BIO PROCESS TECHNOLOGY

CHASE

TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

TU WIEN

Master Thesis

Digital Twin based feedback control of biomass specific rates in a Fab-fragment producing E. coli Fed-batch process

ausgeführt zum Zwecke der Erlangung des akademischen Grades einer Diplomingenieurin der Verfahrenstechnik
unter der Leitung von

Univ.Prof. Dr. Christoph Herwig
E166 Institut für Verfahrenstechnik, Umwelttechnik und Technische Biowissenschaften
Forschungsgruppe: Bioprocess Technology
und Mitbetreuung von
Dipl.-Ing. Dr.techn. Julian Kager
Competence Center CHASE GmbH

eingereicht an der Technischen Universität Wien

Fakultät für Technische Chemie

von

Horst Nora-Sophie
Matr.Nr.: 1326784

Wien, am 29. September 2022

Horst Nora-Sophie

Ich erkläre eidesstattlich, dass ich die Arbeit selbständig angefertigt habe. Es wurden keine anderen als die angegebenen Hilfsmittel benutzt. Die aus fremden Quellen direkt oder indirekt übernommenen Formulierungen und Gedanken sind als solche kenntlich gemacht. Diese schriftliche Arbeit wurde noch an keiner Stelle vorgelegt.

Wien, am 29. September 2022 Horst Nora-Sophie

Acknowledgement

I would like to express my sincere gratitude to Prof. Dr. Christoph Herwig who enabled me to write my thesis. He has always encouraged me to focus on the novel aspects of my topic, which improved my work. Thanks to him I was able to gain deeper insight into bioprocesses and control which I consider valuable for my professional future.

A special thanks goes to Dr. Julian Kager for his invaluable patience and time that I needed to get on track with the subject. Without him it would not have been possible to dive so deeply into this topic and acquire knowledge on so many new aspects. Owing to his profound knowledge, he was able to give me constructive advise and motivated me a lot to carry on.

On a more personal level I would like to thank my partner and friends who were always there when I needed someone to let of steam. At last but not least I'm grateful to my father who supported me financially and passed his interests and talents on to me. I am sad that he will not be able to see me graduate.

This work was conducted within the COMET Centre CHASE, funded within the COMET Competence Centers for Excellent Technologies programme by the BMK, the BMDW and the Federal Provinces of Upper Austria and Vienna. The COMET programme is managed by the Austrian Research Promotion Agency (FFG).

Abstract

To ensure stable product quality and batch to batch reproducibility of fermentation processes, bioprocess control is the matter of choice. A majority of recombinant protein production processes are carried out in fed-batch mode, which enables control by the feed addition. Models are needed for the establishment and the investigation of more complex control objectives as a representation of the real process. These models could be extended to applicable Digital Twins if they are integrated in the monitoring and control loop of the fermentation plant.

The aim of the thesis is to find control laws to keep biomass specific rates of a recombinant protein producing *E.coli* fed-batch process constant and investigate the impacts and the usability of this new control approaches. Therefore an already existing process model was extended and fitted to historical datasets. Subsequently the model was used to carry out a nonlinear feedback linearization to obtain the demanded control laws. In addition to the well known and often used control of the specific growth rate control laws were derived to control the biomass specific substrate uptake rate, the recombinant protein production rate as well as the biomass specific oxygen uptake and carbon dioxide production rate.

To judge the applicability of these control laws an implementation study was carried out. Therefore an independent verification plant was simulated as a substitute for real experiments. To apply the control laws in a meaningful way a setpoint optimization was performed beforehand and the controllers were tuned.

The control laws could successfully be used to keep the biomass specific rates at constant setpoints. Additionally the implementation study showed an theoretical potential to improve process performance using the determined optimized setpoints. Eventually the potential to apply this novel control approaches and the model to real experiments is discussed with regard to an applicable digital twin.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Biotechnological production

Bioprocesses are widely used to produce chemicals, bulk enzymes, foods, and pharmaceuticals. Commonly used industrially strains are already highly developed and assure high productivity. Additional process optimization focuses on the optimal conditions during the fermentation. Besides physico-chemical parameters such as pH, temperature, dissolved oxygen and nutrient levels the course in time of the process, which can be manipulated via the feed addition is also investigated as it influences the performance of the process.[1] Bioprocesses can be categorized according to their mode of operation, a distinction is made between batch, fed-batch and continuous processes.[2] The simplest process mode is batch fermentation, where the nutrients for the entire process are provided from the beginning. The disadvantage of the batch mode is the fact that growth and product formation rates cannot be regulated and happen at uncontrolled, rates. Continuous operation is intended to maintain the system in a steady state, this is done by keeping identical mass flows in feed and out-flow rates. Even if very effective processes can be achieved with this mode of operation, they have the disadvantage on an industrial scale, in addition to the higher contamination risk, that they must be monitored extremely strictly and that changing cell metabolism and degeneration hampers their long term stability.[1] Therefore, Batch and fed-batch cultivations are used more frequently than continuous processes. A large number of processes are run in fed-batch mode, which will be described in more detail below.
The fed-batch process is characterized by the fact that the reactor volume changes over time as medium is continuously added to the process vessel. The method is used for processes in which growth conditions must be kept low in order to enable product formation.[2] In many cases, the fed-batch production process is preceded by a batch process to increase biomass accumulation in the first step.[1] The continuous feeding method has the advantage over batch operation that substrate inhibition or catabolite suppression of product synthesis can be prevented. [3] In addition, the fed-batch offers the possibility of process optimization by regulating the feed rate, as it directly influences the metabolic activity and the volume of the system. More specifically, the substrate concentration contained in the feed or the

applied feed rate has an influence on the growth rate, the specific product formation rate, and the oxygen uptake rate. In addition feeding has an impact on any other concentration in the system due to dilution effects. [1]

1.2 Recombinant protein production processes

Recombinant proteins make up a large proportion of new important pharmaceutical agents.[4] The name recombinant proteins is derived from the fact that the DNA of the host cell that encodes them has been recombined or manipulated, because the host in which they are synthesized often belongs to a different species than the one from which the proteins originate. [5]
Since the approval of human insulin as the first recombinant product, many other products have been approved for human use, including thrombolytics, hormones, growth factors, interferons and antibody fragments. [5] Just like but not limited to monoclonal antibodies, antibody fragments can also bind antigens. Since full length antibodies are glycosylated, they must be produced in mammalian cells. In contrast, antibody fragments can be produced in microorganisms, which are easier to cultivate.[6] Because of the application advantages such as low doubling times, lower contamination risk and cheap media, E.coli is popular for the production of recombinant proteins. [7]
The production of recombinant proteins requires strict safety requirements and operational restrictions, as their clinical efficacy is influenced by deviations in the production process. To ensure consistent product quality, batch-to-batch reproducibility is of particular importance.[4] Desired productivity and consistent product quality were traditionally achieved mainly through genetic engineering and media development. [8]

1.3 Digital twins in biotechnology

Currently, the focus to improve biotechnological manufacturing processes is increasingly on process system engineering (PSE). PSE mainly comprised creation of mathematical models and implementation of model-based methods for experimental designs, process optimization and real-time monitoring and control strategies.[8] While process control can improve reproducibility[4], automated control is not easy to implement due to the complexity of bioprocesses and requires holistic approaches for monitoring and control. This involves replacing manual and therefore discontinuous measurements and control strategies with a systematic process that takes the entire process into account. A digital twin is suitable here consisting of soft-sensors, real-time mass balances, state observers or model based or model predictive controllers.[9] A digital twin is a digital replica of the actual process that represents the physical state and contains all relevant information of the system. A fully developed digital twin consists of three components: a digital model of the system, soft sensors for real-time acquisition and integration of process data, and automated communi-

cation between the system and the twin via sufficient model-based control algorithms.[10]

Digital Twins as a tool for process control or for process optimization as well as for simulation are increasingly used in all kinds of industries.[11] In the pharmaceutical industry, the technical use of control programs lags a bit behind other fields due to the complexity of the biotechnological processes. Therefore, the technology of the digital twin and the possibilities for processing huge amounts of data, model based techniques are becoming more and more interesting in this field. [11] Digital twin technology is so applicable because they offer a way to represent the complexity of a biotechnological process well enough. In bioprocesses, attention must be paid to the physical, chemical and kinetic peculiarities of the system, as well as the mechanical effects of the process environment on the organisms.[12] A holistic process model such as a digital twin can provide a very precise description of the complex structure of biotechnological processes and thus also significantly improve process understanding and process control strategies.[13]

1.3.1 Process modeling

Models serve to describe real behavior as closely as possible. In particular, the aim is to describe important system properties in order to forecast the original system dynamics.[2] Typically, either data-driven or mechanistic models are used in bioprocess modelling. The advantage of data-driven models is that they are quick and easy to implement, as available regression algorithms such as artificial neural networks (ANN) do not require a detailed understanding of the process. The disadvantage is that the models often do not represent causal relationships, which can lead to error-prone conclusions. In contrast, mechanistic models are based on physical relationships, but do not include smaller process variations. Mechanisitc models are less accurate on the training data but their transferability and prediction ability is often better than purely data driven approaches.[14]
The generic process model used in this work is a physiological model. It is based on the metabolic kinetics and differential equations obtained by the unsegregated kinetic model with closing elemental balances. In this way, relationships between substance concentrations in the system, process parameters and the kinetic parameters can be described. Therefore, model-based investigations can replace experiments of the original process. This opens up many possibilities in research and development, as the less complex models can significantly improve the understanding of the process. In addition, costs and time factors are of course much lower than with real experiments.[2] For this reason, models are a popular tool for biotechnological process development. Due to the better understanding and predictability of molecular processes that can be obtained using process models, they are also used for optimization and control purposes.[15] Considering the many variables and non-linear kinetics of a bioprocess, modelling is particularly interesting in this field especially with regard to control. Adaptive and non-linear control strategies based on mathematical mod-

els offer good alternatives to conventional methods, which often reach their limits in this application.
Models are usually created gradually adding more and more information that is either theoretical or the results of practical experiments.[2] This procedure can be subdivided into several steps; defining the overall purpose of the model, defining the network structure of the model, identification of kinetic rate expressions, completion of the model structure, parameter determination and validation of the model.[15]
Eventually, the model is intended to describe the temporal progressions of the state variables of a system. The state variables describe the mass and energy accumulations of the system. Using the input variables that are not system-dependent, output variables and the states can be calculated via the model. The model thus relates all physical variables of a system to each other via mathematical descriptions. In addition to states, input and output variables, these mathematical descriptions also contain other model-describing variables called parameters.[2]
After a model is described mathematically, it must be calibrated, analyzed and validated.[16]

1.3.2 From classical to digital twin based process control

Digital twins have the potential to be utilized for the development and implementation of conventional controllers, advanced control strategies and holistic bioprocess control eventually. In contrast to the chemical industry, where digital twins are already used in the development of control strategies, the application possibilities for bioprocesses are still in their infancy. This is due to the fact that in bioprocesses many reactions take place simultaneously and small deviations of process variables can already have a strong influence on the kinetics.[12] Due to this complexity of bioprocesses there are still various problems to tackle until fully developed digital twins can be realized for industrial processes.[10] In process control, digital twins can be used to determine the appropriate controller type, to improve control performance and to enhance the overall process performance through appropriate control strategies. Using a digital twin conventional control strategies yield comparable results as more complex nonlinear model predictive control. Advanced and model-based control strategies are already investigated for various processes.[12] Model-based and model predictive control can be based on simple mathematical models [15] and are interesting for bioprocesses because reduced response times of these control strategies lead to higher product concentrations than conventional controllers.[12]
In this work, methods shall be developed in order to be able to control and investigate correlations between feed rate and other typical rates in Fab-Fragment producing *E. coli* Fed-Batch process. These control models shall than be examined and improved by extending a process model towards an applicable digital twin, which in addition to product formation and growth kinetics shall include constraints such as mass transfer limitations and elemental balances. The controller and the process model have to be adapted and improved to be able to make better process predictions and thus to be able to answer concrete

questions about the investigated system. In particular, the various relationships between feed design and process-relevant biomass specific rates such as the substrate uptake rate, the growth rate, the product formation rate, as well as the oxygen uptake rate and the CO2 emission rate will be examined.

The method of non-linear feedback linearization is applied to create control laws to keep the specific rates constant during the production phase. A non-linear relationship between the controlled variable and the input variable can be transformed by this method into a linear form that compensates for the non-linearities of the system and can thus be integrated into classical linear controller designs.

1.4 Novelty

Digital twins are an interesting approach to improve process efficiency, product quality and safety in industrial bioprocesses. Despite the many sources and articles appearing recently on Digital Twins in the biochemical industry, most approaches are very general and conceptual. Literature regarding implementation of the concept is rare.[17]

In industrial applications advanced control strategies like multivariate control, model-based control and adaptive control is still rarely used for bioprocesses.[18] On a scientific scale model predictive control is a widely investigated approach [19][20]. In [21] it is shown that the simpler approach of model-based control cannot compete with model predictive control due to the non-linearities and the limitation to the linearized control space. These two problems can be addressed with the method of nonlinear feedback linearization used in this work.

To generate control laws which incorporate the nonlinearities of a fermentation process the method of nonlinear feedback linearization has recently been used by [22] and [23]. In my work, the method of feedback linearization in combination with a generic model is used to control biomass specific rates of a fed batch fermentation process. After the performance of the control laws has been checked and optimized, it is discussed and elaborated which advantages and disadvantages the control of the individual rates possesses and, above all, an outlook on implementing the laws for real processes is given with regard to existing measurement methods.

To control biomass specific substrate uptake rate, growth rate, oxygen uptake or carbon dioxide emission rate via the feedrate in a fed-batch fermentation process, the approach of feedback linearization has not been used so far. As the respiratory rates could be determined quite straight forward via mass balances over aeration and off-gas measurements, the control of the oxygen uptake rate or the carbon dioxide emission rate could be interesting approaches regarding the implementation of evolved digital twins for this type of fermentation process.

Chapter 2

Goals and workplan of the thesis

2.1 Goals

This thesis aims to elucidate possibilities to control relevant biomass specific rates directly or indirectly of a Fab-fragment producing *E. coli* fed-batch fermentation process. To achieve this overall goal, the following milestones had to be worked on successively:

1. An recently developed model had to be extended and parameterized to a broader set of data. The extension of the model involved the integration of off-gas dynamics such as oxygen consumption and CO_2 emission by elemental balancing.

2. Subsequently, various control laws were obtained from the nonlinear dynamic model by the method of feedback linearization. Control laws to control biomass specific rates via feed supply for substrate uptake, growth, product formation, oxygen uptake and carbon dioxide emission were received for the intention to keep this rates at constant setpoints and investigate the resulting process behavior.

3. To examine the performance of this control laws a setpoint optimization was carried out

4. An independent experimental data set was parameterized to establish a virtual plant for the simulation study to gain knowledge on the potential applicability

2.2 Workplan

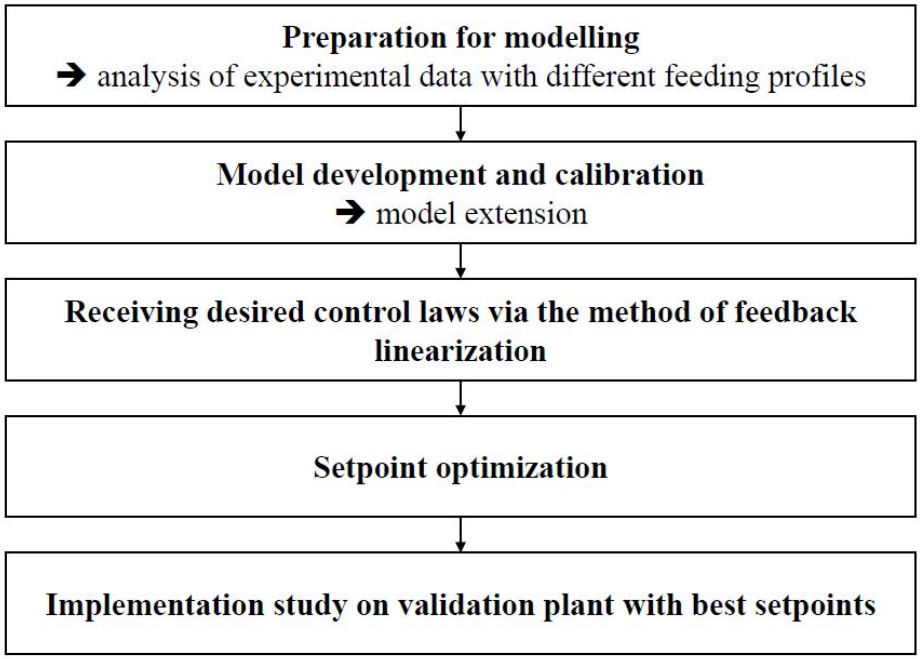

Figure 2.1: Workplan

Chapter 3

Material and methods

3.1 Process model

To set up a model of a bioprocess, one needs differential equations that can be derived from the material balances 3.1, 3.2 of the system. In combination with the kinetics of the system, a process can be described in a simplified way. The resulting descriptions promote the understanding of the processes and can be used for control and optimization of the processes. [2]
The material balances of a fed-batch bioreactor can be written as follows:

$$\frac{dm}{dt} = \dot{m}_{in} - \dot{m}_{out} \tag{3.1}$$

and further the material balance for a single component i in the broth:

$$\frac{dm_i}{dt} = \dot{m}_{i,in} - \dot{m}_{i,out} + r_i \tag{3.2}$$

With m_i, in and m_i, out being the mass in- and outfluxes of a component i and the reaction r_i within the reactor. Accordingly, the following general balance results for the concentration of any component i 3.3 in the reactor. V_R is the reactor volume, c_i describes the related concentration, $c_{i,in}$ and $c_{i,out}$ describe inflow and outflow of component i respectively, $\dot{V}_{in}$ the volumetric flux into the system and r_i the reaction rate of component i

$$\frac{d(c_i V_R)}{dt} = c_{i,in}\dot{V}_{in} - c_{i,out}\dot{V}_{in} + V_R r_i \tag{3.3}$$

3.1.1 System kinetics

Substrate consumption

The biomass specific substrate uptake rate q_S can be described using a Monod Kinetic.

$$q_S = q_{S,max}\frac{c_S}{c_S + K_S}, \tag{3.4}$$

in 3.4 $q_{S,max}$ describes the maximal biomass specific substrate uptake rate, c_S is the mass concentration of substrate and K_S is the saturation constant for substrate uptake.

Biomass growth rate

The biomass growth rate μ is described by 3.5

$$\mu = Y_{X/S}(q_S - M) \tag{3.5}$$

where $Y_{X/S}$ is the coefficient which describes the metabolization of substrate into biomass. The maintenance coefficient M describes the limitation of the growth rate by the energy demand of the cells required for maintenance and repair, so in the case of strong substrate limitation, substrate uptake can occur without any growth at all.[2] The yield can then be further expressed by 3.6 where the maximum possible yield $Y_{X/S,max}$ and an asymptotic decay $K_{Y_{X/S}}$ of the metabolized sugar S_{met} to account for an decreasing yield during the recombinant protein formation phase.

$$Y_{X/S} = Y_{X/S,max} e^{(-S_{met} K_{Y_{X/S}})} \tag{3.6}$$

Product formation

The product formation is composed of two terms, a monod term and a Haldane term, which describes the product inhibition. Inhibition in biotechnological processes is often referred to as the reverse phenomenon of substrate limitation, a reduction in growth, substrate uptake rate, or as in the scenario here, a reduction in product formation occurs as the concentration of an acting substance increases. [2]

$$q_P = q_{P,max} \frac{q_S}{q_S + K_{S_{q_S}}} \frac{S_{met}}{\frac{S_{met}^{Hal}}{K_{I_{q_P}}} + S_{met} + K_{S_{q_P}}} \tag{3.7}$$

In 3.7 the Monod term the dependence of qp_{max}, the maximum product formation rate on q_S with the half saturation term $K_{S_{q_S}}$ is described. The Haldane term describes the start and decay phase with S_{met} as the trigger. $K_{S_{q_P}}$ is the delay coefficient $K_{I_{q_P}}$ the decay coefficient and Hal is the Haldane exponent.

Product release

In addition to the product formation, an expression for the product release has to be considered since product is also observed outside the cells. This product release 3.8 can be determined based on the metabolized sugar and the substrate uptake rate.

$$q_{P_r} = K_{q_s rel}(q_S - M)(1 - e^{\frac{S_{met}}{K_{S_{Ps}}}}) \tag{3.8}$$

Where $K_{q_S rel}$ describes the half-speed constant, and $K_{S_{Ps}}$ the delay coefficient.

Respiration

The specific oxygen rate q_{O_2} and the specific CO_2 rate q_{CO_2} can be obtained via the material balances of carbon and the degree of reduction balance. In order to draw conclusions from the material balance, the compositions of substrate, biomass and product are essential.[24] The material balance can be written as follows.

$$\mathrm{CH_mO_n} + a\,\mathrm{O_2} - b\,\mathrm{NH_3} \longrightarrow c\,\mathrm{CH_\alpha O_\beta N_\delta} + d\,\mathrm{CH_xO_yN_z} + e\,\mathrm{H_2O} + f\,\mathrm{CO_2}$$

The first term CH_mO_n describes the substrate used which was Glycerol in the case of the considered fermentation. The oxygen and ammonia inputs are not relevant for the C-balance. Ammonia can also be neglected for the DoR balance, since the DoR is $\gamma_{NH_3} = 0$. This also applies to H_2O and CO_2 on the right hand side of the chemical equation. The composition of the biomass $CH_\alpha O_\beta N_\delta$ was available, the composition of the product $CH_xO_yN_z$ for which a standard Fab-fragment composition was used. To obtain expressions for the respiration-specific rates the C-balance 3.9 and the DoR-balance 3.10 are written as shown below. The rates must behave in such a way that their sum is always zero.

$$r_S + r_X + r_P + r_{CO_2} = 0 \tag{3.9}$$

$$r_S\gamma_S + r_{O_2}\gamma_{O_2} + r_X\gamma_X + r_P\gamma_P = 0 \tag{3.10}$$

The stoichiometric rates here represent the substrate uptake rate r_S, the biomass formation rate r_X and r_P the product formation rate, the oxygen uptake rate (OUR) r_{O_2} and the carbon dioxide emission rate (CER) r_{CO_2}. The biomass specific rates for oxygen uptake and CO_2 release can now be written in terms of the specific rates for biomass and product formation and substrate uptake previously explained. The biomass specific rates for oxygen uptake 3.11 and CO2 release 3.12 can now be expressed by the specific rates previously explained and the molar masses for the substrate, biomass, and product which are shown in table 3.1

$$q_{O_2} = -\frac{1}{\gamma_{O_2}}\left(\frac{\mu\gamma_X}{M_X} + \frac{q_P\gamma_P}{M_P} - \frac{q_S\gamma_S}{M_S}\right) \tag{3.11}$$

$$q_{CO_2} = \frac{\mu}{M_X} + \frac{q_P}{M_P} - \frac{q_S}{M_S} \tag{3.12}$$

	Substrate	**O_2**	**Biomass**	**Product**	**Dim**
DoR	4.667	-4	4.159	–	–
M	30.67	16	25	23.25	$gmol^{-1}$

Table 3.1: Degree of reduction γ and molar mass of every compound used in the stoichiometric balance to get the expression for the biomass specific oxygen uptake and CO_2 emission rate depending on the specific rates of growth, product formation and substrate consumption

3.1.2 Differential equations

To describe a simple fed-batch fermentation process, a dynamic model must consist at least of three conservation mass balances, one for the cells, one of the substrate and the overall mass or mass concentration respectively.[25] The fermentation process employed here can be expressed by the following differential equations 3.13 for the changing reactor volume, the mass concentrations of the biomass, of the substrate and the product both intracellular and released as well as the metabolized substrate. These equations are derived from the mass concentration balance of any component i (3.3).

$$\begin{aligned}
\frac{dV_R}{dt} &= \dot{V}_{in} \\
\frac{dc_X}{dt} &= \mu c_X + \frac{\dot{V}_{in}}{V_R}(c_{X,in} - c_X) \\
\frac{dc_S}{dt} &= q_S c_X + \frac{\dot{V}_{in}}{V_R}(c_{S,in} - c_S) \\
\frac{dc_P}{dt} &= q_P c_X - q_{P_R} c_P - \frac{\dot{V}_{in}}{V_R} c_P \\
\frac{dc_{P_R}}{dt} &= q_{P_R} c_P - \frac{\dot{V}_{in}}{V_R} c_{P_R} \\
\frac{dS_{met}}{dt} &= q_S
\end{aligned} \tag{3.13}$$

The change in reactor volume V_R results from operating in fed-batch mode, where a continuous feed but no outflow is presumed to $\frac{dV_R}{dt} = \dot{V}_{in}$. The concentration changes of the cell mass c_X, the substrate in broth c_S, the intracellular product c_P as well as the released product c_{P_R} and the metabolized substrate S_{met} in the system can be described by the associated specific rates and the dilution effect caused by the substrate feed $\dot{V}_{in}$.

3.2 Experimental data

The experimental data used was generated by a previous master thesis (by Julian Kager 2015) [26] and provided for the realization of this work. For the process, a modified K12 E. coli strain was used for the production of a recombinant protein which was a Fab-fragment of an Antibody. The strain has a rhamnose-inducible expression system (rhaBAD promoter) as an expression vector. The fermentations were run in a DASGIP multi-bioreactor system (Eppendorf, Hamburg, Germany). DASGIP modules were used to control components such as the stirrers and pumps for feeding and base addition, to control aeration and to measure pH, pO2 and temperature. Throughout the process, the temperature was kept at 35°C, the stirrer speed at 1400, the gassing at 1.4 vvm and the pH at 7 using a 12.5% NH_4OH solution. The standardized synthetic media used, contained Glycerol as carbon source. The batch phase was started from a pre-culture of 2.5% of the 1L batch volume. For the 12 hour

batch phase, a 20 $gL^{-}1$ glycerol feed was exponentially fed. With the one-time addition of a sterile rhamnose solution (1.5 g L-rhamnose) at the end of the exponential feed phase, the recombinant protein production was started and the glycerol feed was set to a constant value or linearly increased. For each experiment, 8 offline samples were taken during the induction phase and evaluated manually. The biomass cell dry weight was determined by centrifuging 2 ml of broth (4500g, 10 min at 4°C), washing the pellet with 5 ml of water and weighing it after drying. An analogue pellet was disrupted and re-buffered to determine the intracellular product and the supernatant to determine the extracellular product. A protein G affinity column (HiTrap ProtG, GE Healthcare, USA) was used for this purpose. Acetate and glycerol concentrations were determined from the supernatant by enzymatic, photometric principle in a robotic system (BioHT, Roche, Germany) and were not possible to measure because of the detection limit and the limiting conditions during induction. Specific rates were calculated using a predefined algorithm.[26]

3.3 Parameter estimation

Mechanistic models partially contain terms that are based on experience or assumptions about the behavior of the described system. The parameters of these terms must be estimated, measured or determined from experimental data.[16][2]
Since biological processes are very complex and models do not describe all factors that influence the system, the determination of parameters is more difficult than in physical models.[16]
The process of gaining parameters that describe a system quantitatively is referred to as parameter estimation (identification), model calibration or parameter fitting. Parameters of a model are determined using statistical methods which provide approached values for the investigated system.[27]
These statistical methods usually describe offsets between preliminary estimations and data as well as how reliable resulting values are. Parameter estimation therefore tries to minimize a distance criterion between prediction and data.[16] For example, parameters could be estimated using the log likelihood function, which is a very common method. [27] A well-known algorithm for optimizing a bioprocess model is the Nelder-Mead simplex algorithm.[16]
A further approach to fit predicted parameters is the weighted sum of squared residuals.[28] It is used to validate an error over the trial time span of one single state.[29]
The strategy used in [16] to estimate the parameters consists of systematic testing, model analysis, model calibration and model validation. The presented workflow for model calibration consists of 3 steps. After an initial set of parameters has been estimated, the first step consists of analyzing the resulting model. This analysis step includes a sensitivity analysis of all parameters, colinearities are determined and an identification analysis is performed. In the second step, the information obtained will be used to restrict the limits between which parameters are estimated. Once a new parameter estimation has been carried out,

the model obtained must be validated. Error calculations are used to determine whether the model sufficiently describes the processes under investigation.
A similar workflow is illustrated in [30] and shown in 3.1

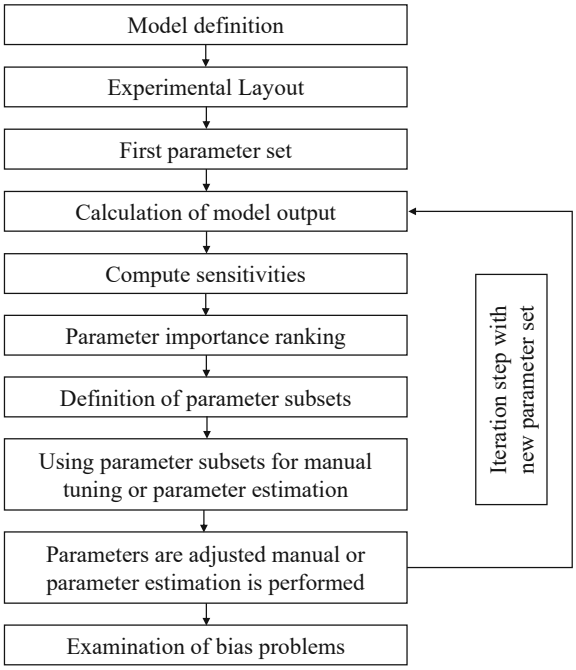

Figure 3.1: Workflow of systematic parameter subset selection and tuning procedure [30]

In the first step the model is defined as explained before, specifying process balances, inputs and initial conditions. Step 2 contains the definition of the experimental layout. The third step is based on expert knowledge and essential for the quality of the resulting conclusions. It shall provide a first set of parameters for iteration including quantified uncertainties and an assessment of the predictability from process data. Furthermore scales of the different outputs shall be estimated. Step 5 to 7 are the calculation of sensitivities and identifiability of parameter subsets. Parameter subsets are identifiable if they are linear independent firstly and secondly made up of sufficiently sensitive parameters. Choosing parameter subsets for manual tuning (step 9) is carried out in Step 8, which is also depending on expertise. Subsequently potential bias problems of the parameter values gained by the estimation process could be examined. [30]

NRMSE

The normalized root-mean-square error 3.14 can be used to examine how high the deviation between an individual state is compared to its model prediction. It is often used as an

acceptance criterion.[29] If the deviations of all measuring points from a predicted value and the corresponding actual value are summed up, i.e. the errors, divided by the number of measuring points and then the square root of this average prediction error is calculated one obtains the root-mean-square error (RMSE). Usually this value is normalized by dividing it by the difference between the largest and the smallest value measured.

$$NRMSE(c_i) = \frac{\sqrt{\frac{1}{d}\sum_{j=1}^{d}(c_{i,j} - c_{i,j}^*)^2}}{max(c_i) - min(c_i)} \tag{3.14}$$

c_i is the concentration of any state i, d the number of measuring points an c_i^* the model prediction of the state i. The smaller the error, the better the model fits to describe the corresponding state. In this study the NRMSE is additionally used to describe the control quality of the received control laws as shown in 3.15

$$NRMSE(q_i) = \frac{\sqrt{\frac{1}{d}\sum_{j=1}^{d}(q_{i,j} - q_{i,j}^*)^2}}{(\bar{q}_i)} \tag{3.15}$$

In this case the RMSE for any controlled rate q_i is normalized by the mean value of the concerning rate $\bar{q}_i$.

3.4 Process control

Even though the scope of bioprocess control does not only concern the fermentation process but also various product recovery and purification steps, the bioreactor itself is the biggest challenge in terms of process control.[31] In general, control can be described as the manipulation of a controlled variable to reduce or eliminate measured deviations from a desired status.[32] Bioprocess control can be defined in this context as the creation and maintenance of a close to optimal environment for microorganisms to grow, multiply and produce the desired products. For the further approach it is important to underline that the process control of fermentations is based on influencing internal dynamics by manipulating the external environment. To create this environment, important cellular parameters such as temperature and pH must be kept at an ideal level for the biomass, and the optimum nutrient supply must be ensured. Benefits that can be achieved through better control strategies include reduction of process variability, higher productivity and product yields, and better on-line monitoring and troubleshooting for more automated processes. [31]

3.4.1 Common control strategies for bioprocesses

It is very common way to control a fed-batch fermentation process by manipulating the feed-rate as it directly affects the proceedings.[1] Feeding profiles should be smooth to generate stable processes. Smooth feeding profiles have important significance for process stability because abrupt changes in the environment of the cells affect cell metabolism in a way that negatively impacts productivity.[3]
Control strategies used to regulate the feed rate are open loop control, adaptive control, model predictive control (MPC), fuzzy control, artificial neural networks (ANN), probing control and statistical process control (SPC).[1]

For open loop control, a defined feed is added to the process. Such a profile is determined beforehand in such a way that a desired process behavior can be expected. An example of how predefined feed profiles could be used in industrial processes is the control of the growth phase, where exponential feed profiles can be defined taking into account the initial conditions and strain-specific parameters. Even if disturbances cannot be taken into account, open loop control is an interesting way to control processes due to its simplicity.[1]

In closed-loop control, online information is additionally used to correct the manipulated variable in such a way that firstly the output variable is kept at the required setpoint and secondly the manipulated variable can be regulated to the target setpoint as accurately as possible and with little time delay. With open loop control, a distinction can be made between feedback and feedforward control. The difference is that feedback control measures the controlled variable and reacts to its deviations, whereas feedforward control measures the disturbance and thus, in the best case, influences the deviation of the controlled variable in advance. Since feedforward control requires a model that can predict the controlled variable and in application usually does not make a perfect estimate, feedback control is often the preferred strategy.[32] Feedback control with classical PI or PID controllers is often used in biotechnological processes to control variables such as temperature, pH or dissolved oxygen. When controlling other process variables, problems can arise due to a lack of measurement methods or non-linear process behavior. This problem can be addressed with kinetic models. Unmeasured conditions can be estimated and thus regulated.[21]

3.4.2 Feedback linearization

Unlike conventional linearization, the feedback linearization method used here is not simply a linear approximation of a non-linear behavior in a chosen operating point.[33] In feedback linearization, a nonlinear system is linearized by a nonlinear coordinate transformation.[34] The purpose of this method is to generate a nonlinear feedback controller which compensates the nonlinearity of the system exactly and therefore can be integrated in a linear control loop.[33] Here, an input affine single-input single-output (SISO) system is considered, which

can be described as shown in 3.16 due to [33]

$$\begin{aligned} \dot{\boldsymbol{x}} &= \boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})u \\ y &= h(\boldsymbol{x}) \end{aligned} \tag{3.16}$$

In the first equation, the state equation $\boldsymbol{x}$ describes the the state vector, u the input and $\boldsymbol{f}(\boldsymbol{x})$ and $\boldsymbol{g}(\boldsymbol{x})$ the nonlinear mappings. The second equation (output equation) y describes the output of the system and $\boldsymbol{h}(\boldsymbol{x})$ the nonlinear mapping.[25] Now the derivative of the output results in 3.17

$$\begin{aligned} \dot{y} &= \frac{dh(\boldsymbol{x})}{dt} \\ &= \frac{\partial h(\boldsymbol{x})}{\partial \boldsymbol{x}}\dot{\boldsymbol{x}} \\ &= \frac{\partial h(\boldsymbol{x})}{\partial \boldsymbol{x}}\Big(\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{g}(\boldsymbol{x})u\Big) \\ &= L_{\boldsymbol{f}}h(\boldsymbol{x}) + L_{\boldsymbol{g}}h(\boldsymbol{x})u \end{aligned} \tag{3.17}$$

The terms $L_{\boldsymbol{f}}h(\boldsymbol{x})$ and $L_{\boldsymbol{g}}h(\boldsymbol{x})$ are the so called Lie-derivatives along $\boldsymbol{f}(\boldsymbol{x})$ and $\boldsymbol{g}(\boldsymbol{x})$ of $h(\boldsymbol{x})$.[33]
With $L_{\boldsymbol{g}}h(\boldsymbol{x}) = 0$ which is the case in the vast majority of technical applications, the derivative of the output 3.18 can be written

$$\dot{y} = L_{\boldsymbol{f}}h(\boldsymbol{x}) \tag{3.18}$$

The relative degree δ of the system describes the order of the higher derivative of y where the term $L_{\boldsymbol{g}}{L_{\boldsymbol{f}}}^{\delta-1}h(\boldsymbol{x}) \neq 0$. A system where $\delta = dim(\boldsymbol{x})$ is said to have full relative degree n. Every system with a relative degree $\delta < dim(\boldsymbol{x})$ comprise so called internal dynamics. Unlike external dynamics which describe the changes of the states and therefore influence the system output, internal dynamics have no effect on the output variable and can not be observed either.
The nonlinear transformation z of a system with internal dynamics leed to 3.19

$$\boldsymbol{z} = \begin{bmatrix} z1 \\ z_2 \\ \vdots \\ z_\delta \\ z_{\delta+1} \\ \vdots \\ z_n \end{bmatrix} = \boldsymbol{t}(\boldsymbol{x}) = \begin{bmatrix} h(\boldsymbol{x}) \\ L_{\boldsymbol{f}}h(\boldsymbol{x}) \\ \vdots \\ {L_{\boldsymbol{f}}}^{\delta-1}h(\boldsymbol{x}) \\ t_{\delta+1}(\boldsymbol{x}) \\ \vdots \\ t_n(\boldsymbol{x}) \end{bmatrix} \tag{3.19}$$

where the elements of $t_{\delta+1}(\boldsymbol{x}), ..., t_n(\boldsymbol{x})$ are chosen arbitrarily so that $\boldsymbol{t}$ satisfies the requirement to be continuously differentiable and additionally has an inverse function 3.20

$$\boldsymbol{x} = \boldsymbol{t}^{-1}(\boldsymbol{z}) \tag{3.20}$$

for which the same applies.
For the SISO system introduced in the beginning of the chapter the derivative transformation yields

$$\dot{\boldsymbol{z}} = \begin{bmatrix} \dot{z_1} \\ \dot{z_2} \\ \vdots \\ \dot{z_{\delta-1}} \\ \dot{z_\delta} \\ \dot{z_{\delta+1}} \\ \vdots \\ \dot{z_n} \end{bmatrix} = \boldsymbol{t}(\boldsymbol{x}) = \begin{bmatrix} L_{\boldsymbol{f}} h(\boldsymbol{x}) \\ {L_{\boldsymbol{f}}}^2 h(\boldsymbol{x}) \\ \vdots \\ {L_{\boldsymbol{f}}}^{\delta-1} h(\boldsymbol{x}) \\ {L_{\boldsymbol{f}}}^{\delta} h(\boldsymbol{x}) + L_{\boldsymbol{g}} {L_{\boldsymbol{f}}}^{\delta-1} h(\boldsymbol{x}) u \\ \dot{t}_{\delta+1}(\boldsymbol{x}) \\ \vdots \\ \dot{t}_n(\boldsymbol{x}) \end{bmatrix} \tag{3.21}$$

For 3.21 the derivatives for the terms t_i for $i = \delta + 1, \ldots, n$ are determined such that the dependence on u disappears and therefore leads to 3.22

$$\begin{aligned} \dot{t}_i(\boldsymbol{x}) &= \frac{\partial t_i(\boldsymbol{x})}{\partial \boldsymbol{x}} \dot{\boldsymbol{x}} \\ &= \frac{\partial t_i(\boldsymbol{x})}{\partial \boldsymbol{x}} \Big(\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{f}(\boldsymbol{x}) u \Big) = L_f t_i(\boldsymbol{x}) + \underbrace{L_g t_i(\boldsymbol{x})}_{=0} u \\ &= L_f t_i(\boldsymbol{x}) \end{aligned} \tag{3.22}$$

The internal dynamics are not linearizable and therefore not controllable and as mentioned before the internal dynamics can not be observed and have no impact on the output $y(t)$. The external dynamics on the other hand can be linearized exactly by transforming the input variable u to a new input v.
For this new input v a classical state vector feedback control law 3.23 is used

$$v = -\boldsymbol{k}^T z + K_w w \tag{3.23}$$

where the vector $k^T = [\alpha_0 \ \alpha_1 \ \ldots \ \alpha_{\delta-1} \ 0 \ \ldots \ 0]$ describes the gains. This results in the following for the actual nonlinear control law 3.24

$$u = -\frac{L_{\boldsymbol{f}}^{\delta} h(\boldsymbol{x}) + \boldsymbol{k}^T z}{L_{\boldsymbol{g}} L_{\boldsymbol{f}}^{\delta-1} h(\boldsymbol{x})} + \frac{K_w}{L_{\boldsymbol{g}} L_{\boldsymbol{f}}^{\delta-1} h(\boldsymbol{x})} w \tag{3.24}$$

which can also be written in the form 3.25

$$u(\boldsymbol{x}, w) = -r(\boldsymbol{x}) + q(\boldsymbol{x}) w \tag{3.25}$$

The term $q(\boldsymbol{x})$ which depends on the reference input w can be considered as pre-filter. In the closed loop system as shown in figure 3.2 the term $r(\boldsymbol{x})$ describes the feedback part.

With $Kw = \alpha_0$ the terms can be written as 3.26

$$\begin{aligned} r(\boldsymbol{x}) &= \frac{L_f^{\delta} h(\boldsymbol{x}) + \boldsymbol{k}^T z}{L_g L_f^{\delta-1} h(\boldsymbol{x})} \\ q(\boldsymbol{x}) &= \frac{\alpha_0}{L_g L_f^{\delta-1} h(\boldsymbol{x})} \end{aligned} \tag{3.26}$$

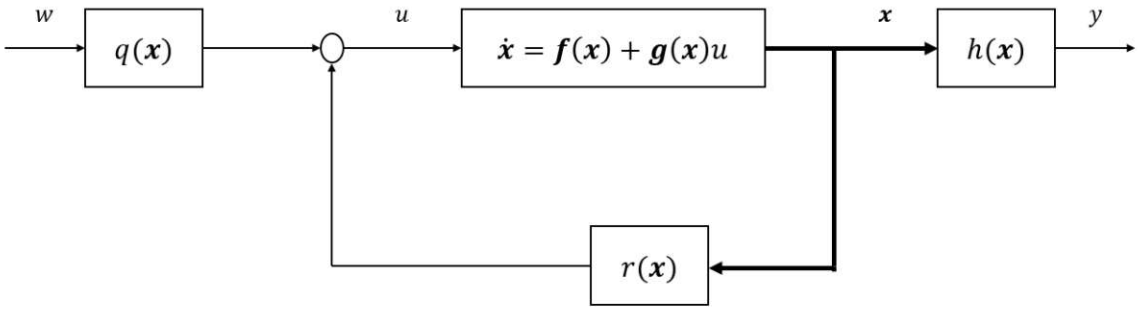

Figure 3.2: Closed loop block diagram of non-linear SISO system [33]

3.4.3 Controller design

Since control laws obtained by feedback linearization compensate for all non-linearities, they can therefore be integrated into simple open-loop or closed-loop controllers.

Open-loop controller

For the open-loop control, a suitable feed profile was calculated via the model and handed over to the process without integrating any information regarding the course of the process into the controller. The open-loop controller can be represented as shown in 3.3.

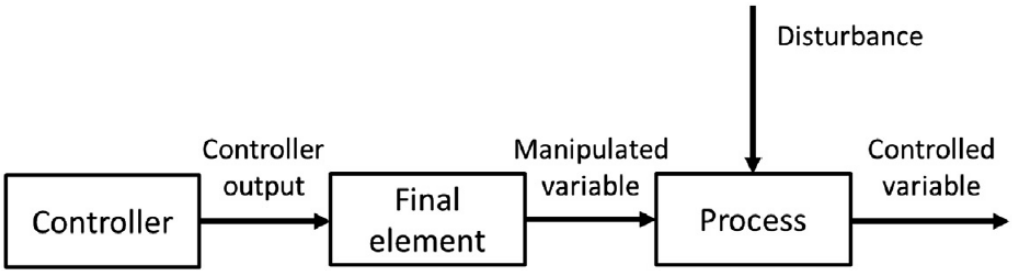

Figure 3.3: Example of open loop control system [32]

Closed-loop controller

A two-degrees-of-freedom controller was chosen as the closed-loop controller. Degree of freedom in this context is described as the number of independently adjustable transfer functions. Thus, the two previously described closed-loop strategies can be combined, as shown in figure 3.4. The controller thus still has an output variable but an additional input, the setpoint r and the controlled variable y, on the basis of which the manipulated variable u is determined. The transfer function of the feedback part $C(s)$ is called the serial compensator, $C_f(s)$ is referred to as the feedforward compensator. [35]

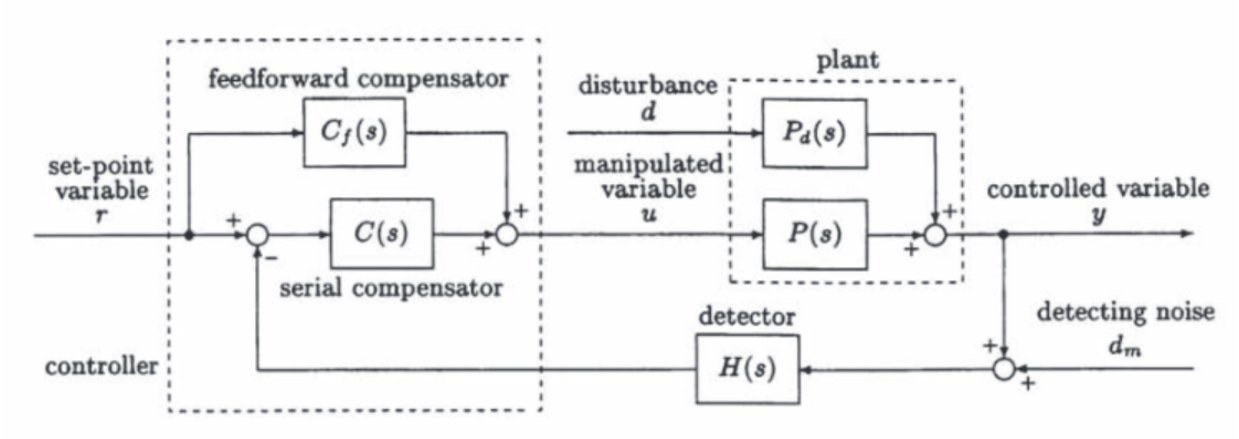

Figure 3.4: Example of two-degree-of-freedom control system. $C(s)$ describes the function for the serial compensator, $C_f(s)$ the feedforward compensator, together they respresent the controller. The function $H(s)$ incorporates the feedback and noise detection, $P(s)$ is the mathematical description of the system with the corresponding disturbances $P_d(s)$ [35]

3.5 Software

The computational work to analyze data as well as to evaluate the process behavior using a model were performed in Matlab R2020b (The MathWorks, Inc., USA). Using an ODE solver, the time-dependent behavior of the states as subjects of the feed profile can be calculated by solving the system of differential equations in 3.13.

Chapter 4

Results and discussion

4.1 Model description

A recently developed kinetic, unstructured model for the induction phase of the fermentation process was extended and parameterized to a broader set of data. Therefore the experimental data described in 3.2 on page 11 was used. The extension of the model involves the integration of off-gas dynamics such as oxygen consumption and CO_2 emission. The respiratory rates were included in the model using elemental balancing described in chapter 3.1.1 on page 10. This should enable the model to calculate the specific rates for the substrate consumption q_S, the growth μ, the product formation q_P, the oxygen demand q_{O_2}, and the CO_2 emission q_{CO_2} in the course of the process and to subsequently control them.
To fit model parameters, 6 experiments were selected from the series of trials. Care was taken to ensure that the experiments were run at feed rates that varied as much as possible during the induction phase of the fermentation and that they were stable processes that could be described as well as possible with the known kinetics. Three experiments with constant feed rate at three different setpoints were selected, one experiment with very low feed rate, one with medium and one with high feed rate. In addition, three experiments with dynamically changing feed rates were selected, and profiles with different start values and gradients were chosen. The feed profiles of the selected experiments on which the model was fitted are shown in the graph below 4.1.

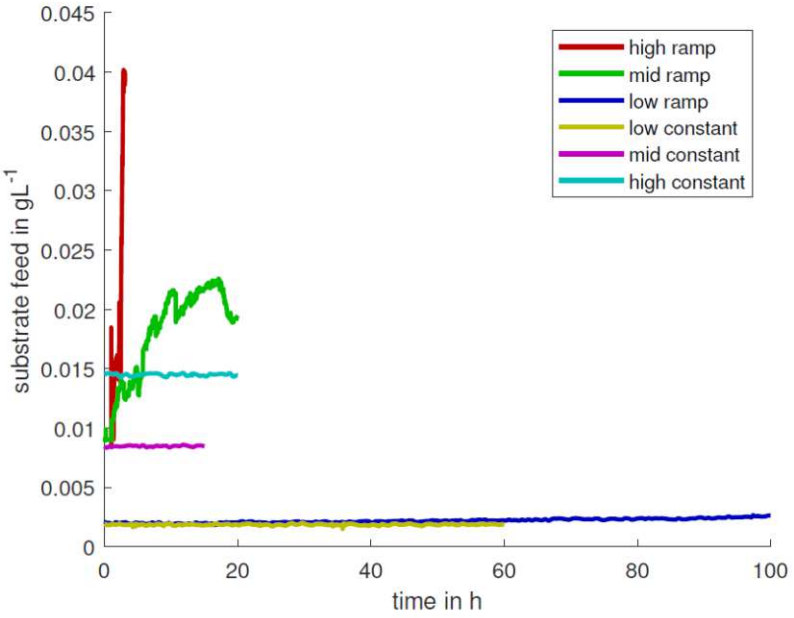

Figure 4.1: Feed profiles of Experiments chosen from the data set provided

4.1.1 Model evaluation

The parameters for each experiment as well as of the resulting model were now calculated as described in chapter 3.3. Graphs 4.3, 4.2, 4.7, 4.6, 4.5, 4.4 show the deviations of the final model estimations from the data of the individual experiments. The model estimates are compared with the data of the biomass concentration, the product concentrations in the cell and outside the cell as well as the total product concentration. In addition, the deviations of the specific product formation rate, the substrate uptake rate and the respiratory rates CER and OUR are presented.
For the experiment with the very high and rapidly increasing feed rate, the model underestimates the measured biomass and consequently also the product concentrations formed and the associated product formation rate. For the other two experiments with increasing feed rates, all process variables shown are described very well by the model. The measured biomass concentrations are slightly underestimated for all experiments. Especially for the experiment with constant and very low feed rate the growth within the experiment is significantly higher than the model is expecting. The product concentrations are generally very well represented by the model.

It should be noted that the model does not accurately describe the process behavior after the maximum has been reached. This poor description towards the end of the process can

be explained by the fact that the dynamics for cell lysis and product degradation are not incorporated in the model.

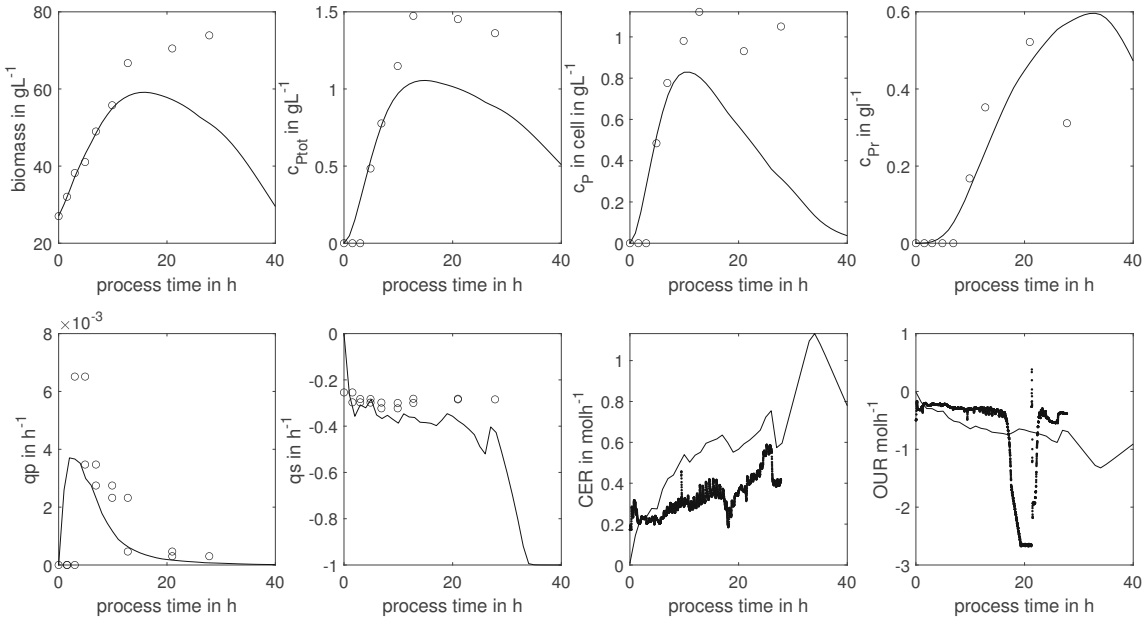

Figure 4.2: Experiment with increasing feed rate and high starting point: comparison between experimental data and model estimation, illustrated in red color in figure 4.1

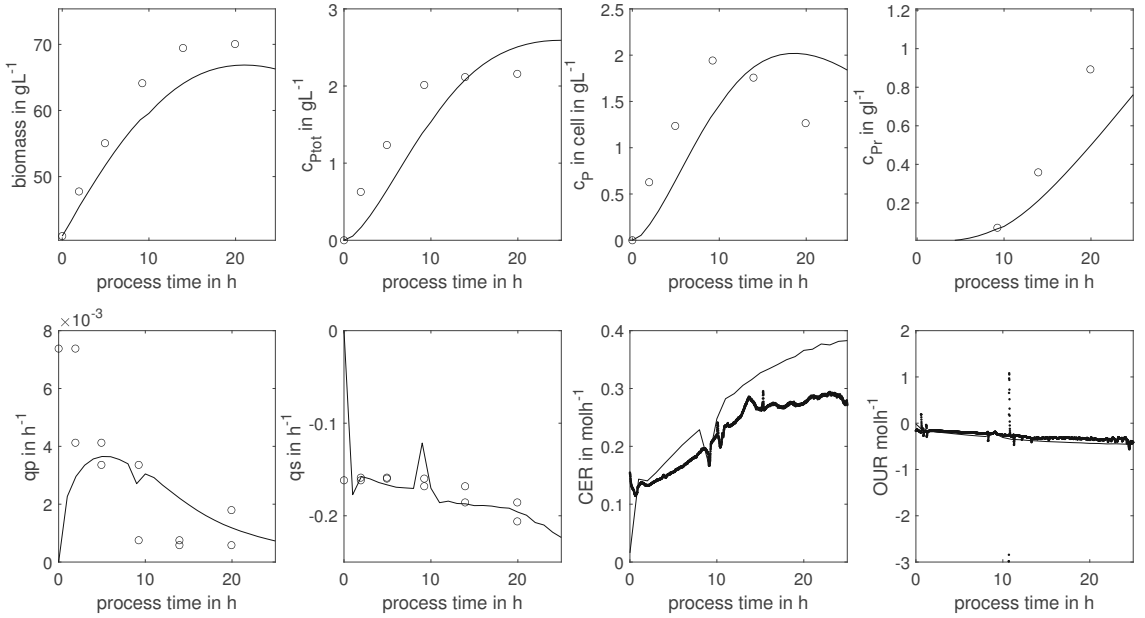

Figure 4.3: Experiment with increasing feed rate and medium starting point: comparison between experimental data and model estimation, illustrated in green color in figure 4.1

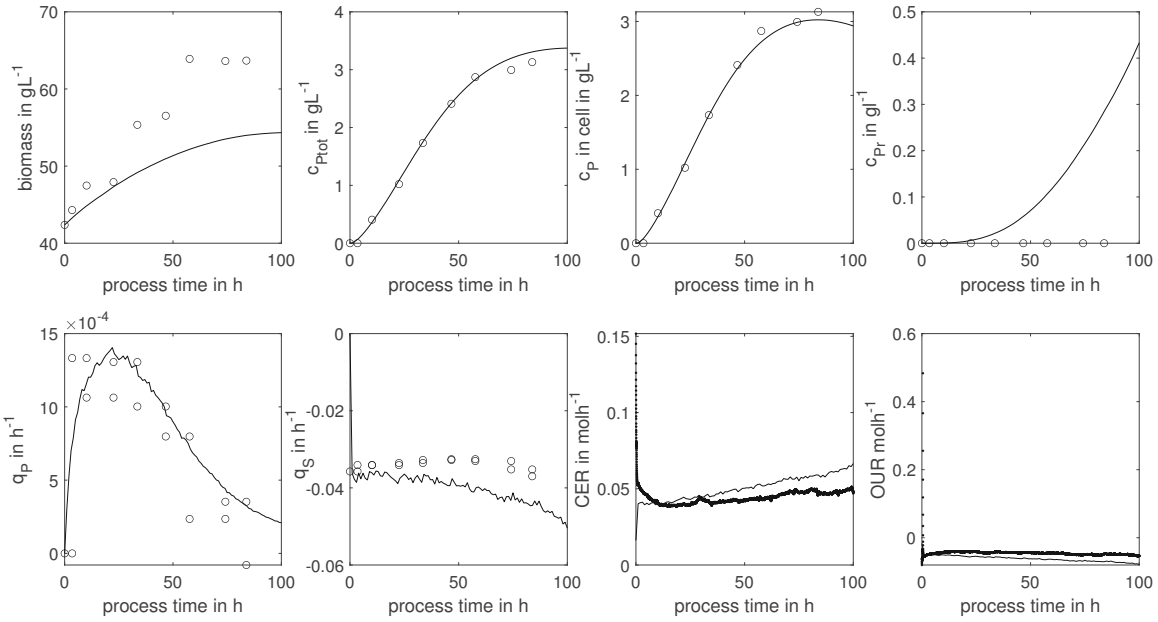

Figure 4.4: Experiment with increasing feed rate and low starting point: comparison between experimental data and model estimation, illustrated in blue color in figure 4.1

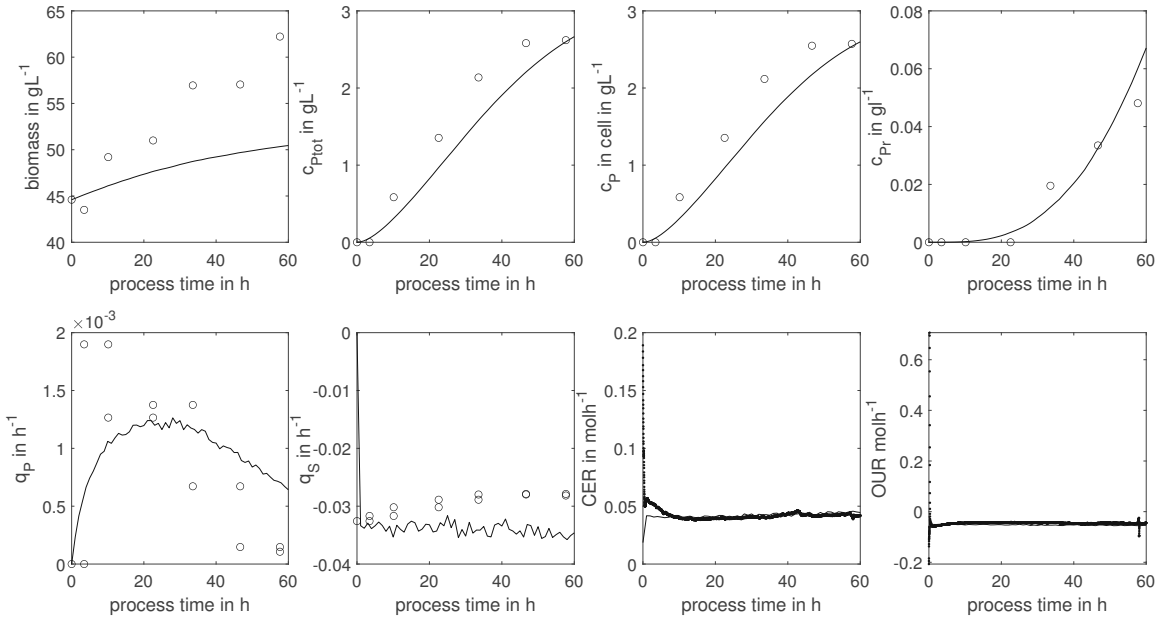

Figure 4.5: Experiment with constant feed rate and low set point: comparison between experimental data and model estimation, illustrated in yellow color in figure 4.1

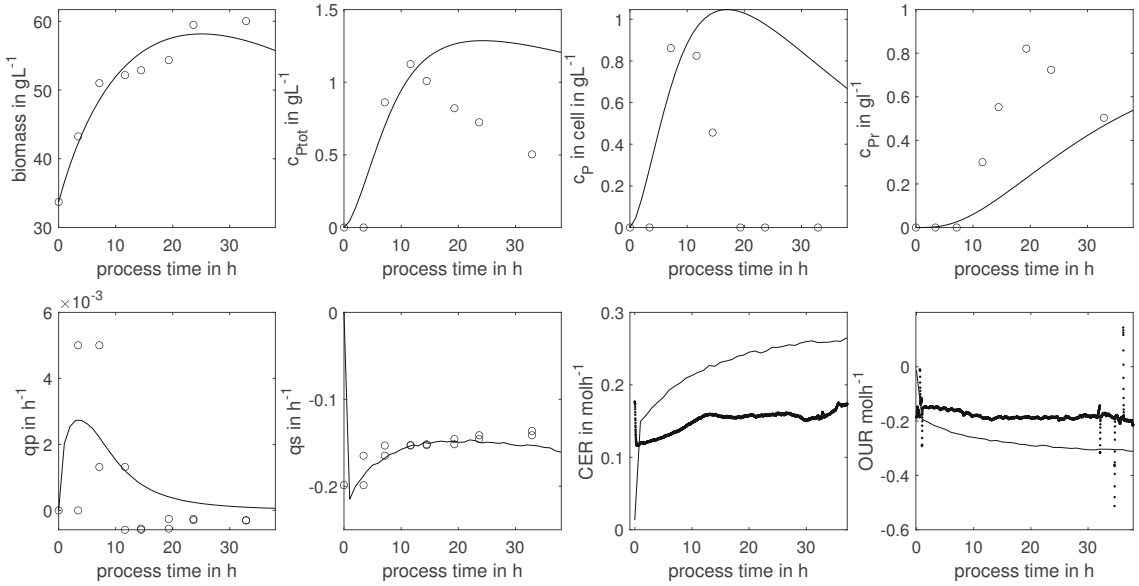

Figure 4.6: Experiment with constant feed rate and medium set point: comparison between experimental data and model estimation, illustrated in pink color in figure 4.1

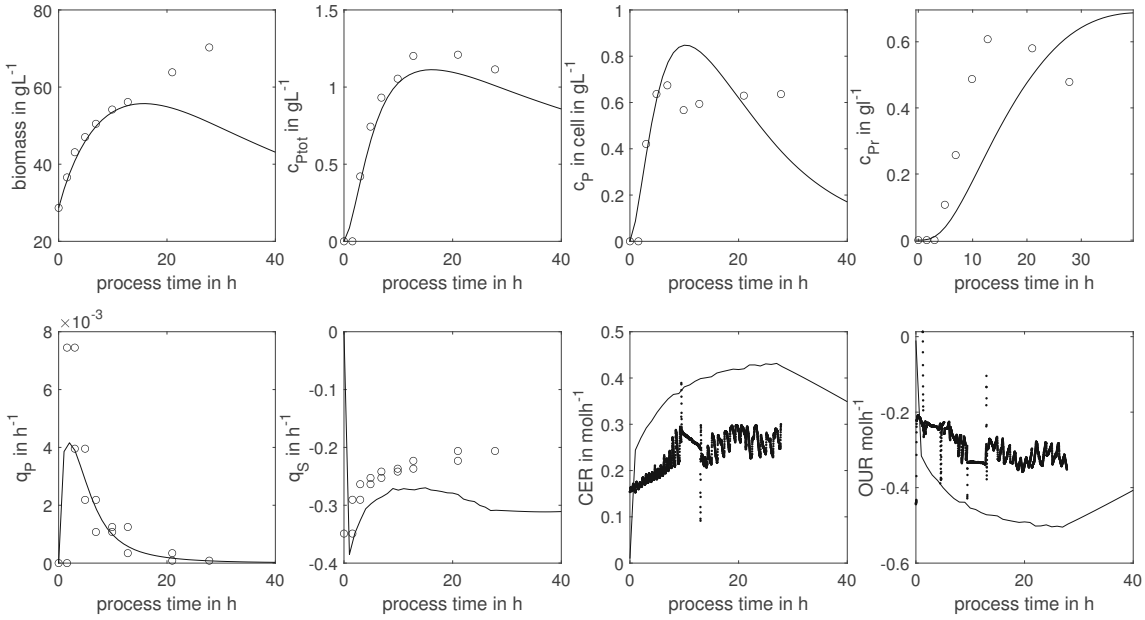

Figure 4.7: Experiment with constant feed rate and high set point: comparison between experimental data and model estimation, illustrated in cyan color in figure 4.1

In figure 4.8 the deviations between the measurements and the model predictions for important process variables are illustrated. The experiments used for this evaluation are those on which the model is based. The calculated error (RMSE) of the biomass concentration X, the concentration of the product inside P and outside P_R the cells and the total product P_{tot} is presented. Also shown is the RMSE of important and later controlled rates such as growth rate μ, substrate uptake rate q_S, product formation rate q_P and the respiratory rates OUR and CER. The overall error for every variable of all experiments can also be derived from the plot. Since the control laws are eventually to be evaluated in terms of their influence on productivity, it is important that the model provides good forecasts of the productivity behavior.

The errors on biomass and product concentration as well as inside the cells and outside are lower than 15%, the error on the total product is beneath 10%. Especially with regard to product formation and the very closely related biomass development, the model shows relatively small errors and should thus serve its purpose. The errors on the specific rates are acceptable with around 25%. The growth rate of the experiments with high and mid increasing feed deviate significantly from the model which leads to an overall error of 28% on the growth rate. The carbon dioxide emission is described well by the model with an overall error of 15.2%. The error on the oxygen uptake rate is higher because oxygen was dynamically added to the air during the processes.

The accuracy and level of detail of a model are driven to a large extent by the model objective, i.e. what the model is usually intended to do. In process control, dynamic models that capture the time courses of the variables properly are needed; they do not require very accurate models [25], since deviations from the behavior are taken into account in the process control and are compensated by the controller.

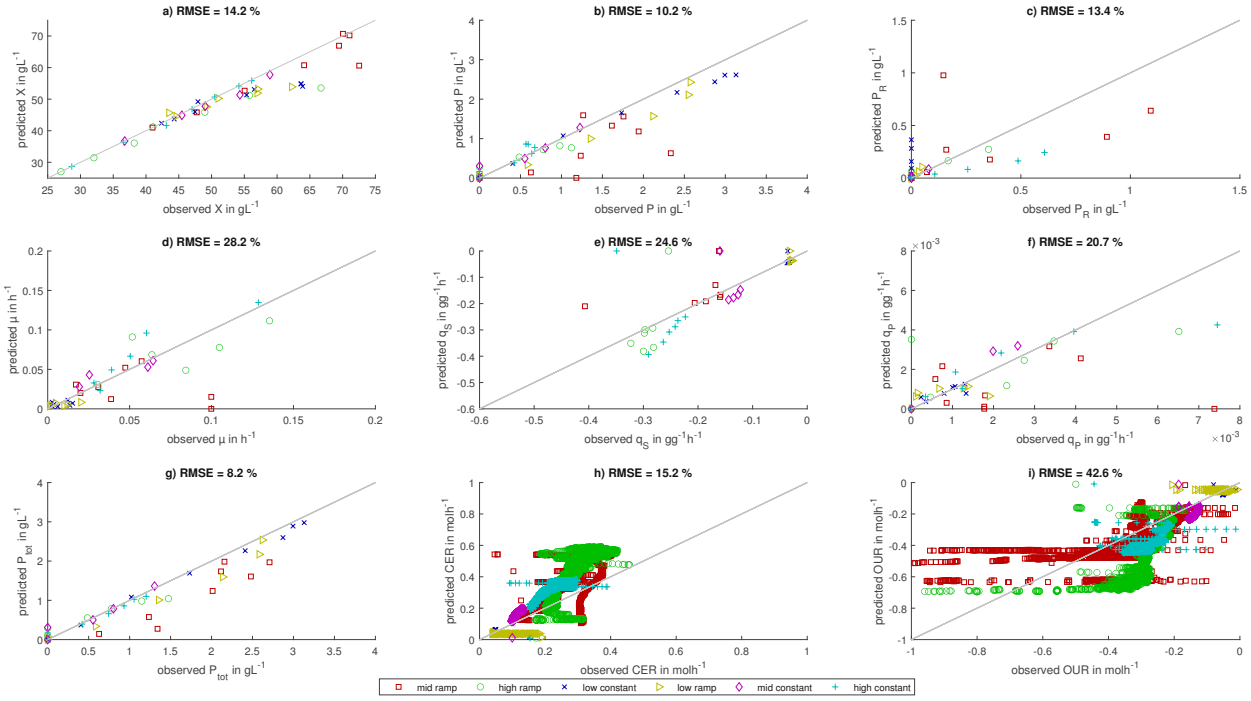

Figure 4.8: Modelfit of all experiments: shown are the deviations between the overall model prediction and experimental data for the biomass (X), the product concentration (P_{tot}), in cell (P) and released product (P_R), for the specific growth rate (μ), the substrate uptake rate (q_S), the product formation rate (q_P), and the respiratory rates (q_{O_2} and q_{CO_2})

The parameters of the model were estimated as described in 3.3. Instead of a Nelder-Mead simplex algorithm mentioned on page 12, parameters were obtained using (fmincon, MATLAB) an other simplex optimization algorithm with estimation boundaries.

Table 4.1 shows the obtained parameter values for the model and the experimental data. It can be seen that the parameters of the individually fitted experiments do not deviate strongly from the overall model. The model parameters usually result as an average value between the parameters of the experiments with low feed rates and those with high feed rates, as it is the case for $q_{P,max}$. This parameter is higher for lower feed settings and lower for the fast processes with the high feeds.

Table 4.1: Parameters and feed settings for the model and the six experiments (high constant (hc), high ramp (hr), mid constant (mc), mid ramp (mr), low constant (lc), low ramp (lr)). Model parameters were investigated for every data set separately and the whole data set was used to fit the model.

	Model	**hc**	**hr**	**mc**	**mr**	**lc**	**lr**	**Dim**
$\dot{V}_{in}$	–	0.015	hr	0.01	mr	0.002	lr	Lh^{-1}
$c_{S,in}$	850	931	900	841	796	853	853	gL^{-1}
$q_{S,max}$	1	1	1	1	1	1	1	$gg^{-1}h^{-1}$
$q_{P,max}$	0.008	0.00794	0.00766	0.00720	0.0088	0.00869	0.00867	$gg^{-1}h^{-1}$
K_S	0.0050	0.0050	0.0050	0.0050	0.0050	0.0050	0.0050	gL^{-1}
M	0.02	0.018098	0.018	0.020617	0.018098	0.021981	0.020147	h^{-1}
$Y_{X/S,max}$	0.47	0.47	0.47	0.47	0.47	0.47	0.47	gg^{-1}
$K_{Y_{X/S}}$	0.39525	0.42263	0.35575	0.43474	0.43477	0.35766	0.35579	gg^{-1}
$K_{S_{q_S}}$	0.1591	0.16317	0.16654	0.17501	0.14319	0.14874	0.14688	$gg^{-1}h^{-1}$
Hal	3.833	3.9087	3.8051	4.2161	3.5313	4.2117	4.0057	–
$K_{I_{q_P}}$	6.9	6.8991	7.0742	6.21	7.3056	6.6868	6.9710	–
$K_{S_{q_P}}$	0.19513	0.17879	0.21464	0.20769	0.17563	0.19048	0.17562	gg^{-1}
$K_{q_{S,rel}}$	0.18898	0.20786	0.17009	0.20765	0.20788	0.17014	0.17008	h^{-1}
$K_{S_{P_S}}$	2.2924	2.0639	2.5216	2.0744	2.0632	2.5153	2.5104	–

4.2 Controller development

4.2.1 Feedback linearization

In order to feedback linearize the state space model, presented in Chapter 3.1.2, can be written according to equation 3.16 in chapter 3.4.2 using the state vector $\boldsymbol{x} = [V_R\ c_X\ c_S\ c_P\ c_{P_R}\ S_{met}]^T$ with the associated differential equations given in 3.1.2 in the following form

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \end{bmatrix} = \begin{bmatrix} 0 \\ \mu\, x_2 \\ q_S\, x_2 \\ q_P\, x_2 - q_{P_R}\, x_4 \\ q_{P_R}\, x_4 \\ q_S \end{bmatrix} + \begin{bmatrix} 1 \\ \frac{c_{X,in} - x_2}{x_1} \\ \frac{c_{S,in} - x_3}{x_1} \\ \frac{x_4}{x_1} \\ \frac{x_5}{x_1} \\ 0 \end{bmatrix} u \tag{4.1}$$

The biomass-specific rates $(\mu, q_S, q_P, q_{O2}, q_{CO2})$ are intended to be controlled, therefore the respective rate is selected as the output variable of the control law. In the following, the procedure is described as an example for the substrate uptake rate (q_S), therefore

$y = q_S$. The flow rate into the reactor $\dot{V}_{in}$ is the controlled variable u. Inserting the kinetic correlations that are outlined in 3.1.1 the system can be written as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \end{bmatrix} = \begin{bmatrix} 0 \\ Y_{X/S,max} e^{\left(-x_6 K_{Y_{X/S}}\right)} \left(\frac{q_{S,max}\, x_3}{x_3+K_S} - M\right) x_2 \\ \frac{q_{S,max}\, x_3}{x_3+K_S}\, x_2 \\ \frac{q_{P,max}\, \frac{q_{S,max}\, x_3}{x_3+K_S}}{q_S+K_{S_{q_S}}} \frac{x_6}{\frac{x_6^{Hal}}{K_{I_{q_P}}}+x_6+K_{S_{q_P}}} x_2 - K_{q_s,rel} \left(\frac{q_{S,max}\, x_3}{x_3+K_S} - M\right) \left(1 - e^{\frac{x_6}{K_{S_{P_S}}}}\right) x_4 \\ K_{q_S,rel} \left(\frac{q_{S,max}\, x_3}{x_3+K_S} - M\right) \left(1 - e^{\frac{x_6}{K_{S_{P_S}}}}\right) x_4 \\ \frac{q_{S,max}\, x_3}{x_3+K_S} \end{bmatrix} + \begin{bmatrix} 1 \\ \frac{c_{X,in}-x_2}{x_1} \\ \frac{c_{S,in}-x_3}{x_1} \\ \frac{x_4}{x_1} \\ \frac{x_5}{x_1} \\ 0 \end{bmatrix} u \tag{4.2}$$

Now the Lie-derivative along $\boldsymbol{g}(\boldsymbol{x})$ of $h(\boldsymbol{x}) = y = q_S$ can be calculated to determine the relative degree of the system.

$$L_{\boldsymbol{g}} h(\boldsymbol{x}) = \frac{\partial h(\boldsymbol{x})}{\partial \boldsymbol{x}}\, g(\boldsymbol{x}) = [0\ 0\ \frac{q_{S,max}\, K_S}{(x_3+K_S)^2}\ 0\ 0\ 0] \begin{bmatrix} \mathbf{1} \\ \frac{\boldsymbol{c_{X,in}-x_2}}{\boldsymbol{x_1}} \\ \frac{\boldsymbol{c_{S,in}-x_3}}{\boldsymbol{x_1}} \\ \frac{\boldsymbol{x_4}}{\boldsymbol{x_1}} \\ \frac{\boldsymbol{x_5}}{\boldsymbol{x_1}} \\ \mathbf{0} \end{bmatrix} \neq 0 \tag{4.3}$$

Because $L_{\boldsymbol{g}} h(\boldsymbol{x}) \neq 0$ the relative degree of the system $\delta = 1$. With the formula 3.24 given in 3.4.2 the law for the controlled variable can be determined as follows

$$u = -\frac{L_{\boldsymbol{f}}^1 h(\boldsymbol{x}) + \boldsymbol{k}^T z}{L_{\boldsymbol{g}} L_{\boldsymbol{f}}^0 h(\boldsymbol{x})} + \frac{K_w}{L_{\boldsymbol{g}} L_{\boldsymbol{f}}^0 h(\boldsymbol{x})}\, w = \frac{-L_{\boldsymbol{f}} h(\boldsymbol{x}) - \alpha_0\, q_S + \alpha_0\, w}{L_{\boldsymbol{g}} h(\boldsymbol{x})} \tag{4.4}$$

The control laws for every other specific rate can be obtained analogously. The relative degree is always $\delta = 1$ because $L_{\boldsymbol{g}} h(\boldsymbol{x}) \neq 0$ for all cases considered as can be seen in chapter 4.2.2.

4.2.2 Control laws

Due to 4.4 the nonlinear system can now be written in the following form of a simple proportional control law 4.5 where w describes the reference input (the setpoint) and q for the concerning controlled rate.

$$\dot{V}_{in} = \frac{\alpha_0 * (w - q) - L_f h(x)}{L_g h(x)} \tag{4.5}$$

The control law for the respective rate is thus fully specified by the defined setpoint, the behavior of the rate over time predicted by the model and the associated Lie derivatives.

The proportional term α_0 is obtained by simulating an experiment several times and is thus also determined for each rate separately. The Lie derivatives of the various control laws are given in equation 4.6, 4.7, 4.8, 4.9 and 4.10.

Substrate rate q_S

$$\begin{aligned} L_f h(\boldsymbol{x}) &= -q_S c_X \frac{q_{Smax} - q_S}{K_S + c_S} \\ L_g h(\boldsymbol{x}) &= \frac{(c_{S,in} - c_S)}{V_R} \frac{(q_{Smax} - q_S)}{K_S + c_S} \end{aligned} \tag{4.6}$$

Growth rate

$$\begin{aligned} L_f h(\boldsymbol{x}) &= q_S \left(\mu \, K_{Y_{X/S}} - Y_{X/S} \, q_S \, c_X \left(1 - \frac{1}{K_S + c_S} \right) \right) \\ L_g h(\boldsymbol{x}) &= \frac{(c_{S,in} - c_S)}{V_R} \frac{Y_{X/S} \, q_{Smax} \, K_S}{(K_S + c_S)^2} \end{aligned} \tag{4.7}$$

Product formation rate

$$\begin{aligned} L_f h(\boldsymbol{x}) &= q_S \, q_P \, (b_1 \, c_X + b_2) \\ L_g h(\boldsymbol{x}) &= \frac{(c_{S,in} - c_S)}{V_R} b_1 \, q_P \\ b_1 &= \frac{1}{K_S + c_S} \left(\frac{q_{Smax} - q_S}{K_{Sqs} + q_S} - \frac{K_S}{c_S} \right) \\ b_2 &= \frac{K_S \, q_S - \frac{(Hal-1) S_{met}^{Hal}}{K_{I_{q_P}}}}{S_{met} \left(K_S \, qp + S_{met} + \frac{S_{met}^{Hal}}{K_{I_{q_P}}} \right)} \end{aligned} \tag{4.8}$$

Oxygen uptake rate

$$
\begin{aligned}
L_f h(\boldsymbol{x}) &= \frac{q_S}{4}(c_X\, a_1 + \mu\, a_2) \\
L_g h(\boldsymbol{x}) &= \frac{(c_{S,in} - c_S)}{V_R}\, a_1 \\
a_1 &= \frac{\gamma_X\, Y_{X/S} \frac{q_{S,max} - q_S}{K_S + c_S}}{M_X} + \frac{\gamma_S\,(q_S - q_{S,max})}{M_S\,(K_S + c_S)} \\
a_2 &= \frac{\gamma_X\, K_{Y_{X/S}}}{M_X}
\end{aligned}
\tag{4.9}
$$

Carbon dioxide emission rate

$$
\begin{aligned}
L_f h(\boldsymbol{x}) &= -q_S\, A - q_S\, c_X\, (B + q_P\, C) \\
L_g h(\boldsymbol{x}) &= \frac{(c_{S,in} - c_S)}{V_R}(B + q_P\, C) \\
A &= -\frac{\mu\, K_{Y_{X/S}}}{M_X} + \frac{q_P}{M_P}\left(\frac{1}{Smet} - \frac{\frac{Hal * S_{met}^{(Hal-1)}}{K_I qp} + 1}{K_{Sqp} + S_{met} + S_{met}^{\frac{Hal}{K_I qp}}}\right) \\
B &= \frac{q_{Smax} - q_S}{K_S + c_S} * (\frac{1}{M_S} - \frac{Y_{X/S}}{M_X}) \\
C &= \frac{1}{M_P(K_S + c_S)} - \frac{1}{M_P c_S} + \frac{\frac{q_{Smax} - q_S}{K_S + c_S}}{M_P(K_{Sqs} + q_S)}
\end{aligned}
\tag{4.10}
$$

4.2.3 Control circuits

The proportional controller described in equation 4.5 showed a simple proportional control law. First we use the model calculated rate q^* as a feedback to generate a feed trajectory to keep the desired rate at the predefined setpoint w. This resulting feed trajectory will further be used to control the rates of a virtual process and due to the model based character will further be referenced as feedforward controller and is shown in equation 4.11. The control loop 4.9 shown below describes why this controller could be interpreted as an feedforward controller. The feedback linearized control law referred to as feedforward controller in the graph compares the desired setpoint w with the output of the parallel running model q^* and thus counteracts disturbances that could be described by the model.

$$
\dot{V}_{in} = \frac{\alpha_0 * (w - q^*) - L_f h(x)}{L_g h(x)} \tag{4.11}
$$

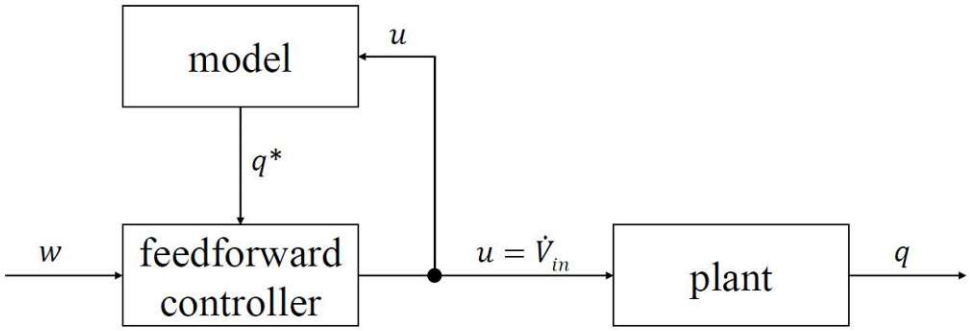

Figure 4.9: Feedforward control circuit: cf. Fig. 3.4, the feedback term $H(s)$ is not applicable, the controller just incorporates the feedforward compensator $C_f(s)$

In application the simulation will most likely not be able to compensate for all disturbances of the real process. To compensate mismatches from the model, the feedforward control law above will now be extended by a feedback term. According to [31] bioprocess control loops often use PI controllers. Therefore, a PI term is also used here for the feedback part. The resulting law compares the setpoint with the model $(w - q)$ and the output of the model with that of the real process $(q^* - q)$. The control law thus represents a two degrees of freedom controller shown in 4.10 and described in chapter 3.4.3.

$$\dot{V}_{in} = \frac{\alpha_0 * (w - q^*) + \alpha_1 \int_0^t (q^* - q)dt + \alpha_2(q^* - q) - L_f h(x)}{L_g h(x)} \tag{4.12}$$

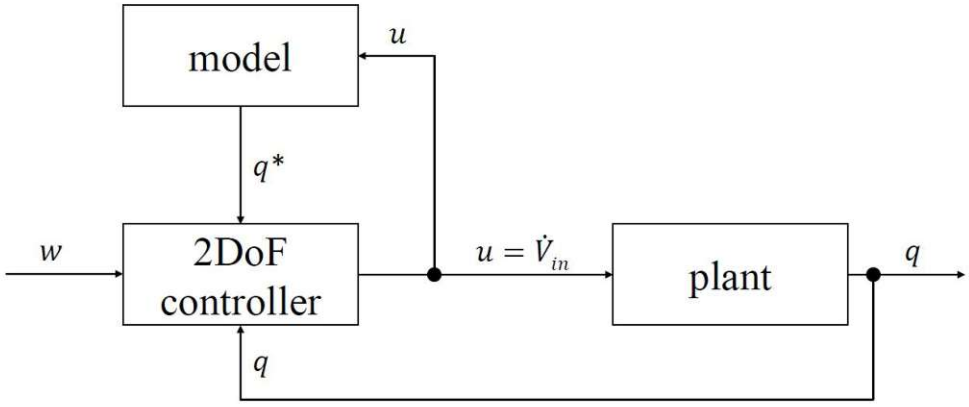

Figure 4.10: Feedback contol circuit: cf. Fig. 3.4 and Eq. 4.12 the feedforward part compares the mismatch of the model prediction and the set point $(w-q^*)$ and the feedback part incorporates the mismatch of the model and the real plant behavior $(q^* - q)$

With these control laws, the individual rates can now be kept constant. The resulting substrate feed profiles are shown in figure 4.18. The feedforward control loops will be used to find optimal setpoints, the 2DoF controllers in the following referred to as feedback controllers are calibrated and used to control an independent validation plant to assess the applicability of the different control approaches.
In table 4.2 the tuning parameters α_0 for the feedforward part, α_1 and α_2 for the feedback part of the control laws are shown for each controlled rate.

Table 4.2: Control law tuning parameters

controlled rate	$\boldsymbol{\alpha_0}$	$\boldsymbol{\alpha_1}$	$\boldsymbol{\alpha_2}$
$\dot{V}_{in}$	20	10	7
q_S	1	100	5
μ	2000	1000	7
q_P	3500	500	20
q_{O_2}	6000	100	0.01
q_{CO_2}	2500	70	60

4.3 Setpoint optimization

In order to apply the obtained control laws to the process in a meaningful way, to compare them and to analyze the applicability, setpoints have to be defined. The setpoints should be chosen in a way that the processes can run physiologically stable. Since the control laws will be compared with the original process runs in terms of their performance, setpoints should also be chosen such that the predicted productivities or product yields are optimal. Typically, feasible setpoints for bioprocesses are determined by trial and error in pilot plants, which is a slow and labor-intensive process. Modeling can not only help develop optimal control strategies through better process understanding, but can also be used to generate better process setpoints or even optimal setpoint trajectories.[31]
In order to generate comparable setpoints of the different control laws for the implementation study, a setpoint optimization was performed.
In fed-batch processes, the down times for cleaning, preparation and sterilization of a new batch, play a decisive role with regard to the overall yield of the process. In the industry, the focus is therefore often on optimizing the space-time yield in order to generate more effective processes.[7]
Therefore, for each control objective two setpoints were defined using the optimal space-time yield and the maximal achievable product concentration predicted as performance criteria. First the setpoints with the highest space-time yield were determined.

4.3.1 Setpoints for optimal space-time yield

The process down-times were estimated. In [7] the duration of sterilization in place (SIP) and cleaning in place (CIP) are given as 3 and 6 hours respectively, the duration of the batch phase as 6h and the non-induced fed-batch as 8h. In this work, the batch process is considered to take 8h and the non-induced fed-batch 12h, these estimates are based on the experimental data used. For set-up, dismantling and cleaning, 16 hours were taken into account, resulting in a non-productive process time of 34 hours to be considered in the calculation of the space-time yield.
Perfect model assumption and the feedforward control circuit were used to simulate the process outcomes.
Graph 4.11, 4.12, 4.13,4.14, 4.15 and 4.16 show the contour plots for every constant specific rate used as a control objective. The optimal space-time yield is indicated by an x. The space-time yield is given for the combination of different constant setpoints and the respective harvest points in time.

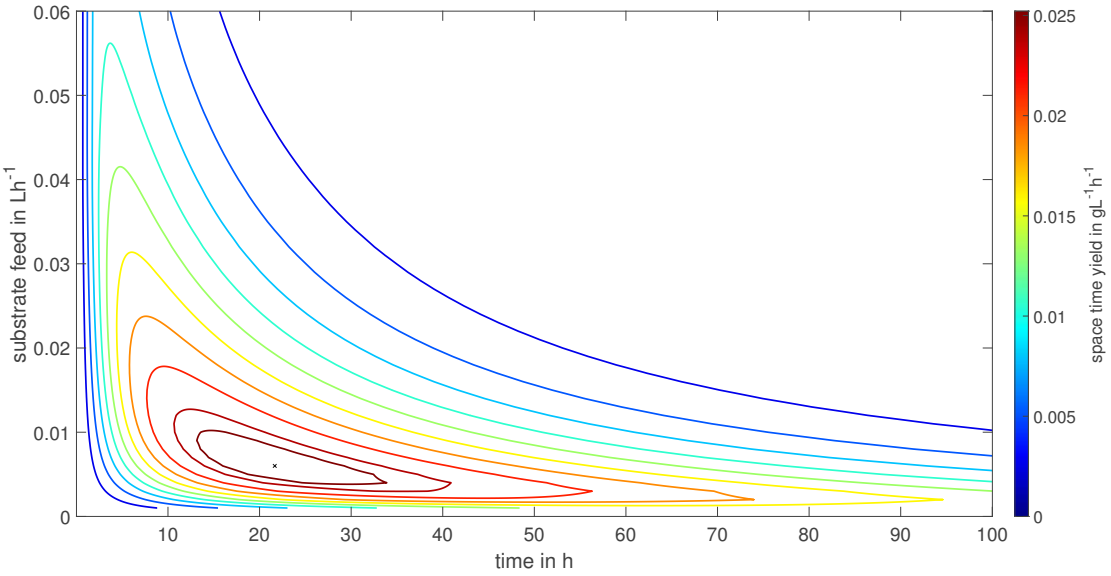

Figure 4.11: Setpoint optimization constant substrate feed and the harvest time (time): the first contour line shows the limits in which 95% of the maximum possible space-time yield can be achieved, the second the 90% limit and then in 10% steps downwards.

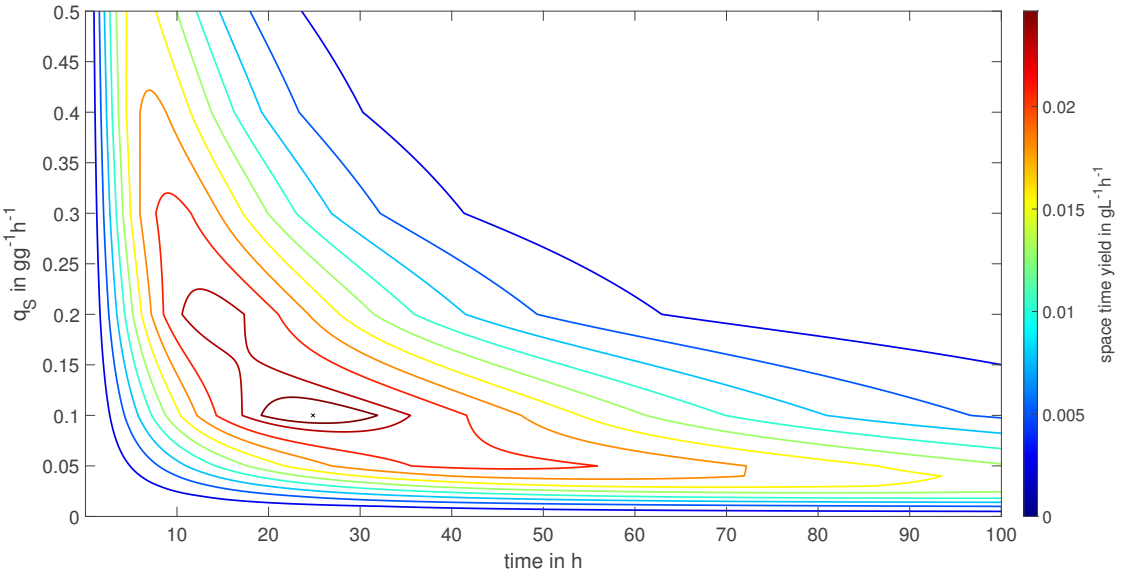

Figure 4.12: Setpoint optimization constant q_S: the first contour line shows the limits in which 95% of the maximum possible space-time yield can be achieved, the second the 90% limit and then in 10% steps downwards.

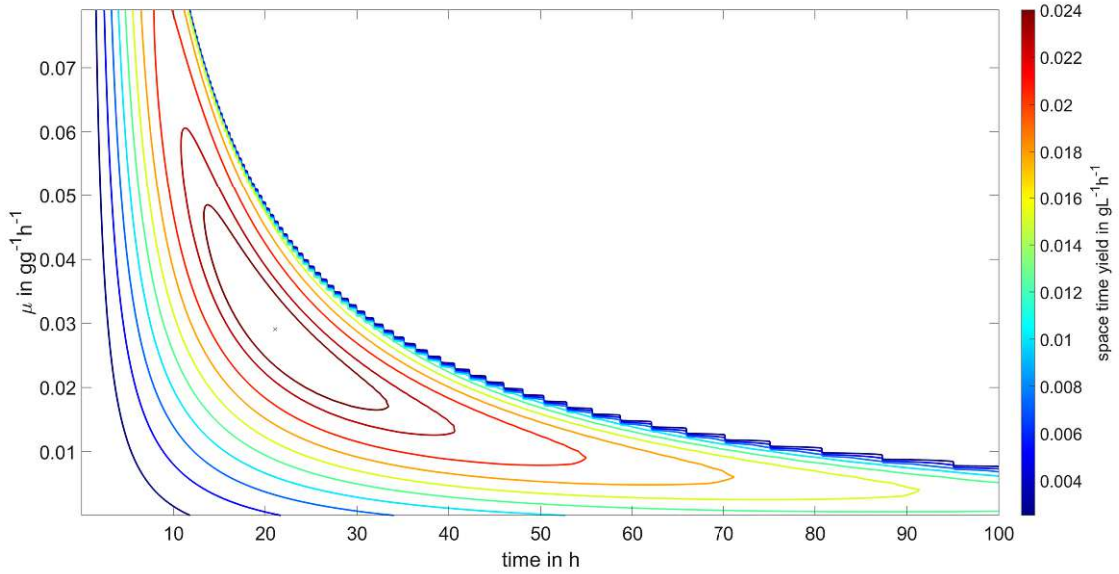

Figure 4.13: Setpoint optimization constant μ: the first contour line shows the limits in which 95% of the maximum possible space-time yield can be achieved, the second the 90% limit and then in 10% steps downwards.

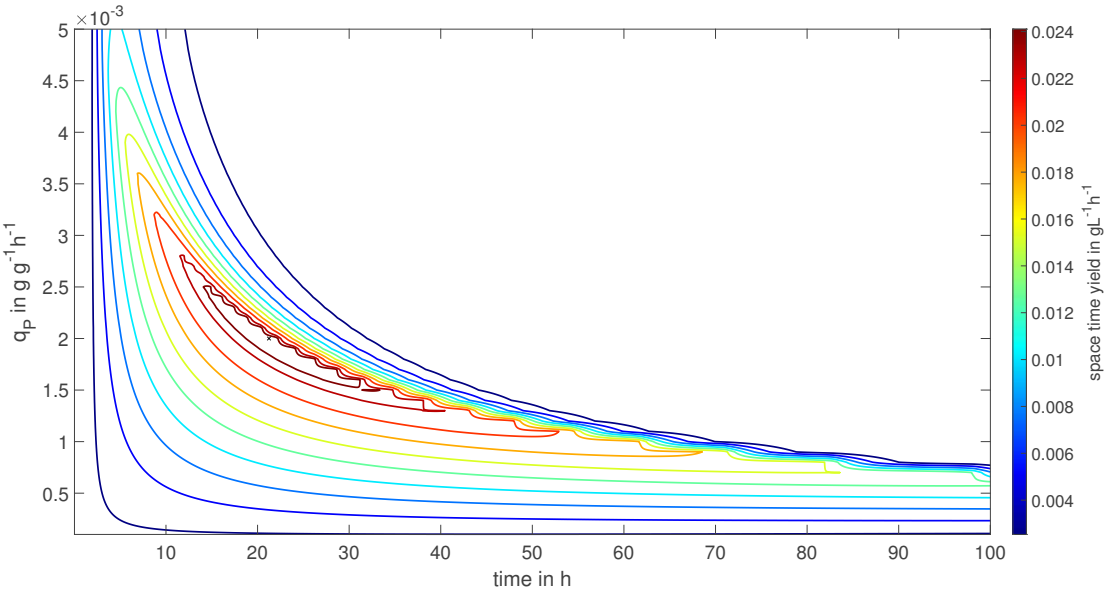

Figure 4.14: Setpoint optimization constant q_p: the first contour line shows the limits in which 95% of the maximum possible space-time yield can be achieved, the second the 90% limit and then in 10% steps downwards.

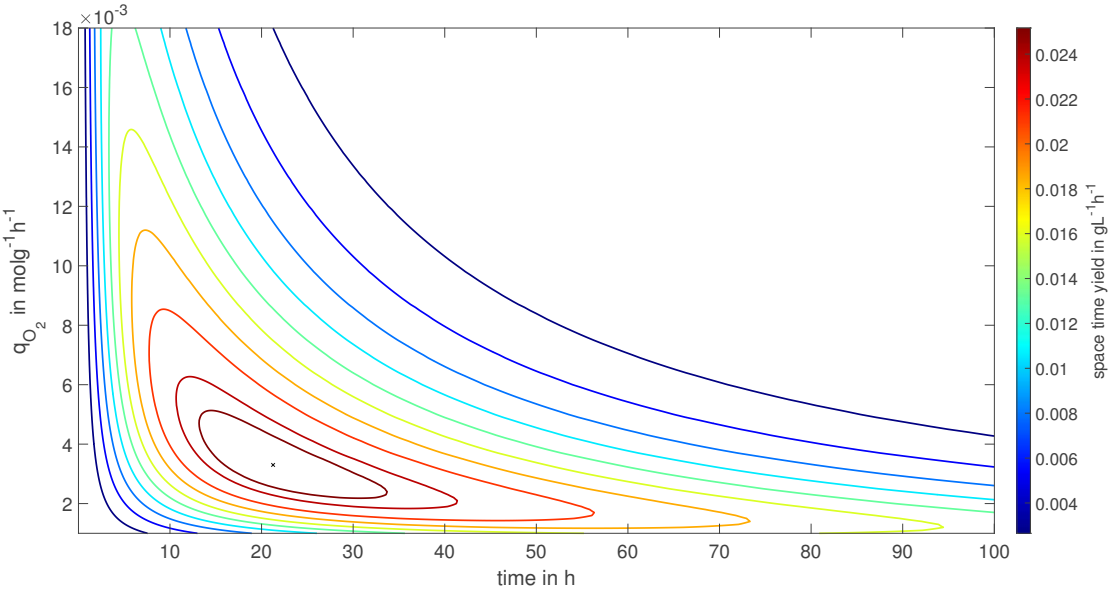

Figure 4.15: Setpoint optimization constant q_{O_2}: the first contour line shows the limits in which 95% of the maximum possible space-time yield can be achieved, the second the 90% limit and then in 10% steps downwards.

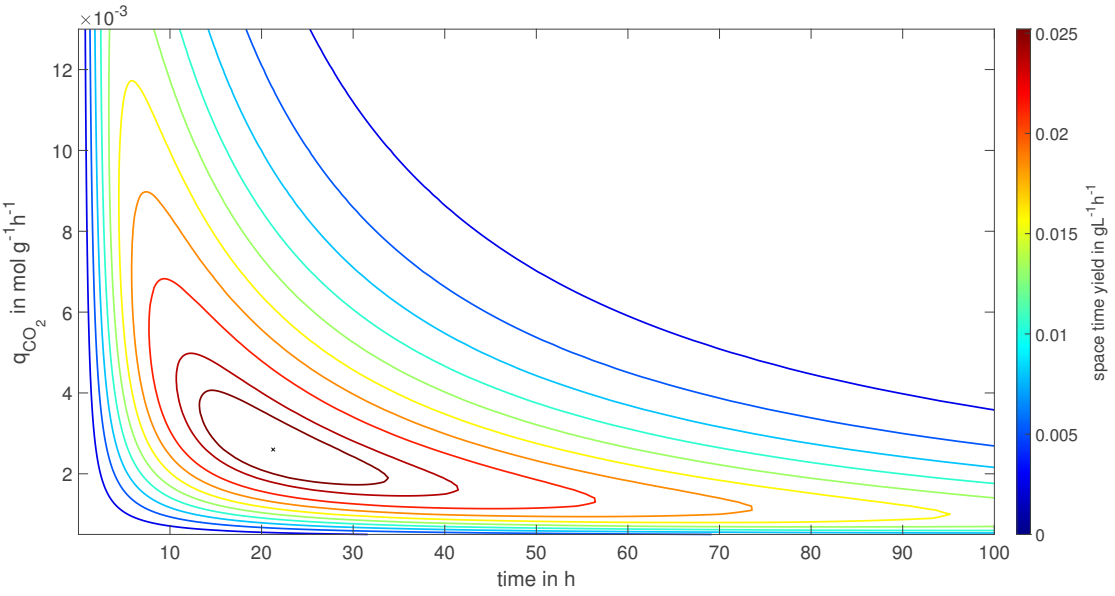

Figure 4.16: setpoint optimization constant q_{CO_2}: the first contour line shows the limits in which 95% of the maximum possible space-time yield can be achieved, the second the 90% limit and then in 10% steps downwards.

The optimal setpoints can be derived from the contour plots, which show the relationship between harvest time, rate setpoint and space-time yield. The first dark red contour line, which describes the range in which 95% of the highest possible space-time yields are possible, shows that different levels of setpoint achieve good results in terms of time dependent product yield. The optimal space-time yield occurs at about the same time for all the control laws examined. This is due to the fact that the preparation and post-processing times of the process have a significant influence on the space-time yield and are the same for all control laws. The yields are also comparable, whereby the direct control of the feed rate together with the control of the oxygen uptake rate performs best. The respiratory rates have the advantage that the time window for harvesting is the largest at around 15 hours. Maintaining the sweet spot of the product formation rate as a control variable turns out to be the most difficult to ensure.

4.3.2 Setpoints for maximum product concentration

The virtual experiments implemented later should behave differently from the model due to their parameterisation. This is to investigate the applicability of the model. In order to be able to compare the processes and make possible forecasts about the improvement of the processes, they are evaluated with regard to the maximum productivity. For this reason, in addition to the setpoints for the optimal time-space yields, the setpoints at which the maximum possible product concentration could be reached during the process were also identified. As could be seen in figure 4.17 very low setpoints lead to a higher maximum product concentrations. This correlation is valid for every examined rate. Therefore the chosen setpoints do not just serve the selected process goals, they also represent the boundaries of the feedrate that lead to reasonable processes.

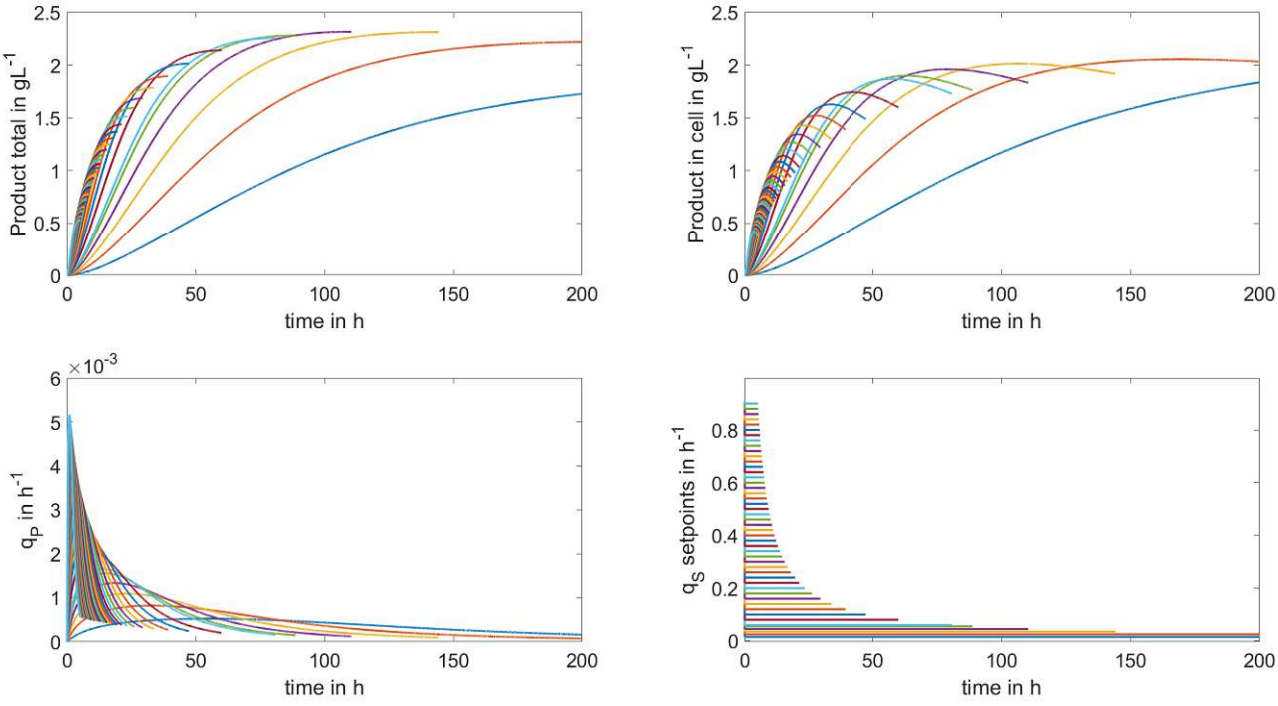

Figure 4.17: setpoint optimization to reach maximum possible product concentration during the process shown for the specific substrate uptake rate q_S as an example: it can be seen that lower setpoints and slower processes achieve the highest product concentrations

4.3.3 Selected setpoints

The optimal setpoints identified for all investigated control laws are listed in table 4.3. As discussed before, very low setpoints lead to the best overall product output. Setpoints to achieve the maximum space-time yield are way higher. These setpoints assure processes with harvest times between 21 hours (for the controlled growth rate) and 25 hours (for the substrate uptake rate control) from the beginning of the induction process. The processes run on the lower setpoints to maximize the product concentration take between 200 hours and up to nearly 400 hours.

Table 4.3: Setpoint values for every rate used as a control objective to achieve maximum space-time yield or maximum product concentration respectively. The setpoints for the maximum space-time yield can be read out from the counter plots in chapter 4.3.1. The optimal setpoints to generate the maximum product concentration were obtained via simulations of the process as shown in the example for q_S as control variable in 4.17.

	maximum space-time yield	**maximum c_P**	**Dim**
$\dot{V}_{in}$	0.006	0.001	Lh^{-1}
q_S	0.1	0.035	$gg^{-1}h^{-1}$
μ	0.02905	0.00005	$gg^{-1}h^{-1}$
q_P	0.002	0.0007	$gg^{-1}h^{-1}$
q_{O_2}	0.003	0.0012	$molg^{-1}h^{-1}$
q_{CO_2}	0.0026	0.0009	$molg^{-1}h^{-1}$

4.4 Feed profiles of control laws

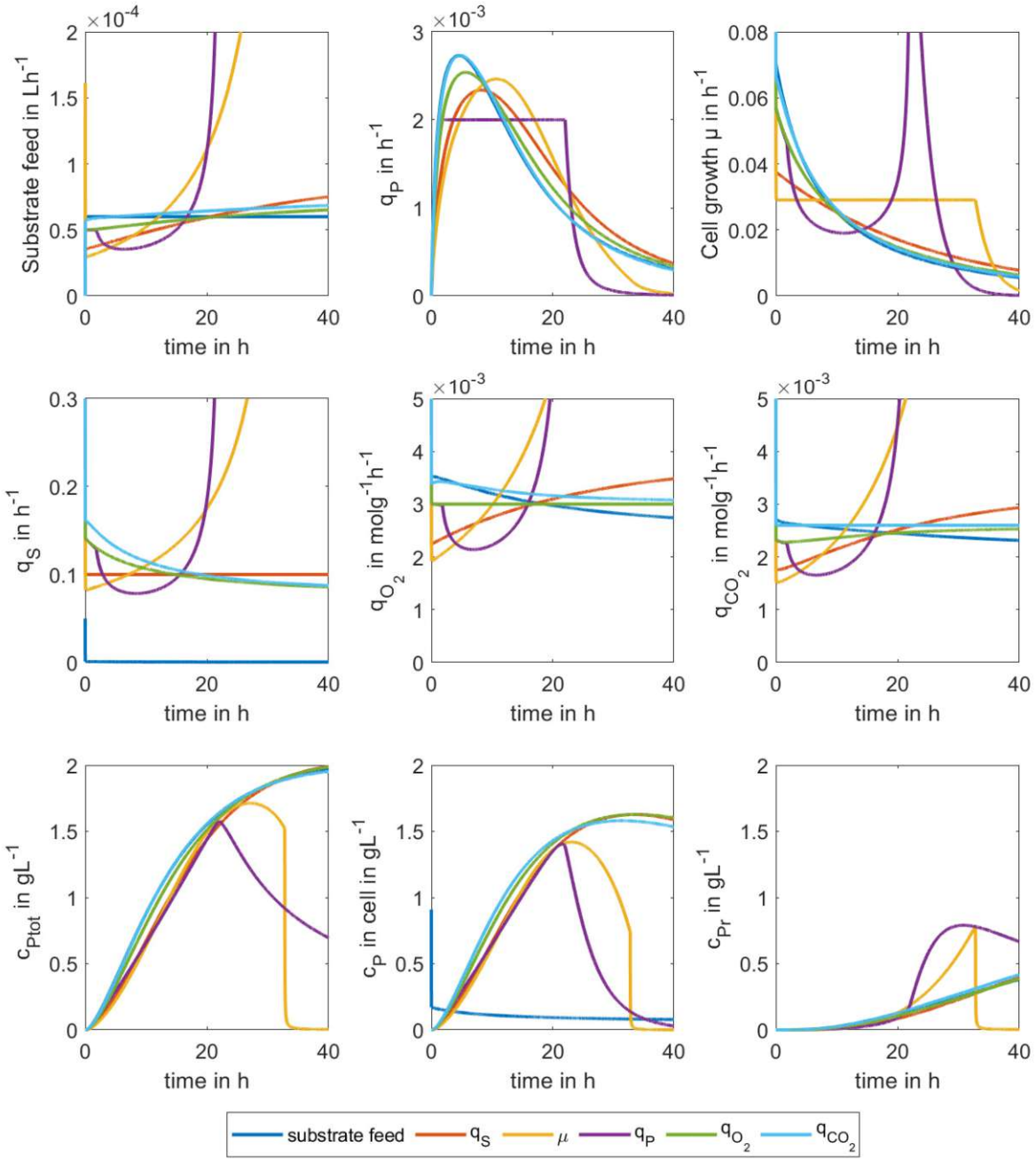

Figure 4.18: Feedprofiles of different control laws as well as courses of product formation rate q_P, growth rate μ and product concentration c_{Ptot}, c_P and c_{Pr} over time.

The different specific rates that are kept constant result in different feeding profiles that have to be supplied to the process in order to maintain the desired setpoint. This causes changes in the process behavior and productivity. With the setpoints defined in the previous chapter where the maximum space-time yield occurred (See Table 4.3), an feedforward simulated to illustrate the different rates and process behavior was executed.

Graphic 4.18 shows the feed profiles associated with the respective specific rate as well as the product formation rate, the specific growth rate and the progress of the product concentration over the time course of the process.
Compared to the constant feeding rate, especially the profiles for the product formation rate and the growth rate deviate strongly. Considerable feed adjustments are needed to keep these two rates at a constant level. This was to be expected, as constant growth can only be guaranteed by an exponentially increasing feedrate. Since productivity is strongly related to the growth, it is only logical that a exponential feed is also required in this case to keep the rate at a constant level. For the specific substrate uptake rate, an approximately linear course of the feed rate results, which was also to be expected. For the respiration rates, flat feed curves result, which are quite similar.

4.5 Performance of control laws with model-plant mismatches

An virtual plant was created to evaluate the control performance. The plant behaves mathematically like the model used. The parameters of the plant were adapted to the experiment with low and constant feed from the data set and are shown in 4.1. Simulations were carried out with this experimental plant to tune the control laws for the desired setpoint.

As can be seen in graphs 4.19, 4.20, 4.21, 4.22 and 4.23 the model (shown in blue color) and the simulated plant (shown in orange for the feedback controlled and in yellow for the feedforward controlled system) deviate from each other. The deviation is referred to as model-plant mismatch and has impact on the productivity and the temporal evolvement of the process. The figures show the controller performances using the setpoints for the highest space-time yield. The time curves of the constant rates for both the open-loop control and the closed-loop control as well as the corresponding curves of the product concentration over time are displayed.

In blue, the model (as presented in chapter 4.1) behavior holding the simulated setpoint is shown. In orange the control performance of the feedback controller while simulating the virtual plant shows, that the setpoint could be held well on the desired (blue) value. The corresponding productivity is also shown in orange color. The plant was also simulated for using a feedforward controller (depicted in yellow color) and shows that it does not perform as well as the feedback controller in keeping the desired setpoint. However, since there is no time delay in the control at the beginning of the process, the feedforward controller performs slightly better than the feedback controller for the respiratory rates, since these seem to be less sensitive to setpoint deviations for the investigated process.

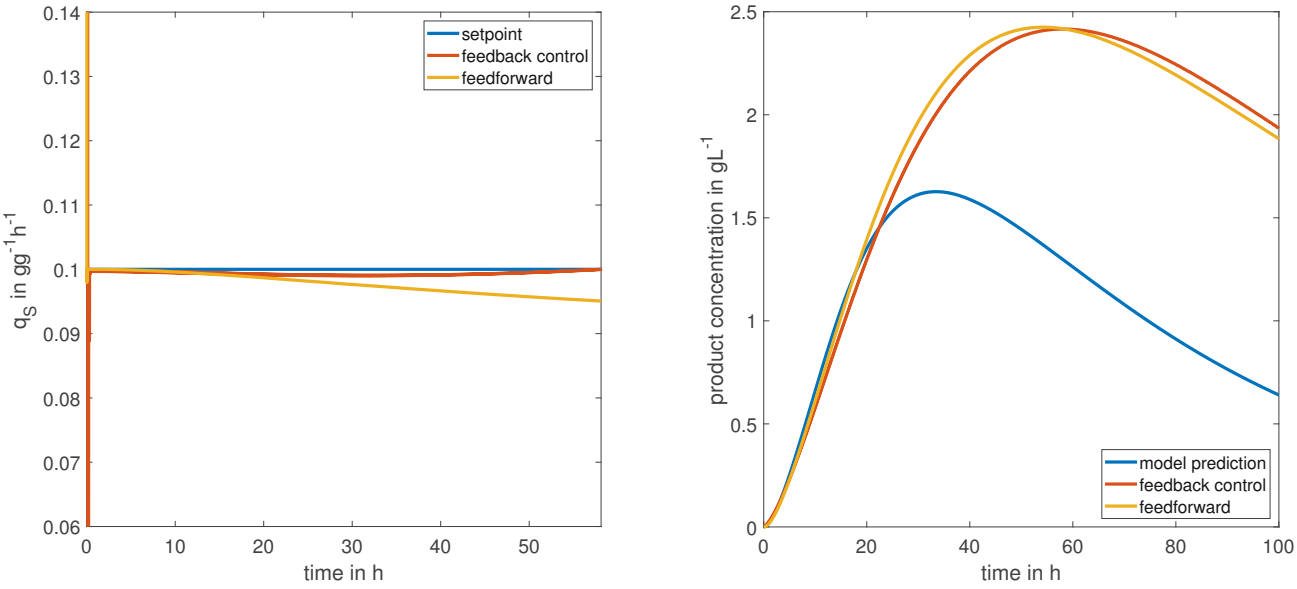

Figure 4.19: controller performance q_S: on the left hand side the course in time of the controlled rates for feedforward and feedback control is shown. On the right hand side the courses of the product formation are shown to assess the effects of deviations on the process performance.

In figure 4.19 on the left hand side it is shown how well the specified setpoint could be kept by the control laws. With the feedforward control law the desired setpoint could not be held constant during the whole process. The necessary adjustment to maintain the setpoint is achieved by the feedback control law which performed very well for the substrate rate. On the right hand side the course of the product concentration during the process is depicted. The product concentration is increasing faster at the beginning for the feedforward control law. This is due to the fact that the feedrate for the feedforward control is a bit higher at the beginning of the process. Because of the minor deviations between feedback and feedforward control similar product concentrations could be reached during the course of the process.

Unlike the control of the substrate uptake rate, the deviation between the feedforward and the feedback control law is significant for both the control of the specific growth rate 4.20 and the specific product formation rate 4.21. The feedforward control do not achieve constant rates. The control performance for both control approaches can therefore be improved

considerably. Feedback control achieves satisfactorily constant rates which additionally has a positive impact on the productivity.

The control performance of the specific oxygen uptake rate is shown in figure 4.22 and the behavior is similar to the substrate rate controller. The feedforward law does not fulfill the desired constant setpoint but the deviation could be compensated very well with the feedback controller. Again the increase of the product concentration is grater because of higher feedrates at the beginning of the investigated process which is also the case for the specific CO_2 emission rate. For this control objective the feed trajectory does not supply the right feedrate during the whole process to keep the setpoint. Feedback control works better but in the course of the process it could be seen that the controller is not able to compensate disturbances perfectly as the curve of the carbon dioxide emission rate exceeds the setpoint at the end of the process.

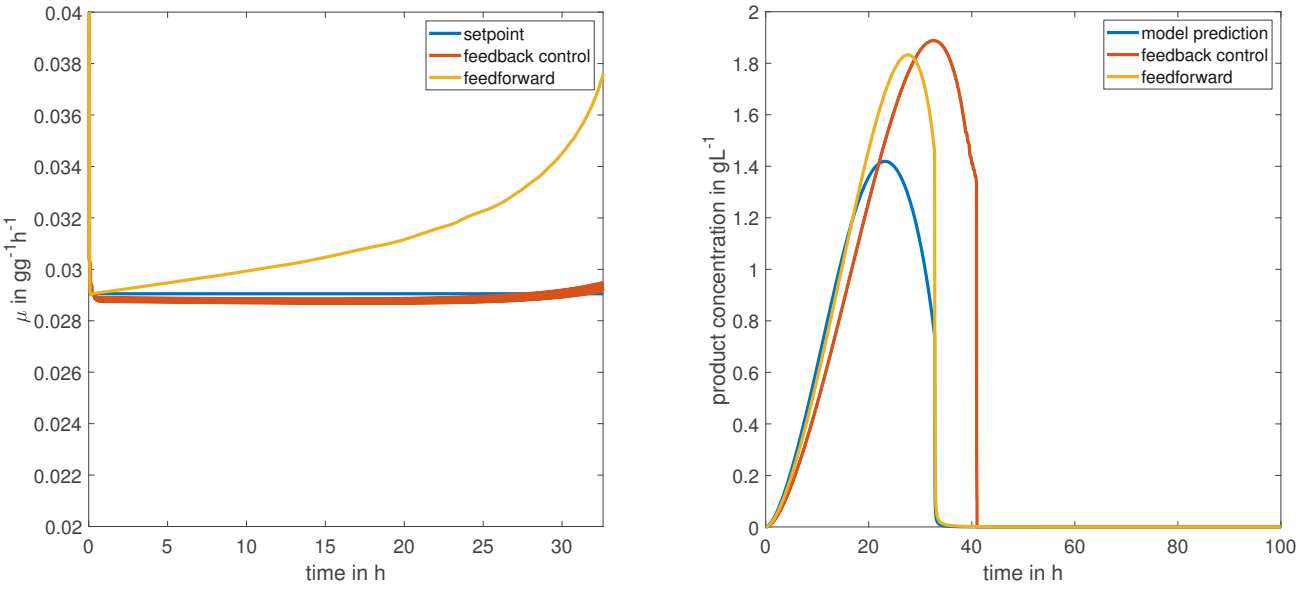

Figure 4.20: controller performance μ: on the left hand side the course in time of the controlled rates for feedforward and feedback control is shown. On the right hand side the courses of the product formation are shown to assess the effects of deviations on the process performance.

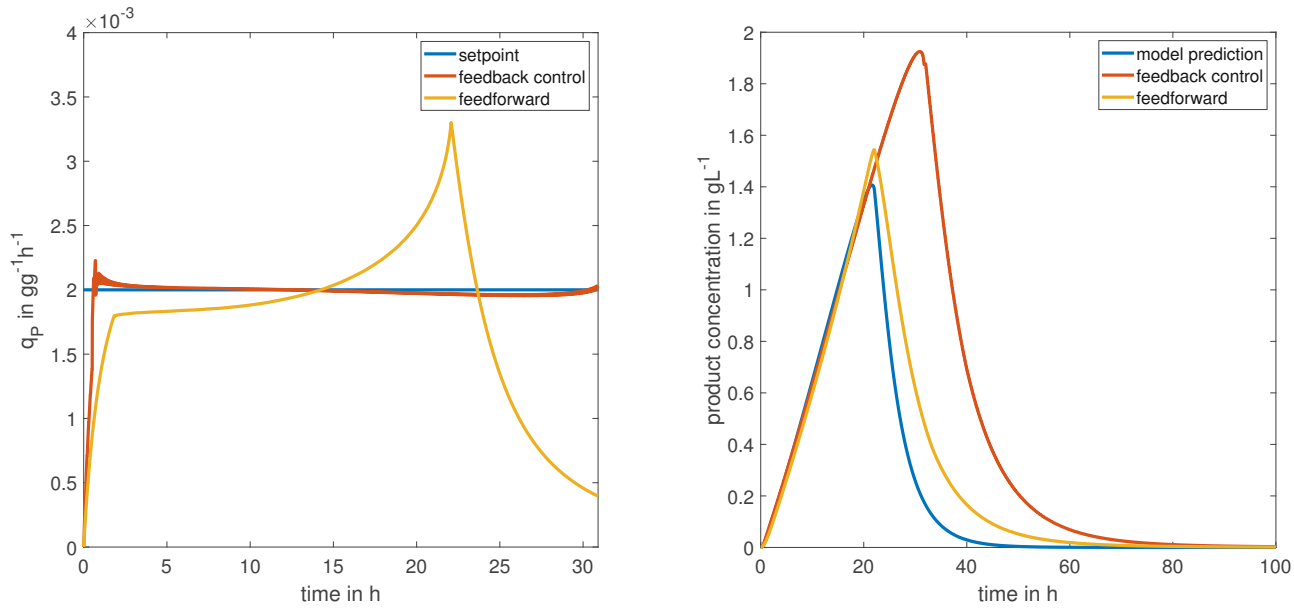

Figure 4.21: controller performance q_P: on the left hand side the course in time of the controlled rates for feedforward and feedback control is shown. On the right hand side the courses of the product formation are shown to assess the effects of deviations on the process performance.

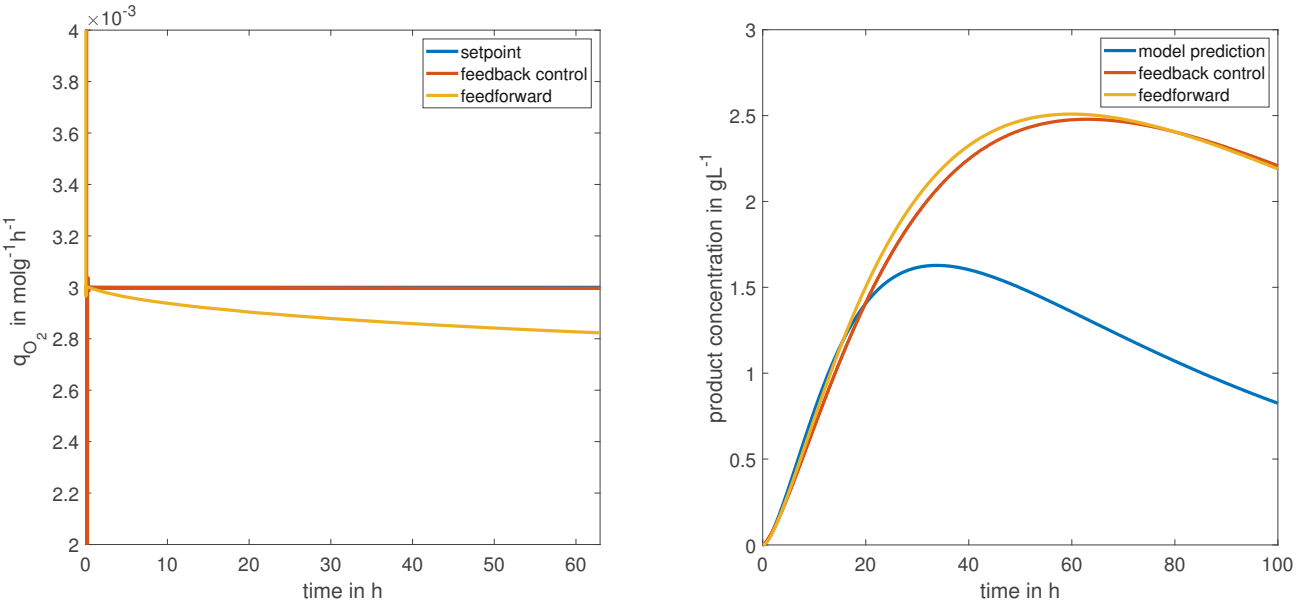

Figure 4.22: controller performance q_{O2}: on the left hand side the course in time of the controlled rates for feedforward and feedback control is shown. On the right hand side the courses of the product formation are shown to assess the effects of deviations on the process performance.

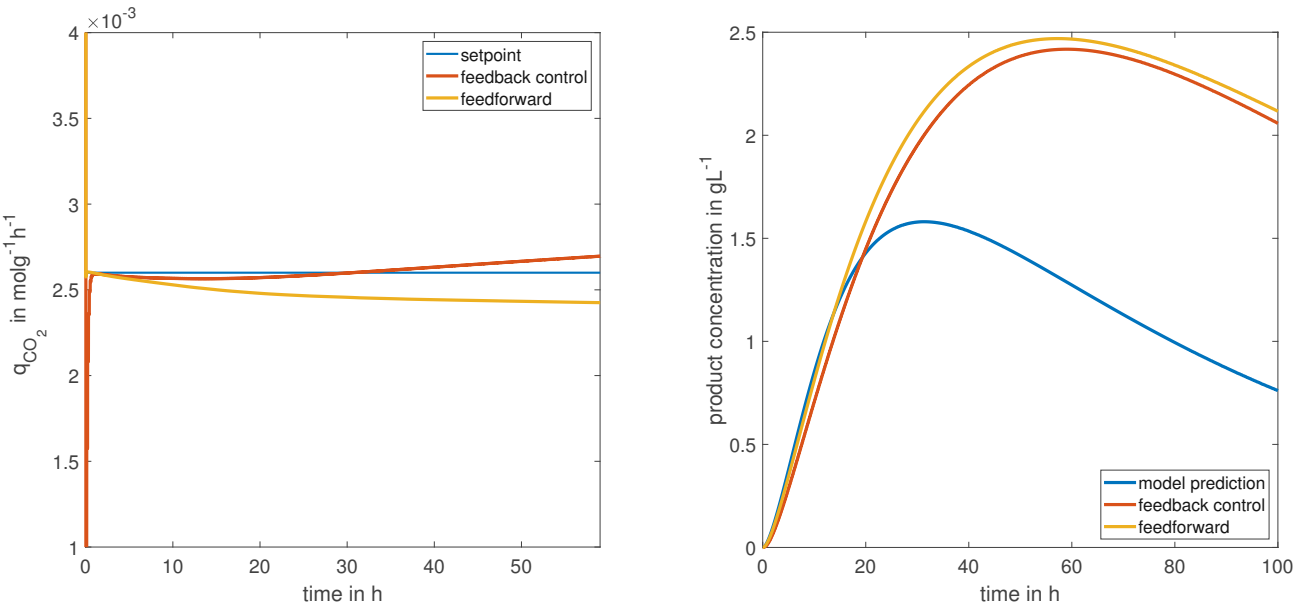

Figure 4.23: controller performance q_{CO2}: on the left hand side the course in time of the controlled rates for feedforward and feedback control is shown. On the right hand side the courses of the product formation are shown to assess the effects of deviations on the process performance.

Apparently all the graphs above show that the examined experiment performed better concerning productivity than the model on which the controller is based on.

In table 4.4 the NRMSE of the control laws is given for the feedforward as well as for the feedback control laws. It can be seen, that the feedback control law using the presented two degrees of freedom controller could improve the control performance. Anyway due to its simplicity the model based feedforward controller is also an interesting approach for further investigations. Of course, as discussed before especially for the control of the specific product formation and growth rate the courses obtained using feedforward control deviate strongly from the setpoint which has negative effects on the productivity and the target to keep the rates constant.

Table 4.4: Normalized root mean square errors between setpoint and the concerning specific rate held constant via feedforward or feedback control

control objective	**feedforward control**	**feedback control**	**Dim**
q_S	2.83	0.70	%
μ	9.18	0.88	%
q_P	38.94	8.30	%
q_{O_2}	4.28	1.24	%
q_{CO_2}	5.38	2.02	%

Additionally, the control errors for the maximum product formation setpoints are shown in table 4.5. The errors for the feedforward controller exceeds those of the space-time yield setpoints as the controller was tuned on those. Besides this higher errors the feedback control is still effective and reduced the tracking error compared to the feedforward control.

Table 4.5: Normalized root mean square errors between setpoint and the concerning specific rate held constant via feedforward or feedback control

control objective	**feedforward control**	**feedback control**	**Dim**
q_S	1.36	1.73	%
μ	59.64	313.34	%
q_P	69.69	4.22	%
q_{O_2}	1.85	0.25	%
q_{CO_2}	9.88	10.66	%

Table 4.5 shows that the constant growth rate controller does not work properly at the low setpoint. The growth rate becomes negative due to the very low feed and the model behaves physiologically incorrectly. Despite the negative growth rate, the product concentration continues to increase. Thus, keeping the growth rate constant at the specified level is difficult to control and results on the product yield are not conclusive for this case.

4.6 Implementation study

The implementation study is intended to provide first insights on the applicability of the developed model-based control laws. For realistic model-plant mismatches the controller used the generic model parameters 4.1 and the plant behavior was simulated with parameters from the single Experiments. In a addition to that an "new" verification plant was included which contained model parameters from an experiment, which was not included in the model parameterisation procedure.

4.6.1 Selected plant within the calibrated range

For the implementation study the experiment with the low constant feed rate 4.6 was chosen because it was the expermient with the highest product yields and concentrations. The graph 4.24 below shows the maximum achievable product concentrations for each control law at the setpoint for the maximum space-time yield. To provide a reference, the product yield of the experiment for the described data set is also given. It can be seen that the model predicts less product yield than can be achieved with the parameter set of the selected experiment in the simulation. As explained in the previous chapter, the setpoints can be well maintained with both the feedforward and the feedback controller. Because of this and because the experiment was used to fit the model and thus behaves very similar to the model, product yields using both control strategies are comparable. Just in the case of controlling the specific product formation rate, the feedback controller has a decisive advantage compared to the feedforward control strategy. At the setpoints for the optimal space-time yields, the maximum achievable product concentration is not to be expected and in graph 4.24 is also not evident. The results are shown in figure 4.25 and demonstrate the potential of the control strategies to improve the process.

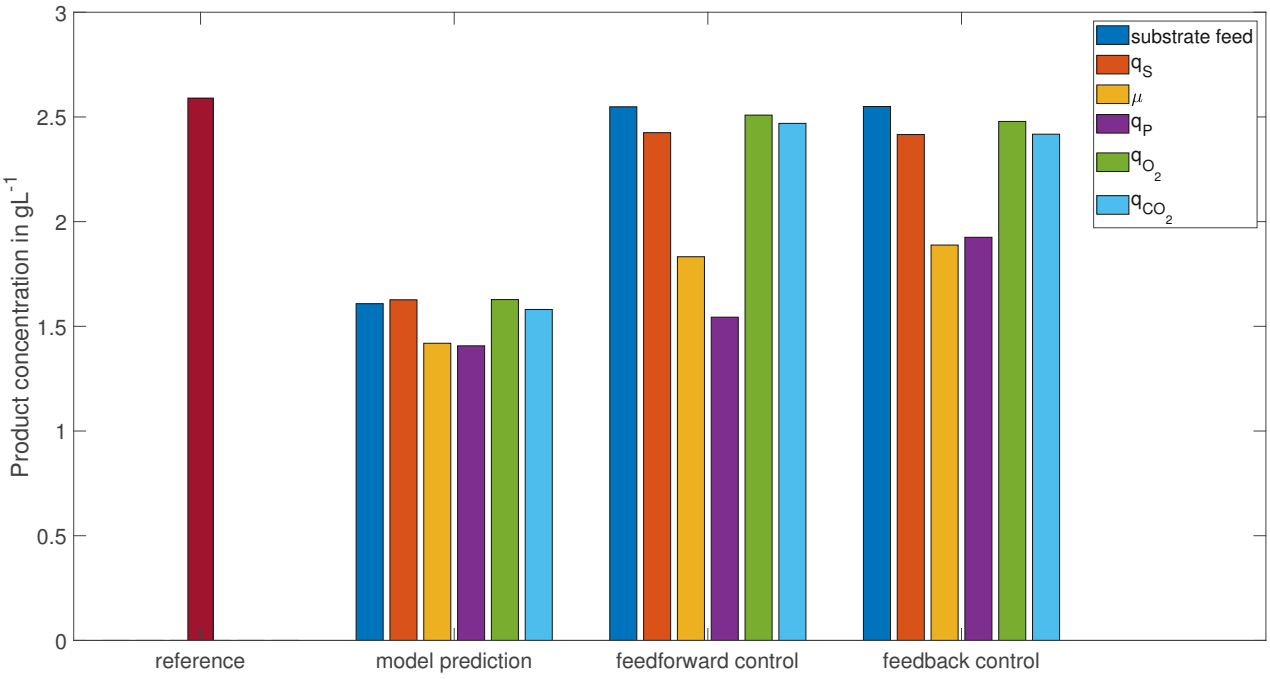

Figure 4.24: Achieved product concentration of different control laws with set-points for maximum space-time yield to be expected

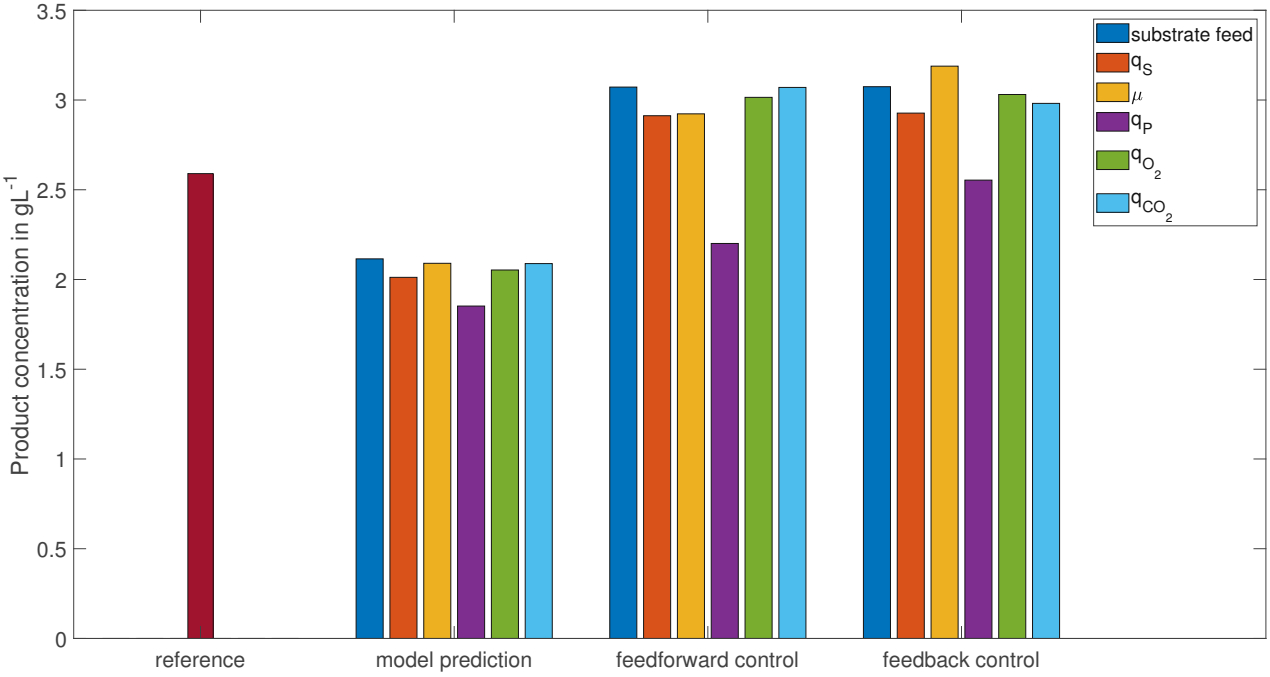

Figure 4.25: Achieved product concentration of different control laws with set-points for maximum product concentration to be expected

4.6.2 Independent verification plant

To get an idea of the applicability of the obtained control laws, the plant behavior of the independent verification experiment was simulated as a substitute for a real experiment. This independent verification plant was created by selecting an experiment from the data set, that was not included into model parametrization. The model parameters of the verification plant were estimated in the same way as for the model calibration. Mass balances and kinetic expressions to describe the behavior of plant were kept the same.

The parameters of this verification plant are listed in table 4.6. They deviate significantly from the actual model compared to the experiments that were included in the model. Especially parameters related to the product formation like $q_{P,max}$, $K_{S_{q_S}}$ the Haldane coefficient Hal and $K_{I_{q_P}}$ deviate considerably from the parameters obtained by the calibration data sets 4.1.

Table 4.6: Parameter set for the new plant gained fitting an independent experiment

	Model	**New plant**	**Dim**
$\dot{V}_{in}$	–	–	Lh^{-1}
$c_{S,in}$	850	840	gL^{-1}
$q_{S,max}$	1	1	$gg^{-1}h^{-1}$
$q_{P,max}$	0.008	0.00686	$gg^{-1}h^{-1}$
K_S	0.0050	0.0050	gL^{-1}
M	0.02	0.022	h^{-1}
$Y_{X/S,max}$	0.47	0.47	gg^{-1}
$K_{Y_{X/S}}$	0.39525	0.42263	gg^{-1}
$K_{S_{q_S}}$	0.1591	0.04013	$gg^{-1}h^{-1}$
Hal	3.833	5.6	–
$K_{I_{q_P}}$	6.9	4.82	–
$K_{S_{q_P}}$	0.19513	0.14	gg^{-1}
$K_{q_{S,rel}}$	0.18898	0.1134	h^{-1}
$K_{S_{P_S}}$	2.2924	2.0829	–

The developed controllers, the predetermined, optimal setpoints were used to run several simulations using the new independent plant. The results are shown in the bar charts below. Figure 4.26 shows the product concentrations achieved with the setpoints for the maximum space-time yield and figure 4.27 the expected maximum product concentrations.

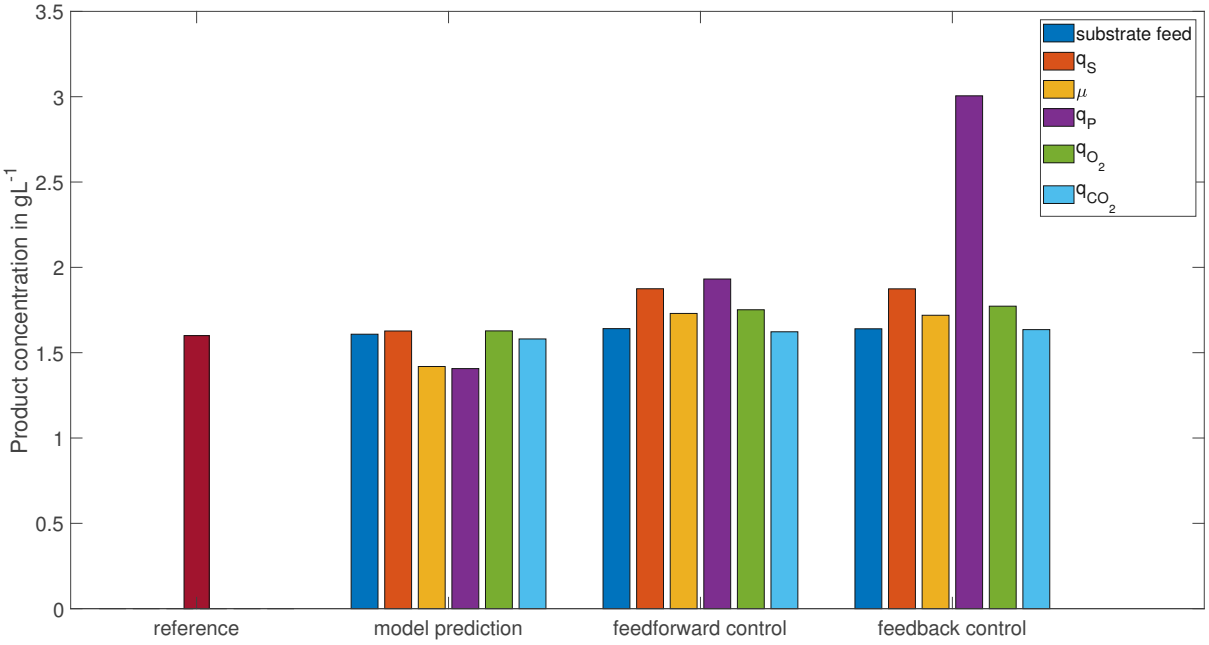

Figure 4.26: Achieved product concentration of different control laws with set-points for maximum space-time yield to be expected for the independent experiment

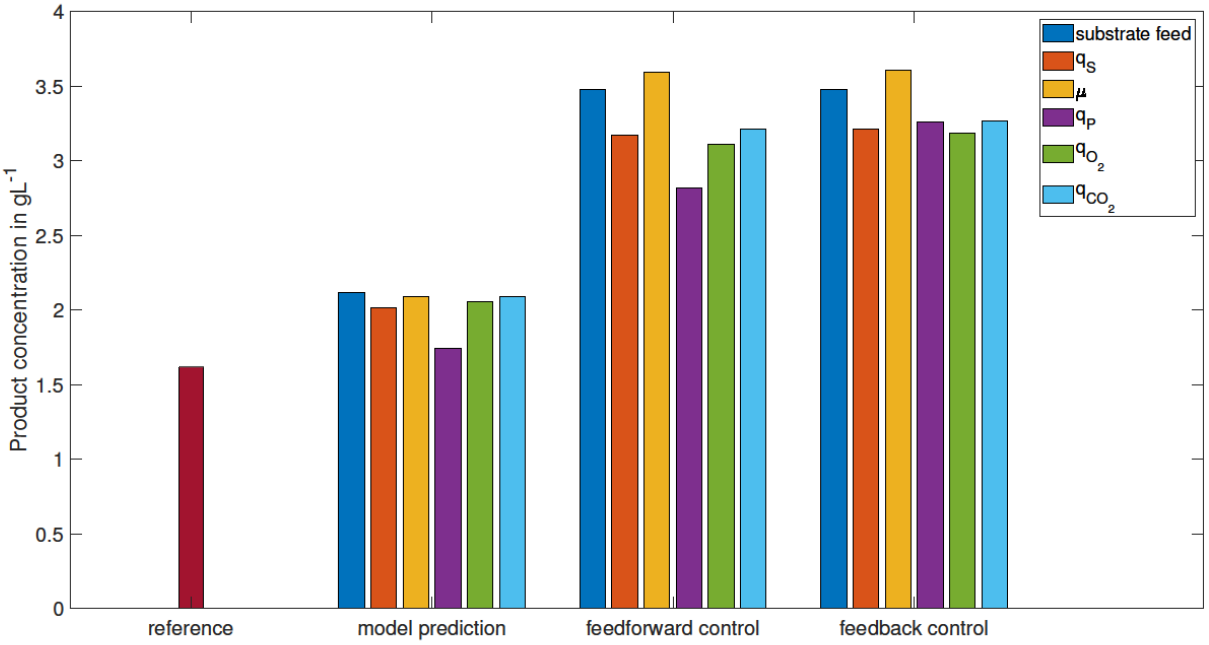

Figure 4.27: Achieved product concentration of different control laws with set-points for maximum product concentration to be expected for the independent experiment

Overall the independent plant performed better than the model concerning productivity. In 4.26 it could be seen that with the direct control of q_P (violet bar) the highest product concentrations could be achieved with the optimal space-time yield setpoint. For this rate the feedforward and feedback control differ significantly from each other. Nevertheless the performance of the feedforward control was comparable to the results of other controlled rates even if it deviates strongly from the desired trajectory. With the feedback control the space-time yield setpoint for q_P could be held constant quite well with an error < 7 % and the highest maximum product concentration could be reached with about 3 g/l. The feedforward control performed well for all the other rates and the setpoints for the optimal space-time yield, therefore the maximum product concentrations for both feedforward and feedback control were quite the same.

4.27 shows the performance of the setpoints for the maximum product concentration. Feedback control improved the control performance of all controlled rates and therefore higher product titers were reached compared to feedforward control. This was especially the case for the product formation rate. Again the low setpoint for the growth rate can not be held constant for the same reason described before. The high product concentrations predicted are not reliable as the model behaves incorrectly in the case of negative growth rates.

The errors of the control laws are depicted for the virtual experiment as well in table 4.7 for the setpoints for the maximum space-time yield and in table 4.9 for the setpoints for the maximum product concentration.

Table 4.7: Normalized root mean square errors between desired and the actual value of the concerning specific rate held constant due to simulation of the virtual experiment for the setpoint of the maximum space-time yields for feedforward and feedback control respectively

control objective	**feedforward control**	**feedback control**	**Dim**
q_S	0.96	0.34	%
μ	2.44	1.34	%
q_P	175.17	6.96	%
q_{O_2}	2.03	1.95	%
q_{CO_2}	2.52	2.03	%

Table 4.8: Normalized root mean square errors between desired and the actual value of the concerning specific rate held constant due to simulation of the virtual experiment for the setpoint of the maximum product concentration for feedforward and feedback control respectively

control objective	**feedforward control**	**feedback control**	**Dim**
q_S	2.09	0.43	%
μ	365.8	137.5	%
q_P	85.81	74.21	%
q_{O_2}	3.48	0.37	%
q_{CO_2}	3.99	3.97	%

Table 4.9: Normalized root mean square errors between desired and the actual value of the concerning specific rate held constant due to simulation of the virtual experiment for the setpoint of the maximum product concentration for feedforward and feedback control respectively

The control laws for the different constant rates could be successfully applied to control the independent plant. The feedforward control approach reaches satisfactory results concerning the product output. However the feedback control of the product formation rate achieves the best result with the maximum space-time yield setpoint. The mismatches for the maximum product formation rate setpoints are high for the product formation rate as well as for the growth rate. While the product formation could be improved again via feedback control, the setpoint for the growth rate could not be reached as described above. In general the different control laws and setpoints lead to improved processes compared to the reference experiments for the simulation. Of course, this results have to be investigated further in experiments. However the implementation study shows the potential of the investigated control laws, especially for the feedforward control approach. In application the feedback controller might nevertheless be the better choice to compensate disturbances.

4.7 Applicability of the different controllers

The simulation of the experiment, which is independent of the model, shows that the control laws created are functional. Furthermore, an increase in productivity can be expected through setpoint optimization using the developed model. This model can also be used to find feed trajectories or to derive control laws to control biomass specific rates.
The effects on keeping the desired setpoint at a constant level and on productivity achieved by optimization using the model can be additionally improved by feedback control. However, it must be noted that both specific rates and the process variables necessary for calculation cannot be determined straightforwardly online. Many important process variables are de-

termined by sampling and are therefore not available in time resolved form, which makes feedback control challenging.[21]
As the specific rates are not directly measurable, the assumption of continuous information about the controlled variable used for the simulation of the virtual experiment is a simplification and does not usually apply to practical applications. In the following, the applicability of the individual specific rates as controlled variables will be further discussed.

4.7.1 Growth rate

The growth rate is one of the key variables in a fed-batch fermentation process, as it has a direct influence on cell metabolism and thus on product formation.[36]
Especially when it comes to process control the growth rate is of great importance. This results from the strong impact on the productivity and product quality, as the biosynthesis of many products depend on certain growth profiles and overflow metabolism that decreases biomass productivity can be prevented controlling the growth rate below a critical value.
However the control of the specific growth rate in fed-batch processes depends on the accuracy of biomass estimation.[37]
Although biomass concentration can be monitored using various methods, the signal obtained is often noisy and unstable. This is caused due to the sensitivity of the probes to process conditions such as flow and aeration. These inaccuracies in biomass estimation make it difficult to control the specific growth rate. After all, since feedback control depends on the capability to quantify the biomass trend in real time. Thus, to obtain useful feedback on biomass concentration, appropriate data processing methods are needed in addition to measurement technologies to smooth and filter the signals. Signal filtering can reduce signal noise, but unfortunately at the cost of measurement delays, which in turn negatively affect controllability.[36]
Usually the growth rate is estimated using reference models that take substrate consumption or the oxygen uptake rate into account because measuring the biomass directly faces difficulties. Because of uncertainties caused by the continually changing state of the culture models are not able to predict the biomass concentration and therefore the specific growth rate precisely. Highly complex adaptation mechanisms to face this problem by minimizing model mismatches often lack the necessary robustness to be applicable.[37]
The simple approach of controlling μ by using a predefined exponential feed profile for open-loop control is therefore often used, but the simple exponential feed cannot cope with unpredictable process dynamics.[36]
Using conventional proportional or proportional-integral controllers to face this problem is difficult because of the exponential nature of process kinetics and disturbances and therefore lack the required robustness. Dabros et al. used a simple feed-forward/feedback principle where the proportional and integral terms are included directly in the exponential term which led to a sufficiently stable control. Dielectric spectroscopy was used to monitor the the biomass which than was subjected to online balances and reconciled in real time against

metabolite concentrations and off-gas compositions.[37]
As dependable online measurements of biomass concentration, using online balancing and data reconciliation for verification can be achieved [37], the presented control strategy for the growth rate could be an interesting approach.

4.7.2 Substrate uptake rate

The biomass-specific substrate uptake rate is also an interesting control variable. The control of the substrate concentration close to a value corresponding to the critical osmotic capacity can be used as well as the control of the specific growth rate to prevent overflow metabolism.[22] Compared to the growth rate, it features a larger control range at lower values. In [21] the biomass specific uptake rate of limiting substrate could be successfully controlled using an PI(D) controller as well as an open loop model based controller. The occurring instabilities of the PI controller as well as the limitation of the MBC controller described could be tackled using the approach detailed in this work. Of course, this assumption has to be verified in practice, but it could possibly be a simpler alternative to a model predictive control.

4.7.3 Product formation rate

Increasing the overall product concentration or the productivity of a fed-batch fermentation process is of course a common control objective. However, the absence of robust on-line measurements of these variables still hamper the direct control of the product specific quantities.[1]

The presented approach to control the specific product formation rate via an open loop control based on the model presented could therefore be an interesting strategy to be investigated further in laboratory experiments. In addition, as the future focus in the biotech industry moves towards continuous processes [38], the approach mentioned here to keep productivity constant is particularly interesting.

4.7.4 Oxygen uptake and carbon dioxide emission rate

Especially for high cell-density cultures, the oxygen transfer rate plays a major role. In such processes, the limiting oxygen transfer capacities are problematic, as accumulated cell mass decreases the mass transfer of oxygen, and thus the specific growth rate, decreases.[39] Dissolved oxygen is one of the few robust and reliable measured variables in the bioprocesses. In addition, the variables measured in the exhaust gas have the considerable advantage that they are very accurate, as they are not affected by concentration fluctuations such as those measured in the reactor. The online gas data therefore provide important information on the CO2 emission rate (CER) the oxygen uptake rate (OUR) and the respiratory quotient

(RQ).[1] Since these values are available in good quality and in real time, it is reasonable to consider them as potential control variables. Gustavsson et. al could show that a fed-batch culture producing a recombinant protein can be controlled by the CO2 emission rate. already the simple feedback control used in this work could achieve good results for process control via qCO2. This variant of bioprocess control via alternative rates will presumably be further improved with more precise process models and thus become interesting for industrial applications.[40] Like CER, OUR is also a potential control variable. In a study from 2010, it was shown that a process for the biotechnological production of a plant protection agent could be successfully controlled on an industrial scale via the oxygen uptake rate.[41] In [42] a respiratory quotient (RQ) based feedback control was applied for glucose feeding. A PID controller was used to automatically adjust the flow rate to maintain RQ at a set point to control metabolism and avoid glucose overflow. This resulted in higher product titers and lower by-product production.

Chapter 5

Conclusion

The aim of the thesis was to obtain and investigate control laws to control the biomass specific rates of a fed-batch fermentation process in the induction phase. For this purpose, an already existing nonlinear process model was extended and fitted on selected experiments of a provided data set. Using the method of nonlinear feedback linearization control laws to keep the specific rates constant were obtained. These control laws which compensate the nonlinearities of the system were than be integrated into a linear control loop. An optimization was performed simulating the process at various setpoints for the controlled variable to receive interesting setpoints for further investigations. The conclusive implementation study showed the theoretical potential of the control laws concerning controller robustness as well as optimization potential.

In this work new possibilities to control the induction phase of a fed-batch fermentation process are shown. The described approach has two advantages with respect to process control; (1) a relatively simple model is used that can be easily adapted to other processes with sufficient knowledge of process kinetics, (2) despite or because of the simple model that does not use complicated adaptation algorithms, a simple and robust feedback control can be achieved that compensates for process instabilities. These aspects, in combination with sufficiently accurate measurement data, seem to bring a direct control of interesting process parameters such as the growth or product formation rate within reach. Especially when it comes to continuous processing keeping the specific product formation rate or specific rates strongly related to product formation like the growth rate constant could be an interesting applicable approach to control a fermentation process.

However, the applicability of the presented control laws depends strongly on the availability of sufficiently accurate measurement methods. The simplifications used with regard to the availability of time-resolved measurements for the individual rates are not permissible, as in reality online measurement methods or analytical methods for data preparation represent a bottleneck in process realization.[38] Model-based control addresses this problem in part, of course, by predicting deviations, but it cannot completely replace measurements. Since

the model used behaves physiologically analogously to the simulated experimental process, the controller using model-based feedforward control is similarly effective to the closed loop feedback control. Feedforward control for the process described here could therefore have great potential in application, which needs to be tested in practice.

Bibliography

[1] Lisa Mears, Stuart M. Stocks, Gürkan Sin, and Krist V. Gernaey. A review of control strategies for manipulating the feed rate in fed-batch fermentation processes. *Journal of Biotechnology*, 245:34–46, 2017.

[2] Karl Schügerl and Karl-Heinz Bellgardt. *Bioreaction Engineering: Modeling and Control*. Springer-Verlag, Berlin Heidelberg, 2000.

[3] Silvia Ochoa. A new approach for finding smooth optimal feeding profiles in fed-batch fermentations. *Biochemical Engineering Journal*, 105:177–188, 2016.

[4] Rimvydas Simutis Andreas Lübbert Stefan Gnoth, Marco Jenzsch. Control of cultivation processes for recombinant protein production: a review. *Bioprocess and Biosystems Engineering*, 31:21–39, 2008.

[5] Tim W. Overton. Recombinant protein production in bacterial hosts. *Drug Discovery Today*, 19(5):590–601, 2014.

[6] Oliver Spadiut, Simona Capone, Florian Krainer, Anton Glieder, and Christoph Herwig. Microbials for the production of monoclonal antibodies and antibody fragments. *Trends in Biotechnology*, 32(1):54–60, 2014.

[7] Julian Kopp, Stefan Kittler, Christoph Slouka, Christoph Herwig, Oliver Spadiut, and David J. Wurm. Repetitive fed-batch: A promising process mode for biomanufacturing with e. coli. *Frontiers in Bioengineering and Biotechnology*, 8, 2020.

[8] Yu Luo, Varghese Kurian, and Babatunde A Ogunnaike. Bioprocess systems analysis, modeling, estimation, and control. *Current Opinion in Chemical Engineering*, 33:100705, 2021.

[9] Denes Zalai, Julian Kopp, Bence Kozma, Michael Küchler, Christoph Herwig, and Julian Kager. Microbial technologies for biotherapeutics production: Key tools for advanced biopharmaceutical process development and control. *Drug Discovery Today: Technologies*, 38:9–24, 2020.

[10] Harini Narayanan, Mattia Sponchioni, and Massimo Morbidelli. Integration and digitalization in the manufacturing of therapeutic proteins. *Chemical Engineering Science*, 248:117159, 2022.

[11] Yingjie Chen, Ou Yang, Chaitanya Sampat, Pooja Bhalode, Rohit Ramachandran, and Marianthi Ierapetritou. Digital twins in pharmaceutical and biopharmaceutical manufacturing: A literature review. *Processes*, 8(9), 2020.

[12] Christian Appl, André Moser, Frank Baganz, and Volker C. Hass. *Digital Twins for Bioprocess Control Strategy Development and Realisation*, pages 63–94. Springer International Publishing, Cham, 2021.

[13] Jens Smiatek, Alexander Jung, and Erich Bluhmki. Towards a digital bioprocess replica: Computational approaches in biopharmaceutical development and manufacturing. *Trends in Biotechnology*, 38(10):1141–1153, 2020. Special Issue: Therapeutic Biomanufacturing.

[14] Benjamin Bayer, Roger Dalmau Diaz, Michael Melcher, Gerald Striedner, and Mark Duerkop. Digital twin application for model-based doe to rapidly identify ideal process conditions for space-time yield optimization. *Processes*, 9(7), 2021.

[15] Joachim Almquist, Marija Cvijovic, Vassily Hatzimanikatis, Jens Nielsen, and Mats Jirstrand. Kinetic models in industrial biotechnology – improving cell factory performance. *Metabolic Engineering*, 24:38–60, 2014.

[16] Sophia Ulonska, Paul Kroll, Jens Fricke, Christoph Clemens, Raphael Voges, Markus M. Müller, and Christoph Herwig. Workflow for target-oriented parametrization of an enhanced mechanistic cell culture model. *Biotechnology Journal*, 13(4):1700395, 2018.

[17] Isuru A. Udugama, Pau C. Lopez, Carina L. Gargalo, Xueliang Li, Christoph Bayer, and Krist V. Gernaey. Digital twin in biomanufacturing: challenges and opportunities towards its implementation. *Systems Microbiology and Biomanufacturing*, 1:257–274, 2021.

[18] Anurag S. Rathore, Somesh Mishra, Saxena Nikita, and Priyanka Priyanka. Bioprocess control: Current progress and future perspectives. *Life*, 11(6), 2021.

[19] Magdalena Pappenreiter, Sebastian Döbele, Gerald Striedner, Alois Jungbauer, and Bernhard Sissolak. Model predictive control for steady-state performance in integrated continuous bioprocesses. *Bioprocess and Biosystems Engineering*, 45:1499–1513, 2022.

[20] Wan Ying Chai, Kenneth Tze Kin Teo, Min Keng Tan, and Heng Jin Tham. Model predictive control in fermentation process – a review. *AIP Conference Proceedings*, 2610(1):070008, 2022.

[21] Julian Kager, Andrea Tuveri, Sophia Ulonska, Paul Kroll, and Christoph Herwig. Experimental verification and comparison of model predictive, pid and model inversion

control in a penicillium chrysogenum fed-batch process. *Process Biochemistry*, 90:1–11, 2020.

[22] Merouane Abadli, Laurent Dewasme, Sihem Tebbani, Didier Dumur, and Alain Vande Wouwer. An experimental assessment of robust control and estimation of acetate concentration in escherichia coli bl21(de3) fed-batch cultures. *Biochemical Engineering Journal*, 174:108103, 2021.

[23] Julian Kager, Johanna Bartlechner, Christoph Herwig, and Stefan Jakubek. Direct control of recombinant protein production rates in e. coli fed-batch processes by non-linear feedback linearization. *Chemical Engineering Research and Design*, 182:290–304, 2022.

[24] Michael L. Shuler and Fikret Kargi. *Bioprocess engineering : basic concepts / Michael L. Shuler, Fikret Kargi.* Prentice Hall international series in the physical and chemical engineering sciences. Prentice-Hall, Englewood Cliffs, N.J., 1992.

[25] J. Bokor K.M. Hangos and G. Szederkényi. *Analysis and Control of Nonlinear Process Systems.* Springer-Verlag, Berlin Heidelberg, 2004.

[26] Julian Kager. Physiological bioprocess control: Constant specific substrate uptake rate in an antibody producing e. coli expression platform.

[27] Joseph DiStefano III.

[28] A. Raue, C. Kreutz, T. Maiwald, J. Bachmann, M. Schilling, U. Klingmüller, and J. Timmer. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics*, 25(15):1923–1929, 06 2009.

[29] Sven Daume, Sandro Kofler, Julian Kager, Paul Kroll, and Christoph Herwig. *Generic Workflow for the Setup of Mechanistic Process Models*, pages 189–211. Springer US, New York, NY, 2020.

[30] Roland Brun, Martin Kühni, Hansruedi Siegrist, Willi Gujer, and Peter Reichert. Practical identifiability of asm2d parameters—systematic selection and tuning of parameter subsets. *Water Research*, 36(16):4113–4127, 2002.

[31] Alford Joseph S. Bioprocess control: Advances and challenges. *Computers Chemical Engineering*, 30(10):1464–1475, 2006. Papers form Chemical Process Control VII.

[32] J.A. Baeza. 18 - principles of bioprocess control. In Christian Larroche, Maria Ángeles Sanromán, Guocheng Du, and Ashok Pandey, editors, *Current Developments in Biotechnology and Bioengineering*, pages 527–561. Elsevier, 2017.

[33] Jakubek Stefan. *Feedback Control*. Institut für Mechanik und Mechatronik, Wien, 3 edition, 2019.

[34] Nikolaus Euler-Rolle, Igor ***S***krjanc, Christoph Hametner, and Stefan Jakubek. Automated generation of feedforward control using feedback linearization of local model networks. *Engineering Applications of Artificial Intelligence*, 50:320–330, 2016.

[35] Hidefumi Taguchi and Mituhiko Araki. Two-degree-of-freedom pid controllers — their functions and optimal tuning. *IFAC Proceedings Volumes*, 33(4):91–96, 2000. IFAC Workshop on Digital Control: Past, Present and Future of PID Control, Terrassa, Spain, 5-7 April 2000.

[36] Yann Brignoli, Brian Freeland, David Cunningham, and Michal Dabros. Control of specific growth rate in fed-batch bioprocesses: Novel controller design for improved noise management. *Processes*, 8(6), 2020.

[37] Michal Dabros, Moira Monika Schuler, and Ian W. Marison. Simple control of specific growth rate in biotechnological fed-batch processes based on enhanced online measurements of biomass. *Bioprocess and Biosystems Engineering*, 33:1109–1118, 2010.

[38] Julian Kopp, Christoph Slouka, Oliver Spadiut, and Christoph Herwig. The rocky road from fed-batch to continuous processing with e. coli. *Frontiers in Bioengineering and Biotechnology*, 7, 2019.

[39] Yu-Cheng Chung, I-Lung Chien, and Der-Ming Chang. Multiple-model control strategy for a fed-batch high cell-density culture processing. *Journal of Process Control*, 16(1):9–26, 2006.

[40] Robert Gustavsson, Cornelia Lukasser, and Carl-Fredrik Mandenius. Control of specific carbon dioxide production in a fed-batch culture producing recombinant protein using a soft sensor. *Journal of Biotechnology*, 200:44–51, 2015.

[41] Ju Chu Yong-Hong Wang Ying-Ping Zhuang Jian-Guang Liang, Xiao-He Chu and Si-Liang Zhang. Oxygen uptake rate (our) control strategy for improving avermectin b1a production during fed-batch fermentation on industrial scale (150 m3). *African Journal of Biotechnology*, 9(42), 2010.

[42] Zhi-Qiang Xiong, Mei-Jin Guo, Yuan-Xin Guo, Ju Chu, Ying-Ping Zhuang, Nam Sun Wang, and Si-Liang Zhang. Rq feedback control for simultaneous improvement of gsh yield and gsh content in saccharomyces cerevisiae t65. *Enzyme and Microbial Technology*, 46(7):598–602, 2010.