygen once antifoam was added, see Figure 5.4, the latter explanation can be discarded because $\Delta c^{\mathrm{n}}_{\mathrm{O_2,l}}$ increased rather what is indicated by measured drops of $p_{\mathrm{O_2,l}}$ at the relevant time instances, i.e., when antifoam is fed. In summary, adding antifoam agent to the cultivation broth caused decreased gas transfer because it affected the $\mathrm{k_La}$ negatively by changing viscosity parameters of the cultivation broth. In contrast to O_2, a drop of the dissolved CO_2 concentration could not be observed once antifoam was added, which can be explained when investigating the transport parameters. Carbon dioxide transfer is very slow and afflicted with a comparatively large time constant that shadows the response towards antifoam adding. Moreover, the gas component CO_2 was consumed less by the cells than O_2 and accordingly, the effects of reduced gas transport were not noticeable in dissolved CO_2. Additionally, CO_2 was not controlled in the gas phase as explained in Section 4.4.2 on page 46, and hence pseudo-steady state could not be assumed for this component. Nonetheless, for all gases including carbon dioxide, suddenly reduced feed flows $q_{\mathrm{gas,v}}$ and associated pressure overshoots were observed as a reaction towards pulsing antifoam feeding. However, these effects were not included in process model (I) and a constant $\mathrm{k_La}$ was used as presented in the next section.

5.2 General process model (I) for H16

In this section, the medium-sized structured general process model (I) is introduced. It is derived from mass balances of all relevant compounds plus balances of liquid volumes, mainly. Some parts of the model have been developed by Rossner (2014).
According to the model, schematically depicted in Figure 5.5, cells, abstrahized as an ellipse, consist of active biomass, the internal carbon storage compound PHB (abstrahized as an ellipse), membrane-bound hydrogenases (MBH) and soluble hydrogenases (SH). To grow, the cells consume ammonium (N), phosphate (P) and dissolved gases. All nutrients are directly fed either as liquids or gas flows. Hence, liquid and gaseous phases have to be considered as well as the gas transport via the interface. Due to pH-control, base and acid are fed when dissolved carbon dioxide fluctuates, but the major cause for base feeding is biologic ammonium uptake. Ammonium is required for building metabolites, such as proteins and DNA/RNA. Bacterial cells assimilate NH_3 instead of ammonium (NH_4^+) as reported in Ritchie (2013), and therefore a proton is split off during the uptake leading to a base feed flow to compensate for the pH drift. Consequently, the difference of fed base and acid is an indirect measure for the formation of active biomass corrected for carbon dioxide that serves as an additional state.
The model of Figure 5.5 encompasses equations to describe the evolution of states $\underline{x}(t)$, manipulating variables $\underline{u}(t)$ and equations for the measured quantities $\underline{y}(t)$ as detailed in Sections 5.2.4, 5.2.1 and 5.2.5, respectively. In the model, biomass compartments, dissolved nutrients as well as the difference of fed base and acid belong to the state vector. The

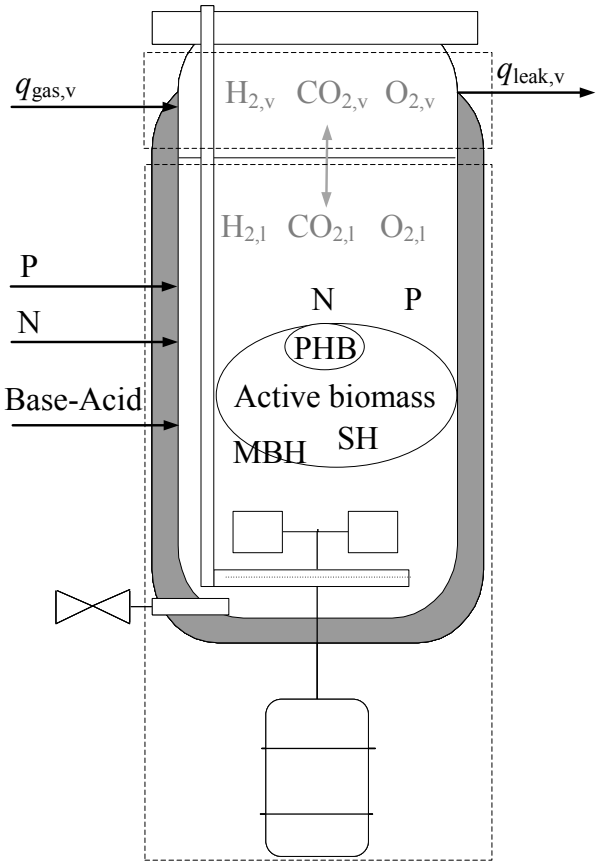


Figure 5.5: Simplified scheme of process model (I) in which the biomass (large ellipse) and its compartments are represented. System boundaries used for modeling are given in dashed lines. Liquid phase and gaseous model are connected via gas transport (gray).

dynamical behavior of the states relies on reaction rates, which depend on kinetic functions to be introduced in Section 5.2.2. In the biological context, kinetic functions are often used to predict reaction rates depending on extra- or intracellular conditions such as pH, nutrient concentrations, temperature. Gas transport between liquid and vapor phase is part of model (I) and supplementary equations for gas solubility are presented in Section 5.2.3. Identified model parameters are discussed in Section 5.2.6. The model does only apply for cultivations carried out under the general conditions of Section 4.1 that are listed in Table 5.2.

Table 5.2: General conditions of the general process model (I) for H16

Temperature	30 °C
Stirring speed	500 rpm
Fe feed flow	identical with the N feed flow
Compressor flowrate	approximately 7 L min^{-1}

5.2.1 Input vector for modeling

For defining model equations, an input vector must be defined. According to the system's input vector, eq. (4.3) on page 41, the gaseous and liquid flows are inputs. Consequently, gaseous volume flows ($q_{\mathrm{gas,v}}$) would be used to calculate the headspace gas composition and the pressure, which affects gas transport and dissolved gas concentrations, and thus the outcome of kinetic functions in terms of reaction rate values. Calculating the gas phase composition correctly is crucial to obtain valid simulation results. Peripheral, not quantifiable disturbances, such as gas phase leakages and gas phase control errors, would lead to wrong calculation of the gas phase composition. Therefore, rather pseudo-steady state for the gas phase is postulated and the measured gas fractions $x_{\mathrm{gas,v}}$ as well as the system's pressure P are directly used as inputs, which leads to more accurate simulations of dissolved gases. In summary, the input vector of model (I) reads as

$$\underline{u}^T_{\text{model I}} = (u_{\mathrm{N}} \;\; u_{\mathrm{Fe}} \;\; u_{\mathrm{P}} \;\; x_{\mathrm{H_2,v}} \;\; x_{\mathrm{CO_2,v}} \;\; x_{\mathrm{O_2,v}} \;\; P \;\; q_{\mathrm{leak,v}} \;\; u_{\mathrm{base}} \;\; u_{\mathrm{acid}} \;\; u_{\mathrm{antifoam}}). \tag{5.41}$$

5.2.2 Reaction rates

From a system biologic perspective, cells produce a certain compound via a specific metabolic pathway inheriting many metabolic intermediates and interfaces to different pathways. Each production rate of a metabolic intermediate depends on the presence (or absence) of certain proteins as well as other metabolic intermediates. Since these compounds often cannot be measured, neither in industrial applications nor in the laboratory, in this study, metabolic pathways are simplified by lumping metabolic steps together. Finally, most reaction rates describe the conversion of measurable substrates to measurable products. In this section, utilized kinetic functions are outlined and formulated reaction rates discussed.

Kinetic functions

Reaction rates r, often referred to as μ in biotechnology, can be described by the maximum rate $\mu_{\max}$ and the product of kinetic functions

$$\mu = \mu_{\max} \prod g_i. \tag{5.42}$$

The values of kinetic functions g_i are calculated by kinetic parameters k_i and substrate (or compound) concentrations c_i. They should adopt values between 0 and 1 so that the reaction rate μ cannot exceed its maximum value and is not less than zero. Therefore, some kinetic functions that occur in the process models of this thesis had to be normalized. Table 5.3 lists all kinetic functions used in this thesis along with their normalization term (if required). On a qualitative level, it is possible to distinguish between limiting and inhibiting dependencies

between substrates and reaction rates. The Michealis-Menten (MiMe) kinetic describes production for a limiting substrate, meaning, the production rate grows with increasing substrate availability. An inhibiting relation for one single substrate is expressed by an Aiba (Ai) kinetic. Both, limiting and inhibiting effects of the same substrate are realized by the kinetic functions Specific$_1$ (Ro$_1$), which employs one parameter, and Specific$_2$ (Ro$_2$) exhibiting two parameters, as initially developed by Rossner (2014). The Moser kinetic (Mo) equals an approximated step function. It is implemented to describe limitations that steeply increase once a critical substrate concentration is reached. Additionally, a new specific kinetic function (Spec) is proposed to approximate the inhibitory effects of two substrates on the reaction rate by returning the maximum of two Aiba kinetics

$$g_{\mathrm{Spec}} = \max\left(g_{\mathrm{Ai}}(c_1, k_1), g_{\mathrm{Ai}}(c_2, k_2)\right). \tag{5.43}$$

The maximum-function of eq. (5.43) was approximated because the employed simulation software tool Advanced Batch Control, described in Herold et al. (2017), requires differentiable expressions. The approximation strategy is explained in the following. A maximum of two values is defined as the mean plus half of the absolute difference

$$\max\left(g_{\mathrm{Ai}}(c_1, k_1), g_{Ai}(c_2, k_2)\right) = \frac{g_{\mathrm{Ai}}(c_1, k_1) + g_{\mathrm{Ai}}(c_2, k_2) + |g_{\mathrm{Ai}}(c_1, k_1) - g_{\mathrm{Ai}}(c_2, k_2)|}{2}. \tag{5.44}$$

In principle, the absolute term above can be reformulated by first squaring and then calculating the square root. However, numerical inaccuracies might lead to imaginary numbers when calculating the square root that cannot be evaluated by the software tool. Hence, the root function was approximated with the curve fitting toolbox from `MATLAB` by a third degree polynominal yielding

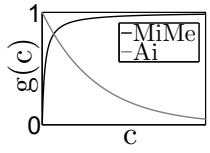
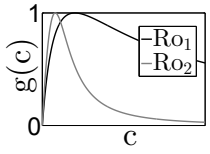
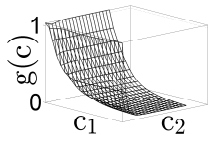
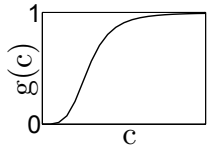
$$f(x) = \frac{12}{5}x - \frac{58}{20}x^2 + \frac{3}{2}x^3 + \frac{34}{500}. \tag{5.45}$$

For this approximation, the root function was only considered in the range from 0 to 1, since the absolute term to be approximated depends on kinetic functions that cannot adopt other values. To obtain a more accurate approximation at low substrate concentrations, the interval was first divided equidistantly in steps of 0.1 and then the obtained values were squared. At these discrete instances the root function was evaluated and resulting points were then used for the polynomial approximation. Replacing the absolute term of eq. (5.44) by eq. (5.45) with the argument $x = [g_{\mathrm{Ai}}(c_1, k_1) - g_{\mathrm{Ai}}(c_2, k_2)]^2$ results in a software-compliant

description for the specific kinetic

$$g_{\mathrm{Spec}} = \frac{1}{2\,\mathrm{e}^{c_1\,k_1}} + \frac{1}{2\,\mathrm{e}^{c_2\,k_2}} + \frac{6\left(\mathrm{e}^{-c_1\,k_1} - \mathrm{e}^{-c_2\,k_2}\right)^2}{5} - \frac{29\left(\mathrm{e}^{-c_1\,k_1} - \mathrm{e}^{-c_2\,k_2}\right)^4}{20} + \\ + \frac{3\left(\mathrm{e}^{-c_1\,k_1} - \mathrm{e}^{-c_2\,k_2}\right)^6}{4} + \frac{17}{500}\,. \tag{5.46}$$

Table 5.3: List of used kinetic functions $g(c_i)$ that depend on substrate concentrations c_i and constant parameters k_i.

Name, Abbreviation	Kinetic function $g(c)$	Normalization term $g_{\max}^{-1}$	Type	Trend curve
Michaelis–Menten, MiMe(c, k)	$\frac{c}{c+k}$	-	lim	
Aiba, Ai(c, k)	e^{-kc}	-	inh	
Specific$_1$, Ro$_1$(c, k)	$\frac{g_{\max}^{-1}\cdot c}{1+c+\left(\frac{c}{k}\right)^2}$	$1+\frac{2}{k}$	lim, inh	
Specific$_2$, Ro$_2$(c, k_1 k_2)	$\frac{g_{\max}^{-1}\cdot c}{1+c+\left(\frac{c}{k_1}\right)^{k_2}}$	$1+\frac{k_2(k_2-1)^{\frac{1}{k_2}}}{(k_2-1)k_1}$	lim, inh	
Specific, Spec(c_1 c_2, k_1 k_2)	$\max(\mathrm{e}^{-k_1\cdot c_1}, \mathrm{e}^{-k_2\cdot c_2})$	-	inh	
Moser, Mo(c, k_1 k_2)	$\frac{c^{k_1}}{c^{k_1}+k_2}$	-	step	

The kinetic functions introduced were used to describe the reaction rates that determine growth. Modeling of these reaction rates was an iterative process in which the hypotheses were constantly examined and compared with the data from cultivations. For most reaction rates, the initial hypotheses were based on biological considerations. Exception was the biomass formation rate μ_{X} proposed by Rossner (2014), which was first assumed, then examined and modified. Examination in this context means that functions with kinetic parameters that cannot be estimated due to limited experimental data were removed or

replaced by other kinetic functions. The relevant and final reaction rates for each state are explained in the following paragraphs.

Biomass production

R. e. produces active biomass with the growth rate μ_{X}, which depends on the concentrations of the substrates needed for growth and on the presence of inhibitors

$$\begin{aligned}\mu_{\mathrm{X}} =&\mu_{\mathrm{X,max}} \cdot \mathrm{MiMe}\left(c_{\mathrm{N}}, k_{\mathrm{N}}^{\mathrm{X}}\right) \cdot \mathrm{MiMe}\left(c_{\mathrm{P}}, k_{\mathrm{P}}^{\mathrm{X}}\right) \cdot \mathrm{MiMe}\left(c_{\mathrm{H_2,l}}, k_{\mathrm{H_2}}^{\mathrm{X}}\right) \cdot \\ &\cdot \mathrm{Ro_2}\left(c_{\mathrm{O_2,l}}, k_{1,\mathrm{O_2}}^{\mathrm{X}}\ k_{2,\mathrm{O_2}}^{\mathrm{X}}\right) \cdot \mathrm{Ro_1}\left(c_{\mathrm{CO_2,l}}, k_{\mathrm{CO_2}}^{\mathrm{X}}\right). \end{aligned} \tag{5.47}$$

Nutrients required for growth are ammonium, phosphate, hydrogen, oxygen and carbon dioxide as well as iron together with trace elements. Iron and trace elements are assumed to be available in sufficient amounts at all times and therefore not affecting the growth rate. Hydrogen has a positive effect on μ_{X}, i.e., increasing dissolved hydrogen concentration leads to faster growth. In consequence, $c_{\mathrm{H_2,l}}$ is regarded as limiting and follows $\mathrm{MiMe}(c_{\mathrm{H_2,l}}, k_{\mathrm{H_2}}^{\mathrm{X}})$. Also N and P are limiting and modeled in a similar fashion via the terms $\mathrm{MiMe}(c_{\mathrm{N}}, k_{\mathrm{N}}^{\mathrm{X}})$ and $\mathrm{MiMe}(c_{\mathrm{P}}, k_{\mathrm{P}}^{\mathrm{X}})$, even though it is reasonable to assume that very high salt concentrations inhibit growth. However, during the conducted cultivations an inhibiting effect was not observed for c_{N} up to $4\,\mathrm{g\,L^{-1}}$ and c_{P} up to $5.5\,\mathrm{g\,L^{-1}}$. Dissolved carbon dioxide and oxygen are other required substrates. Oxygen, as an electron acceptor, is involved in energy production and carbon dioxide is fixed anabolically. In consequence, both gases are of limiting nature. On the other hand, high oxygen concentrations are known to inhibit the activity of the enzyme hydrogenase, which is involved in energy production, see Ludwig et al. (2009). Inhibitory effects at large concentrations have been described by Shang et al. (2003) for carbon dioxide as well. A limiting and inhibiting effect of one substrate is approximated by the first and second specific kinetic functions. For oxygen, the specific two-parameter function $\mathrm{Ro_2}(c_{\mathrm{O_2,l}}, [k_{1,\mathrm{O_2}}^{\mathrm{X}}\ k_{2,\mathrm{O_2}}^{\mathrm{X}}])$ was selected. Two parameters allow for a highly dynamic response towards oxygen variations. For carbon dioxide, in contrast, the specifc function exhibiting one parameter $\mathrm{Ro_1}(c_{\mathrm{CO_2,l}}, k_{\mathrm{CO_2}}^{\mathrm{X}})$ sufficed, because the cultivation data indicated that the growth rate is less sensitive towards fluctuations of $c_{\mathrm{CO_2,l}}$.

PHB production and degradation

The PHB formation rate reads as follows

$$\begin{aligned}\mu_{\mathrm{PHB}} =&\mu_{\mathrm{PHB,max}} \cdot \mathrm{Spec}(c_{\mathrm{N}}, c_{\mathrm{P}}, k_{\mathrm{N}}^{\mathrm{PHB}}, k_{\mathrm{P}}^{\mathrm{PHB}}) \cdot \mathrm{Ai}\left(x_{\mathrm{PHB}}, k_{\mathrm{xPHB}}^{\mathrm{PHB}}\right) \cdot \\ &\cdot \mathrm{MiMe}\left(c_{\mathrm{H_2,l}}, k_{\mathrm{H_2}}^{\mathrm{X}}\right) \cdot \mathrm{Ro_2}\left(c_{\mathrm{O_2,l}}, k_{1,\mathrm{O_2}}^{\mathrm{X}}\ k_{2,\mathrm{O_2}}^{\mathrm{X}}\right) \cdot \mathrm{Ro_1}\left(c_{\mathrm{CO_2,l}}, k_{\mathrm{CO_2}}^{\mathrm{X}}\right). \end{aligned} \tag{5.48}$$

It is known that *R. e.* produces the carbon storage polymer PHB when phosphate respectively ammonium limitation occurs, compare Raje and Srivastava (1998) and Ryu et al. (1997). A stimulating effect of low salt concentrations on the PHB production rate μ_{PHB} was implemented by two Aiba kinetic functions $\mathrm{Ai}(c_{\mathrm{N}}, k_{\mathrm{N}}^{\mathrm{PHB}})$ and $\mathrm{Ai}(c_{\mathrm{P}}, k_{\mathrm{P}}^{\mathrm{PHB}})$, the outputs of which were evaluated in the Specific function eq. (5.43). Mozumder et al. (2015) postulated a decreasing μ_{PHB} at high cellular PHB content. In consequence, an inhibiting effect of the present PHB fraction

$$x_{\mathrm{PHB}} = \frac{c_{\mathrm{PHB}}}{m_{\mathrm{X}}/V_{\mathrm{l}} + c_{\mathrm{PHB}}} \tag{5.49}$$

on μ_{PHB} was implemented by an Aiba function $\mathrm{Ai}(x_{\mathrm{PHB}}, k_{\mathrm{xPHB}}^{\mathrm{PHB}})$, with $m_{\mathrm{X}}/V_{\mathrm{l}}$ being the concentration of active biomass.

In autotrophic metabolism, carbon dioxide is primarily fixed through the Calvin-Benson-Bassham (CBB) pathway as described in Park et al. (2011) and then either processed to biomass or to the storage polymer PHB. Required energy to run the CBB is derived by oxidizing hydrogen to water. We assumed that the rate-limiting steps for PHB and biomass production are at the beginning of the CBB metabolic pathway. If it were different, metabolic intermediates would accumulate, similar to acid production in overflow-metabolism of *E. coli* or ethanol production induced by the Crabtree effect in yeast, which has not been observed. Since the CBB pathway is needed for PHB and biomass formation, kinetic functions describing the effects of dissolved gases on the PHB production rate are assumed to be similar as for μ_{X}. Consequently, all kinetic functions and parameters regarding dissolved gases are directly copied from μ_{X} to formulate μ_{PHB}.

According to Schlegel et al. (1961) and Bartha (1962), assimilated PHB can be converted to biomass and for this we formulated the conversion rate

$$\begin{aligned}\mu_{\mathrm{X,PHB}} =& \mu_{\mathrm{X,PHB,max}} \cdot \mathrm{Ai}\left(c_{\mathrm{CO_2,l}}, k_{\mathrm{CO_2}}^{\mathrm{X,PHB}}\right) \cdot \\ & \cdot \mathrm{MiMe}\left(c_{\mathrm{N}}, k_{\mathrm{N}}^{\mathrm{X,PHB}}\right) \cdot \mathrm{MiMe}\left(c_{\mathrm{P}}, k_{\mathrm{P}}^{\mathrm{X,PHB}}\right) \cdot \mathrm{MiMe}\left(x_{\mathrm{PHB}}, k_{\mathrm{PHB}}^{\mathrm{X,PHB}}\right).\end{aligned} \tag{5.50}$$

Degradation of PHB is active, once dissolved carbon dioxide is absent and ammonium as well as phosphate are present. Accordingly, an Aiba kinetic for carbon dioxide $\mathrm{Ai}(c_{\mathrm{CO_2,l}}, k_{\mathrm{CO_2}}^{\mathrm{X,PHB}})$ and Michaelis–Menten kinetics for N and P, $\mathrm{MiMe}(c_{\mathrm{N}}, k_{\mathrm{N}}^{\mathrm{X,PHB}})$ and $\mathrm{MiMe}(c_{\mathrm{P}}, k_{\mathrm{P}}^{\mathrm{X,PHB}})$, were implemented. The conversion of the storage polymer is only possible if sufficient PHB has accumulated. A premise that is realized by a limiting kinetic $\mathrm{MiMe}(x_{\mathrm{PHB}}, k_{\mathrm{xPHB}}^{\mathrm{X,PHB}})$ depending on the present fraction of PHB. When hydrogen respectively oxygen are absent, PHB is degraded and used as energy source. Since this case has not been investigated yet, it is not included in the present model.

MBH expression

MBH is assumed to be produced partly growth-associated, more precisely proportional to μ_{X} by the factor $K_{\mathrm{MBH,X}}$ that is to be introduced with the state in eq. (5.68). Additionally, a formation rate μ_{MBH} was implemented that is highest at an optimal dissolved oxygen concentration that is realized by a specific kinetic with two parameters

$$\mu_{\mathrm{MBH}} = \mu_{\mathrm{MBH,max}} \cdot \mathrm{Ro}_2 \left(c_{\mathrm{O_2,l}}, k^{\mathrm{X}}_{1,\mathrm{O_2}}\ k^{\mathrm{X}}_{2,\mathrm{O_2}} \right). \tag{5.51}$$

According to Cracknell et al. (2009), very large oxygen concentrations not only slow down the MBH production by affecting μ_{MBH}, but also degrade or inactivate MBH. Thus, the degradation rate depends on oxygen in a limiting fashion and is modeled by

$$\mu_{\mathrm{MBH,deg}} = \mu_{\mathrm{MBH,deg,max}} \cdot \mathrm{MiMe} \left(c_{\mathrm{O_2,l}}, k^{\mathrm{MBH}}_{\mathrm{deg}} \right). \tag{5.52}$$

SH production and degradation

Similar to μ_{X} and μ_{MBH}, we assumed that the SH formation rate μ_{SH} depends on dissolved oxygen. At an optimal oxygen level, the SH expression is highest, which is realized by a two-parameter specific function

$$\mu_{\mathrm{SH}} = \mu_{\mathrm{SH,max}} \cdot \mathrm{Ro}_2 \left(c_{\mathrm{O_2,l}}, k^{\mathrm{SH}}_{1,\mathrm{O_2}}, k^{\mathrm{SH}}_{2,\mathrm{O_2}} \right). \tag{5.53}$$

If the oxygen concentration is very high, similar to MBH in eq. (5.52), not only the SH formation rate decreases, but also the degradation of SH takes place. This is expressed by a Michaelis–Menten function in

$$\mu_{\mathrm{SH,deg}} = \mu_{\mathrm{SH,deg,max}} \cdot \mathrm{MiMe} \left(c_{\mathrm{O_2,l}}, k^{\mathrm{SH}}_{\mathrm{deg}} \right). \tag{5.54}$$

Inhibition caused by oxygen

Experimental data showed that long-term exposures to high dissolved oxygen concentrations may affect the formation of biomass by decreasing μ_{X} and μ_{PHB}. As a mechanistic explanation could not be given, a relation between high oxygen and decreased rates was therefore modeled by introducing a fictitious inhibitory state I_{In}, quantifying the amount of inhibitory action. The formation rate of inhibitory action r_{In} increases strongly once the oxygen concentration $c_{\mathrm{O_2,l}}$ reaches a certain level that is approximately equivalent to an oxygen partial pressure of atmospheric air ($p_{\mathrm{O_2,l}}$=100 %). Hence, a Moser kinetic function was utilized for implementation, which yields

$$r_{\mathrm{In}} = k_{\mathrm{In,max}} \cdot \mathrm{Mo}(c_{\mathrm{O_2,l}}, k^{\mathrm{In}}_{1,\mathrm{O_2}}, k^{\mathrm{In}}_{2,\mathrm{O_2}}). \tag{5.55}$$

5.2.3 Gas solubility

The dynamics of the present autotrophic system are not only determined by the reaction rates, but are also influenced by gas transport. This has been modeled as presented in Rossner (2014). Gas transport via the liquid-gas interface is indicated by gray arrows in Figure 5.5 and assuming no consumption, the differential equation for dissolved gas reads

$$\dot{m}_{\mathrm{gas,l}} = \dot{m}_{\mathrm{trans,gas}}. \tag{5.56}$$

Driving force for gas transport $\dot{m}_{\mathrm{trans,gas}}$ is the concentration difference between saturation and present dissolved gas. This is multiplied by a gas-specific volumetric gas transfer coefficient $\mathrm{k_L a_{gas}}$ and the liquid volume V_{l}, yielding

$$\dot{m}_{\mathrm{trans,gas}} = \mathrm{k_L a_{gas}} \cdot (c_{\mathrm{gas,sat}} - c_{\mathrm{gas,l}}) \cdot V_{\mathrm{l}}. \tag{5.57}$$

Henry's law determines the saturation concentration by multiplying the temperature compensated coefficient $H_{\mathrm{gas,30}}$, the partial pressure of the component in the gas phase $p_{\mathrm{gas,v}}$ and the molar mass M_{gas}

$$c_{\mathrm{gas,sat}} = H_{\mathrm{gas,30}} \cdot p_{\mathrm{gas,v}} \cdot M_{\mathrm{gas}}, \tag{5.58}$$

with $p_{\mathrm{gas,v}} = x_{\mathrm{gas,v}} \cdot P$. Although $\mathrm{k_L a_{gas}}$ of eq. (5.57) was assumed to be constant for this thesis, a dynamic description for $\mathrm{k_L a_{gas}}$ should be used for future cultivations so that the stirring speed can be varied as described in Section 5.1.2.

5.2.4 Evolution of states

The mass of substance in a fed-batch bioprocess changes because of inflows, conversions and outflows in terms of sampling. Here, samples were considered in discrete sampling compensations, and thus are not included as outflows in the ordinary differential equation (ODE) system. Discrete sampling compensation means that the ODE system is integrated up to the point in time of sampling. Then, the states are corrected for the amount of taken sample and passed as new state vector to the ODE solver to continue integration. In this section, the differential equations of cell compounds (X, PHB, MBH, SH), dissolved substrates (P, N, H_2, CO_2, O_2) and volumes (V_{l}, V_{BaAc}) are explained. But first the inhibitory state I_{In} describing inhibition of μ_{X} caused by long-term oxygen exposures is introduced.
It was observed that dissolved partial pressures of oxygen higher than those of air in equilibrium ($p_{\mathrm{O_2,l}} = 100$ %), led to reduced maximum growth rates not only for that particular time instance but also in future. An effect of exposure to high oxygen on the maximum production rate of PHB metabolism was also observed.
These reduced rates were modeled with the help of the fictitious inhibitory state I_{In}, which

can increase with the rate r_{In} from eq. (5.55) and decrease proportional to the production of active biomass with factor $K_{\mathrm{In,deg}}$. This can be expressed through

$$\dot{I}_{\mathrm{In}} = r_{\mathrm{In}} - K_{\mathrm{In,deg}} \cdot \dot{m}_{\mathrm{X}}. \tag{5.59}$$

An increase is caused by high $c_{\mathrm{O_2,l}}$ levels and a reduction can be explained because new (reproduced) cells have not been exposed to high oxygen and consequently are not affected. To compensate for I_{In}, the maximum rates, e.g., $\mu_{\mathrm{X,max}}$, are lowered accordingly

$$\mu_{\mathrm{X,max,In}} = \mu_{\mathrm{X,max}} - I_{\mathrm{In}} \cdot \mu_{\mathrm{X,max}}. \tag{5.60}$$

In a similar manner, $\mu_{\mathrm{PHB,max}}$ is corrected to

$$\mu_{\mathrm{PHB,max,In}} = \mu_{\mathrm{PHB,max}} \cdot (1 - I_{\mathrm{In}}), \tag{5.61}$$

and thus the following state equations for PHB and biomass do not depend on the maximum rates from eq. (5.47) and (5.48), but on those from eq. (5.60) and (5.61), instead and read

$$\begin{aligned}\mu_{\mathrm{X,In}} =& \mu_{\mathrm{X,max}} \cdot (1 - I_{\mathrm{In}}) \cdot \mathrm{MiMe}\left(c_{\mathrm{N}}, k_{\mathrm{N}}^{\mathrm{X}}\right) \cdot \mathrm{MiMe}\left(c_{\mathrm{P}}, k_{\mathrm{P}}^{\mathrm{X}}\right) \cdot \mathrm{MiMe}\left(c_{\mathrm{H_2,l}}, k_{\mathrm{H_2}}^{\mathrm{X}}\right) \cdot \\ & \cdot \mathrm{Ro_2}\left(c_{\mathrm{O_2,l}}, k_{1,\mathrm{O_2}}^{\mathrm{X}}\ k_{2,\mathrm{O_2}}^{\mathrm{X}}\right) \cdot \mathrm{Ro_1}\left(c_{\mathrm{CO_2,l}}, k_{\mathrm{CO_2}}^{\mathrm{X}}\right) \end{aligned} \tag{5.62}$$

$$\begin{aligned}\mu_{\mathrm{PHB,In}} =& \mu_{\mathrm{PHB,max}} \cdot (1 - I_{\mathrm{In}}) \cdot \mathrm{Spec}(c_{\mathrm{N}}, c_{\mathrm{P}}, k_{\mathrm{N}}^{\mathrm{PHB}}, k_{\mathrm{P}}^{\mathrm{PHB}}) \cdot \mathrm{Ai}\left(x_{\mathrm{PHB}}, k_{\mathrm{xPHB}}^{\mathrm{PHB}}\right) \cdot \\ & \cdot \mathrm{MiMe}\left(c_{\mathrm{H_2,l}}, k_{\mathrm{H_2}}^{\mathrm{X}}\right) \cdot \mathrm{Ro_2}\left(c_{\mathrm{O_2,l}}, k_{1,\mathrm{O_2}}^{\mathrm{X}}\ k_{2,\mathrm{O_2}}^{\mathrm{X}}\right) \cdot \mathrm{Ro_1}\left(c_{\mathrm{CO_2,l}}, k_{\mathrm{CO_2}}^{\mathrm{X}}\right). \end{aligned} \tag{5.63}$$

To prevent negative entries for corrected maximum rates, I_{In} is restricted to values between 0 and 1. If I_{In} equals 1, no growth is notable. We also observed that the uptake of salts was affected by lowered reaction rates, but the gas consumption was not. Apparently, the cells consume the same amount of gas/potential energy, but produce less biomass so that the cell growth efficiency is decreased. It is assumed that the conversion rate of PHB to active biomass $\mu_{\mathrm{X,PHB}}$ is also affected by the overall inhibition. However, experiments were not run aiming at both, conversion of PHB and rate limitation caused by persistently high oxygen leading to $I_{\mathrm{In}} > 0$. Thus, the rate $\mu_{\mathrm{X,PHB}}$ is not corrected for I_{In} in this thesis.
The major cell compound, active biomass (m_{X}), is built when cells grow, which is determined by the compensated growth rate ($\mu_{\mathrm{X,In}}$)

$$\dot{m}_{\mathrm{X}} = \mu_{\mathrm{X,In}} \cdot m_{\mathrm{X}} + (U_{\mathrm{N,X,PHB}} + U_{\mathrm{P,X,PHB}} + 1) \cdot \mu_{\mathrm{X,PHB}} \cdot m_{\mathrm{X}}. \tag{5.64}$$

The second term on the right side of eq. (5.64) approximates a conversion of PHB to active biomass. It is governed by the rate $\mu_{\mathrm{X,PHB}}$, depends on the amount of active biomass m_{X} and is connected to an uptake of phosphate (P) and ammonium (N). Specific substrate uptake coefficients $U_{\mathrm{N,X,PHB}}$ and $U_{\mathrm{P,X,PHB}}$ describe how much N or P is required in order to convert one gram of PHB.
PHB, as a carbon storage polymer, is the second most important cell compound. It is built in proportion to active biomass with the rate $\mu_{\mathrm{PHB,In}}$ and it is degraded through conversion to biomass with the rate $\mu_{\mathrm{X,PHB}}$:

$$\dot{m}_{\mathrm{PHB}} = \mu_{\mathrm{PHB,In}} \cdot m_{\mathrm{X}} - \mu_{\mathrm{X,PHB}} \cdot m_{\mathrm{X}}. \tag{5.65}$$

In cultivations, when comparing measurements and simulations, we observed a difference between the measured biomass concentration and the sum of simulated PHB and active biomass, while the values for PHB and OD were congruent to their measurements. Hypotheses are discussed below to explain our observation before giving an additional state equation to compensate for these discrepancies.
One hypothesis was that the analysis of PHB did not isolate the total fraction of PHB from the cells, meaning that the actual PHB values were higher than those measured (subsript "meas").

$$c_{\mathrm{PHB}} = c_{\mathrm{PHB,meas}} + c_{\mathrm{PHB,additional}} \tag{5.66}$$

If we had corrected all PHB measurements for this loss based on the analysis or inserted the corresponding equation above containing the additional PHB amount, the description of the PHB formation and the calculation of the OD in the model would have had to be changed to remain congruent with the measurements. However, the equation for OD introduced in Section 5.2.5 and its identified parameters in Section 5.2.6 are well compatible with the literature, so this hypothesis was not pursued further. Another hypothesis was that the second component of the measured biomass, i.e., the actual amount of active biomass, was higher than predicted by the model. However, since this would affect both phases with and without PHB formation, but the discrepancies only occurred in phases with PHB, this hypothesis was rejected. A third hypothesis was based on the assumption that an additional (inactive) compound proportional to PHB was produced, which had an influence on the measured biomass. However, by implementing a suitable factor, the measured biomass concentration could not yet be simulated because the discrepancy between measurements and simulation increased with increasing PHB content. The last hypothesis, which later led to the modeling of the granulate state, also assumed an additional substance. It is known from literature that PHB is stored at the cellular level in so-called granules, which were analysed by Gebauer (2009) and Beeby et al. (2012). These granules are coated with non-PHB molecules, e.g., proteins, which were considered in the model with a corresponding state ($m_{\mathrm{PHB,gr}}$). In experiments it was observed that this component must increase particularly

strongly at high PHB contents, i.e., its formation was assumed to be proportional to m_{PHB}. In order to limit the number of new parameters to the minimum necessary, it was assumed that the production rate of the granules is proportional to the PHB formation. Therefore, the dynamics of the granulus state of eq. (5.67) depends on the difference of $\mu_{\mathrm{PHB,In}}$ and $\mu_{\mathrm{X,PHB}}$ multiplied by m_{PHB}.

$$\dot{m}_{\mathrm{PHB,gr}} = (\mu_{\mathrm{PHB,In}} - \mu_{\mathrm{X,PHB}}) \cdot m_{\mathrm{PHB}} \cdot F_{\mathrm{PHB,gr}}, \tag{5.67}$$

with $F_{\mathrm{PHB,gr}}$ being a constant factor that needs to be identified. The surface layer of granules was taken into account because it contributed to the measured cell dry weight but since it consists mainly of proteins, it does not affect OD as will be explained below. However, since not much is known about the formation of this granulate surface and it is only relevant in cultivations with PHB formation, no gas consumption or salt uptake was modeled for it.
In contrast to PHB, the influence of membrane-bound hydrogenases on measured total cell dry weight is regarded as negligible. Nonetheless, the amount of MBH (m_{MBH}) is balanced because it represents a target compound. It is assumed that MBH can be build and degrades according to

$$\dot{m}_{\mathrm{MBH}} = K_{\mathrm{MBH,X}} \cdot \dot{m}_{\mathrm{X}} + (\mu_{\mathrm{MBH}} - \mu_{\mathrm{MBH,deg}}) \cdot m_{\mathrm{X}}. \tag{5.68}$$

The MBH expression is dependent on a growth-dependent and a growth-independent term, and is therefore based on $\dot{m}_{\mathrm{X}}$ and $\mu_{\mathrm{MBH}} \cdot m_{\mathrm{X}}$. By contrast, MBH degradation or inactivation is exclusively caused by substrates. Since the concentrations of MBH cannot be measured but instead the activities can be quantified, eq. (5.68) must be converted. Assuming a constant relationship between the total activity A_{MBH} and amount of MBH m_{MBH} via the conversion factor $K_{\mathrm{MBH,a}}$ with units $\mathrm{kU\,g^{-1}}$

$$A_{\mathrm{MBH}} = K_{\mathrm{MBH,a}} \cdot m_{\mathrm{MBH}}, \tag{5.69}$$

the MBH mass balance equation (5.68) becomes eq. (5.70), an expression in the activities

$$\underbrace{K_{\mathrm{MBH,a}} \cdot \dot{m}_{\mathrm{MBH}}}_{\dot{A}_{\mathrm{MBH}}} = K_{\mathrm{MBH,a}} \cdot (K_{\mathrm{MBH,X}} \cdot \dot{m}_{\mathrm{X}} + (\mu_{\mathrm{MBH}} - \mu_{\mathrm{MBH,deg}}) \cdot m_{\mathrm{X}}) \tag{5.70}$$

As the value for $K_{\mathrm{MBH,a}}$ is unknown, after inserting the reaction rates of eq. (5.51) and eq. (5.52) into eq. (5.70), only the products of $K_{\mathrm{MBH,a}} \cdot \mu_{\mathrm{MBH,deg,max}}$ and $K_{\mathrm{MBH,a}} \cdot \mu_{\mathrm{MBH,max}}$ and $K_{\mathrm{MBH,a}} \cdot K_{\mathrm{MBH,X}}$ can be identified, but not the individual parameters.
Similar to MBH, the soluble hydrogenase SH does not contribute to the modeled total cell dry weight. But since it is a target component, SH is also modeled. A mathematical description for SH resulted from experiments that were controlled by multi-model online

Optimal Experimental Design (OED) presented in Neddermeyer et al. (2016). The obtained SH description considers substrate dependent formation and degradation:

$$\dot{m}_{\mathrm{SH}} = (\mu_{\mathrm{SH}} - \mu_{\mathrm{SH,deg}}) \cdot m_{\mathrm{X}}. \tag{5.71}$$

Multiplying the mass balance equation (5.71) with $K_{\mathrm{SH,a}}$, which is the specific conversion factor in $\mathrm{kU\,g^{-1}}$, analogous to $K_{\mathrm{MBH,a}}$ in eq. (5.69), gives an expression in the activities,

$$\dot{A}_{\mathrm{SH}} = K_{\mathrm{SH,a}} \cdot (\mu_{\mathrm{SH}} - \mu_{\mathrm{SH,deg}}) \cdot m_{\mathrm{X}}. \tag{5.72}$$

Gaseous substrates in the liquid phase, namely dissolved hydrogen, carbon dioxide and oxygen, increase with gas transfer $\dot{m}_{\mathrm{trans,gas}}$ into the medium and decrease due to gas consumption for the production of major compartments that are active biomass and PHB:

$$\begin{aligned}\dot{m}_{\mathrm{H_{2,l}}} =& \dot{m}_{\mathrm{trans,H_2}} - U_{\mathrm{H_2,X}} \cdot \mu_{\mathrm{X}} \cdot m_{\mathrm{X}} - U_{\mathrm{H_2,PHB}} \cdot \mu_{\mathrm{PHB}} \cdot m_{\mathrm{X}} - \\ & U_{\mathrm{H_2,X,PHB}} \cdot \mu_{\mathrm{X,PHB}} \cdot m_{\mathrm{X}}\end{aligned} \tag{5.73}$$

$$\dot{m}_{\mathrm{CO_{2,l}}} = \dot{m}_{\mathrm{trans,CO_2}} - U_{\mathrm{CO_2,X}} \cdot \mu_{\mathrm{X}} \cdot m_{\mathrm{X}} - U_{\mathrm{CO_2,PHB}} \cdot \mu_{\mathrm{PHB}} \cdot m_{\mathrm{X}} \tag{5.74}$$

$$\begin{aligned}\dot{m}_{\mathrm{O_{2,l}}} =& \dot{m}_{\mathrm{trans,O_2}} - U_{\mathrm{O_2,X}} \cdot \mu_{\mathrm{X}} \cdot m_{\mathrm{X}} - U_{\mathrm{O_2,PHB}} \cdot \mu_{\mathrm{PHB}} \cdot m_{\mathrm{X}} - \\ & U_{\mathrm{O_2,X,PHB}} \cdot \mu_{\mathrm{X,PHB}} \cdot m_{\mathrm{X}}.\end{aligned} \tag{5.75}$$

Here, the uncompensated rates are employed as mentioned above, because I_{In} did not seem to affect gas uptake. Specific substrate uptake coefficients, e.g., $U_{\mathrm{gas,X}}$, indicate how much substance of gas (in milligram) is required for one gram of biomass. In the conversion of PHB to biomass due to carbon dioxide absence, hydrogen and oxygen provide the required energy, and therefore are added as consumption terms in eq. (5.73) and (5.75).
During cultivation, *R. e.* consumes ammonium for biomass production and the conversion of PHB to biomass, which is implemented in eq. (5.76) employing consumption factors $U_{\mathrm{N,X}}$ and $U_{\mathrm{N,X,PHB}}$. Any uptake of ammonium dedicated to the production of MBH, SH of granule surface is neglectable. Ammonium is also fed at the feed rate u_{N} and feed stock concentration $c_{\mathrm{N,feed}}$ to compensate for the growth-associated uptake in the process, resulting in the state equation

$$\dot{m}_{\mathrm{N}} = -U_{\mathrm{N,X}} \cdot \mu_{\mathrm{X,In}} \cdot m_{\mathrm{X}} - U_{\mathrm{N,X,PHB}} \cdot \mu_{\mathrm{X,PHB}} \cdot m_{\mathrm{X}} + u_{\mathrm{N}} \cdot c_{\mathrm{N,feed}}. \tag{5.76}$$

Similarly to ammonium in (5.76), phosphate is balanced

$$\dot{m}_{\mathrm{P}} = -U_{\mathrm{P,X}} \cdot \mu_{\mathrm{X,In}} \cdot m_{\mathrm{X}} - U_{\mathrm{P,X,PHB}} \cdot \mu_{\mathrm{X,PHB}} \cdot m_{\mathrm{X}} + u_{\mathrm{P}} \cdot c_{\mathrm{P,feed}}. \tag{5.77}$$

Carbon dioxide reacts to carbonic acid when being dissolved. To maintain a set-point pH of 6.8, NaOH or H_2SO_4 (both in a 3 normal concentration) are fed, once carbon dioxide dissolves or gasses out, expressed by the derivative ($\dot{m}_{CO_2,l}$). These carbonic acid associated correction fluid flows are considered by the factor K_{BaAc,pCO_2} as second term in the balance

$$\dot{V}_{BaAc} = \frac{(U_{N,X} \cdot \mu_{X,In} + Y_{N,X,PHB} \cdot \mu_{X,PHB}) \cdot m_X}{18 \frac{g_N}{mol_N} \cdot 3 \frac{mol_{OH^-}}{1000\,mL}} + K_{BaAc,pCO_2} \cdot \dot{m}_{CO_2,l}. \tag{5.78}$$

However, the major reason for a base feed flow is not CO_2 induced, but of metabolic nature. As stated above, Ritchie (2013) reported that in bacteria ammonium is transported into the cells *after* cleaving a proton, which would lower the pH in the fermentation broth. However, pH control counteracts a pH drop by feeding base instead. In consequence, the consumption of ammonium in eq. (5.76) can be expressed as a base feed flow dividing by the molar mass of ammonium ($18\,g_N\,mol_N^{-1}$) and the molarity of NaOH ($3\,mmol_{OH^-}\,mL^{-1}$), which is the first term of eq. (5.78).
The liquid volume changes over time because of liquid feed flows and correction fluids (u_i), which is described in the volume balance with the first and last three summands

$$\begin{aligned}\dot{V}_l = u_N + u_{Fe} + u_P + \frac{1}{\rho_{H_2O,30}} (U_{H_2O,X} \cdot \mu_{X,In} \cdot m_X) - \dot{n}_{H_2O,leak} \cdot M_{H_2O} + \\ + \frac{1}{\rho_{H_2O,30}} \cdot Y_{H_2O,PHB} \cdot \mu_{PHB,In} \cdot m_X + u_{antifoam} + u_{acid} + u_{base}.\end{aligned} \tag{5.79}$$

In addition, the balance also includes a term for water leakage since water evaporates into the headspace and is transferred out by the leakage flow. Besides, water is produced when *R. e.* generates energy by oxidizing hydrogen with oxygen to water. Energy is required in proportion to growth, and hence water production also depends on the compensated cell mass formation rates $\mu_{X,In}$ and $\mu_{PHB,In}$ with $U_{H_2O,PHB} = U_{H_2O,X} = 3.45\ g_{H_2O}\,g_X^{-1}$, originally suggested by Bongers (1970). Due to simplification reasons, for PHB production and formation of active biomass the same substrate consumption coefficient is postulated in this thesis and the conversion of PHB to active biomass is not included in the liquid volume balance in eq. (5.79).

$$\begin{aligned}\underline{x}^T_{\text{model I}} = (&I_{In} \quad m_X \quad m_{PHB} \quad m_{PHB,gr} \quad A_{MBH} \quad A_{SH} \quad m_{H_{2,l}} \quad m_{CO_{2,l}} \quad m_{O_{2,l}} \\ &m_N \quad m_P \quad V_{BaAc} \quad V_l).\end{aligned} \tag{5.80}$$

5.2.5 Measured quantities

In order to compare model simulations with real data, measurement equations need to be defined as not all states of eq. (5.80) are accessible directly. The output vector of the model reads as

$$\begin{aligned}\underline{y}^T_{\text{model I}} =(&c_{\text{X}} \;\; \text{OD} \;\; c_{\text{PHB}} \;\; a_{\text{MBH}} \;\; a_{\text{SH}} \;\; \Delta P \;\; q_{\text{H}_2,\text{v}} \;\; q_{\text{CO}_2,\text{v}} \;\; q_{\text{O}_2,\text{v}} \;\; c_{\text{CO}_2,\text{l}} \;\; p_{\text{O}_2,\text{l}} \\ &c_{\text{N}} \;\; c_{\text{P}} \;\; V_{\text{BaAc}} \;\; V_{\text{l}}),\end{aligned} \tag{5.81}$$

and is obtained below.
The first entry of the measurement vector is the total biomass concentration c_{X}

$$y_1 = c_{\text{X}} = \frac{m_{\text{X}} + m_{\text{PHB}} + m_{\text{PHB,gr}}}{V_{\text{l}}}. \tag{5.82}$$

According to the model, weight relevant cell compounds are active biomass (m_{X}), internal PHB (m_{PHB}) and its granule surfaces ($m_{\text{PHB,gr}}$).
Another method of determining the concentration of biomass is by inspection of the optical density OD of the cultivation broth. The OD grows linearly with the sum of active biomass concentration ($m_{\text{X}}/V_{\text{l}}$) multiplied by a factor $K_{\text{OD,X}}$ and the PHB concentration multiplied by $K_{\text{OD,PHB}}$

$$y_2 = \text{OD} = K_{\text{OD,X}} \cdot \frac{m_{\text{X}}}{V_{\text{l}}} + K_{\text{OD,PHB}} \cdot c_{\text{PHB}}. \tag{5.83}$$

Different factors for active biomass and PHB are employed, because their light scattering properties differ as well (see Wilde (1962)). For this thesis, it was assumed that PHB granules are coated with proteins that affect the total biomass by $m_{\text{PHB,gr}}$. Which molecules make up the surface of PHB granules is not entirely clear but the latest investigations suggest that these are exclusively proteins (see Bresan et al. (2016)). Proteins are measured photometrically at 200–300 nm because at these wavelengths the peptide bonds and aromatic amino acids absorb light (see, e.g., Sizer and Peacock (1947) and Porterfield and Zlotnick (2010)). Based on these investigations, it was assumed that $m_{\text{PHB,gr}}$ does not have an impact on OD that is measured at 436 nm.
The PHB concentration c_{PHB} is the state m_{PHB} divided by V_{l}

$$y_3 = c_{\text{PHB}} = \frac{m_{\text{PHB}}}{V_{\text{l}}}. \tag{5.84}$$

For both hydrogenases, MBH and SH, mass balances were established yielding absolute activity state equations after some reformulations. Since only specific activities a are measurable, the absolute activity of MBH A_{MBH} needs to be divided by the amount of membrane

protein m_{MP}, resulting

$$y_4 = a_{\mathrm{MBH}} = \frac{A_{\mathrm{MBH}}}{m_{\mathrm{MP}}}. \tag{5.85}$$

An empiric model for the amount of membrane protein was determined by Rossner (2014)

$$m_{\mathrm{MP}} = 0.033 \cdot m_{\mathrm{X}}. \tag{5.86}$$

Analogously, A_{SH} is divided by the amount of total proteins m_{Pr}, which has been preliminary analyzed and averaged to $m_{\mathrm{Pr}} = 0.6 \cdot m_{\mathrm{X}}$, yielding

$$y_5 = a_{\mathrm{SH}} = \frac{A_{\mathrm{SH}}}{m_{\mathrm{Pr}}}. \tag{5.87}$$

To calculate the excess pressure ΔP in mbar the environmental pressure P_0 in Pa is subtracted from the measured pressure of the model input vector P in eq. (5.41) and divided by 100 $\mathrm{Pa\,mbar^{-1}}$

$$y_6 = \Delta P = \frac{P - P_0}{100 \frac{\mathrm{Pa}}{\mathrm{mbar}}}. \tag{5.88}$$

Assuming pseudo-steady state for the gas phase as described in Section 5.2.1, measurement equations for the required gas flow rates ($q_{\mathrm{gas,v}}$) in $\mathrm{L\,h^{-1}}$ can be derived. To this end, we postulate that the amount of gas compounds in the headspace does not change, and that the inlet gas flow is solely transferred into the liquid phase or transported out by the leak flow

$$\dot{m}_{\mathrm{leak,gas,v}} = \frac{P}{R \cdot T} \cdot q_{\mathrm{leak,v}} \cdot 10^{-3} \cdot x_{\mathrm{gas,v}} \cdot M_{\mathrm{gas}}, \tag{5.89}$$

where M_{gas} is the molar mass of a gas component. Hence, summarizing the transported gas flow $\dot{m}_{\mathrm{trans,gas}}$, see eq. (5.57), and the leakage flow $\dot{m}_{\mathrm{leak,gas,v}}$ yields

$$y_7 = q_{\mathrm{H_2,v}} = (\dot{m}_{\mathrm{H_2,leak,v}} + \dot{m}_{\mathrm{trans,H_2}}) \cdot \frac{R \cdot T}{M_{\mathrm{H_2}} \cdot P_0} \cdot 1000 \frac{\mathrm{L}}{\mathrm{m^3}} \tag{5.90}$$

$$y_8 = q_{\mathrm{CO_2,v}} = (\dot{m}_{\mathrm{CO_2,leak,v}} + \dot{m}_{\mathrm{trans,CO_2}}) \frac{R \cdot T}{M_{\mathrm{CO_2}} \cdot P_0} \cdot 1000 \frac{\mathrm{L}}{\mathrm{m^3}} \tag{5.91}$$

$$y_9 = q_{\mathrm{O_2,v}} = (\dot{m}_{\mathrm{O_2,leak,v}} + \dot{m}_{\mathrm{trans,O_2}}) \frac{R \cdot T}{M_{\mathrm{O_2}} \cdot P_0} \cdot 1000 \frac{\mathrm{L}}{\mathrm{m^3}}. \tag{5.92}$$

Equations (5.90) to (5.92) summarize leak flow and gas transport across the interface. To convert mass per hour into liters per hour the ideal gas law is utilized. Like all other calculated measured values, these simulated gas flows are used to compare them with the real measured values and thus carry out parameter identifications.
Ammonium and phosphate concentrations (c_{N}, c_{P}) in $\mathrm{g\,L^{-1}}$ are derived from the states and

divided by the liquid volume (V_l)

$$y_{10} = c_\mathrm{N} = \frac{m_\mathrm{N}}{V_\mathrm{l}} \tag{5.93}$$

$$y_{11} = c_\mathrm{P} = \frac{m_\mathrm{P}}{V_\mathrm{l}}. \tag{5.94}$$

The dissolved CO_2 concentration in $\mathrm{mg\,L^{-1}}$ is calculated by converting the state CO_2 in eq. (5.74) into a concentration and multiplying with a factor to match the sensor value units,

$$y_{12} = c_{\mathrm{CO_2,l}} = \frac{m_{\mathrm{CO_2,l}}}{V_\mathrm{l}} \cdot 1000\,\frac{\mathrm{mg}}{\mathrm{g}}. \tag{5.95}$$

Dissolved oxygen is measured in percent, referenced to a state with atmospheric O_2 in equilibrium at cultivation conditions. Therefore, the dissolved O_2 amount state of eq. (5.75) is divided by V_l and the reference concentration according to Henry's law,

$$y_{13} = p_{\mathrm{O_2,l}} = \frac{m_{\mathrm{O_2,l}}/V_\mathrm{l} \cdot 100\,\%}{c_{\mathrm{O_2,sat,air,30}}}. \tag{5.96}$$

Measurement equations (5.97) and (5.98) are identical to the states V_BaAc and V_l, respectively,

$$y_{14} = V_\mathrm{BaAc} \tag{5.97}$$

$$y_{15} = V_\mathrm{l}. \tag{5.98}$$

5.2.6 Identified parameters of the general process model (I)

The number of parameters of the general process model (I) adds up to 48. Forty-six out of the 48 were estimated by optimization. Resulting optimized parameter values are listed in Table 5.7 and 5.8 in this section. Only for $K_\mathrm{X,OD}$ and $K_\mathrm{PHB,OD}$ from the measurement eq. (5.83) for OD the values of a multilinear regression were adopted and retained during all parameter estimations. If these two parameters were estimated together with all the others, unrealistic values would result from multiple correlations as we have found out. The 46 parameters, estimated by optimizations, are too correlated to be identified in a single step. To decrease correlation, parameters affecting metabolism were split into four groups according to their relevance for the production of active biomass, PHB, MBH and SH. The parameters of a group were meant to be estimated with the experiments listed in Table 5.4, leading to four estimation runs. But even after grouping the parameters, the degree of correlation was still too high. To further reduce the number of correlations, some parameters were excluded from the optimization at first. These parameters were set to fixed values that were either reported in the literature or identified by additional experiments. Detailed

Table 5.4: Overview of the data sets used for the model descriptions of active biomass, PHB, MBH and SH. The crosses indicate which experiment was used to identify which parameter group. Cultivations REatc11b–REatc15c were carried out within the thesis of Rossner (2014).

	Active biomass	PHB	MBH	SH
REatc11b	x			
REatc11c				x
REatc12a				x
REatc12c	x			
REatc13a	x	x		
REatc13b	x		x	
REatc13c	x			
REatc14b	x		x	
REatc15c	x			
REatc16	x	x		
REatc17	x	x		
REatc18	x			
REatc19		x		
REatc21		x		
REatc22			x	x
REatc23			x	x
REatc25a			x	x
REatc25b			x	x

information about which parameters were involved and if they had been re-integrated into a subsequent optimization is given in the paragraphs below.
Summarized, four estimation runs were carried out to estimate the parameter values involved in the description of the corresponding compartments, i.e., active biomass, PHB, MBH and SH. Each of the four estimation runs consisted of two optimizations: In the first one, some parameter values were set to fixed values. In the second optimization, all parameters of the parameter group were estimated. In a final run, all parameters of Tables 5.7–5.8, including metabolically non-relevant ones and initially fixed parameters, were estimated together to check for convergence.
In parameter estimation, different measurement quantities were sampled at different frequencies. Therefore, the weighting matrix $\mathbf{W}$ was included in the cost function as mentioned in Chapter 3. It compensated measuring frequencies and measurement tolerances σ in parameter estimation and the following explains how the matrix was created. For manual analysis methods, the tolerances were approximated in a linear manner by

$$\hat{\sigma}_i = \sigma_{0,i} + \theta_{\sigma,i} \cdot \bar{y}_i, \tag{5.99}$$

as described in Chapter 3. With the information of the manufacturers in combination with empirical values, absolute $e_{i,\min}$ and relative tolerances $e_{i,\mathrm{rel}}$ for the online (automated)

sensors could be given. The maximum of the two was then used by evaluating the following equation for each measured value:

$$\hat{\sigma}_i = \max\left(e_{i,\min}, e_{i,\mathrm{rel}}\right) \tag{5.100}$$

Parameters or equations for calculating the tolerances of manual and automated measurements used in parameter identification are listed in Table 5.5. To calculate the weighing

Table 5.5: Absolute and relative tolerances for the measurement vector

Measurement	Symbol	Unit	Absolute tolerance $e_{\min}$	Relative tolerance e_{rel}	Approximated tolerance $\hat{\sigma}$
Biomass concentration	c_{X}	$\mathrm{g\,L^{-1}}$			$10.08 \cdot y_1$
Optical density	OD	-			$0.39 + 0.02 \cdot y_2$
PHB concentration	c_{PHB}	$\mathrm{g\,L^{-1}}$			$0.06 + 0.08 \cdot y_3$
Specific MBH activity	a_{MBH}	$\mathrm{U\,mg_{MP}^{-1}}$			$0.05 + 0.07 \cdot y_4$
Specific SH activity	a_{SH}	$\mathrm{U\,mg_{Pr}^{-1}}$			$0.03 + 0.13 \cdot y_5$
Excess pressure	ΔP	mbar	4		
Volume flow H_2	$q_{\mathrm{H_2,v}}$	$\mathrm{L\,h^{-1}}$	0.5	$0.04 \cdot y_7$	
Volume flow CO_2	$q_{\mathrm{CO_2,v}}$	$\mathrm{L\,h^{-1}}$	0.1	$0.04 \cdot y_8$	
Volume flow O_2	$q_{\mathrm{O_2,v}}$	$\mathrm{L\,h^{-1}}$	0.2	$0.04 \cdot y_9$	
Dissolved CO_2	$c_{\mathrm{CO_2,l}}$	$\mathrm{mg\,L^{-1}}$	1	$0.08 \cdot y_{10}$	
Dissolved O_2	$c_{\mathrm{O_2,l}}$	%	1	$0.06 \cdot y_{11}$	
Ammonium concentration	c_{N}	$\mathrm{g\,L^{-1}}$			$0.03 + 0.07 \cdot y_{12}$
Phosphate concentration	c_{P}	$\mathrm{g\,L^{-1}}$			$0.09 + 0.03 \cdot y_{13}$
Volume correction fluids	V_{BaAc}	mL	3	$0.02 \cdot y_{14}$	
Volume	V_{l}	L	0.2	$0.1 \cdot y_{15}$	

matrix $\mathbf{W}$, measuring tolerance influences among different measurement quantities were neglected so that all off-elements of $\mathbf{W}$ were set to zero. For each measurement vector $\underline{y}$ the diagonal elements of the matrix $\mathbf{W}$ were calculated with

$$\mathbf{W}_{i,i} = \frac{1}{\hat{\sigma}_i} \cdot \frac{N_{y,\max}}{N_{y,i}}, \tag{5.101}$$

where $N_{y,\max}$ is the sample number of the measurement quantity most frequently measured in the entire cultivation and $N_{y,i}$ is the sample number of each individual measurement quantity. In parameter estimation, the first term of eq. (5.101) had to be calculated for each induvidual measurement because it changed with respect to the measured value as described above. In addition, the weighting matrix allowed to exclude incorrect measured values or series of measured values from parameter estimation, for example due to temporarily incorrectly calibrated sensors, by setting the corresponding values in $\mathbf{W}$ to zero. When identifying the parameters, the values that they were allowed to assume during opti-

mization were bounded at the upper and lower ends (ub, lb). All cultivations used for the parameter estimation (autovalidation) carried out in this thesis are presented together with the simulations in the figures of the Appendix C.

Gas transfer rates

During the first parameter identification runs, the constant gas transport coefficients for H_2, CO_2 and O_2 were not optimized. Later, they were included in the parameter estimation. For the initial estimations, the volumetric gas transfer rate for oxygen ($k_La_{O_2}$) was defined as determined by experiments in Rossner (2014). For hydrogen, $k_La_{H_2}$ was calculated according to the film theory that relates diffusion coefficients and volumetric gas transfer rates. It is described in Garcia-Ochoa and Gomez (2009). The $k_La_{CO_2}$ for carbon dioxide could not be calculated with the film theory due to chemical dissolving of CO_2 in the buffered medium. Therefore, additional, organism-free experiments were carried out with the cultivation medium to approximate the gas transfer rate for carbon dioxide, see Section 5.1.1.

Specific consumption coefficients

Morinaga et al. (1978) identified specific consumption coefficients that quantify the uptake of gases involved in biomass production. In their investigations, one gas component in each case was growth-limiting. Bongers (1970) also determined the gas consumption coefficients and additionally specific ammonium uptake coefficients. All cited consumption coefficients are listed in Table 5.6. For PHB formation, the gas uptake rates were investigated by Tanaka

Table 5.6: Consumption coefficients of hydrogen, oxygen, carbon dioxide and ammonium for the production of biomass and PHB in continuous cultures. The values are converted into grams and rounded to two digits.

Literature	Limitation	$U_{H_2,X}$ ($g\,g_X^{-1}$)	$U_{O_2,X}$ ($g\,g_X^{-1}$)	$U_{CO_2,X}$ ($g\,g_X^{-1}$)	$U_{H_2O,X}$ ($g\,g_X^{-1}$)	$U_{N,X}$ ($g\,g_X^{-1}$)
Bongers (1970)		0.44	2.1	1.9	-3.45	0.13
Morinaga et al. (1978)	H_2	0.67	3.3	2.5		
Morinaga et al. (1978)	O_2	0.43	2.9	2.3		
Morinaga et al. (1978)	CO_2	0.59	2.3	1.9		
Ishizaki and Tanaka (1990)		0.41–0.52	1.6–1.8	1.8–2		
		$U_{H_2,PHB}$ ($g\,g_X^{-1}$)	$U_{O_2,PHB}$ ($g\,g_X^{-1}$)	$U_{CO_2,PHB}$ ($g\,g_X^{-1}$)	$U_{H_2O,PHB}$ ($g\,g_X^{-1}$)	$U_{N,PHB}$ ($g\,g_X^{-1}$)
Tanaka et al. (1995)		0.77	4.5	2.1	-6.25	-

et al. (1995). When active biomass and PHB are produced, water is generated that is taken into account by a negative consumption rate. Values for hydrogen uptake rates, for active

biomass and for PHB production, were set to averaged values from Table 5.6 during the first parameter estimation step.

Kinetic parameters

In Ludwig et al. (2009) hydrogenases were exposed to different levels of oxygen. According to their results the hydrogenase in vitro maintains an activity of 75 % exposed to 30 % atmospheric oxygen. The parameters for the specific kinetic function Ro_2 ($k^{\mathrm{X}}_{1,\mathrm{O}_2}$, $k^{\mathrm{X}}_{2,\mathrm{O}_2}$) were set randomly so that $g(p_{\mathrm{O}_2,\mathrm{l}} = 30\ \%) = 0.75$ for the first parameter identifications. In the later parameter estimations, they also were allowed to vary.
All CO_2 is fixed in Ralstonia via the Calvin-Benson-Bassham (CBB) cycle as described by Bowien and Schlegel (1981). The enzymes involved in energy production with oxyhydrogen and CO_2 fixation are supposed to be the limiting step in metabolism. This is assumed for all kind of metabolites, also for PHB that is formed from Acetyl-Coenzyme A as a product of the CBB, see Pohlmann et al. (2006) and Lee et al. (2009). Therefore, parameter values of the gaseous kinetic functions describing the formation rate of PHB are identical with the ones for PHB-free growth. If the fixation of CO_2 and the generation of energy with oxyhydrogen were not the bottlenecks of velocity in the metabolic pathways, metabolic intermediates would accumulate and that does not happen as far as we know from our analysis and the literature.

Other Parameters

PHB has, as a fat compound, light scattering properties, as reported by Wilde (1962). In this thesis, this property is addressed by employing a multi-linear dependency among the inputs c_{PHB}, active biomass concentration and the output OD in eq. (5.83) on page 84. It is assumed that PHB and active biomass are not correlated, i.e., c_{PHB} can rise and active biomass remains constant. For uncorrelated inputs, multiple linear regression based on least squares is a suitable tool to obtain the above mentioned factors $K_{\mathrm{OD,X}}$ and $K_{\mathrm{OD,PHB}}$. Regression results are depicted in Figure 5.6. Since the active biomass concentration is not measured directly, it is approximated by the difference $c_{\mathrm{X}} - c_{\mathrm{PHB}}$, and thus calculated for each sample. This approximation is not entirely correct as in samples with high PHB content the granule layer should also be subtracted since it is assumed to affect the measured biomass but not the OD. For this work, though, there was no measurement method for the granule layer. However, this is only relevant in cultures with high PHB and even there accounts for a maximum of 5 % of the total biomass in the simulation with process model (I). Therefore, when calculating the active biomass, the amount of granules was neglected, knowing that some calculated values of the active biomass are subject to small errors, which can be regarded as synthetic measurement errors.
In total, 181 samples were analyzed for c_{X}, c_{PHB} and OD, and served for the regression

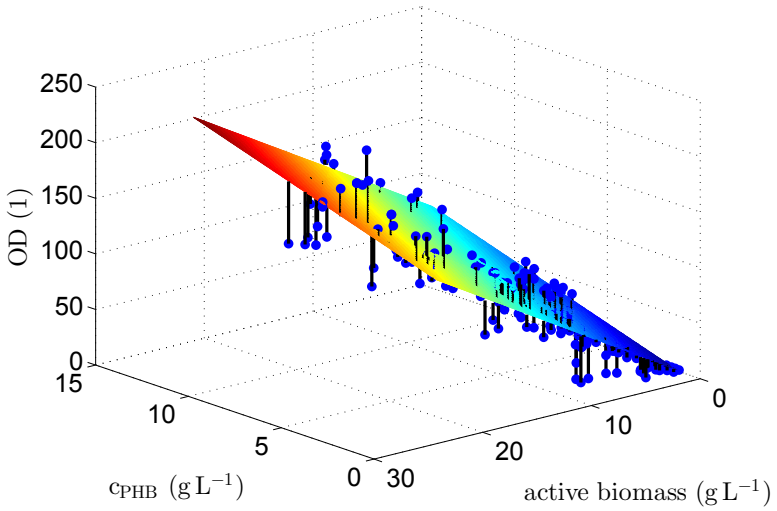


Figure 5.6: Linear relationship between OD and c_{PHB} and active biomass approximated by a plane. Data points are given in blue and residues are indicated with black bars.

analysis leading to values of 5.8 $\mathrm{L\,g^{-1}}$ and 6.2 $\mathrm{L\,g^{-1}}$ with relative standard deviations of 4.3 % and 14.5 % for the parameters $K_{\mathrm{OD,X}}$ and $K_{\mathrm{OD,PHB}}$, respectively. When the biomass concentration increases by 1, the OD changes by 5.8 but when the PHB content changes by 1 $\mathrm{g\,L^{-1}}$, the OD changes by 6.2 and dividing $K_{\mathrm{OD,PHB}} = 6.2$ by $K_{\mathrm{OD,X}} = 5.8$ yields 1.07. In other words, the light scattering properties of PHB are 1.07 times better than those of active biomass, i.e., cells without PHB. This value is in accordance with the observations of Wilde (1962). They reported that when PHB is produced the extinction increased faster than the cell dry weight by the factor 1.1. The following example illustrates that their observations support our result: A certain biomass concentration $c_{\mathrm{X},0}$ is assumed, which consists exclusively of active cells and with $K_{\mathrm{OD,X}}$ corresponds to an OD of $5.8\,c_{\mathrm{X},0}$. Now it is presumed that the cells exclusively form PHB in a defined time and thus increase the biomass by the factor 0.3. This results in a biomass concentration of $1.3\,c_{\mathrm{X},0}$. After this time interval the OD is calculated according to our results with $5.8\,c_{\mathrm{X},0} + 0.3 \cdot 6.2\,c_{\mathrm{X},0}$. In this example, the OD increases by 0.32 with the formation of PHB, which, similar to Wilde's results, is 1.07 times faster than the increase of biomass concentration.
The identified parameter values for the OD justify modeling the effect of the granule surface $m_{\mathrm{PHB,gr}}$ only for measured biomass, as shown above, because only in this way can the OD parameter values presented in literature be satisfied and the measured values in our cultivations be met. If the granule surface was modeled as a part of PHB, different values for the parameters of OD would have been identified leading to a factor other than ≈ 1.1,

Table 5.7: Estimated parameter values of model (I)—continued on next page, applied lower and upper boundaries (lb, ub) and relative standard deviation (rel. std. d.) calculated with the Fisher information matrix

Parameter name	Unit	Parameter value	lb	ub	rel. std. dev. (%)
$K_{\mathrm{BaAc,pCO_2}}$	$\mathrm{mL\,mg^{-1}}$	0.04	0.016	0.2	0.5
$F_{\mathrm{PHB,gr}}$	-	2.5	0	50	0.5
$\mathrm{k_La_{H_2}}$	$\mathrm{h^{-1}}$	273	30	900	0.1
$\mathrm{k_La_{CO_2}}$	$\mathrm{h^{-1}}$	38	10	580	0.2
$\mathrm{k_La_{O_2}}$	$\mathrm{h^{-1}}$	194	150	800	0.2
$U_{\mathrm{N,X}}$	$\mathrm{g\,g_X^{-1}}$	0.13	0.08	0.3	0.09
$U_{\mathrm{P,X}}$	$\mathrm{g\,g_X^{-1}}$	0.08	0.004	0.3	0.09
$U_{\mathrm{H_2,X}}$	$\mathrm{g\,g_X^{-1}}$	0.4	0.3	0.7	0.1
$U_{\mathrm{CO_2,X}}$	$\mathrm{g\,g_X^{-1}}$	1.4	1.3	4	0.3
$U_{\mathrm{O_2,X}}$	$\mathrm{g\,g_X^{-1}}$	1.6	1.6	3.8	0.3
$U_{\mathrm{N,X,PHB}}$	$\mathrm{g\,g_X^{-1}}$	0.69	0.01	4	1.4
$U_{\mathrm{P,X,PHB}}$	$\mathrm{g\,g_X^{-1}}$	0.46	0.02	2	2.3
$U_{\mathrm{H_2,X,PHB}}$	$\mathrm{g\,g_X^{-1}}$	2.6	0.4	3.6	1.5
$U_{\mathrm{O_2,X,PHB}}$	$\mathrm{g\,g_X^{-1}}$	15.5	3.2	48	1.3
$U_{\mathrm{H_2,PHB}}$	$\mathrm{g\,g_X^{-1}}$	0.45	0.2	0.7	0.3
$U_{\mathrm{CO_2,PHB}}$	$\mathrm{g\,g_X^{-1}}$	2.9	1.3	5.3	0.5
$U_{\mathrm{O_2,PHB}}$	$\mathrm{g\,g_X^{-1}}$	1.7	1.3	6.4	0.4
$\mu_{\mathrm{X,max}}$	$\mathrm{h^{-1}}$	0.35	0.1	0.7	0.9
$k_{\mathrm{H_2}}^{\mathrm{X}}$	$\mathrm{mg\,L^{-1}}$	0.004	$2 \cdot 10^{-4}$	0.61	11
$k_{\mathrm{CO_2}}^{\mathrm{X}}$	$\mathrm{mg\,L^{-1}}$	44.6	4.4	176	0.8
$k_{\mathrm{1,O_2}}^{\mathrm{X}}$	$\mathrm{mg\,L^{-1}}$	1.9	1.6	2.6	0.1
$k_{\mathrm{2,O_2}}^{\mathrm{X}}$	-	2.7	2	8	0.05
$k_{\mathrm{N}}^{\mathrm{X}}$	$\mathrm{g\,L^{-1}}$	0.34	0.1	0.8	2.1
$k_{\mathrm{P}}^{\mathrm{X}}$	$\mathrm{g\,L^{-1}}$	0.72	0.4	1.5	1.2

which was not wanted.

The identified parameters $K_{\mathrm{OD,X}}$ and $K_{\mathrm{OD,PHB}}$ were not allowed to vary in the parameter estimation of the remaining 46 parameters. For simulations with the obtained process model, defining a multilinear dependency for the OD implies the following: In cultivations with increased PHB formation the simulations of OD and measured biomass concentration c_{X} do not develop identically.

Table 5.8: Estimated parameter values of model (I)—continuation, applied lower and upper boundaries (lb, ub) and relative standard deviation (rel. std. dev.) calculated with the Fisher information matrix

Parameter name	Unit	Parameter value	lb	ub	rel. std. dev. (%)
$\mu_{\mathrm{X,PHB,max}}$	h^{-1}	0.20	0.1	2.7	65
$k_{\mathrm{CO_2}}^{\mathrm{X,PHB}}$	$\mathrm{L\,mg}^{-1}$	783	88	1980	3.9
$k_{\mathrm{xPHB}}^{\mathrm{X,PHB}}$	-	9.96	0.01	10	68
$\mu_{\mathrm{PHB,max}}$	h^{-1}	0.11	0	0.85	1.4
$k_{\mathrm{xPHB}}^{\mathrm{PHB}}$	-	3.4	0.1	10	1.3
$k_{\mathrm{P}}^{\mathrm{PHB}}$	$\mathrm{L\,g}^{-1}$	42.9	1	300	1
$k_{\mathrm{N}}^{\mathrm{PHB}}$	$\mathrm{L\,g}^{-1}$	1.8	0.2	20	1
$K_{\mathrm{MBH,a}} \cdot \mu_{\mathrm{MBH,max}}$	$\mathrm{kU\,g}^{-1}\,\mathrm{h}^{-1}$	0.022	0.02	0.4	0.3
$K_{\mathrm{MBH,a}} \cdot K_{\mathrm{MBH,X}}$	$\mathrm{kU\,g}^{-1}$	0.1	0.01	10	160
$k_{1,\mathrm{O_2}}^{\mathrm{MBH}}$	$\mathrm{mg\,L}^{-1}$	2.8	0.64	3.2	< 0.1
$k_{2,\mathrm{O_2}}^{\mathrm{MBH}}$	-	6.8	4	30	4.4
$K_{\mathrm{MBH,a}} \cdot \mu_{\mathrm{MBH,deg,max}}$	$\mathrm{kU\,g}^{-1}\,\mathrm{h}^{-1}$	0.06	0.003	0.3	415
$k_{\mathrm{deg}}^{\mathrm{MBH}}$	$\mathrm{mg\,L}^{-1}$	4.7	0.003	26	21
$k_{\mathrm{In,max}}$	h^{-1}	0.01	0.001	0.9	7.4
$K_{\mathrm{In,deg}}$	g^{-1}	$1.1 \cdot 10^{-5}$	10^{-9}	1	83
$k_{1,\mathrm{O_2}}^{\mathrm{In}}$	$\mathrm{mg\,L}^{-1}$	2.1	0.001	400	19
$k_{2,\mathrm{O_2}}^{\mathrm{In}}$	-	0.58	10^{-6}	400	3.1
$K_{\mathrm{SH,a}} \cdot \mu_{\mathrm{SH,max}}$	$\mathrm{kU\,g}^{-1}\,\mathrm{h}^{-1}$	0.16	0.001	1	0.8
$k_{\mathrm{deg}}^{\mathrm{SH}}$	$\mathrm{mg\,L}^{-1}$	0.064	0.003	32	16
$K_{\mathrm{SH,a}} \cdot \mu_{\mathrm{SH,deg,max}}$	$\mathrm{kU\,g}^{-1}\,\mathrm{h}^{-1}$	0.033	0.001	0.4	3.7
$k_{1,\mathrm{O_2}}^{\mathrm{SH}}$	$\mathrm{mg\ L}^{-1}$	1.25	0.64	3.2	1
$k_{2,\mathrm{O_2}}^{\mathrm{SH}}$	-	5.33	4	10	8.6

5.2.7 Cross-validation

After the parameter identification of the 46 parameters, the obtained general process model (I) was evaluated by means of cross-validation and these results will be shown in the following to highlight model deficiencies. As listed in Table 5.4, data of cultivations REatc22 and REatc23 were only used to estimate the parameters of hydrogenase production, which has no effect on biomass formation. Thus, both cultivations were used to evaluate the growth description (including PHB). The corresponding measurement data and the simulation results of cultivations REatc22, REatc23 and four additional experiments not used for parameter estimation (REatc20, REatc26, REatc33a, REatc33b) are plotted together with the measurement data in Figures 5.7–5.12. Parts of REatc33a and REatc33b were already shown and discussed in Section 4.4.4 because the controller from Section 4.4 adjusted most of the time H_2, O_2 and ΔP in the headspace. In the following figures, the complete cultivations are presented including the reference trajectories $r_{H_2,v}$ and $r_{O_2,v}$ for the periods in which the gas phase controller operated.
Measured V_l is not plotted because these measurements are synthetic: They are based on an internal software routine that considers liquid feed flows u_i and evaporation. Instead of V_l, the inhibitory state I_{In} is depicted to provide a better understanding of changing growth rates. However, water is also produced by growth in the cultivation of *R. e.*, see eq. (5.79) on page 83, and thus slight deviations are to be expected. For this reason, the synthetic volume measurements are weighted with 10 % errors, as listed in Table 5.5, for the parameter estimation that roughly does justice to the increase in volume through growth.
In some experiments used in cross- and autovalidations, e.g., REatc33a and REatc19 (see Appendix C) displayed in Figures 5.11 and C.4, c_X does not agree with the simulations to a satisfactory degree. Hence, the growth model needs refinement. The same holds true for the description of ammonium consumption as measured c_N is in accordance with the simulations in REatc20, REatc23, REatc26, see Figures 5.7, 5.9 and 5.10, respectively, although growth is overestimated. Moreover, almost in all cultivations, the simulations of $q_{gas,v}$ and $p_{O_2,l}$ are a poor approximation of the measurements. To improve this part of the model, a more precise description of the gas transport and consumption must be added to the model.

However, the aim of this work was to develop and test methods for model adaption. To this end, a first model was presented above that describes the autotrophic growth of *R. e.* H16, which can be adapted to mutant strains, although it has some deficiencies and is structurally complex. When adapting this complex model (I) to mutant strains, the parameters must be re-estimated with the data obtained from mutant cultivations. However, estimation is a challenge because due to its complexity, it takes a very long time and depending on the optimizer, the resulting parameter values may be wrong because the optimizer may not have found the global minimum. But the optimizer can be supported by providing start

values for the parameters that are close to the optimum. To find these start parameters, the process model (II) is introduced in the following sections. Its structure is identical to that of model (I), but gas transport is not considered and therefore model (II) is less complex. Thus, it can be utilized to perform faster parameter identifications. As stated above, the resulting estimated parameter values can then be used as starting values for a subsequent optimization with model (I). Since the gas transport is not described by model (II), the dissolved gas concentrations must be directly taken as input because they are not calculated by the model. We could only measure O_2 and CO_2 but not H_2, although we developed a probe for it, which is explained in Section 5.3. However, since the probe did not provide reliable values, a method for data-driven calculation of $c_{H_2,l}$ was developed, which is introduced in Section 5.4 before the process model (II) is described in Section 5.5. Finally, its limitations are pointed out in Section 5.6.

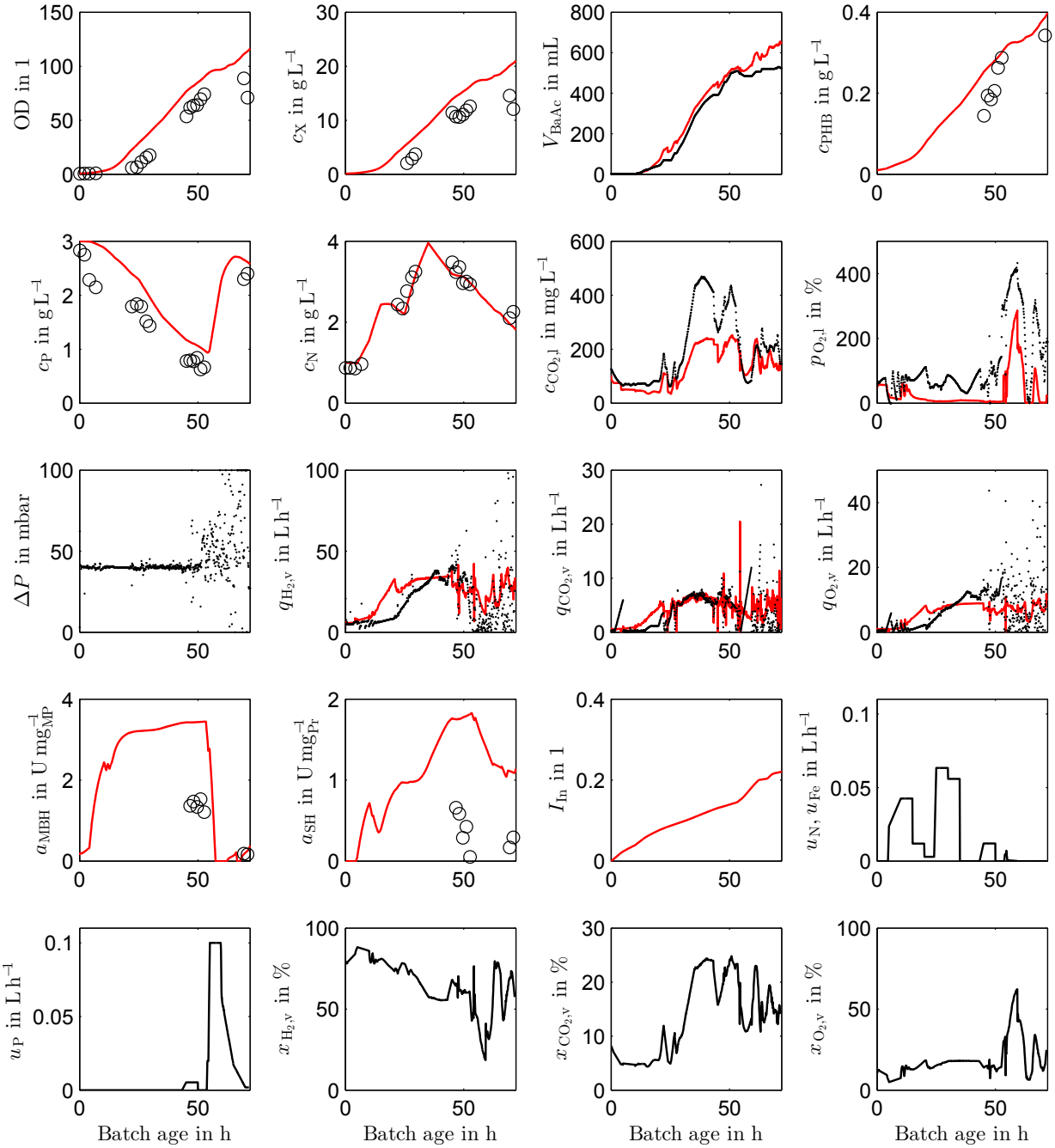


Figure 5.7: Cross-validation for model (I) of cultivation REatc20. Measurement data are given in black circles or dots, simulations are red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs.

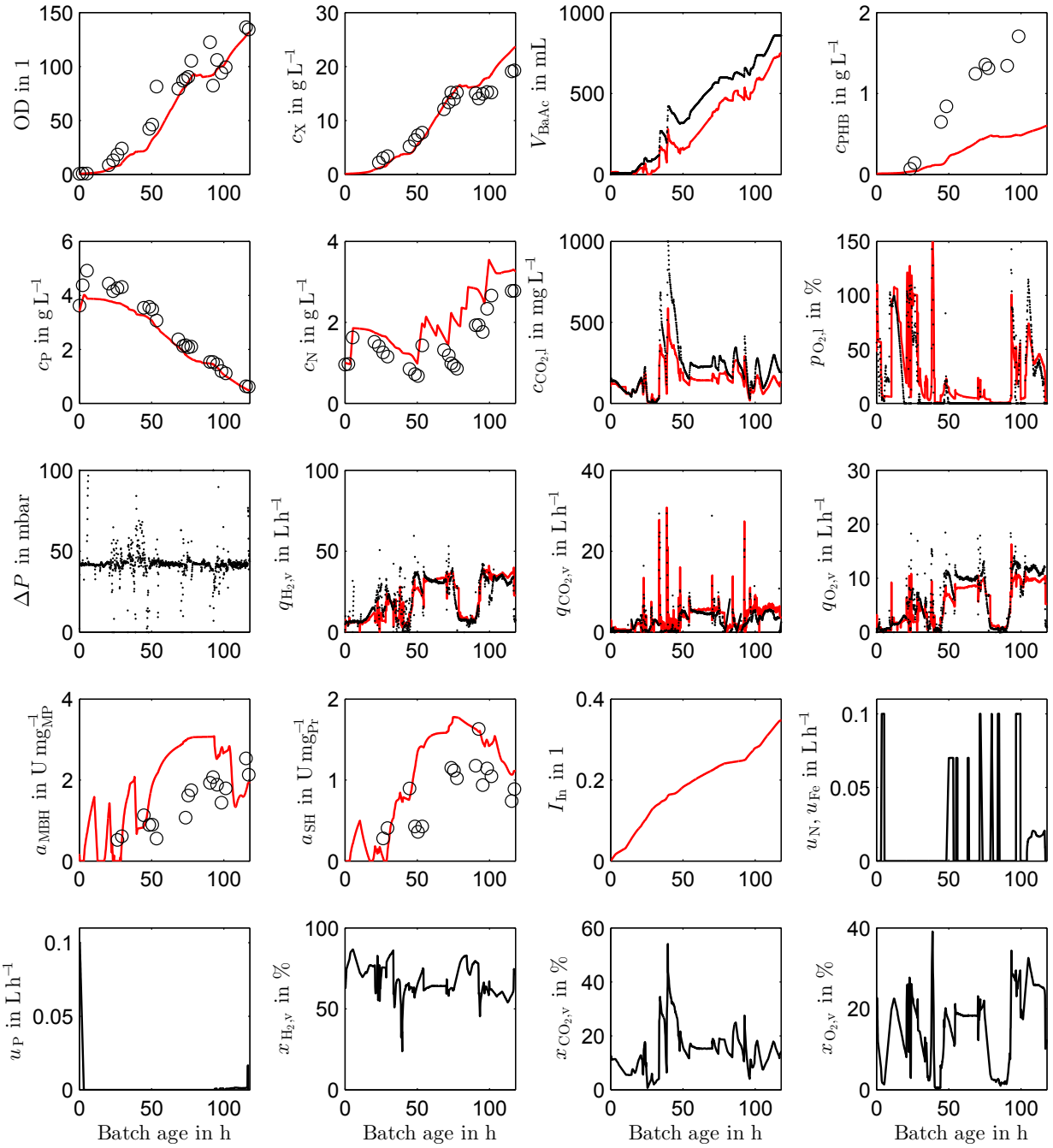


Figure 5.8: Cross-validation for model (I) of cultivation REatc22. Measurement data are given in black circles or dots, simulations are red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs.

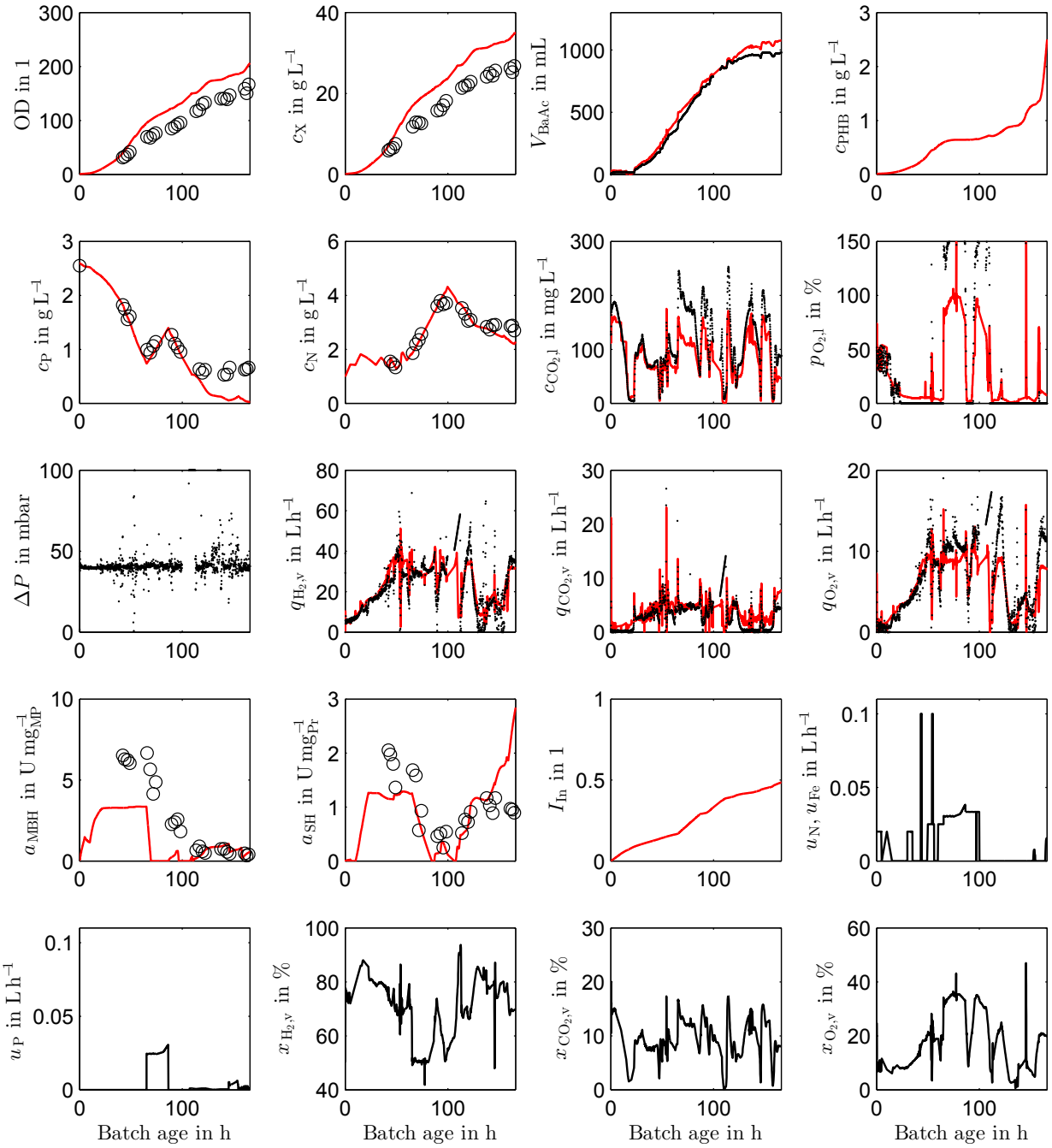


Figure 5.9: Cross-validation for model (I) of cultivation REatc23. Measurement data are given in black circles or dots, simulations are red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs.

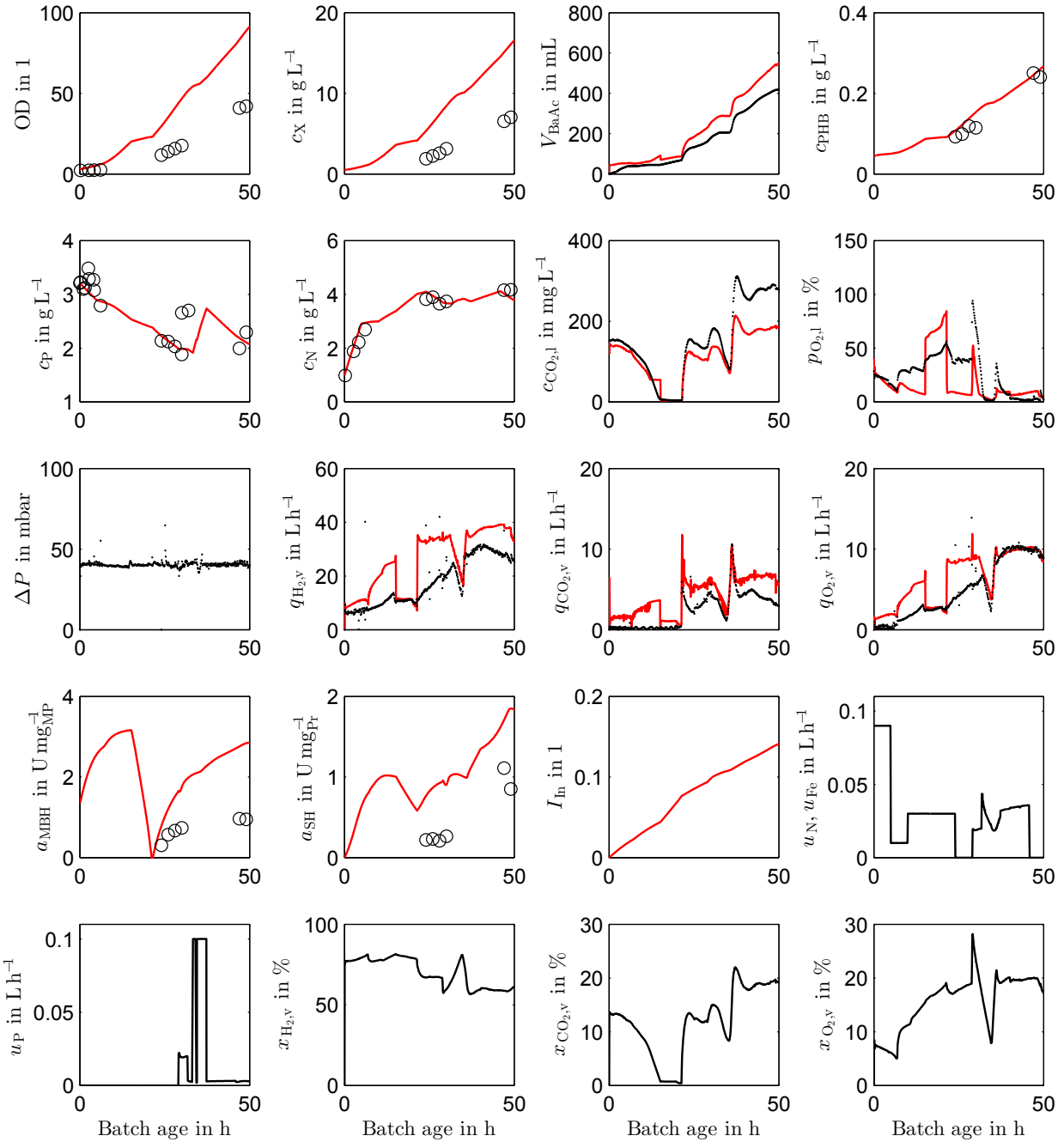


Figure 5.10: Cross-validation for model (I) of cultivation REatc26. Measurement data are given in black circles or dots, simulations are red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs.

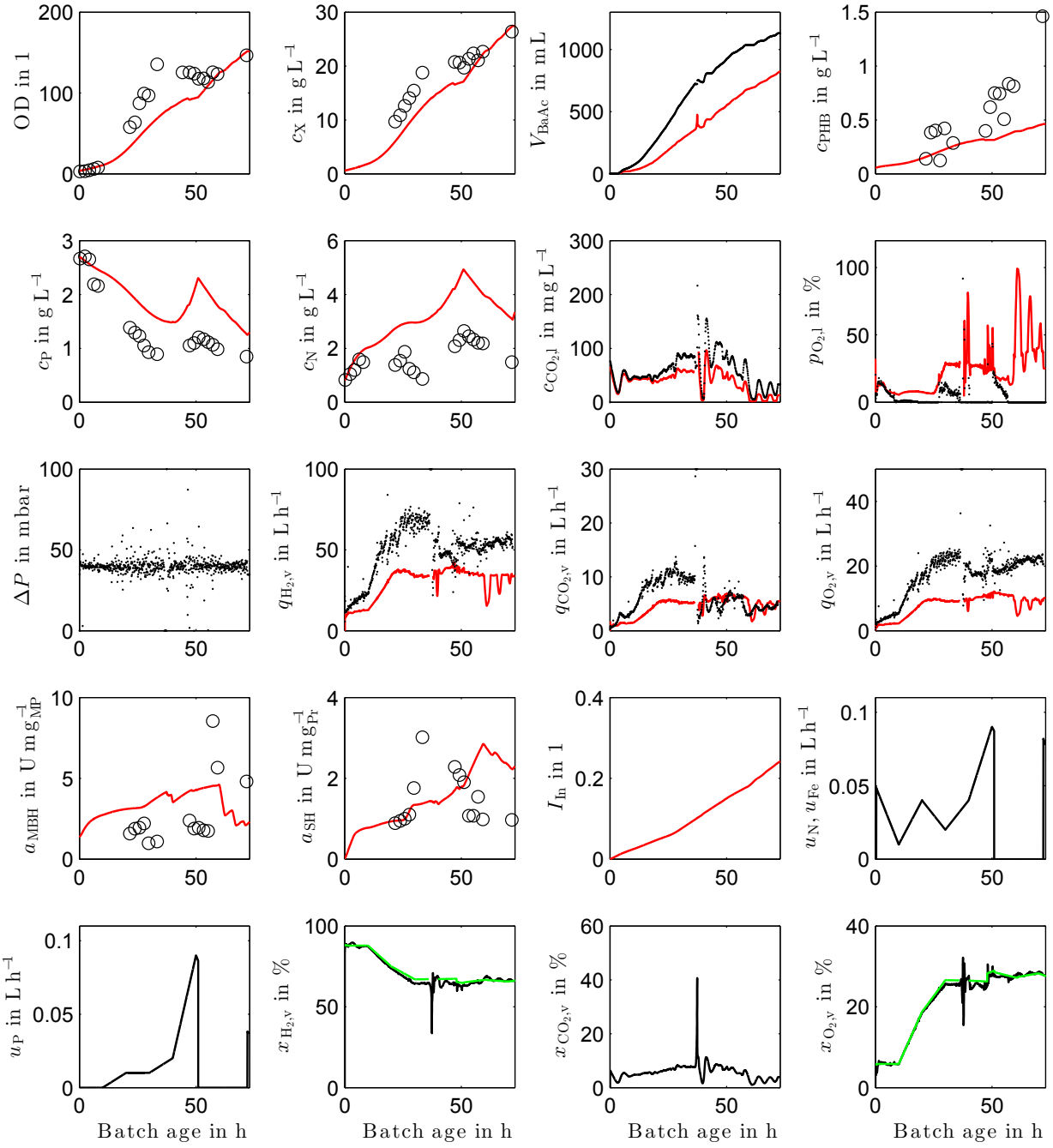


Figure 5.11: Cross-validation for model (I) of cultivation REatc33a. Measurement data are given in black circles or dots, simulations are red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs, of which $x_{H_2,v}$ and $x_{O_2,v}$ were adjusted by the gas controller. Reference values for control are given in green.

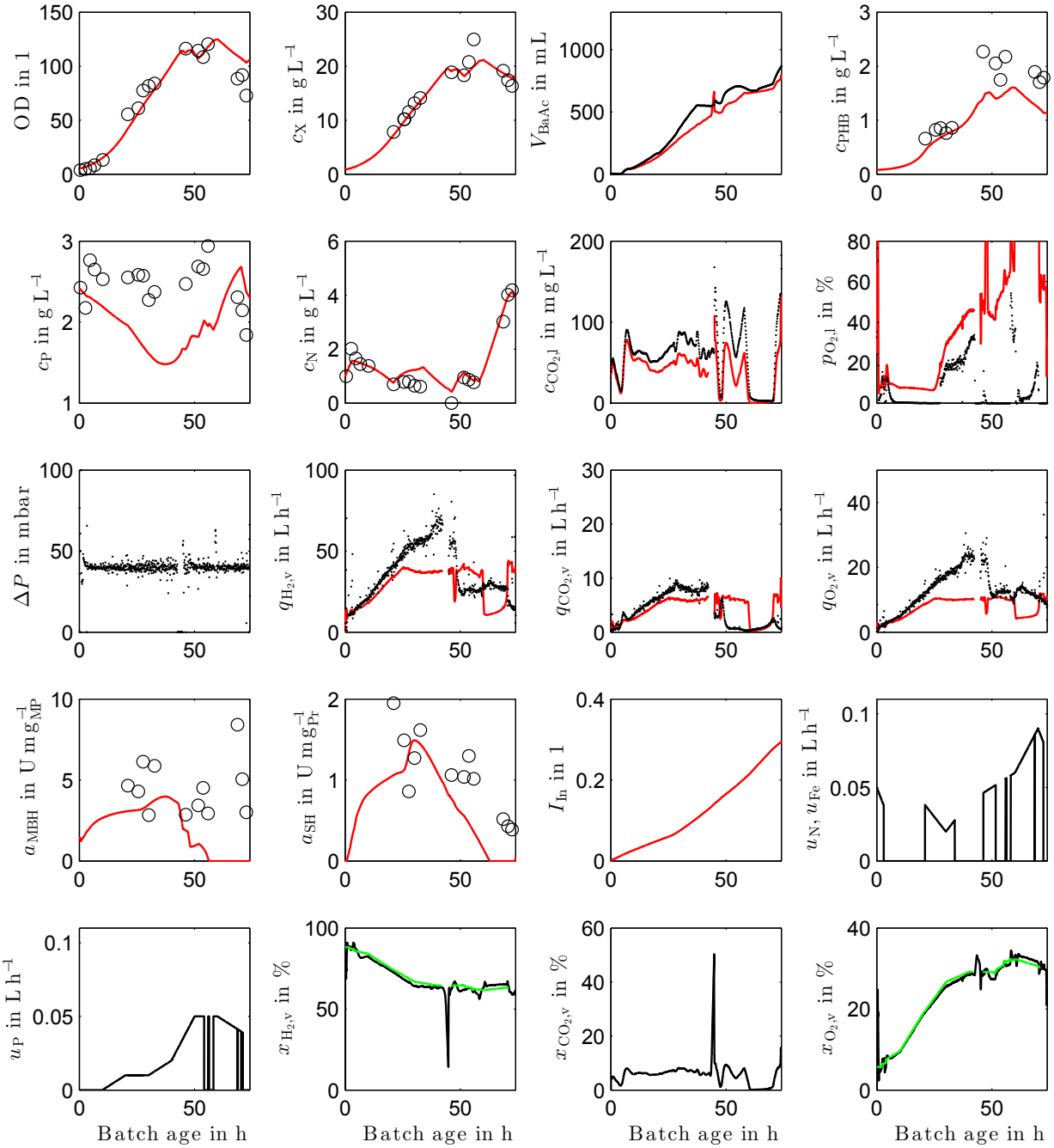


Figure 5.12: Cross-validation for model (I) of cultivation REatc33b. Measurement data are given in black circles or dots, simulations are red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs, of which $x_{H_2,v}$ and $x_{O_2,v}$ were adjusted by the gas controller. Reference values for control are given in green.

5.3 Development of a probe for dissolved hydrogen

The process model (II), to be introduced in Section 5.5, requires dissolved gas concentrations as inputs for simulation. As described in Chapter 4, the system contains sensors for dissolved carbon dioxide and oxygen. For hydrogen, commercially available short term sensors are based on amperometric measurements similar to the Clark electrode for oxygen. From the company UNISENS (Denmark), an H_2 microsensor based on the Clark principle is offered. Unfortunately, the product is comparatively costly because of its short lifecycle due to the depletion of the reactants. Schill (1996) and Takeshita et al. (1993) reported successful methods to transform an oxygen electrode into a probe for dissolved hydrogen. The standard Clark electrode to detect dissolved oxygen works as follows: Oxygen is reduced at the platinum (Pt) cathode,

$$\mathrm{O_2} + 2\,\mathrm{H_2O} + 4\,\mathrm{e^-} \longrightarrow 4\,\mathrm{OH^-}, \tag{5.102}$$

while the silver (Ag) anode is oxidized to silver-cloride (AgCl),

$$4\,\mathrm{Ag} + 4\,\mathrm{Cl^-} \longrightarrow 4\,\mathrm{AgCl} + 4\,\mathrm{e^-}. \tag{5.103}$$

Utilizing this Clark electrode in cultivation, between cathode and anode a voltage of 675 mV is applied to ensure a full reduction of O_2. The same electrode can be used for oxidation of H_2 when a different voltage is applied and the silver electrode is coated. To convert a Clark electrode for O_2 into a more sensitive probe for H_2 by coating, the company Hansatech (England) used to sell a "Hydrogen Plating box" that is not available any more. For this thesis, the probe was modified without a Plating box following the instructions by Schill (1996).

Originally, the silver electrode used for oxygen measurements is covered with a silver-chloride layer that is always renewed by the anode reaction. If the sensor is utilized as a hydrogen probe, small quantities of chloride will interfere and reduce the long-term stability according to Schill et al. (1996). For this reason, the silver-chloride layer was polished with fine emery paper and later oxidized electrochemically to silver-(I)-oxide as proposed. For the modified electrode, when measuring H_2 in cultivation, the anodic reaction at the Pt-electrode is

$$\mathrm{H_2} \longrightarrow 2\,\mathrm{H^+} + 2\,\mathrm{e^-}, \tag{5.104}$$

and the cathodic reaction reads

$$\mathrm{Ag_2O\,(s)} + 2\,\mathrm{e^-} + \mathrm{H_2O} \longrightarrow 2\,\mathrm{Ag\,(s)} + 2\,\mathrm{OH^-}. \tag{5.105}$$

Through this transformation, the former anode became cathode and vice versa. When measuring dissloved H_2 with this transformed electrode, the electrons therefore move from the platinum wire to the silver-oxide electrode.
In order to determine the voltage required for an oxidation of the original silver anode, cyclovoltammetry was performed using Ag/AgCl or platinum (Pt) as reference electrode as shown in Figure 5.13. To prevent the formation of oxidized silver molecules other than

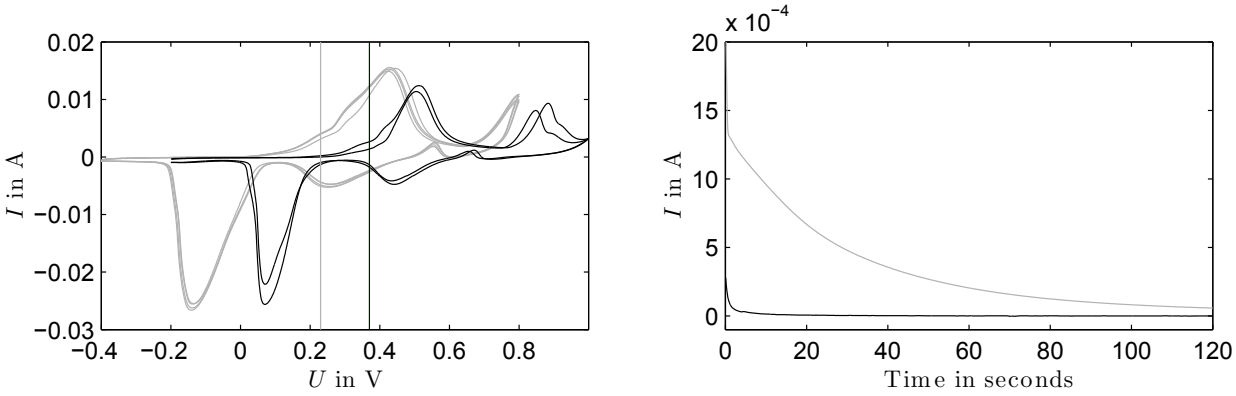


Figure 5.13: Left: Cyclic voltammograms (several runs each) for the oxidation of silver with Pt as counter and Ag/AgCl (gray) or Pt (black) as reference electrode. Right: Observed currents for the oxidization at a potential of 0.23 V and 0.37 V for references Ag/AgCl (gray) and Pt (black), respectively.

silver-(I)-oxide, electric voltages of the oxidation peak beginning were selected, as marked with vertical lines. Accordingly, the selected voltages for oxidization of the probe were 0.23 V and 0.37 V for Ag/AgCl and Pt reference electrodes, respectively. Under the observation of the current flow, two test trials were run, with a Pt and AgCl electrode serving as reference. Already after 120 seconds it was obvious that the Pt reference cannot be used because the potential (displayed in Figure 5.13, right graph) seems to shift. The measured current for the oxidization of Pt (black line) is very small and at some time points it adopts small negative values, indicating unwanted reactions. In contrast, the current of the Ag/AgCl reference (gray line) remains above zero at all times.
Consequently, for the oxidization of the silver electrode (which is also named working electrode in this context) an Ag/AgCl reference electrode was used and platinum wire served as counter electrode. For oxidation, a voltage of 0.23 V was applied for 5000 seconds between reference and working electrode and 1 molar NaOH was the electrolyte. The entire experimental setup was protected from light because radiation might change the energetic level and consequently the degree of oxidization that could have led to an electrode coating other than silver-(I)-oxide.
Once the newly coated electrode is used in cultivations, it must be polarized to be able to measure hydrogen electrochemically. Schill (1996) suggested to polarize the modified electrode with 100 mV so that hydrogen is completely reduced and as few other substances

(gases/salts) as possible interfere with the obtained signal. A solution of 3 molar KOH served as electrolyte. Step experiments with the same setting as was described in Section 5.1.1 were performed to test functionality of the electrode and repeatability of the measurements (results not shown). Conditioning the liquid phase with different hydrogen concentrations and recording the probe responses allowed to relate generated currents and dissolved hydrogen concentrations in a linear way. The obtained relation was very similar to the one of Schill (1996). However, if the probe was used in a cultivation with microorganisms, it did not seem to work: A correlation between generated current and simulated hydrogen or H_2 in the gas phase could not be observed as shown in Fig 5.14. Dissolved hydrogen was simulated with

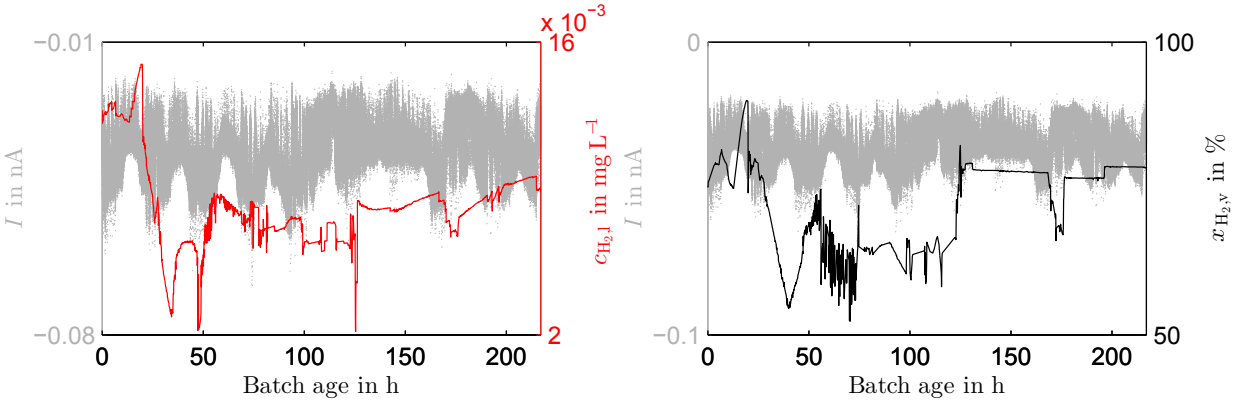


Figure 5.14: Left: measured current strength in a cultivation plotted against the dissolved hydrogen concentration simulated with process model (I). Right: the measured current strength plotted against the fraction of hydrogen in the headspace.

process model (I) that was introduced above. Assuming a correct process model, on the left of Figure 5.14, the current measured by the probe (gray) would have developed similarly to the simulated values (red). In any circumstance, however, in periods with low microbial hydrogen uptake, the development of measured hydrogen in the headspace would have been the same as the course of the measured current strength (see Figure 5.14, right). This applies in particular to the beginning of a cultivation, because the number of gas-consuming organisms is low. But also between batch age 0–30 h the dynamics of $x_{H_2,v}$ and I differ significantly. Possibly, salts or trace elements of the medium or compounds related to the cells disturb the measurements, or the signal to noise ratio is too high. For this reason, dissolved H_2 could not be measured and used as an additional input for process model (II). However, a data-driven calculation of $c_{H_2,l}$ was developed, which is introduced in the next section.

5.4 Data-driven modeling of dissolved hydrogen

For process model (II), in this section a data-driven calculation of $c_{H_2,l}$ will be suggested that can also be used for the "dependency analysis of approximated rates" in Section 6.1

and enables to utilize the modeling tool "phenomena recognition" that will be described in Section 6.2.
The data-driven calculation of dissolved hydrogen is based on three assumptions. First, a defined amount of hydrogen and a defined amount of oxygen are required to produce a certain amount of biomass and water,

$$21.36\,H_2 + 6.21\,O_2 + 4.09\,CO_2 + 0.76\,NH_3 \longrightarrow C_{4.09}H_{7.13}O_{1.89}N_{0.76} + 18.7\,H_2O. \quad (5.106)$$

The stochiometric equation above for active biomass production in H16 was suggested by Ishizaki and Tanaka (1990). For *R. e.* mutant strains, the stochiometries might differ. The second assumption states that the transport of oxygen and hydrogen between the liquid and gas phase takes place at the same velocity as identified by Rossner (2014), i.e.:

$$k_L a_{O_2} \approx k_L a_{H_2}. \quad (5.107)$$

This assumption is consistent with the film theory described in Garcia-Ochoa and Gomez (2009). And third, for each time step a pseudo-steady state is assumed,

$$\dot{n}_{\text{trans,gas}} = \nu_{\text{gas}}, \quad (5.108)$$

so that the amount of gas crossing the interface equals the rate of gas consumption ν.
Accordingly, in stages of a cultivation in which the saturation concentration $c_{O_2,\text{sat}}$, calculated according to Henry's law as in eq. (5.58) on page 78, is higher than the indirectly measured concentration $c_{O_2,l}$, derived from $p_{O_2,l}$ measurements, oxygen is being consumed. It is to be assumed that this also holds for hydrogen. The stoichiometric coefficient of the growth reaction in eq. (5.106) is used to calculate the difference between the supposed molar hydrogen concentration and the saturation concentration. With the steady-state oxygen consumption rate denoted by ν_{O_2} and the required molar ratio $H_2/O_2 \approx 3$ as suggested in eq. (5.106), follows

$$\nu_{H_2} = 3 \cdot \nu_{O_2}. \quad (5.109)$$

For the ratio the exact value of 3.44 was not used, because Morinaga et al. (1978) reported different consumption coefficients for H16 that were also considered. According to Morinaga et al. (1978), gas consumption changed with the limiting gaseous substrate and converting the consumption coefficients of Table 5.6 to molar values, yields ratios between 2.4 and 4.1. Because in different cultivations at varying times another substrate was limiting, the approximate and in a sense averaged value of 3 was used for the ratio. For mutant strains, however, this ratio must be adjusted by evaluating only the gas consumption in cultivations, since the stoichiometric biomass equations are usually unknown. Employing eq. (5.109) and

the steady-state relation yields

$$\dot{n}_{\mathrm{trans,H_2}} = \nu_{\mathrm{H_2}} = 3 \cdot \nu_{\mathrm{O_2}} = 3 \cdot \dot{n}_{\mathrm{trans,O_2}} = 3 \cdot \mathrm{k_L a_{O_2}} \cdot V_\mathrm{l} \cdot \left(c^\mathrm{n}_{\mathrm{O_2,sat}} - c^\mathrm{n}_{\mathrm{O_2,l}}\right), \tag{5.110}$$

with the molar saturation concentration being calculated according to eq. (5.35) on page 65. Replacing the $\mathrm{k_L a_{O_2}}$ in eq. (5.110) by using eq. (5.107), an alternative description for the molar gas transport of hydrogen results:

$$\dot{n}_{\mathrm{trans,H_2}} = \mathrm{k_L a_{H_2}} \cdot V_\mathrm{l} \cdot 3 \cdot \left(c^\mathrm{n}_{\mathrm{O_2,sat}} - c^\mathrm{n}_{\mathrm{O_2,l}}\right). \tag{5.111}$$

Equating eq. (5.111) with the molar standard description for the gas transport of hydrogen

$$\dot{n}_{\mathrm{trans,H_2}} = \mathrm{k_L a_{H_2}} \cdot V_\mathrm{l} \cdot (c^\mathrm{n}_{\mathrm{H_2,sat}} - c^\mathrm{n}_{\mathrm{H_2,l}}) \tag{5.112}$$

gives

$$c^\mathrm{n}_{\mathrm{H_2,sat}} - c^\mathrm{n}_{\mathrm{H_2,l}} = 3 \cdot \left(c^\mathrm{n}_{\mathrm{O_2,sat}} - c^\mathrm{n}_{\mathrm{O_2,l}}\right) \tag{5.113}$$

$$c^\mathrm{n}_{\mathrm{H_2,l}} = c^\mathrm{n}_{\mathrm{H_2,sat}} - 3 \cdot \left(c^\mathrm{n}_{\mathrm{O_2,sat}} - c^\mathrm{n}_{\mathrm{O_2,l}}\right). \tag{5.114}$$

As mentioned above, both saturation concentrations can be calculated and $p_{\mathrm{O_2,l}}$ is measured and converted to the molar concentration $c^\mathrm{n}_{\mathrm{O_2,l}}$, and therefore $c^\mathrm{n}_{\mathrm{H_2,l}}$ is then calculated with eq. (5.114). Multiplying $c^\mathrm{n}_{\mathrm{H_2,l}}$ with the molar mass $M_{\mathrm{H_2}}$ yields the mass concentration $c_{\mathrm{H_2,l}}$. However, the stochiometry in eq. (5.106) applies only to the formation of active biomass. When the storage compound PHB is formed, according to Tanaka et al. (1995), its stochiometric equation reads

$$33\,\mathrm{H_2} + 12\,\mathrm{O_2} + 4\,\mathrm{CO_2} + 0.76\,\mathrm{NH_3} \longrightarrow \mathrm{C_4H_6O_2} + 30\,\mathrm{H_2O}. \tag{5.115}$$

As for active biomass, the consumption ratio for PHB is $\mathrm{H_2/O_2} \approx 3$ and thus equals eq. (5.109). Therefore, regardless of whether active biomass or PHB is formed, the method described above can be used to approximately calculate dissolved hydrogen. A comparison between data-driven calculations of $c_{\mathrm{H_2,l}}$ and simulations with the general process model (I) is shown in Figure 5.15. Both simulations are quite similar for large parts of the cultivations. If they do not match, there are usually differences between the measured $p_{\mathrm{O_2,l}}$ values and the ones simulated by process model (I) at the same time. This can be explained by the fact that data-driven calculation and model-based simulation calculate gas input and gas consumption very similarly. Discrepancies between measurements and simulations of dissolved gasses might be caused by changing gas transport. It is probable that the $\mathrm{k_L a}$ value changes during the cultivation due to viscosity modifications or alteration of the compressor performance as the membrane of the compressor loses elasticity over time and needs

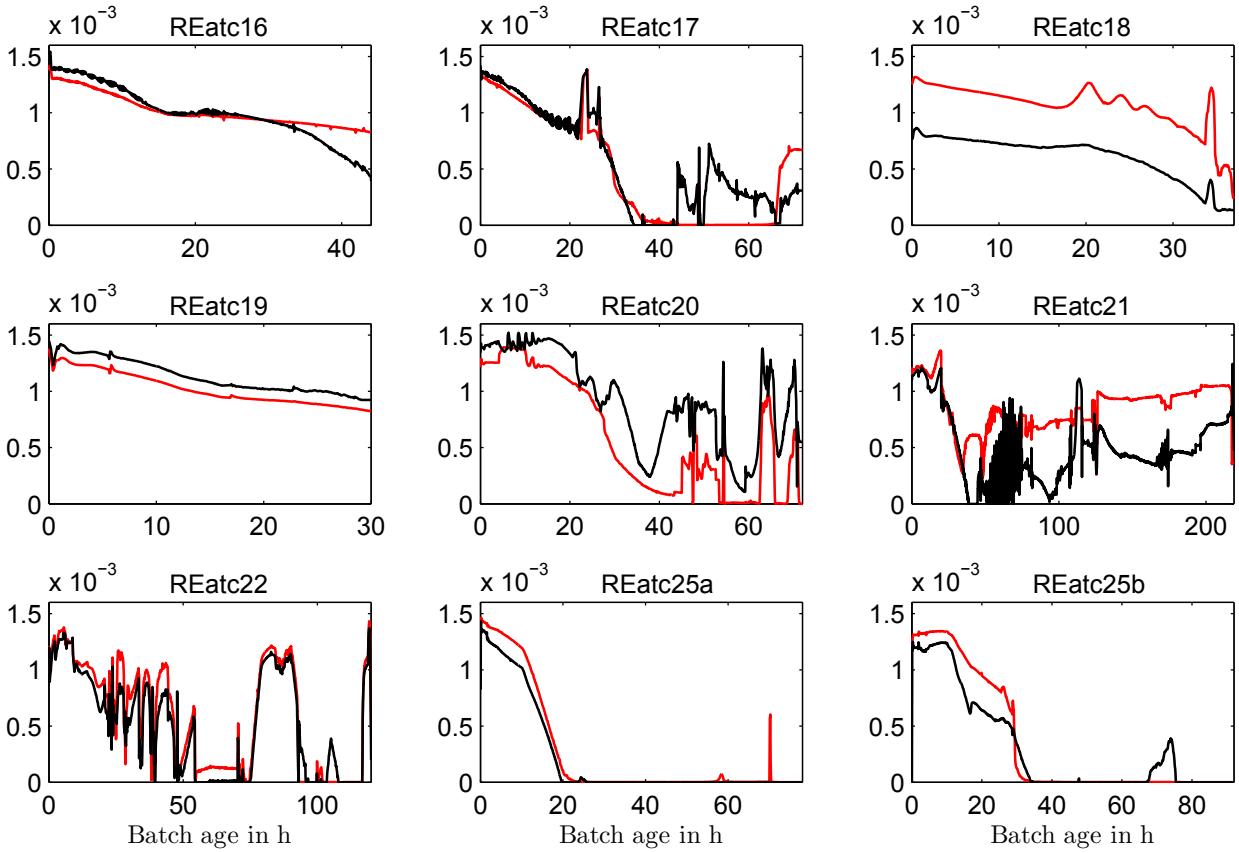


Figure 5.15: Data-driven calculated dissolved hydrogen $c_{H_2,l}$ (black) compared to the simulation (red) with process model (I), both given in mg L^{-1}

to be replaced in certain intervals. These issues affect all dissolved gas concentrations and are neither incorporated in the process model nor in the data-driven hydrogen calculation. Another explanation is provided by the underlying reaction equations. In order to calculate the consumed hydrogen on a data-driven basis, stochiometric coefficients of the growth equations taken from literature are employed. In the model, not stoichiometric coefficients but consumption coefficients are used that indicate a similar ratio of gas consumption. It is to be assumed that the coefficients are not correct at all stages of the cultivation since the cells might change their metabolic behavior. For example, it is known that cells produce phosphate storage in the absence of certain nutrients (see Doi et al. (1989)), which certainly influences gas consumption.

The presented data-driven hydrogen calculations are necessary for the model adaption. Employing them as inputs for dissolved hydrogen, model (II) can be used for an initial parameter estimation and calculated $c_{H_2,l}$ serves the adaption routines "dependency analysis of approximated rates" and "phenomena recognition" presented in Chapter 6. If the calculations of $c_{H_2,l}$ are compared with the simulations of process model (I), model deficiencies of

model (I) are pointed out in case the values do not match. This typically occurs when the measurements for $c_{O_2,l}$ differ from the simulated ones. In this case it might be useful to use process model (II) for the development of reaction rates, as these are then developed based on the measured dissolved gases. This way, the derived kinetic functions are more accurate than if they were based on incorrectly simulated gas concentrations calculated by process model (I).

5.5 Process model (II) without gas transport

In this section, the structure of process model (II) is given. The main difference between model (I) and (II) is that the latter considers only the liquid phase. Thus the model (II) is less complex than the model (I), therefore kinetic parameters can be estimated relatively quickly and these estimation results serve as initial parameters for parameter identification with the model (I).
In the general process model (I) of Section 5.2, the dissolved gas concentrations are simulated and serve as inputs for kinetic functions to determine reaction rates. As soon as discrepancies between measured and simulated dissolved gas concentrations occur, error-prone reaction rates result, and hence kinetic parameters cannot be estimated correctly. Depending on the cultivation, dissolved gases cannot always be modeled accurately. In order to identify relationships between products and gaseous substrates and to identify kinetic parameters, the process model (II) without gas transport was formulated which uses the measured dissolved carbon dioxide and oxygen concentrations as inputs

$$\underline{u}^T_{\text{model II}} = (u_{\text{N}} \quad u_{\text{Fe}} \quad u_{\text{P}} \quad c^{\text{n}}_{H_2,l} \quad c_{CO_2,l} \quad p_{O_2,l} \quad u_{\text{base}} \quad u_{\text{acid}} \quad u_{\text{antifoam}}). \tag{5.116}$$

Dissolved hydrogen is also a nutrient in autotrophic cultivations, and hence represents an additional input, but in contrast to $p_{O_2,l}$ and $c_{CO_2,l}$ it could not be measured as discussed in Section 5.3. In order to still be able to use the dissolved hydrogen as model input, a data-driven calculation was developed, as described in the previous section.
To use the measured and calculated values of the gases as inputs, the following had to be considered: Since low carbon dioxide and oxygen measurements were unreliable, these values had to be adjusted. For carbon dioxide, the sensor limit was 5 $\text{mg}\,\text{L}^{-1}$. Thus, all measured data hitting or below this threshold had to be manipulated post-experimentally with respect to the inlet carbon dioxide flow. If no carbon dioxide had been fed, the measured dissolved concentration was set to zero.
When comparing the simulations of process model (I) to those of process model (II) regarding general growth behavior, it became obvious that process model (II) underestimates growth when dissolved O_2 or H_2 were limiting. If, for example, the measured dissolved oxygen

concentration is zero, the cells cannot grow according to process model (II) because a relevant substrate is missing. But if $q_{O_2,v} > q_{leak,v} \cdot x_{O_2,v}$ at the same time, it is known that oxygen is only measured as zero but the cells are exposed to higher oxygen concentrations than indicated by the electrode. One possible cause for the discrepancy between measured and real $p_{O_2,l}$ could be that the surface-loving bacteria have accumulated on the sensor membrane and no oxygen diffuses to the electrode. Another explanation for the discrepancy would be that, contrary to our assumption, the reactor is not completely gradient free despite fast stirring. In this case, at the $p_{O_2,l}$ measuring point, i.e., at the edge of the reactor, the oxygen concentration would be lower than at the point where O_2 enters the system, i.e., centrally via the sparger. An additional explanation would be that the sensor for dissolved oxygen is not accurate enough to measure very small concentrations, however these low values of $c_{O_2,l}$ are sufficient for growth of *R. e.* In a nutshell, although oxygen is present, which is indicated by $q_{O_2,v}$ and cell growth, the sensor measures for $p_{O_2,l}$ a value of zero. Consequently, for very low measurements of dissolved oxygen the inlet gas flow has to be considered to correct the measured value accordingly.
Since the dissolved hydrogen is calculated as a function of $c_{O_2,l}$ that is derived from $p_{O_2,l}$, as explained in Section 5.4, $c_{H_2,l}$ must also be adjusted. After correcting the minimum $p_{O_2,l}$ to 1 % and the minimum dissolved hydrogen concentration to $2 \cdot 10^{-4}$ mol L^{-1}, growth was no longer underestimated using the process model (II). Both minimum values were found empirically to equal roughly one and five percent of the saturation gas concentrations when the gas composition consists of 20 % O_2 and 50 % H_2. For carbon dioxide, a positive correction is not necessary because there were no experiments where CO_2 was measured as zero or almost zero and a positive $q_{CO_2,v}$ was present at the same time. As stated before, in this process model (II), the calculated dissolved H_2 and the measured dissolved CO_2 and O_2 represent nutrients, and thus determine the values of kinetic functions. All state equations and kinetic functions are structurally identical to those in the general process model (I). But here, in contrast to the general model (I), dissolved gas concentrations are not derived by mass balances and the impact of dissolving CO_2 on the consumed base is neglected. Thus, the state vector shortens to

$$\underline{x}^T_{\text{model II}} = (I_{\text{In}} \;\; m_X \;\; m_{PHB} \;\; A_{MBH} \;\; A_{SH} \;\; V_{BaAc} \;\; m_N \;\; m_P \;\; V_l). \tag{5.117}$$

Since gas transport is not regarded, required gas volume flows and dissolved gas concentrations are not calculated and the measurement vector of model (II) shortens to the components

$$y_1 = c_X = \frac{m_X + m_{PHB} + m_{PHB,gr}}{V_l} \tag{5.118}$$

$$y_2 = \text{OD} = K_{OD,X} \cdot \frac{m_X}{V_l} + K_{OD,PHB} \cdot \frac{m_{PHB}}{V_l} \tag{5.119}$$

$$y_3 = c_{\mathrm{N}} = \frac{m_{\mathrm{N}}}{V_{\mathrm{l}}} \tag{5.120}$$

$$y_4 = c_{\mathrm{P}} = \frac{m_{\mathrm{P}}}{V_{\mathrm{l}}} \tag{5.121}$$

$$y_5 = a_{\mathrm{SH}} = \frac{A_{\mathrm{SH}}}{m_{\mathrm{Pr}}} \tag{5.122}$$

$$y_6 = V_{\mathrm{BaAc}} \tag{5.123}$$

$$y_7 = a_{\mathrm{MBH}} = \frac{A_{\mathrm{MBH}}}{m_{\mathrm{MP}}} \tag{5.124}$$

$$y_8 = c_{\mathrm{PHB}} = \frac{m_{\mathrm{PHB}}}{V_{\mathrm{l}}} \tag{5.125}$$

$$y_9 = V_{\mathrm{l}} \tag{5.126}$$

As the simulated dissolved gases from model (I) often do not correspond to the measured ones, the parameter values of the kinetic functions cannot simply be copied. A new parameter estimation of model (II) is necessary before it can be utilized for developing new model branches. Therefore, in the model adaption for mutants, the parameters of model (II) were re-estimated for each strain before it was used for adaption.

5.6 Limitations of the process model (II)

One drawback of the process model (II) is that it requires calculations for dissolved hydrogen as no measurements are available. A calculation works with ratios of consumed gases, either obtained from stoichiometric equations or cultivation data. Hence, the usage of model (II) requires prior knowledge of the organism, which is not always available when new strains are involved, or cultivations of the new strain must have already been carried out at this stage. But in general, measurements of $c_{\mathrm{H_2,l}}$ would be more accurate than the approximate calculations that are used in this thesis.
Moreover, according to the process model (II), a change of dissolved carbon dioxide does not lead to a feed flow of base or acid. However, the dissolving of CO_2 produces acid as discussed with reaction equations in Section 5.1.1 and gassing out of CO_2 produces hydroxide. These reactions are compensated for by pH control during cultivations. To include the dependency of the base on CO_2, the model would have to be extended with another input namely the numerically obtained time derivative of the dissolved carbon dioxide concentration.

Despite of these limitations, the process model (II) presented above can be employed to support model adaption. It serves two purposes: First, it can be used for a preliminary parameter estimation to find good initial values, so that the following estimation with model (I) is faster. On the other hand, it can be used for modeling new components, since the dissolved gases serve directly as model inputs. This way, effects of dissolved gases on these components can be detected and related. In addition, kinetic parameters can be estimated more accurately compared to model (I), as the measured dissolved gas concentrations are always congruent to the measurements. Details on how model (II) is integrated into the model adaption are given in the next chapter.

Chapter 6

Model adaption routines

When a mutant strain is cultivated autrotrophically, its metabolic behavior can be different compared to the wild-type H16, and therefore the process model (I) has to be adapted. For this purpose, process model (II) is employed together with adaption methods to be introduced. Process model (II) without gas transport, as described in Section 5.5, is a part of the general process model (I) that was presented in Section 5.2. An overview of the models' subunits and their interfaces is shown in Figure 6.1. Model (I) consists of three subunits: Unit 1 considers gas transport including microbial gas consumption to simulate dissolved gas concentrations $c_{\mathrm{gas,l}}$. The latter and the liquid feedrates ($\underline{u}_{\mathrm{l}}$) are inputs of the second subunit. Outputs of the second unit are the fraction of active biomass x_{active} and the concentrations of biomass, ammonium and phosphate (c_{X}, c_{N}, c_{P}). All compartments that significantly affect the first measurement quantity of model (I) that is y_1, i.e., the measured biomass (for H16: active biomass and PHB including granule surfaces), are simulated within the growth model unit. In this context, compartment refers to all cell components that are described in the model by state equations and represent parts of the modeled total biomass. The third subunit calculates strain-dependent products, e.g., SH and MBH, which do not have an impact on the measured biomass concentration. Their formation and degradation depends on the outputs of Unit 2, which are x_{active}, c_{X}, c_{N} and c_{P}, as well as on $c_{\mathrm{gas,l}}$. The state relevant rows of the input vector $\underline{u}^T_{\mathrm{model\,I}}$ from eq. (5.41) on page 72 are split into liquid and gaseous inputs $\underline{u}_{\mathrm{l}}$ and $\underline{u}_{\mathrm{v}}$, respectively. Liquid inputs are the correction fluids, salt as well as iron feed flows that are required to calculate the liquid volume V_{l}. On the other hand, V_{l} is needed in Unit 1 to convert the quantity of gases from the mass balances into concentrations $c_{\mathrm{gas,l}}$, see eq. (5.73)–(5.75). According to these balances, consumed gas changes the dissolved gas concentrations and must therefore be considered within Unit 1. For this purpose, the reactions rates μ_i for gas-consuming reactions (for H16: μ_{X}, μ_{PHB} and $\mu_{\mathrm{X,PHB}}$) and information of the amount of active cells are transferred from Unit 2 to Unit 1. The gaseous inputs $\underline{u}_{\mathrm{v}}$ are composed of $x_{\mathrm{gas,v}}$ and P.

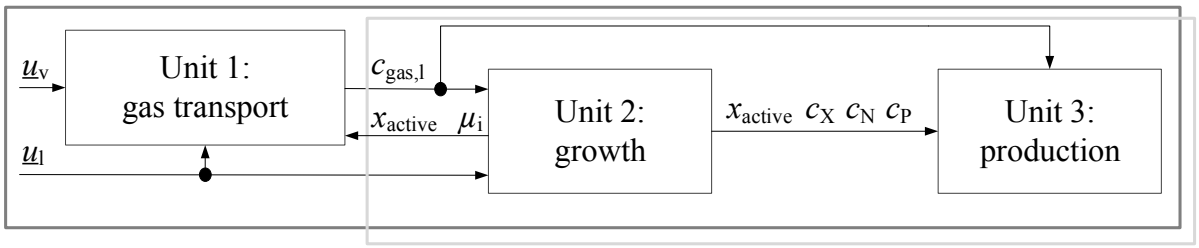


Figure 6.1: Subunits of the simplified general process model (I) framed in dark gray and process model (II) framed in light gray, describing the autotrophic cultivation of *R. e.* Growth-associated water production is not taken into consideration.

In order to adapt the general process model (I) for mutant strains, the growth model Unit 2 has to be adjusted on a structural level, first. Since the development of new strains may lead to metabolic changes, they must be taken into account in the model. As soon as such a change affects the compartments of model (I), they must first be adapted. For instance, a deletion of particular genes might cause that PHB cannot be formed and the model equation for PHB would have to be removed.
After adjusting the model structurally by adapting the compartment descriptions, all remaining parameters of the gas transport and growth model Units 1 and 2 are estimated by an optimization employing cultivation data of the mutant strain. In case the optimizer is stuck in a local minimum, it is beneficial to reduce the model complexity. This is done by exclusively utilizing Unit 2 of process model (II) for parameter estimation. Therein, the dissolved gases together with the liquid feedrates serve directly as inputs. Once the growth model unit simulations of model (II) are in accordance with the cultivation data, the estimated parameters of Unit 2 are used as initial parameter values for a subsequent estimation with process model (I). Then, the parameters of the gas transport Unit 1 are estimated separately before the parameters of gas transport and growth Units 1 and 2 are finally identified together.
Once a product other than those described in the model (I) is relevant, it must be added to the production unit. For this purpose, three methods were developed or adjusted for autotrophic cultivations and employed for the mutant strains. Of the three, "dependency analysis of approximated rates", to be introduced in Section 6.1, is presented first, followed by a brief explanation of "phenomena recognition" in Section 6.2 that was developed by Herold and King (2014). Both methods were applied to cultivation data of strain HF951 and the results will also be presented. The third method to be introduced is online Optimal Experimental Design (OED) in Section 6.3 that was tested in a simulation study. With exception of the last method, the methods only provide structural proposals for the production Unit 3, and therefore a subsequent parameter estimation is required. To do so, model (II) is again consulted for a parameter identification in which the prior estimated parameters

values of the growth Unit 2 are used. Utilizing model (II) for a parameter estimation of the production Unit 3 leads to more accurate results and thus to a better understanding of the process, since the measured dissolved gas concentrations are used instead of the simulated values of the gas transport Unit 1, which may not be congruent to the measurements. Subsequently, the estimated parameter values are transferred to model (I) and one final estimation step for all parameters belonging to Units 1, 2 and 3 follows.

6.1 Dependency analysis of approximated rates

If the existing growth model of Unit 2 is to be extended by another quantity, e.g., a product of Unit 3, at least one further state and one additional measurement equation must be introduced. A description for the formation and/or degradation rate of this extra state has to be found. As the approximated reaction rates in mechanistic bioprocess models depend on monitored process quantities, such as substrate concentrations, the modeler's work can be automatized, as schematically shown in the figure below. The procedure is as follows: First,

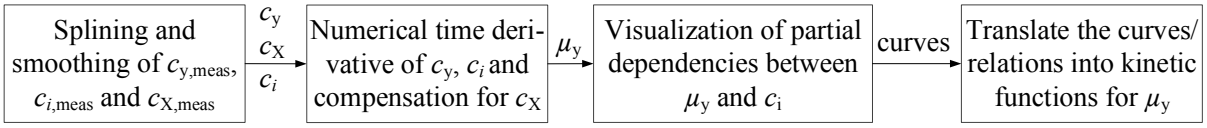


Figure 6.2: Workflow of dependency analysis of approximated reaction rates. The working steps are indicated in the boxes and their outcome is shown by an inscribed arrow. For this scheme, a product formation depending on the biomass concentration c_X was assumed.

the relative rate μ_y of the new state must be approximated. In Kristensen et al. (2003), it is estimated with an extended Kalman filter (EKF). To employ an estimator, such as an EKF, the process model needs to be known. Before running the EKF, the reaction rate to be estimated is defined as an additional state. In this thesis, an alternative, non model-based approximation is presented, which, similar to state estimation, requires measurements of a cultivation including measured values of the additional quantity $c_{y,meas}$, nutrient concentrations $c_{i,meas}$ and the biomass concentration $c_{X,meas}$. Usually, the new compound as well as the biomass concentration and some substrates cannot be measured continuously so that the measurement gaps are bridged by splining and smoothing, yielding c_y, c_i and c_X. To determine the absolute production rate, first, the time derivative of c_y is obtained. Then, to calculate the relative rate μ_y of the biomass related production, the rate values $\dot{c}_y(t_i)$ for different time instances t_i are divided by $c_X(t_i)$ to compensate for the change of biomass. If the production was suspected to depend on cell compartments instead of biomass, e.g., DNA, the concentration of DNA would have to be measured, splined and smoothed and then utilized for division. Finally, μ_y is plotted against the nutrient concentrations c_i that are suspected to affect the production rate. The shapes of the curves visualize dependencies

that are then translated into kinetic functions. For instance, if μ_y increases with an increasing nutrient concentration, a limiting relation is probable. If the rate rises with a falling substrate concentration, an inhibitory relation is suggested. The dependency analysis works best when the experimental data meet certain requirements. Ideally, the production of the target compound is influenced by just one substrate over a period of time, while the remaining substrate concentrations are constant or in a region where their influence is minor. If several substrates are identified by the figures as influences for the rate, the modeler has several options for translation. Kinetic functions can be multiplied or summed to define the reaction rate. Or combined kinetic functions (e.g., the specific kinetic function that was introduced in Section 5.2.2 on page 72) can be used or rather developed.
Below, the dependency analysis procedure is clarified with an example. The strain HF951 forms the product cyanophycin and the dependency analysis is employed to define the production rate. Assuming a biomass related production of the new compound cyanophycin (cy) yields

$$\frac{dm_{\mathrm{cy}}(t)}{dt} = \mu_{\mathrm{cy}}(t) \cdot m_{\mathrm{X}}(t) \tag{6.1}$$

$$\mu_{\mathrm{cy}}(t) = \frac{1}{m_{\mathrm{X}}(t)} \cdot \frac{dm_{\mathrm{cy}}(t)}{dt} \tag{6.2}$$

$$\mu_{\mathrm{cy}}(t) = \frac{1}{c_{\mathrm{X}}(t)} \cdot \frac{dc_{\mathrm{cy}}(t)}{dt}. \tag{6.3}$$

Cyanophycin can only be measured indirectly as a function of optical density (OD) and biomass concentration (c_{X})

$$K_{\mathrm{OD,cy}} \cdot c_{\mathrm{cy}} = \mathrm{OD} - K_{\mathrm{OD,X}} \cdot c_{\mathrm{X}}, \tag{6.4}$$

with $K_{\mathrm{OD,cy}}$ and $K_{\mathrm{OD,X}}$ being constant factors that will be described in detail in Section 7.2. Considering the indirect measurement method, $K_{\mathrm{OD,cy}} \cdot c_{\mathrm{cy}}$ is used instead of c_{cy} and eq. (6.3) becomes

$$K_{\mathrm{OD,cy}} \cdot \mu_{\mathrm{cy}}(t) = \frac{1}{c_{\mathrm{X}}(t)} \cdot \frac{K_{\mathrm{OD,cy}} \cdot dc_{\mathrm{cy}}(t)}{dt}, \tag{6.5}$$

Thus, the time dependent reaction rate $K_{\mathrm{OD,cy}} \cdot \mu_{\mathrm{cy}}(t)$ can be calculated by multiplying the inverted biomass concentration c_{X} and the numerically differentiated indirect cyanophycin measurements. In Figure 6.3, the smoothing results of indirect cyanophycin values in cultivation HF951A (visualized in Figure 7.3 on page 139) are shown together with the resulting development of the reaction rate. The approximated rate of cyanophycin (see Figure 6.3) is plotted against the measured nutrients c_{N}, c_{P}, $c_{\mathrm{CO_2,l}}$ and the calculated $c^{\mathrm{n}}_{\mathrm{H_2,l}}$ in Figure 6.4. If the plotted values tend to be monotonously increasing or decreasing or a combination of both, dependencies of the rate on the corresponding nutrient can be derived. When

interpreting the curve that correlates CO_2 and cynaophycin, the values on the gray background should be excluded because measurements of $c_{CO_2,l}$ below 5 mg L^{-1} are unreliable. Values marked as gray circles were also excluded from the evaluation because they belong to time periods with rather fast dynamics of dissolved CO_2 and concentrations higher than 5 mg L^{-1} (see batch ages 15.6–18.6 h, 23.3–26.7 h and 30.5–31.5 h in Figure 7.3). Such rapid gas changes may result in a rapid change in the cyanophycin formation rate, but this

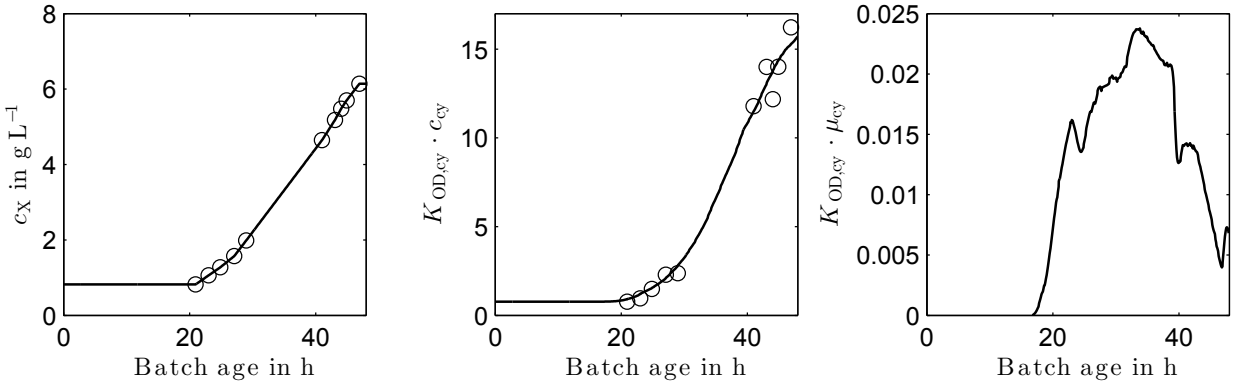


Figure 6.3: The approximated cyanophycin rate $K_{OD,cy} \cdot \mu_{cy}$ over time (right) calculated with smoothed and interpolated indirect cyanophycin measurements (middle) and the interpolated biomass concentration c_X (left). Measurements are given in circles.

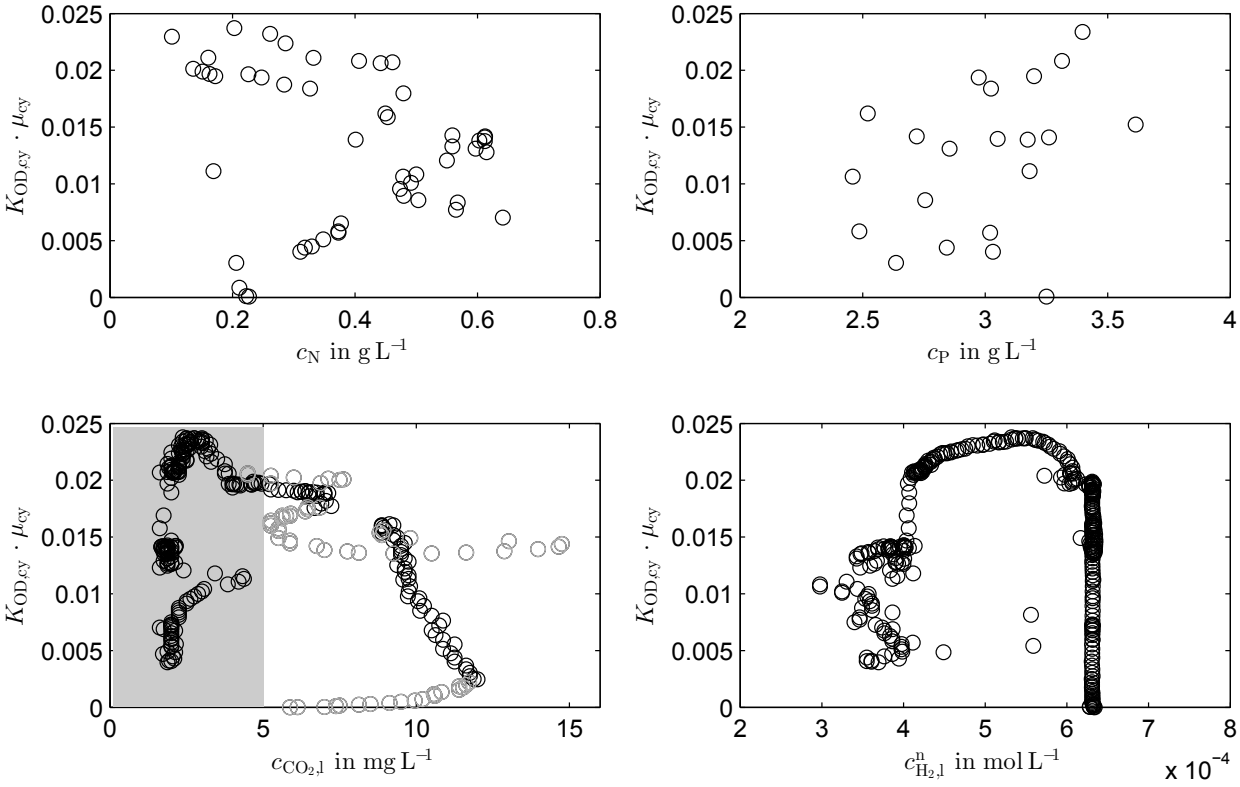


Figure 6.4: Partial dependency analysis based on a numerically derived reaction rate $K_{OD,cy} \cdot \mu_{cy}$ in cultivation HF951A

could not be detected from the measurements due to the rather rare manual sampling, i.e., low frequent cyanophycin values. In the evaluation of the remaining measurements, the cyanophycin rate tends to decrease with increasing dissolved carbon dioxide. Thus, according to the experiment, a negative impact of carbon dioxide on the formation of cyanophycin is possible. For ammonium, phosphate and hydrogen the tendency is not clear and there seems to be no correlation between the concentrations of these substrates and the indirect measurement quantity $K_{\mathrm{OD,cy}} \cdot \mu_{\mathrm{cy}}$. The results of this dependency analysis are resumed when the adapted model of the cyanophycin-producing strain HF951 will be presented in Section 7.2.

6.2 Phenomena recognition

An automated modeling tool evaluating heterotrophic cultivation data was developed at the chair of Measurement and Control at the TU Berlin (published in Herold and King (2014)). It features "phenomena recognition" and model coding. Based on nutrient measurements in cultivations, it identifies suitable reaction rate expressions. The phenomena recognition tool works with possible metabolic relationships and a database of kinetic functions created by the user. In phenomena detection, the dynamics of measured substrates are related to the ones of analyzed products, which is quite similar to the "dependency analysis of approximated rates" above.
For *R. e.* the dissolved gases $c_{\mathrm{gas,l}}$ and salts c_{N}, c_{P} are substrates that are known to influence growth and product formation. In this thesis, the phenomena detection part of this software tool was extended for the autotrophic cultivation with hydrogen, carbon dioxide and oxygen as gaseous substrates. Ammonium and phosphate were measured manually and the dissolved substrates O_2 and CO_2 were measured automatically as $p_{\mathrm{O_2,l}}$ and $c_{\mathrm{CO_2,l}}$, respectively. Since no sensor for dissolved hydrogen was available, a method for approximating these concentrations on a data-driven basis was introduced in Section 5.4. To elucidate whether the dynamics of measured concentrations were caused by reactions or simply by dilution due to feedings, all measurements were reconciled by the software. Since water was produced by the organisms in the autotrophic cultivation, it needed to be considered as a diluting factor in addition to the liquid feedings. Metabolically produced water as a side product was calculated according to the description that is presented in the next section. Using the measurements from experiments together with the calculations of $c_{\mathrm{H_2,l}}$ and produced water, the data from HF951A-E cultivations were evaluated by the tool phenomena recognition to find a description for cyanophycin. Found phenomena will be listed and discussed in Section 6.2.2.

6.2.1 Data-driven modeling of microbially produced water

In addition to the calculation of $c_{H_2,l}$, the quantification of metabolically produced water is also needed for "phenomena recognition". To calculate the water production in the cultivations investigated, the measured hydrogen inflow $q_{H_2,v}$ corrected by the leakage volume flow rate $q_{leak,v}$ was evaluated,

$$q_{H_2,v}^{corrected} = q_{H_2,v} - (q_{leak,v} \cdot x_{H_2,v}). \tag{6.6}$$

Then, the volume flow $q_{H_2,v}^{corrected}$ was converted into a molar flow,

$$\dot{n}_{H_2} = q_{H_2,v}^{corrected} \cdot \frac{P_0}{R \cdot T \cdot 10^3}. \tag{6.7}$$

For H16 it was assumed that 10 % of PHB and 90 % of active biomass were produced because we found this composition in cultivations without nutrient limitations, i.e., without enhanced PHB production. As the ratio of H_2O/H_2 equals 18.7/21.36 for active biomass as in eq. (5.106) and 30/33 for PHB as in eq. (5.115), it follows

$$\dot{n}_{H_2O} = \dot{n}_{H_2} \left(0.9 \cdot \frac{18.7}{21.36} + 0.1 \cdot \frac{30}{33} \right). \tag{6.8}$$

Stoichiometric equations and consumption coefficients were taken into account and averaged when determining the ratio for the calculation of $c_{H_2,l}^{n}$ in Section 5.4. As far as the calculation of the water is concerned, no production coefficients were available, only two stoichiometric equations were used to calculate this ratio. Thus, averaging is not possible and the exact values of these equations were employed. It is to be expected, however, that the assumed ratio is not always correct and strongly depends on the nutrient supply, i.e., PHB formation. In order to compensate for varying proportions of PHB in the total biomass, the analysis results should ideally be used. However this is only possible for experiments in which PHB was measured. The strain HF951 is unable to produce PHB. Accordingly, the equation above was modified such that the water production reads

$$\dot{n}_{H_2O} = \dot{n}_{H_2} \cdot \frac{18.7}{21.36}. \tag{6.9}$$

Employing the microbial water production and the calculated $c_{H_2,l}$ together with the manual and automated measurements, allowed the usage of the "phenomena recognition" tool to evaluate the autotrophic cultivations of strain HF951.

6.2.2 Phenomena in cultivations of *R. e.* HF951

In order to detect phenomena with the software tool, the data of the cultivations HF951A–E were used, which will be presented in the Section 7.2. The relevant phenomena found were:

1. Dissolved carbon dioxide inhibits the formation of cyanophycin (batch age 44.2–47.4 h in HF951C, see Figure 7.4 on page 140),
2. The measured total biomass and cyanophycin grow simultaneously (batch age 124.2–141 h in HF951E, see Figure 7.7 on page 144),
3. Ammonium inhibits the formation of cyanophycin (batch age 44.2–58.2 h in HF951C, see Figure 7.4 on page 140).

The first phenomenon has also been detected by the "dependency analysis of approximated rates" of the previous section. It is remarkable, though, that only a very short period was found in one of five cultivations where cyanophycin was formed at low dissolved carbon dioxide concentrations. This will be discussed in the following. Phenomena recognition relates numerical derivatives of measurements in a defined timeframe, which is 2 h per default and this frame is shifted through the data of the entire cultivation. To identify an inhibitory effect of carbon dioxide on cyanophycin, carbon dioxide must have decreased and cyanophicin must have increased in a time interval of 2 h. Analysis of the cultivation data showed, see Figures 7.3–7.7 on pages 139–144, that often cyanophycin increases 6–8 h after carbon dioxide dropped almost to zero. Hence, the time frame must be adjusted for this strain. Moreover, the different sampling rates of $c_{CO_2,l}$ and cyanophycin led to the fact that potential periods for the detection of correlations were not listed. This was always the case when cyanophycin was incorrectly interpolated, so that a change in concentration was erroneously recorded before the concentration of the relevant substrate, i.e. $c_{CO_2,l}$, changed. In addition, $c_{CO_2,l}$ fluctuated in low concentration ranges, resulting in positive and negative derivative values, while cyanophycin was formed. At the same time, this contradicted the possible phenomenon of inhibition. For gases, the derivatives might have to be smoothed more than for conventional substrates, as they are subject to greater fluctuations in absolute terms.

The second phenomenon found seems to be also useful, since cyanophycin as an internal compound contributes to the total measured biomass. In contrast, the third phenomenon found is unlikely because ammonium is a main component of cyanophycin and therefore required for its formation. The algorithm probably erroneously established a causal relationship between decreasing c_N and simultaneously increasing cyanophycin.

Once these results are meant to be transferred to the adapted process model (I), what from here on is referred to as process model HF951, it has to be ensured that the data-

driven values for $c_{H_2,l}$ resemble the dissolved hydrogen concentration simulated by process model HF951. To do so, at this stage growth Unit 2 of the process model HF951 has to be validated. In order to provide a validated model, the model structure, major gas consuming reactions and its stoichiometries must be known. Therefore, calculating dissolved hydrogen and employing the obtained values in "phenomena recognition" cannot be used to develop a process model from scratch. It is mainly a tool for extending the adapted process model (I), when new products are investigated and Unit 3 needs to be altered. Although the process model also provides simulations for $c_{H_2,l}$ it is favored to use the data-driven ones in "phenomena recognition", because the simulated ones are more error-prone and thus less accurate.

6.3 Multi-model online Optimal Experimental Design

Another method to quickly extend an existing growth model for an additional model branch, e.g., the description of a specific product, is multi-model online Optimal Experimental Design (OED). It was developed for the autotrophic cultivation and tested in a simulation study first that will be presented in Section 6.3.1. When applied in cultivations, multi-model online OED caused problems described and discussed in Sections 6.3.2 and 6.3.3.
Model-based optimal experimental design was first addressed by Goodwin and Payne (1977). Since then, OED has also been used in bioprocess development. In order to deal with strong uncertainties in the structural model, which is often the case with bioprocesses, the method was adapted in Baltes et al. (1994). They included a term for model validity in the cost function. In this thesis, the usage of multi-models accounts for structural uncertainties. Model-based OED is also named Model-Based Design of Experiments and well reviewed in Franceschini and Macchietto (2008).
In contrast to traditional experimental planning, OED is model-based, and therefore requires a model that already includes the new model extension. In order for OED to make sense, even if little is known about the new model branch, the algorithm is used with different model candidates, i.e., different possible model extensions.

6.3.1 Simulation study

In the following simulation study, three models that are identical in the general growth description, but show different extensions for the dynamical behavior of the soluble hydrogenase (SH) production, were employed in an "online" routine. Based on maximizing the information content of the SH measurements with regard to a subsequent parameter identification, the system's ideal stimuli were calculated using the process models. In this way, the output sensitivities were maximized at the pre-determined sampling times.

For this simulation study, after a trajectory planned (TP) growth phase, OED was run online, which means it repeatedly calculated optimal ramp-shaped future input trajectories during the cultivation utilizing actual initial states that were estimated by DEKF or SPKF. In order to keep the study as realistic as possible, the number of cost function evaluations for OED was limited and thus the time needed for optimization as well. For OED the A-criterion was taken as a cost function in the minimization problem

$$\phi = \arg \min_{\underline{u}(t)} \left[\text{trace} \left(\mathbf{F} \right)^{-1} \right], \tag{6.10}$$

with $\mathbf{F}$ being the Fisher information matrix of eq. (3.8) on page 26 that was determined by summing over all sampling times.
In online OED, first, all model candidates were employed and, therefore, the cost function was calculated as in eq. (6.10) for the individual models, then the cost function values were summed and divided by the number of models for averaging. With the first manual SH measurements, all parameters of the new branch were estimated for each model. By direct comparison of the model error, which was measured by the cost function value as in eq. (3.6) on page 26, the best model was selected. The latter was then used to update the state estimator and utilized for the following OED. An overview of the algorithm structure is shown in the Figure below. To reduce the computational cost during OED, the operator

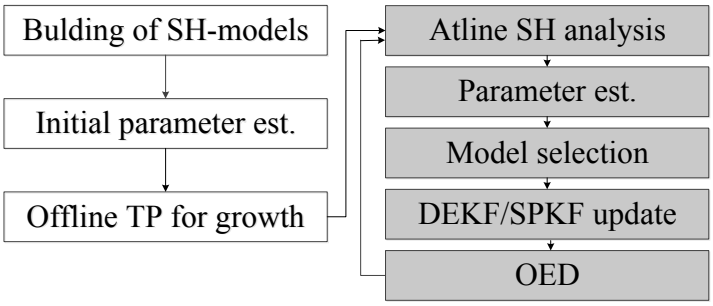


Figure 6.5: Multi-model OED workflow for modeling the SH. The repetitive online cycle is marked in gray. Further execution notes for an experimental application will be given in the Appendix A.

has to decide, which model inputs are to be optimized. For this study, it was assumed that the dissolved gas concentrations as well as the biomass itself influence the SH activity, and hence the gas fractions were allowed to be optimized by OED. As phosphate, ammonium and iron plus trace elements are required for growth and in turn might be toxic when present in large amounts, all three liquid feedings were controlled by P-controllers and optimal concentrations of 2 $\mathrm{g\,L^{-1}}$ phosphate and 3 $\mathrm{g\,L^{-1}}$ ammonium were maintained.

Proposed models

All three models contain the equations of the general process model (I) introduced in Section 5.2 and different extensions for the description of SH. In all models, the amount of the enzyme SH (m_{SH}) depends on active biomass m_{X} and the formation rate μ_{SH} and, optional, on the degradation rate $\mu_{\mathrm{SH,deg}}$,

$$\dot{m}_{\mathrm{SH}} = (\mu_{\mathrm{SH}} - \mu_{\mathrm{SH,deg}}) \cdot m_{\mathrm{X}}. \tag{6.11}$$

Both rates are defined by candidate kinetic functions that were listed in Table 5.3 on page 74. To convert the state to the measured activity, as in eq. (5.87) on page 85, $\dot{m}_{\mathrm{SH}}$ in eq. (6.11) is transformed into an expression in the activities according to eq. (5.72) on page 82 after the integration. The three model candidates differ in the descriptions for μ_{SH} and $\mu_{\mathrm{SH,deg}}$.

SH model 1

The enzyme SH regenerates reduction equivalents that are needed for growth. Hence, for SH model 1 it is assumed that SH is expressed in a growth-associated manner, and an optimal concentration of dissolved oxygen also enhances the production resulting in

$$\mu_{\mathrm{SH}} = \mu_{\mathrm{SH,max}} \cdot \mu_{\mathrm{X,In}} \cdot \mathrm{Ro}_2(c_{\mathrm{O}_2,\mathrm{l}}, k^{\mathrm{SH}}_{1,\mathrm{O}_2}\ k^{\mathrm{SH}}_{2,\mathrm{O}_2}). \tag{6.12}$$

Oxygen concentrations beyond the optimum are not beneficial for SH expression. According to the equation above, expression can only be at its maximum if the cells grow at their maximum rate, i.e. $\mu_{\mathrm{X,In}} = \mu_{\mathrm{X,max}}$, which is only the case if all substrates are present in optimal concentrations and the inhibition state is zero (see also eq. (5.62)). The latter is the case if no critically high oxygen concentrations have occurred in cultivation up to this point.
The enzyme SH is also assumed to be degraded or inactivated at high dissolved oxygen concentrations

$$\mu_{\mathrm{SH,deg}} = \mu_{\mathrm{SH,deg,max}} \cdot \mathrm{MiMe}(c_{\mathrm{O}_2,\mathrm{l}}, k_{\mathrm{SH,deg}}). \tag{6.13}$$

The rates μ_{SH} eq. (6.12) and $\mu_{\mathrm{SH,deg}}$ eq. (6.13) are inserted into eq. (5.72) to obtain an expression in the activities, as mentioned above. As $K_{\mathrm{SH,a}}$ of eq. (5.72) is also unknown, only the products $K_{\mathrm{SH,a}} \cdot \mu_{\mathrm{SH,max}}$ and $K_{\mathrm{SH,a}} \cdot \mu_{\mathrm{SH,deg,max}}$ can be identified.

SH model 2

SH model 2 also postulates an optimum of dissolved oxygen and implies that there is a limit of SH production in a cultivation, e. g., due to an inhibition caused by SH directly or

indirectly:

$$\mu_{\mathrm{SH}} = \mu_{\mathrm{SH,max}} \cdot \mathrm{Ro}_2(c_{\mathrm{O_2,l}}, k^{\mathrm{SH}}_{1,\mathrm{O_2}}\ k^{\mathrm{SH}}_{2,\mathrm{O_2}}) \cdot \mathrm{Ai}(a_{\mathrm{SH}}, k^{\mathrm{SH}}_{\mathrm{SH}}). \tag{6.14}$$

As in SH model 1, the SH degradation rate is determined by dissolved oxygen similar to eq. (6.13) and the parameters $K_{\mathrm{SH,a}}$ and $\mu_{\mathrm{SH,max}}$ can only be identified together. That also applies to $K_{\mathrm{SH,a}}$ and $\mu_{\mathrm{SH,deg,max}}$.

SH model 3

The third model was utilized to generate synthetic measurements for this study. It assumes a minimal SH amount ($m_{\mathrm{SH,min}}$) at high oxygen concentrations and a maximum SH amount ($m_{\mathrm{SH,max}}$) at low oxygen concentrations. This model resembles a description for MBH proposed by Rossner (2014). Due to cellular regulations, it is postulated that the cell aims for a certain value of SH between these boundaries depending on the availability of O_2. Naming this value a saturation amount $m_{\mathrm{SH,sat}}$, the kinetic function is given by

$$\mu_{\mathrm{SH}} = \mu_{\mathrm{SH,max}} \cdot (m_{\mathrm{SH,sat}} - m_{\mathrm{SH}})\,. \tag{6.15}$$

Multiplying this equation with $K_{\mathrm{SH,a}}$ gives an expression in the specific activities:

$$K_{\mathrm{SH,a}} \cdot \mu_{\mathrm{SH}} \quad = \quad \mu_{\mathrm{max,SH}} \cdot (a_{\mathrm{SH,sat}} - a_{\mathrm{SH}})\,. \tag{6.16}$$

The saturation activity $a_{\mathrm{SH,sat}}$ depends on the dissolved oxygen concentration,

$$a_{\mathrm{SH,sat}} \quad = \quad a_{\mathrm{SH,min}} + (a_{\mathrm{SH,max}} - a_{\mathrm{SH,min}}) \cdot \mathrm{Ro}_2(c_{\mathrm{O_2,l}}, k^{\mathrm{SH}}_{1,\mathrm{O_2}}\ k^{\mathrm{SH}}_{2,\mathrm{O_2}}). \tag{6.17}$$

In contrast to SH model 1 and 2, a maximum amount of SH per cell exists ($a_{\mathrm{SH,max}}$). The activity of SH converges towards the saturation activity, and thus can increase and decrease similar to the given models above. Hence, all three proposed models are able to describe rising and falling SH activity trajectories from a structural point of view.

Sequence of steps

To obtain reliable SH analysis results, a minimal biomass concentration is assumed to be beneficial in the simulation study. Therefore, the cultivation starts with a growth phase of 24 h that has been optimized via offline trajectory planning (TP) aiming at maximum biomass concentration. In Figure 6.6, the growth phase is marked with a blue background. Simulated manual sampling, which serves for OD and SH analysis, is shown in black circles. Due to limitations in sample procession equipment, a maximum of four samples is assumed during daytime, the OD is measured straight away and the SH is analyzed the next day. That means, if the SH samples are collected during 12 hours a day, then another 24 hours (night plus measuring period) passes until all samples have been measured. Hence, the

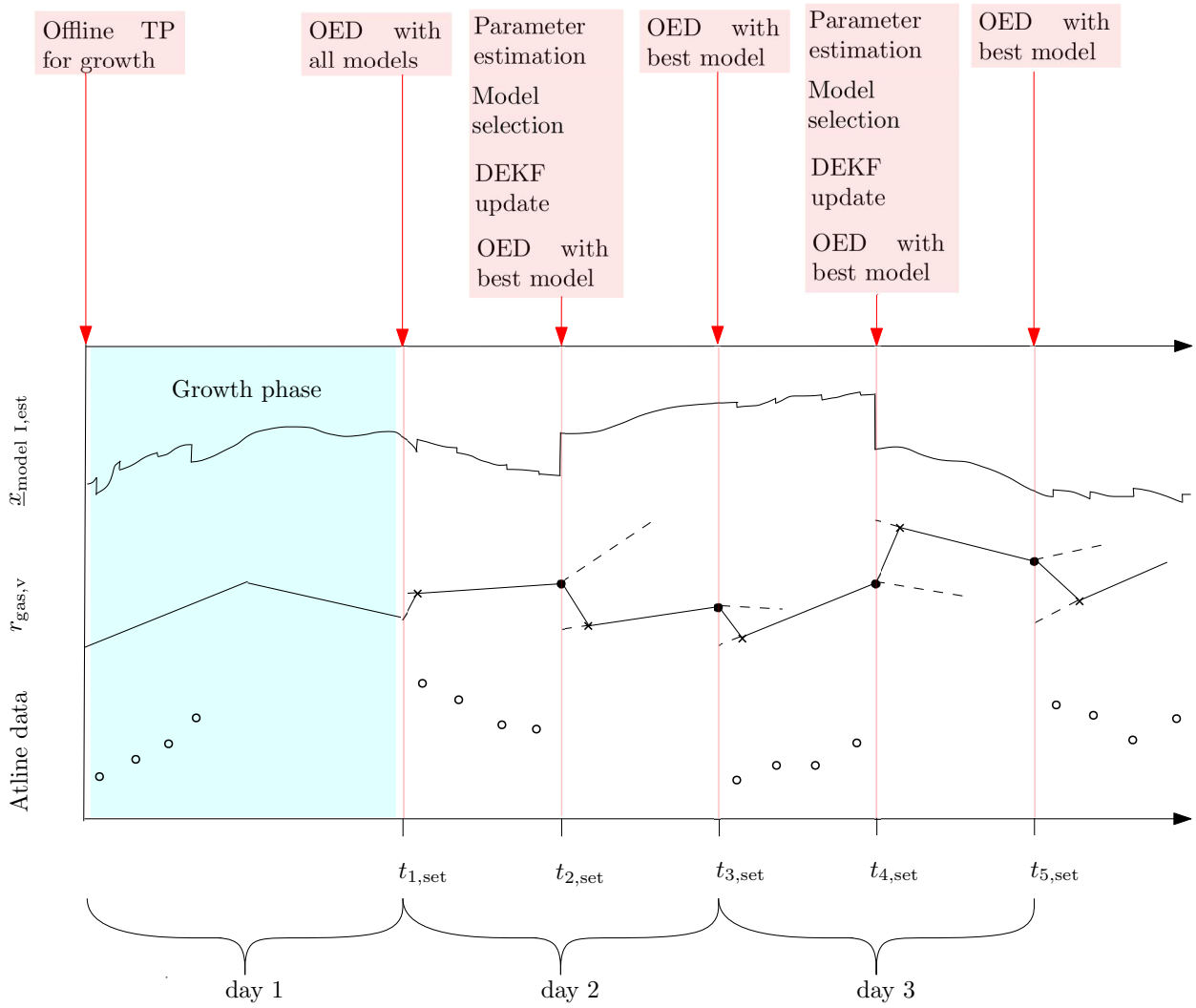


Figure 6.6: Scheme of multi-model online OED in a simulated cultivation

measurements come with a maximum delay of 36 h. A DEKF using the initially chosen model, online and time-dependent atline measurements estimates the state vector $\underline{x}_{\text{model I}}$. After the growth phase at $t_{1,\text{set}}$, the optimal gas fractions of the headspace $r_{\text{gas,v}}(t_{i,\text{set}})$ for the time instances $t_{i,\text{set}}$ are calculated by OED. Between $t_{i,\text{set}}$ the reference values are ramped because the gas phase controller cannot realize stepwise changes. With these ramp-shaped reference values, the value of the cost function, see eq. (6.10) for each individual model is calculated and then these values are summed and averaged, as described above, before being evaluated by the optimizer. After 36 h, at $t_{2,\text{set}}$, the first SH analysis data of day 1 is evaluated in parameter estimations of the different SH model extensions. The resulting cost function values, which are calculated with eq. (3.6) on page 26, are compared, the best model is selected and the model number as well as its newly identified parameters are passed to the DEKF. An Optimal Experimental Design based on the best model follows, which calculates the optimal gas composition for all setpoint changes from $t_{2,\text{set}}$ to the end of the cultivation.

In this simulation study, every 12 hours at $t_{\mathrm{i,set}}$ optimal set gas fractions for the remaining cultivation time were calculated by online OED. In Figure 6.7, it is shown exemplified, how the optimized reference trajectory of a prior OED (filled, black circles) is joint to the subsequent OED result (filled, red circles) in order to avoid stepwise changing reference values and to fulfill the demands of the gas phase controller. By the time the subsequent

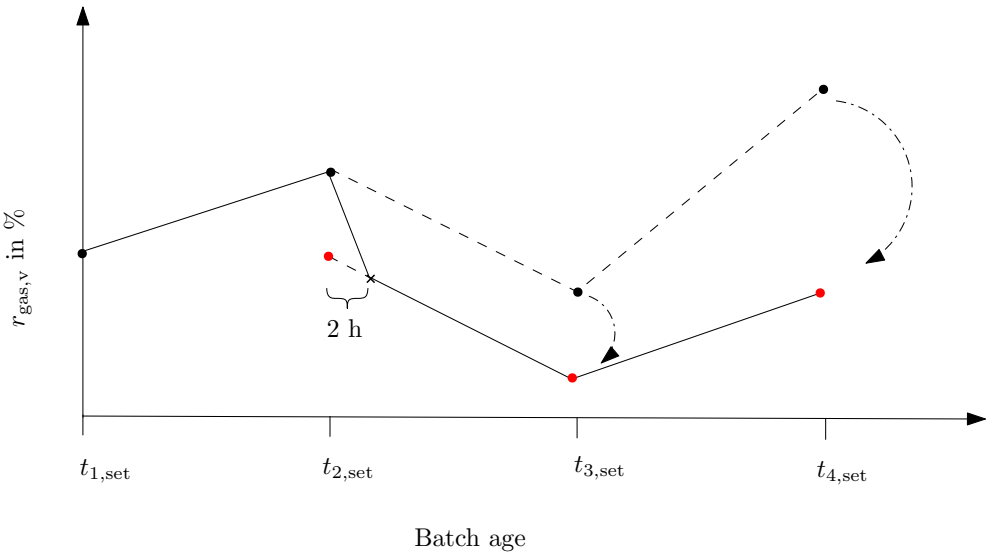


Figure 6.7: Smooth changes of reference values calculated by online OED

OED delivers optimized trajectories (here at $t_{2,\mathrm{set}}$), the actual reference value of the prior OED is linked in a linear manner to the value of the subsequent OED (cross), which is two hours in the future. The resulting black solid line is the relevant reference trajectory for the gas phase controller.
For the simulation study, a total cultivation time of 108 h was assumed. After the initial growth phase, multi-model OED was run for the remaining 84 h so that in the first OED, seven reference gas compositions were to be optimized. In the course of the experiment the number of remaining composition setpoints to be optimized decreased.

Simulation results

In this study, hydrogen, carbon dioxide and oxygen were allowed to vary between 55–100 %, 9–17 % and 17–27 %, respectively, for the first OED run at $t_{1,\mathrm{set}}$ (batch age 24 h). For the following OEDs, the range for oxygen was increased to 17–35 %. An adjustment of boundaries was required because long exposures to high dissolved oxygen as well as long starving periods harm the organism and those effects are not included in the process model (I). Instead of adjusting the reference boundaries, an alternative approach would be to run the optimization with constraints on the states for dissolved gases. However, it would increase computational costs drastically, and therefore boundaries for the references instead of state constraints were employed. Depending on the OED intervention batch age, resulting refer-

ence trajectories are given in Figure 6.8. The red lines resulted from the first OED at batch

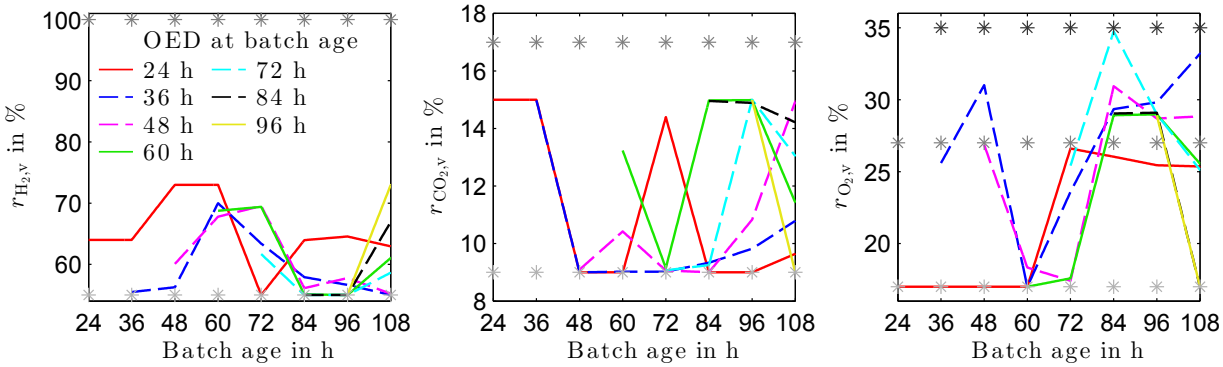


Figure 6.8: Reference trajectories of the gas fractions in the headspace that were optimized by OED. Upper reference boundaries are marked with dark gray asterisk symbols and lower ones with light gray.

age 24 h, in which an average cost function of all models was employed. After each parameter estimation step, the best model was selected and utilized for subsequent OEDs until the parameters were re-estimated. In Table 6.1, the parameter estimation results are listed. Parameters that belong to the selected model candidate are highlighted in gray. As the relevant models, respectively parameters, employed in OED changed at different $t_{i,\text{set}}$, the resulting reference trajectories differ in their course of development (see Fig 6.8). Another reason for different trajectory courses might be a maximum number of allowed cost function evaluations (which has to be defined by the operator) that implies a maximum optimization duration. In some optimizations, the OED was probably stopped before a global minimum was found. However, the more the cultivation proceeded, the lower were the computational costs and the more likely it was that the global minimum was found by the optimizer in the maximum time available. This is one of the reasons for convergence of reference trajectories that were calculated at a later stage of cultivation, which are yellow, black, turquoise and green in the figure. Besides a decrease in computational costs, a stable model number and converging model parameter values (see gray highlighted values in Table 6.1) are also assumed to cause convergence.

Evolving references are given in Figure 6.9 together with the simulated noisy process measurements (black) and corresponding estimated values (red) over time. Estimations were calculated with the best model candidate and updated parameters as in the real application. To this end, the covariance matrix of estimation error as well as the initial values of the states and the updated model were transferred to the Kalman filter at the time points after parameter estimation that were after 36 h, 60 h and 84 h. By passing the information to the updated model instead of running estimation with all model candidates in parallel

Table 6.1: Estimated parameter values during multi-model OED simulation study using SH models 1–3. Values in bold mark hit boundaries and gray backgrounds mark parameter sets of models that were temporarily selected by the routine.

Model	Name	Unit	Estimated values at batch age				Boundaries		Original values
			36 h	60 h	84 h	108 h	lb	ub	
1	$K_{\mathrm{SH,a}}\cdot$	$\mathrm{kU\cdot g^{-1}\cdot}$	0.68	0.66	0.69	0.96	0.001	1	
	$\mu_{\mathrm{SH,max}}$	-							
	$k^{\mathrm{SH}}_{1,\mathrm{O_2}}$	$\mathrm{mg\cdot L^{-1}}$	0.67	0.7	0.75	1.51	0.64	3.2	
	$k^{\mathrm{SH}}_{2,\mathrm{O_2}}$	-	5.8	**12**	5.7	4.88	4.5	12	
	$K_{\mathrm{SH,a}}\cdot$	$\mathrm{kU\cdot g^{-1}\cdot}$	0.026	**0.001**	0.026	**0.001**	0.001	0.4	
	$\mu_{\mathrm{SH,deg,max}}$	$\mathrm{h^{-1}}$							
	$k_{\mathrm{SH,deg}}$	$\mathrm{mg\cdot L^{-1}}$	0.8	**32**	0.12	1.84	0.032	32	
2	$K_{\mathrm{SH,a}}\cdot$	$\mathrm{kU\cdot g^{-1}\cdot}$	0.14	0.05	0.06	0.07	0.001	1	
	$\mu_{\mathrm{SH,max}}$	$\mathrm{h^{-1}}$							
	$k^{\mathrm{SH}}_{1,\mathrm{O_2}}$	$\mathrm{mg\cdot L^{-1}}$	**0.64**	0.81	0.76	0.74	0.64	3.2	
	$k^{\mathrm{SH}}_{2,\mathrm{O_2}}$	-	7.7	**4.5**	**4.5**	**4.5**	4.5	12	
	$k^{\mathrm{SH}}_{\mathrm{SH}}$	$\mathrm{mg_{Pr}\cdot U^{-1}}$	0.43	**0.001**	0.26	0.53	0.001	10	
	$K_{\mathrm{SH,a}}\cdot$	$\mathrm{kU\cdot g^{-1}\cdot}$	0.04	**0.01**	**0.01**	**0.01**	0.01	0.4	
	$\mu_{\mathrm{SH,deg,max}}$	$\mathrm{h^{-1}}$							
	$k_{\mathrm{SH,deg}}$	$\mathrm{mg\cdot L^{-1}}$	1.5	2.9	**3.2**	**3.2**	0.032	3.2	
3	$\mu_{\mathrm{SH,max}}$	$\mathrm{h^{-1}}$	0.003	0.01	0.01	0.01	0.001	1	0.01
	$a_{\mathrm{SH,max}}$	$\mathrm{U\cdot mg_{Pr}^{-1}}$	0.67	**5**	3.8	4.3	1	5	4.7
	$a_{\mathrm{SH,min}}$	$\mathrm{U\cdot mg_{Pr}^{-1}}$	0.026	**0.4**	0.03	0.003	0.001	0.4	0.03
	$k^{\mathrm{SH}}_{1,\mathrm{O_2}}$	$\mathrm{mg\cdot L^{-1}}$	0.65	0.66	0.66	0.66	0.32	3.2	0.65
	$k^{\mathrm{SH}}_{2,\mathrm{O_2}}$	-	5.44	2.41	**2.4**	2.41	2.4	12	2.59

and switching between the estimated trajectories, jumping estimates for SH were avoided. To obtain synthetic noisy measurements, a Gaussian noise term based on the measurement inaccuracies of Table 5.5 on page 88 was added to each simulated value. To avoid large calculation efforts for the OED routine, the gas controller dynamics were not included in this study. Consequently, the reference gas fractions are assumed to be realized without delay and thus equal the present gas fractions.

Until the first parameter estimation, a randomly selected model (here SH model 3) was used for state estimation. As mentioned above, after each model selection step, the new parameters and model number were passed on to the Kalman filter as an update.

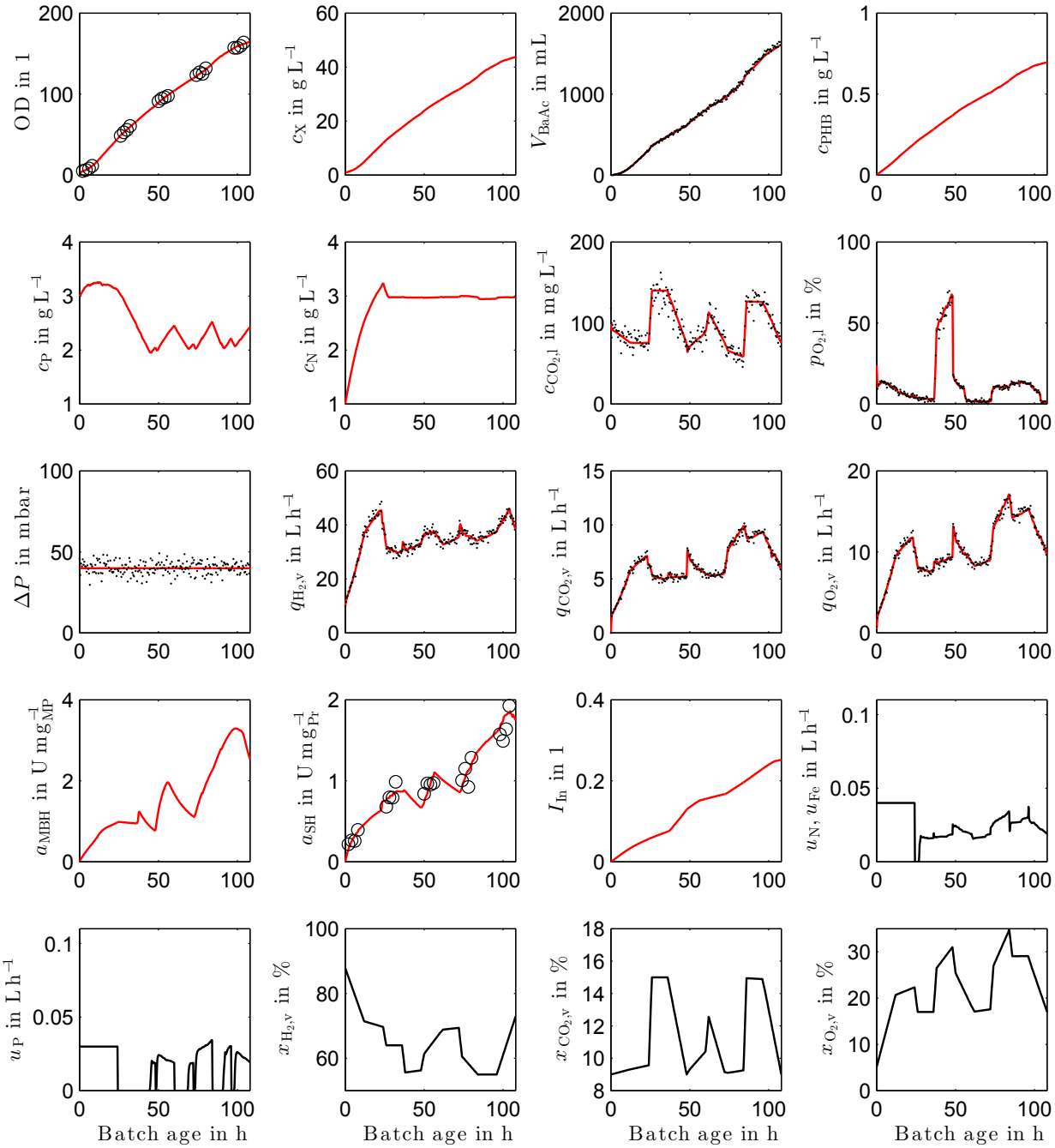


Figure 6.9: Noisy measurements (black circles or dots) versus estimated measurements (red), both derived by process model (I) combined with SH model 3, in the course of the multi-model OED cultivation. Model inputs are given in black lines.

Due to the SH description in all three models, fluctuating dissolved oxygen represents optimal stimuli, which can be seen by the courses of gas composition and dissolved oxygen concentration $p_{O_2,l}$ in the simulated cultivation, see Figure 6.9. For parameter estimation, relevant boundaries were set with respect to initially estimated parameter values using the data of prior experiments. After the second parameter estimation, SH model 2 was selected (see also Table 6.1). Apparently, its model structure is very flexible, and therefore it can

approximate best the noisy measurements that were originally generated by SH model 3. However, an evaluation of these identified parameters showed that three of the six parameters hit boundaries, generally indicating model structure deficits. The finally identified parameter values are very close to the original ones of SH model 3 that were used to generate SH measurements (see Table 6.1). However, in the course of the cultivation, different parameter values are adopted for SH model 3 and even boundaries were hit, indicating that the parameters were difficult to identify. Identifiability problems may result from correlated and/or little sensitive parameters. But due to perfect stimuli, the sensitivities were increased and these parameters became identifiable. So they converged towards their original values.

6.3.2 Experimental results

In the simulation study presented above, multi-model online OED was run successfully. But when employing it in reality, various difficulties occurred in these cultivations. In one experiment, as described in Neddermeyer et al. (2016), the general growth part of process model (I) at that time had major deficiencies, and therefore a final modeling step was required and OED led to unsatisfactory results. Another OED controlled cultivation failed, because the optimizer calculated long periods with concentrations of low dissolved oxygen and the cells supposedly starved to death. In another cultivation, the yield of ammonium per biomass was higher than foreseen by the process model (I). Therefore, ammonium was limited in periods of cultivation and PHB was produced, which interfered with the SH analysis. Due to time limitations, no further experiments could be run. Therefore, the promising simulation results still need to be confirmed in real experiments. Despite this open issue, detailed execution notes for multi-model online OED in cultivations will be given in the Appendix A.

6.3.3 Possible improvements for multi-model OED

In hindsight, a wrong decision was made to test multi-model OED by means of an extension of the existing model (I) with respect to SH production. As outlined in Section 3.2, analysis of SH is a very elaborate, time-consuming step that did not allow for frequent and fast measurements and limited the number of possible test runs. To run the algorithm successfully, high frequent measurements would be useful to show differences in the structures of the model candidates for model selection. Moreover, the measurement values should be present with little time delay. Spectroscopy may be an attractive alternative, which is why strain H798, introduced in Section 2.3, was investigated for potential online OED cultivations. This strain produces an online measurable fluorescent protein together with SH and the investigation results will be given in Section 7.3.
During this thesis, the online OED could often only be run with an inaccurate version of

the process model (I), which neglects gas controller dynamics. When taking the gas controller behavior into account as described in the Appendix B, step size limitations occurred although the solver `ode15s` was used that is specialized on stiff systems. To overcome this limitation, a powerful environment equipped with even better specialized solvers for differential equation systems should be integrated into the software used at the Chair of Measurement and Control.
In the experiments, reference boundaries had to be adjusted frequently, which makes a permanent supervision by an experienced operator indispensable. By permanent adjustments, drifts into unmodelled process states are prevented. Generally, the optimizer tended to push the relevant references (in the simulation study $r_{O_2,v}$) towards their boundaries (see Figure 6.8) and, if not prevented by the operator, also towards the definition limits of the model. It would be desirable if the reference boundaries were adjusted automatically.
Major problems concerning the model structure are that the growth model (here growth part of process model (I)) needs to be accurate for a successfully working multi-model online OED, otherwise optimizer and estimator cannot give satisfying results. Moreover, only model branches defined a priori can be tested, since the modeling step is not updated dynamically during the experiment. In case the initial model extensions are structurally wrong, the algorithm cannot succeed. Thus, multi-model online OED depends strongly as well on the modeling experience of the operator.

Once the Online OED has been modified so that it can be successfully used in cultivations, it is a method that allows new products to be modeled comparatively quickly. In other words, if the data from cultivations with mutant strains have been used to adapt growth and gas transport units of process model (I), online OED can be used to derive models for the production unit of these strains. Also, the methods of "phenomena recognition" and "dependency analysis of approximated rates" presented above can be utilized for autotrophic cultivation of *R. e.* mutants to develop models for the formation of product compounds. Both methods were employed for the strain HF951 to find a description for the product cyanophycin. The resulting model for HF951 and adaption results of process model (I) for the strains HF805 and H798 are presented and discussed in the next chapter.

Chapter 7

Adapted models

The mutant strains of *R. e.* from Chapter 2 were cultivated in the system of Chapter 4 to collect data as described in Chapter 3 for the adaption of process model (I) presented in Chapter 5. For the adaption, parameters of the growth and gas transport units were re-estimated and the methods "dependence analysis of approximated rates" and "phenomena recognition" of the Chapter 6 were employed to model the production unit. The resulting adapted models are given and discussed in the following sections.

7.1 Adapted model of the strain HF805

The strain *R. e.* HF805 cannot express SH, translates a tagged MBH and is known to grow more slowly than the wild-type H16. One cultivation had been carried out with this strain. Figure 7.1 visualizes the cultivation data together with the simulations (red) of the adapted model.

In order to adapt the process model (I), first, the growth metabolism parameters concerning active biomass and the formation of PHB were adjusted to the HF805 cultivation data. Since during the entire experiment the dissolved oxygen concentration was below the threshold for long-term growth inhibition, all parameter values of the inhibitory state of eq. (5.59) on page 79 were taken from process model (I). Moreover, the experiments were carried out in such a way that PHB could not be converted to active biomass, and therefore, the corresponding parameters were kept constant at the values of process model (I). Identified parameters of the adapted model for HF805 are given in Table 7.1.

The parameter $k_{\mathrm{xPHB}}^{\mathrm{PHB}}$ of eq. (5.48) on page 75, defining the inhibition of PHB formation for high amounts of PHB, was identified in the parameter estimation to 0.1005 which is very close to the set minimum. When the identified value is inserted into the Aiba function, it has practically no influence on the rate μ_{PHB} any more. It can therefore be concluded that the cultivation data at hand do not provide enough information to determine the

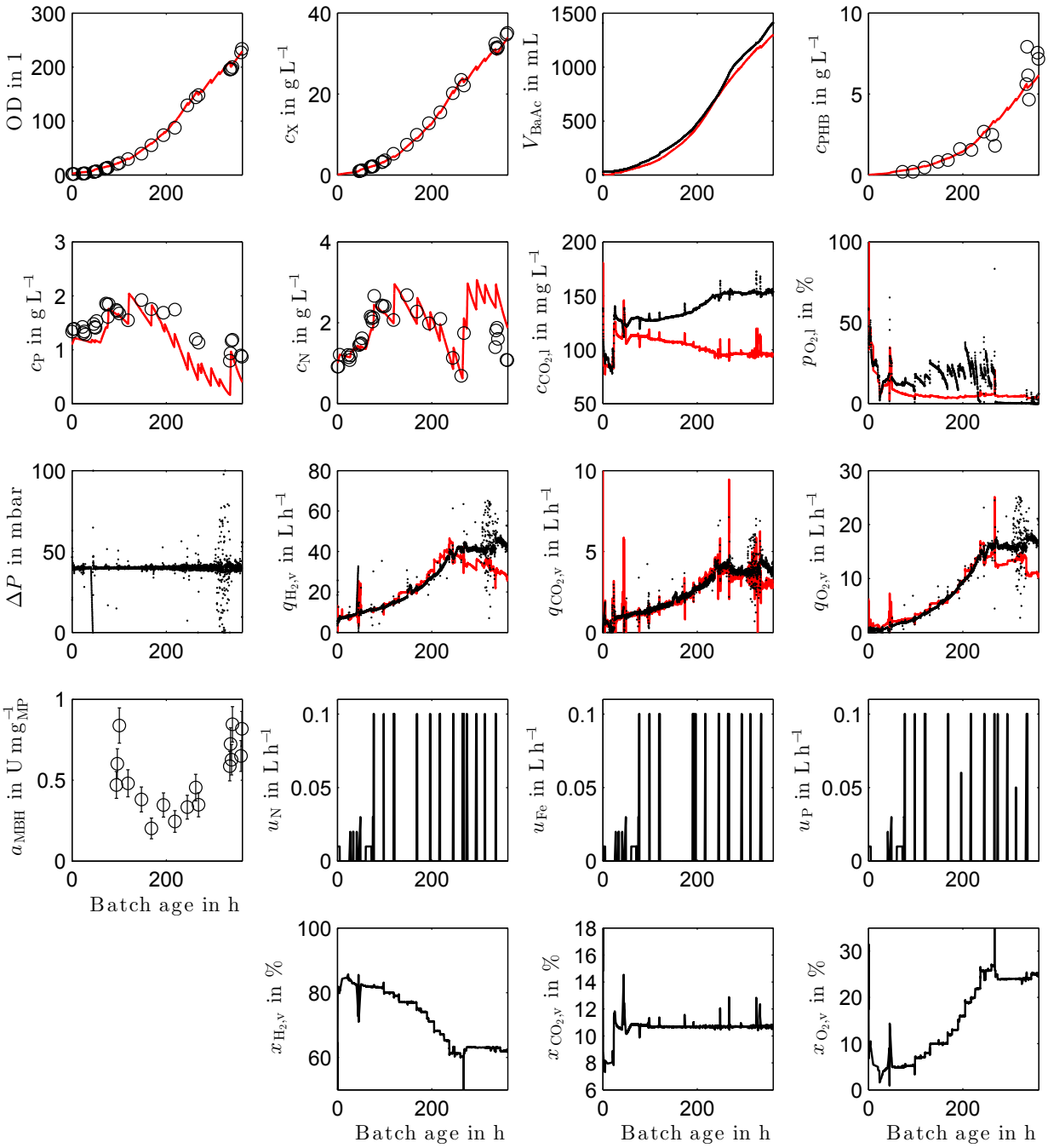


Figure 7.1: Cultivation data (black) of strain HF805 and the adapted model simulations (red) derived from process model (I). The figures in the lower two rows with exclusively black lines show relevant model inputs.

parameter k_{xPHB}^{PHB}, because high PHB amounts were never reached during the cultivation. Also the parameters $k_{H_2}^X$, $k_{CO_2}^X$ and k_N^X are quite close to their lower or upper limit and beyond that these parameters are strongly coupled in the equation for biomass production (see eq. (5.47)). With a more dynamic stimulation of the corresponding substrates, the parameters could probably have been estimated more precisely.

Table 7.1: Estimated parameter values of the adapted model HF805 (derived from process model (I)) and lower and upper boundaries (lb, ub) applied as constraints for the parameter estimation

Parameter name	Unit	Parameter value	lb	ub
$F_{\mathrm{PHB,gr}}$	-	0.31	0	50
$U_{\mathrm{N,X}}$	$\mathrm{g\,g_X^{-1}}$	0.20	0.08	0.3
$U_{\mathrm{P,X}}$	$\mathrm{g\,g_X^{-1}}$	0.12	0.004	0.3
$U_{\mathrm{H_2,X}}$	$\mathrm{g\,g_X^{-1}}$	1.13	0.3	1.4
$U_{\mathrm{CO_2,X}}$	$\mathrm{g\,g_X^{-1}}$	1.37	0.4	4
$U_{\mathrm{O_2,X}}$	$\mathrm{g\,g_X^{-1}}$	4.2	1.6	16
$U_{\mathrm{H_2,PHB}}$	$\mathrm{g\,g_{PHB}^{-1}}$	2.7	0.1	3.6
$U_{\mathrm{CO_2,PHB}}$	$\mathrm{g\,g_{PHB}^{-1}}$	8.9	1.3	14
$U_{\mathrm{O_2,PHB}}$	$\mathrm{g\,g_{PHB}^{-1}}$	25	1.3	58
$\mu_{\mathrm{X,max}}$	$\mathrm{h^{-1}}$	0.32	0.1	2
$k^{\mathrm{X}}_{\mathrm{H_2}}$	$\mathrm{mg\,L^{-1}}$	0.002	$2 \cdot 10^{-4}$	1.2
$k^{\mathrm{X}}_{\mathrm{CO_2}}$	$\mathrm{mg\,L^{-1}}$	432	4.4	440
$k^{\mathrm{X}}_{1,\mathrm{O_2}}$	$\mathrm{mg\,L^{-1}}$	3.1	0.16	6.4
$k^{\mathrm{X}}_{2,\mathrm{O_2}}$	-	8.5	1	15
$k^{\mathrm{X}}_{\mathrm{N}}$	$\mathrm{g\,L^{-1}}$	0.11	0.1	0.8
$k^{\mathrm{X}}_{\mathrm{P}}$	$\mathrm{g\,L^{-1}}$	2	0.04	3.5
$\mu_{\mathrm{PHB,max}}$	$\mathrm{h^{-1}}$	0.52	0	0.85
$k^{\mathrm{PHB}}_{\mathrm{xPHB}}$	-	0.1005	0.1	10
$k^{\mathrm{PHB}}_{\mathrm{P}}$	$\mathrm{L\,g^{-1}}$	28	1	300
$k^{\mathrm{PHB}}_{\mathrm{N}}$	$\mathrm{L\,g^{-1}}$	18	0.2	20

In a cultivation, the concentration of a dissolved gas that is consumed can at most correspond to the saturation concentration. When the organisms consume gas, $c_{\mathrm{gas,l}}$ is often lower than $c_{\mathrm{gas,sat}}$. In this experiment, $c_{\mathrm{CO_2,l}}$ equals $c_{\mathrm{CO_2,sat}}$ only in the beginning until $x_{\mathrm{CO_2,v}}$ was increased from 8 % to 11 %. Remarkably, from then, $c_{\mathrm{CO_2,l}}$ increased in the course of the cultivation although the carbon dioxide fraction in the headspace, and, as a consequence, also the dissolved saturation concentration $c_{\mathrm{CO_2,sat}}$ were constant. Therefore, the course of dissolved CO_2 could hypothetically be explained by microbial production, e.g., from an unknown heterotrophic process. However, if the organisms had released carbon dioxide into the liquid phase, it would have been gassed out. The controller would have counteracted a change of the gas composition in the gas phase by reducing the fraction of CO_2 in the total feed flow. But the proportion of CO_2 in the total feed flow was constant at approximately 7 %. Hence, an increase of $c_{\mathrm{CO_2,l}}$ can only be explained with a drift of the probe. Such

drift may result from a changing composition of the liquid broth (organic and anorganic), as described by Zosel et al. (2011).
When comparing the simulations of the adapted model to the measurements (see Figure 7.1), oxygen and hydrogen gas flows seem underestimated by the model from hour 250 h on. Possibly, the strain HF805 required energy for maintenance, which was not included in the model structure of process model (I). Therefore, after model adaption, i.e., parameter estimation of the model (I) with cultivation data from HF805, maintenance cannot be simulated as structural changes would have been necessary. Moreover, $p_{O_2,l}$ was not in accordance with the measurements. Thus, it was decided to proceed with process model (II) in order to find a mathematical description for the product tagged MBH. Before it was used for MBH modeling, model (II) was adapted for HF805. This procedure will now be briefly presented and the resulting model (II) for HF805 discussed.
First, the concentration of dissolved hydrogen was calculated as explained in Section 5.4. The measured values of $p_{O_2,l}$ from batch age $\approx$ 270 h to the end were beyond the threshold of 1 %. Before using them as input for model (II), they were corrected together with the calculated values of $c^{n}_{H_2,l}$ according to Section 5.5 and then the parameters of process model (II) were estimated. In Figure 7.2, the simulations (red) of model (II) are plotted together with the measurements and model inputs. At batch age 335 h, in a phase during which the $p_{O_2,l}$ values were corrected, the base volume flow had increased slightly, as can be seen from the slope of the base. As base is considered as a soft sensor for active biomass formation, the cells seem to have grown faster. This can be explained by examining $x_{O_2,v}$ and $q_{gas,v}$ of model (I) in Figure 7.1 at batch age 335 h. The fraction of O_2 was slightly increased that led to higher gas feed flows, which indicates an improved production of active biomass. This effect, though, cannot be captured by the model (II), since all gas values below a certain threshold were corrected to the same value. Instead of a uniform gas correction as used in this thesis, a correction in dependence of the gas feed flows would be more accurate. All in all, however, biomass, ammonium and phosphate were simulated to a satisfactorily degree so that this model (II) was employed to find a description for tagged MBH by manual modeling. But neither by this, nor via the adaption tools "phenomena recognition" and "dependency analysis of approximated rates" a suitable model for the tagged MBH was found. High measurement to noise ratios of the values of tagged MBH made modeling challenging. The approximated measurement uncertainties are indicated in the figures and were calculated as explained in Section 3.2. Additional experiments, in which the MBH develops more dynamically, are required to allow for a mathematical formulation and a subsequent parameter estimation. As in the last chapter, the decision to test these new methods with substances that are difficult to measure turned out to be not the best choice for a first experimental application. Nevertheless, major parts of the growth model of HF805 could be adapted to a satisfactory level with an acceptable workload starting with process model (I).

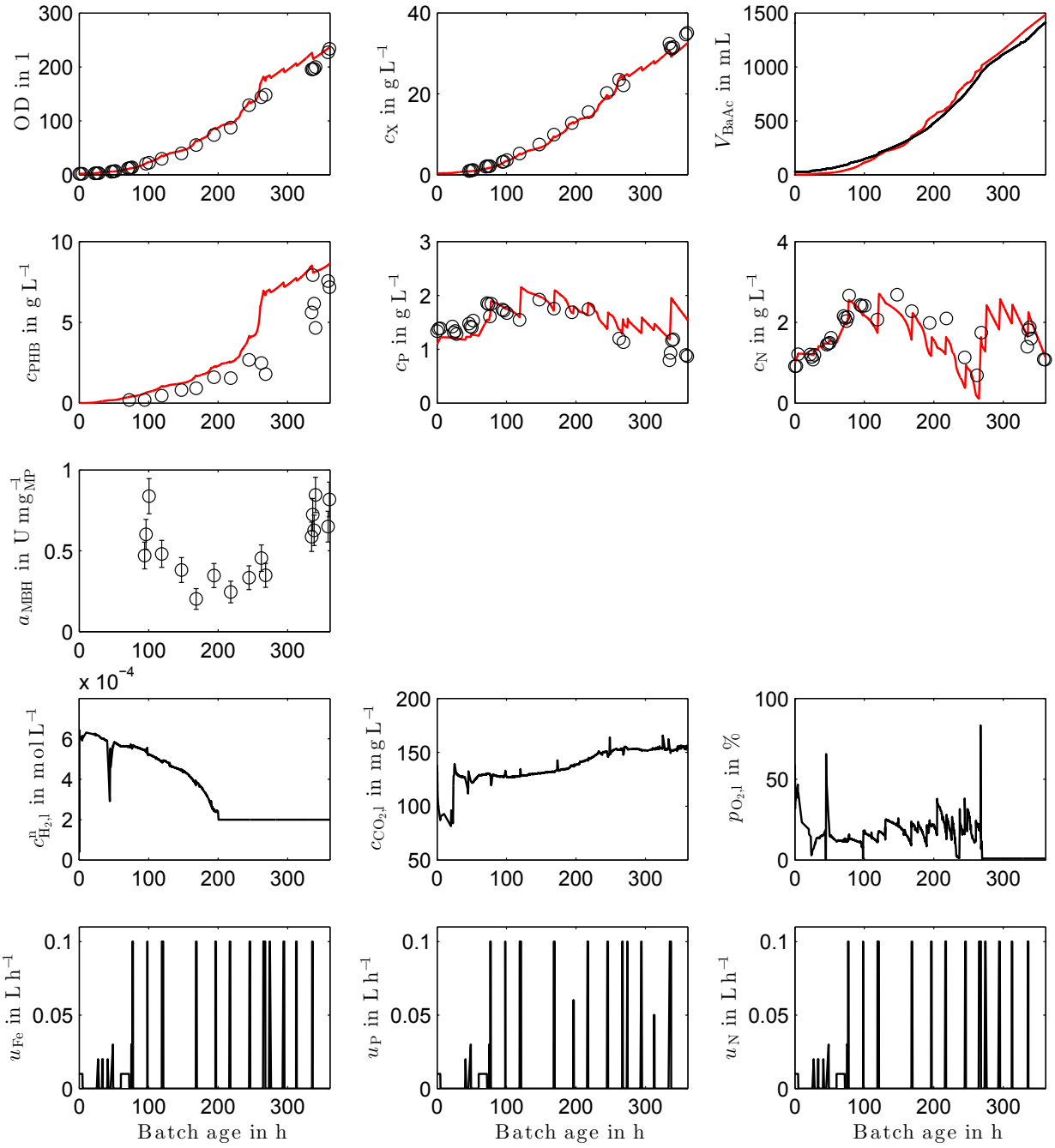


Figure 7.2: Cultivation data of strain HF805 (black) and the adapted model simulations (red) derived from process model (II). Figures of the two lower rows show relevant model inputs.

7.2 Adapted model of the strain HF951

At the chair of Measurement and Control, prior to this thesis, the strain HF951 was grown autotrophically in cultivations HF951A–E that were utilized for parameter estimations and "phenomena recognition". Furthermore, HF951A was employed in "dependency analysis of approximated rates". HF951 produces cyanophycin (cy), an internal molecule consisting mainly of nitrogen and carbon. *R. e.* HF951 cannot synthesize the carbon storage compound

PHB. For model adaption, the first step were structural alterations of the growth unit. PHB was removed for the aforementioned metabolic reason. Then, the remaining growth parameters, including long-term inhibition x_{In} caused by $c_{\mathrm{O_2,l}}$, were estimated leading to the values listed in Table 7.2. Similar to H16, the simulation results were better when the inhibition state x_{In} was included. In contrast to HF805, there were phases with high oxygen concentrations during cultivations of the strain HF951, which made it possible to estimate the relevant parameters for x_{In}. The adapted process model was then extended for

Table 7.2: Estimated parameter values of the adapted model HF951 (derived from process model (I)) and lower and upper boundaries (lb, ub) applied as constraints for the parameter estimation

Parameter name	Unit	Parameter value	lb	ub
$U_{\mathrm{N,X}}$	$\mathrm{g\,g_X^{-1}}$	0.18	0.08	0.3
$U_{\mathrm{P,X}}$	$\mathrm{g\,g_X^{-1}}$	0.05	0.004	0.3
$U_{\mathrm{H_2,X}}$	$\mathrm{g\,g_X^{-1}}$	0.44	0.3	1.4
$U_{\mathrm{CO_2,X}}$	$\mathrm{g\,g_X^{-1}}$	1.4	0.4	4
$U_{\mathrm{O_2,X}}$	$\mathrm{g\,g_X^{-1}}$	3.3	1.6	16
$\mu_{\mathrm{X,max}}$	$\mathrm{h^{-1}}$	1.6	0.1	2
$k_{\mathrm{H_2}}^{\mathrm{X}}$	$\mathrm{mg\,L^{-1}}$	0.70	$2 \cdot 10^{-4}$	1.2
$k_{\mathrm{CO_2}}^{\mathrm{X}}$	$\mathrm{mg\,L^{-1}}$	14.7	4.4	440
$k_{\mathrm{1,O_2}}^{\mathrm{X}}$	$\mathrm{mg\,L^{-1}}$	3.1	0.16	6.4
$k_{\mathrm{2,O_2}}^{\mathrm{X}}$	-	8.7	1	15
$k_{\mathrm{N}}^{\mathrm{X}}$	$\mathrm{g\,L^{-1}}$	0.58	0.1	0.8
$k_{\mathrm{P}}^{\mathrm{X}}$	$\mathrm{g\,L^{-1}}$	0.041	0.04	3.5
$k_{\mathrm{In,max}}$	$\mathrm{h^{-1}}$	0.008	0.001	0.01
$K_{\mathrm{In,deg}}$	$\mathrm{g^{-1}}$	$2 \cdot 10^{-5}$	10^{-6}	10^{-3}
$k_{\mathrm{1,O_2}}^{\mathrm{In}}$	$\mathrm{mg^{0.58}\,L^{-0.58}}$	3	0.75	5
$k_{\mathrm{2,O_2}}^{\mathrm{In}}$	-	0.58	0.07	0.65

the production unit describing the formation of cyanophycin. As cyanophycin could not be quantified directly, the values were calculated on base of measured OD and c_{X}. Similar to PHB in eq. (5.83) on page 84, it was therefore assumed that cyanophycin affects the optical density in a linear fashion,

$$\mathrm{OD} = K_{\mathrm{OD,X}} \cdot \frac{m_{\mathrm{X}}}{V_{\mathrm{l}}} + K_{\mathrm{OD,cy}} \cdot \frac{m_{\mathrm{cy}}}{V_{\mathrm{l}}}. \tag{7.1}$$

The proportion of cellular cyanophycin only accounts for up to 10 % of total biomass and the relative measurement error of c_{X} is about 8 % (see Table 5.5). The contribution of

cyanophycin to c_X almost disappears in the inaccuracy of the measurement. Thus, the measured total biomass concentration c_X was assumed to equal active biomass m_X/V_l to calculate an indirect cyanophycin concentration, i.e.,

$$K_{OD,cy} \cdot c_{cy} = OD - K_{OD,X} \cdot c_X, \tag{7.2}$$

with $K_{OD,cy}$ being the constant factor indicating to what extent cyanophycin affects OD. The inaccuracies of this calculation method, which is based on the assumption that cyanophycin has a minor effect on the measured cell dry weight, can be evaluated as measurement noise of the indirect and calculated cyanophycin measurements.
Furthermore, it was assumed that for H16 and HF951, the same linear relationship between active biomass and OD applies. Hence, $K_{OD,X}$ from eq. (7.2) is set to a value of 5.8 $L\,g^{-1}$ similar to H16. Thus, for samples in which c_X and OD have been analyzed, the indirect cyanophycin concentration $K_{OD,cy} \cdot c_{cy}$ can be calculated. Analogously, the state equation for indirect cyanophycin is set up as follows

$$K_{OD,cy} \cdot \dot{m}_{cy} = K_{OD,cy} \cdot \mu_{cy} \cdot m_X. \tag{7.3}$$

By means of "phenomena recognition" (see Section 6.2) and "dependency analysis of approximated rates" for cyanophycin shown in Figure 6.4 on page 116, the formation rate μ_{cy} was assumed to depend on $c_{CO_2,l}$ in an inhibiting manner

$$\mu_{cy} = \mu_{cy,max} \cdot (1 - I_{In}) \cdot Ai(c_{CO_2,l}, k^{cy}_{CO_2}) \cdot Ro_2(c_{O_2,l}, [k^{cy}_{1,O_2}\ k^{cy}_{2,O_2}]). \tag{7.4}$$

Additionally, similar to general growth and product formation of MBH, PHB, SH in H16, a limiting and inhibiting dependency on $c_{O_2,l}$ was implemented using the kinetic function Ro_2. As for the production of active biomass and PHB in H16, including the long-term inhibition I_{In} for cyanophycin caused by high oxygen leads to better results. After multiplying both sides of eq. (7.4) with $K_{OD,cy}$,

$$K_{OD,cy} \cdot \mu_{cy} = K_{OD,cy} \cdot \mu_{cy,max} \cdot (1 - I_{In}) \cdot Ai(c_{CO_2,l}, k^{cy}_{CO_2}) \cdot Ro_2(c_{O_2,l}, [k^{cy}_{1,O_2}\ k^{cy}_{2,O_2}]), \tag{7.5}$$

this can be inserted into the cyanophycin state eq. (7.3). The parameters $K_{OD,cy} \cdot \mu_{cy,max}$ can only be identified together. Degradation of cyanophycin is not included although a conversion of cyanophycin to active biomass is possible from a metabolic point of view when nutrient limitation occurs. However, the present experimental data are not sufficient to allow for a model including degradation. To account for nitrogen used to build cyanophycin, the ammonium balance of process model (I) as in eq. (5.76) was changed to

$$\dot{m}_N = -U_{N,X} \cdot \mu_{X,In} \cdot m_X - U_{N,cy} \cdot K_{OD,cy} \cdot \mu_{cy} \cdot m_X + u_N \cdot c_{N,feed}. \tag{7.6}$$

Table 7.3: Estimated parameters of the cyanophycin metabolism in HF951

Parameter	Unit	Parameter value	lb	ub	rel. std. dev. (%)
$K_{\mathrm{X,cy}}$	$\mathrm{g\,L^{-1}}$	0.07	10^{-6}	0.5	0.45
$K_{\mathrm{OD,cy}} \cdot \mu_{\mathrm{cy,max}}$	$\mathrm{L\,g^{-1}h^{-1}}$	0.21	10^{-5}	0.5	0.25
$k^{\mathrm{cy}}_{\mathrm{CO_2}}$	$\mathrm{mg\,L^{-1}}$	15.5	$4.4 \cdot 10^{-4}$	4401	0.45
$k^{\mathrm{cy}}_{1,\mathrm{O_2}}$	$\mathrm{mg\,L^{-1}}$	1.6	0.32	3.2	0.19
$k^{\mathrm{cy}}_{2,\mathrm{O_2}}$	-	8.3	2	12	<0.1
$U_{\mathrm{N,cy}}$	$\mathrm{g_N}$	$1.1 \cdot 10^{-6}$	0	0.5	3670

When calculating indirect cyanophycin measurements by eq. (7.2), the impact of cyanophycin on cell dry weight was neglected and the resulting inaccuracies were treated as measurement noise. However, for the sake of an accurate description in the model, the simulated cell dry weight is assumed to be affected by cyanophycin, resulting in

$$y_1 = \frac{m_{\mathrm{X}}}{V_{\mathrm{l}}} + K_{\mathrm{X,cy}} \cdot \frac{K_{\mathrm{OD,cy}} \cdot m_{\mathrm{cy}}}{V_{\mathrm{l}}}. \tag{7.7}$$

The second model output y_2 is the OD from eq. (7.1). Estimated parameter values of the cyanophycin production unit are given in Table 7.3. The parameter standard deviations were calculated with the Fisher information matrix as in eq. (3.8) on page 26 and give only a lower bound of the true uncertainties. In this uncertainty analysis, the sum of the quadratic simulation errors corrected for the number of cyanophycin-parameters was employed as measurement noise. The simulations of the obtained model are presented in Figure 7.3–7.7.

Although the identified value for $k^{\mathrm{X}}_{\mathrm{N}}$, which refers to the ammonium concentration and growth rate, has a small standard deviation and is within the boundaries, it is questionable with respect to large discrepancies between simulated and measured ammonium. Hence, the development of ammonium cannot be well described by considering consumption for growth and cyanophycin as in eq. (7.6). Therefore, it seems as if there was an additional consumer not captured by the model. Evaluating the approximated standard deviation of $U_{\mathrm{N,cy}}$, the consumption coefficient parameter is classified as insignificant and cyanophycin related ammonium uptake can be neglected. Even without having calculated the standard deviations of the growth model parameters, conclusions of their reliability can be drawn. The parameter $k^{\mathrm{X}}_{\mathrm{P}}$ was estimated to a value of 0.041 $\mathrm{g\,L^{-1}}$, while the measured phosphate concentrations in the cultivations were between 1.5–4.5 $\mathrm{g\,L^{-1}}$. Such a low parameter value close to the lower boundary at comparatively large c_{P} implies that phosphate is not limiting growth during the entire experiment according to the model. Further experiments with lower phosphate concentrations would allow a more accurate estimation of $k^{\mathrm{X}}_{\mathrm{P}}$.

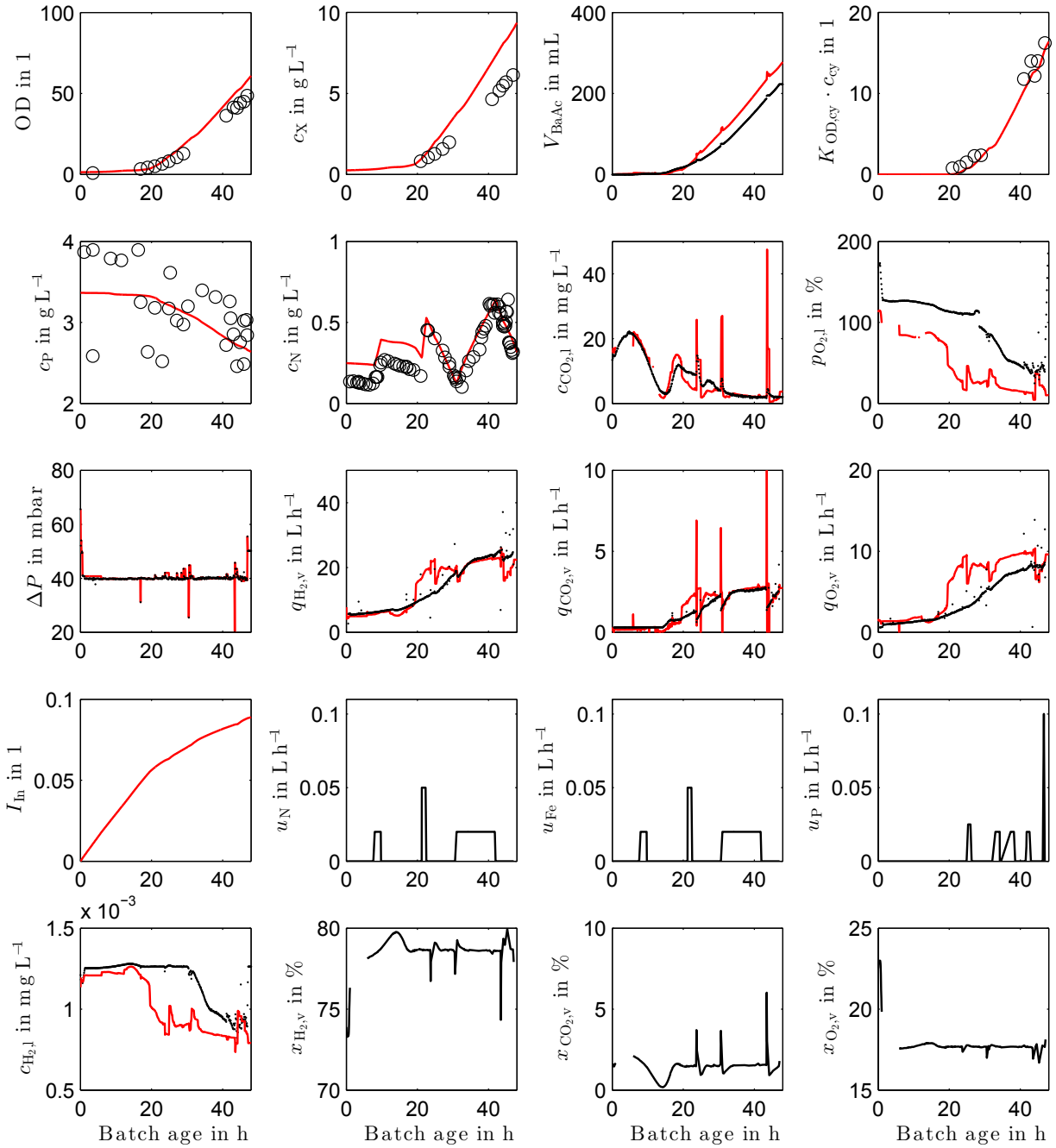


Figure 7.3: Cultivation A of the strain HF951. These data were employed in "dependency analysis of approximated rates" to model cyanophycin. Measurements are given in black and the simulation of the adapted model in red. The synthetic measurements of $c_{H_2,l}$ were calculated on a data-driven basis. The figures with exclusively black lines show relevant model inputs.

To employ "phenomena recognition" and "dependency analysis of approximated rates", $c_{H_2,l}$ had to be calculated according to Section 5.4. To this end, the gas consumption ratio of eq. (5.109) on page 105, which applies for H16, had to be customized. The factor 3 was replaced by the quotient of estimated consumption rates from the adapted process model (I)

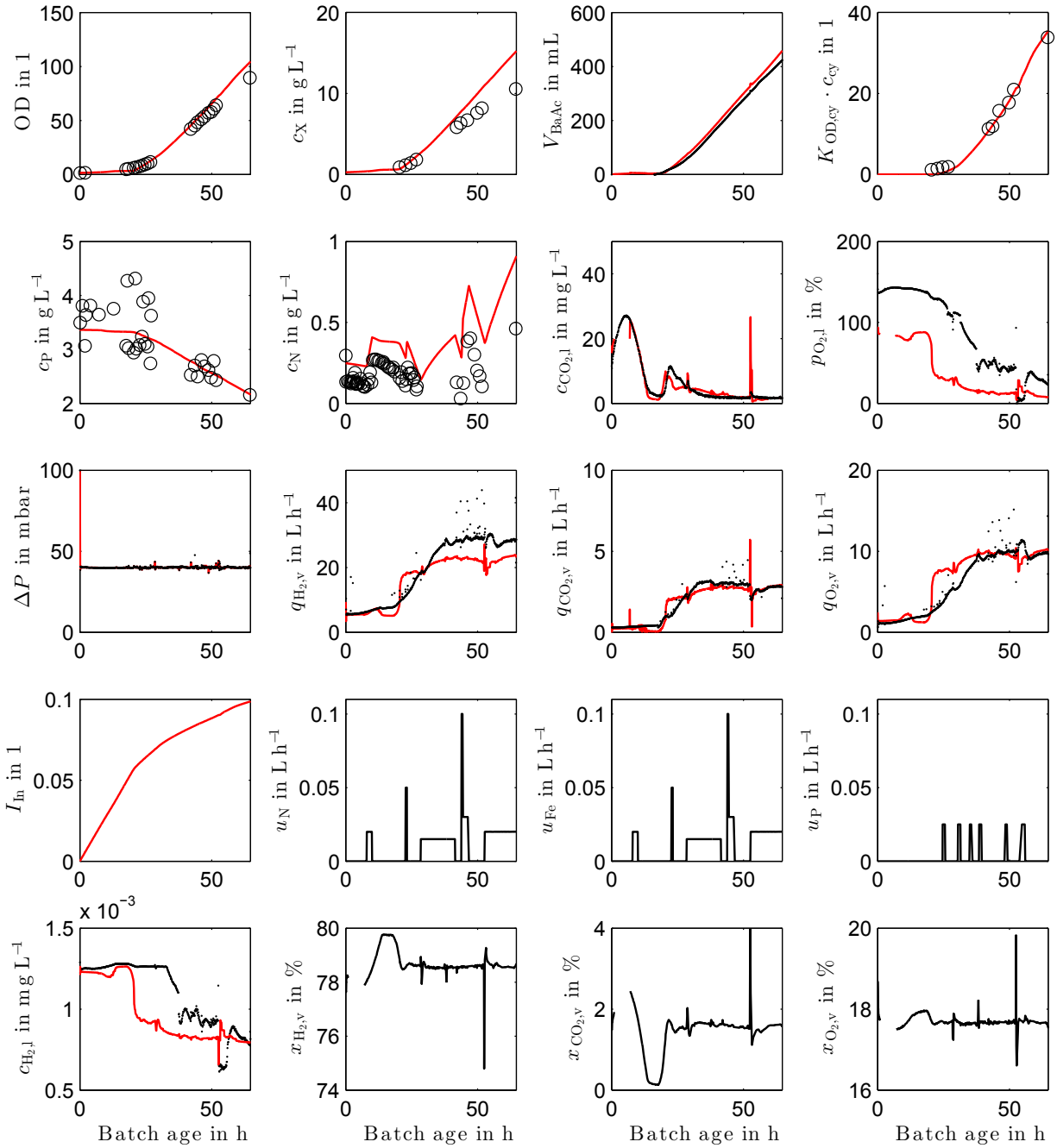


Figure 7.4: Cultivation B of the strain HF951. Measurements are given in black and the simulation of the adapted model in red. The synthetic measurements of $c_{H_2,l}$ were calculated on a data-driven basis. The figures with exclusively black lines show relevant model inputs.

for the strain HF951, i.e., the quotient of $U_{H_2,X}/U_{O_2,X}$ (given in Table 7.2), with the coefficients converted to molar values before division. This way, the H_2/O_2 ratio yielded 2 for HF951 that was considered when calculating $c_{H_2,l}$.

The cultivation data of HF951A shown in Figure 7.3 served for "dependency analysis of approximated rates". In this experiment and HF951B (see Figure 7.4), the simulated oxygen

decreased faster than the measured one. As a consequence, simulated $c_{H_2,l}$ also decreased faster than the calculated one, which indicates model deficiencies concerning gas transport respectively consumption. This comparison is shown for HF951A in the lower row on the left with simulated and calculated $c_{H_2,l}$ in red and black, respectively. But these large differences between simulated and measured $p_{O_2,l}$ also might have been caused by an incorrectly calibrated oxygen sensor. This assumption is based on the fact that, at the beginning of a cultivation, the microbial consumption rate is very low, and therefore $p_{O_2,l}$ should have equaled $p_{O_2,sat}$. However, since in both experiments measured $p_{O_2,l}$ is higher than the saturation concentration, an incorrect sensor calibration seems likely. In the other cultivations of HF951, the dissolved gases are fairly well approximated by the simulations.

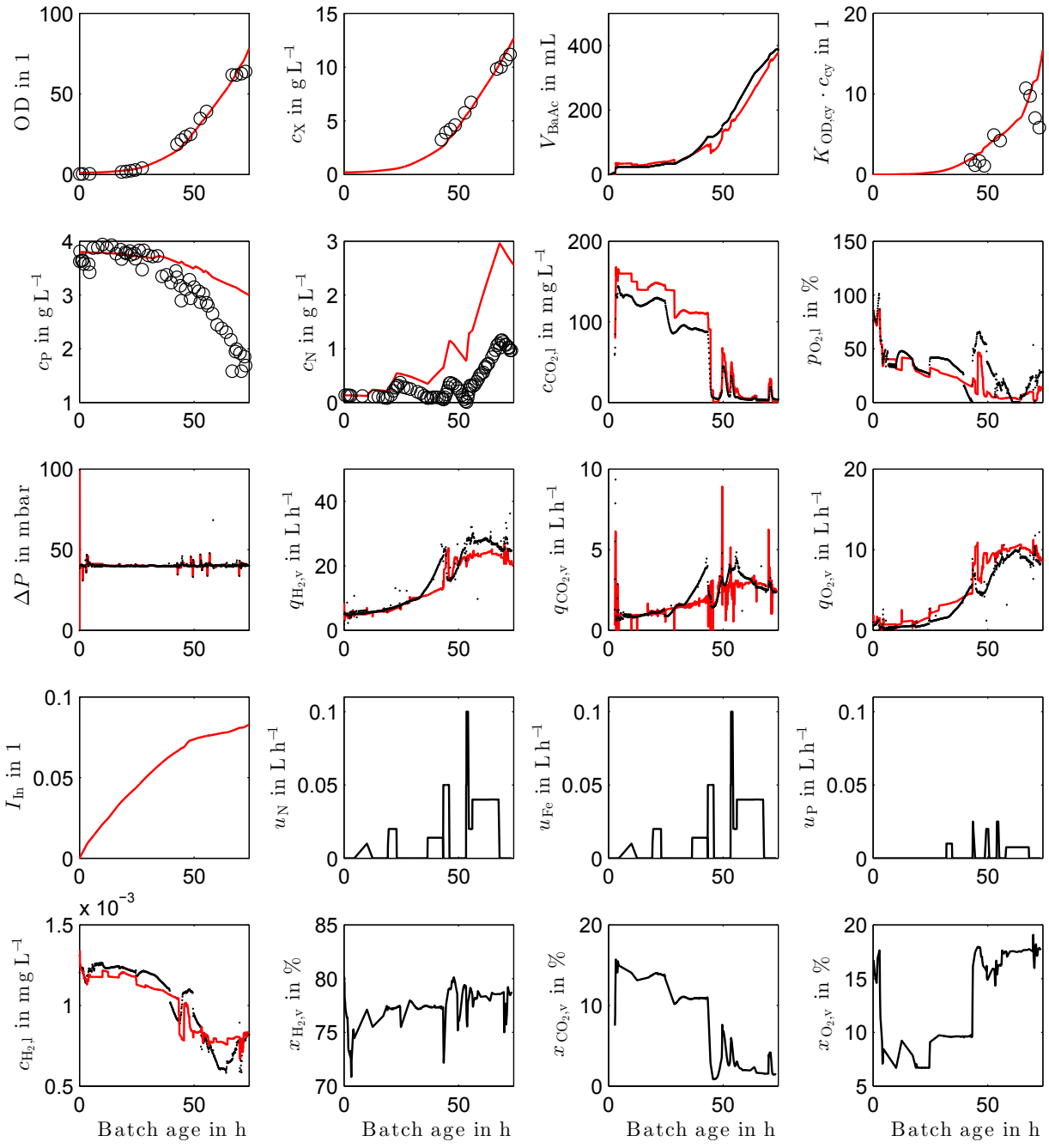


Figure 7.5: Cultivation C of the strain HF951. Measurements are given in black and the simulation of the adapted model in red. The synthetic measurements of $c_{H_2,l}$ were calculated on a data-driven basis. The figures with exclusively black lines show relevant model inputs.

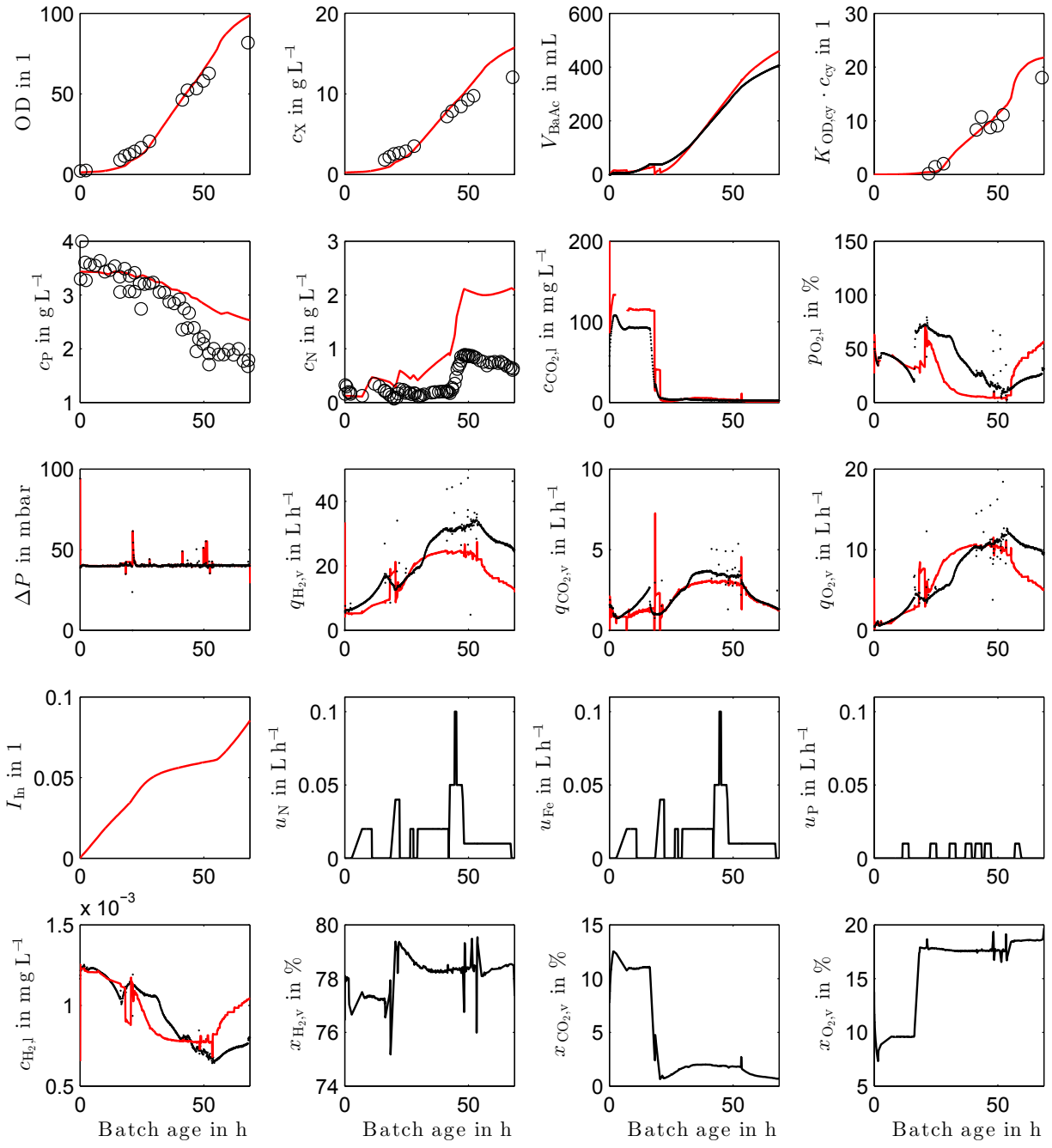


Figure 7.6: Cultivation D of the strain HF951. Measurements are given in black and the simulation of the adapted model in red. The synthetic measurements of $c_{H_2,l}$ were calculated on a data-driven basis. The figures with exclusively black lines show relevant model inputs.

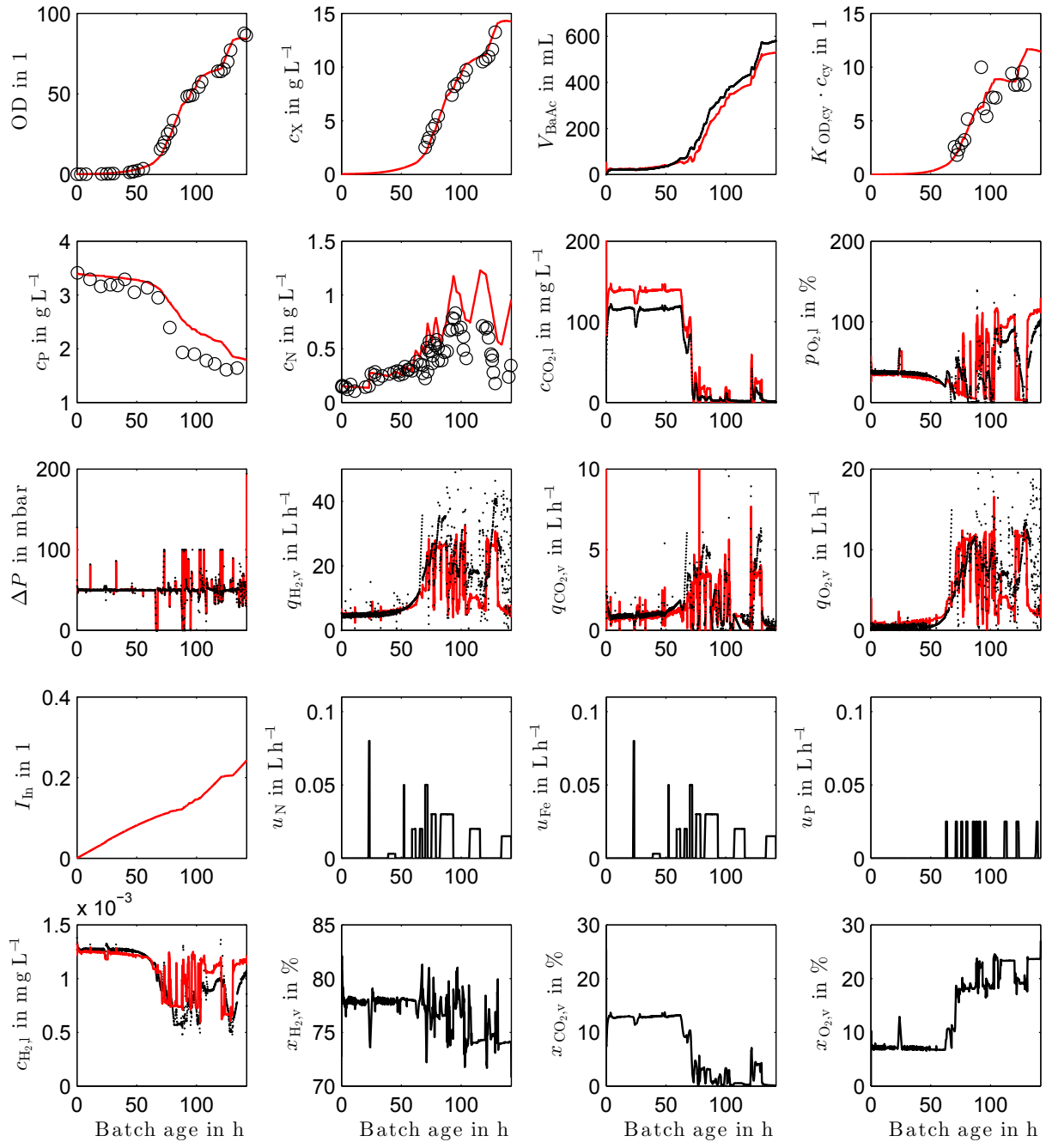


Figure 7.7: Cultivation E of the strain HF951. Measurements are given in black and the simulation of the adapted model in red. The synthetic measurements of $c_{H_2,l}$ were calculated on a data-driven basis. The figures with exclusively black lines show relevant model inputs.

7.3 Adapted model of the strain H798

Process model (I) was also adapted for a genetically modified strain H798 expressing SH simultaneously with the protein Frex that couples to NADH. This strain was selected, because in online OED for H16 the measurement situation was held responsible for poor experimental results. As in H798 the SH expression is coupled to the production of the fluorescing protein Frex, it was aimed to use fluorescence to measure the activity of SH indirectly by means of spectroscopy and employ H798 to test online OED experimentally. First cultivations indicated, though, that there is no proportional relation between measured fluorescence and SH activity. Hence, the strain could not be used to test the online OED, but the results of the adaption of process model (I) are presented here nevertheless. Two data sets H798A and H798B, displayed in Figures 7.8–7.9, served for model adaption. In cultivation H798B, the gas phase controller that was introduced in Section 4.4 controlled hydrogen, oxygen and pressure in the headspace of the reactor. For adaption, the growth parameters of the process model (I) were re-estimated. However, since there were no phases with high oxygen during cultivation, the parameters of the inhibition state x_{In} were taken from process model (I) listed in Table 5.8 on page 93. During most of the cultivation, an increased formation of PHB was counteracted by sufficient supply of nutrients and PHB was not measured post-experimentally. Parameters of the PHB pathway were therefore taken from process model (I) as well. For the remaining parameters, reasonable values had been identified and are listed in Table 7.4. However, to rely on the model for process control,

Table 7.4: Estimated parameter values of model H798 (derived from process model (I)) and lower and upper boundaries (lb, ub) applied as constraints for the parameter estimation

Parameter name	Unit	Parameter value	lb	ub
$U_{\mathrm{N,X}}$	$\mathrm{g\,g_X^{-1}}$	0.23	0.08	0.3
$U_{\mathrm{P,X}}$	$\mathrm{g\,g_X^{-1}}$	0.11	0.004	0.3
$U_{\mathrm{H_2,X}}$	$\mathrm{g\,g_X^{-1}}$	0.21	0.20	0.71
$U_{\mathrm{CO_2,X}}$	$\mathrm{g\,g_X^{-1}}$	1.3	0.4	4
$U_{\mathrm{O_2,X}}$	$\mathrm{g\,g_X^{-1}}$	1.1	0.32	3.8
$\mu_{\mathrm{X,max}}$	$\mathrm{h^{-1}}$	0.36	0.1	0.7
$k^{\mathrm{X}}_{\mathrm{H_2}}$	$\mathrm{mg\,L^{-1}}$	0.22	$2 \cdot 10^{-4}$	0.61
$k^{\mathrm{X}}_{\mathrm{CO_2}}$	$\mathrm{mg\,L^{-1}}$	65	4.4	176
$k^{\mathrm{X}}_{1,\mathrm{O_2}}$	$\mathrm{mg\,L^{-1}}$	1.6	$3.4 \cdot 10^{-4}$	2.6
$k^{\mathrm{X}}_{2,\mathrm{O_2}}$	-	3	0.8	9
$k^{\mathrm{X}}_{\mathrm{N}}$	$\mathrm{g\,L^{-1}}$	0.1	0.01	0.8
$k^{\mathrm{X}}_{\mathrm{P}}$	$\mathrm{g\,L^{-1}}$	0.45	0.4	1.5

further cultivations are required and the model should be cross-validated.

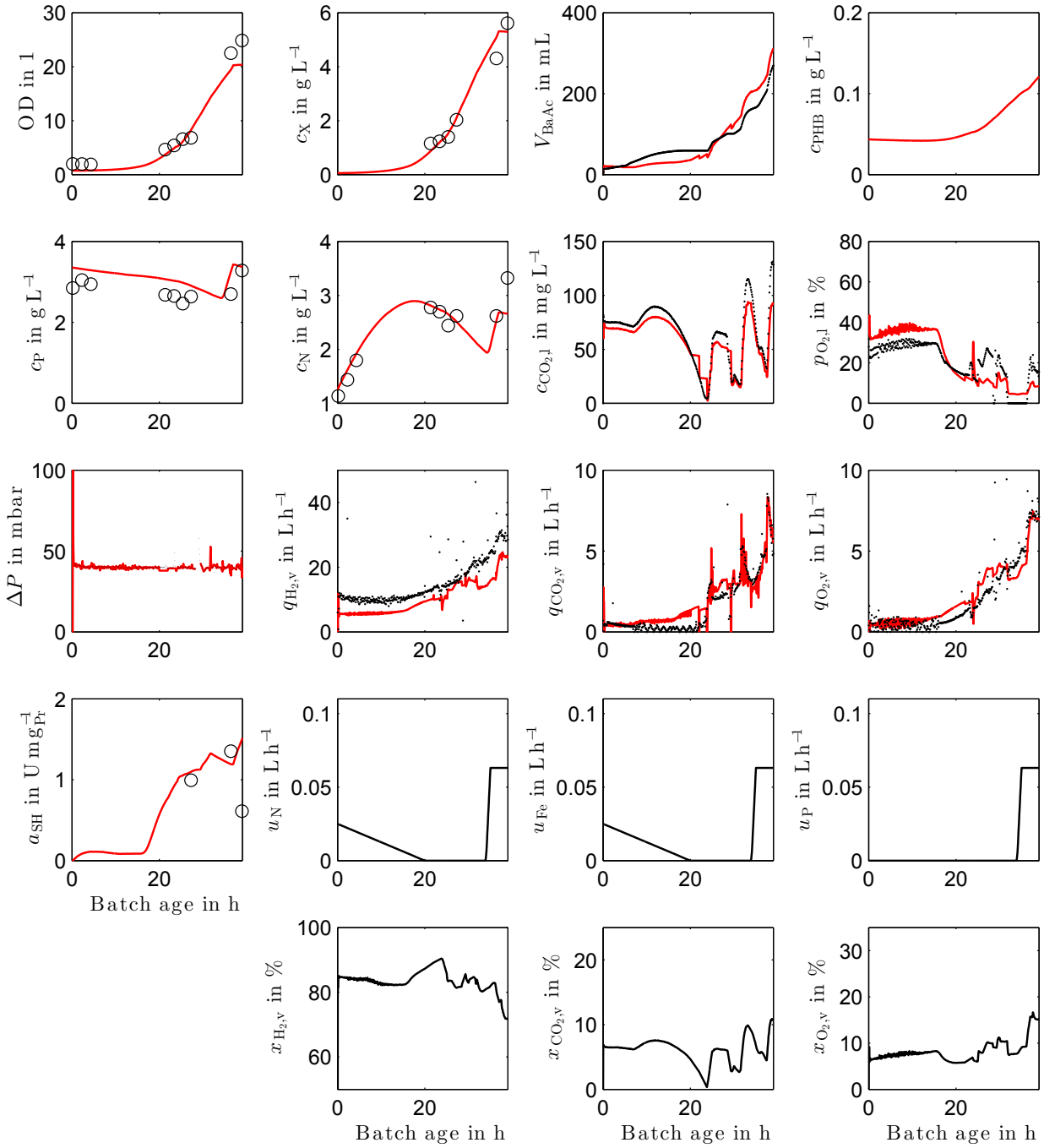


Figure 7.8: Cultivation A of the strain H798. Measurements and model inputs are given in black and the simulation of the adapted model in red. The figures with exclusively black lines show relevant model inputs.

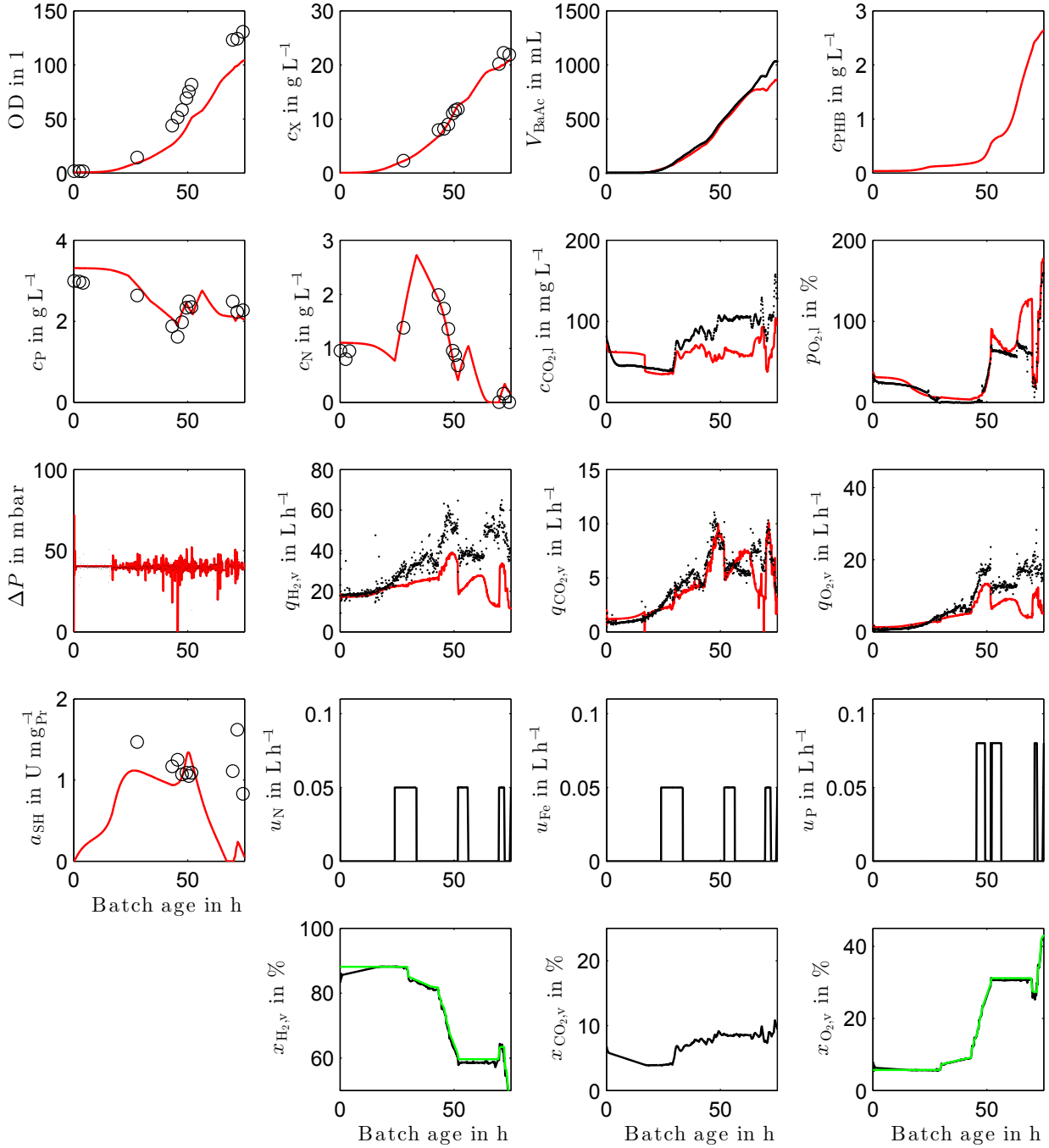


Figure 7.9: Cultivation B of the strain H798. Measurements are given in black circles or dots and the simulation of the adapted model in red. Black lines in graphs of the two bottom rows represent relevant model inputs, of which $x_{H_2,v}$ and $x_{O_2,v}$ were adjusted by the gas controller. Reference values for control are given in green.

Chapter 8

Summary and conclusion

Bioprocesses are employed to produce biotechnological compounds in the red (e.g., pharmaceuticals), white (e.g., enzymes for washing detergents) and green (e.g., fertilizers) sectors. Compared to the respective classical chemical process, production with microorganisms is more selective and the conditions for cultivation are generally environmentally friendlier. When developing an industrial bioprocess, the engineering of the strain and the process (including process control) aims to improve the yield of target compounds. However, advanced controllers depend on the process model that changes once the strain is genetically modified. In order to start with the controller synthesis, although the final strain to be used is not yet known, strategies are required for model adaption. In this thesis, a model adaption framework for the autotrophic cultivation of *R. e.* was presented. The first part dealt with the setup of a gas controlled cultivation system that is suitable for autotrophic organisms. Here, a strain independent gas phase control was developed and its performance results were exemplary shown for cultivations with H16 and H798. It allows to realize trajectories of gas compositions in the reactor headspace given by the operator, while maintaining the desired excess pressure. In contrast to the control suggested by Rossner (2014), it can be applied to any autotrophic strain growing on gases without further knowledge about its metabolism. Running gas controlled cultivations is necessary to generate data suitable for later process model adaption.

In the second part of this thesis, an adaptable general process model (I) was presented followed by model-based and model-free methods for its adaption. Process model (I) considers the liquid and the gaseous phase in a cultivation. It is schematically divided into three subunits: gas transport via the liquid-gas interface, microbial growth and product formation. The dynamics of the latter do not have an effect on the growth and gas transport units. In contrast, dissolved gases, which are important states of the gas transport unit, influence growth and vice versa. To adapt the process model (I), growth parameters are re-estimated using cultivation data of the new strain. Then, if required, the model is extended

for a product description of the new strain. The methods presented in this thesis for the construction of the production unit are based on an exact description of the growth unit. Deficiencies in the gas transport unit can be compensated for by using process model (II), which is similar to model (I) but it does not consider the gas phase but depends on the dissolved gas concentrations directly.

Difficulties of model adaption, as pointed out in Chapters 6–7, are partly based on structural inaccuracies of the gas transport and growth units of the process model (I) as well as the insufficient measurement situation. With a more precise and more complete model, the online OED, for example, would not try to approach states in which the organisms might die. And if more substances were measured, the "dependency analysis of approximated rates" could be used to analyse further dependencies, for example those between product and cell compartments such as DNA or metabolic intermediates. If target substances as well as substrates were measured more frequently, further periods in the cultivations could be evaluated and interpreted with the "phenomena recognition" and the "dependency analysis of approximated rates".

Some model deficits do not become apparent with adaption, but are already recognizable by the cross-validation of H16, which is only partially consistent with the experimental data, as described in Section 5.2.7. Incorrect simulations may result from inaccurate information about the initial states. In particular, the amount of active biomass in the inoculum is not known. If a faulty initial biomass is used for simulation, the measured quantities do not correspond to the experimental data. However, incorrect simulations may also be caused by structural model deficiencies and wrong parameter values. All parameters for which large uncertainties were estimated, are either not identifiable from a structural point of view, or require more data for a subsequent parameter estimation. Determining uncertainties by employing the Fisher information matrix, the parameters of membrane-bound hydrogenase (MBH) and polyhydroxybutyrate (PHB) formation are classified as uncertain above all. To decrease their uncertainties, the model structure needs to be altered and eventually metabolic intermediates must be described in more detail. However, most of the approximated relative uncertainties were very small, so that the parameter values appear quite accurate. Since the measured values are not always met, though, the model structure does not seem to be appropriate. A more complex model structure would be complimented with additional measurement quantities. In conclusion, more experimental data would be required. Specifically, the parameters of growth-associated formation and degradation of MBH show large uncertainties and their estimated values are questionable. However, MBH is part of the production unit of process model (I) and structural deficiencies in product description are not important for model adaption.

Process model (II), which is a part of model (I), can be beneficial for model extension and for the estimation of kinetic growth parameters in case dissolved gases are approximated poorly by model (I). Before employing model (II), the parameters must be estimated with

the cultivation data of the mutant strain. Although process model (II) was developed to circumvent shortcomings of the gas transport unit, it shows other gas-related inaccuracies that were discussed in Section 5.6.
Adapting the growth unit of process model (II) for the investigated, genetically modified strains (HF805, HF951, H798) gives satisfactory results. Cross-validations and parameter uncertainty analysis would be beneficial to evaluate the qualities of the adapted models. To this end, more experimental data, i.e., cultivations of the respective strains are required. To obtain good simulation results, an adaption of the growth unit of model (I) for the mutant strains was done. For the same parameters different values were identified depending on the strain that justifies an adaption of the growth unit parameters.
Model adaption methods in terms of model extension, in order to describe the formation of product, only led to satisfactorily results for cyanophycin of HF951 employing "phenomena recognition" and "dependency analysis of approximated rates". However, an additional manual modeling step of the production unit, evaluating and integrating the results of both methods, was necessary.
For HF805, neither by the introduced "dependency analysis of approximated rates" nor by the "phenomena recognition" tool, a description for MBH was found. A drawback of both methods is that they rely on an educated guess of the operator. The operator decides what quantities are investigated in dependency analysis and fills the phenomena pool based on past experience. Occurring phenomena beyond this pool, though, will never be found by this method. Also in "dependency analysis of approximated rates", quantities will only be related and evaluated if the operator suspects a dependency. Moreover, the measurements for MBH are very elaborate and error-prone and are therefore subject to high noise levels, which makes modeling even more difficult. For multi-model online OED, the problems are similar. Besides the poor measurement situation of SH, the initially proposed models, again, depend on the knowledge and experience of the operator. If none of the models is structurally able to approximate reality, also this method fails in experiments. The algorithms of multi-model online OED work well as tested in the simulation study, but they fall short practically because they optimized the gaseous inputs in such a way that the system was pushed beyond the scope of its model as discussed in Section 6.3.3. In multi-model online OED cultivations, metabolic effects occurred that were not included in the process model. Consequently, the model-based methods of online OED no longer worked reliably.
The adaption methods presented in this thesis were trialed on the cultivation data of three genetically modified strains. Hence, no generalized statement about the success of the introduced model adaption framework can be made. Whether model adaption (including growth and production unit) works, depends to a large extent on the metabolic effects of the genetic modifications carried out. It is conceivable that genetic alterations have such a serious influence that the model structure must be fundamentally altered.
Generally, the reliability and success of the introduced adaption methods depend on the

quality as well as quantity of measurements, the accuracy of the initial process model and the knowledge of the operator. It is conceivable that the latter two limitations can be overcome by hybrid modeling. Mechanistic models with reaction rates described by neural networks might lead to more accurate system approximations. Allowing neural networks as reaction rates in model adaption methods, the operator's knowledge is partly replaced by computational power. A drawback is that neural networks require a lot of experimental data. Thus, more online analytical methods would need to be established when working with artificial intelligence.

Appendix A

Notes on the execution of multi-model online OED

Multi-model online Optimal Experimental Design (OED) was developed for autotrophic cultivations and it uses some routines of the Advanced Batch Control (ABC) software (see Herold et al. (2017)). In the following sections, the workflow for the application of online OED is outlined. To get an overview of the folder structure (gray boxes) and the data exchange between the main scripts (ellipses), these elements are shown in Figure A.1. Folders marked in light gray accommodate model data and results of routines that are exchanged respectively stored. Manual interaction by the user is given in a white rectangle. Solid lines represent connections of the tree-like folder structure and labeled broken arrows present data transfer.

The temporal sequence of the multi-model OED routines in online mode is explained below using the three SH models of Section 6.3. To increase the accuracy of the algorithm, the gas controller dynamics can be taken into account as described in the Appendix B if fast solvers are available for stiff differential equation systems.

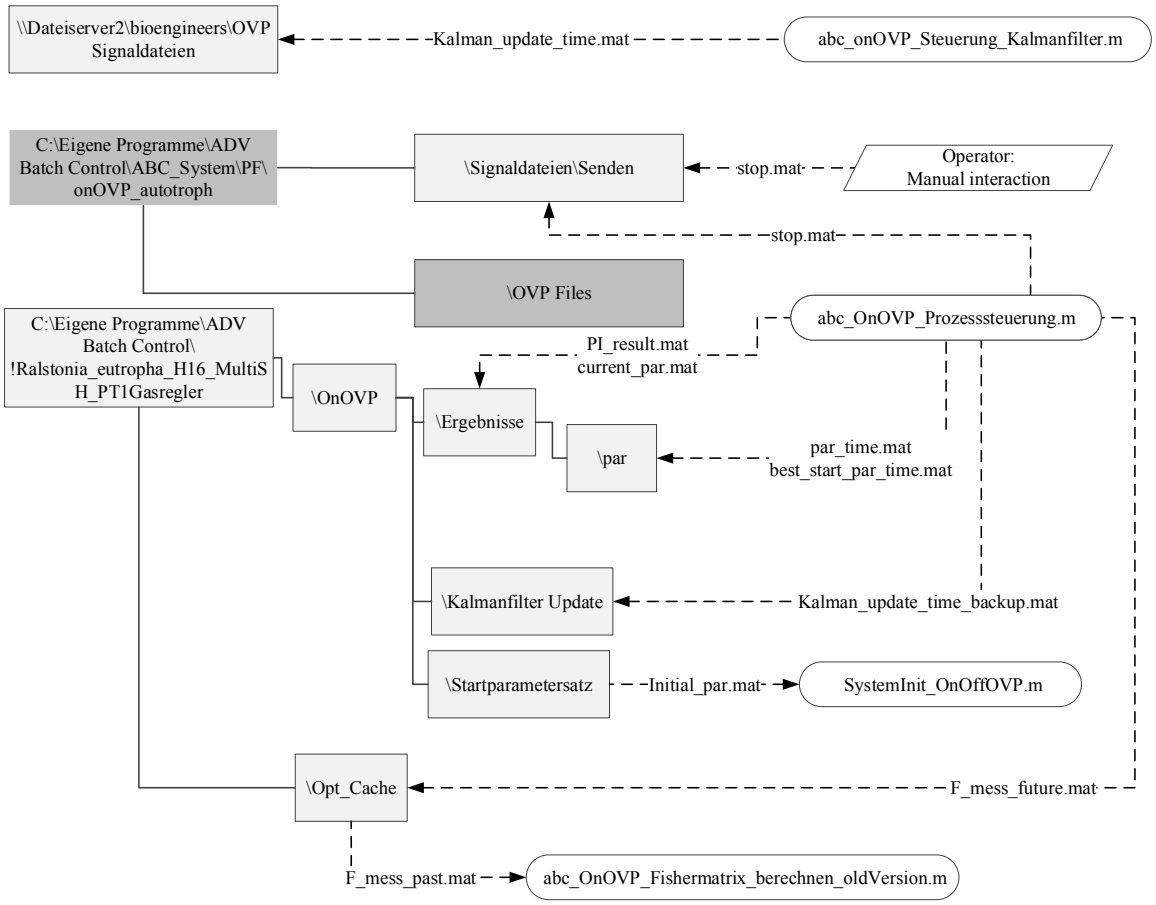


Figure A.1: Data exchange and folder structure of online OED. Folders in light gray accommodate model data and results of routines to be exchanged respectively stored. Manual interaction by the operator is given in a white rectangle. Solid lines represent connections of the tree-like folder structure and labeled broken arrows present data transfer.

A.1 Preparations to start of the experiment

All N models (here: N=3) that differ in target compound description are located in one model folder. If the gas controller dynamics are to be included, models (N+1:2N) consider the behavior of the underlying controller for OED and models (1:N) work without gas phase controller dynamics for state estimation and parameter identification. For the given example of Section 6.3, the gas controller was not included, and therefore the models N+1:2N were identical with models 1:N. But for the sake of completeness, in this section the gas controller dynamics are incorporated.

All manipulating variable constraints must be defined in the ABC script `SystemInit.m` of the corresponding model folder. In the file `SystemInit_OnOffOVP.m`, settings for linear constraints of manipulating variables, model number N, solver algorithm for the differential equation system, initial model for state estimation, experiment duration, start state vector,

maximal allowed optimization duration and sampling times have to be made.
As SH can only be measured reliably in the laboratory at high cell densities, the first period of the experiment is devoted to biomass production. Therefore, the optimal trajectories assigning growth during the first 24 hours are calculated offline and transferred to the database using standard ABC routines. Nevertheless, it is useful to take and analyze samples during the growth phase in order to obtain as many measured values as possible.

A.2 Preparations during growth phase

In this period, no online OED is running. Instead, the prior calculated optimal trajectories for fast growth are tracked by employing online trajectory planning that is an ABC standard routine. During growth phase, the states are estimated with models (N+1:2N) in parallel utilizing the standard estimation scripts in ABC so that later on, during multi-model online OED, a warm start for state estimation is possible for all N models that reduces computational cost. However, during online OED, only one state estimator may run. Hence, before online OED is started, all N estimators are stopped. Then, the model with which the state estimation is to be started (selected by the operator) is passed to `abc_onOVP_Steuerung_Kalmanfilter.m`, which provides a communication interface between multi-model online OED and ABC state estimation. In the same script, prior to this call, required folder paths must be set.
After 24 h, at the end of the growth phase and shortly before online OED is started, the file `abc_onOVP_Fishermatrix_berechnen_oldVersion` calculates and saves initial Fisher information matrices for all models (1:N). This script is usually executed before online OED. During online OED, it is called automatically, but if, for some reason, a sample was taken while online OED was not running, calling it manually updates and saves the Fisher information matrix for the missed sampling time.

A.3 Start of the online OED phase

Calling the script `onOVP_Prozessfuehrung.m` is the beginning of online OED. Automatically, the settings defined in `SystemInit_OnOffOVP.m` are loaded as well as the prior calculated Fisher information matrices. In this example, the script calculates optimal trajectories for the future gas fractions in the headspace. In contrast to state estimation, no model of 1:N has been selected directly after starting the routine, and thus the cost function, calculated according to eq. (6.10) on page 121, is averaged with respect to a factor defined in `SystemInit_OnOffOVP.m`. Since online OED requires a certain set value data format, which is no standard in ABC, reference trajectories of the gases defined in `SystemInit_OnOffOVP.m` satisfy these requirements for the inital optimization and serve for the first OED cycle, in-

stead of those stored in the database. For the following optimizations, reference values of the database are employed. Multi-model online OED runs automatically in a cycle using the estimated state vector as the system's start vector for optimizing the gas composition in the headspace. To this end, the inverted Fisher information matrices for future sampling times are summed in eq. (6.10) on page 121 and form the cost function that is minimized. While developing the scripts, we found out that the employed optimizer does not always return the trajectories with the lowest cost function. In order to support the work of the optimizer, the result of each iteration step (cost function value) is compared to the minimal value that is stored in a cache file. If the value is lower, it will be stored in the cache. As soon as no lower value can be found or the maximum duration is reached, the calculated trajectories of reference gas fractions and liquid feedrates for the entire further cultivation are stored in the database.

A.4 Manual interaction

When atline measurements, which are relevant for the new model branch, are entered into the database, a parameter estimation and subsequent model selection is possible. The operator interacts with `onOVP_Prozessfuehrung.m` by the signaling file `stop.mat` that is located in a regularly checked folder "Senden". As soon as the file contains the value "true", the online OED cycle is interrupted, the previous cultivation data of the database are downloaded by ABC routines and converted into an ABC-compatible data structure for parameter estimation. The employed ABC scripts work with prompts, and hence the operator has to be present.

A.5 Parameter estimation and model selection

As soon as the downloaded and processed measurement data structure is present in the workspace, the `onOVP_Prozessfuehrung.m` continues with a parameter estimation for each model (N+1:2N). In order to start the parameter identification with the best initial values, all past estimated parameter vectors are used to evaluate and compare the cost function in eq. (3.4) on page 26 of the new measurements data structure. For each model, the parameter vector with the lowest cost function value serves as initial vector for a subsequent parameter estimation. After the parameters of all models have been estimated, the one with the lowest cost function value is selected for the following online OEDs. Identified parameters and the number of the selected model are passed to the shared folder "Ergebnisse" to communicate with the state estimation routine by updating the utilized model. Before resuming online OED, the initial Fisher information matrix has to be calculated based on the newly se-

lected model with new parameters by `abc_onOVP_Fishermatrix_berechnen_oldVersion` considering all past sampling points.

Appendix B

Gas phase controller dynamics

The process model (I) from Chapter 5 does not include the controller dynamics of the gas controller presented in Section 4.4. In this section it is explained how the gas controller dynamics were integrated and why the resulting model was not used in closed-loop process control, e.g., online OED, although this would have been more accurate. For integration, the dynamics were modeled and delay parameter estimated employing data of previous, gas phase controlled cultivations.

The gas phase controller (see Section 4.4) operates correctly with ramp-like reference trajectories of $x_{\mathrm{gas,v}}$, but the cultivation control software (MFCS) requires step-wise reference changes for the liquid feed flows. A solution for closed-loop control (see Section 4.3) was sought that would allow simultaneous optimization of the liquid and gaseous inputs. It was decided to run optimizations with ramp-like input changes every 10–12 h for all inputs and to later discretize the resulting liquid reference feed flow ramps, which are marked in red in Figure B.1, in steps of 0.1 h. In this way, the ramp-like, optimized liquid input trajectories were approximated as steps that can be realized by MFCS.

The time required by the gas phase controller to achieve a reference point for the gas composition depends strongly on the total gaseous volume flow conveyed. This in turn is calculated on the basis of the consumption of the organisms. Accordingly, there are discrepancies between the reference values $r_{\mathrm{gas,v}}$ and the actual values $x_{\mathrm{gas,v}}$ that disappear at different rates depending on the total volume flow $q_{\mathrm{t,v}}$ being pumped. These discrepancies do not matter post-experimentally, i.e., when an experiment is evaluated, since the measured values of the headspace components $x_{\mathrm{gas,v}}$ are entries of the process model input vector in eq. (5.41) on page 72. In contrast, the input vector for planning and optimizations includes $r_{\mathrm{gas,v}}$ as reference values as described in Section 4.3. For an accurate planning, the process model was extended by a description of the delay caused by gas phase control that was assumed to be of first order:

$$\dot{x}_{\mathrm{gas,v}} = \frac{1}{T_{1,\mathrm{gas}}} \cdot (r_{\mathrm{gas,v}} - x_{\mathrm{gas,v}}). \tag{B.1}$$

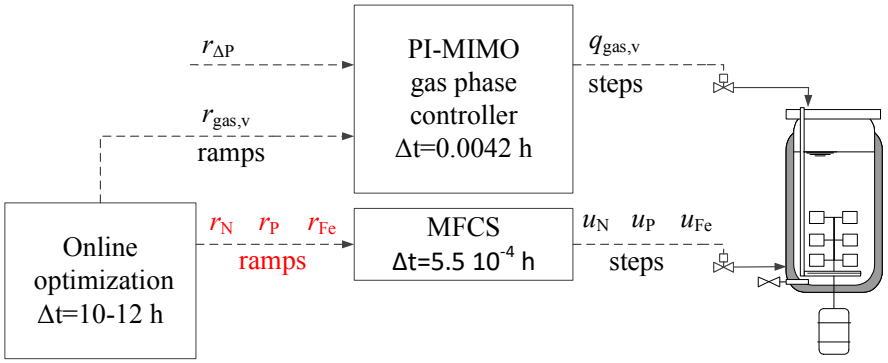


Figure B.1: Signal adjustment of reference and manipulating variables for communication among the different process units during closed-loop process control

For the controlled gases H_2 and O_2, the differential eq. (B.1) was formulated and $\dot{x}_{gas}$ was added to the state vector of process model (I) in eq. (5.80) on page 83. Carbon dioxide was not considered because for the evaluated experiments, H_2 and O_2 were the controlled gases. The delay factor $T_{1,gas}$ for each gas component depends on the reciprocal transferred gas flow $q^{-1}_{gas,trans}$. Assuming a pseudo-steady state, it is proportional to the reciprocal gas consumption rate ν^{-1}_{gas}, see eq. (5.108) on page 105, and hence changes during the cultivation:

$$T_{1,gas} = \frac{F_{gas}}{q_{gas}} = \frac{F_{gas}}{q_{gas,leak} + q_{gas,trans}}. \tag{B.2}$$

According to eq. (B.2), the higher the gas consumption rate, the shorter the delay time. The parameter F_{gas} (see Table B.1) were identified utilizing data of a cultivation using a previous version of the gas phase controller described in Section 4.4. As $T_{1,gas}$ sometimes adopts very small values in the course of the cultivation and the differential equations (B.1) are solved together with $\dot{x}_{model\ I}$, the combined system becomes very stiff. Orthogonal collocation or established solvers for ode systems like `ode45` or `ode15s` fail, although the latter is specialized on stiff systems. This problem was not addressed further as the focus of this work was different. For the same reason, no parameter identifications were made with the data set of a cultivation using the final gas controller of Section 4.4. If the dynamics of

Table B.1: Estimated values of the delay parameters caused by gas phase control

Parameter	Unit	Parameter value	lb	ub	rel. std. dev. (%)
F_{H_2}	L	0.052	0.01	50	7
F_{O_2}	L	0.013	10^{-4}	0.5	15

the gas phase controller are to be included in the future, the ABC software must be given an interface to another platform that can solve stiff systems even faster.

Appendix C

Autovalidation of the general process model (I)

All *R. e.* H16 cultivations used for parameter estimation (see Table 5.4) and carried out as part of this thesis are presented here. Cultivations REact11–REatc15 were performed by Rossner (2014). Since not all parameters were estimated together, some results displayed below can be interpreted as cross-validation for special parameter groups. For instance REatc19 (see Figure C.4) was only used to identify the parameters of PHB metabolism that belongs to the growth unit. Hence, measurements of the product MBH can be compared to the simulation in terms of cross-validation. Vice versa, for REatc25a and REatc25b the pictured biomass and PHB can be interpreted as cross-validation because only parameters of the production unit were estimated using these cultivations.

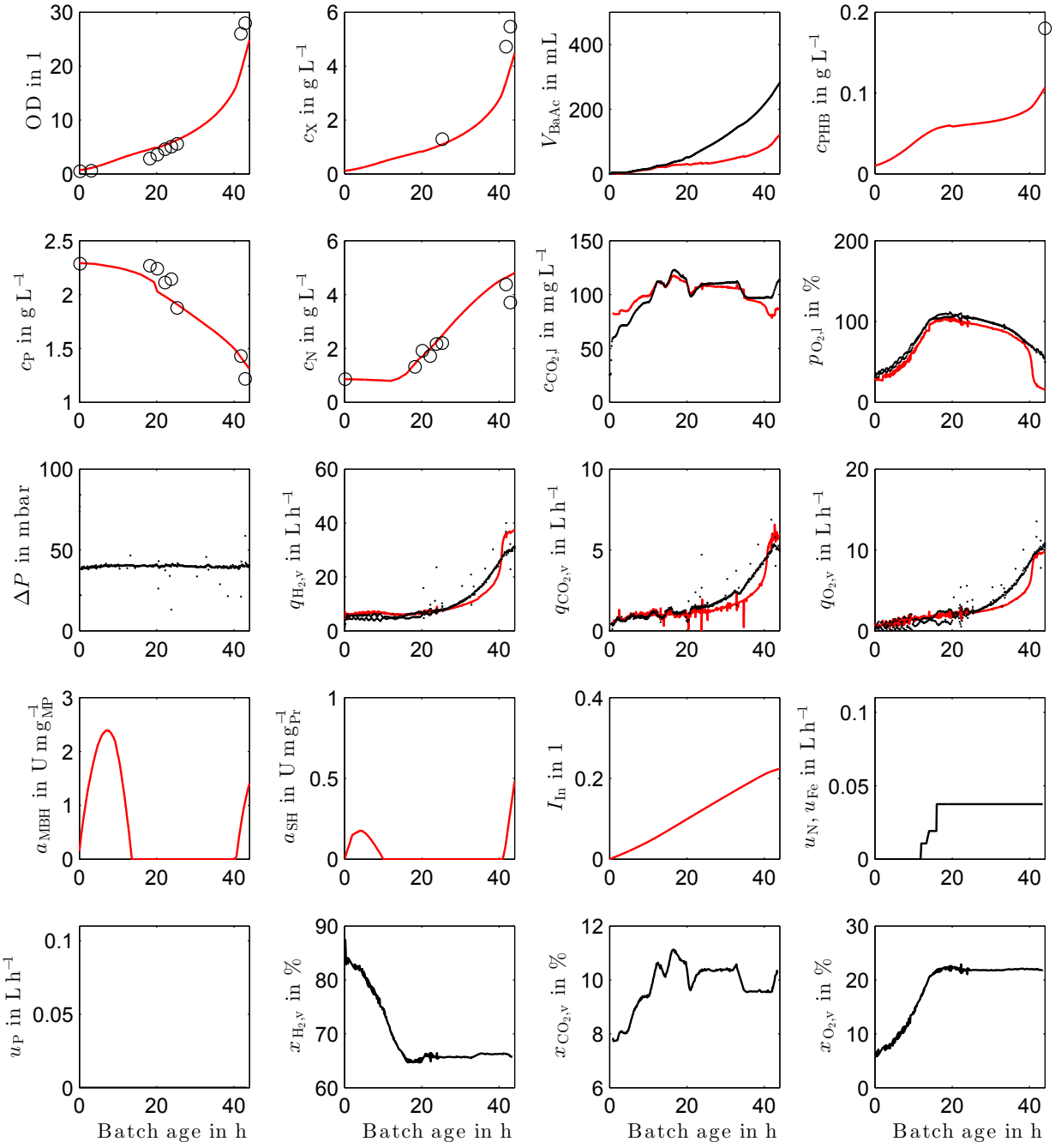


Figure C.1: Autovalidation for model (I) of cultivation REatc16. Measurement data are given in black circles or dots and simulations in red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs.

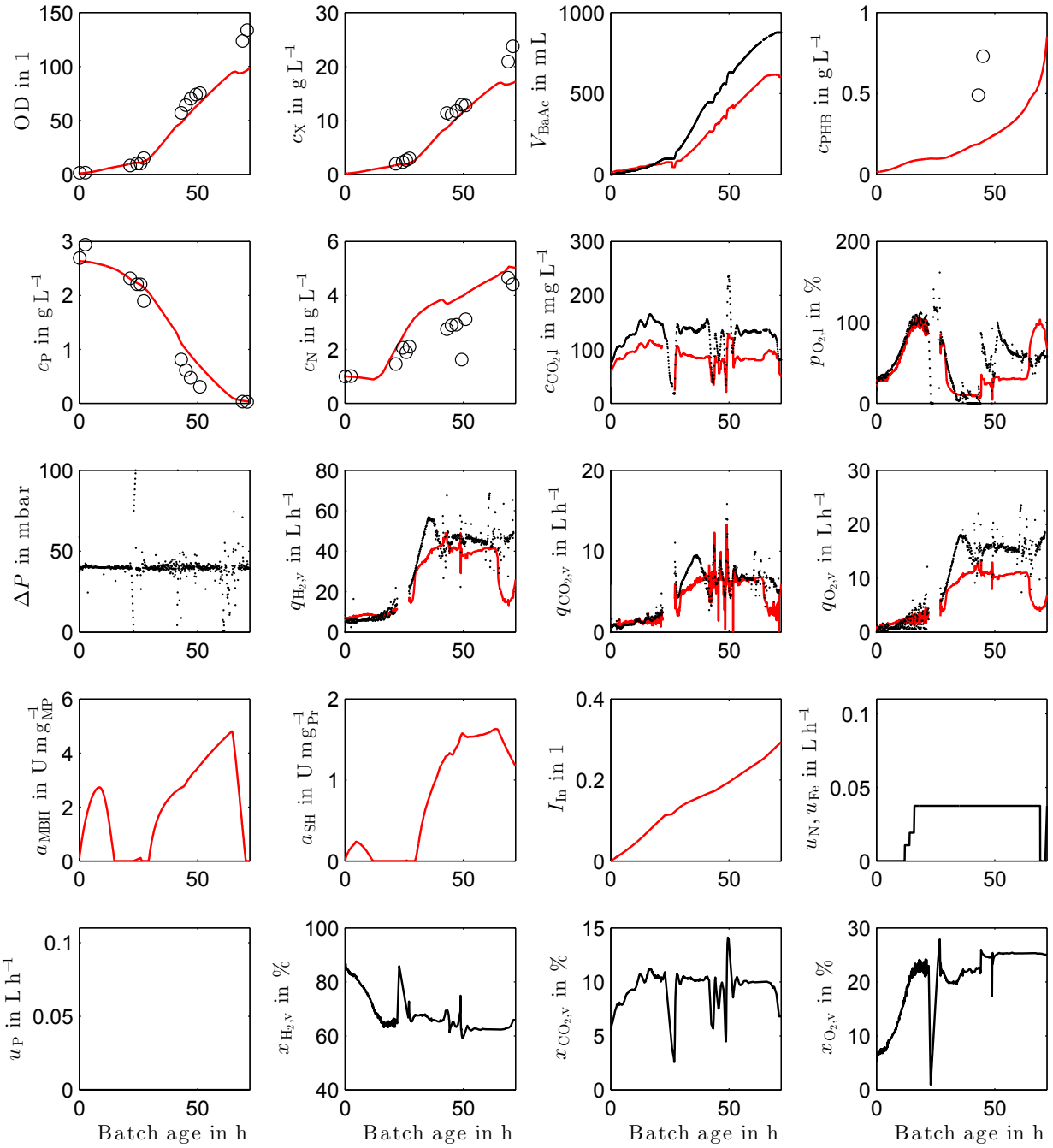


Figure C.2: Autovalidation for model (I) of cultivation REatc17. Measurement data are given in black circles or dots and simulations in red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs.

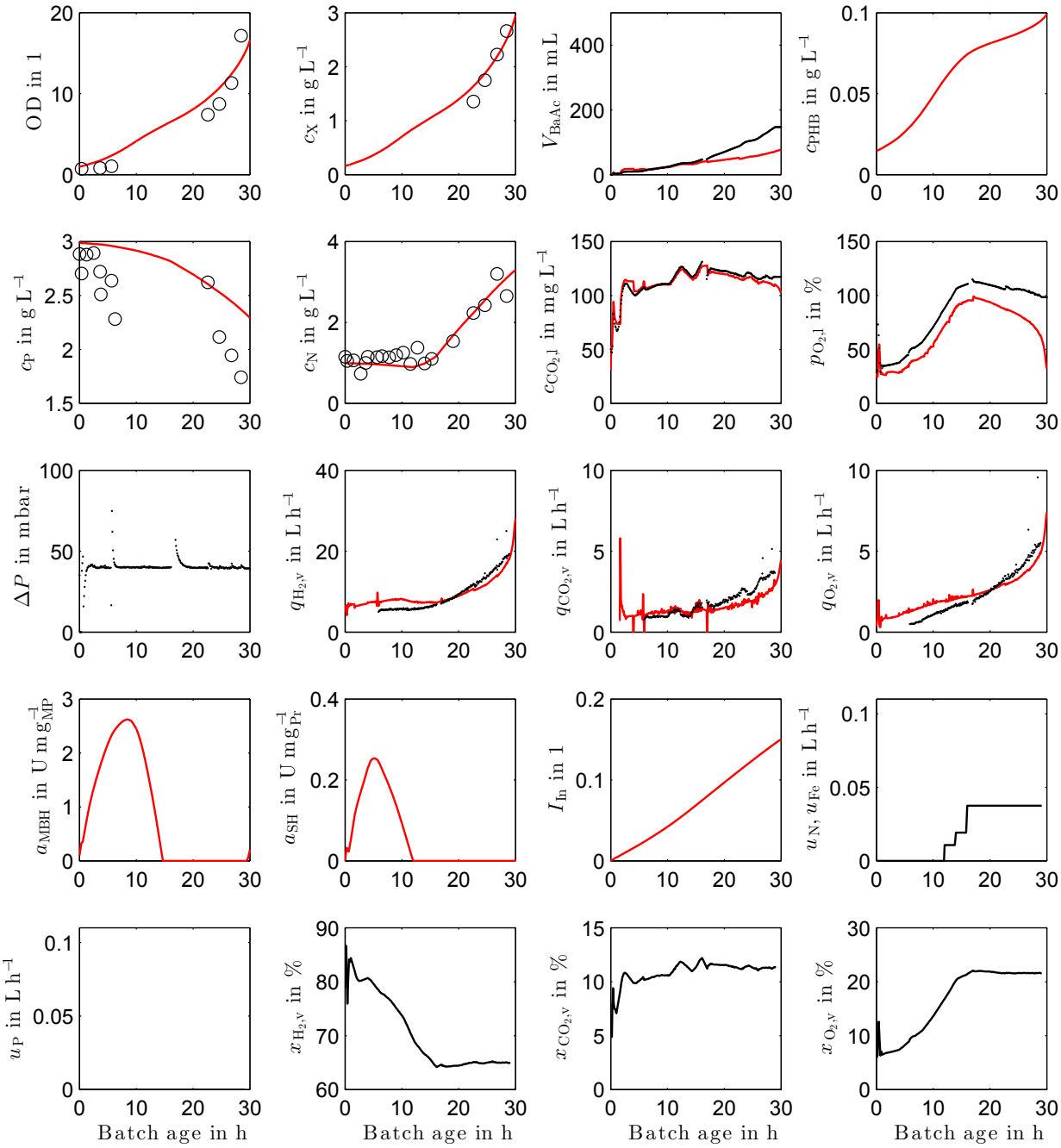


Figure C.3: Autovalidation for model (I) of cultivation REatc18. Measurement data are given in black circles or dots and simulations in red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs.

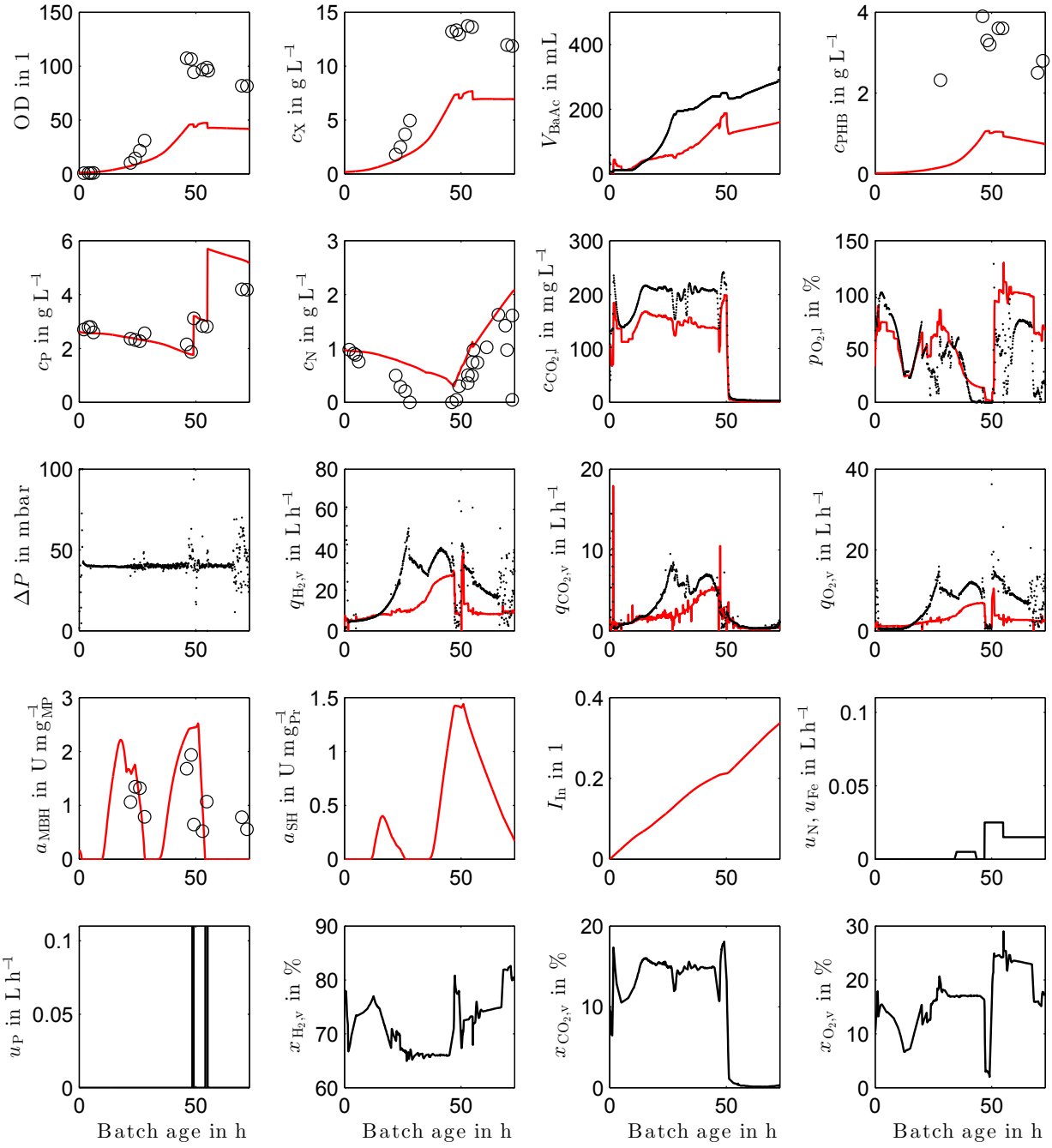


Figure C.4: Autovalidation for model (I) of cultivation REatc19. Measurement data are given in black circles or dots and simulations in red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs.

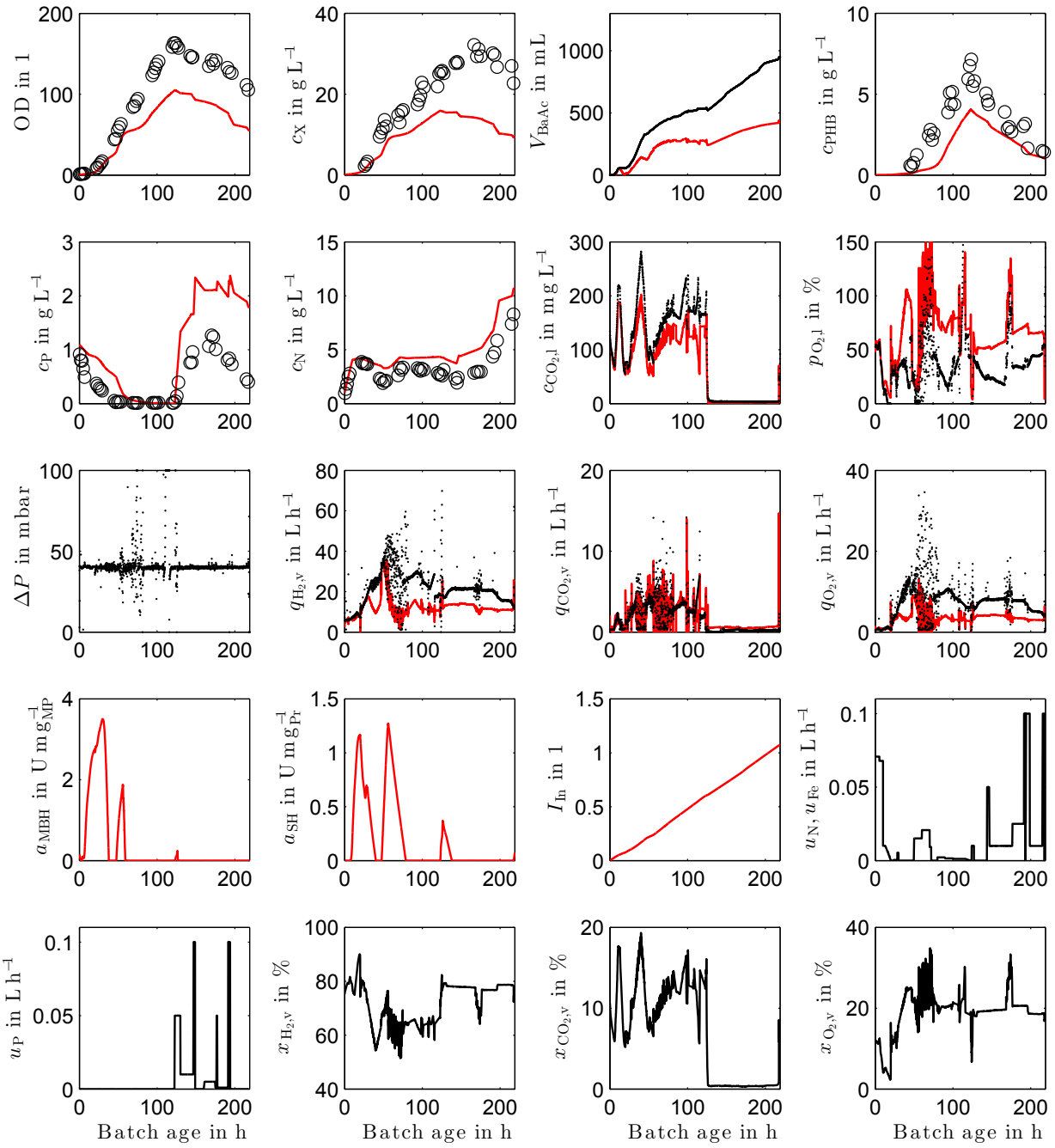


Figure C.5: Autovalidation for model (I) of cultivation REatc21. Measurement data are given in black circles or dots and simulations in red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs.

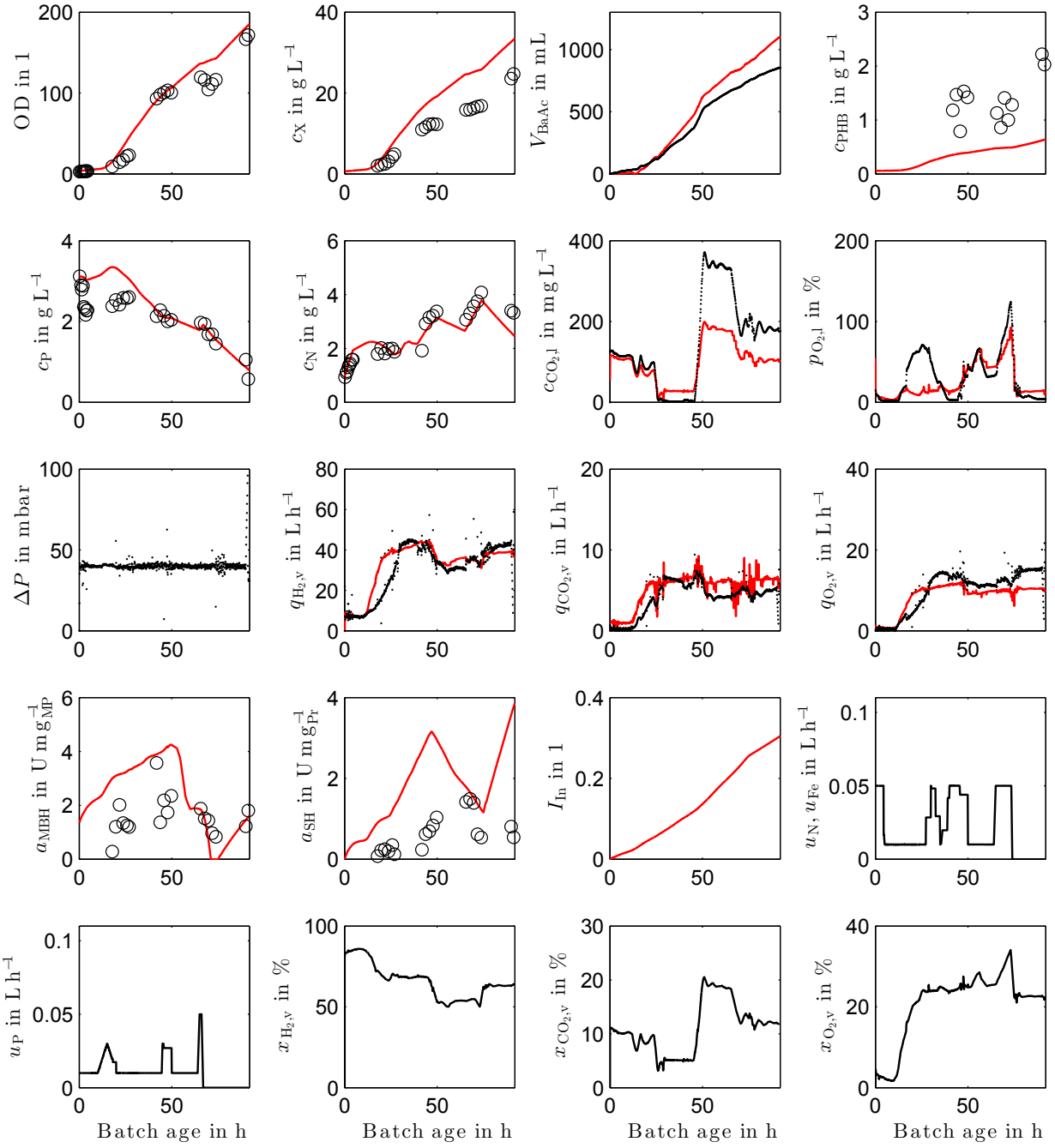


Figure C.6: Autovalidation for model (I) of cultivation REatc25a. Measurement data are given in black circles or dots and simulations in red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs.

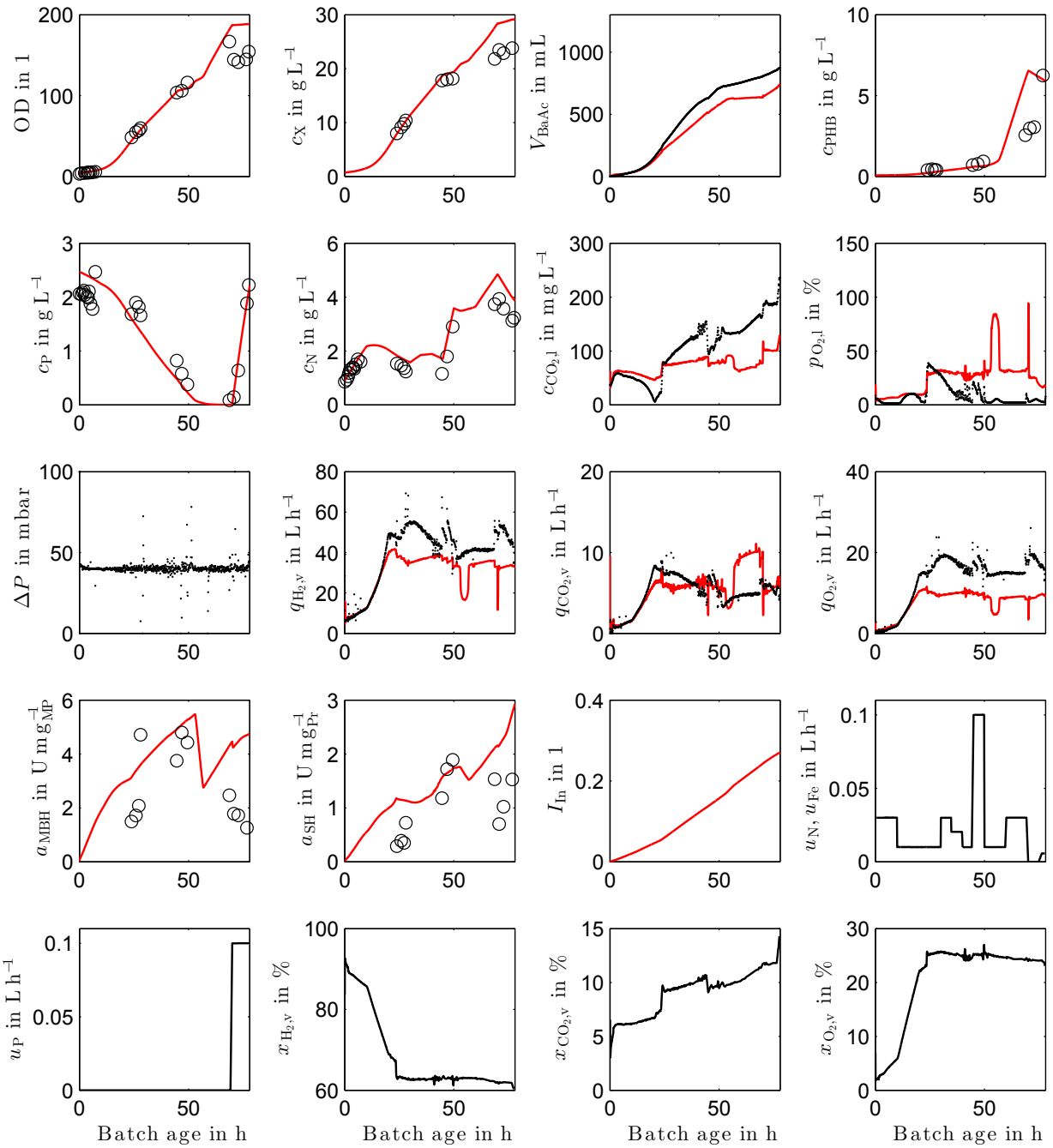


Figure C.7: Autovalidation for model (I) of cultivation REatc25b. Measurement data are given in black circles or dots and simulations in red. Black solid lines in the graphs of the two bottom rows represent relevant model inputs.

Bibliography

Arjunwadkar, S. J., Sarvanan, K., Kulkarni, P. R., and Pandit, A. B. 1998. Gas-liquid mass transfer in dual impeller bioreactor. *Biochem. Eng. J.* 1:99–106.

Baltes, M., Schneider, R., Sturm, C., and Reuss, M. 1994. Optimal experimental design for parameter estimation in unstructured growth models. *Biotechnol. Progr.* 10:480–488.

Bartha, R. 1962. Physiologische Untersuchungen über den chemolithotrophen Stoffwechsel neu isolierter Hydrogenomonas-Stämme. *Arch. Mikrobiol.* 41:313–350.

Beeby, M., Cho, M., Stubbe, J., and Jensen, G. J. 2012. Growth and Localization of Polyhydroxybutyrate Granules in Ralstonia eutropha. *J. Bacteriol.* 194:1092–1099.

Bongers, L. 1970. Energy generation and utilization in hydrogen bacteria. *J. Bacteriol.* 104:145–151.

Bowien, B. and Schlegel, H. G. 1981. Physiology and biochemistry of aerobic hydrogen-oxidizing bacteria. *Annu. Rev. Microbiol.* 35:405–452.

Bresan, S., Sznajder, A., Hauf, W., Forchhammer, K., Pfeiffer, D., and Jendrossek, D. 2016. Polyhydroxyalkanoate (PHA) Granules Have no Phospholipids. *Sci Rep* 6:1–13.

Calik, P., Ileri, N., Erdinç, B. I., Aydogan, N., and Argun, M. 2005. Novel antifoam for fermentation processes: Fluorocarbon-hydrocarbon hybrid unsymmetrical bolaform surfactant. *Langmuir* 21:8613–8619.

Cracknell, J. A., Wait, A. F., Lenz, O., Friedrich, B., and Armstrong, F. A. 2009. A kinetic and thermodynamic understanding of O2 tolerance in [NiFe]-hydrogenases. *Proc. Natl. Acad. Sci. U.S.A.* 106:20681–20686.

Crépin, L., Lombard, E., and Guillouet, S. E. 2016. Metabolic engineering of Cupriavidus necator for heterotrophic and autotrophic alka(e)ne production. *Metab. Eng.* 37:92–101.

Doi, Y., Kawaguchi, Y., Nakamura, Y., and Kunioka, M. 1989. Nuclear Magnetic Resonance Studies of Poly(3-Hydroxybutyrate) and Polyphosphate Metabolism in Alcaligenes eutrophus. *Appl. Environ. Microbiol.* 55:2932–2938.

Dubois, M., Gilles, K., Hamilton, J., Rebers, P., and Smith, F. 1956. Colorimetric method for determination of sugars and related substances. *Anal. Chem.* 28:350–356.

Franceschini, G. and Macchietto, S. 2008. Model-based design of experiments for parameter precision: State of the art. *Chem. Eng. Sci.* 63:4846–4872.

Fritsch, J. 2011. Molekulare Grundlagen der Sauerstofftoleranz der membrangebundenen [NiFe]-Hydogenase aus Ralstonia eutropha H16. Dissertation, Humboldt-Universität, Berlin.

Garcia-Ochoa, F. and Gomez, E. 2009. Bioreactor scale-up and oxygen transfer rate in microbial processes: an overview. *Biotechnol. Adv.* 27:153–176.

Gebauer, B. 2009. Das Abbauverhalten des PHB-Granulumkomplexes aus Ralstonia eutropha: Charakterisierung mit Hilfe neu entwickelter Methoden. Dissertation, Universität, Stuttgart.

Gelb, A., Kasper, J. F., Nash, R. A., Price, C. F., and Sutherland, A. A. 1974. Applied Optimal Estimation. MIT Press, Cambridge, MA.

Goldet, G., Wait, A. F., Cracknell, J. A., Vincent, K. A., Ludwig, M., Lenz, O., Friedrich, B., and Armstrong, F. A. 2008. Hydrogen production under aerobic conditions by membrane-bound hydrogenases from Ralstonia species. *J. Am. Chem. Soc.* 130:11106–11113.

Goodwin, G. C. and Payne, L. 1977. Dynamic System Identification. Mathematics in Science and Engineering. Elsevier.

Goris, T., Wait, A. F., Saggu, M., Fritsch, J., Heidary, N., Stein, M., Zebger, I., Lendzian, F., Armstrong, F., Friedrich, B., and Lenz, O. 2011. A unique iron-sulfur cluster is crucial for oxygen tolerance of a NiFe-hydrogenase. *Nat. Chem. Biol.* 7:310–318.

Heine, T. 2004. Modellgestützte Überwachung und Führung von Fed-Batch-Prozessen zur Antibiotikaproduktion. Dissertation, Technische Universität, Berlin.

Henzler, H.-J. 1982. Verfahrenstechnische Auslegungsunterlagen für Rührbehälter als Fermenter. *Chem. Ing. Tech.* 54:461–476.

Herold, S. and King, R. 2014. Automatic identification of structured process models based on biological phenomena detected in (fed-)batch experiments. *Bioprocess Biosyst. Eng.* 37:1289–1304.

Herold, S., Krämer, D., Violet, N., and King, R. 2017. Rapid process synthesis supported by a unified modular software framework. *Eng. Life Sci.* 17:1202–1214.

Humphreys, C. M. and Minton, N. P. 2018. Advances in metabolic engineering in the microbial production of fuels and chemicals from C1 gas. *Curr. Opin. Biotech.* 50:174–181.

Ishizaki, A. and Tanaka, K. 1990. Batch culture of Alcaligenes eutrophus 5ATCC6 17697T using recycled gas closed circuit culture system. *J. Ferment. Bioeng.* 69:170–174.

Jacquel, N., Lo, C.-W., Wei, Y.-H., Wu, H.-S., and Wang, S. S. 2008. Isolation and purification of bacterial poly(3-hydroxyalkanoates). *Biochem. Eng. J.* 39:15–27.

Kraume, M. 2012. Transportvorgänge in der Verfahrenstechnik. Springer Berlin Heidelberg, Berlin, Heidelberg.

Krieg, T., Sydow, A., Faust, S., Huth, I., and Holtmann, D. 2018. CO2 to Terpenes: Autotrophic and Electroautotrophic α-Humulene Production with Cupriavidus necator. *Angew. Chem. Int. Edit.* 57:1879–1882.

Kristensen, N. R., Madsen, H., and Jørgensen, S. B. 2003. Developing phenomena models from experimental data, pp. 1091–1096. *In* Kraslawski, Turunen (eds.) 2003 – European Symposium on Computer Aided, volume 14.

Lauterbach, L. 2013. Die Analyse der Sauerstofftoleranz und biotechnologische Anwendung der NAD+-reduzierenden Hydrogenase aus Ralstonia eutropha H16. Dissertation, Humboldt-Universität, Berlin.

Lauterbach, L. and Lenz, O. 2013. Catalytic production of hydrogen peroxide and water by oxygen-tolerant NiFe-hydrogenase during H2 cycling in the presence of O2. *J. Am. Chem. Soc.* 135:17897–17905.

Lauterbach, L., Lenz, O., and Vincent, K. A. 2013. H_2-driven cofactor regeneration with $NAD(P)^+$-reducing hydrogenases. *FEBS J.* 280:3058–3068.

Lee, S. E., Li, Q. X., and Yu, J. 2009. Diverse protein regulations on PHA formation in Ralstonia eutropha on short chain organic acids. *Int. J. Biol. Sci.* 5:215–225.

Liew, F., Martin, M. E., Tappel, R. C., Heijstra, B. D., Mihalcea, C., and Köpke, M. 2016. Gas Fermentation-A Flexible Platform for Commercial Scale Production of Low-Carbon-Fuels and Chemicals from Waste and Renewable Feedstocks. *Front. Microbiol.* 7:694.

Ludwig, M., Cracknell, J. A., Vincent, K. A., Armstrong, F. A., and Lenz, O. 2009. Oxygen-tolerant H2 oxidation by membrane-bound [NiFe] hydrogenases of ralstonia species. Coping with low level H2 in air. *J. Biol. Chem.* 284:465–477.

Morão, A., Maia, C. I., Fonseca, M. M. R., Vasconcelos, J. M. T., and Alves, S. S. 1999. Effect of antifoam addition on gas-liquid mass transfer in stirred fermenters. *Bioprocess Eng.* 20:165.

Morinaga, Y., Yamanaka, S., Ishizaki, A., and Hirose, Y. 1978. Growth characteristics and cell composition of Alcaligenes eutrophus in chemostat cultures. *Agric. Biol. Chem.* 42:439–444.

Mozumder, M. S. I., Garcia-Gonzalez, L., de Wever, H., and Volcke, E. I. 2015. Poly(3-hydroxybutyrate) (PHB) production from CO2: Model development and process optimization. *Biochem. Eng. J.* 98:107–116.

Neddermeyer, F. and King, R. 2019. Advanced nonlinear multi-layer process control for autotrophic cultivations. *IFAC-PapersOnLine* 52:733–738.

Neddermeyer, F., Marhold, V., Menzel, C., Krämer, D., and King, R. 2016. Modelling the production of soluble hydrogenase in Ralstonia eutropha by on-line optimal experimental design. *IFAC-PapersOnLine* 49:627–632.

Park, J., Kim, T., and Lee, S. 2011. Genome-scale reconstruction and in silico analysis of the Ralstonia eutropha H16 for polyhydroxyalkanoate synthesis, lithoautotrophic growth, and 2-methyl citric acid production. *BMC Syst Biol* 5:101.

Pohlmann, A., Fricke, W. F., Reinecke, F., Kusian, B., Liesegang, H., Cramm, R., Eitinger, T., Ewering, C., Potter, M., Schwartz, E., Strittmatter, A., Voss, I., Gottschalk, G., Steinbuchel, A., Friedrich, B., and Bowien, B. 2006. Genome sequence of the bioplastic-producing "Knallgas" bacterium Ralstonia eutropha H16. *Nat. Biotechnol.* 24:1257–1262.

Porterfield, J. Z. and Zlotnick, A. 2010. A Simple and General Method for Determining the Protein and Nucleic Acid Content of Viruses by UV Absorbance. *Virology* 407:281–288.

Raje, P. and Srivastava, A. K. 1998. Updated mathematical model and fed-batch strategies for poly-β-Hydroxybutyrate (PHB) production by Alcaligenes eutrophus. *Bioresour. Technol.* 64:185–192.

Ratzka, J. 2011. Charakterisierung und Stabilisierung der löslichen Hydrogenase aus Ralstonia eutropha H16 für den technischen Einsatz bei der reduktiven Cofaktorregenerierung. Dissertation, Technische Universität, Berlin.

Rhine, E. D., Mulvaney, R. L., Pratt, E. J., and Sims, G. K. 1998. Improving the Berthelot Reaction for Determining Ammonium in Soil Extracts and Water. *Soil. Sci. Am. J.* pp. 473–480.

Ritchie, R. J. 2013. The ammonia transport, retention and futile cycling problem in cyanobacteria. *Microb. Ecol.* 65:180–196.

Rossner, N. 2014. Robuste modellbasierte Prozessführung auf Basis von Gaußschen Mischdichten am Beispiel der Bray-Liebhafsky-Reaktion und der autotrophen Kultivierung von Ralstonia eutropha H16. Dissertation, Technische Universität, Berlin.

Ryu, H. W., Hahn, S. K., Chang, Y. K., and Chang, H. N. 1997. Production of poly(3-hydroxybutyrate) by high cell density fed-batch culture of Alcaligenes eutrophus with phospate limitation. *Biotechnol. Bioeng.* 55:28–32.

Saggu, M., Zebger, I., Ludwig, M., Lenz, O., Friedrich, B., Hildebrandt, P., and Lendzian, F. 2009. Spectroscopic insights into the oxygen-tolerant membrane-associated NiFe hydrogenase of Ralstonia eutropha H16. *J. Biol. Chem.* 284:16264–16276.

Schill, N., van Gulik, W. M., Voisard, D., and von Stockar, U. 1996. Continuous cultures limited by a gaseous substrate: Development of a simple, unstructured mathematical model and experimental verification with Methanobacterium thermoautotrophicum. *Biotechnol. Bioeng.* 51:645–658.

Schill, N. A. 1996. Continuous cultures of Methanobacterium Thermoautotropicum: Thermodynamic analysis, growth physiology and modeling. PhD Thesis, Polytechnische Hochschule, Lausanne.

Schlegel, H. G., Gottschalk, G., and Bartha, R. 1961. Formation and utilization of poly-beta-hydroxybutyric acid by Knallgas bacteria (Hydrogenomonas). *Nature* 191:463–465.

Schlink, B. and Schlegel, H. G. 1979. The membrane-bound hydrogenase of Alcaligenes eutrophus: I. Solubilization, purification, and biochemical properties. *Biochim. Biophys. Acta* 567:315–324.

Schneider, R. and Georgakis, C. 2013. How To NOT Make the Extended Kalman Filter Fail. *Ind. Eng. Chem. Res.* 52:3354–3362.

Seletzky, J. M. 2009. Process development and scale-up from shake flask to fermenter of suspended and immobilized aerobic microorganisms. Dissertation, RWTH, Aachen.

Shang, L., Jiang, M., Ryu, C. H., Chang, H. N., Cho, S. H., and Lee, J. W. 2003. Inhibitory effect of carbon dioxide on the fed-batch culture of Ralstonia eutropha: Evaluation by CO2 pulse injection and autogenous CO2 methods. *Biotechnol. Bioeng.* 83:312–320.

Sizer, I. W. and Peacock, A. C. 1947. The ultraviolet absorption of serum albumin and of its constituent amino acids as a function of pH. *J. Biol. Chem.* 171:767–777.

Söderström, T. 1989. System identification. Prentice Hall international series in systems and control engineering. 1. publ edition.

Sultana, N. 2012. Production and characterization of tissue engineering scaffolds based on polyhydroxybutyrate-co-hydroxyvalerate polymers. *In* Biomedical Engineering (ICoBE), 2012 International Conference on, pp. 129–132.

Takeshita, T., Tanaka, K., Ishizaki, A., and Stanbury, P. F. 1993. Studies on Dissolved Hydrogen Behavior in Autototrophic Culture of Alcaligenes eutrophus ATCC 17697T. *J. Fac. Agric. Kyushu Univ.* 38:55–64.

Takors, R., Kopf, M., Mampel, J., Bluemke, W., Blombach, B., Eikmanns, B., Bengelsdorf, F. R., Weuster-Botz, D., and Dürre, P. 2018. Using gas mixtures of CO, CO2 and H2 as microbial substrates: The do's and don'ts of successful technology transfer from laboratory to production scale. *Microb. Biotechnol.* 11:606–625.

Tanaka, K., Ishizaki, A., Kanamaru, T., and Kawano, T. 1995. Production of poly(D-3-hydroxybutyrate) from CO2, H2, and O2 by high cell density autotrophic cultivation of Alcaligenes eutrophus. *Biotechnol. Bioeng.* 45:268–275.

Villaverde, A.F., Evans N.D., Chappell M.J., and Banga J.R. 2019. Input-dependent structural identifiability of nonlinear systems. *IEEE Contr. Syst. Lett.* 3:1–6.

Walter, É. and Pronzato, L. 1997. Identification of Parametric Models from Experimental Data. Communications and Control Engineering. Springer-Verlag, Berlin, 1 edition.

Wilde, E. 1962. Untersuchungen über Wachstum und Speicherstoffsynthese von Hydrogenomonas. *Arch. Mikrobiol.* 43:109–137.

Zhao, Y., Jin, J., Hu, Q., Zhou, H.-M., Yi, J., Yu, Z., Xu, L., Wang, X., Yang, Y., and Loscalzo, J. 2011. Genetically encoded fluorescent sensors for intracellular NADH detection. *Cell Metab.* 14:555–566.

Zosel, J., Oelssner, W., Decker, M., Gerlach, G., and Guth, U. 2011. The measurement of dissolved and gaseous carbon dioxide concentration. *Meas. Sci. Technol.* 22:072001.