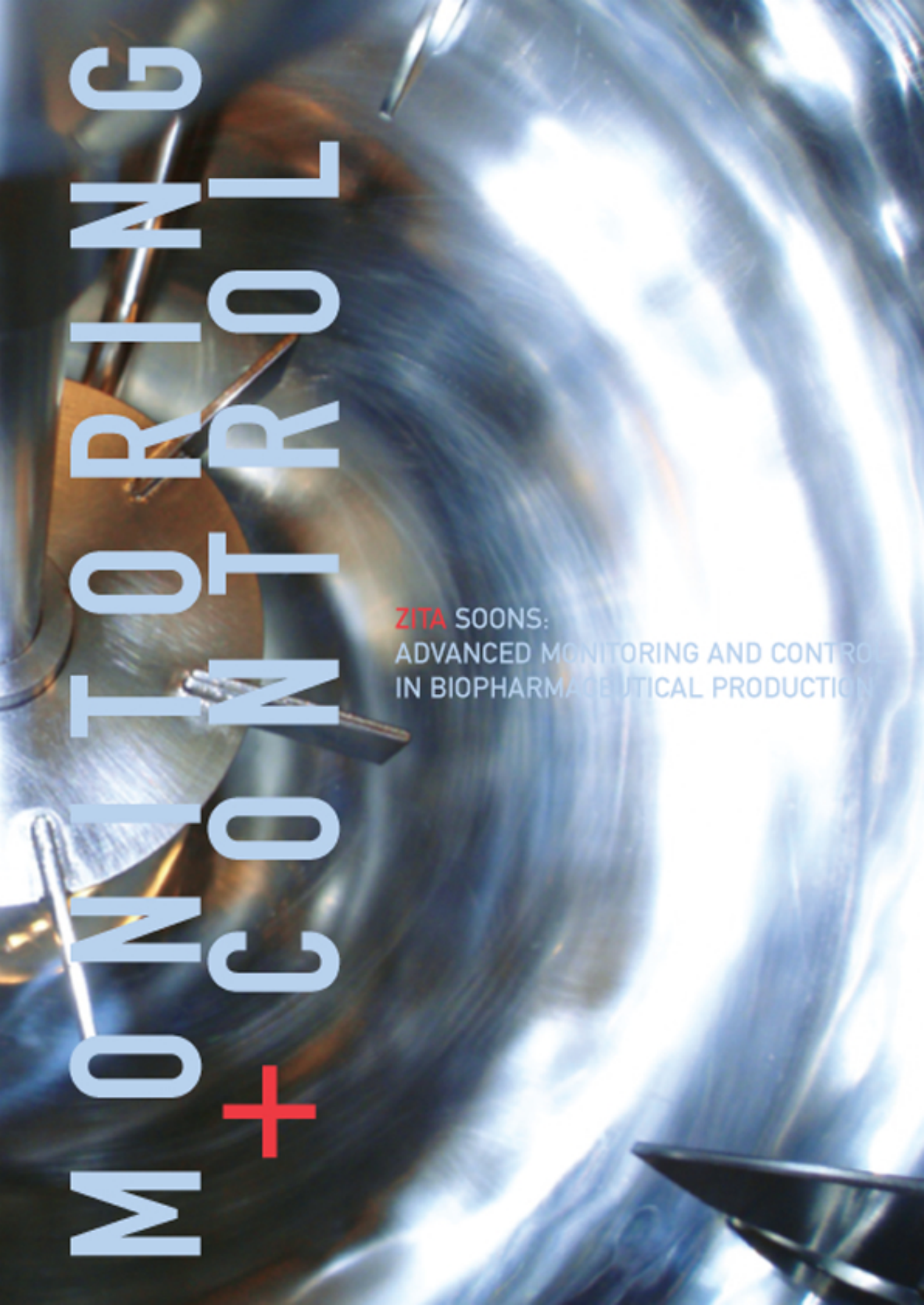


MONITORING + CONTROL

ZITA SOONS:
ADVANCED MONITORING AND CONTROL
IN BIOPHARMACEUTICAL PRODUCTION

Advanced monitoring and control in biopharmaceutical production

Zita I. T. A. Soons

Promotor:	***Prof. dr. ir. G. van Straten*** *Hoogleraar Meet-, Regel- en Systeemtechniek* *Wageningen Universiteit*
Copromotor:	***Dr. ir. A. J. B. van Boxtel*** *Universitair docent, leerstoelgroep Meet-, Regel- en Systeemtechniek* *Wageningen Universiteit*
Promotiecommissie:	***Prof. A. J. Morris FREng.*** *Newcastle University, United Kingdom*
	Prof. dr. ir. R.H. Wijffels *Wageningen Universiteit*
	Prof. dr. ir. J. van Amerongen *Universiteit Twente*
	Dr. ir. R. Neeleman *Xendo, Leiden*

Dit onderzoek is uitgevoerd binnen de onderzoekschool VLAG

Advanced monitoring and control in biopharmaceutical production

Zita I. T. A. Soons

Proefschrift

ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
Prof. dr. M. J. Kropff
in het openbaar te verdedigen
op vrijdag 13 juni 2008
des namiddags te half twee in de Aula

Advanced monitoring and control in biopharmaceutical production

Zita I. T. A. Soons

PhD thesis, Wageningen University, The Netherlands, 2008
With summaries in English and Dutch

ISBN: 978-90-8504-933-3

PREFACE

This thesis represents the work of my PhD at the Systems and Control Group from Wageningen University in cooperation with the Process Development group of the Netherlands Vaccine Institute. Since February 2004 I worked in a project called “PaRel” or “Parametric Release” denoting the ultimate goal of the project: release of the product on the basis of real-time measurements of quality parameters during the upstream process. The experiments and most of the work were performed at the Netherlands Vaccine Institute. This thesis contains eight chapters in the multidisciplinary field of systems and control theory, biology, and chemometrics in upstream biotechnology. Six chapters have been published or submitted for publication in international journals.

Many people have contributed to the completion of my dissertation. It is my pleasure to thank the people, which were most involved to my work: my promoter and copromotor Ton van Boxtel and Gerrit van Straten for the stimulating discussions and the critical comments on my manuscripts; Mathieu Streefland for the discussions on the biology and the critical comments; Bas van de Waterbeemd for doing the micro-arrays in the laboratory; Paul van Herpen for being my room mate; Leo van der Pol for giving me the opportunity to perform the experiments in the laboratory; Coen Beuvery for his enthusiasm as the initiator of the “PaRel” project; and the students Nicole Larmonie and Jing Shi for their major contributions to the work involved in my thesis. I wish to thank all people from the systems and control group and the unit process and product development for providing me a pleasant environment to work. Last but not least I want to thank my parents for their support and my daddy for designing the beautiful cover of my thesis.

ABSTRACT

Bioprocesses are characterised by natural variability in raw materials, initial conditions, human intervention, and varying properties of the micro-organism. In traditional biopharmaceutical production quality of the product is currently tested at the end of the production process only. Recently the Food and Drug Administration (FDA) released guidelines on process analytical technology (PAT) aiming at reducing variability in the end-products by introducing timely measurement and control of quality variables during the manufacturing process.

This thesis intends to be a step forward towards the introduction of real-time feedback control within the frame of PAT by combining several techniques from a multidisciplinary field of systems and control theory, biology, and chemometrics. The dual substrate cultivation of *Bordetella pertussis* for the production of bulk whole cell whooping cough vaccine was chosen as test case.

A software sensor was developed to reconstruct biomass and specific growth rate from easy to obtain measurements during cultivation. The software sensor accurately monitors biomass growth over the whole range from low to high biomass concentrations during different types of cultivation (batch, fed-batch, and continuous cultivations). It has also been investigated how near infra red spectra can be used to yield an alternative measurement of biomass. In the current state of development the software sensor is the preferred choice when it comes to feed-back control.

A model-reference based controller was derived that adapts its settings in response to the time-varying conditions in order to control the specific growth rate during fed-batch cultivation, thus obtaining higher bulk vaccine concentrations compared to the standard production process. Two methods to automate controller tuning were proposed and have successfully been implemented.

To obtain an indication which controller set-points are needed to obtain the best vaccine quality, the effect of specific growth rate on the formation of virulence factors has been investigated. The data and a proposed preliminary model are in favour of choosing the highest feasible growth rate.

Finally, with some precautions the monitoring and control tools developed on lab-scale using headspace aeration turned out to be transportable to medium size pilot-scale cultivation using sparger aeration.

CONTENTS

Chapter 1	*State of the art and challenges*	*page 11*
Chapter 2	*Observer design and tuning for biomass growth and k_La using online and offline measurements*	*page 21*
Chapter 3	*Towards PAT monitoring and control: near infrared and software sensor*	*page 43*
Chapter 4	*Constant specific growth rate in fed-batch cultivation of Bordetella pertussis using adaptive control*	*page 63*
Chapter 5	*Online automatic tuning and control for fed-batch cultivation*	*page 91*
Chapter 6	*Scaling-up vaccine production: implementation aspects of a biomass growth observer and controller*	*page 115*
Chapter 7	*Modelling the cultivation of Bordetella pertussis: the effect of specific growth rate on expression of virulence factors on gene and protein level*	*page 135*
Chapter 8	*Conclusions and perspectives*	*page 153*
References		*page 161*
Summary		*page 169*
Samenvatting		*page 173*
Curriculum Vitae		*page 177*
List of publications		*page 179*
Training activities		*page 181*

CHAPTER 1

State of the art and challenges

Developments in the pharmaceutical industry

The current Food Manufacturing Practiced (GMP) regulations for pharmaceutical manufacturing involve the production of products following strict rules, a fixed protocol, and fixed procedures (Fig 1A). Quality of the product is only tested at the end. Bioprocesses naturally contain variability due to the raw materials, initial conditions, human intervention, and varying properties of the micro-organism. By setting fixed protocols the current GMP regulations do not encompass the reduction of such deviations. This leads to products containing a considerable amount of variability, and possibly out of specifications. Already two decades ago Deming (1986) stated that "quality can not be proven into products" and "manage the cause and not the result". Recently the FDA released guidelines on process analytical technology (PAT) (FDA, 2004) - a framework for innovative pharmaceutical development, manufacturing, and quality assurance - in which they stress the need for reducing variability and aim at improving manufacturing processes by introducing timely measurement and control of quality variables (Fig. 1B). It requires a science-based approach to understand, monitor and control the process and its critical variables in real-time.

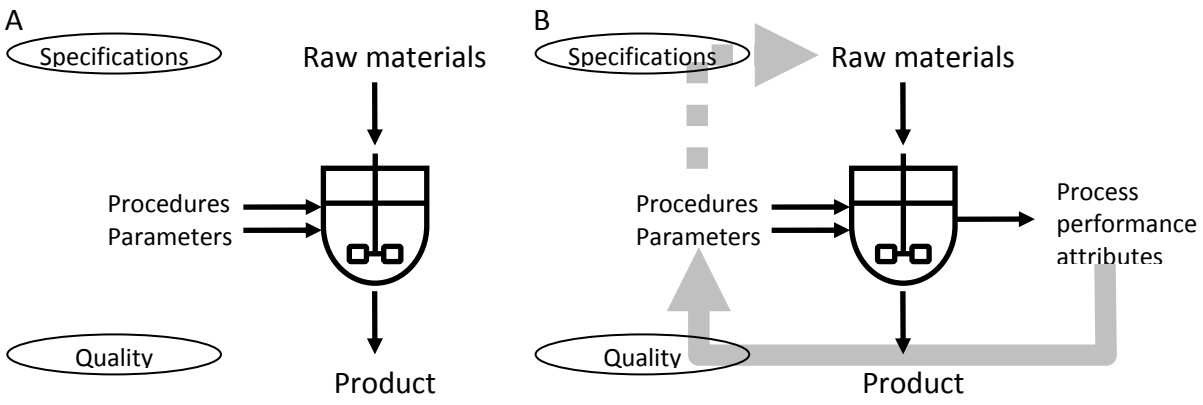


Figure 1A. *Classical fixed procedures to obtain fixed quality, which is only tested in the end-product* ***B.*** *PAT based procedures involving online measurement and feedback control of key variables for process performance and product quality.*

Developments in biopharmaceutical and vaccine production

The biopharmaceutical industry tends to move slowly and cautiously unlike many other industries with respect to new approaches in advanced technologies, such as feedback control algorithms (Alford, 2006). The biopharmaceutical is released on the basis of offline tests on the final product, e.g. sterility, homogeneity, and potency and safety tests in animals, involving large variability, delayed release, high costs and letting aside the possibility for feedback. Feedback means that deviations are corrected in real-time and can only be applied if measurements become available fast. In the newest draft of Pharmeuropa (Anonymous, 2007) the final bulk of whole cell pertussis vaccine is released on the basis of sterility, antimicrobial preservative, the presence of fimbriae, opacity (the bacterial concentration should not exceed a certain level per human dose), identity (based on an immunological reaction), specific toxicity (survival of healthy mice), aluminium (maximum per single dose), and free formaldehyde (maximum concentration). The requirements for the production process are: "The production method must be validated to yield consistently vaccines comparable with the vaccine of proven clinical efficacy and safety in man". Cultures must be checked at different stages for purity, identity, cell opacity, and pH. Cell harvest must comply with requirements for consistency based on the following criteria: growth rate, pH, yield, phase I characteristics (such as the presence of fimbriae), purity, identity, cell opacity, and free pertussis toxin, dermonecrotic toxin, or tracheal cytotoxin (the free toxins must be absent). So, the release of bulk vaccine is mainly based on offline data, and thus subjected to the drawbacks mentioned above (Fig. 1A). One handicap for the (bio)pharmaceutical industry's' reluctance to pursue advanced technologies is the accuracy and reliability by which the critical parameters of the complex cultivation processes can be monitored and controlled online. This thesis aims to develop methodologies for online monitoring and control that contribute to enhanced consistency and quality for the cultivation of *Bordetella pertussis* being the basis for the bulk vaccine against whooping cough (Fig. 1B).

Most biopharmaceuticals and vaccines are produced in batch cultivation, where cells grow until the main nutrients are depleted. This step in the production process mainly determines the quality of the final product (Streefland et al., 2007). Formation of virulence factors - the basis for a proper immune response against infection with whooping cough - is growth-associated (among others Licari et al., 1991; and Rodriguez et al., 1994). Deviations in growth rate therefore may lead to deviations in virulence factors and vaccine quality. To obtain a high quality-vaccine and to ensure batch-to-batch consistency, it is important to control the growth rate at a constant level. As a consequence, it is favourable to extend the

traditional batch cultivation with continuous or fed-batch cultivation. Advanced feed strategies can be applied to control the process at a desired state, thereby obtaining higher biomass concentrations and regulating the quality of the end-product. Ultimately, online monitoring and control towards quality and consistency should be the basis for product release and replace or diminish the tests in laboratory animals.

The PaRel project: a multidisciplinary team

In 2003-2007 a project was executed with a multidisciplinary approach to enhance the manufacturing of biopharmaceuticals. The project was named "PaRel" or parametric release, denoting the ultimate goal of the project: release of the product on the basis of real-time measurements of quality parameters during the upstream process. The project involved four partners, funded by the Dutch Ministry of Economical Affairs (TSGE3067). The four disciplines in the project were:

- Designing the hardware and software architecture of the PAT infrastructure (SIPAT) of bioreactors for process development and production of biopharmaceuticals. Specific aspects were the automation of the data collection, data interpretation and experimental design based on a near infrared sensor in one expert system (Siemens NV, The Hague, The Netherlands).
- The development of industrial measurement and control systems. Three tools were developed to facilitate the real-time application of PAT: Instrumentation enabling monitoring and control of standard variables (temperature, dissolved oxygen and pH) named "*i*-Control"; a parallel screening bioreactor for design of experiments on micro-scale (µ-24) ; and novel sensor techniques: redox, oxygen, and viable cell sensors (Applikon Biotechnology BV, Schiedam, The Netherlands).
- Enhanced process understanding through the development of key quality indicators for improved accuracy and consistency towards the release of a well characterized product. Four aspects were considered: reducing the variance of batch cultivations; definition of whole cell vaccine quality on the basis of virulence gene expression through micro-arrays; definition of critical parameters through design of experiments; and the application of the near infrared sensor for fingerprinting the batch cultivation (Netherlands Vaccine Institute, The Netherlands).
- Enhanced monitoring and control through the application of advanced feed strategies. A five-fold higher biomass concentration compared to standard batch cultivation is obtained by performing specific growth rate controlled fed-batch

cultivations. Four aspects were considered: monitoring using a software sensor and a near infrared sensor; auto-tuning feedback control to facilitate the implementation and tuning of controllers; modelling towards predicting vaccine quality; and scale-up from laboratory to pilot-scale production. (Wageningen University and Netherlands Vaccine Institute, The Netherlands).

In this thesis, the focus is on the last discipline: designing advanced modelling, monitoring, and control systems in computer simulations and its application to bacterial cultivations of *B. pertussis* on laboratory and pilot-scale (Fig. 2). Prior to discussing the research goals and scientific challenges elaborated in this thesis, first a brief overview is given about these three components.

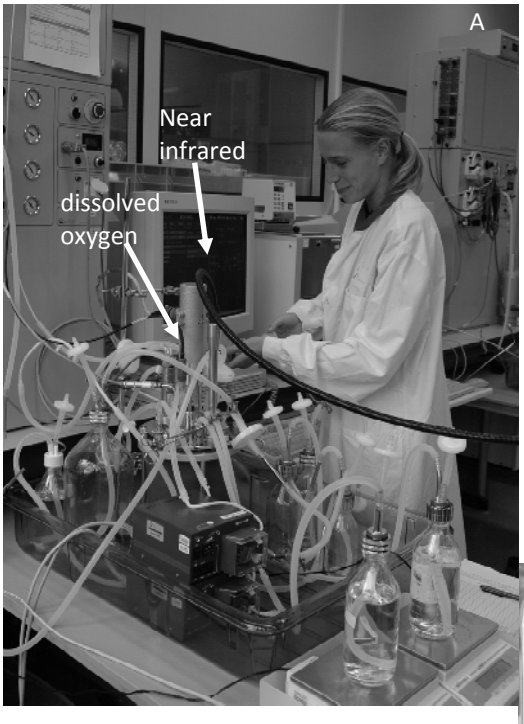


Figure 2A. *5-L bioreactor set-up for fed-batch cultivation of B. pertussis. The sensors for dissolved oxygen and near infrared are indicated with white arrows. The bottles containing the feed with the substrates and the pumps, which are computer controlled, are shown in front the bioreactor.*

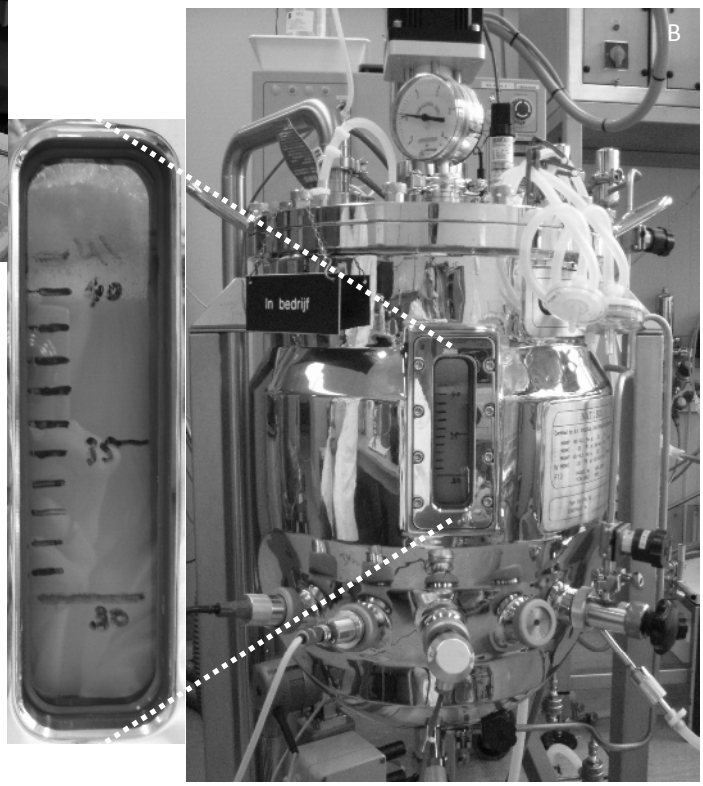


Figure 2B. *60-L bioreactor set-up for fed-batch cultivation of B. pertussis.*

Model

Batch and fed-batch processes are more challenging than stationary processes for modelling and control purposes, because batch processes show a dynamic range of operating conditions rather than moving towards steady states. Given these challenges, the analysis, monitoring, and control of batch processes tend to rely more on the data-driven generation of empirical models (Bakshi et al., 1994). In addition, the values of measured variables reflect the contributions from a variety of external factors such as sensor noise, time degradation of equipment, pump failure or variance of biological assays, etc. The modelling challenge is to develop models for steady-state experiments and for dynamic batch and fed-batch experiments. The models should be a solid basis for the ensuing model-based process design, monitoring, control, and predicting vaccine quality. Two extreme approaches to modelling systems are as follows (Sontag, 2006):

- *White model.* All relevant variables, such as substrates, biomass, proteins, and metabolites are specified. The forms for all reactions and reaction constraints are described. Physical, chemical, or biological phenomena dictate the model structure.
- *Black-box model.* The black box model ignores all mechanistic information and characterizes behaviour solely in terms of stimulus-response data. Statistical models that correlate the inputs and outputs, neural networks and other approximate models are black-box.

The grey*-box model* is in between these two extremes. It combines available mechanistic information with input-output data. Many biological models are this type.

Monitoring

Karim et al., 2003 stated: “One of the challenges when dealing with biological processes is how to extract biologically significant information from the relatively few process states measured online. These measurements are typically of physical bioreactor conditions (e.g., temperature, pH, agitation rate, bioreactor and exhaust gas compositions), whereas crucial information about the production of a desired metabolite or product often lies in the physiological condition of the cell, requiring time-consuming offline measurements and assays.” This work aims to monitor the “quality” state of the cells by using physical measurements. We distinguish three types of non-invasive bioprocess monitoring:

- *Standard Sensors*. Measurements from standard sensors like temperature, dissolved oxygen, and pH can be used to monitor the evolution and consistency of the cultivation.
- *Software Sensor*. The task of the soft-ware sensor is to monitor states, which can not be measured online due to the lack of available sensors. The software sensor combines process measurements from standard sensors with a process model (white-, black-, or grey-box) to observe the unknown states and to reduce noise on the measurements, which is often present on data from sensors in the bioreactor.
- *Statistical Process Monitoring.* Cultivation processes usually generate numerous data, which are often not used. As an example, a near infrared sensor may generate over 2000 measurements every minute. Multivariate techniques (e.g. principal component analysis or partial least squares) can reduce data, monitor process performance, or predict relevant parameters.

Control and tuning

Bioprocess control is defined as providing a near optimal environment for microorganisms to grow, multiply, and produce a desired product (Alford, 2006). Alternatively, biological reactors use the cellular metabolic pathways to produce a desired product and as such require conditions that optimize either cellular growth, metabolite or protein production (Karim et al., 2003). In bioreactors only standard physical or thermodynamic measurements, such as dissolved oxygen, temperature, and pH, are used for feedback control in some cases. The more relevant parameters for product quality are usually not measured nor controlled. A challenge is to regulate an indicator for quality (specific growth rate) at a desired level by using feedback control in this thesis.

- *Standard Control.* Based on a survey of over eleven thousand controllers in industries (Desborough and Miller, 2001), 97% of regulatory controllers use a proportional-integral derivative (PID) feedback control algorithm. A PID controller reacts on deviations after they occur and/or persist.
- *Adaptive Control*. The performance of standard control methods is limited because of the nonlinear and time-varying characteristics in bioprocesses. Fed-batch cultivations involve ever changing process conditions due to the exponentially increasing biomass. A model-based and adaptive approach, therefore, has been proven most effective to control complex process during a range of operating conditions (Åström et al., 1993). And so the aim is to develop a model-based controller that is capable of

adapting the control output based on the changing conditions in dynamic fed-batch cultivation, e.g. biomass and volume in a bioreactor.

- *Auto-tuning.* Another fact from industry in Desbourough and Miller (2001) is that the performance of only one third of 26.000 controllers was acceptable or excellent and two thirds was poor or fair and had opportunities for significant improvements. Auto-tuning has the potential to upgrading performance online if performance is not acceptable. Moreover, auto-tuning may enhance the application of advanced control techniques in industry, because it is makes the tuning of controllers easier and hence more user-friendly.

Thesis challenges and outline

In different parts of science and technology the word "control" has different meanings (Wold, 2006). In the world of process analytical technology or statistical process control "control" often means "check" whether the process is still within the desired specifications by means of monitoring or classifying the process. In the engineering sense "control" means to ensure that the process stays close to the set-points. Wold (2006) states that "PAT process control" in the engineering sense is still a challenge of the future. However, the philosophy of this these is to pursue just that. The research is intended to be a step forward to the "PAT process control" and aims to monitor and control the cultivations step of *B. pertussis* for the production of whole cell whooping cough vaccine towards desired quality specifications and enhanced batch-to-batch consistency. This thesis concerns the following research questions:

- How to improve monitoring of complex bioprocesses?
- How can the cultivation process be controlled at a desired level while coping with the nonlinear and time-varying characteristics of bioprocesses?
- How to automate controller tuning?
- What set-points are needed to obtain the best-quality vaccine?
- Can the developed monitoring & control tools be applied to cultivation on production-scale?

In order to answer the research questions the thesis comprises modelling, monitoring, and control in small- and pilot- scale bioreactors, which are operated in batch, fed-batch and continuous mode. The analysis of this work is supported not only by simulations but also in real laboratory experiments.

- Chapter 2 illustrates the design, development, and evaluation of a software sensor that observes specific growth rate and biomass using standard and cheap online dissolved oxygen measurements. Offline biomass measurements are incorporated to improve and safeguard the biomass estimation and to observe the effect of biomass growth on the oxygen transfer coefficient.
- As an alternative for the software sensor of chapter 2 and to safeguard the cultivation by introducing an additional sensor, chapter 3 describes the use of a near-infrared (nIR) sensor towards online monitoring and controlling the cultivation. Multivariate techniques and wavelength selection are used to monitor the evolution of the biomass concentration in batch cultivations. The potential and challenges of nIR monitoring are evaluated and compared with the software sensor of chapter 2.
- Chapter 4 uses the observed biomass growth of chapter 2 to control the specific growth rate in fed-batch cultivation. The dual substrate model developed in previous work (Neeleman et al., 2001) is combined with specifications on control performance to develop a model-based controller that is adaptive for the time-varying characteristics of bioprocesses. The adaptive feed-strategy is applied to obtain higher biomass concentrations in a controlled way.
- Chapter 5 concerns the automation of tuning of the specific growth rate controller of chapter 4, which is often a difficult task in bioprocesses. The performance of several methods for auto-tuning is evaluated.
- Chapters 1-5 concern monitoring & control in small-scale bioreactors. However, the value of the methodology needs to be asserted for larger scales in order to be useful for production of biopharmaceuticals. Chapter 6 makes the advanced monitoring and control system for fed-batch cultivation plausible to near-production-scale by showing the application to pilot-scale cultivations. In addition, one of the problems in scale-up “the calculation of the oxygen consumption counteracting additional off-gas dynamics due to the mixing and delays in the headspace and other equipment” are successfully tackled.
- The chapters so far concerned the monitoring and control of obtain biomass growth in a controlled way. Chapter 7 provides further insights in what set-point of the specific growth rate to choose to obtain the product that meets the desired specifications. A preliminary model is proposed to predicting formation of virulence factors on gene and protein level as function of the specific growth rate.
- Chapter 8 reflects the achievements obtained in this thesis and presents some future perspectives.

Note that the presented constants and parameters may differ in the subsequent chapters, because the order of the chapters in this thesis was not identical to the chronological order in which the chapters were written. This thesis is a sequel of the work of Neeleman (2002). The state of the art in relation to literature is given in the thesis chapters. A combination of the methods developed in chapters 2-7 can be used to predict, monitor, and control the cultivation and finally the end-quality of the vaccine. Thus, meeting the science-based approach of PAT to understand, monitor and control the process and its critical variables in real-time; and being a step forward to obtain quality by design and finally to reduce (animal) tests on the end-products.

CHAPTER 2

Observer design and tuning for biomass growth and k_La using online and offline measurements

Published as: Z. I. T. A. Soons, J. Shi, J. D. Stigter, L. A. van der Pol, G. van Straten, A. J. B. van Boxtel, Journal of Process Control, in press.

Abstract

Measurement of the key process variables is essential during biopharmaceutical production. These measurements are often not available online. This work combines frequent online measurements (oxygen uptake rate) with infrequent offline measurements (biomass) to estimate the specific growth rate, biomass, and the oxygen transfer coefficient (k_La) online. The system consists of an Extended Kalman filter and parameter adaptation for the time-varying k_La. Tuning is based on minimization of the error between the simulation and the estimation. Although the process itself is not stable, stability of the observer is evaluated heuristically by application of the Routh criterion. Performance and convergence of the observer are shown in both simulations and experiments in continuous and fed-batch cultivations of *Bordetella pertussis*.

Keywords

Extended Kalman filter, online and offline measurements, bioreactor

Nomenclature

A, B, C	system matrices
c	relative k_La
C_G, C_G^{in}	glutamate concentration in the medium and in the feed [$mmol.l^{-1}$]
C_L, C_L^{in}	lactate concentration in the medium, respectively in the feed [$mmol.l^{-1}$]
C_X	biomass concentration [OD]
C_O^{in}	oxygen concentration in inlet stream (liquid based) [$mmol.l^{-1}$]
C_O^{h}	equilibrium oxygen concentration in headspace (liquid based) [$mmol.l^{-1}$]
C_O^{L}, $c_{O,sensor}^{L}$	oxygen concentration in the medium, measured by the sensor [$mmol.l^{-1}$]
d	constant
D	dilution rate [h^{-1}]
E	performance criterion
F_S, F_{out}	substrate feed rate into reactor, feed rate out the reactor [$l.h^{-1}$]
I	identity matrix
ISS	input-to-state stability
K	Kalman gain
K_G, K_L	Monod constant on glutamate, respectively lactate [$mmol.l^{-1}$]
k_La, k_La^{Stir}	oxygen transfer coefficient, k_La as function of stirrer speed [h^{-1}]
m_G, m_L, m_O	maintenance coefficient on glutamate, lactate, and oxygen [$mmol.OD^{-1}.h^{-1}$]
OUR, OTR	oxygen uptake rate, oxygen transfer rate [$mmol.l^{-1}.h^{-1}$]
O	observability matrix
P	variance of the states
Q, Q^{tune}	system noise, system noise calculated by tuning
R	output noise
R^{Stir}	stirrer speed [rpm]
t	cultivation time [h]
T	frequent sampling interval [h]
T_m	time interval between two offline samples (not constant) [h]
u, x, y	inputs, states, outputs
V	liquid volume [l]
v, w	measurement noise, system noise
Y_{G1}, Y_{G2}	biomass yield on glutamate over pathway 1 and pathway 2 [$OD.mmol^{-1}$]
Y_L, Y_O	biomass yield on lactate and oxygen respectively [$OD.mmol^{-1}$]

Greek letters

β, γ	tuning parameters
ε	error of the states
μ	specific growth rate [h^{-1}]
μ_{max}, μ_{enh}, μ_{set}	maximum, respectively enhanced, set-point specific growth rate [h^{-1}]
σ	standard deviation
τ_{sensor}	response time for dissolved oxygen sensor [h]

Superscripts and subscripts

^, *EKF*	observed values
..$^{-}$	predicted values
Error! Objects cannot be created from editing field codes.	average values
m	measured values
mod	model values
$_0$, $_k$	initial values, values on time instant *k*

Introduction

Most biopharmaceuticals are produced in a batch or fed-batch cultivation. The quality of the product is formed in this step and is the result of the metabolic state of the micro-organisms. It is therefore essential to measure the physiological state of the process. Metabolic activity is difficult to measure directly due to the lack of sensors, but respiration can be monitored by the oxygen mass balance. The oxygen uptake rate can in turn be used to estimate the specific growth rate and biomass. The specific growth rate and biomass concentration are key parameters.

In this application, production of a vaccine for whooping cough by cultivation of *Bordetella pertussis,* only dissolved oxygen measurements were available. In previous work (Neeleman et al., 2004) a two step asymptotic observer has been used to observe the biomass and specific growth rate. The oxygen uptake rate (*OUR*) was used as input for the observer and in contrast with other observers (see amongst others Bastin and Dochain, 1990) measurement of (part of) the state variables such as biomass and substrates - which are often not available online or require use of additional sensors and equipment - is not used. Although, the performance in Neeleman et al., (2004) was good there was still a need for faster convergence and a higher accuracy. Therefore, a software sensor based on an Extended Kalman filter (EKF) was developed to observe specific growth rate and biomass every minute

using the oxygen uptake rate as input (Soons et al., 2007). The choice for an EKF is motivated, amongst other things, by the fact that the measurement noise of the sensors is well known and this knowledge can be included in the algorithm.

In industrial production environments Extended Kalman filters are uncommon due to the tuning load (Lübbert and Jørgensen, 2001; Jenzsch et al., 2006; Wilson et al., 1998). The EKF may be difficult to tune and can suffer from stability and convergence problems. Another problem is the explicit use of process kinetics, because this kinetics is often highly uncertain (Perrier et al., 2000). This paper demonstrates that the EKF yields good performance for biomass and specific growth rate estimation – even by using a rather generic observer model that does not rely on a kinetic model. An evaluation tool is used to gain (heuristic) insight in the stability of the EKF observer.

In most applications, the oxygen transfer coefficient k_La is measured in advance of the cultivation in medium and is assumed to depend on stirrer speed and volume only. However, the formation of cells, proteins, and other molecules, which absorb at gas-liquid interfaces, cause interfacial blanketing and reduce the oxygen transfer coefficient (Doran, 1995). Because concentrations of cells, substrates, and products change during (fed-)batch cultivation, the value of k_La also changes. Examples of change in k_La due to these factors are given in (Lübbert and Jørgensen, 2001; Sabra et al., 2002; Galaction et al., 2004). Changing k_La causes errors in the *OUR* calculation and the estimation of the specific growth rate and biomass. It is therefore essential to deal with time-varying k_La.

Offline measurements are mostly considered as not suited for control and estimation purposes, because they become available with a delay and at infrequent and irregular times. These measurements however contain valuable information about the states of the system and can make the observer more robust (Dondo and Marqués, 2003). Amongst the literature on bioprocess monitoring (e.g. Bastin and Dochain, 1990) the use of offline information for online estimation is limited. Myers et al. (1995), Tatiraju et al. (1999), and Mutha et al. (1997) use offline and online measurements for state estimation, but do not estimate parameters; Lubenova et al. (2003); Gudi et al. (1995, 1997); Dondo and Marqués (2003); and Ignatova et al. (2003) estimate parameters in addition. Gudi et al. (1995) and Mutha et al. (1997) increase state observability by using the offline measurements multiple times. In Dondo and Marqués (2003), Lubenova et al. (2003), Gudi et al. (1995, 1997), and Ignatova et al. (2003) the parameters are *directly* related to the online measurements and they are part of the measurement and/or state equations. In this work, however, the parameter k_La is neither in the measurement nor in the state equation and is only observable *indirectly* from the offline biomass measurements. Figure 1 shows an example of such types of systems.

Two time scales play a role in the estimation scheme in this work (Fig. 2). This work combines frequent online measurements (which are measured every minute) with infrequent offline measurements (which are sampled and analyzed irregularly and with a delay). These estimates can be used to gain better control of a bioreactor. Figure 2 shows an overview of the system, in which two types of observers are involved to accommodate the two time scales: a *frequent* observer using the online data (the oxygen uptake rate to estimate the specific growth rate and biomass) and an *infrequent* observer activated by sampled offline data (biomass to observer biomass more accurately). The offline biomass measurements are also used to adapt the time-varying k_La in the infrequent adaptive k_La observer.

Figure 1. *Interaction plot for biomass (C_X), specific growth rate (μ), oxygen uptake rate (OUR), maintenance coefficient on oxygen (m_O), and oxygen transfer coefficient (k_La).*

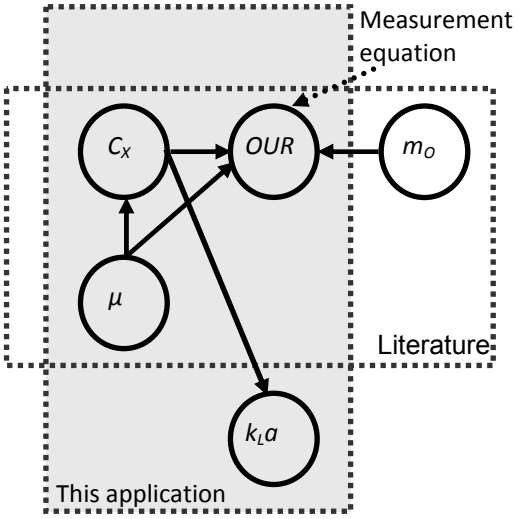
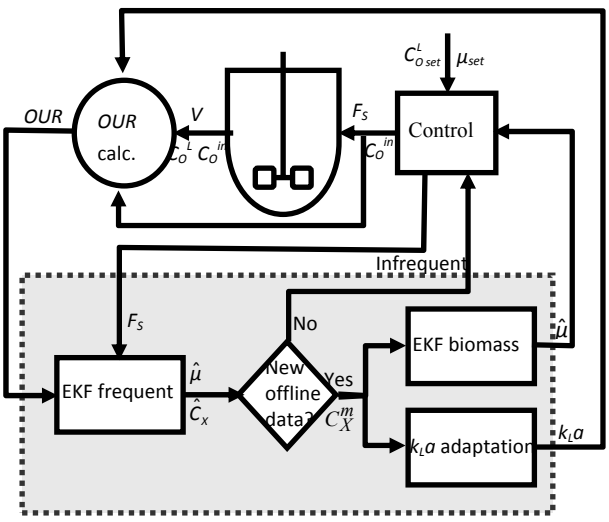


Figure 2. *System configuration.*



Bioreactor model for data generation by simulation

The observer is tested in both simulations and laboratory experiments. The purpose of the simulations is to generate data as if it is reality to tune and evaluate observer performance. A validated model (Eqs. 1-12, see Neeleman et al., 2001; Neeleman, 2002; Soons et al., 2006) is used in the tuning and simulation sections. To obtain realistic tests disturbances have been introduced in the simulation.

Growth of *B. pertussis* is limited by two substrates (Thalen et al., 1999). The model (Eqs. 1-12) applies for batch, fed-batch, and continuous-flow stirred-tank (CSTR) cultivations, depending on the feed going in and out the system. Monod kinetics and oxygen excess are assumed. The parameters in this model were taken from independent experiments (see Neeleman et al., 2001; Neeleman, 2002 and appendix A).

$$\mu(C_G, C_L) = \mu_{\max} \cdot \frac{C_G}{K_G + C_G} + \mu_{enh} \cdot \frac{C_G}{K_G + C_G} \cdot \frac{C_L}{K_L + C_L} \qquad 1$$

$$\frac{dV}{dt} = F_S - F_{out} \qquad 2$$

$$\frac{dC_X}{dt} = \left(\mu(C_G, C_L) - \frac{F_S}{V} \right) C_X \qquad 3$$

$$\frac{dC_G}{dt} = \frac{F_S}{V}\left(C_G^{in} - C_G\right) - \left(\frac{\mu_{\max} \frac{C_G}{K_G + C_G}}{Y_{G1}} - \frac{\mu_{enh} \frac{C_G}{K_G + C_G} \frac{C_L}{K_L + C_L}}{Y_{G2}} + m_G \right) C_X \qquad 4$$

$$\frac{dC_L}{dt} = \frac{F_S}{V}\left(C_L^{in} - C_L\right) - \left(\frac{\mu_{enh} \frac{C_G}{K_G + C_G} \frac{C_L}{K_L + C_L}}{Y_{L2}} + m_L \right) C_X \qquad 5$$

Changes of dissolved oxygen in time are equal to the oxygen transfer rate minus the oxygen uptake rate minus the dilution:

$$\frac{dC_O^L}{dt} = OTR - OUR - \frac{F_S}{V} C_O^L \qquad 6$$

The oxygen uptake rate (*OUR*) is the sum of oxygen used for growth (μ/Y_O) and oxygen used for maintenance (m_O):

$$OUR = \left(\frac{\mu}{Y_o} + m_o \right) C_x \qquad 7$$

The bioreactor is aerated using headspace aeration only. The liquid phase oxygen concentration (C_O^h) in equilibrium with the gas phase in the headspace following Henry's law Neeleman (2002), the dissolved oxygen concentration in the medium (C_O^L), and the oxygen transfer coefficient k_La determine the oxygen transfer rate (*OTR*):

$$OTR = k_l a \cdot (C_O^h - C_O^L) \qquad 8$$

The dynamics of oxygen is much faster than the dynamics of the other relevant processes (e.g. growth of biomass, consumption of substrates) (Wang and Stephanopoulos, 1984) and the contribution of the dilution term is small (third term Eq. 6) compared to the rate of change of dissolved oxygen. As a consequence, Eq. 6 is considered in steady-state and *OUR* is calculated every minute during the experiment using Eq. 9:

$$OUR \approx OTR \qquad 9$$

The oxygen concentration in the headspace is assumed to be equal to the oxygen concentration in the inlet stream because of the high aeration rate:

$$C_O^h \approx C_O^{in} \qquad 10$$

Dissolved oxygen is controlled by changing the incoming oxygen fraction. Dissolved oxygen measured by the sensor is modelled as a first order system:

$$\frac{dC_{O,sensor}^L}{dt} = \frac{C_O^L - C_{O,sensor}^L}{\tau_{sensor}} \qquad 11$$

In practice, dissolved oxygen measurements contain estimation noise. In simulations used for design and tuning, noise is introduced by adding white noise with an intensity proportional to *OUR* (0.1 initially to 3% at end). This intensity was chosen to mimic the observed fact that *DO* is noisier towards the end:

$$C_{O,sensor}^{L,m}(t_k) = C_{O,sensor}^L(t_k) + v(t_k) \qquad 12$$

Extended Kalman Filter

The observer is based on a nonlinear continuous-time model:

$$\frac{dx}{dt} = f(x,u) \qquad y = h(x) \tag{13}$$

With *f* a nonlinear function of the states *x* and inputs *u*, and *y* the output.

Lewis (1986) gives a good explanation of an Extended Kalman filter. The application in biotechnological applications is amongst others discussed in Gudi et al. (1997), Neeleman (2002), and Keesman (2002). The structure of a discrete time EKF is based on the following equations following Lewis (1986):

$$x_{k+1} = A_k x_k + B_k u_k + w_k \qquad y_k = C_k x_k + v_k \tag{14}$$

Where A_k, C_k and B_k at each instant follow from discretization and linearization of Eq. 14 for a time step *T*:

$$F = \left[\frac{\partial f}{\partial x}\right]_{\hat{x},u}, \quad C_k = \left[\frac{\partial h}{\partial x}\right]_{\hat{x},u} \qquad A_k = I + FT, \quad B_k = \left[\frac{\partial f}{\partial u}\right]_{\hat{x},u} T \tag{15}$$

with u_k the input vector and *I* the identity matrix. The initial states x_0 are stochastic variables with average $\bar{x}_0$ and variance P_0: $x_0 \sim (\bar{x}_0, P_0)$; $w_k \sim (0, Q_k)$ is system noise and consists of model errors an unknown inputs; and $v_k \sim (0, R_k)$ is measurement noise. The algorithm has two steps. The time update and the measurement update.

Time update

When a sample becomes available at time *k*, first the time update *k+1* is calculated using the original nonlinear model

$$x^-_{k+1} = f(\hat{x}_k, u_k) \qquad y^-_{k+1} = h(x^-_{k+1}) \tag{16}$$

giving the predicted states x^-_{k+1} and the output y^-_{k+1}. The prediction of the variance of the states P^-_{k+1} is based on the system A_k and system noise (Q_k).

$$P_{k+1}^{-} = A_k \hat{P}_k A_k^T + Q_k \qquad 17$$

Measurement update

Next the measurement update calculates new estimates using the predicted model states x_{k+1}^{-} and outputs y_{k+1}^{-}, the actual measurements y_{k+1}, and the predicted variance P_{k+1}^{-}:

$$\begin{aligned} \hat{P}_{k+1} &= P_{k+1}^{-} - P_{k+1}^{-} C_{k+1}^T R_{k+1} C_{k+1} P_{k+1}^{-} \\ K_k &= \hat{P}_{k+1} C_{k+1}^T R_{k+1}^{-1} \\ \hat{x}_{k+1} &= x_{k+1}^{-} + K_k \left(y_{k+1} - y_{k+1}^{-} \right) \end{aligned} \qquad 18$$

Where the Kalman gain K_k is the magnitude of the correction and aims to minimize the covariance $\hat{P}_k$.

Frequent observer

Observer model

Instead of on the full process model (Eqs. 1-12) the frequent observer is based on a generic model (Eqs. 19 and 20) for biomass growth with μ as an additional state.

$$\begin{aligned} \mu_{k+1}^{-} &= \hat{\mu}_k \\ C_{X,k+1}^{-} &= \hat{C}_{X,k} \cdot e^{\left(\hat{\mu}_k - \frac{F_S}{V}\right)} \end{aligned} \qquad 19$$

$$OUR_k = \left(\frac{\mu_k}{Y_O} + m_O \right) C_{X,k} \qquad 20$$

The model contains only two parameters: the yield (Y_O) and maintenance (m_O) coefficient on oxygen. The performance of the observer depends on accurate knowledge of these parameters. Appendix A gives the calculation of these parameters from separate experiments and shows that the parameters are not dependent on time.

The states x are specific growth rate and biomass (μ, C_X), the output y is the oxygen uptake rate (OUR). The volume V is assumed to be exactly known and doesn't need to be observed. F_S is the feed rate. The prediction of the states follows every instant (one minute) by calculation of Eq. 19, the measurement update from calculation of Eq. 18.

The observability matrix O

$$O_K = \begin{bmatrix} C_k \\ C_k A_k \end{bmatrix} = \begin{bmatrix} \frac{C_{X,k}}{Y_O} & \frac{\mu_k}{Y_O} + m_O \\ \left(\frac{\mu_k}{Y_O} + m_O\right) \cdot C_{X,k} & \left(\frac{\mu_k}{Y_O} + m_O\right) \cdot \left(\mu_k - \frac{F_S}{V}\right) \end{bmatrix} \qquad 21$$

is full rank, which renders the linearized system observable if the oxygen uptake rate and feed rate are not negligible, which is the case under the conditions after inoculation of the cells.

Tuning

In the implementation of the observer, system noise Q can be regarded a degree of freedom for the designer. The initial state covariance matrix P_0 was chosen zero, because simulations with high and low P_0 showed that it converges quickly and only weakly influences the estimation. The covariance of the measurements R is derived from variation in real process measurements for a worst case scenario with a high level of noise. This leaves the elements of Q as the tuning parameters. In general, the Q to R ratio is a measure of confidence in the model against the measurements. A large ratio indicates that the measurements are better than the model, and the measurements are therefore weighted more heavily, and vice versa (Gee and Ramirez, 1996). Given the good predictive performance of the model, we expect Q to be small relative to R. The tuning of Q is based on simulation. The system noise matrix (Q) with Q_μ the intensity of the system noise on the specific growth rate and Q_{C_X} on the biomass

$$Q = \begin{bmatrix} Q_\mu & 0 \\ 0 & Q_{C_X} \end{bmatrix} \qquad 22$$

was chosen in such a way that the deviation (E) between estimated states (or time-varying parameters) (^) and “true” model-based data ($_{mod}$ Eqs. 1-12) is minimal.

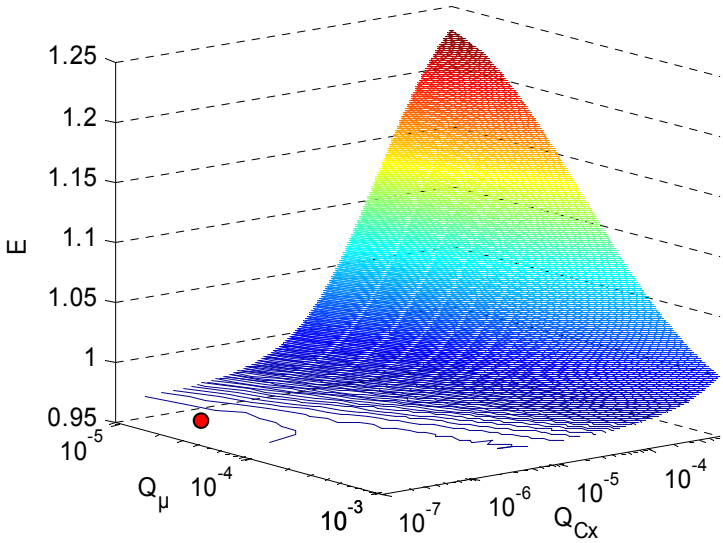
$$\min_{Q_\mu, Q_{C_X}} E = \min_{Q_\mu, Q_{C_X}} \frac{\sqrt{\sum_{k=1}^{N} \left[\hat{\mu}(k) - \mu_{\text{mod}}(k)\right]^2}}{\bar{\mu}_{\text{mod}}(N-1)} + \frac{\sqrt{\sum_{k=1}^{N} \left[\hat{C}_X(k) - C_{X,\text{mod}}(k)\right]^2}}{\bar{C}_{X,\text{mod}} \cdot (N-1)} \qquad 23$$

Both terms are divided by their average to make them equally important.

Figure 3 shows a contour plot for E for a range of Q_μ and Q_{Cx} (note the log scale). The minimum E is flat, indicating the performance of the EKF is hardly sensitive for reasonable errors on the tuning parameters Q_μ and Q_{Cx}. The figure also shows that the system noise on

biomass is small compared to the system noise on specific growth rate in the minimum, as disturbances on biomass are mainly caused by disturbances on μ.

***Figure 3**. Contour lines of the objective function E (Eq. 23) as function of tuning parameters Q_μ and Q_{Cx} during batch cultivation. The red dot indicates the minimum Q^{tune}.*



A heuristic approach to assure the stability of the observer

In the previous section the system noise Q in the EKF is tuned on the basis of a performance criterion. Although stability properties based on a linearized system may have limited utility in a nonlinear context (Sontag, 2001), it does give insight in the effect of the choice of the gain matrix K on the stability of the observer. Since the sampling frequency, and so the linearization frequency, are high compared with the process time-scale we may be able to monitor the stability of the observer by a simple linear method. Other methods, such as input-to-state stability (ISS) (Sontag, 2001), require the unforced system to be stable and this is not the case in our fed-batch system. Indeed, since the observer model is unstable due to the exponentially increasing biomass, ISS cannot be easily applied. Sontag (2001) states that: "perhaps the most interesting set of open problems concerns the construction of feedback laws that provide ISS stability with respect to observation errors.", and so we limit ourselves to a heuristic approach to stability. The error dynamics for the continuous time linearized model are (Ljung, 1979; Dochain, 2003):

$$\frac{d\varepsilon_k}{dt} = \left(A_k(\hat{x}) - K_k(\hat{x})C_k(\hat{x})\right)\varepsilon_k \qquad 24$$

where

$$\varepsilon_k = \hat{x}_k - x_k \qquad 25$$

and A_k and C_k as defined in Eq. 15. The Kalman gains are:

$$K_k(\hat{x}) = [K_{k,\mu}(\hat{x}) \quad K_{k,C_x}(\hat{x})]^T \qquad 26$$

The linearized system is stable if $A_k(\hat{x}_k) - K_k(\hat{x}_k)C_k(\hat{x}_k)$ have strictly negative real eigenvalues and if $\hat{x}_k$ is assumed constant over the (short) sampling interval [k,k+1):

$$\mathrm{Re}\left(\lambda_i[A_k(\hat{x}_k) - K_k(\hat{x}_k)C_k(\hat{x}_k)]\right) < 0 \,\forall\, \hat{x}_k \; and \; i = 1 \; to \; n \qquad 27$$

with n the dimension of the states x. The Routh stability criterion (see e.g. Nise, 2000) is applied to gain insight in stability of the observer as a means to arrive at analytical expressions for the Kalman gain matrix elements that guarantee stability on the time-interval [k,k+1). After substituting the analytical expressions for the linearized system (Eqs. 19-20) in the Routh stability criterion, we find

$$K_{k,C_x} > \frac{Y_O \cdot \mu_k \cdot V_k - Y_O \cdot F_S - K_\mu \cdot C_{X,k} \cdot V_k}{V_k\left(\mu_k + Y_O \cdot m_O\right)} \qquad 28$$
$$K_{k,\mu} > 0$$

Infrequent observer

The infrequent observer updates the biomass and k_La estimation using infrequent offline and frequent online measurements (Fig. 2). Since inaccurate values of k_La may cause biased observations for the frequent observer, k_La is also updated when offline samples are available.

The following three-step algorithm is designed. First the frequent observer estimates μ and C_X using frequent online available data. Next, if samples are available, the biomass observer estimates biomass and finally the adaptive observer estimates k_La. The main advantage of this approach is that convergence of k_La- adaptation is determined by one tuning parameter which can be enforced *a priori* as will be shown in the sequel.

EKF biomass

The biomass observer contains the following discrete model, in which the sample and analysis delay is assumed to be constant and is incorporated by addition of *n* fictitious delayed states (*n = 10*) (Gudi et al., 1997):

$$\begin{aligned} C_{X_1,k+1} &= \left(1 + T(\mu_k - \frac{F_{S,k}}{V_k})\right) C_{X_1,k} \\ C_{X_2,k+1} &= C_{X_1,k} \\ &. \\ &. \\ C_{X_{n+1},k+1} &= C_{X_n,k} \end{aligned} \qquad 29$$

where *T*, as before, is the frequent sampling interval. The states of the biomass observer are the (delayed) biomass concentrations ($C_{X1}, C_{X1}, \ldots C_{Xn+1}$); the output is biomass delayed with *n* frequent samples (C_{Xn+1}). The prediction of the states follows every infrequent instant from calculation of Eq. 29, the measurement update from calculation of Eq. 18 (if an offline sample is available, Fig. 2).

The observability matrix has full rank. Tuning and stability analysis are performed similar to the frequent observer by optimizing the covariances of *Q*. The extra states represent a shift in time for which the covariances elements (Q_{Cx2} to Q_{Cxn+1}) are assumed zero. This leaves optimization of the first element in *Q* (Q_{Cx1}) only.

Adaptive k_La observer

During cultivation the microorganisms and their products may affect the physical properties of the liquid and as a result k_La may change. Deviations between the measured and actual biomass are assumed to be caused by errors in k_La. The adaptive observer adapts k_La when samples are available. The relative correction *c* is calculated every infrequent instant and assumed constant until the next sample becomes available:

$$
\begin{aligned}
&\textit{if} \;\; C_{X,k}^{m} \;\; \textit{exists} \\
&\quad k_L a_k^{Stir} = c \cdot f(V, R^{Stir}) \\
&\quad k_L \hat{a}_k = k_L a_k^{Stir} + \gamma (C_{X,k}^{m} - \hat{C}_{X,k}) \cdot T_m \\
&\quad c = \frac{k_L \hat{a}_k}{k_L a_k^{Stir}} \\
&\textit{else} \\
&\quad k_L a^{Stir} = c \cdot f(V, R^{Stir}) \\
&\quad k_L \hat{a}_k = k_L a_k^{Stir} \\
&\quad c = c
\end{aligned}
\tag{30}
$$

Depending on the oxygen demand during the cultivation, agitation speed is adjusted. The oxygen transfer coefficient k_La is measured in advance of the cultivation in biomass-free medium and depends on the stirrer speed and on the volume as specified in the function $f(V,R^{Stir})$ in Eq. 30. This function is used to adjust k_La *a priori* for known changes stirrer speed.

The tuning parameter γ determines the convergence speed of the observer. T_m is the infrequent sampling interval, which can be irregular and infrequent.

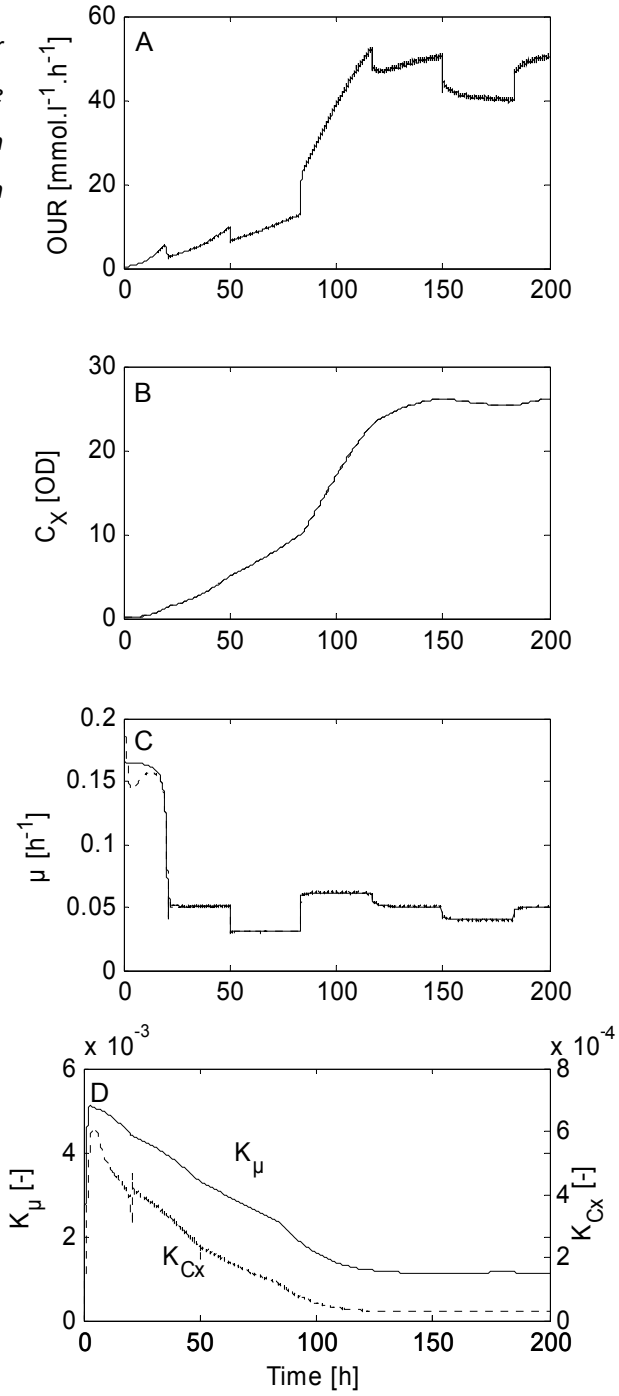
Results and discussion

Simulation results and discussion

Figure 4 shows a long-term simulation of a fed-batch cultivation using the frequent observer. The simulation consists of a batch and fed-batch phase. The batch cultivation lasts until the limiting substrates are depleted at about $t = 20h$; next, the fed-batch phase takes place, in which different levels of specific growth rate were achieved by adjusting the feed rate in order to evaluate tracking/convergence of the observer.

Figure 4A gives the *OUR* obtained in this system. The course of this variable varies strongly due the induced variations in the specific growth rate. The *OUR* signal is used as an input for the observer to estimate biomass (C_X) and specific growth rate (μ).

***Figure 4**. Simulation using the frequent observer for fed-batch cultivation with Q^{tune} **A.** Oxygen uptake rate data generated by the simulation **B.** Biomass from simulation and estimation. **C.** Specific growth rate from simulation and estimation. **D.** Kalman gains.*



Biomass and specific growth rate were initialized with +20% errors (Fig. 4BC). The observed specific growth rate and biomass converged to the real simulated values within 15 hours. During the fed-batch (*t = 20-200h*) the observed biomass and specific growth rate correspond to the data generated by the model.

Changes in specific growth rate were induced five times during the fed-batch phase (Fig. 4C). The EKF accurately tracks these changes. The Kalman gains converge to non-zero values (Fig. 4D) preventing the observer to become “lazy” and therefore guaranteeing good tracking performance, also during long-term experiments.

For the purpose of the test, additional simulations were performed with up to a five-fold increase respectively decrease in maximum specific growth rate (figures not shown). Tuning parameters Q and R were taken identical to the previous simulations. Performance was as good as the previous case, showing that the EKF can be applied to faster or slower-growing micro-organisms without retuning.

The effect of uncertain maintenance and yield coefficients was also tested by simulations. The simulations showed that the observed specific growth rate converges to the actual value if maintenance and/or yield coefficients on oxygen were uncertain. Biomass estimations, however, tend to be biased if these coefficients are not properly known. If yields and maintenance coefficients are time-varying, this effect should be incorporated. In the case of cultivation with *B. pertussis*, however, this effect is not under discussion, because the coefficients are accurately known and time-varying behaviour of these coefficients has not been observed (Appendix A).

The observer model (Eq. 19) is unstable due to the exponentially increasing biomass. Applying the Routh criterion (Eq. 28) showed that the EKF stabilized the system for $t = 18\text{-}200h$ and that the error system was not stable in the first 18 hours. Nevertheless the observer performed well. To enhance stability the values of the systems noise Q can be increased (putting more trust on the measurements compared to the model). However, as a consequence the system becomes more noise sensitive. To avoid increased noise sensitivity over the whole operational time, the Kalman gains can be adapted during the period of instability ($t = 0\text{-}18h$) on the basis of the conditions derived from the Routh criterion. This stabilization method shows good performance in simulations. Relative to using a fixed Q^{tune}, the performance criterion E (Eq. 23) is only 10% larger using the stabilization based on the Routh criterion (Eq. 28), but E is 40% larger using an increased Q chosen such that the EKF was stable over the whole fed-batch

In practice, there is a little more room for tuning the parameters without jeopardizing stability, because the stability analysis is based on a simplified and linearized model for biomass growth (Eq. 19), which can be different from the real nonlinear process.

Simulations with the infrequent observer showed that the biomass observer was not stable due to the lack of frequent offline samples (for a case in which samples were available about

every two hours as in the experiments in the next section). If frequent samples (e.g. every 20 minutes) are available, the observer does become stable after an initialization phase of several hours. Also November and Van Impe (2002) state that the convergence rate is highly correlated with the sample time.

Experimental results and discussion

Fed-batch and CSTR cultivations with the dual substrate consuming bacterium *Bordetella pertussis* were performed with glutamate and L-lactate as the main carbon sources. The cells were grown in a seven-litre bioreactor containing 2.88 litres medium for the CSTR cultivations and 3 to 4 litres medium for the fed-batch cultivations. A six-bladed impeller was used to agitate the medium. Temperature was controlled at 34ºC, agitation speed at 300 to 750 rpm depending on the oxygen demand, and dissolved oxygen at 30% air saturation by headspace aeration only (by changing the incoming oxygen fraction in an oxygen/nitrogen mixture). The total gas flow was kept constant at 1 l/min. Further materials and methods (bioreactor conditions, medium, analysis, software, and hardware) were applied as reported by Soons et al. (2006).

In bioreactors, aerated by a high air flow entering the headspace, the difference between the inlet and exhaust oxygen fraction is small and can therefore not be measured accurately. Hence *OUR* is calculated using the oxygen balance in the liquid phase:

$$OUR \approx k_L a \cdot (C_O^h - C_O^L) \qquad 31$$

The CSTR experiments were performed to evaluate the frequent EKF only, the fed-batch experiments to evaluate the combination of frequent EKF, infrequent EKF, and k_La adaptation. Fed-batch or CSTR cultivations started with batch cultivation. Next, the fed-batch phase started automatically when the limiting substrates were depleted; the CSTR phase started when the biomass reached an optical density of 1.0. Before the cultivation k_La was observed in organism-free medium for a range of stirrer speeds and bioreactor volumes. All experiments were performed using a fixed value for the system noise Q^{tune}, because good performance was obtained in simulation experiments.

Maintenance and yield coefficients were calculated on the basis of a series of four CSTR experiments with low biomass (Appendix A). Blanketing effects are therefore small so that k_La can be assumed independent of biomass and product formation, which allows calculation of proper maintenance and yield coefficients.

Figure 5 shows that the observed biomass coincides well with the offline biomass measurements during CSTR. The observed specific growth rate converged exactly to the preset dilution rate ($D = F_S.V^{-1}$). Uncertainty on μ was small throughout the cultivation.

Blanketing effects however do occur during fed-batch cultivation with higher biomass concentrations, resulting in overestimation of the biomass if only the frequent observer is used (Fig. 6). Figure 6 compares the frequent observer with the infrequent observer for fed-batch cultivation. An exponentially increasing feed was added to the bioreactor to keep the specific growth rate constant, resulting in an increase in volume and a decrease in k_La; sampling causes a small decrease in volume and therefore slight increase in k_La. The abrupt changes of k_La were caused by changes of stirrer speed (to meet the increasing oxygen demands). Biomass was overestimated using only the frequent observer (Eqs. 19 and 20). Biomass was observed more accurately when the infrequent observer (Eqs. 19, 20, 30, and 31) corrected the k_La and C_X estimations for biomass and product formation.

The performance of the EKF is also compared to an asymptotic observer designed in previous work (Neeleman et al., 2004) and proved to be significantly better by showing faster convergence and a higher accuracy and stability.

Figure 5. *Frequent observer for CSTR cultivation* ***A.*** *Biomass.* ***B.*** *Dilution rate, estimated specific growth rate and its standard deviation (σ).*

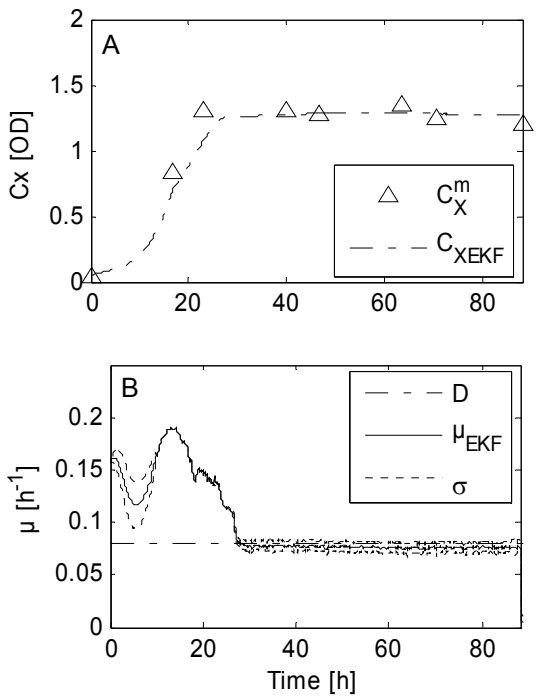
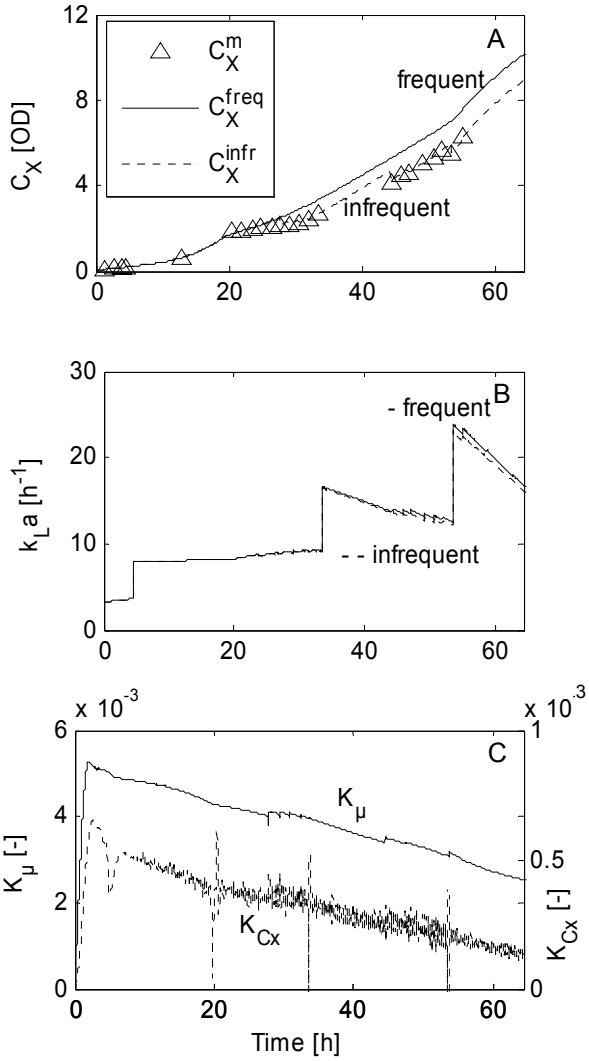


Figure 6. *Combined frequent and infrequent observers for fed-batch cultivation.* ***A.*** *Biomass.* ***B****. Oxygen transfer coefficient.* ***C.*** *Kalman gains.*

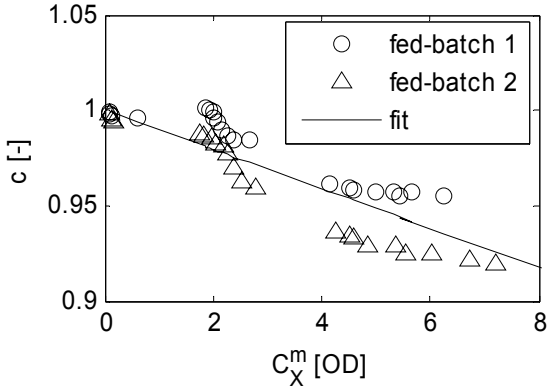


Performance of the infrequent observer was good even when sparser biomass measurements were available. Fewer measurements will result in larger sampling intervals T_m and larger k_La adaptation per offline measurement (Eq. 30). The effect of more sparse measurements on observer performance will therefore be small, however, the observer becomes more noise sensitive when the measurements contain a high level of noise. Obviously, some measurements are required to find the proper k_La pattern.

Figure 7 shows the values of the relative k_La ($c = k_L\hat{a} / k_La^{Stir}$) for two fed-batch cultivations at different biomass concentrations. The relative k_La decreased by up to 8% for a cultivation grown to 8 *OD* due to biomass and product formation. This effect can be modelled by a linear relation (d=0.01, straight line in Fig. 7):

$$c(C_X) = 1 + d \cdot C_X \qquad 32$$

Figure 7. *Relative k_La (k_La during fed-batch compared to k_La in pure medium, $c = k_L\hat{a} / k_La^{Stir}$) as function of the biomass concentration for two fed-batch experiments.*



For high cell density cultivation (e.g. *E. coli*), in which cells can grow up to approximately 150 *OD*, k_La will drastically decrease. It illustrates the potential of incorporating this effect in the estimation (assuming that the maintenance and yield coefficients are constant or time-varying in a known way).

Conclusions

Application of the presented observer design and tuning rules improves bioprocess monitoring in biopharmaceutical production. The observer consists of two parts operating on two time-scales: a *frequent* observer that estimates specific growth rate and biomass online; and an *infrequent* observer, activated by offline measurements, that estimates the biomass and the oxygen transfer coefficient online. The observer model is based on generic equations for biomass growth and is therefore widely applicable. Tuning is based on a performance criterion that minimizes the error between the generated bioreactor data and the estimations on biomass and specific growth rate. Fine-tuning is based on the Routh stability criterion and enhances stability by forcing the Kalman gains to stable values.

Simulations and experimental data showed that the EKF converged to the real values for specific growth rate and biomass and is robust for tuning and initialization errors, kinetics, and observation noise.

Performance of the observers depends on exact knowledge of the maintenance and yield coefficients, which are accurately known for the cultivation with *B. pertussis*.

Changes in oxygen transfer coefficients caused by increasing biomass and product concentrations are accurately observed. The observer thereby gives new information on the model for the time-varying k_La. This ability offers potential for high cell density cultivations.

Acknowledgments

Wageningen University and the Netherlands Vaccine Institute (NVI) work together in a project with Applikon Biotechnology BV and Siemens NV in order to improve the production process by release of biopharmaceuticals on the basis of new online techniques. Senter supports the project under project name TSGE3067. Laboratory experiments were performed at NVI, The Netherlands.

Appendix A

Since the performance of the observer depends on the exact knowledge of the maintenance and yield coefficients on oxygen, separate experiments were performed to calculate these parameters. In contrast to batch or fed-batch cultivation, CSTR cultivations allow independent calculation of these parameters. Figure A1 shows the relation of the specific oxygen consumption (q_{O2}) to μ as in (Pirt, 1982). These data allow accurate calculation of the maintenance and yield coefficients ($R^2=0.98$). Obviously, they were constant and independent on the specific growth rate. These values were used in the simulation and observer model: m_O = 0.68 $mmol.OD^{-1}.h^{-1}$, Y_O = 0.033 $OD.mmol^{-1}$.

In some bioprocesses, e.g. high cell density cultivations, the maintenance and yield coefficients tend to be time-varying due to a shift in carbon for maintenance metabolism. If present, this time-varying nature can be observed from the oxygen consumption in relation with the biomass in the following way:

The overall yield is defined as the ratio between the biomass production and the oxygen consumption rate (Van't Riet and Tramper, 1991):

$$Y_O' = -\frac{r_{C_X}}{r_O} = \frac{\frac{dC_X}{dt}}{OUR} \qquad \text{A1}$$

Integration of Eq. A1 gives Eq. A2 if Y_O' is constant:

$$Y_O' = \frac{C_X}{\int OUR \ dt} \qquad \text{A2}$$

Note that changes of the specific growth rate in time influence the overall yield, but can be assumed negligible, since μ is approximately constant in batch and constant in fed-batch. Figure A2 shows that the overall yield Y_O' is constant. So, the yield and maintenance

coefficients were neither dependent on μ nor on time for the cultivations with *B. pertussis* in this work.

Figure A1. *Calculation of the biomass yield and maintenance on oxygen using data on specific oxygen consumption rate and specific growth rate, R^2 = 0.98.*

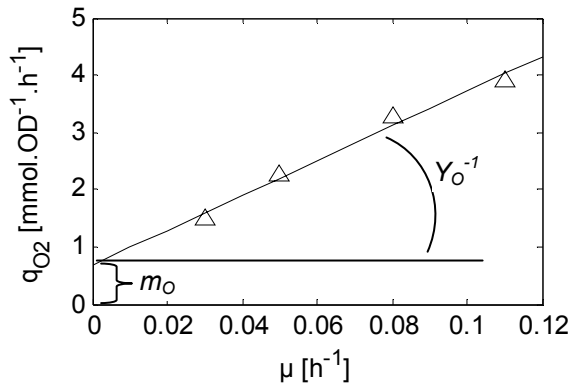
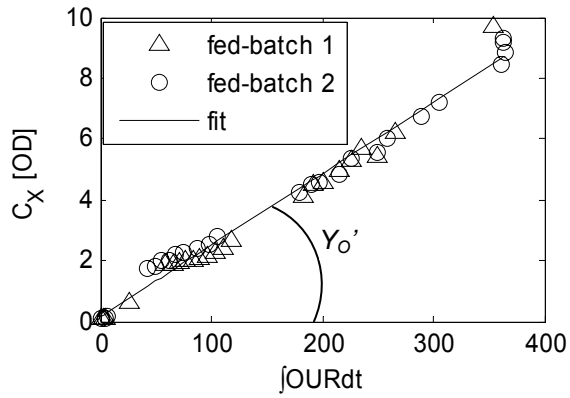


Figure A2. *Biomass versus integral oxygen uptake rate.*



CHAPTER 3

Towards PAT monitoring and control: near infrared and software sensor

Submitted to Chemometrics and Intelligent Laboratory Systems: Z. I. T. A. Soons, M. Streefland, G. van Straten, A. J. B. van Boxtel. Towards PAT monitoring and control: near infrared and software sensor

Abstract

Spectroscopic instrumentation is often seen as promising for process analytical technology (PAT) to enhance control of manufacturing (bio)pharmaceuticals. The interpretation of near infrared spectra is challenging due to the large number of wavelengths recorded and the overlapping absorbance features of near infrared spectroscopy. This work applies a controlled random search procedure to select an optimal window of wavelengths giving a good calibration model for biomass concentrations during cultivation of *Bordetella pertussis*, the causative agent of whooping cough. The proposed wavelengths selection procedure outperforms the traditional calibration procedures. In the second half of the paper, the near infrared based predictions are compared with the estimations obtained from a software sensor for biomass and specific growth rate based in standard measurements of oxygen consumption. The near infrared predictions depend on the quality of the training dataset, which needs to encompass all possible sources of temporary disturbances like pH and dissolved oxygen. The accuracy and robustness of the near infrared predictions with the current linear partial least squares model are less favourable than those of the software sensor. Although near infrared has the potential to provide more information than just biomass, the software sensor is the preferred choice for feedback control of biomass and specific growth rate.

Keywords

wavelengths selection, partial least squares; near infrared spectroscopy, *Bordetella pertussis*, software sensor, process analytical technology.

Nomenclature

b	calibration coefficients
C_X	biomass concentration [*OD*]
C_O^L	oxygen concentration in bioreactor [$mmol.OD^{-1}.h^{-1}$]
C_O^*	oxygen concentration at gas-liquid interface [$mmol.OD^{-1}.h^{-1}$]
DO	dissolved oxygen fraction
E	spectral residuals
EKF	Extended Kalman filter
f	biomass residuals
F	substrate feed rate [$l.h^{-1}$]
J	number of wavelengths
k_La	oxygen transfer coefficient
m_O	maintenance coefficient on oxygen [$mmol.OD^{-1}.h^{-1}$]
nIR	near infrared
N	number of nIR measurements
OD	optical Density at 590nm in 1ml broth
OUR, OTR	oxygen uptake rate, oxygen transfer rate [$mmol.l^{-1}.h^{-1}$]
P	loadings matrix
PLS	partial least squares
q	weight matrix
R	number of latent variables
RSS	residual sum of squares
T	scores matrix
V	broth volume [*l*]
W	weight matrix
w	weights
X	predictor matrix, near infrared measurements
Y	output matrix, biomass concentrations
Y_O	biomass yield on oxygen [$OD.mmol^{-1}$]
μ	specific growth rate [h^{-1}]

Superscripts and subscripts

m	measured values
^, OUR	observed values using *OUR* measurements
$..^-$	values calculated from time update software sensor

nIR predicted using near infrared measurements

Introduction

Cultivation processes usually generate numerous offline and online data. Amongst the offline data are product quality assessments and biomass measurements. Online measurements are obtained from e.g. sensors for dissolved oxygen, pH, temperature, redox, or near infrared spectroscopy. Standard measurements, such as dissolved oxygen, temperature, and pH are not used to evaluate process and product performance, but rather serve to assure certain physical or thermodynamic properties. The offline measurements are often more relevant for product quality, but they become available with a (large) delay (Karim et al., 2003) and hence are not available for real-time monitoring and control purposes. Vaccines, for instance, are produced in batch cultivation and released on the basis of offline tests on the final product, e.g. potency tests requiring animal models, involving large variability, delayed release or off-spec losses, and high costs. The challenge is to extract biologically relevant information from online measurements and to control these in real-time.

To stimulate understanding, monitoring and control of manufacturing processes, the FDA released the PAT guidance (FDA, 2004). Wold et al. (2006) defined four levels of chemometrics analysis for PAT ranging from multivariate calibration to combining several online-monitored critical process variables in real-time to asses quality of the final product. Here, the biomass and the specific growth rate are regarded as key variables that need to be controlled online, because biomass growth influences production of virulence factors (Licari et al., 1991; Rodriguez et al., 1994), which are indicators for vaccine quality. The natural fifth level would be PAT-control: including real-time feedback based on the online monitored variables. Wold et al. (2006) states that “PAT process control” in sense of feedback control is still a challenge of the future. Nevertheless, feedback control in the frame of PAT have already been reported (Gnoth et al., 2007; Soons et al., 2006). The use of spectroscopic instrumentation, although, is often seen as promising for PAT (e.g. Lopes et al., 2004; Triadaphillou et al., 2007; Wold et al., 2006), in the frame of feedback control it is less well developed. This paper describes a step towards real-time feedback control for PAT by comparing two technologies for online monitoring of biomass growth during the cultivation step in vaccine production for *Bordetella pertussis*. The first is based on online near infrared spectroscopy; the second is a software sensor using measurements of oxygen consumption based on observer theory from the field of control engineering.

Near infrared spectroscopy is often addressed in the context of PAT and is widely used to predict the chemical composition of raw materials (Lopes et al., 2004). Abundant literature is available about near infrared spectroscopy for bioprocess monitoring. Scarff et al. (2006) provide an overview on calibration of near infrared models for cultivation of micro-organisms; Roggo et al. (2007) on near infrared spectroscopy and chemometrics in pharmaceutical production.

The near infrared sensor generates a spectrum that consists of a large number of transmittance signals for wavelengths between 833 and 2500 nm at a high frequency. The evolvement of the spectra in time contains information on the course of the cultivation process. Partial least squares (PLS) is commonly used to regress a response variable (component concentrations) on a set of predictor spectra (near infrared measurements) (Sprang et al., 2007). Not all wavelengths in the spectrum, however, contain information relevant to the component of interest. Moreover, absorbance ranges of different components may overlap and substances in a complex mixture may contribute to signals that are spread across the complete spectral range. To counteract these effects, wavelength selection is a common tool to gather wavelengths that do contain relevant information. The complexity of the near infrared spectrum, however, precludes to some extend the use of the classical wavelength selection approach. Most works select spectral regions on the basis of prior knowledge on the absorbance of specific chemical components (amongst others Sprang et al., 2007 and Vaidyanathan et al., 1999). As an alternative, automated search methods are suggested by amongst others Triadaphillou et al. (2007). A further overview of automatic search methods is given in the section “automatic wavelengths selection”. The currently available wavelength search methods are not effective in finding global minima. To overcome this problem, in this paper, the controlled random search procedure presented by Price (1977) is applied to obtain the most informative wavelengths.

The other technology to which near infrared spectroscopy is compared, is a software sensor that uses data from standard oxygen consumption measurements. Software sensors are based on observer techniques which follow from the discipline of systems and control theory (Lewis, 1986; Ljung, 1979). Major work on software sensors for bioprocess control is presented by Bastin and Dochain (1995) and Stephanopoulos and San (1984), more recent work by e.g. Gudi (1997) and Keesman (2002). The design and application of a software sensor for specific growth rate control during fed-batch cultivation of *Bordetella pertussis* is given by Soons et al., (2006, 2008).

Cimander and Mandenius (2004) show the potential of near infrared spectroscopy in real-time feedback control, but did not yet compare it to other techniques like observer based

software sensors. In this paper the monitoring performance of these sensors is compared and from this conclusions are drawn for the potential of feedback control of the specific growth rate.

Summarizing, this work concerns two aspects: first, improving the linear fit of near infrared spectra and biomass measurements by enhanced automatic wavelengths selection; and second, comparing the merits of near infrared spectroscopy versus an oxygen uptake rate based software sensor for controlling specific growth rate.

Materials and Methods

Measurements

Online measurements

A polarographic electrode (Applikon, the Netherlands) was used for measuring dissolved oxygen in the medium; a pH electrode (Mettler Toledo, Udorf, Switzerland) for pH; and a Pt100 for temperature. A Bruker Optics Matrix F near infrared sensor with 5mm path length and 4cm^{-1} resolution was mounted in the broth of the bioreactor to obtain transmittance near infrared spectra. Each spectrum contained 2074 wavelengths, calculated as the mean over 16 scans. The measurements of the online sensors were stored every minute.

Offline measurements

Offline measurements were gathered approximately every hour. Glutamate and L-lactate concentrations were measured offline with a YSI 2750 select analyzer (Yellow Springs Instruments, Yellow Springs, USA). Biomass was obtained by measuring the optical density (OD) at 590 nm of 1ml suspension using a Vitalab 10 (Vital Scientific, the Netherlands).

Bioreactor conditions

Data of the near infrared sensor and data from other sensors (dissolved oxygen, pH, etc.) together with offline measurements of biomass concentration are taken from eight batch cultivations. The cultivations with the dual substrate consuming bacterium *Bordetella pertussis* were performed in a 7-liter bioreactor containing four litres of medium with glutamate and L-lactate as the main carbon sources (Thalen et al., 1999). A six-bladed

turbine impeller was used to agitate the medium. All cultivations were performed using identical operating conditions. Temperature was controlled at 34ºC and dissolved oxygen was kept at 30% air saturation by first increasing agitation speed from 450 to 650 rpm and next increasing the oxygen fraction in the headspace. The total gas flow was kept constant at 1 l/min. We refer to Van Sprang et al. (2007) for more details on the cultivation methods.

The nutrient and biomass profiles were similar for all experiments. In four batches periods of deviations from the intended profiles were introduced to enhance the model robustness in counteracting the natural deviations of bioprocesses and to test PLS prediction performance:

- PAB0003: Dip in dissolved oxygen concentration around *t = 2h*.
- PAB0004: Poor pH control at *t = 0-10h*.
- PAB0005: Biased biomass measurements at *t = 15h* and *t = 17h*.
- PAB0007: Dissolved oxygen limitation and lowered pH at *t=0-9h*.

Estimation procedure

Spectral Data Pre-Treatment

All spectra were used to predict the biomass evolution in time during the batch cultivation. The calibration of the biomass against near infrared spectra was done in three steps.

- The first step of the spectral data pre-treatment was to partition the batches into two groups, a training set (for calibration) and a validation set (for prediction). Six batches were assigned for training (PAB0003, PAB0005, PAB0006-1, PAB0006-2, PAB0007, and PAB0009-2), two for validation (PAB0004 and PAB0009-1). This distribution was chosen to make sure that batches with temporary deviations are present in both sets.
- The second pre-treatment step was to apply Savitsky-Golay smoothing with a 45-point window and a second order polynomial to reduce noise, whilst maintaining signal information content. The noise reduction was found particularly important in the later stages of the cultivation where vibrational effects caused by agitation and gas phase effects cause observation noise on the near infrared signal.
- The third step was a consequence of the offline measurements of the biomass concentration, which were sampled approximately 15 times per batch cultivation. Triadaphillou et al., (2007) and Montague and Martin (2007) interpolate the offline measurements to obtain values of concentrations at the same sampling times as the spectral measurements. To avoid the introduction of interpolation errors, we

preferred to use the near infrared spectra at the sample times of the offline measurements for calibration, leaving *k*=90 samples for calibration and 28 for validation.

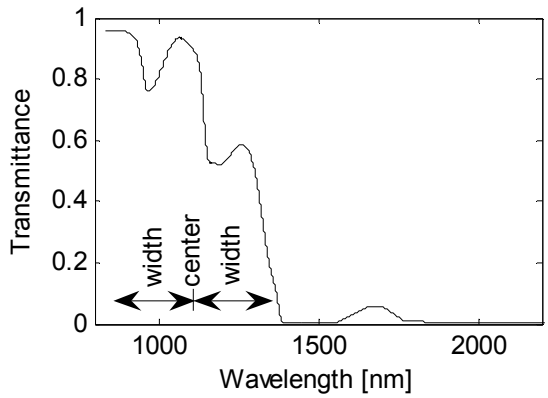
Partial Least Squares (PLS)

The assumed model is that the biomass *Y* depends linearly on some of the wavelengths of *X*, which is composed of a spectral block with a width and center (Fig. 1). The choice for partial least squares over e.g. multiple least squares is motivated by the possibility to reduce the large amount of data and to calculate weights such that the slopes at the uninformative wavelengths are close to zero (Eqs. 5-6). The presence of observation noise, however, degrades the performance of the calibration model. This is counteracted wavelengths selection to select the window of wavelengths that is informative for biomass concentrations.

The calibration dataset is used to build the PLS model for biomass on the near infrared spectra in the *calibration* step. Both the calibration and the validation dataset are used to predict biomass concentrations (based on the PLS model) using all available near infrared measurements in the *prediction* step.

The matlab toolbox LIBRA contains a library of robust statistical methods (Verboven and Hubert, 2005). RSIMPLS is a function that performs partial least squares (PLS) and is used for calibration of the biomass model. Linear PLS is a quite common technique and is outlined briefly in the sequel. Further details can be found in e.g. Haaland and Thomas (1988) or Helland (1988).

Figure 1. *Example of spectral width and center selection.*



Calibration

The dataset consists of an output vector *Y* (Nx1) containing the *N* biomass concentrations over time and a predictor matrix *X* (NxJ) containing the spectral measurements. The number of columns J is equal to the number of wavelengths. The number of wavelengths J used for PLS calibration depends on the width of the wavelengths window, which is selected in the controlled random search procedure (Fig. 1):

$$J = 2 \cdot width + 1 \qquad 1$$

or may contain all wavelengths (2074) if the full spectrum is used. The matrix *X* is decomposed as the product of the scores *T* (NxR) and the loadings *P* (JxR) plus the spectral residuals *E* (NxJ); the matrix *Y* as the scores *T* (containing R columns or latent variables) and a weight matrix *q* (Rx1) plus the biomass residuals *f* (Nx1):

$$X = TP^T + E \qquad 2$$

$$Y = Tq + f \qquad 3$$

$$T = XW(P^T W)^{-1} \qquad 4$$

where *W* (JxR) is an orthogonal weight matrix, calculated from a least squares solution.

Prediction

PLS prediction using an unknown spectrum involves the calculation and use of the final calibration coefficients *b* (Jx1)*:*

$$b = W(P^T W)^{-1} q \qquad 5$$

to predict the response variables *Y* (biomass) online:

$$\hat{Y} = Xb \qquad 6$$

Automatic Wavelength Selection

Abrahamsson et al. (2003) compares four methods of automatic wavelength selection: genetic algorithms (GA), iterative PLS (IPLS), uninformative variable elimination by PLS (UVE-PLS), and interactive variable selection for PLS (EVE-PLS). Problems observed by these authors were: lack of convergence to the same optimal solution (GA); disappointingly small improvements (less than 7% for IPLS and UVE-PLS); lack of freedom to place the windows anywhere in the wavelength range (e.g. the windows start with 1, 101, 201, etc.) and slow

progress (IVS-PLS). To overcome the problem of different results for different runs in genetic algorithms, Ferreira and Cardoso de Menzes (2007) propose a weighted genetic algorithm, but still multiple runs are necessary to select the wavelengths with the highest average weight to incorporate in the PLS model; so not yet an optimal solution.

Triadaphillou et al. (2007) use a spectral window selection (SWS) algorithm to obtain multiple models by selecting spectral windows centers and widths based on random increments. The models are next combined by stacking. Although the method selects similar wavelengths more often than genetic algorithms, optimal solutions are not obtained.

Price (1977) developed a controlled random search procedure, which is effective in searching for global minima using an iterative procedure. This method is applied here to automatically select the center and width of a spectral window (Fig. 1) and the number of latent variables *R* (in the scores *T*) on the basis of good prediction of biomass. The window of wavelengths may vary in size, ranging from a single wavelength to 221 wavelengths.

The CRS algorithm consists of three steps:

- A set of *N* trial points is generated at random using *n* inputs (number of latent variables *R* and wavelengths width and center). The weighted residual sum of squares (RSS)

$$RSS = \sqrt{\frac{\sum_{i=1}^{k} \frac{\left(C_{X,k}^{m} - C_{X,k}^{nIR}\right)^2}{w_k}}{k}} \qquad 7$$

  is evaluated for each trial point *M* and stored in an array *A* together with the inputs; w_k is a weight indicating whether the data point is an outlier or not; *k* indicates the biomass samples, $C_{X,k}^{m}$ the biomass measurements, and $C_{X,k}^{nIR}$ the biomass concentrations fitted using the near infrared spectra in the PLS model ($\hat{Y}$ in Eq. 6).
- A new point is generated by choosing *n+1* random distinct points R_1, R_2, … R_{n+1} from the set of *N* stored points. The next trial point *TP* is computed from the centroid *G* of the *n* points $R_,$,… R_n minus the last point R_{n+1}:

$$TP = 2 \cdot G - R_{n+1} \qquad 8$$

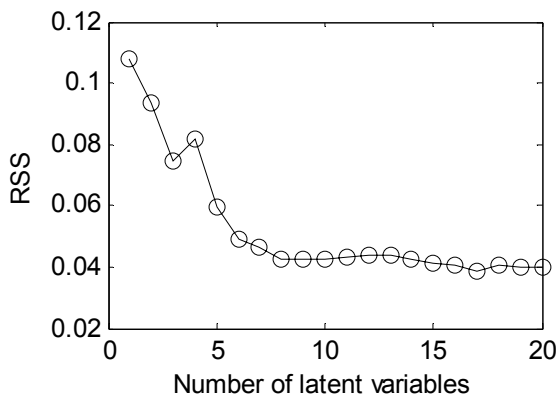
- The stored point *M* with the greatest RSS is determined. The RSS_{TP} is evaluated in point *TP* and compared with the RSS_M in point *M*. If $RSS_{TP} < RSS_M$, *M* is replaced by *TP* in *A*.

Step two and three are repeated until the stop criterion is satisfied (a maximum number of function evaluations or the maximum RSS is smaller than a certain value). For more details on the method we refer to Price (1977).

Constraints on the number of latent variables are applied to prevent over-fitting. Figure 2 shows an example on the choice of the number of latent variables *R*. From this figure, 17 latent variables give the smallest RSS and thus without constraints 17 latent variables would be selected by the algorithm. With a constraint of nine latent variables, eight latent variables would be selected, and the figure shows that with this constraint the RSS is only slightly larger (0.042 compared to 0.038).

The number of windows was fixed at one as it was found that two or more windows did not improve the PLS model.

***Figure 2**. Number of latent variables selection plot*



Software sensor for biomass growth

The software sensor uses a process model and process measurements to estimate noise-reduced variables that can not be measured online. The software sensor is based on an Extended Kalman Filter (EKF) and estimates the specific growth rate ($\hat{\mu}$) and biomass ($\hat{C}_X$) using the online oxygen uptake rate (*OUR*) measurements (Soons et al., 2008). Lewis (1986) and Ljung (1979) give a good explanation of an Extended Kalman filter. The application in biotechnological applications is amongst others discussed by Stephanopoulos and San (1984), Gudi et al (1997), Keesman et al. (2002), and Neeleman and van Boxtel (2001). The EKF calculations are recursive as shown schematically in Fig. 3. When a sample becomes available at time instant *k*, first the time update is calculated using a nonlinear model; next the measurement update calculates new estimates using the model, the actual

measurements, and variance *P*. Depending on the accuracy of the model and of the measurements, the measurement update relies more, or less on the model compared to the measurement. The applied EKF is based on a model with only two parameters (m_O and Y_O) that are required to be known accurately. It estimates the evolution of the state variable biomass (C_X) and the extended state specific growth rate (μ):

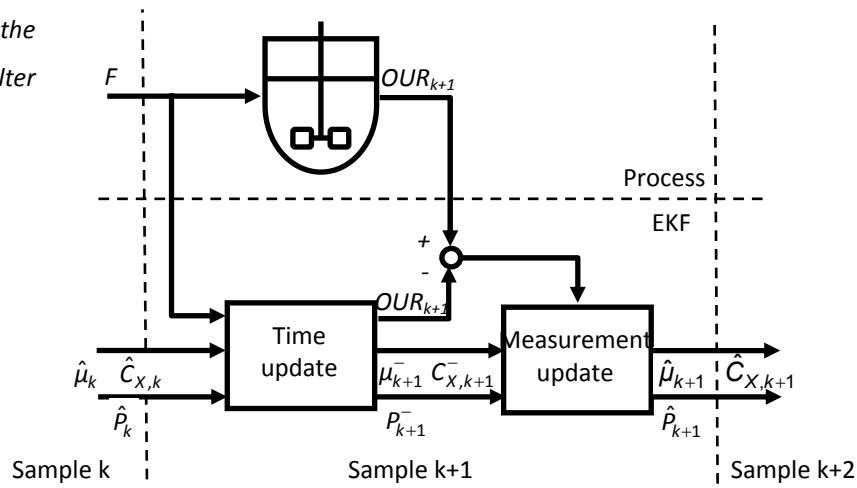
$$\frac{dC_X}{dt} = \left(\mu - \frac{F}{V}\right)C_X$$
$$\frac{d\mu}{dt} = 0 \qquad 9$$
$$OUR = \left(\frac{\mu}{Y_O} + m_O\right)C_X$$

where *F* is the feed rate and *V* the liquid volume in the bioreactor. The oxygen uptake rate is the input for the EKF and is calculated every time instant using Eq. 10 (e.g. Wang and Stephanopoulos, 1984):

$$OUR \approx OTR = k_L a(C_O^* - C_O^L) \qquad 10$$

where C_O^* and C_O^L are the oxygen concentrations at the gas-liquid interface and in the cultivation broth.

Figure 3. *Schematic of the Extended Kalman Filter calculations.*



Results

Experimental data

Figure 4A shows the typical evolution of a batch cultivation captured by the near infrared analyzer. The increasing biomass concentration gives more absorption and hence a decreasing transmittance signal as the batch evolves. During the major part of the batch cultivation a linear relationship between the transmittance signal and the biomass concentrations is a plausible assumption. Above a certain biomass concentration, saturation effects occur and the near infrared-biomass relation becomes nonlinear, which is approximately indicated by the dotted line. Eventually at even higher cell densities the transmittance signal would become negligible.

Transmittance at 800-1400 nm was linear during the main part of the batch cultivation. Wavelengths above 1400nm were in the nonlinear range of the spectrum. Higher cell densities that would occur in fed-batch cultivations would give even lower transmittance signals, indicating that the chosen path length of the near infrared sensor is not be appropriate for such situations. A limitation of current submerged near infrared measurement is that the path length cannot be adjusted online to anticipate the changing biomass concentrations during batch operation. Nonlinear techniques, such as a transformation of the near infrared data (Robertsson, 2001), might be applied to capture the complete cultivation from inoculation to high cell densities. Due to the limited number of samples in this range, nonlinear PLS is not feasible here.

The software sensor based on the calculation of the oxygen uptake rate using dissolved oxygen measurements (Fig. 4B) gives appropriate signals for the whole range of possible operations and can be applied without any modification to fed-batch cultivations (Soons et al., 2008).

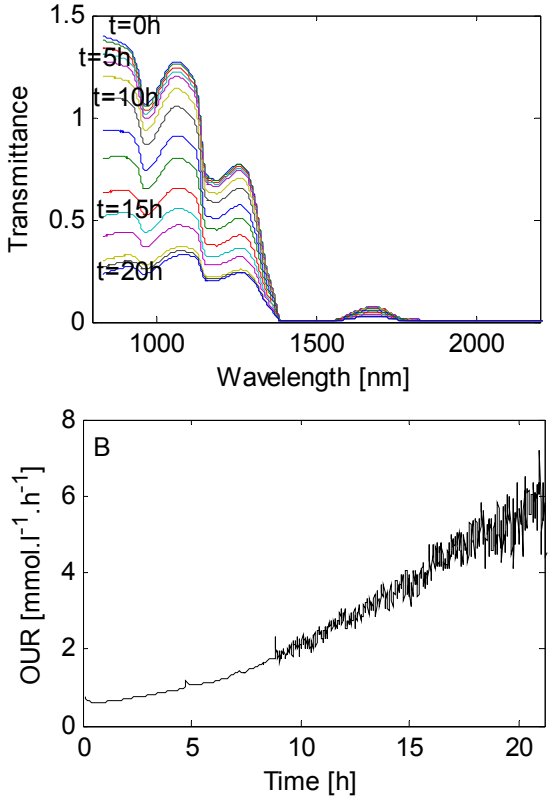


Figure 4. *Measurement signals of both methods:* ***A.*** *Typical evolution of the transmittance near infrared spectra for one batch run (PAB0004). The shown spectra are taken at the time instants where also offline measurements are available.* ***B.*** *Typical evolution of the oxygen uptake rate for a batch run (PAB0009-2).*

Controlled random search of wavelengths

The controlled random search procedure was applied several times to the training dataset and converged to the same spectral window and number of latent variables during each run. Figure 5 indicates that a window of 976-1053 nm gives the best prediction of the biomass concentration. The selection of a wide window consisting of 195 wavelengths makes that the model is not too specific for the training dataset and is able to predict biomass concentrations on other batches.

Figure 6 shows the measured (C_X^m) versus the predicted biomass concentrations (C_{XnIR}). The results show good correspondence for the low and medium values of biomass in the training set. Predicted and measured values in the high biomass range deviate due to non-linearity in the transmittance. Out of six batches in the training dataset three contained “natural” deviations in pH and/or dissolved oxygen (PAB0003, PAB0005, and PAB0007). Despite of these hick-ups the near infrared model is capable to estimate biomass.

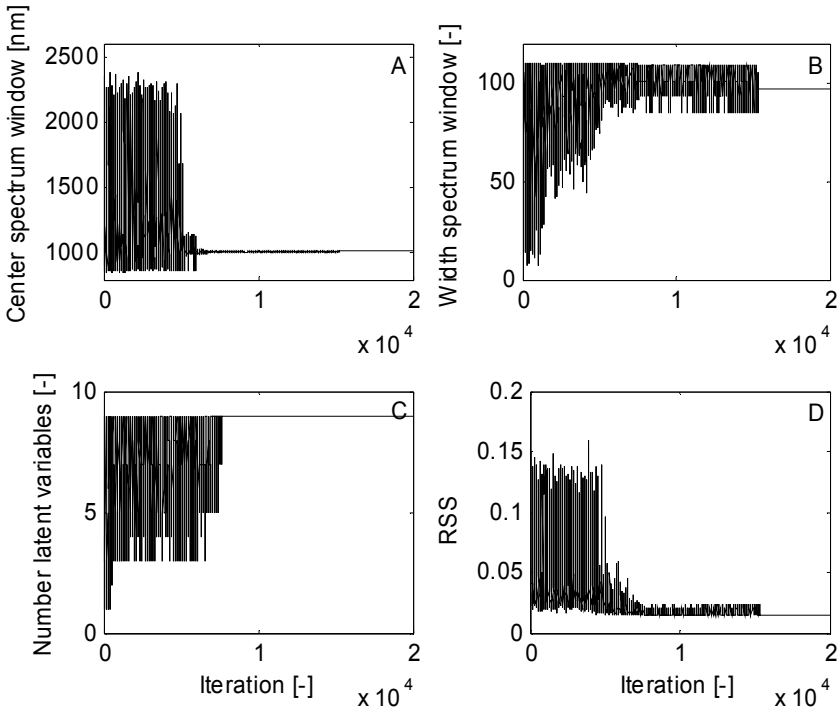


Figure 5. *Automatic wavelength selection by applying a controlled random search on the training dataset for the trial points.* ***A.*** *Center spectrum window* ***B.*** *Width spectrum window* ***C.*** *Number of latent variables* ***D.*** *Weighted residual sum of squares.*

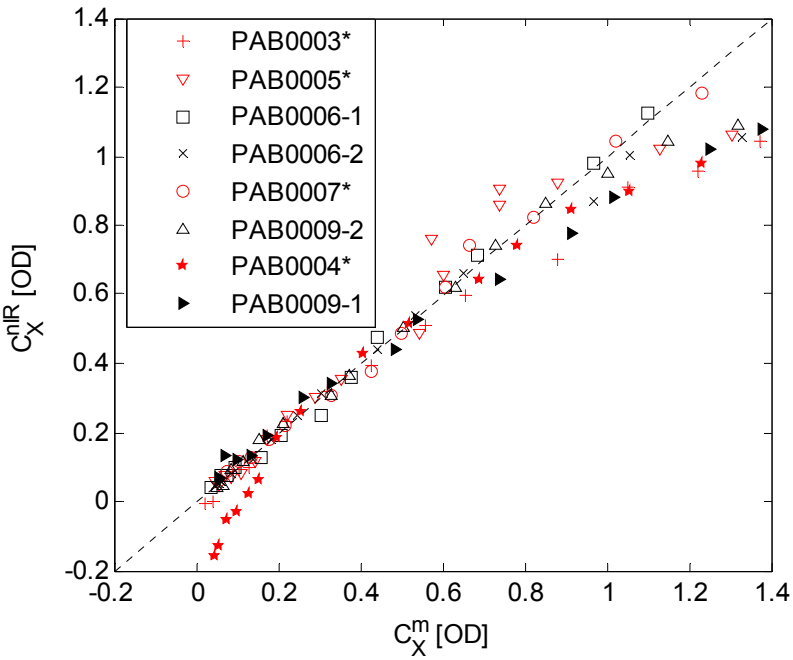


Figure 6. *Measured biomass concentrations versus concentrations predicted using the near infrared sensor. The training set is shown in open symbols, the validation set in filled symbols (PAB0004 and PAB0009-1); the batches in red and indicated with an asterix (*) contain temporary deviations in pH and dissolved oxygen, the batches in black do not contain deviations.*

Experiment PAB0004 of the validation set contained a temporary deviation that was not in the training set. pH control in the first ten hours of batch cultivation was poor and as a result the prediction error of the biomass concentration was significant in the initial phase of the cultivation. So, if deviations are not present in the calibration set, estimations are likely to be inaccurate. For future calibration of PLS models, it is important to assure that the training

dataset contains the natural variability that might be expected during production runs; for instance by performing “design of experiment” batches.

The performance of the PLS model obtained from the controlled random search procedure was compared with a PLS model built on the full spectrum and on wavelengths chosen from literature using the same training dataset (Table I). The PLS model for the whole wavelength region shows moderate estimation errors. This model is probably not representative for the biomass alone, but for the entire cultivation broth as a whole. Vaidyanathan et al. (1999) reported that biomass of five different micro-organisms absorbs prominently between 2270 and 2350 nm and 1650-1800 nm. Our model predictions based on the wavelengths between 2270 and 2350 nm were poor due to the absorbance of water and other components in those regions, so that the transmittance signal was negligible (Fig. 4). Selection of the wavelengths 1650-1800 nm weakly deteriorated the prediction accuracy of the PLS model, because the transmittance was in the nonlinear range. The selection of single wavelengths by Cimander and Mandenius (2004) also deteriorated the estimation for the same reason. The manual selection of Sprang et al. (2007) improved the estimation of biomass with 17%. Further improvements up to 35% were obtained by automatically selecting wavelengths using the controlled random search procedure.

Table I. *Comparing different PLS calibration results for the training dataset.*

Selected wavelengths [nm]	Method	RSS [OD]	Improvement [%]	Number of latent variables *R*
833-2400	Full spectrum	0.0223	-	9
2270-2350	Manual selection (Vaidyanathan et al., 1999)	0.1358	-509	7
1650-1800	Manual selection (Vaidyanathan et al., 1999)	0.0248	-11	9
1498, 1590, 1724, 1732, 1794	Automatic selection (Cimander and Mandenius, 2004)	0.0422	-89	3
1111-1397	Manual selection (Sprang et al., 2007)	0.0185	17	8
976-1053	Automatic wavelength selection (Controlled random search)	0.0146	35	9

Comparison near infrared monitoring and *OUR* based software sensor

To test the usefulness of near infrared monitoring during batch cultivations the near infrared biomass estimates are compared to those of the software sensor for three batch cultivations (Figs 7-11). The biomass estimations (C_X^{nIR}) can in turn be used for online estimation of the specific growth rate, which is useful in situations where the specific growth rate is controlled (Soons et al., 2006). Previous research showed that the software sensor accurately observed both the biomass and the specific growth rate from the oxygen uptake rate (Soons et al., 2008). In Figs. 7-9 it can be seen that the sensor is hardly affected by the presence of hick-ups, and that it performs well over the entire range from start to end. Also the normally distributed noise in OUR shown in Fig. 4 was effectively filtered out by the software sensor.

In the cultivations that were consistent with the cultivations of the training dataset the near infrared based estimator estimated biomass growth fairly well. Significant deviations, however, occurred during cultivations containing unseen deviations. Even negative estimations occur during the initial phase of the cultivation when pH control failed to track the set-point (Fig. 9). The near infrared estimations deteriorated at the end of the cultivation due to nonlinearities in the transmittance signal. Furthermore, the estimations were relatively inaccurate and noisy in the initial phase of the cultivation due to the low absorbance level (Figs. 10-11). The software sensor was more accurate in this low range.

The erratic biomass estimations from the near infrared sensor has important consequences for specific growth rate estimation and control. The variations were not normally distributed and therefore more difficult to filter out, resulting in a delayed reconstruction of the specific growth rate (and biomass). These erratic properties of the spectroscopic biomass growth monitoring deteriorate performance in feedback control.

Figure 9B shows the consequences of the biased and variable biomass estimations using the near infrared data. The estimated specific growth rate was initialized at the expected value of 0.16 h^{-1} for all cultivations. As a result of the negative biomass estimations and the required slow filtering, the specific growth rate converged first to negative values and subsequently increased only slowly as the cultivation proceeded. It is obvious that these estimates would give very poor regulation if they were used for feedback control.

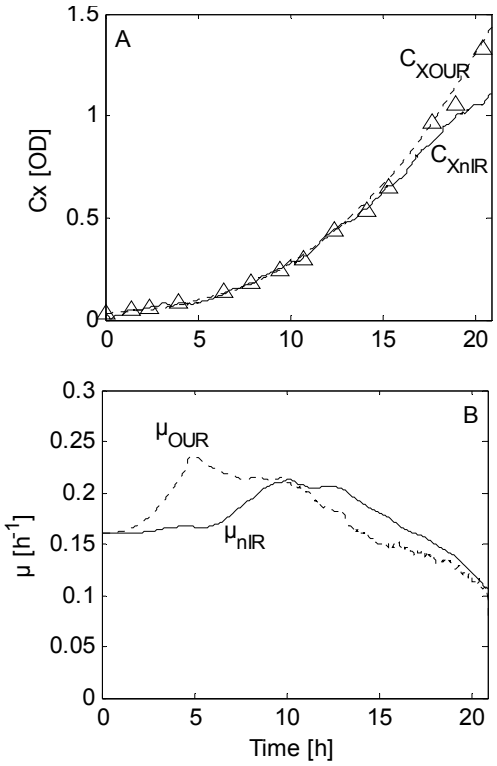


Figure 7. *Bioprocess monitoring based on the near infrared and on the dissolved oxygen sensor for PAB0006-2 of the training set.* ***A.*** *Biomass* ***B.*** *Specific growth rate.*

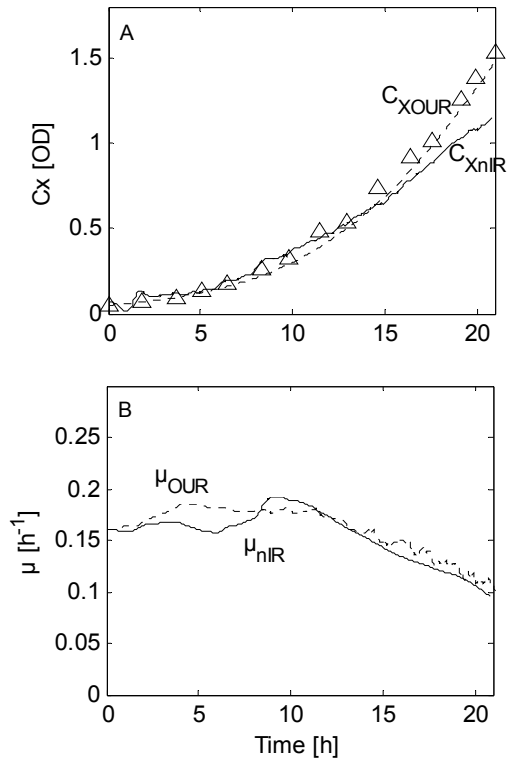


Figure 8. *Bioprocess monitoring based on the near infrared and on the dissolved oxygen sensor for PAB0009-1 of the validation set.* ***A.*** *Biomass* ***B.*** *Specific growth rate.*

Figure 9. *Bioprocess monitoring based on the near infrared and on the dissolved oxygen sensor for PAB0004 of the validation set.* ***A.*** *Biomass* ***B.*** *Specific growth rate.*

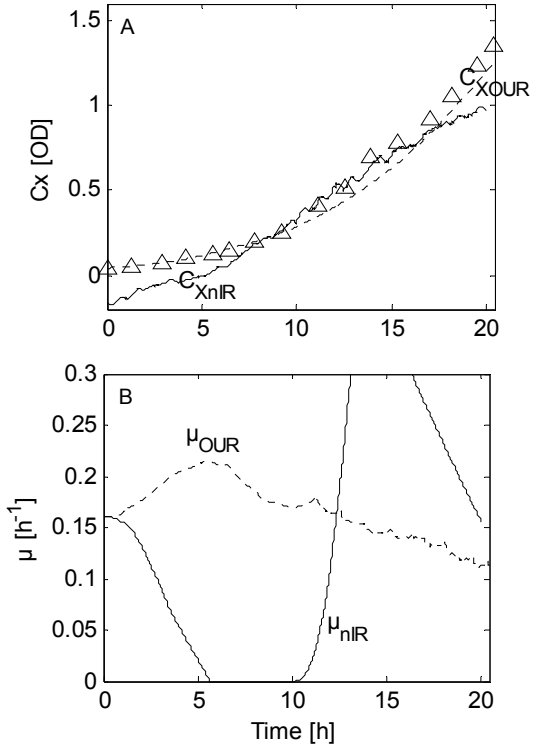


Figure 10. *Detail of Fig. 7. Biomass estimations in the initial phase of the batch cultivation for PAB0006-2 (training set).*

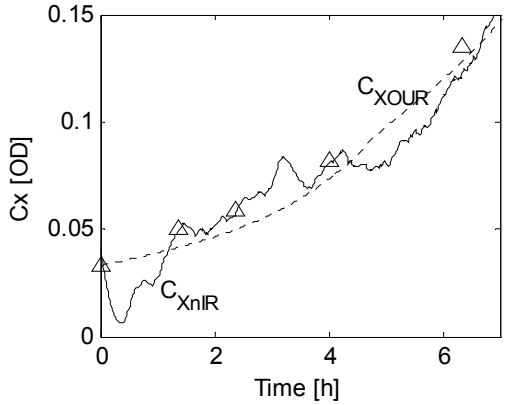
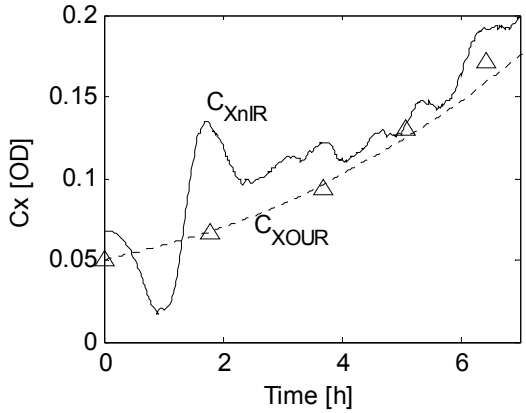


Figure 11. *Detail of Fig. 8. Biomass estimations in the initial phase of the batch cultivation for PAB0009-1 (validation set).*



Discussion and conclusions

Whilst the accuracy of the PLS model has improved by application of wavelength selection, offsets were found in batches, in which deviations in pH and/or dissolved oxygen occurred (Fig. 9). So, if near infrared estimations are intended for feedback control, special care must be taken. It is suggested to check whether the behaviour of the current batch is consistent with the past experience to detect deviating behaviour, e.g. by performing a principal component analysis on the online process data. One might suspect that an obvious improvement of the near infrared regression can be achieved by ensuring better control over batches, so that there are less hick-ups. However, in that case information about the robustness of the method in situations that do happen to occur in practice is lost. Seen this way, a better alternative may be to try to incorporate dissolved oxygen and pH as independent factors in the regression, perhaps in combination with nonlinear regression methods based on a larger number of batches. In the current application, noise was reduced by applying Savistky-Golay smoothing. The small delay introduced by this method might be reduced by improving the accuracy of the spectral data by taking a higher number of scans, but as this can only be done at the expense of a larger sampling interval, the benefit is probably small.

Considering the current limitations of near infrared monitoring and the challenges to be solved, such as a fixed path length, a linear PLS model, and a limited number of batches, at present the software sensor is the preferred choice for monitoring and feedback control of biomass and specific growth rate. This finding is in line with the work of Triadaphillou et al. (2007) who state: “spectroscopic instrumentation will more likely be used for retrospective analysis and development stages of a product than for direct feedback control purposes in the near future. The proposed spectral analysis strategy serves to enhance accuracy by which key component concentrations can be determined but practical implementation and interpretation challenges remain to be addressed.” Additional advantages of the software sensor are the use of a standard and cheap sensor for dissolved oxygen and the fact that only a few batches are needed to obtain the required maintenance and yield coefficients on oxygen. The software sensor is furthermore less sensitive for deviations of other operational conditions and nonlinearities. The near infrared data, on the other hand, have potential in monitoring other main components in the cultivation broth (e.g. substrates) and in monitoring consistency of the overall production process compared to a predefined standard. The use of data from two different types of sensors can yield complementary information and can thereby safeguard monitoring if one sensor fails.

Wold et al. (2006) state that “PAT process control” is still a challenge of the future. This work is intended to be a step towards real-time feedback control for PAT by implementing two technologies for online monitoring of biomass growth during the cultivation step in vaccine production for *B. pertussis* with a view on control.

Acknowledgements

The Dutch Ministry of Economical Affairs (TSGE3067) funds the project. We would like to thank Remko van Brakel for the helpful discussion about PLS modelling and Bas Speetjens for providing the matlab code of the controlled random search procedure. The experiments were performed at the Netherlands Vaccine Institute

CHAPTER 4

Constant specific growth rate in fed-batch cultivation of Bordetella pertussis using adaptive control

Published as: Z. I. T. A. Soons, J. A. Voogt, G. Van Straten, A. J. B. Van Boxtel (2006), Constant specific growth rate in fed-batch cultivation of *Bordetella pertussis* using adaptive control, Journal of Biotechnology 125: 252-268.

Abstract

Monitoring and control of production processes for biopharmaceuticals have become standard requirements to support consistency and quality. In this paper, a constant specific growth rate in fed-batch cultivation of *Bordetella pertussis* is achieved by a newly designed specific growth rate controller.

The performance of standard control methods is limited because of the time-varying characteristics due to the exponentially increasing biomass and volume. To cope with the changing dynamics, a stable model reference adaptive controller is designed which adapts the controller settings as volume and biomass increase. An important asset of the design is that dissolved oxygen is the only required online measurement.

An original design without considering the dissolved oxygen dynamics resulted experimentally in oscillatory behaviour. Hence, in contrast to common believes, it is essential to include dissolved oxygen dynamics. The robustness of this novel design was tested in simulation. The validity of the design was confirmed by laboratory experiments for small-scale production of *Bordetella pertussis.* The controller was able to regulate the specific growth rate at the desired set-point, even during a long fed-batch cultivation time with exponentially increasing demands for substrates and oxygen.

Keywords

model-reference adaptive control, biopharmaceuticals, fed-batch cultivation, specific growth rate, dissolved oxygen, *Bordetella pertussis*

Nomenclature

a, b, c, d	constants for dual substrate model of *B. pertussis*
C	nominal value of controller [$mmol.l^{-1}$]
C_G	glutamate concentration [$mmol.l^{-1}$]
C_{G0}	initial glutamate concentration [$mmol.l^{-1}$]
C_G^{in}	glutamate concentration in the feed [$mmol.l^{-1}$]
C_L	lactate concentration [$mmol.l^{-1}$]
C_{L0}	initial lactate concentration [$mmol.l^{-1}$]
C_L^{in}	lactate concentration in the feed [$mmol.l^{-1}$]
C_X	biomass concentration [OD]
C_{X0}	initial biomass concentration [OD]
$\hat{C}_X$	software sensor biomass concentration [OD]
DO	dissolved Oxygen concentration in the medium [$mmol.l^{-1}$]
DO_{set}	set-point for oxygen concentration in the medium [$mmol.l^{-1}$]
DO_{sensor}	oxygen concentration measured by the sensor [$mmol.l^{-1}$]
E	objective function
E_μ	relative variation of μ
E_{DO}	relative variation of DO
f_G	Monod kinetics for glutamate
f_L	Monod kinetics for lactate
F_{O2}	(enriched) airflow through the headspace [$l.h^{-1}$]
F_{G+L}^{in} F_{tot}	total substrate feed rate (glutamate + lactate) [$l.h^{-1}$]
F_1	“proportional” correction substrate feed rate [$l.h^{-1}$]
F_2	feed rate by prior calculation [$l.h^{-1}$]
F_3	“integral” correction substrate feed rate [$l.h^{-1}$]
ISE	integral squared error [h^{-1}]
K_1, K_2	gains for specific growth rate control
K_C	gain for dissolved oxygen control
K_C^h	gain for headspace control
K_G	Monod constant on glutamate [$mmol.l^{-1}$]
K_L	Monod constant on lactate [$mmol.l^{-1}$]
k_La	oxygen transfer coefficient [h^{-1}]
m_G	maintenance coefficient on glutamate [$mmol.OD^{-1}.h^{-1}$]
m_L	Maintenance coefficient on lactate [$mmol.OD^{-1}.h^{-1}$]
m_O	maintenance coefficient on oxygen [$mmol.OD^{-1}.h^{-1}$]

OD	optical Density at 590nm [$OD_{590}.ml^{-1}$]
OTR	oxygen transfer rate [$mmol.l^{-1}.h^{-1}$]
OUR	oxygen uptake rate [$mmol.l^{-1}.h^{-1}$]
OUR_{noise}	noise on the oxygen uptake rate [$mmol.l^{-1}.h^{-1}$]
O_2^a	auxiliary oxygen concentration [$mmol.l^{-1}$]
O_2^h	oxygen concentration in the headspace [$mmol.l^{-1}$]
O_2^{in}	oxygen concentration in the incoming air [$mmol.l^{-1}$]
PRN	pertactin
rpm	rounds per minute
t	cultivation time [h]
V	liquid volume [l]
$\hat{V}$	software sensor liquid volume [l]
V_h	volume of the headspace [l]
v, w	constants for normalised Monod equations
Y_{G1}	yield on glutamate over pathway 1 [$OD.mmol^{-1}$]
Y_{G2}	yield on glutamate over pathway 2 [$OD.mmol^{-1}$]
Y_L	yield on lactate [$OD.mmol^{-1}$]
Y_O	yield on oxygen [$OD.mmol^{-1}$]

Greek letters

β	ratio between normalised Monod equations
β_1, β_2	convergence speed of reference model
γ_1, γ_2	tuning parameters for MRAC
μ	specific growth rate [h^{-1}]
$\hat{\mu}$	software sensor specific growth rate [h^{-1}]
$\hat{\mu}_0$	initial software sensor specific growth rate [h^{-1}]
μ_{enh}	enhanced specific growth rate [h^{-1}]
μ_{max}	maximum specific growth rate [h^{-1}]
μ_{set}	set-point specific growth rate [h^{-1}]
τ_I	reset time for dissolved oxygen control [h]
τ_I^h	reset time for headspace control [h]
τ_{sensor}	response time for dissolved oxygen sensor [h]

Introduction

Monitoring and control of production processes for biopharmaceuticals have become standard requirements to support consistency and quality. Recently, FDA encourages with its view on Process Analytical Technology (PAT) the introduction of new technology, amongst others methods for monitoring and control, to improve manufacturing and quality assurance in pharmaceutical processes (FDA, 2004).

Currently, most biopharmaceuticals are produced in batch cultivation, where cells grow until the main nutrients are depleted. In such a cultivation system, only dissolved oxygen (DO), stirrer speed, pH, and temperature are controlled. In a typical batch cultivation, specific growth rate cannot be controlled. With the decreasing substrate concentration in time, the specific growth rate (μ) correspondingly changes and eventually becomes zero due to depletion.

In order to ensure a high level of batch-to-batch consistency, it is preferable to control metabolic activity. However, it is not possible to measure metabolic activity directly. To cope with this limitation, indirect measurements can be used. Pörtner et al. (2004) apply the oxygen mass balance to monitor respiration. Levisauskas et al. (1996) make use of the fact that the biosynthesis of many intracellular components is closely related to the particular specific growth rate. Control of the specific growth rate as a measure for metabolic activity, implies constant substrate concentrations and thus, partly, control of the environment of the cells. So, to realise specific growth rate control a feed of substrate is necessary and as a consequence the batch-wise operations must be shifted to a fed-batch operation mode.

Antigen level (Westdijk et al., 1997) and lipopolysaccharide content (Rodriguez et al., 1994) are important aspects for the quality of *B. pertussis* suspensions used for vaccine production against infection with whooping cough. Westdijk et al. (1997) demonstrated that, during batch cultivation the growth phase determines the antigen production and release. No conclusion, however, is given about the relation between antigen production and specific growth rate or nutrient composition.

In literature, limited information is available on the effect of specific growth rate on product quality. Rodriguez et al. (1994) and Licari et al. (1991) show that the production of pertussis toxin (which is one of the important antigens) is strongly growth-associated and that a high specific growth rate is an effective way for producing pertussis toxin. Although this favours the traditional batch cultivation, the work of Rodriguez also indicates that with increasing growth rate lipopolysaccharide production increases. This is regarded as a drawback because lipopolysaccharide is suspect of causing adverse reactions and should therefore be reduced

in the vaccine. So, in order to obtain a high quality-vaccine, it is important to restrict lipopolysaccharide content of the vaccine, while still providing efficient pertussis toxin. This is possible by controlling the cultivation at constant level for the specific growth rate. Pörtner et al. (2004) implemented several feeding strategies for fed-batch cultures and discussed the characteristics:

- fixed feed trajectory,
- *a priori* calculation of the feed trajectories based on a kinetic model,
- predictive control with feedback in intervals,
- feedback control via *OUR*.

The first two methods are open-loop methods and do not correct for deviations, arising from model mismatches. These options are not suitable to control metabolic activity. The third method is a predictive control method with a prediction update interval equal to the interval between manually taken samples. So, feedback takes only place at the sampling moments. In bioreactors, the time between samples may go up to several hours, thus jeopardizing the effectively of the feed-back. The fourth method, control via *OUR*, has the capability of responding fast to deviations from the set-point, regardless of the cause and duration of the disturbance. Metabolic activity is regulated regardless of external disturbances and varying yields and kinetics by adjusting the feed rate every minute. Feedback control using *OUR* measurements is applied in this work because it is able to correct errors, does not require intensive process development, and is applicable in all areas, provided that *OUR* can be measured (Pörtner et al. (2004)).

Neeleman (2002) and Neeleman et al. (2004) used a standard “PI” controller to control the specific growth rate during fed-batch cultivation of *Bordetella pertussis*. However, the performance of such standard control methods is limited due to time-varying characteristics of a fed-batch process. To incorporate exponentially changing biomass and volume, the controller needs to have adaptive properties.

In the literature on adaptive control two approaches can be distinguished. The adaptation of controller settings in the first group is based on controller performance, where the error between desired and actual behaviour (which can be caused by uncertainties in the kinetic parameters or mismatch in the total model) is the driver for adaptation. Babuška et al (2003) apply this method to control dissolved oxygen, Frahm et al (2002) to control substrate concentrations, and Akay et al (2002) to control medium temperature in bioreactors. To achieve enough adaptation in this approach, errors must be introduced by applying systematic excitation signals to the system. The other group of adaptive controllers is

represented by e.g. Bastin and Dochain (1990), Van Impe and Bastin (1995), and Smets et al. (2002, 2004). In this approach, the adaptive expressions for the controller parameters are linked to measured or estimated process states.

The work of Chang (2003) fits to the second group. Gain scheduling for a PI^2D controller is based on the measured states. Levisauskas et al. (1996) apply the same methodology by using a controller that adapts its parameters to the changing states. A drawback of these two controllers is the requirement for explicit measurement of part of the states (e.g. substrates, biomass, or ethanol). Alvarez and Simutis (2004) present a control law for regulation of the specific growth rate for a single substrate model, where estimation of specific growth rate and biomass is performed by a Kalman filter. This work is based on simulations, and is not qualified by laboratory experiments.

Applying excitation signals is not desired in cultivation systems for biopharmaceutical production, because it may affect critical variables. In this work, therefore, adaptation according to the second category is used to cope with the time-varying biomass and volume. Robustness and accuracy are evaluated by simulations with time-varying or drifting kinetics, changes in set-point, and external disturbances and by evaluating controller performance in laboratory experiments with the dual substrate-using *B. pertussis*.

In this work, bioreactor control is based on two loops: the standard *DO* control loop (which was not allowed to change) and the adaptive specific growth rate control loop (Fig. 1). Oxygen and substrate consumption are strongly coupled and in the same order of magnitude. If dissolved oxygen dynamics are ignored, tuning of the specific growth rate controller may lead to oscillations (Fig. 2). So, in contrast to common believes, incorporation of dissolved oxygen dynamics is essential for the design of specific growth rate control in order to avoid interactions between dissolved oxygen and specific growth rate control. The performance specifications for the design of the specific growth rate control are:

- It has to cope with changing process dynamics during fed-batch cultivation.
- Set-point tracking and disturbance rejection must be realised during a long cultivation time.
- The only required measurement is dissolved oxygen.
- Oscillations due to interaction with other controllers must be absent.
- Batch-to-batch consistency and quality has to be improved.

In the next sections, first the model is given; then the adaptive control law is derived and evaluated by simulation. Finally, performance of the controller is confirmed in laboratory experiments.

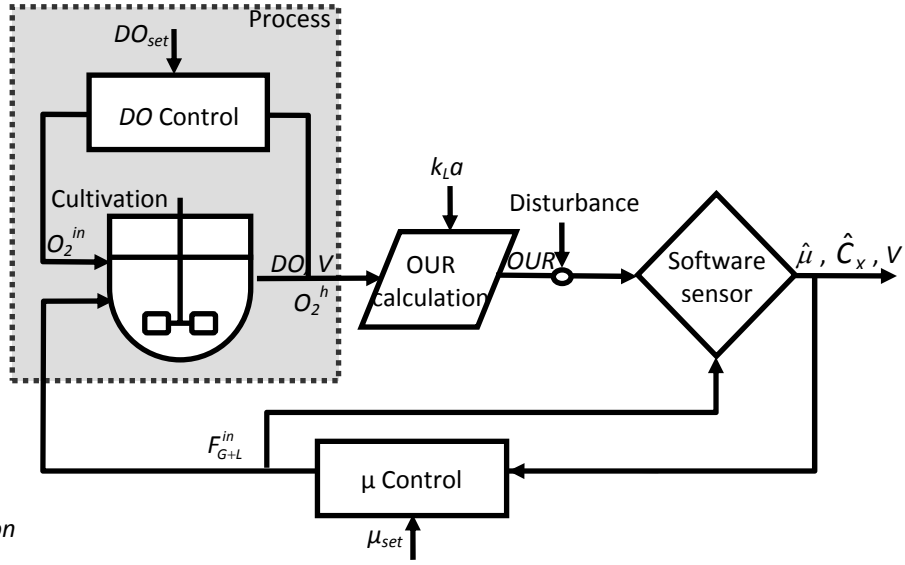
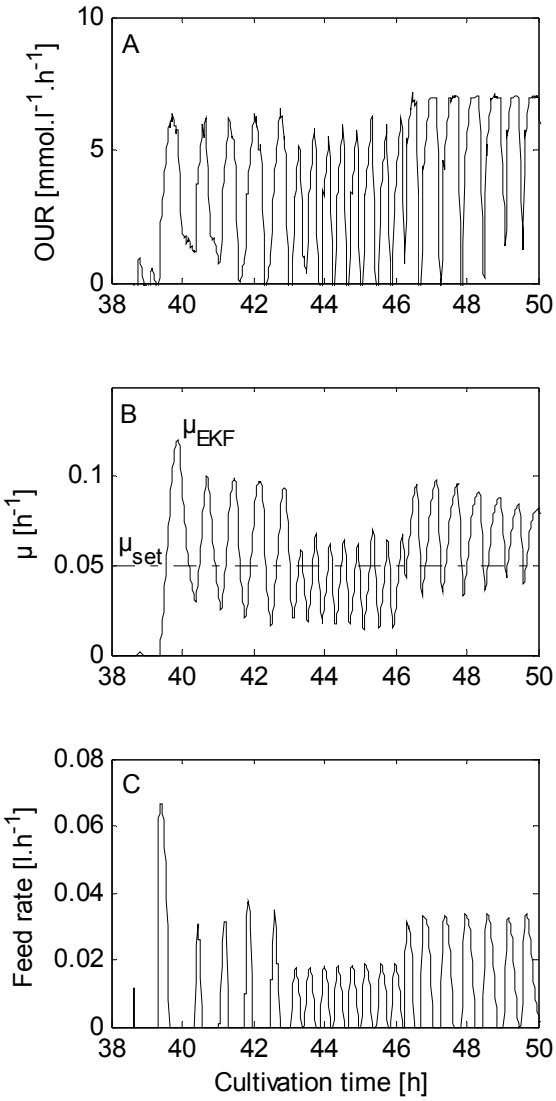


Figure 1. *System configuration*

Figure 2. *Oscillations in measured and software sensor results of the fed-batch phase due to interactions between substrate and oxygen consumption.* ***A.*** *Oxygen uptake rate (OUR)* ***B.*** *Software sensor specific growth rate and set-point,* ***C.*** *Total substrate feed rate*



Materials and methods

Strain and culture media

B. pertussis strain 509 is grown in chemically defined medium containing glutamate and L-lactate as the main carbon sources (Thalen et al., 1999).

Bioreactor conditions

In the PAT framework, process control is one of the four tools with the goal of ensuring final product quality (FDA, 2004). To obtain a constant process, the cultivation must meet three requirements to ensure a constant environment:

- Medium must be of constant quality.
- Gases must be of constant quality.
- Cultivation must be controlled within tight bounds for among others dissolved oxygen, temperature, and specific growth rate.

The cells were grown in five-litre bioreactor containing three litres medium. A six-bladed impeller was used to agitate the medium. Temperature was controlled at 34ºC, agitation speed at 300 to 750 rpm depending on the oxygen demand, and dissolved oxygen at 20% air saturation by headspace aeration only (by changing the incoming oxygen fraction in an oxygen/nitrogen mixture). The total gas flow was kept constant at 1 l/min. Figure 1 shows the configuration of the experimental set up.

Cultivation was performed in two phases. First, batch cultivation was performed until the limiting substrates are depleted. Second, controlled fed-batch cultivation was performed to ensure batch-to-batch consistency and quality by adding a feed with limiting substrates. The fed-batch automatically started when the specific growth rate dropped to the set-point. The set-point for specific growth rate was chosen at 0.05 h^{-1} (Neeleman, 2002). In the fed-batch phase, two 500 ml concentrated stock solutions of glutamate and lactate were used for the two separate feeds and placed on analytical balances. The feeds were added to the bioreactor with a fixed ration by two pumps (101U/R 32 rpm, Watson Marlow ltd., Cornwall UK) connected to the bioreactor control system.

Analysis

A polarographic electrode (Applikon, the Netherlands) was used to measure dissolved oxygen in the medium. A pH electrode (Mettler Toledo, Udorf, Switzerland) was used to measure pH, and temperature was measured with a Pt100 temperature sensor. Glutamate and L-lactate were measured offline with a YSI 2750 select analyser (Yellow Springs Instruments, Yellow Springs, USA). Biomass was measured offline by measuring optical density (OD) at 590 nm of 1 ml suspension using a Vitalab 10 (Vital Scientific, the Netherlands). Quantification of cell-associated protein pertactin (PRN) in whole-cell suspensions was done using ELISA developed by Westdijk et al. (1997).

Hard and software set-up

All sensors were connected to the bioreactor control system ADI1040 (Applikon, Schiedam, The Netherlands), which in turn was connected to a UNIX machine with BCSV (Compex, Belgium (Wieten et al, 1995)). BCSV-software performed the basic control-loops (dissolved oxygen, temperature, and agitation speed) and logged these control-data on the UNIX machine. Using FTP, a Windows XP machine downloaded new data every minute. A routine written in Matlab (Mathworks, Massachusetts, USA) processed data to estimate specific growth rate and biomass, and to calculate feed rate and new set-points for the substrate pumps. BCSV then uploaded the new set-points for the pumps to the UNIX machine (and to the pumps).

Matlab 6.5 was used for the development and application of the software sensor (Extended Kalman Filter) and the adaptive control law. Figure 1 shows an overview of the cultivation process. The software sensor is based on an Extended Kalman Filter and estimates every minute the specific growth rate and biomass using the oxygen uptake rate as an input. A simple model is used with only two parameters: the yield and maintenance coefficient on oxygen. The parameterisation of the software sensor is based on simulation. System and output noise were chosen in such a way that the deviation between estimation and simulation data is minimal. Afterwards the software sensor was evaluated on experimental data. The software sensor showed fast convergence and fitted well to the data. Therefore, in the design of the controller estimations of the software sensor are considered as ideal measurements.

Because dissolved oxygen is controlled, accumulation of dissolved oxygen is small, and the oxygen uptake rate (*OUR*) can be set equal to the oxygen transfer rate (*OTR*). *OUR*,

therefore, was assumed to be equal to *OTR*. *OTR* was calculated every minute using the average of the last four incoming oxygen concentrations (O_2^{in}) and dissolved oxygen values, and from k_La, measured in advance of the cultivation at a range of agitation speeds assumed to be time invariant. The calculated *OTR* will be named "*OUR* measured" from now on.

Simulation of the process

Model for cultivation of *B. pertussis*

Growth of *B. pertussis* is limited by two substrates. The metabolism of *B. pertussis* is described in detail by Thalen et al. (1999) and can be generalised to the formation of biomass from glutamate and lactate by two major pathways: glutamate alone (pathway 1) or glutamate and lactate (pathway 2). The organism can grow on glutamate only, but growth on lactate alone is not possible. In the current medium, glutamate is an essential, and lactate is an enhancing substrate. Growth via these two pathways is assumed to be parallel, and thus the individual growth rates can be added. Neeleman et al. (2001 and 2004) described the cultivation by the following dual substrate model assuming Monod kinetics and oxygen excess. Glutamate is essential for both pathways.

$$\mu(C_G, C_L) = \mu_{\max} f_G(C_G) + \mu_{enh} f_G(C_G) f_L(C_L) \qquad 1$$

$$f_G(C_G) = \frac{C_G}{K_G + C_G} \text{ and } f_L(C_L) = \frac{C_L}{K_L + C_L} \qquad 2$$

$$\frac{dV}{dt} = F_{G+L}^{in} \qquad 3$$

$$\frac{dC_X}{dt} = \left(\mu - \frac{F_{G+L}^{in}}{V}\right) C_X = \left(\mu_{\max} f_G(C_G) + \mu_{enh} f_G(C_G) f_L(C_L) - \frac{F_{G+L}^{in}}{V}\right) C_X \qquad 4$$

$$\frac{dC_G}{dt} = \frac{F_{G+L}^{in}}{V}\left(C_G^{in} - C_G\right) - \left(\frac{\mu_{\max} f_G(C_G)}{Y_{G1}} - \frac{\mu_{enh} f_G(C_G) f_L(C_L)}{Y_{G2}} + m_G\right) C_X \qquad 5$$

$$\frac{dC_L}{dt} = \frac{F_{G+L}^{in}}{V}\left(C_L^{in} - C_L\right) - \left(\frac{\mu_{enh} f_G(C_G) f_L(C_L)}{Y_{L2}} + m_L\right) C_X \qquad 6$$

Glutamate consumption is subject to the following constraint:

$$\frac{\mu_{enh} f_G(C_G) f_L(C_L)}{Y_{G2}} < \frac{\mu_{\max} f_G(C_G)}{Y_{G1}} \qquad 7$$

where C_G and C_L are the glutamate respectively lactate concentration, Y_{G1}, Y_{G2}, and Y_{L2} the biomass yields over the different pathways. μ_{max} is the maximum specific growth rate over pathway 1, μ_{enh} the "enhancing" specific growth rate over pathway 2. K_G and K_L are Monod constants. The biomass growth rate is directly related to biomass (C_X), specific growth rate (μ), and dilution rate (which is described by the incoming substrate feed rate (F_{G+L}^{in}) divided by the volume of the broth (V)). Values used for simulation are given in Table 1. The oxygen uptake rate (*OUR*) is the sum of oxygen used for growth and oxygen used for maintenance.

$$OUR = \left(\frac{\mu}{Y_O} + m_O \right) C_X \qquad 8$$

where Y_O is the yield for biomass on oxygen, and m_O is the maintenance coefficient for biomass on oxygen. The oxygen transfer rate between gas phase and liquid phase is proportional to the concentration gradient in the interfacial area and the mass transfer coefficient (k_La):

$$OUR \approx OTR = k_l a \cdot (O_2^h - DO) \qquad 9$$

where O_2^h is the oxygen concentration in the headspace at the gas-liquid interface following Henry's law (Neeleman, 2002). The output of the process (*OUR*) is calculated from dissolved oxygen, O_2^h, and k_La. In practice, *DO* measurements contain a high level of noise increasing with *OUR*. In simulations used for design and tuning, noise is introduced in the model by adding uniformly distributed random noise as function of *OUR*:

$$\begin{aligned} OUR &= OUR + OUR_{noise} \\ OUR_{noise}(OUR) &= 0.05 \cdot OUR \cdot rand(1) \end{aligned} \qquad 10$$

Table 1. *Constants and initial values for simulation of the dual substrate model for B. pertussis*

V_0 = 3 *l*	Y_L = 0.018 $OD.mmol^{-1}$	Y_{G1} = 0.055 $OD.mmol^{-1}$
C_{X0} = 0.03 *OD*	m_L = 0 $mmol.OD^{-1}.h^{-1}$	Y_{G2} = 0.061 $OD.mmol^{-1}$
C_{G0} = 9.61 $mmol.l^{-1}$	μ_{max} = 0.12 h^{-1}	m_G = 0 $mmol.OD^{-1}.h^{-1}$
C_G^{in} = 500 $mmol.l^{-1}$	μ_{enh} = 0.055 h^{-1}	Y_O = 0.041 $OD.mmol^{-1}$
C_{L0} = 16.24 $mmol.l^{-1}$	K_G = 0.5 $mmol.l^{-1}$	m_O = 0.41 $mmol.OD^{-1}.h^{-1}$
C_L^{in} = 835 $mmol.l^{-1}$	K_L = 0.5 $mmol.l^{-1}$	$\hat{\mu}_O = 0.16\ h^{-1}$

Model for dissolved oxygen control

The model for the cultivation of *B. pertussis* (Neeleman, 2001 and Neeleman et al., 2004) is extended with the dissolved oxygen control loop. The development of dissolved oxygen in time is described by the oxygen transfer rate minus the oxygen uptake rate minus the dilution due to the substrate feed rate:

$$\frac{dDO}{dt} = OTR - OUR - \frac{F_{G+L}^{in}}{V} \cdot DO \qquad 11$$

Control of dissolved oxygen is based on a master-slave control system. The inner or slave loop uses values of the oxygen concentration in the headspace (O_2^h) calculated from the oxygen mass balance. The control variable is the oxygen concentration in the gas flow to the headspace (O_2^{in}). The outer or master loop uses the dissolved oxygen measurement and controls the set-point of the slave controller via an 'auxiliary' variable (O_2^a). The situation is given in Fig. 3. The oxygen concentration in the headspace (O_2^h) calculated from the mass balance. O_2^h is the incoming minus the outgoing amount of oxygen minus the oxygen transfer rate:

$$\frac{dO_2^h}{dt} = \frac{F_{O_2}}{V_h} \cdot (O_2^{in} - O_2^h) - OTR \qquad 12$$

The 'auxiliary' variable becomes (master PI control, see Fig. 3):

$$O_2^a = K_C \cdot (DO_{set} - DO_{sensor}) + \int_0^t \frac{K_C}{\tau_I} \cdot (DO_{set} - DO_{sensor}) dt + C \qquad 13$$

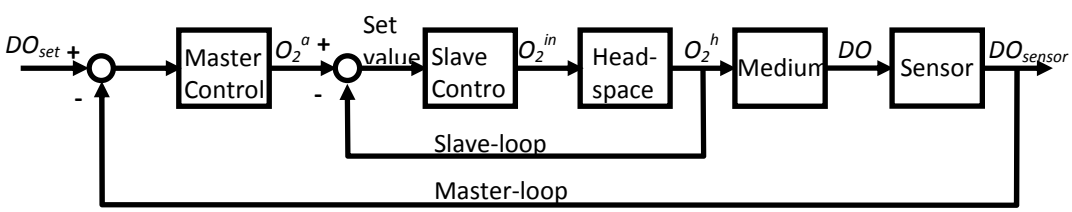


Figure 3. *Master-slave control for dissolved oxygen.*

where DO_{sensor} is the output of the sensor for dissolved oxygen and C is the nominal value. The incoming oxygen fraction is calculated from a calculated "auxiliary" variable (O_2^a). The incoming oxygen concentration (O_2^{in}) becomes (slave PI control, see Fig. 3):

$$O_2^{in} = K_C^h \cdot (O_2^a - O_2^h) + \int_0^t \frac{K_C^h}{\tau_I^h} \cdot (O_2^a - O_2^h) dt + C \qquad 14$$

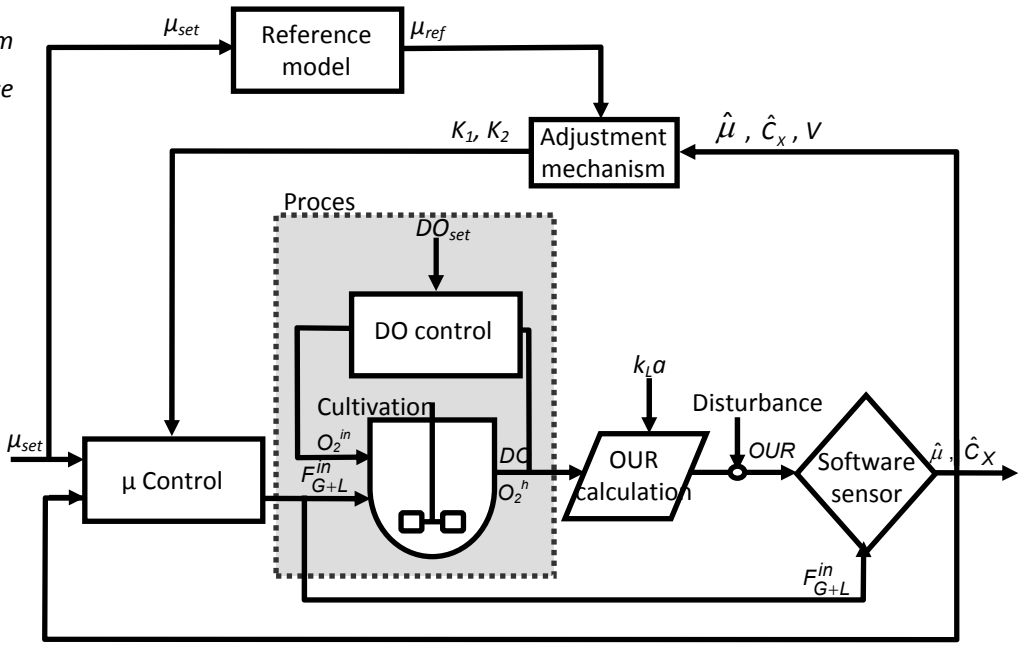
Finally, the response time of the sensor (about 20 s) is modelled by a first order system:

$$\frac{dDO_{sensor}}{dt} = -\frac{1}{\tau_{sensor}} \cdot DO_{sensor} + \frac{1}{\tau_{sensor}} \cdot DO \qquad 15$$

Specific growth rate control

In order to obtain batch-to-batch consistency and constant quality, the cultivation of *B. pertussis* must be controlled. The controller has to maintain the specific growth rate close to set-point in the presence of various uncertainties including external disturbances and time-varying or not exactly known parameters. Model-reference adaptive control (Berber (1998), Ioannou and Sun (1996), and Åström and Wittenmark (1995)) is applied to control the specific growth rate. Fig. 4 shows the control diagram. The adjustment mechanism updates the controller gains (K_1 and K_2) every minute using the estimated states ($\hat{\mu}$ and $\hat{C}_X$), the set-point and the volume. Because the set-point of specific growth rate is constant, controller settings are updated from the estimated states and the set-point specific growth rate. For derivation of the adaptive control law, the software sensor is considered as an ideal measurement; for simulation of the closed loop and for tuning of the controller, the software sensor is included.

Figure 4. *Diagram for model-reference adaptive control*



Model-reference adaptive control

Performance specifications of the controller are given in the terms of a reference model. This model defines how the specific growth rate ideally should respond. A second-order response is preferable to be able to cope with disturbances and offsets. β_1 and β_2 determine the convergence speed of the model-reference controller.

$$\mu + \beta_1 \frac{d\mu}{dt} + \beta_2 \frac{d^2\mu}{d^2t} = \mu_{set} \qquad 16a$$

$$\frac{d^2\mu}{d^2t} = \frac{\mu_{set} - \mu}{\beta_2} - \frac{\beta_1}{\beta_2} \cdot \frac{d\mu}{dt} \qquad 16b$$

Integration of Eq. 16b gives:

$$\frac{d\mu}{dt} = \int_0^t \frac{\mu_{set} - \mu}{\beta_2} dt - \frac{\beta_1}{\beta_2} \cdot \int_0^t \frac{d\mu}{dt} dt \qquad 17a$$

Fed-batch cultivation is started as soon as the specific growth rate drops to the set-point (t=0). With:

$$\mu_{t=0} = \mu_{set} \qquad 17b$$

the reference model for the specific growth rate becomes:

$$\frac{d\mu}{dt} = \int_0^t \frac{\mu_{set} - \mu}{\beta_2} dt + \frac{\beta_1}{\beta_2} \cdot (\mu_{set} - \mu) \qquad 17c$$

with

$$\gamma_2 = \beta_2 \text{ and } \gamma_1 = \frac{\beta_1}{\beta_2} \qquad 17d$$

$$\frac{d\mu}{dt} = \frac{\mu_{set} - \mu}{\gamma_1} + \int_0^t \frac{\mu_{set} - \mu}{\gamma_2} dt \qquad 17e$$

where γ_1 and γ_2 are tuning parameters of the controller.

Although closed loop stability of the cultivation process is not an issue because of the exponentially increasing biomass and volume, stability of the reference model can be guaranteed. The reference model is stable if γ_1 and γ_2 are strictly positive numbers. Because the dynamics of the process plus controller are forced to follow the dynamics of the

reference model, in that case convergence of the error is guaranteed as well. The control law is derived in appendix A and gives:

$$F_{G+L}^{in} = \underbrace{\frac{ac+bd}{aC_G^{in}+bC_L^{in}}\hat{C}_X V}_{\text{Feed rate calculation}} + \underbrace{K_1 \cdot (\mu_{set} - \hat{\mu})}_{\text{“Proportional” action}} + \underbrace{\int_0^t K_2 \cdot (\mu_{set} - \hat{\mu})dt}_{\text{“Integral” action}} \qquad 18a$$

where K_1 and K_2 are the controller gains directly related to the tuning parameters γ_1 and γ_2 in Eq. 18b. K_1 and K_2 are adapted to the changing process circumstances:

$$K_1 = \frac{V}{\gamma_1(aC_G^{in}+bC_L^{in})} \text{ and } K_2 = \frac{V}{\gamma_2(aC_G^{in}+bC_L^{in})} \qquad 18b$$

In this application (vaccine production) only dissolved oxygen and volume measurements were available. Therefore, a software sensor based on an Extended Kalman Filter (EKF) estimates specific growth rate and biomass every minute. The controller uses the estimated values. The constants *a*, *b*, *c*, and *d* depend on the model parameters as explained in the Appendix.

Note that the laws of model-reference adaptive control plus process differential equations lead to an adaptive “PI” controller (appendix A). In this PI control, the nominal value (“feed rate calculation”) and the gains K_1 and K_2 are functions of the states.

Controller tuning

According to Bastin and Dochain (1990) closed loop stability cannot be guaranteed *a priori*. The state variables follow an exponentially growing trajectory, which is characteristic of an unstable behaviour. Hence the goal of a fed-batch cultivation process is not to stabilise it, but to optimise it while keeping an inherently unstable type of behaviour under control. Therefore tuning is performed by minimisation of the error.

The tuning parameters γ_1 and γ_2 determine the convergence speed of the controller. The controller is stable by definition of Eq. 17e, if γ_1 and γ_2 are strictly positive numbers. The optimal trade-off between tracking behaviour, disturbance rejection, and stability should be pursued. Initially, tuning was performed without incorporation of dissolved oxygen dynamics (Eqs. 12 to 15). At implementation, tuning on the specific growth rate only

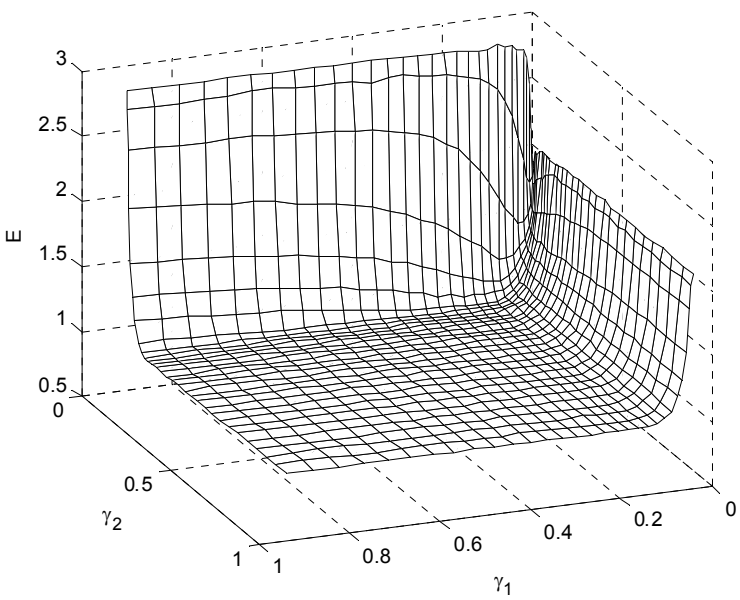
$$E_{\mu} = \frac{1}{\mu_{set}} \cdot \sqrt{\frac{\sum_{k=1}^{N} \left[\hat{\mu}(k) - \mu_{set}(k) \right]^2}{(N-1)}} \qquad 19$$

results in oscillatory bioreactor control due to interactions between dissolved oxygen and specific growth rate control (Fig. 2), because time constants of oxygen and substrate consumption are in the same order of magnitude. Neglecting dissolved oxygen dynamics leads to improper tuning of the specific growth rate controller and to oscillations. In contrast to common believes, simulation of the process, therefore, should incorporate dissolved oxygen control (Eqs. 12 to 15). Controller parameters γ_1 and γ_2 should be tuned by minimising a combination of the errors for $\hat{\mu}$ as well as *DO* in time. The ability to keep the specific growth rate at the desired set-point (E_{μ}) and the ability to reduce oscillations (E_{DO}) are balanced in an objective function (E):

$$E = E_{\mu} + E_{DO} = \frac{1}{\mu_{set}} \cdot \sqrt{\frac{\sum_{k=1}^{N} \left[\hat{\mu}(k) - \mu_{set}(k) \right]^2}{(N-1)}} + \frac{1}{DO_{set}} \cdot \sqrt{\frac{\sum_{k=1}^{N} \left[DO(k) - DO_{set}(k) \right]^2}{(N-1)}} \qquad 20$$

Optimisation of this objective function (Eq. 20) gives the optimal tuning parameters. Calculation of the objective function for a range of γ_1 and γ_2 results in the profile shown in Fig. 5. It is important to choose γ_1 and γ_2 not too small, but the tuning parameters can be chosen larger than the optimal value without significant loss of controller performance. This is in agreement with the practical findings: too small values for the controller tuning parameters lead to severe interaction and oscillations, whereas values larger than the optimal value have no effect on controller performance.

Figure 5. *Objective function (E) as a function of γ_1 and γ_2.*



Simulation results

Robustness and accuracy tests were performed to examine controller performance. Several difficulties have been introduced in the model: time-varying or drifting kinetics, changes in set-point, and disturbances. For the purpose of the test, we assumed a drift plus a sinus with different frequencies on the model parameters (Fig. 6). External disturbances are introduced by observation noise (Eq. 10). In the controller, the parameters were set constant.

Figure 7A-E shows a characteristic simulation with good performance of the controller. As a result of the objective function the interference of the controllers as observed in Fig. 2 for the case without considering *DO* dynamics disappeared. The control was free of oscillations. Furthermore, the controller perfectly tracked the specific growth rate without offsets. The controller coped well with changes in set-point, noise on *OUR*, and time-varying biomass, volume, kinetics, and substrate yields. Concluding, the adaptive controller passed the robustness and accuracy tests with flying colours.

Figure 7 shows also a comparison of the adaptive controller with a standard PI controller, in which the controller gains are constant during the cultivation (Fig. 7F). Performance of both controllers is compared by calculating the integral squared error (ISE). Compared to the standard PI-controller the performance in terms of ISE improves 2.5 times by using the model-reference adaptive controller. The improvement is most pronounced during changes of set-point and with respect to offset compensation.

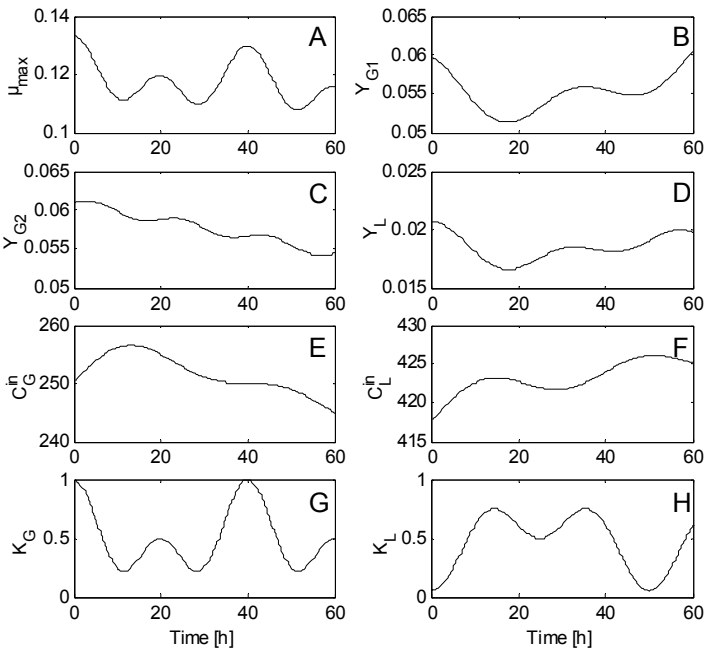


Figure 6. *Injected time-variation of model parameters during simulation for robustness tests* ***A.*** *Maximum specific growth rate (μ_{max})* ***B.*** *Yield on glutamate over pathway 1 (Y_{G1})* ***C.*** *Yield on glutamate over pathway 2 (Y_{G2})* ***D.*** *Yield on lactate (Y_L)* ***E.*** *Glutamate concentration in the feed (C_G^{in})* ***F.*** *Lactate concentration in the feed (C_L^{in})* ***G.*** *Monod constant on glutamate (K_G)* ***H.*** *Monod constant on lactate (K_L)*

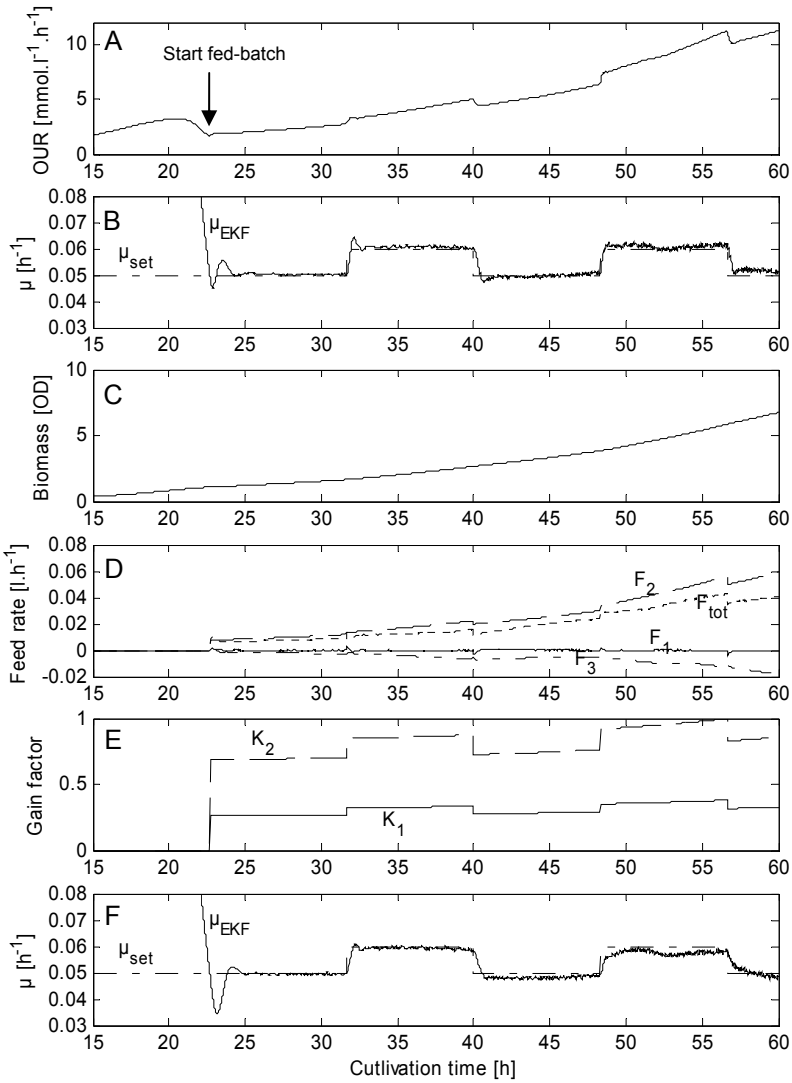


Figure 7. *Simulation results with time-varying parameters (Fig. 6) and changes of set-point.* ***A.*** *Oxygen Uptake Rate (OUR)* ***B.*** *Specific growth rate (μ) using model-reference adaptive control, ISE=1.2.10^{-4}* ***C.*** *Biomass concentration (C_X)* ***D.*** *Substrate feed rate. F_1: proportional action of the controller, F_2: action of the feet rate calculator, F_3: integral action of the controller, and F_{tot}: total feed rate* ***E.*** *Controller Gain. K_1: "proportional"-action, K_2: "integral"-action* ***F.*** *Specific growth rate (μ) using standard PI control, ISE=3.0.10^{-4}*

Experimental results and discussion

In the PAT framework, process control is one of the four tools with the goal of ensuring final product quality (FDA, 2004). The controller design in the previous section is put to a test in a laboratory cultivation. Figures 8, 9, and 10 show the results.

In compliance with the PAT rules, the cultivation was controlled within tight bounds (Figs. 8 and 9) for dissolved oxygen, temperature, and specific growth rate (and therefore substrates).

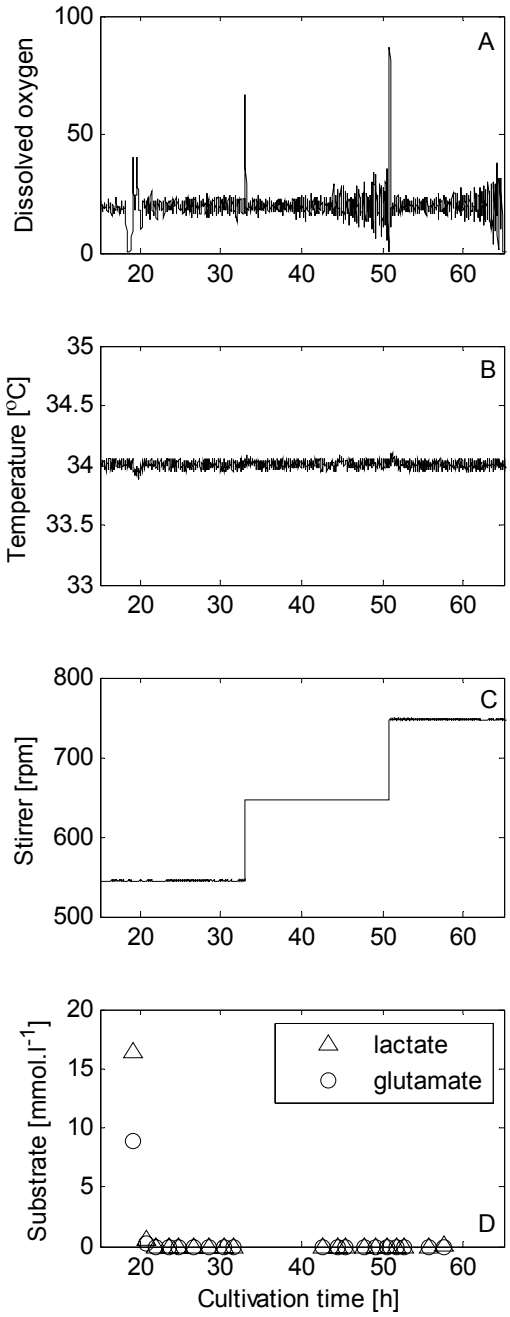
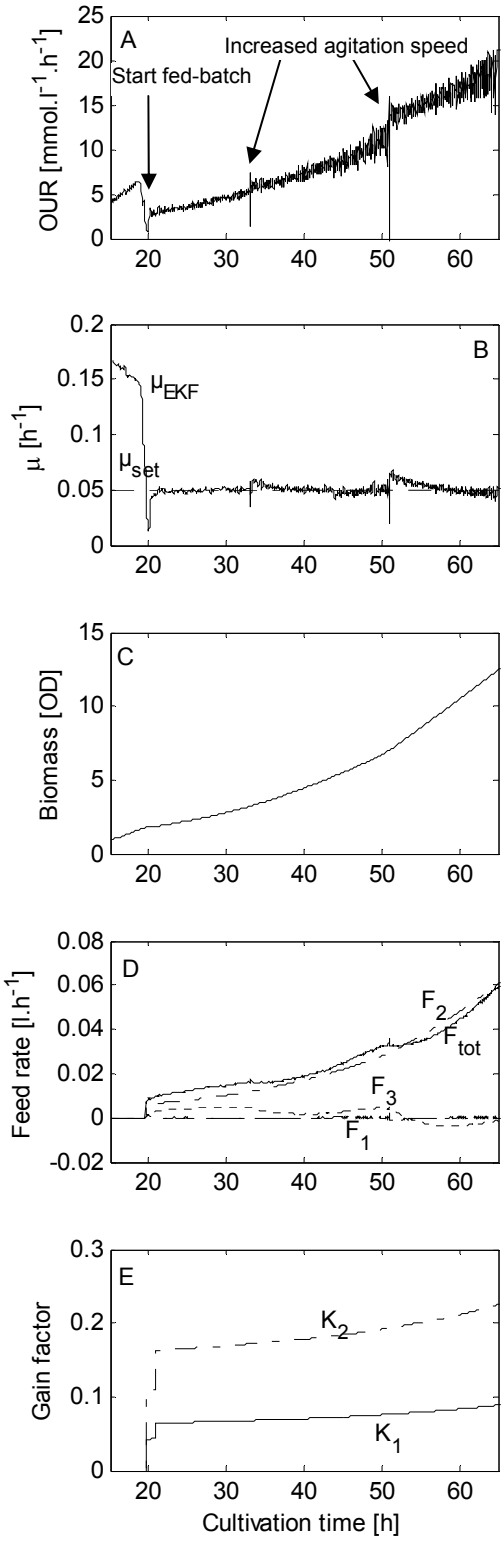


Figure 8. *Experimental cultivation conditions.* ***A.*** *Dissolved oxygen* ***B.*** *Temperature* ***C.*** *Agitation speed* ***D.*** *Substrate concentrations: lactate and glutamate*

The performance of the designed model-reference adaptive controller was satisfactory. The controller responds fast to deviations (Fig. 9). It properly tracked the set point in the presence of various uncertainties including disturbances on dissolved oxygen, improper working substrate pumps, and uncertain parameters. The controller properly coped with exponentially increasing biomass. The estimated specific growth rate was controlled close to the set point. The standard deviation was between 0.004 and 0.012 during the fed-batch phase.

Figure 9. *Experimental results for model- reference adaptive control with doubled controller tuning parameters. A. Measured oxygen uptake rate (OUR) B. Specific growth rate estimated by the software sensor and set-point C. Estimated biomass concentration (C_X) D. Substrate feed rate. F_1: proportional action of the controller, F_2: action of the feet rate calculator, F_3: integral action of the controller, and F_{tot}: total feed rate E. Controller Gain. K_1: "proportional"-action, K_2: "integral"-action*



Time delays change process dynamics and tuning results and may introduce oscillations. To stay away from oscillations, the tuning parameters were doubled or quadruplicated without loss of controller performance in comparison to their optimal values calculated in simulation. The experiments showed good controller performance (see Fig. 9 for doubled controller tuning parameters). As forecasted by simulations oscillations due to interference of the specific growth rate and the dissolved oxygen controller as in Fig. 2 were absent with the chosen controller parameters. Furthermore, the controller tracked the specific growth rate without offsets. Although the specific growth rate was estimated from a noisy *OUR* signal, it was properly estimated and controlled.

At the start of the fed-batch (t = 23 *h*), it happened that only small amounts of substrates were pumped into the bioreactor due to improper working pumps. As can be seen in Fig. 9, the controller reacted fast by immediately increasing the feed rate, and bringing the specific growth rate back to the set-point.

Depending on the oxygen demand, agitation speed was adjusted in steps from 550 to 750 *rpm* during the fed-batch phase (Fig. 8). When oxygen limitation occurred, agitation speed was increased with 100 *rpm*. Two peaks were observed during the cultivation caused by increasing agitation speed (Figs. 8 and 9). The first peak coincides with increasing agitation speed from 550 to 650 *rpm*; the second peak coincides with increasing agitation speed from 650 to 750 *rpm*.

Figure 10 show details of two experimental fed-batch cultivations for *Bordetella pertussis* to compare the performance of the adaptive controller with a fixed PI controller used in previous work (Neeleman et al., 2004). The ISE of the adaptive controller improved up to 10 times compared to the standard PI-controller.

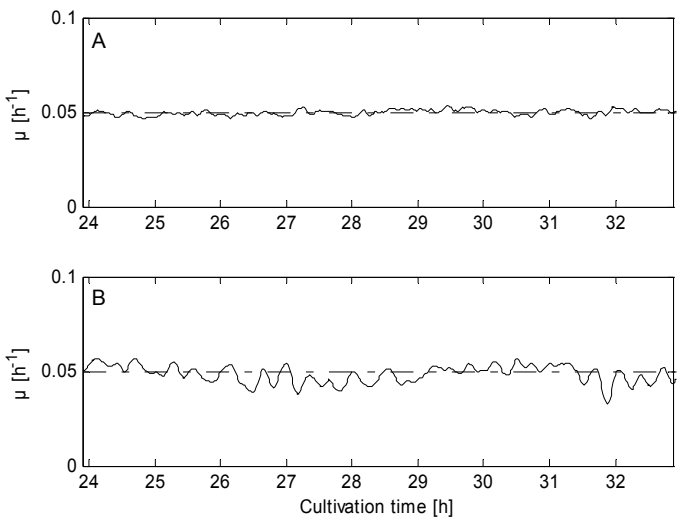


Figure 10. *Detail of two fed-batch experiments with specific growth rate control.* ***A.*** *Using model-reference adaptive control, ISE=2.0.10^{-5}* ***B.*** *Using PI control with fixed tuning parameters (Neeleman et al., 2004), ISE=2.2.10^{-4}.*

PRN-activity was measured at the end of the fed-batch. High cell surface densities of protective proteins are likely to augment the potency of pertussis whole cell vaccine. Therefore, it is important to examine density of outer membrane components per cell besides protein activity per volume (Westdijk et al., 1997). The PRN-activity and density for fed-batch cultivation was compared with the activity and density of this component for traditional batch cultivation in Table 2 (batch performed by Westdijk et al., 1997). First, variation in PRN-activity and density was smaller after fed-batch cultivation despite the high amounts reached. Second, although PRN-density was equal or lower after fed-batch than after batch cultivation, PRN-activity after fed-batch was fourfold of the PRN-activity of the batch cultivation. Concluding, applying controlled fed-batch cultivation instead of traditional fed-batch cultivation increased batch-to-batch consistency, vaccine potency, and quality.

The design of the specific growth rate controller results in a controller of PI type. The main achievement is the broad applicability, because the settings in the adaptive PI controller are the result of the process characteristics. In contrast to standard PI control, retuning is easy when the controller is applied in a different situation by inserting the corresponding process characteristics. Simulations showed that performance of the specific growth rate controller was good for a ten-fold larger scale bioreactor (figures are not shown). Important requirements for good performance of the specific growth rate controller are accurate measurement of the oxygen uptake rate and proper dissolved oxygen control. The oxygen uptake rate can be calculated using the overall mass balance, in which the composition of the gas flows in and out the bioreactor are measured. Attention should be paid where to place the dissolved oxygen sensor in a large bioreactor in order to deal with non-homogeneity on large scale. Also, the design of the bioreactor must be such that supply of oxygen is not a problem. Taking into account these requirements, the specific growth rate controller is well applicable on large scale using identical tuning parameters γ_1 and γ_2.

Table 2. *PRN activity and density after batch and fed-batch cultivation*

	PRN activity [$mg.l^{-1}$]	PRN density [$\mu g.OD^{-1}$]
Fed-batch	58 ± 4	16.5 ± 1.5
Traditional batch	14.5 ± 5.5	26 ± 10

Conclusions

To control the metabolic activity during biopharmaceutical production, a specific growth rate controller is designed and evaluated by simulation and experiments for the production of the vaccine against whooping cough (*B. pertussis*). The qualifications of the controller are summarized below.

- Using the laws of model-reference control in combination with mass balances, a stable adaptive "PI" controller was derived. The controller settings are adapted to the changing process dynamics (volume, biomass, and set-point for specific growth rate).
- The method does not require online model identification, thus avoiding the need for process perturbation and complex implementation. Furthermore, it does not require online measurement of (part of) the states. The designed controller only requires online measurement of the oxygen uptake rate.
- Tuning based on a combination of deviations in specific growth rate and dissolved oxygen leads to bioreactor control without interactions between specific growth rate and dissolved oxygen controller.
- The model-reference adaptive controller has an outstanding performance, both in simulation as well as in experimental conditions. Although the oxygen consumption rate contains a high level of noise, the specific growth rate is controlled accurately
- Simulations show that the controller is highly robust for a large range of disturbances. Experiments showed the controller was good for its ability to cope with time-varying kinetics and states, noise, and external disturbances.
- Control of the specific growth rate at low level results in constant and high antigen activity of the vaccine. Application of the controller, therefore, contributes to batch-to-batch consistency and quality. It can be broadly applied for the production of biopharmaceuticals during fed-batch cultivation.
- The designed control system is a step forward to meet the PAT objectives as defined by the FDA.

Acknowledgements

Wageningen University (WU) and the Netherlands Vaccine Institute (NVI) work in a project with Applikon Biotechnology BV (Schiedam, the Netherlands) and Siemens NV (The Hague, the Netherlands) in order to improve the production process by release of biopharmaceuticals on the basis of new online techniques. The

project is named “PaRel” or Parametric Release. Senter (a subsidiary of the Ministry of Economical Affairs), The Netherlands, supports the project under project name TSGE3067. Experiments are performed at NVI.

Appendix A: derivation of the control law

The adaptive controller is derived from combination of two principles: from the dual substrate model for growth of *B. pertussis* on glutamate and lactate and from the reference model for ideal response of the specific growth rate controller. The purpose is to compare the dynamics of the dual substrate model (Eq. A3) to the dynamics according to the reference model (Eq. A15) and to derive an adaptive control law for the feed rate F_{G+L}^{in} from these equations.

Dual substrate model for *B. pertussis*

The dual substrate model for growth of *B. pertussis* (Neeleman, 2001 and Neeleman et al., 2004):

$$\mu(C_G, C_L) = \mu_{\max} f_G(C_G) + \mu_{enh} f_G(C_G) f_L(C_L) \qquad \text{A1}$$

with Monod kinetics:

$$f_G(C_G) = \frac{C_G}{K_G + C_G} \text{ and } f_L(C_L) = \frac{C_L}{K_L + C_L} \qquad \text{A2}$$

Differentiation of Eq. A1 with respect to time gives the behaviour of μ in response to changes in the substrate concentrations C_G and C_L:

$$\begin{aligned} \frac{d\mu}{dt} &= \frac{\mu_{\max} K_G}{(C_G + K_G)^2} \cdot \frac{dC_G}{dt} + f_L \frac{\mu_{enh} K_G}{(C_G + K_G)^2} \cdot \frac{dC_G}{dt} + f_G \frac{\mu_{enh} K_L}{(C_L + K_L)^2} \cdot \frac{dC_L}{dt} \\ &= \left(\mu_{\max} + f_L \mu_{enh}\right) \cdot \frac{K_G}{(C_G + K_G)^2} \cdot \frac{dC_G}{dt} + f_G \mu_{enh} \frac{K_L}{(C_L + K_L)^2} \cdot \frac{dC_L}{dt} \end{aligned} \qquad \text{A3}$$

where glutamate and lactate concentrations in time are:

$$\frac{dC_G}{dt} = \frac{F_{G+L}^{in}}{V}\left(C_G^{in} - C_G\right) - \left(\frac{\mu_{\max} f_G}{Y_{G1}} - \frac{\mu_{enh} f_G f_L}{Y_{G2}} + m_G\right) C_X \qquad \text{A4}$$

$$\frac{dC_L}{dt} = \frac{F_{G+L}^{in}}{V}\left(C_L^{in} - C_L\right) - \left(\frac{\mu_{enh} f_G f_L}{Y_{L2}} + m_L\right) C_X \qquad \text{A5}$$

Because the glutamate and lactate concentrations cannot be measured online, Eqs. A4 and A5 will be combined with Eq. A3 in steps A6 to A13. Feed concentrations (C_G^{in} and C_L^{in}) are much higher than the substrate concentrations in the bioreactor (C_G and C_G). So the latter can therefore be neglected in Eqs. A4 and A5:

$$C_G^{in} >> C_G \text{ and } C_L^{in} >> C_L \qquad \text{A6}$$

Rewriting Eq. A2 gives:

$$C_G = \frac{f_G K_G}{1 - f_G} \text{ and } C_L = \frac{f_L K_L}{1 - f_L} \qquad \text{A7}$$

Rewriting Eq. A1 gives the normalised equations of Monod:

$$\mu_{\max} f_G = \mu - \mu_{enh} f_G f_L \text{ and } \mu_{enh} f_G f_L = \mu - \mu_{\max} f_G \qquad \text{A8}$$

Substitution of Eqs. A4, A5, A7, and A8 in A3 give:

$$\begin{aligned} \frac{d\mu}{dt} = (\mu_{\max} + f_L \mu_{enh})(1 - f_G)^2 \frac{1}{K_G} \left(\frac{F_{G+L}^{in}}{V} C_G^{in} - \left(\frac{\mu - \mu_{enh} f_G f_L}{Y_{G1}} - \frac{\mu - \mu_{\max} f_G}{Y_{G2}} + m_G \right) C_X \right) \\ + f_G \mu_{enh} (1 - f_L)^2 \frac{1}{K_L} \left(\frac{F_{G+L}^{in}}{V} C_L^{in} - \left(\frac{\mu - \mu_{\max} f_G}{Y_{L2}} + m_L \right) C_X \right) \end{aligned} \qquad \text{A9}$$

In these equations the unknowns are the kinetics f_G and f_L. In principle, an unlimited number of combinations is possible (see Eq. A8). Neeleman (2002) and Neeleman et al. (2004) fixed the ratio (β) between the normalised equations of Monod (f_G and f_L) to obtain metabolic stability:

$$\beta = \frac{f_L}{f_G} \qquad \text{A10}$$

Figure A1 shows the values of f_G and f_L at various β and μ. The white parts of the figure show the admissible regions of the normalised kinetics with the following constraint:

$$\max\left(0, \frac{\mu - \mu_{\max}}{\mu_{enh}}\right) < \beta < \frac{\mu_{\max} + \mu_{enh}}{\mu} \qquad \text{A11}$$

For the simulations and experiments in this study a ratio of 1 was used. At $\mu_{set} = 0.05\ h^{-1}$, the normalised Monod equations are linear approximated by:

$$f_G = v \cdot \mu \text{ and } f_L = w \cdot \mu \qquad \text{A12}$$

where v and w are constants. Substitution of Eq. A12 in A9 gives the behaviour of the

specific growth rate μ in response to the feed rate F_{G+L}^{in}:

$$\frac{d\mu}{dt} = a\left(\frac{F_{G+L}^{in}}{V}C_G^{in} - cC_X\right) + b\left(\frac{F_{G+L}^{in}}{V}C_L^{in} - dC_X\right) \qquad \text{A13a}$$

where

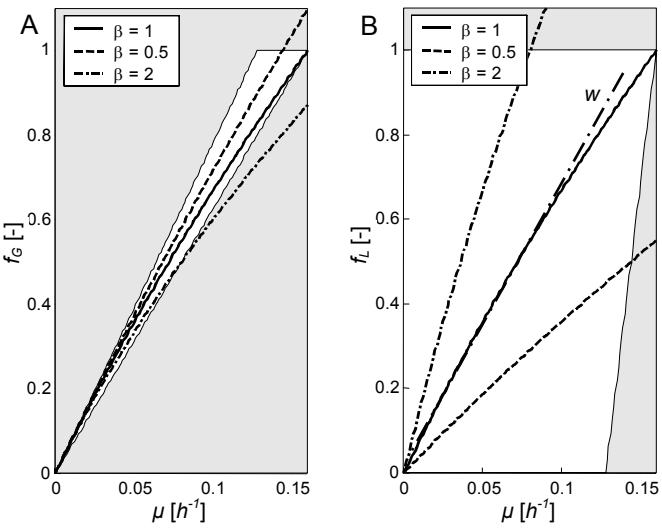
$$a = (\mu_{\max} + w \cdot \mu_{enh}\mu_{set})(1 - v \cdot \mu_{set})^2 \frac{1}{K_G}$$

$$b = v \cdot \mu_{enh}\mu_{set}(1 - w \cdot \mu_{set})^2 \frac{1}{K_L}$$

$$c = \frac{\mu_{set} - v \cdot w \cdot \mu_{enh}\mu_{set}^{\ 2}}{Y_{G1}} - \frac{\mu_{set} - v \cdot \mu_{\max}\mu_{set}}{Y_{G2}} + m_G \qquad \text{A13b}$$

$$d = \frac{\mu_{set} - v \cdot \mu_{\max}\mu_{set}}{Y_{L2}} + m_L$$

Figure A1. *Diagram for Monod kinetics (f_G and f_L) as function of the specific growth rate (μ) at various β.* ***A.*** *f_G For glutamate* ***B.*** *f_L for lactate*



Reference model

Recall the reference model:

$$\frac{d\mu}{dt} = \frac{\mu_{set} - \mu}{\gamma_1} + \int_0^t \frac{\mu_{set} - \mu}{\gamma_2} dt \qquad \text{A14}$$

Combination of dual substrate and reference model

Combining the reference model (Eq. A14) and the dual substrate model (Eq. A13a) by elimination of $d\mu/dt$ and subsequent rewriting give the following adaptive control law:

$$F_{G+L}^{in} = \underbrace{\frac{ac+bd}{aC_G^{in}+bC_L^{in}}\hat{C}_X V}_{\text{Feed rate calculation}} + \underbrace{K_1 \cdot (\mu_{set} - \hat{\mu})}_{\text{“Proportional” action}} + \underbrace{\int_0^t K_2 \cdot (\mu_{set} - \hat{\mu})dt}_{\text{“Integral” action}} \qquad \text{A15a}$$

where the controller gains K_1 and K_2 are directly related to the tuning parameters γ_1 and γ_2:

$$K_1 = \frac{V}{\gamma_1(aC_G^{in}+bC_L^{in})} \text{ And } K_2 = \frac{V}{\gamma_2(aC_G^{in}+bC_L^{in})} \qquad \text{A15b}$$

Eq. 15b shows that the controller gains are automatically adapted to the changing process conditions.

CHAPTER 5

Online automatic tuning and control for fed-batch cultivation

Published as: Z. I. T. A. Soons, G. van Straten, L. A. van der Pol, A. J. B. van Boxtel, Online automatic tuning and control for fed-batch cultivation. Bioprocess and Biosystems Engineering, online first.

Abstract

Performance of controllers applied in biotechnological production is often below expectation. Online automatic tuning has the capability to improve control performance by adjusting control parameters. This work presents automatic tuning approaches for model reference specific growth rate control during fed-batch cultivation. The approaches are direct methods that use the error between observed specific growth rate and its set point; systematic perturbations of the cultivation are not necessary. Two automatic tuning methods proved to be efficient, in which the adaptation rate is based on a combination of the error, squared error and integral error. These methods are relatively simple and robust against disturbances, parameter uncertainties, and initialization errors. Application of the specific growth rate controller yields a stable system. The controller and automatic tuning methods are qualified by simulations and laboratory experiments with *Bordetella pertussis*.

Keywords

online automatic tuning, fed-batch cultivation, specific growth rate, *Bordetella pertussis*, stability

Nomenclature

a, b, c, d	constants for dual substrate model for *B. pertussis*
$a_1, a_2 \ldots a_7$	constants for disturbances on model parameters
A, B, C	system matrices for dual substrate model for *B. pertussis*
C_G, C_G^{in}	glutamate concentration in the medium and in the feed [$mmol.l^{-1}$]
C_L, C_L^{in}	lactate concentration in the medium, respectively in the feed [$mmol.l^{-1}$]
C_N	nominal value of controller [$mmol.l^{-1}$]
C_X	biomass concentration [OD]
c_1, c_2, c_3	constants for the choice of the adaptation mechanisms
DO, DO_{sensor}	dissolved oxygen in the medium and measured by the sensor [$mmol.l^{-1}$]
DO_{set}	set-point dissolved oxygen [$mmol.l^{-1}$]
e	error between reference model and process, $\mu_{ref} - \hat{\mu}$ [h^{-1}]
E	mean absolute error [h^{-1}]
F	state feedback matrix
F_{G+L}^{in}	total substrate feed rate (glutamate + lactate) [$l.h^{-1}$]
k	time instant
k_La	oxygen transfer coefficient [h^{-1}]
K_C, K_I, K_P	controller gains
K_G, K_L	Monod constant on glutamate and lactate [$mmol.l^{-1}$]
m_G, m_L, m_O	maintenance coefficient on glutamate, lactate and oxygen [$mmol.OD^{-1}.h^{-1}$]
N	length of moving window
O	oscillations measure [h^{-1}]
O_2^h, O_2^{in}	oxygen concentration in headspace, and in incoming air (liquid phase) [$mmol.l^{-1}$]
OD	optical Density at 590nm
OUR, OTR	oxygen uptake rate, oxygen transfer rate [$mmol.l^{-1}.h^{-1}$]
p, p^d	model parameters, model parameters with disturbances
t	cultivation time [h]
V	liquid volume [l]
u, x, y	inputs, states, output
Y_{G1}, Y_{G2}	biomass yield on glutamate over pathway 1 and pathway 2 [$OD.mmol^{-1}$]
Y_L, Y_O	biomass yield on lactate, respectively on oxygen [$OD.mmol^{-1}$]

Greek letters

β	adaptation rate

γ_1, γ_2	tuning parameters for μ control
ε	error between set point and process, $\mu_{set} - \hat{\mu}$ [h^{-1}]
μ, μ_{set}	specific growth rate, set point specific growth rate [h^{-1}]
μ_{enh}, μ_{max}	enhanced, maximum specific growth rate [h^{-1}]
τ_I	reset time for dissolved oxygen control [h]
τ_{sensor}	response time for dissolved oxygen sensor [h]

Superscripts and subscripts

^, $_{EKF}$	estimated values

Introduction

Unacceptably sluggish or oscillatory controllers are generally classified as either “fair” or “poor” while controllers with minor performance deviations are classified as “acceptable” or “excellent”. Desbourough and Miller (2002) state that 32% of 26.000 controllers investigated in industry is classified as “poor” or “fair”. Only one third of the controllers were classified as acceptable performers and two thirds had significant improvement opportunity. Figure 1 shows an example of an experimental run for specific growth rate control in vaccine production on lab-scale. The strong oscillations show the poor performance and indicate the relevance of automatic tuning in this field of application. `

In biopharmaceutical production micro-organisms are cultivated in batch or fed-batch systems. Most vaccines, such as whooping cough, are produced in batch cultivation. Production of pertussis toxin (which is one of the important antigens in the vaccine against infection with whooping cough) is strongly growth-associated (Licari et al., 1991; Rodriguez et al., 1994). Deviations in specific growth rate will therefore lead to deviations in antigen levels and vaccine quality. To obtain a high quality-vaccine and to ensure batch-to-batch consistency, it is important to control the specific growth rate at a constant level. Soons et al. (2006) developed a specific growth rate controller by combining a reference model and a cultivation model. The result is a control law, which is adaptive for changes in volume and biomass. It was shown that the choice of the reference model is a main factor determining the performance of the specific growth rate controller.

Apart from the reference model, biological behaviour is essential for the controller performance and as we know biological behaviour is not perfectly known. Despite good performance in Soons et al. (2006), the method does not correct for residual errors due to model-process mismatches (Soons et al., 2006a). Changes in the system, inaccurate

knowledge about kinetics, varying properties of micro-organisms and response times of sensor systems are reasons for model-process mismatches. Online automatic tuning has the capability to deal with the mismatches and to improve controller performance by adjusting the controller parameters (Dagci et al., 2001).

The aim of this paper is to investigate schemes for automating and to speed up the tuning procedures. To this end, the specific growth rate controller for fed-batch cultivation developed in (Soons et al., 2006) is extended with an automatic tuning method that is driven by the persistence of the residual errors. A direct method (in which the adjustment rule makes directly updates of the *controller* parameters) is preferred over an indirect method (in which *process* parameters are estimated to update controller parameters) as in this way a degree of imperfection of the model for model-based control can be handled. Moreover, probing should be avoided, because it may affect the critical variables in biopharmaceutical production.

Although *a priori* closed loop stability is hard to guarantee for a fed-batch system in which the biomass increases exponentially, stability of the control loops under application of the designed control law is evaluated.

This paper is organized as follows. First a brief description of the process and control system is given. Next, controller evaluation criteria are defined. Then, a short literature review and a textbook method for online automatic tuning using the MIT rule are presented. Due to unsatisfactory results of this method three algorithms are derived and evaluated by simulations in the section “Online automatic tuning methods”. Also, stability of the closed loop is considered. Finally, the two best methods are tested in laboratory experiments for *B. pertussis*.

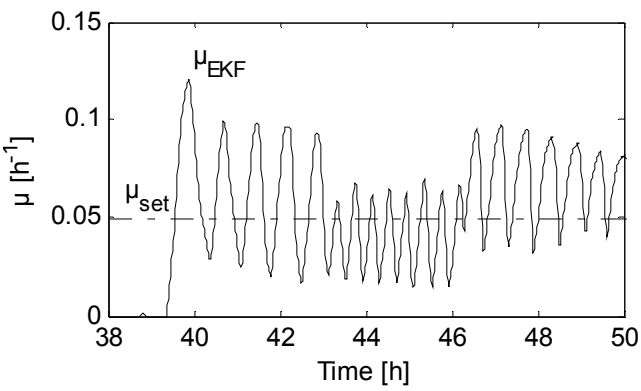


Figure 1. *Illustration of a process showing poor performance. Fed-batch cultivation with specific growth rate control for an experimental run of vaccine production on lab-scale (Soons et al., 2006a).*

Process description

Process model

A process model for vaccine production of *B. pertussis* is used to evaluate the controllers with automatic tuning. The production is performed in fed-batch cultivation, in which the specific growth rate is controlled to achieve a higher level of batch-to-batch consistency. Growth of *B. pertussis* is limited by two substrates (Thalen et al., 1999) and is modelled by the following dual substrate model assuming Monod kinetics and oxygen excess (Neeleman et al., 2001; Soons et al., 2006) during fed-batch cultivation (Eqs. 1 to 13). The changes of dissolved oxygen in time are equal to the oxygen transfer rate minus the oxygen uptake rate minus the dilution due to the substrate feed rate.

$$\mu(C_G, C_L) = \mu_{\max} f_G(C_G) + \mu_{enh} f_G(C_G) f_L(C_L) \qquad 1$$

$$f_G(C_G) = \frac{C_G}{K_G + C_G} \text{ and } f_L(C_L) = \frac{C_L}{K_L + C_L} \qquad 2$$

$$\frac{dV}{dt} = F_{G+L}^{in} \qquad 3$$

$$\frac{dC_X}{dt} = \left(\mu - \frac{F_{G+L}^{in}}{V}\right) C_X = \left(\mu_{\max} f_G(C_G) + \mu_{enh} f_G(C_G) f_L(C_L) - \frac{F_{G+L}^{in}}{V}\right) C_X \qquad 4$$

$$\frac{dC_G}{dt} = \frac{F_{G+L}^{in}}{V}\left(C_G^{in} - C_G\right) - \left(\frac{\mu_{\max} f_G(C_G)}{Y_{G1}} - \frac{\mu_{enh} f_G(C_G) f_L(C_L)}{Y_{G2}} + m_G\right) C_X \qquad 5$$

$$\frac{dC_L}{dt} = \frac{F_{G+L}^{in}}{V}\left(C_L^{in} - C_L\right) - \left(\frac{\mu_{enh} f_G(C_G) f_L(C_L)}{Y_{L2}} + m_L\right) C_X \qquad 6$$

$$\frac{dDO}{dt} = OTR - OUR - \frac{F_{G+L}^{in}}{V} \cdot DO \qquad 7$$

The oxygen uptake rate (*OUR*) is the sum of oxygen used for growth (μ/Y_O) and oxygen used for maintenance (m_O):

$$OUR = \left(\frac{\mu}{Y_o} + m_o\right) C_x \qquad 8$$

The bioreactor is aerated using headspace aeration only. The liquid phase oxygen concentration (O_2^h) in equilibrium with the gas phase in the headspace following Henry's law

(Neeleman, 2002), the dissolved oxygen (*DO*), and the oxygen transfer coefficient k_La determine the oxygen transfer rate (*OTR*):

$$OTR = k_l a \cdot (O_2^h - DO) \qquad 9$$

The dynamics of oxygen is much faster than the dynamics of the other relevant processes (e.g. biomass, substrates) (Wang and Stephanopoulos, 1994) and the contribution of the dilution term is small (third term Eq. 7) compared to the rate of change of dissolved oxygen. As a consequence, Eq. 7 is considered in steady-state and *OUR* is calculated every minute during the experiment using Eq. 10:

$$OUR \approx OTR \qquad 10$$

O_2^h is assumed equal to O_2^{in} because of the high aeration rate:

$$O_2^h \approx O_2^{in} \qquad 11$$

The *DO* sensor has a response time around 20 seconds, and is modelled as a first order system:

$$\frac{dDO_{sensor}}{dt} = -\frac{1}{\tau_{sensor}} \cdot DO_{sensor} + \frac{1}{\tau_{sensor}} \cdot DO \qquad 12$$

Dissolved oxygen is controlled by changing the incoming oxygen fraction (O_2^{in}) in an oxygen/air/nitrogen mixture. The controller is based on an already installed PI controller:

$$O_2^{in} = K_C \cdot (DO_{set} - DO_{sensor}) + \int_0^t \frac{K_C}{\tau_I} \cdot (DO_{set} - DO_{sensor}) dt + C_N \qquad 13$$

Specific growth rate control system

Figure 2 gives an overview of the proposed control system for fed-batch cultivation. It is the same as the scheme presented by Soons et al. (2006), except that it is extended with an automatic tuning method. The control purpose is to regulate the specific growth rate (μ) to a desired value by adding a feed with limiting substrates. The grey block defines the specific growth rate controller, which will be briefly reviewed below.

The definition of the specific growth rate controller given in Soons et al. (2006) started from a second order reference model:

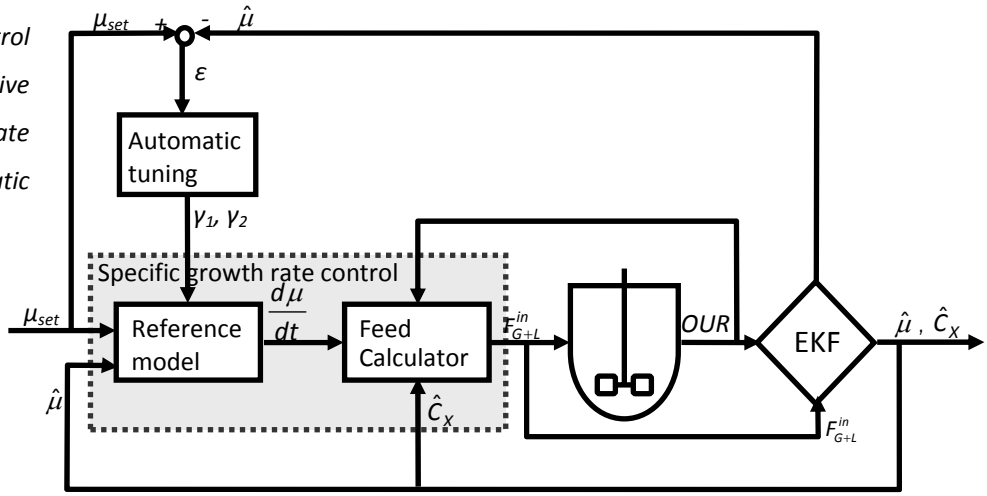


Figure 2. *The control system for adaptive specific growth rate control with automatic tuning.*

$$\frac{d\hat{\mu}}{dt} = \frac{\mu_{set} - \hat{\mu}}{\gamma_1} + \int_0^t \frac{\mu_{set} - \hat{\mu}}{\gamma_2} dt \qquad 14$$

where the tuning parameters γ_1 and γ_2 determine the convergence speed of the controller and therefore controller performance. The reference model is stable if γ_1 and γ_2 are strictly positive numbers. Combination with the process model yields an adaptive "PI" controller:

$$F_{G+L}^{in} = \frac{ac + bd}{aC_G^{in} + bC_L^{in}} \hat{C}_X V + K_P \cdot (\mu_{set} - \hat{\mu}) + \int_0^t K_I \cdot (\mu_{set} - \hat{\mu})\, dt \qquad 15a$$

where the controller gains K_P and K_I are adjusted online to the changing volume:

$$K_P = \frac{V}{\gamma_1 (aC_G^{in} + bC_L^{in})} \text{ and } K_I = \frac{V}{\gamma_2 (aC_G^{in} + bC_L^{in})} \qquad 15b$$

a, b, c, and d are constants depending on the micro-organism (in this work *Bordetella pertussis* (Soons et al., 2006)).

Biomass and specific growth rate measurements are often not available online. Therefore, an Extended Kalman Filter (EKF) estimates specific growth rate ($\hat{\mu}$) and biomass ($\hat{C}_X$) using the measured oxygen uptake rate every minute (Soons et al., 2006, 2007). This software sensor showed fast convergence and fitted well to the data (see Appendix A for an application on continuous cultivation). Therefore, for controller design the estimations of the EKF are taken as replacement of actual measurements ($\mu = \hat{\mu}$ and $C_X = \hat{C}_X$).

Central point of this work is the extension of the specific growth rate controller (Eq. 15) with three methods for automatic tuning, which adapt the controller parameters γ_1 and γ_2 on the

basis of the error in specific growth rate (see Fig. 2). So the approach does not require probing which can disturb the critical variables in biopharmaceutical production.

Controller evaluation

Evaluation methods

The model (Eqs. 1-13) is used to evaluate controller performance in simulations. To obtain realistic tests of the robustness of the controller and performance of the automatic tuning several disturbances have been introduced in the simulation as in Soons et al. (2006): time-varying or drifting kinetics, initialization errors on the controller tuning parameters, and noise. We assumed a drift plus a sinus with different frequencies on the model parameters, whereas in the controller, the model parameters *p* were set constant:

$$p^d = p + a_1 t + a_2 \sin(a_3 \pi t + a_4 \pi) + a_5 \sin(a_6 \pi t + a_7 \pi) \qquad 16$$

Where p^d are the disturbed parameters for μ_{max}, μ_{enh}, Y_{G1}, Y_{G2}, Y_L, C_G^{in}, C_L^{in}, K_G, and K_L. The parameters and initial values are given in Table 1, the disturbances on the parameters in Table 2.

Table 1. *Model parameters of the dual substrate model for B. Pertussis*

V_0 = 3 *l*	Y_L = 0.018 $OD.mmol^{-1}$	Y_{G1} = 0.055 $OD.mmol^{-1}$
C_{X0} = 0.05 *OD*	m_L = 0 $mmol.OD^{-1}.h^{-1}$	Y_{G2} = 0.061 $OD.mmol^{-1}$
C_{G0} = 10 $mmol.l^{-1}$	μ_{max} = 0.12 h^{-1}	m_G = 0 $mmol.OD^{-1}.h^{-1}$
C_G^{in} = 500 $mmol.l^{-1}$	μ_{enh} = 0.055 h^{-1}	Y_O = 0.041 $OD.mmol^{-1}$
C_{L0} = 16.7 $mmol.l^{-1}$	K_G = 0.5 $mmol.l^{-1}$	m_O = 0.41 $mmol.OD^{-1}.h^{-1}$
C_L^{in} = 835 $mmol.l^{-1}$	K_L = 0.5 $mmol.l^{-1}$	$\hat{\mu}_O = 0.16\ h^{-1}$

In practice, dissolved oxygen measurements contain observation noise. In the simulations noise is introduced by adding white noise with an intensity proportional to *OUR* (0.1 initially to 3% at end). This intensity was chosen to mimic the observed fact that *DO* is noisier towards the end:

$$DO^m_{sensor}(t_k) = DO_{sensor}(t_k) + v(t_k) \qquad 17$$

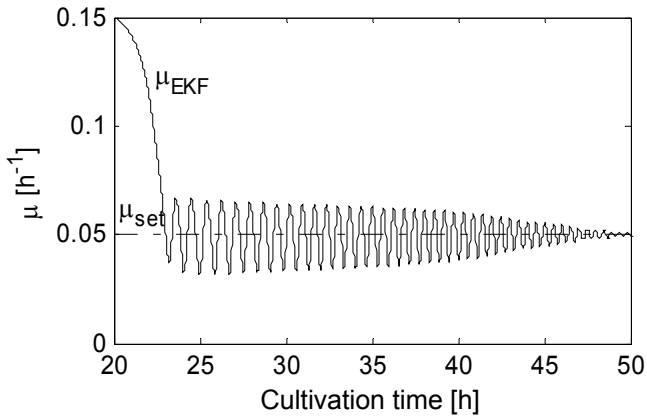
In addition to Soons et al. (2006), controller tuning parameters γ_1 and γ_2 are initialized far below the required values, which results in underdamped behaviour. When the specific

growth rate controller without automatic tuning method is applied with these settings, oscillations were present for almost 30 hours (Fig. 3), underlining the need for additional adaptation. Experimental data showed similar profiles (Fig. 1). Additional simulations are performed with different values for the tuning parameters.

Table 2. *Disturbances on the model parameters for B. Pertussis*

Disturbance parameter	Value From	to
a_1	0	$\pm 1.10^{-3}p$
a_2	0	$\pm 0.2p$
a_3	0	0.1
a_4	0	1.2
a_5	0	$\pm 0.3p$
a_6	0	0
a_7	0	1.2

Figure 3. *Illustration of a simulation with specific growth rate control for poorly tuned and fixed γ_1 and γ_2.*



Evaluation criteria

To evaluate the significance of the automatic tuning system, two criteria are used. The criteria reflect the performance of the specific growth rate controller. In addition the simulation and experimental results are evaluated by comparing graphs.

Babuška et al. (2003) used the following measures to detect a potential loss of performance. The first measure is the mean absolute error over the past *N* points of a moving window at time point *k*:

$$E(k) = \frac{1}{N} \cdot \sum_{k-N}^{k} |\varepsilon(k)| \qquad 18$$

The second is the oscillation measure and is defined by

$$O(k) = \frac{1}{N} \cdot \left(\sum_{k-N}^{k} |\varepsilon(k)| - \sum_{k-N}^{k} \varepsilon(k) \right) \qquad 19$$

where the error ε is the set point minus the estimated specific growth rate ($\mu_{set} - \hat{\mu}$).

Literature

Literature overview

Åström and Wittenmark (1995) give examples of model-based adaptive control to adapt control parameters to changing process behaviour. The adaptation of the controllers for a bioreactor as designed in (Chang, 2003; Levisauskas et al., 1996; Smets et al., 2002, 2004; Van Impe et al., 1995) is also derived from a process model. This solution is limited by the accuracy of the models. Many adaptive controllers require measurement of the state variables such as biomass and substrates (Ignatova et al., 2000; Picó-Marco et al., 2004, 2005; Zlateva, 1997; Krastanov et al., 2003), which are often not available. Other approaches require identification of the system by introducing systematic disturbances (Babuška et al. 2003; Akay et al., 2002). Such “probing” should be avoided in cultivation systems for biopharmaceutical production as it may affect critical variables. Dagci et al. (2001) used sliding mode control to adapt controller parameters. The authors show good control performance during continuous cultivation, but in fed-batch cultivation for vaccine production control performance was not satisfactory.

Classical automatic tuning approaches using PID control also require undesirable disturbances of the process (Åström et al., 1988). Adaptation by a model predictive control approach proposed in Frahm et al. (2002) needs an open-loop model with the disadvantage that accuracy is not guaranteed. Another drawback of some of the reported solutions is the complex implementation in practice.

Industrial applications of adaptive control are often rule-based instead of model-based (Åström et al. 1993). However, for control of complex and difficult industrial processes model-based control approaches have been proven the most effective among the different types of adaptive controllers (Babuška et al., 2003).

Automatic tuning based on the MIT rule

The task of the automatic tuning method is to pursue the best trade-off between tracking behaviour, disturbance rejection, and stability. The automatic tuning method, therefore, should have the following abilities. At one hand, if control parameters γ_1 and γ_2 are initialized too small (Eq. 14), the automatic tuning method must increase the value of γ_1 and γ_2 to avoid underdamped behaviour. On the other hand, the automatic tuning method should reduce the values of γ_1 and γ_2 if the error converges too slowly to zero. The automatic tuning method has to be derived so that a stable system is obtained in the sense that the error remains bounded.

In model-reference adaptive control according to the textbook of Åström and Wittenmark (1995), the MIT rule is used to obtain adaptation. The error between the reference model and process output ($e = \mu_{ref} - \hat{\mu}$) is the driver for the adaptation. Following the MIT rule the parameters γ are adjusted such that a squared error is minimized:

$$\frac{d\gamma}{dt} = -\beta \cdot e \cdot \frac{\partial e}{\partial \gamma} \qquad 20$$

The vector $\partial e / d\gamma$ is the sensitivity of the error with respect to the controller parameter vector γ (error sensitivity). The coefficient β determines the adaptation rate and must be chosen so that the controller parameters change slower than the specific growth rate and dissolved oxygen, but fast enough to obtain adaptation.

Online automatic tuning methods

Adaptation based on the MIT rule is driven by the error e which arises from set-point changes and disturbances. Systematic set-point changes must be avoided in cultivation systems for biopharmaceutical production as it may affect critical variables. As a consequence, in this work, adaptation is only driven by errors caused by disturbances. The model error e equals the set-point error ε and the reference model has become redundant.

The MIT rule requires knowledge of the error sensitivity, which is difficult to derive. Therefore, in literature approximations for the error sensitivity have been made (Åström et al., 1995). Simulations with the outcome of several approximations gave poor results in the cultivation for vaccine production. Therefore alternative adaptation schemes are applied in the next section. Several (weighted) combinations of ε, ε^2, and $\int \varepsilon$ are investigated to approximate the sensitivity derivative:

$$\frac{\partial \varepsilon}{\partial \gamma} = c_1 \cdot 1 + c_2 \cdot \varepsilon + c_3 \cdot \int_0^t \varepsilon \qquad 21$$

Method 1: adaptation based on the error

The most simple mechanism is assuming the sensitivity derivative to be one (Chen et al., 1995; Rocha et al., 2002) ($c_1 = -1$, $c_2 = c_3 = 0$).

$$\frac{d\gamma}{dt} = \beta \cdot \varepsilon \qquad 22$$

which appears to work well for adaptation of estimated process parameters. A test of Eq. 22 revealed that it is not a good mechanism for adaptation of controller parameters, because negative and positive errors will result in opposite adjustments of γ and do not cancel oscillations.

Method 2: adaptation based on the (integral) error

In the second automatic tuning method the error sensitivity $\partial e / d\gamma$ in Eq. 20 is replaced by a combination of the error and the integral error ($c_1 = 0$, $c_2 = -1$, $c_3 = -1$).

$$\begin{aligned} \frac{d\gamma_1}{dt} &= \varepsilon \cdot \left(\beta_2 \varepsilon + \beta_3 \int_0^t \varepsilon dt \right) \\ \frac{d\gamma_2}{dt} &= \varepsilon \cdot \left(\beta_2 \varepsilon + \beta_3 \int_0^t \varepsilon dt \right) \end{aligned} \qquad 23$$

where β_2 and β_3 are a combination of β and c_2 and c_3 ($\beta_2 = \beta c_2$ and $\beta_3 = \beta c_3$). The automatic tuning method (Eq. 23) adapts control parameters γ_1 and γ_2 proportional to ε^2 and $\varepsilon \int \varepsilon$. So, γ_1 and γ_2 are predominately adapted when the quadratic error is large and/or when the error persists. The choice of the adaptation rate β_2 and β_3 is based on simulations. Too large values make adaptation faster than the change of the other variables in the system and make the adaptation mechanism react on instant errors instead of persistent errors, thereby avoiding the "basic" controller (Eq. 15) to counteract on these errors. Too small values of β give slow adaptation, which results in the requirement for multiple tuning runs to upgrade performance. The choice of the adaptation rates for γ_1 and γ_2 is a choice for the designer of the controller, but simulations showed that distinction between the adaptation rates of γ_1 and γ_2 was not necessary.

This online automatic tuning method is applied to simulations of fed-batch cultivations in which γ_1 and γ_2 were initialized too small. By initiating the automatic tuning method at the start of the fed-batch, control performance was significantly upgraded within five hours of cultivation (Fig. 4 shows a typical simulation). The mean absolute error and oscillation measure decreased fast after the start of the fed-batch resulting in upgraded controller performance (Figs. 4C and 4D). Compared to a simulation without adaptation of the controller tuning parameters, oscillations were attenuated five times faster (compare Figs. 3 and 4).

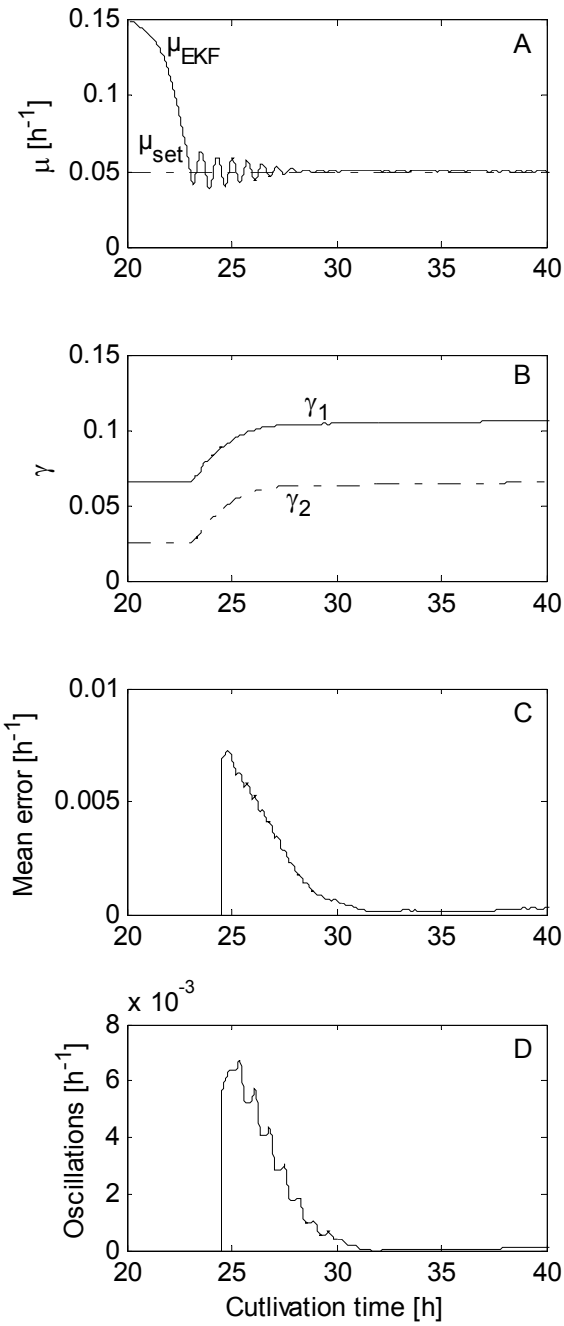


Figure 4. *Simulation results of specific growth rate control with automatic tuning based on the error and the integral error (method 2, Eq. 23). β_2=225, β_3=550. Fed-batch start at t=22 h.* ***A.*** *Specific growth rate (μ)* ***B.*** *Controller tuning parameters.* ***C.*** *Mean absolute error* ***D.*** *Oscillation measure.*

Method 3: adaptation based on the (squared) error

The third automatic tuning scheme consists of a combination of the error and the squared error, normalized by μ_{set}. ($c_1 = -\mu_{set}^{-1}$, $c_2 = -\mu_{set}^{-2}$, $c_3 = 0$).

$$\frac{d\gamma_1}{dt} = \beta_1 \cdot \frac{\varepsilon}{\mu_{set}} + \beta_2 \cdot \frac{\varepsilon^2}{\mu_{set}^2}$$
$$\frac{d\gamma_2}{dt} = \beta_1 \cdot \frac{\varepsilon}{\mu_{set}} + \beta_2 \cdot \frac{\varepsilon^2}{\mu_{set}^2} \qquad 24$$

where β_1 and β_2 are a combination of β and c_1 and c_2 ($\beta_1 = \beta c_1$ and $\beta_2 = \beta c_2$). Analogous to the previous section no distinction is made in β_i-values for the adaptation rates for γ_1 and γ_2. The automatic tuning method (Eq. 24) adapts control parameters γ_1 and γ_2 proportional to the error ε and proportional to the quadratic error ε^2, which is large if the error is large.

The fed-batch phase of the simulation in Fig. 5 started at *22 h*. At the same time the specific growth rate controller was switched on with initial values of γ_1 and γ_2, which were chosen too low as in Fig. 3. The graph shows that these values resulted in underdamped responses (Fig. 5A), but the adjustment of γ_1 and γ_2 (Fig. 5B) was so strong that within 5 hours the deviations were minimal and oscillations were cancelled (Fig. 5D). Figure 5C shows that controller performance was significantly upgraded with respect to the mean absolute error. Note that the mean absolute error and the oscillation measure were delayed 1.5 hours to calculate their measures. Disturbances due to time-varying parameters and noise do not influence controller performance.

Although the MIT rule was the starting point for derivation of the online automatic tuning method, concepts applying directly the MIT rule did not give satisfactory results. The alternatives presented in the last two sections were capable to upgrading performance.

Comparison method 2 and 3 for different β and γ

Adaptation using the integral error as well as adaptation using squared error give similar results for γ_1 and γ_2 (Figs 4-5). Both adaptation mechanisms require five hours to upgrade controller performance and to cancel underdamped behaviour. Controller tuning parameters γ_1 and γ_2 converge to similar values, but adaptation using the integral error is smoother.

Figure 5. *Simulation results of specific growth rate control with automatic tuning based on the error and the squared error (method 3, Eq. 24), $в_1$=0.25, $в_2$=0.5. Fed-batch start at t=22 h.* ***A.*** *Specific growth rate (μ)* ***B.*** *Controller tuning parameters.* ***C.*** *Mean absolute error* ***D.*** *Oscillation measure.*

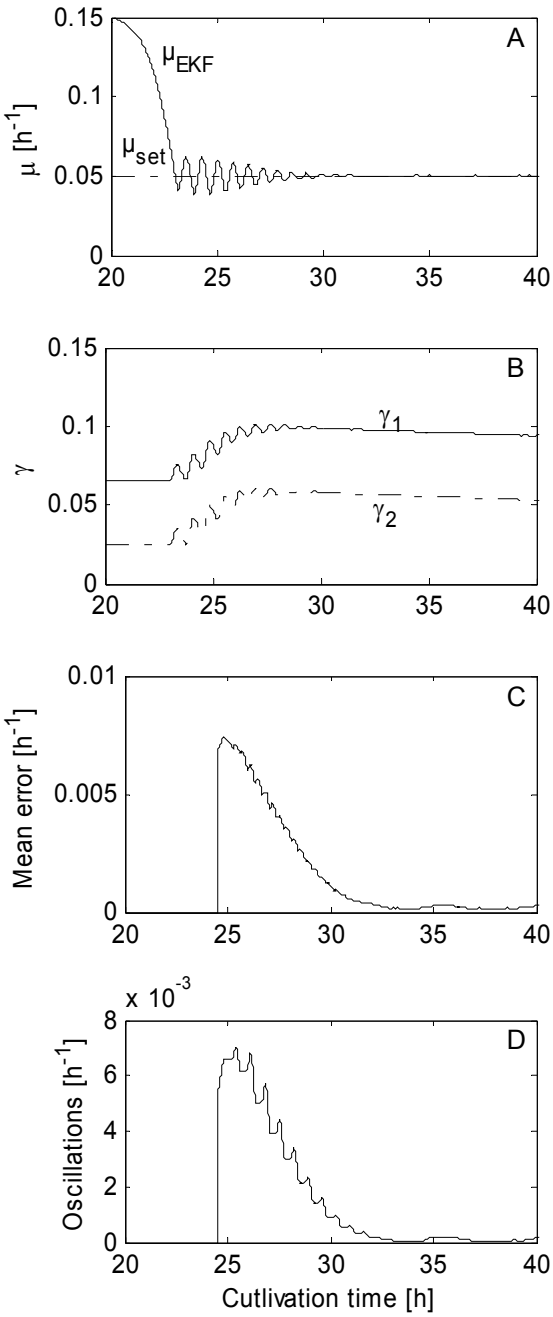


Figure 6 shows simulations with different initialized values of the controller parameters *γ*. Figure 6AB illustrates once again the adaptation of γ_1 and γ_2 initialized with too small values. Simulations initialized close to the end-values of the previous run (Fig. 6AB) showed that the adaptation mechanisms leave the tuning parameters almost unchanged (Fig. 6CD). Similar adaptation results are found for simulations initialized with large initial parameters (Fig. 6EF). Large tuning parameters slow down the controller actions, but only marginally

decrease controller performance. As a result, the driving force (or error) for adaptation is small and so is the adaptation of controller parameters.

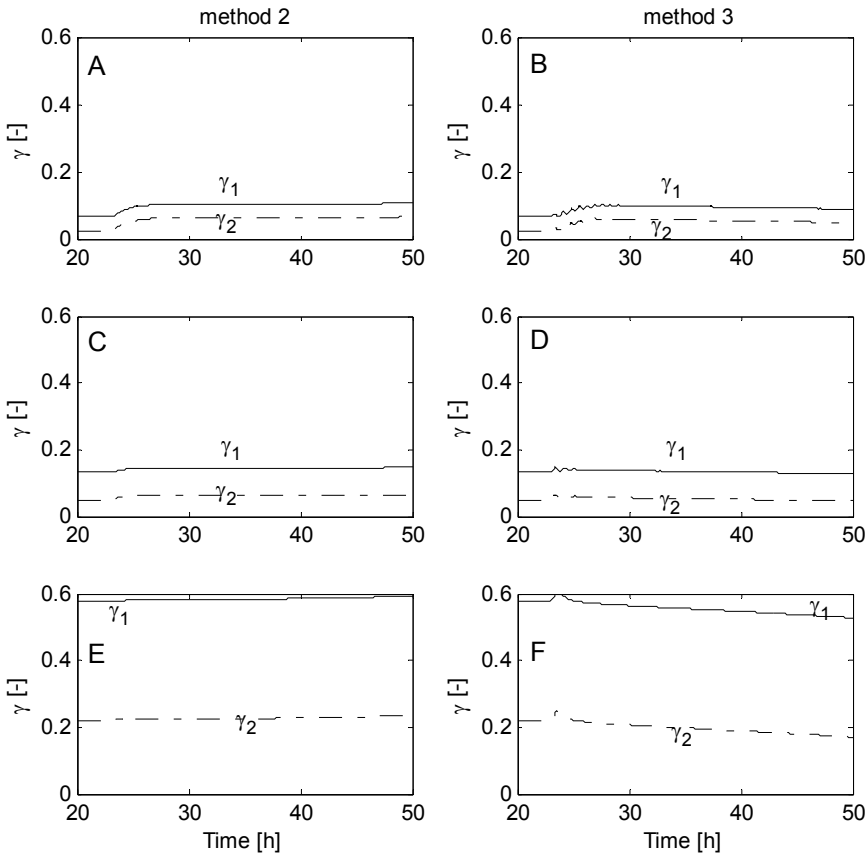


Figure 6. *Tuning parameters γ_1 and γ_2 for different initializations of γ. Fed-batch start at t=22 h.* ***ACE.*** *method 2 (integral errors, Eq. 23, β_2=225, β_3=550.).* ***BDF.*** *method 3 (squared errors Eq. 24, β_1=0.25, β_2=0.5).* ***AB.*** *γ_1 and γ_2 initialized too small* ***CD.*** *γ_1 and γ_2 initialized close to the end-values of the previous run.* ***EF.*** *γ_1 and γ_2 initialized too large.*

Figure 7 shows additional simulations with different values of the adaptation rate β. Figure 7CD illustrates once again the effect of the chosen adaptation rate β on μ and γ as a reference. Small values for the adaptation rate β give slow adaptation of controller parameters γ (Fig. 7AB). The oscillations are attenuated two to three times slower compared with the chosen β and about two times faster relative to fixed γ (β=0, compare with Fig. 3). If β is chosen larger, adaptation becomes faster and larger values for γ are obtained (Fig. 7E). Large tuning parameters slow down the controller actions and may give small offsets. The adaptation rate β in Fig. 7F is chosen too large and gives interference with other dynamics (dissolved oxygen and specific growth rate). So, if the adaptation rate β is not chosen too fast, both methods show robust adaptation.

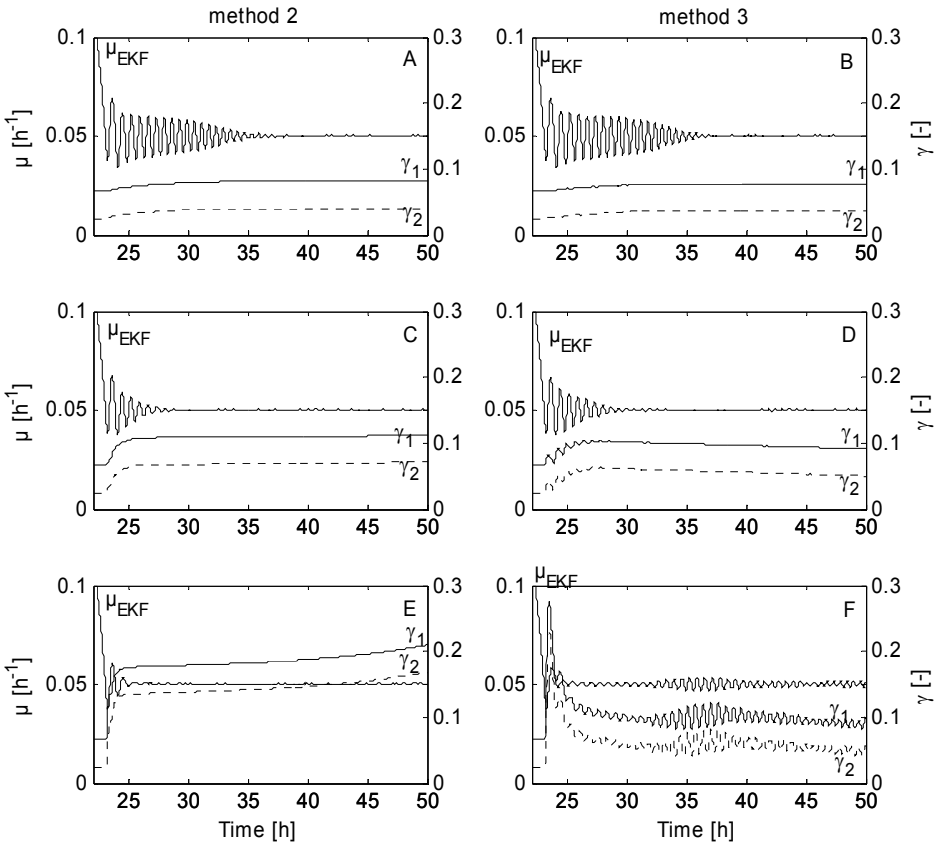


Figure 7. *The effect of different β values on adaptation of tuning parameters γ_1 and γ_2 (right axis) and on μ (left axis). Fed-batch start at t=22 h.* ***CD.*** *β initialized with the chosen values.* ***C.*** *β_2=225, β_3=550. D. β_1=0.25, β_2=0.5.* ***ACE.*** *method 2 (integral errors, Eq. 23).* ***BDF.*** *method 3 (squared errors Eq. 24)* ***AB.*** *β initialized ten times smaller EF. β initialized ten times larger.*

Stability

In previous sections the automatic tuning mechanisms were evaluated by simulations. Although performance is good the effect of changing the tuning parameters on stability is yet unknown. Bastin and Dochain (1990) state that stability of fed-batch cultivations is not an issue because the exponentially increasing biomass and volume make the system inherent unstable and result in positive poles. However, because the positive poles can be cancelled in the controller loops it makes sense to examine stability properties of the control loops.

Dagci et al. (2001) “solve” the stability issue by examining the stability of the control loop; others (Soons et al 2006, 2006a; Smets et al. 2002] refer to the stability of the reference model. In this work, stability is examined by considering the poles of the closed loop. The full system is given by Eqs. 3, 4, 7, 8, 9, 12, and the derivative of Eq. 1 including the controller Eqs. 13 and 15. The following time-varying state space model represents the linearized form of this system along the trajectory of fed-batch cultivation (see appendix B for the equations and the linearization of the system):

$$\frac{dx}{dt} = A(t)x + B(t)u$$
$$y = C(t)x \qquad 25$$
$$u = F(t)x$$

with x the state variables, y the output variables, u the input variables, A the system matrix, B the input matrix, C the output matrix, and F the state feedback matrix.

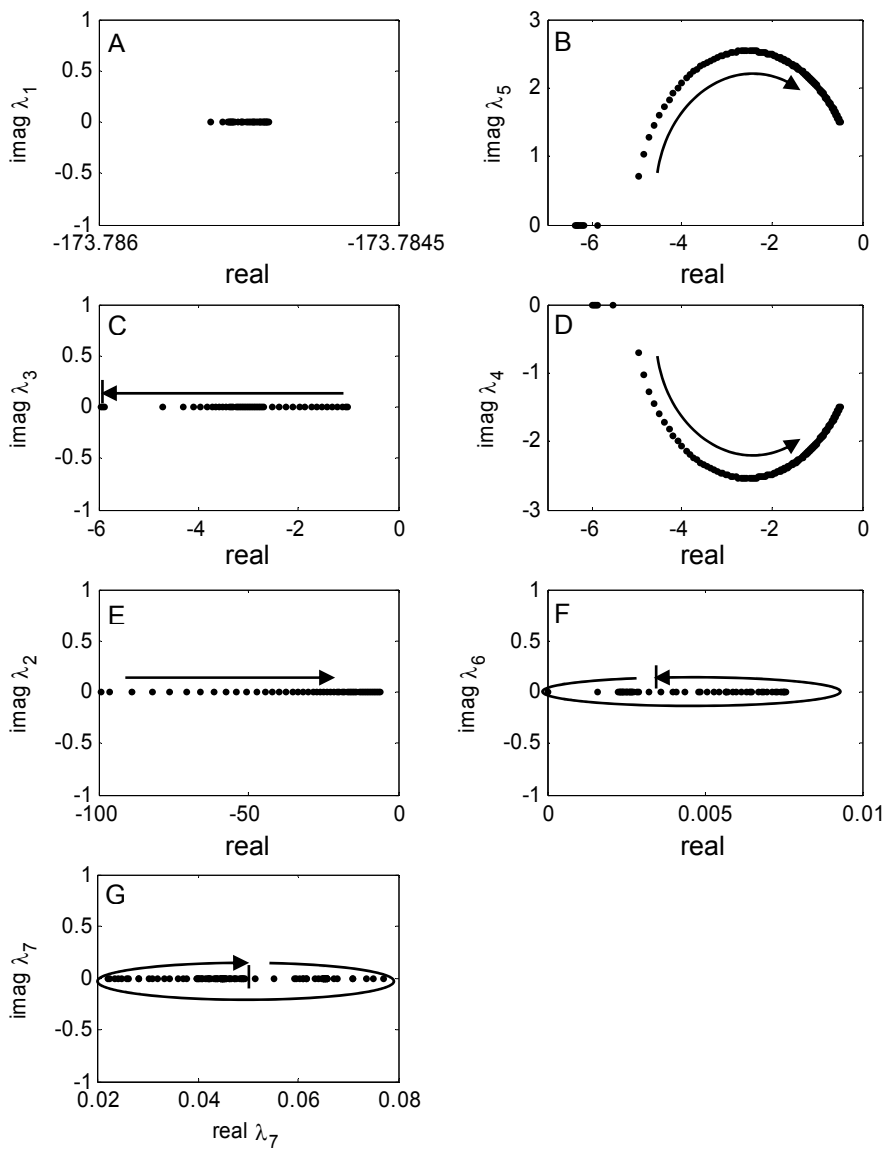


Figure 8. *Example of real and imaginary poles of the closed loop system for a range of γ_1 and γ_2 calculated by offline simulation of the process at one time point along the trajectory. The arrows indicate the direction in which the poles move with increasing controller tuning parameters γ_1 and γ_2.*

Experimental results

Controlled fed-batch cultivations with the dual substrate consuming bacterium *B. pertussis* (causative agent of whooping cough disease) were performed in a 7-liter laboratory bioreactor with a chemically defined medium containing glutamate and L-lactate as the main carbon sources (Thalen et al., 1999). Bioreactor conditions, analysis, and software were applied as reported in Soons et al. (2006). The fed-batch, the μ controller, and the automatic tuning started automatically when the limiting substrates were almost depleted and μ dropped to the set-point.

Control performance of the automatic tuning method and the specific growth rate controller were evaluated for the best methods of the simulations: method 2: based on the (integral) error (Eq. 23) and method 3: based on the (squared) error (Eq. 24).

Method 2: adaptation based on the (integral) error

The automatic tuning method based on the (integral) error (Eq. 23) was implemented and tested by initializing control parameters γ_1 and γ_2 with values, which turned out to be too small in practice. In simulations it turned out that the adaptation parameter γ is slightly increasing as the experiment proceeds towards the end. This makes the γ correction mechanism more sensitive to errors than necessary. Therefore, in order to safeguard the experiment, it was decided to use in the experiment Eq. 26 instead of Eq. 23, which leads to down-tuning of the adaptation towards the end of the cultivation.

$$\begin{aligned} \frac{d\gamma_1}{dt} &= \varepsilon \cdot \left(\beta_2 \frac{\varepsilon}{\gamma_1} + \beta_3 \int_0^t \frac{\varepsilon}{\gamma_2} dt \right) \\ \frac{d\gamma_2}{dt} &= \varepsilon \cdot \left(\beta_2 \frac{\varepsilon}{\gamma_1} + \beta_3 \int_0^t \frac{\varepsilon}{\gamma_2} dt \right) \end{aligned} \qquad 26$$

Evaluation after the experiment showed that the difference with the original γ adaptation is small, so that, on retrospect, the precaution was unnecessary.

The laboratory experiment showed that by initiating the automatic tuning method at the start of the fed-batch at about 20 hours control performance was upgraded (Fig. 9). The mean absolute error and oscillation measure decrease with time. The adaptation of γ_1 and γ_2 continues during the whole fed-batch cultivation, but the changes in the first period are the strongest and at the end the values are close to the steady state.

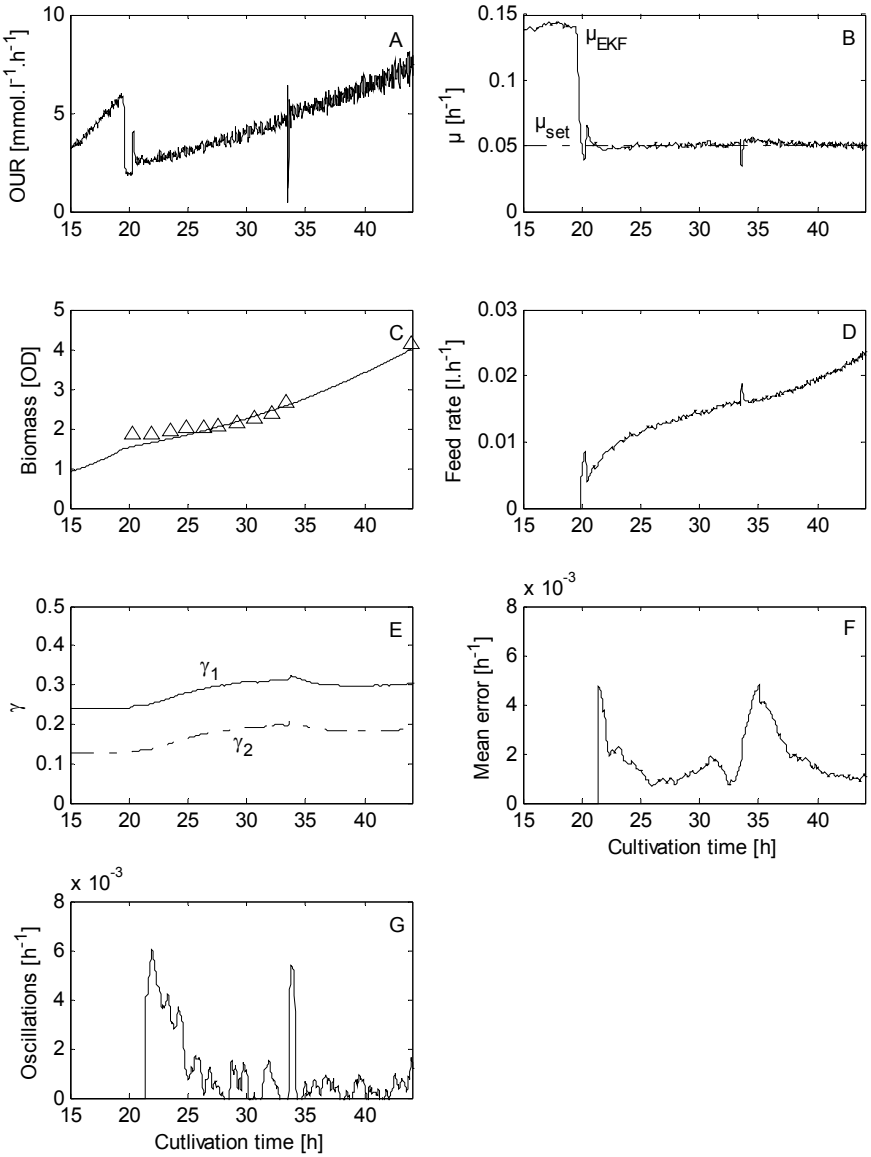


Figure 9. *Results for laboratory experiments with automatic tuning based on (integral) errors (method 2, Eq. 26). Fed-batch start at t=19.7 h, β_2=20, β_3=20.* ***A.*** *Oxygen uptake rate (OUR)* ***B.*** *Specific growth rate (μ)* ***C.*** *Biomass* ***D.*** *Substrate feed rate* ***E.*** *Controller tuning parameters.* ***F.*** *Mean absolute error.* ***G.*** *Oscillation measure.*

A standard deviation of 0.005 h^{-1} on μ is obtained during the fed-batch phase and indicates that the long-term performance is good. The controller maintained μ close to μ_{set} in the presence of various uncertainties including disturbances on dissolved oxygen consumption, uncertain parameters, and initialization errors.

An exponentially increasing feed rate (Fig. 9D) was added into the bioreactor to cope with the exponentially increasing biomass (Fig. 9C). At *t=33.5h* agitation speed was increased with *100rpm* to meet the increasing the oxygen demand during the fed-batch phase. The resulting peak in the dissolved oxygen concentration caused a disturbance in the estimated specific growth rate at *t=33.5h* and in the mean absolute error and oscillation error. The controller parameters were hardly adapted and the specific growth rate was properly controlled until the end of the cultivation.

At *t=40h* the mean absolute error and the oscillation measure became approximately constant. As an option a performance monitor can be added to the system, which can be used to evaluate the performance criteria and as a manager to (de)activate the automatic tuning method. The first function of the performance monitor is to qualify controller performance using the criteria given in Eqs. 18 and 19. Next, based on the qualification, the performance monitor decides to retune the process using an automatic tuning method or to continue with the current settings. If control performance is upgraded and the performance measures are constant, tuning is finished and the specific growth rate controller can continue with the obtained settings. In the experiment at *t=40h* in Fig. 9, the performance monitor would decide to continue the cultivation with the current settings for γ_1 and γ_2 and to deactivate the automatic tuning method. Subsequently, the following cultivations can be performed with the obtained settings.

Method 3: adaptation based on the (squared) error

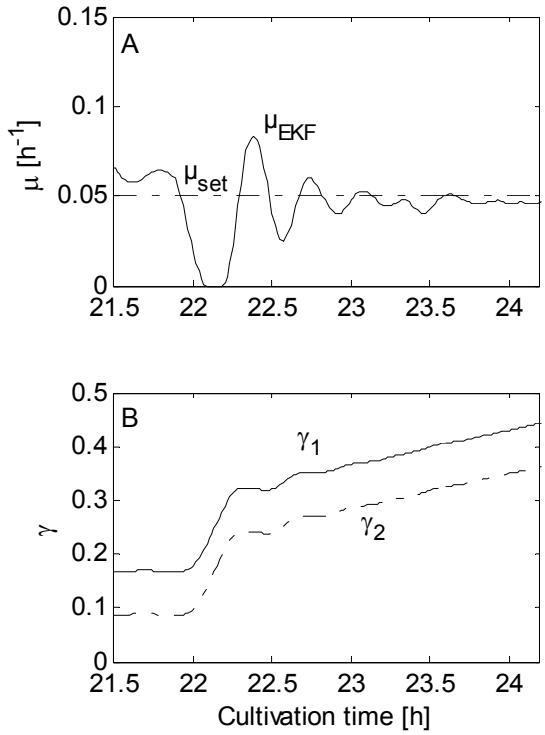
Figure 10 shows a part of a long-term experiment direct after initialization of the automatic tuning controller. The oscillations that occur due to an incorrect choice of γ_1 and γ_2 vanish within three hours. The offset at the end of this period disappears in some hours. The tuning method is doing its task well.

The adaptation of γ_1 and γ_2 continues and has not come to steady state after three hours. The oscillations that were observed in the simulations for this tuning method in the courses of γ_1 and γ_2 are hardly present in the experiment (compare Fig. 10 with Fig. 6)

Comparing the simulation (Fig. 6) with the experimental results (Fig. 10), it can be observed that the tuning parameters γ_1 and γ_2 giving the best performance during laboratory experiments are a factor five to six larger than the tuning parameters giving the best performance during simulations. Mismatches between experimental-simulation are common for biological processes. The experiments show that the automatic tuning method compensates for these mismatches by adapting the controller settings such that the desired behaviour is acquired.

Calculation of the poles of the closed loop for the laboratory experiments shows that the tuning parameters are adapted into the proper direction within the stability area. These results confirm the closed loop stability, which was already calculated for the simulated process.

Figure 10. *Detail of results for laboratory experiments with adaptation based on the (squared) error (method 3, Eq. 24, β_1=0.25, β_2=0.5). Fed-batch start at t=21.7 h.* ***A.*** *Specific growth rate (μ)* ***B.*** *Controller tuning parameters.*



Conclusions

The challenge was to upgrade performance of poorly acting controllers. Online automatic tuning has this capability. Three automatic tuning methods were applied to upgrade control performance for specific growth rate control during fed-batch cultivation. The best two methods were qualified by laboratory experiments for *B. pertussis*. In the first method, the adaptation rate is proportional to the product of the error and the integral error. The second uses an adaptation rate proportional to the error and the squared error. The methods do not require online identification, thus avoiding the need for process perturbation and complex implementation. The qualifications of the controller and automatic tuning method are:

Control performance is evaluated online on the basis of the current mean absolute error and oscillation measure. A performance monitor can be used to activate or deactivate the automatic tuning device.

If control performance is poor, application of automatic tuning yields good performance within five hours by adapting the controller parameters such that the mean absolute error and oscillation measure decrease at least ten-fold. The automatic tuning methods are robust against disturbances among others noise, parameter uncertainties, and initialization errors.

We expect that a cultivation controlled at the desired specific growth rate will result in smaller variations in end quality (vaccine titer) and thus yield a better product (vaccine).

The closed loop with automatic tuning is stable at any point along the trajectory of fed-batch cultivation.

Adaptation of control parameters is straightforward, fast and accurate. This feature, furthermore, is neither specific to fed-batch cultivation nor for specific growth rate control and hence can also be applied to other systems.

The applied automatic tuning methods improve controller performance and reduce the tuning effort by automatically adjusting the tuning parameters in one or more tuning runs.

Acknowledgements

Wageningen University and the Netherlands Vaccine Institute (NVI) work together in a project with Applikon Biotechnology BV and Siemens NV in order to improve the production process by release of biopharmaceuticals on the basis of new online techniques. The Dutch Ministry of Economical Affairs (TSGE3067) funds the project. Experiments are performed at NVI.

CHAPTER 6

Scaling-up vaccine production: implementation aspects of a biomass growth observer and controller

Submitted to Bioprocess and Biosystems Engineering as: Zita I. T. A. Soons, Jan van den IJssel, Leo A. van der Pol, Gerrit van Straten, and Anton J. B. van Boxtel

Abstract

This study considers two aspects of the implementation of a biomass growth observer and specific growth rate controller in scale-up from small- to pilot-scale bioreactors towards a feasible bulk production process for whole-cell vaccine against whooping cough. The first is the calculation of the oxygen uptake rate, the starting point for online monitoring and control of biomass growth, taking into account the dynamics in the gas-phase. Mixing effects and delays are caused by amongst others the headspace and tubing to the analyzer. These gas phase dynamics are modelled using knowledge of the system in order to reconstruct oxygen consumption.

The second aspect is to evaluate performance of the monitoring and control system with the required modifications of the oxygen consumption calculation on pilot-scale. In pilot-scale fed-batch cultivation good monitoring and control performance is obtained enabling a doubled concentration of bulk vaccine compared to the standard batch production.

Keywords

Scale-up, monitoring and control, Bordetella pertussis, gas-phase dynamics, fed-batch, oxygen uptake rate

Nomenclature

a, b, c, d	constants for dual substrate model of *B. pertussis*
C_A	acetoacetate concentration [$mmol.l^{-1}$]
C_G, C_L	glutamate and lactate concentration respectively [$mmol.l^{-1}$]
C_X	biomass concentration [OD_{590}]
C_O^L	oxygen concentration in bioreactor [$mmol.OD^{-1}.h^{-1}$]
C_O^*	oxygen concentration at gas-liquid interface [$mmol.OD^{-1}.h^{-1}$]
CSTR	continuous stirred tank reactor
E	mean absolute error [h^{-1}]
F_{calc}	feed rate by prior calculation [$l.h^{-1}$]
F_G	gas flow rate [$l.h^{-1}$]
F_{G+L}	substrate feed rate (glutamate + lactate) [$l.h^{-1}$]
F_I, F_P	proportional, integral action feed rate [$l.h^{-1}$]
H	Henry coefficient [$Pa.l.mmol^{-1}$]
k_La	oxygen transfer coefficient [h^{-1}]
m_O	maintenance coefficient on oxygen [$mmol.OD^{-1}.h^{-1}$]
M&C	monitoring and control
N	length of moving window
O	oscillation measure [h^{-1}]
O_2	oxygen fraction [-]
OD	optical Density at 590nm
OTR, OUR	oxygen transfer rate, oxygen uptake rate [$mmol.l^{-1}.h^{-1}$]
p	pressure [Pa]
t	cultivation time [h]
T	sampling time [h]
T_D	transport delay [h]
V	liquid volume [l]
V_{tot}	volume total gas-phase [l]
V_{tubing}	volume tubing [l]
u, x, y	inputs, states, outputs
Y_O	biomass yield on oxygen [$OD.mmol^{-1}$]
ε	error between set point and process $\mu_{set} - \hat{\mu}$ [h^{-1}]

Greek letters

β	adaptation rate

γ_1, γ_2	tuning parameters for μ control
μ	specific growth rate [h^{-1}]

Superscripts and subscripts

^, $_{EKF}$	observed values
$_a$	auxiliary
$_{in}$	inlet
$_m$	measured (in outlet stream)
$_{mix}$	outlet stream incl. mixing effect in the total gas phase
$_{out}$	outlet stream of the bioreactor
$_{set}$	set-point
$_p$	predicted values

Introduction

Monitoring and control (M&C) systems in biotechnology are usually designed and tested in laboratory-scale experiments. Its application on production-scale is limited in literature. This work investigates applicability of the monitoring and control system, developed for small-scale in previous work (Soons et al., 2006, 2007a, 2008), on larger-size pilot-scale cultivations. The system for monitoring and controlling fed-batch cultivations is shown in Fig. 1 (Soons et al., 2007a). The system consists of:

- Computation of the oxygen uptake rate (*OUR*) being an essential “measurement” and the starting point for the advanced monitoring and control system.
- Observer for specific growth rate and biomass concentrations. An observer is needed, because there is no sensor available for biomass growth.
- Controller to track the set-point for specific growth rate during the dynamic fed-batch cultivation by manipulating the feed.
- Automatic tuning method to tune the specific growth rate controller automatically without human interference or introduction of excitation signals. Tuning of controllers is necessary to ensure that the process is properly controlled at the desired set-point and does not oscillate or become sluggish due to e.g. interference with other processes like dissolved oxygen.
- Performance monitor to monitor whether the controller is doing its task well and to decide whether to switch the auto-tuning on or off.

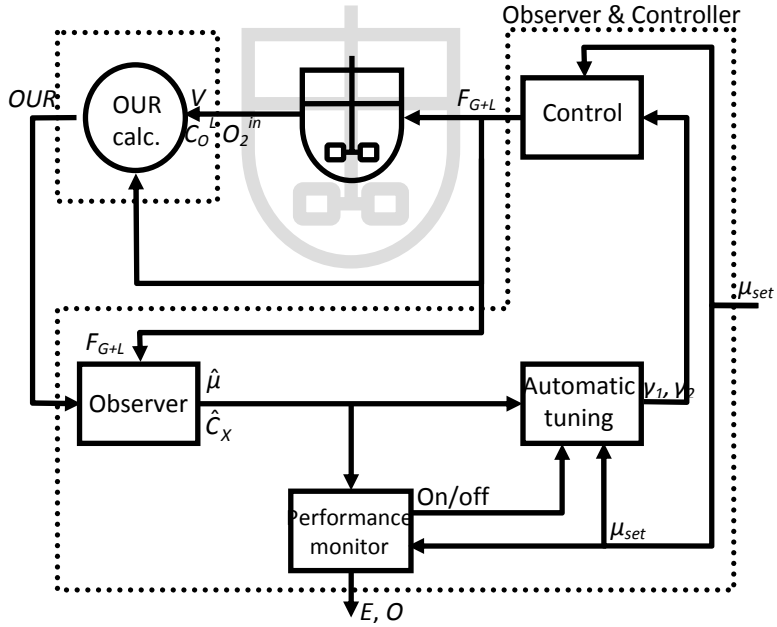


Figure 1. *Schematic monitoring & control system. The dotted boxes indicate the two parts considered for scale-up from laboratory (5L) to pilot-scale cultivations (60L): the OUR calculation and the implementation of the observer and controller for biomass growth.*

The specific growth rate controller with auto-tuning and the observer have already been described in Soons et al. (2006, 2007a, 2008) and are summarized in appendix A. The implementation of the monitoring & control system for scale-up of the vaccine production process to a 60L-fed-batch bioreactor is treated in two parts in this paper (see the dotted boxes in Fig. 1). The first is the computation of the oxygen uptake rate, which needs to be modified before it can be applied on pilot-scale due to mixing effects and delays in the gas-phase, as will be shown in the sequel. The second part is the implementation of the observer and auto-tuning controller for biomass growth – developed and validated on small-scale cultivations – on the pilot-scale system. The cultivated biomass is the basis for vaccines currently being developed at the Netherlands Vaccine Institute.

Dynamic *OUR* calculation

Currently, most biopharmaceuticals are produced in a batch or fed-batch cultivation. The essential quality of the product is formed in this step and is the result of the metabolic state of the micro-organisms. It is therefore essential to measure the metabolic state of the process. Metabolic activity is difficult to measure directly due to the lack of sensors, but respiration can be monitored by the oxygen mass balance. The oxygen uptake rate can in turn be used to reconstruct the specific growth rate and biomass using an observer. The oxygen uptake rate for small-scale systems is usually calculated using Eq. 1, e.g. Wang and Stephanopoulos (1984):

$$OUR \approx OTR = k_L a \cdot \left(C_O^* - C_O^L \right) \qquad 1$$

where C_O^* and C_O^L are the oxygen concentrations at the gas-liquid interface and in the cultivation broth respectively. This method of *OUR* calculation may, however, be inappropriate for pilot- and large-scale systems, because the bioreactor may not be ideally mixed and the oxygen transfer coefficient $k_L a$ is often not accurately known on the large-scale.

Other approaches require switching of the aeration gas composition, e.g. Casa López et al. (2006) and thus introducing systematic disturbances. Such probing should be avoided in cultivation systems for biopharmaceutical production as it may affect critical variables. Furthermore, the approach may be less suited for dynamic systems in the view of online monitoring and control, where *frequent* oxygen consumption "measurements" are necessary.

An alternative that eliminates the need to know $k_L a$ is the computation using off-gas measurements. Often the computation of *OUR* using off-gas measurements is based on a steady-state assumption (for example in Estler, 1995; Levisauskas, 2001; Siegell and Gaden, 1962; Tatiraju et al., 1999; Spriet et al., 1982; Lubenova, 1999; Ghoul et al., 1991; Baart et al., 2007):

$$OUR \approx OTR = \frac{F_G}{V} \cdot \left(O_2^{in} - O_2^m \right) \cdot \frac{p}{H} \qquad 2$$

where F_G is the gas flow rate, O_2^{in} the oxygen fraction in the inlet stream, O_2^m the measured oxygen fraction in the outlet stream, p is the pressure in the bioreactor, and H the Henry coefficient. These off-gas fractions, however, do not reflect the actual concentrations in the bioreactor under dynamic conditions (Bloemen et al., 2003). This is due to the presence of a headspace and tubing, which cause additional dynamics in the off-gas system. Figure 2 shows an overview of the bioreactor and measurement system as it is used in pilot-scale vaccine production. The cells are grown in the liquid-phase of the reactor. The gas that leaves the cultivation broth cannot be analyzed directly. First it enters the headspace (necessary for e.g. foaming), next the centrifuge (for mechanically defoaming), the incinerator (for killing (pathogenic) micro-organisms), the cooler (for drying before the gas can be analyzed) and finally the analyzer. All parts are connected by tubing, which is represented by the tubing box.

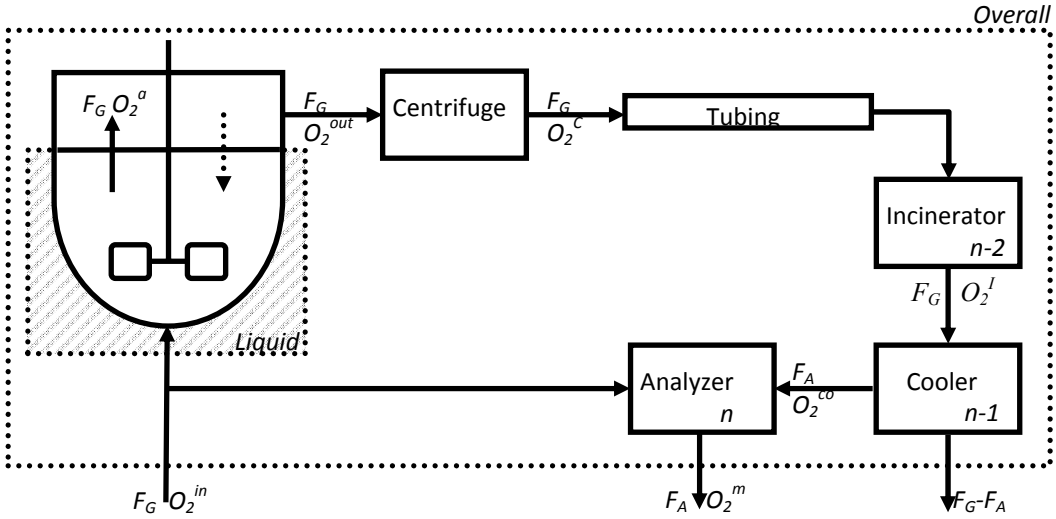


Figure 2. *Schematic of the bioreactor and measurement system.*

Bloemen et al. (2003) include the liquid and gas phase to model the oxygen uptake rate. The off-gas system is modelled as a series of *n* ideally stirred tank reactors (CSTR) plus a time delay. The number of CSTRs and other parameters are fitted using identification experiments. Some parameters of the off-gas system are no longer physically relevant. The method is applied to chemostat cultures, in which the time constants are constant. In batch or fed-batch cultivation, however, the time constants are time-varying due to the exponential growth and the exponential increasing demands for oxygen and substrates. Exact knowledge of k_La is required in Bloemen's approach. In contrast to CSTR, where k_La is approximately constant, k_La is time-varying (fed-)batch cultivations due to variations in stirrer speed, gas flow rate, broth volume, cells, viscosity, products, etc, and is often poorly known in large- or pilot-scale systems.

As an alternative, in this work a method for the calculation of *OUR* from off-gas measurements is applied by only considering the gas phase and its dynamics. It is based on a white box model containing parameters that are easy to retrieve from the system. The method is therefore applicable without the need for new identification runs. Moreover, the method does not require knowledge of k_La. In fact, by combining off-gas data and measurements of dissolved oxygen, k_La can be estimated if needed.

In the *OUR* calculation we are interested in the oxygen fraction at the input of the system, i.e. the auxiliary oxygen fraction entering the headspace, while the measurements take place at the output, i.e. oxygen fraction in the analyzer. The theory and application of the reconstruction of the input using measurements of the output as well as the underlying

model are given in the section "Online reconstruction of the inputs and biomass growth". The values of the (system) parameters are given in appendix B. The biomass growth observer and controller are given in appendix A. The new method for *OUR* computation is an enhancement to the biomass growth observer and controller developed earlier for small-scale laboratory equipment where Eq. 2 is sufficient. Therefore it is relevant to test the performance of the enhance system on real experiments in larger scale. This aspect is described in section "Results and discussion".

Materials and methods for the experiments

Experiments for off-gas model validation

The experiments for validation of the model for the off-gas dynamics (Eqs. 5-7) are performed in a 60-L bioreactor containing 30-40 L demineralised water without bacteria. The inlet gas is let into the headspace to mimic the auxiliary oxygen concentration. The inlet gas is excited with known inputs (by changing the gas composition, the gas flow rates, and the volume of the headspace and liquid). The response of the off-gas system (consisting of the headspace, centrifuge, incinerator, and cooler) is recorded every ten seconds (oxygen fraction in the inlet and outlet gases). Oxygen transfer between the liquid and gas phase is assumed negligible.

Pilot-scale cultivation for vaccine production

Two types of bacteria were grown in the pilot-scale experiments: *Neisseria meningitidis* and *Bordetella pertussis*. The oxygen and carbon dioxide fractions in the inlet and outlet gases were measured on a mass spectrometer (prima White Box 600, Thermo Election, UK). The differences in inlet and outlet gas flows were negligible, because the gas production (CO_2) and consumption (O_2) rates were approximately equal (respiration quotient is 1): $F_G \approx F_G^{in} \approx F_G^{out}$. A pressure test was performed in advance of the cultivation. Biomass was measured offline by measuring optical density (OD) at 590 nm of 1 ml suspension using a Vitalab 10 (Vital Scientific, the Netherlands). Nuclear magnetic resonance NMR (OXFORD NMR AS400, BoveBid, Amsterdam) was applied to measure substrates and metabolites.

The *batch* cultivations with *N. meningitidis* were performed in Baart et al. (2007) and are now used for bioreactor monitoring. For detailed materials and methods we refer to Baart et al., 2007. In short, the bacteria were grown in aerobic conditions on a chemically defined

medium in a 60L stainless steel bioreactor containing 40L of medium. Dissolved oxygen was controlled by increasing stirrer speed, the difference between inlet and outlet gases by adjusting the gas flow rate via the sparger (1-20 $l.min^{-1}$).

The *fed-batch* cultivations with *B. pertussis* were done to test the monitoring and control system on pilot-scale. The bacteria were grown in aerobic conditions on a chemically defined medium in a 60L stainless steel bioreactor containing 30L of medium at the start. After the batch phase, the fed-batch started up to a total additional volume of 5L distributed over the full fed-batch period to obtain higher biomass concentrations. The feed contained highly concentrated substrates (glutamate and L-lactate). The specific growth rate was controlled by an adaptive control law (Soons et al., 2006, 2007a, and appendix A). The tuning parameters of the specific growth rate controller γ_1 and γ_2 were adjusted until the performance measures were satisfactory using automatic tuning. Dissolved oxygen was controlled by increasing the stirrer speed. The bioreactor was aerated using headspace aeration (5 $l.min^{-1}$).

Online reconstruction of the inputs and biomass growth

This section consists of three parts:

- The online reconstruction of the inputs using the outputs is given for the generic system in the section “theory”.
- In our system, it was not possible to measure the oxygen uptake rate directly, e.g. using a sensor. Consequently it is not possible to verify whether the calculated auxiliary oxygen uptake rate fits the real values. The model of the off-gas dynamics that is needed for the reconstruction of *OUR*, on the other hand, can be verified, because the oxygen fraction can be measured accurately. Indirectly, the accuracy of the off-gas model gives an indication of the accuracy of the calculation of the auxiliary oxygen uptake rate. This model and its validation are given in the section “model off-gas dynamics”.
- The application of the theory to the off-gas system of the bioreactor and online reconstruction of biomass growth incorporating the tubing delay are addressed in the section “implementation”.

Theory

In the *OUR* calculations we are interested in the oxygen fraction at the input of the system, while the measurements take place at the output. Therefore, the aim is to reconstruct the input *u* from output *y* (Fig. 3). Suppose a system containing a series of continuous stirred tank reactors. The series of CSTRs are modelled by linear first order systems given by the system matrices *A* and *B*. The model and system matrices are given in the next section "Model for off-gas dynamics". Eq. 3 gives the evolution of the states *x* along the discrete time axis for the generic system:

$$x(t+T) = x(t)e^{AT} + \int_{\vartheta=t}^{\vartheta=t+T} e^{A(\vartheta-t)}Bu(\vartheta)d\vartheta \qquad 3$$

where *y* are the outputs, *t* the time, and *T* the sampling time. The input *u* is assumed piecewise constant. The average input during *t* until *t+T* ($\bar{u}(t)$) is reconstructed by rewriting Eq. 3:

$$\bar{u}(t) = \left(A^{-1}\left(e^{AT}-1\right)B\right)^{-1}\left(x(t+T)-x(t)e^{AT}\right) \qquad 4$$

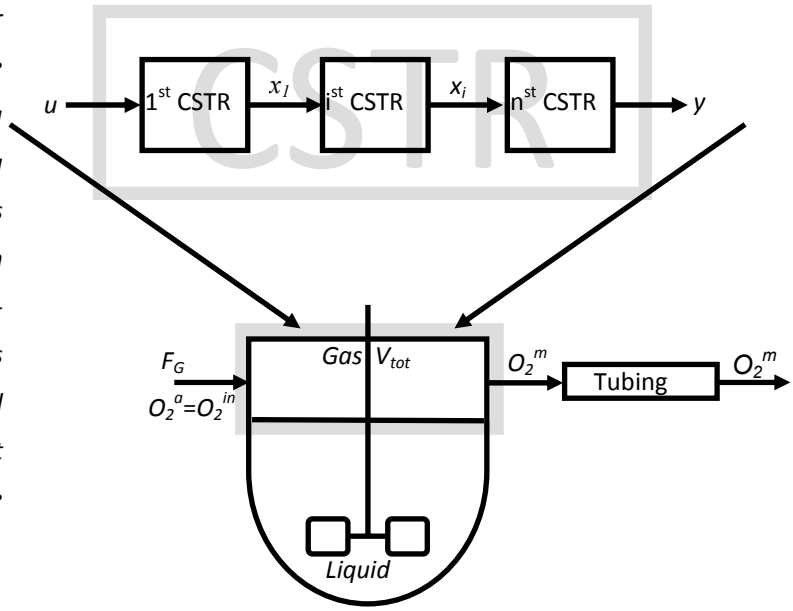


Figure 3. *Schematic of the bioreactor system for modelling the online reconstruction of the input u using output y through a system containing n continuous stirred tank reactors (black) and through a reduced system containing one CSTR (gray). The gas-phase in the reduced model concerns the total gas volume of all compartments V_{tot}. The inlet gas is let into the headspace to mimic the auxiliary oxygen concentration.*

Model for off-gas dynamics

The dynamics of the off-gas system are modelled by a series of *n* ideally stirred tanks plus a time delay (back mixing from the headspace to the cultivation broth is assumed negligible) as in Bloemen et al., 2003:

$$\frac{dO_2^i}{dt} = \frac{F_G}{V^i} \cdot (O_2^{i-1} - O_2^i), i = 1...n \qquad 5$$
$$O_2^m(t) = O_2^n(t - T_D)$$

Where O_2^i is the oxygen fraction in the ith ideally stirred tank (Fig. 2), O_2^0 is the auxiliary oxygen fraction entering the headspace O_2^a, F_G the gas flow rate, and V^i the volume of the gas in the ith ideally stirred tank, which is retrieved from the geometrical properties of the equipment. T_D is the transport delay:

$$T_d = \frac{V_{tubing}}{F_G} \qquad 6$$

In contrast to Bloemen et al., 2003 the dynamics are determined on the basis of physical volume measurements of the stirred tanks (centrifuge, incinerator, and cooler) and tubing instead of identification.

A drawback of Eq. 5 concerning the online calculation of *OUR* is that the signals O_2^i must be differentiated *n* times for the computation of the actual *OUR* giving more measurement noise. To avoid propagation of the measurement noise a reduced model is proposed in which the gas-phase is considered as one ideally stirred reactor (Fig. 3):

$$V_{tot} = \sum_{i=1}^{n} V^i \qquad 7a$$

$$\frac{dO_2^{mix}}{dt} = \frac{F_G}{V^{tot}} \cdot (O_2^a - O_2^{mix}) \qquad 7b$$

$$O_2^m(t) = O_2^{mix}(t - T_D) \qquad 7c$$

where V_{tot} is the total volume of the gas phase consisting of the headspace, centrifuge, incinerator, and cooler. O_2^{mix} is the oxygen concentration after mixing in the total gas phase.

The full model (Eqs. 5-6) and the reduced model (Eqs. 6-7) are verified by calculating the response of the measurement system to varying O_2^a, F_G, and V_h and by comparing the model response with experimental data. In the experiments with water, the inlet gas is let into the

headspace to mimic the auxiliary oxygen concentration (Fig. 3). Mass transfer between the liquid and gas phase is neglected.

The oxygen concentration into the headspace (O_2^a) is varied between 0-21% oxygen and the switching is chosen such that the system is allowed to reach steady state at most instances. Figure 4 shows an example of the response of the model and the system for a headspace of 30L. The mean absolute error was $1.1.10^{-3}$ for the full model and $1.3.10^{-3}$ for the reduced model. Both the full model (Eqs. 5-6) and the reduced model (Eq. 6-7) accurately fit the experimental data. The reduced model is therefore used in the sequel.

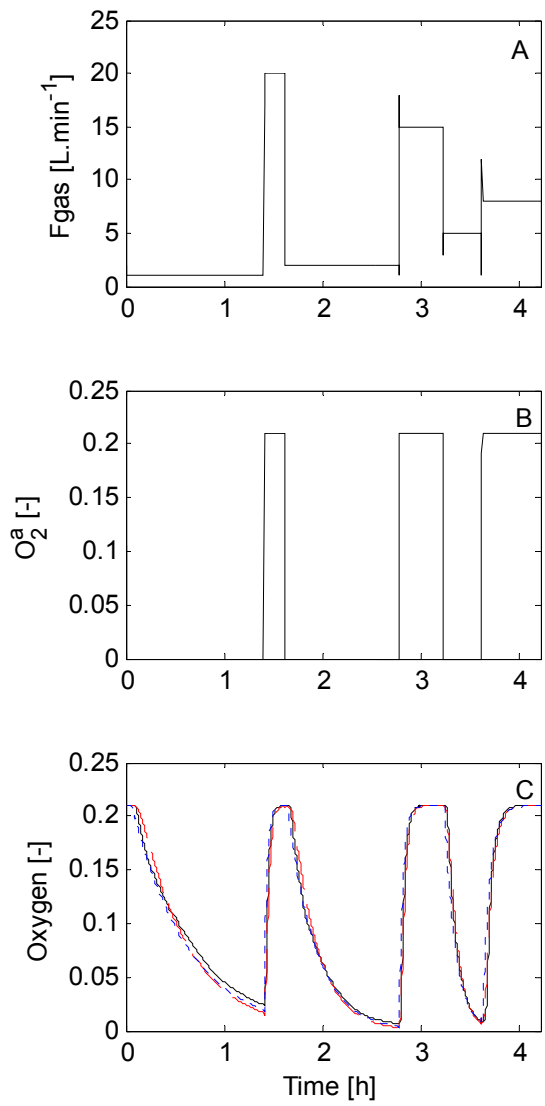


Figure 4. *Dynamics of the gas-phase for a 30L headspace for experimental data.* ***A.*** *Gas flow rate.* ***B.*** *Auxiliary oxygen fraction.* ***C.*** *Measured oxygen fraction (-),oxygen fraction of the full model (four CSTRs plus delay) (red---), and oxygen fraction of the reduced model (one CSTR plus delay) (blue···).*

Implementation

Since the measurements of the oxygen fraction in the outlet stream are delayed due to the tubing and mixing (Figs. 2-3), the following notation is used in the sequel: the subscript $t\text{-}T_D|t$ denotes the value at $t\text{-}T_D$ given all information at t. The reconstruction of the input $O^a_{2,t-T_D|t}$ (circle in Fig. 5) is part of the online reconstruction of biomass growth consisting in four steps (Fig. 5).

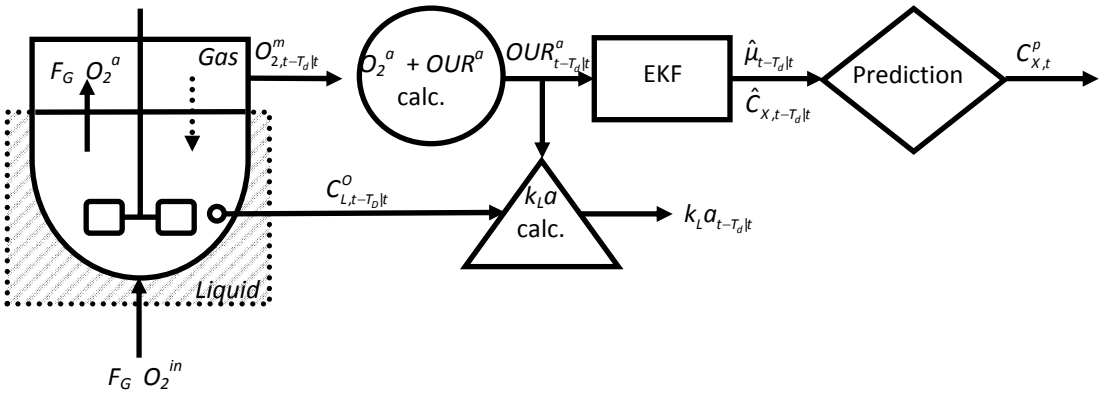


Figure 5. *Schematic of the bioreactor system for online reconstruction of oxygen consumption and growth. The gas-phase includes the gas-phase in the bioreactor and all other devices (centrifuge, incinerator, etc.).*

- *Calculation of* $O^a_{2,t-T_d|t}$ and $OUR^a_{t-T_d|t}$. The oxygen uptake rate that does take into account mixing effects is obtained by replacing O_2^m in Eq. 1 by $O^a_{2,t-T_d|t}$:

$$OUR^a_{t-T_D|t} \approx \frac{F_G}{V} \cdot \left(O^{in}_{2,t-T_D|t} - O^a_{2,t-T_D|t} \right) \cdot \frac{p}{H} \qquad 8$$

The auxiliary oxygen fraction at $t\text{-}T_D$ given all information at t $O^a_{2,t-T_d|t}$ is reconstructed from online $O^m_{2,t-T_d|t}$ measurements using Eq. 4. To reduce the noise propagation, the reduced model is applied (Eqs. 6-7), in which the gas-phase is considered as one ideally mixed tank reactor. With $x = O^m_{2,t-T_D|t}$, $u = O^a_{2,t-T_D|t}$, $A = -F_{G,t-T_D|t} \cdot V_{tot,t-T_D|t}^{-1}$, and $B = F_{G,t-T_D|t} \cdot V_{tot,t-T_D|t}^{-1}$ the auxiliary oxygen fraction becomes:

$$O^a_{2,t-T_D|t} = \frac{O^m_{2,t-T_D|t} - O^m_{2,t-T_D|t-T} \cdot e^{-\frac{F_{G,t-T_D|t}}{V_{tot,t-T_D|t}}T}}{1 - e^{-\frac{F_{G,t-T_D|t}}{V_{tot,t-T_D|t}}T}} \qquad 9$$

The differences in the measured oxygen signals (O_2^m) during one time instant T in Eq. 9 are small and therefore noise sensitive. To further reduce the effect of the noise,

smoothed signals are calculated using Savitsky-Golay smoothing Gorry (1990) for a nine point window and a first order polynomial.

- *Observer for specific growth rate and biomass.* An Extended Kalman Filter (EKF) observes specific growth rate ($\hat{\mu}_{t-T_d|t}$) and biomass ($\hat{C}_{X,t-T_d|t}$) using the auxiliary oxygen uptake rate ($OUR^a_{t-T_d|t}$) every minute (Soons et al., 2008).
- *Biomass prediction.* The tubing delay is cancelled by prediction the current biomass concentration at time *t* ($C^p_{X,t|t}$) given all information at time *t* based on exponential growth during the delay time T_D (assuming μ constant during T_D):

$$\begin{aligned} \mu^p_{t|t} &= \hat{\mu}_{t-T_D|t} \\ C^p_{X,t|t} &= \hat{C}_{X,t-T_D|t} \cdot e^{\left(\mu^p_{t|t} - \frac{F_{S,t-T_D|t}}{V,t-T_D|t}\right)T_D} \end{aligned} \qquad 10$$

- *Computation of oxygen transfer coefficient.* The oxygen transfer coefficient k_La is calculated using the oxygen uptake rate and the measured dissolved oxygen concentration:

$$k_La_{t-T_D|t} = \frac{OUR^a_{t-T_D|t}}{O^a_{2,t-T_D|t} \cdot p/H - C^L_{O,t-T_D|t}} \qquad 11$$

Note that the oxygen transfer rate is assumed to be equal to the oxygen consumption rate, because dissolved oxygen is controlled and oxygen dynamics are much faster than other relevant processes. To reduce the noise, the signal is again smoothed using Savitsky-Golay smoothing (Gorry, 1990) for a nine point window and a first order polynomial.

Results and discussion

The oxygen uptake rate is calculated using the steady-state method (Eq. 2) and using online reconstruction (Eqs. 8-9). Each of these signals were used as input for the Extended Kalman Filter (Soons et al., 2008) to observe biomass and specific growth rate and finally to predict biomass concentrations (Eq. 10) to investigate the effect of the off-gas dynamics during real batch and fed-batch experiments. The auto-tuning controller for specific growth rate, developed and validated for small-scale cultivations in (Soons et al., 2006, 2007a), was applied to the pilot-scale cultivation with *B. pertussis* without any modification.

Eight experiments for *N. meningitides* were evaluated. Figure 6 shows the first five hours of the typical batch cultivation with *N. meningitidis* with sparger aeration. The errors in the calculation of the oxygen uptake rate with the dynamic model (Eqs. 8-9) were up to 30%

smaller than the values calculated from the steady-state method (Eq. 2). As a result the observed biomass concentrations were closer to the samples when the dynamic *OUR* calculation was used (Eqs. 8-10).

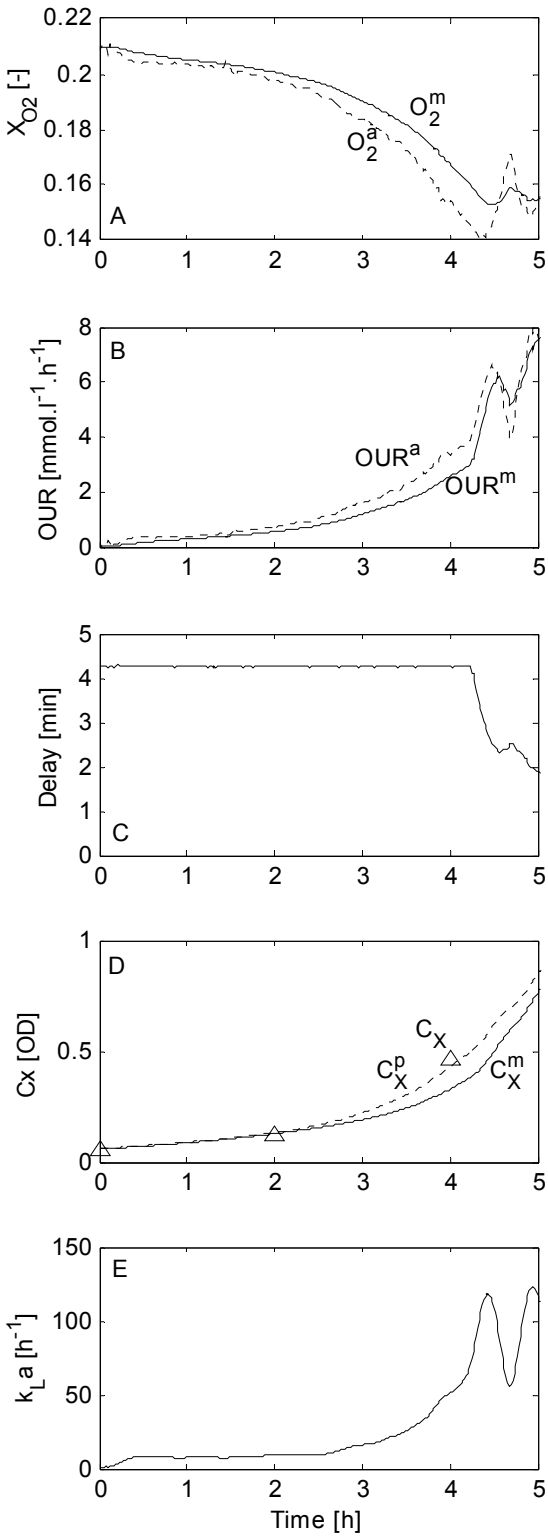


Figure 6. *Experimental results for N. meningitidis.* ***A.*** *Measured (O_2^m) and auxiliary oxygen fraction O_2^a.* ***B.*** *Steady state method (OUR^m) and auxiliary oxygen uptake rate (OUR^a).* ***C.*** *Tubing delay.* ***D.*** *Offline biomass measurements (ΔC_X), biomass observed from OUR^m (C_X^m -), and biomass observed and predicted from OUR^a (C_X^p--).* ***E.*** *Oxygen transfer coefficient.*

The effect of the off-gas dynamics was significant in the first part of the batch cultivation and decreased with time due to the increasing gas flows and thereby decreasing gas mixing times and transport delays.

Figure 6E shows the typical time-varying course of the oxygen transfer coefficient k_La. In the first part of the cultivation, the cells grew exponentially and so did the demand for oxygen. To cope with this demand, the stirrer speed and gas flow rate increased and as a consequence k_La increased.

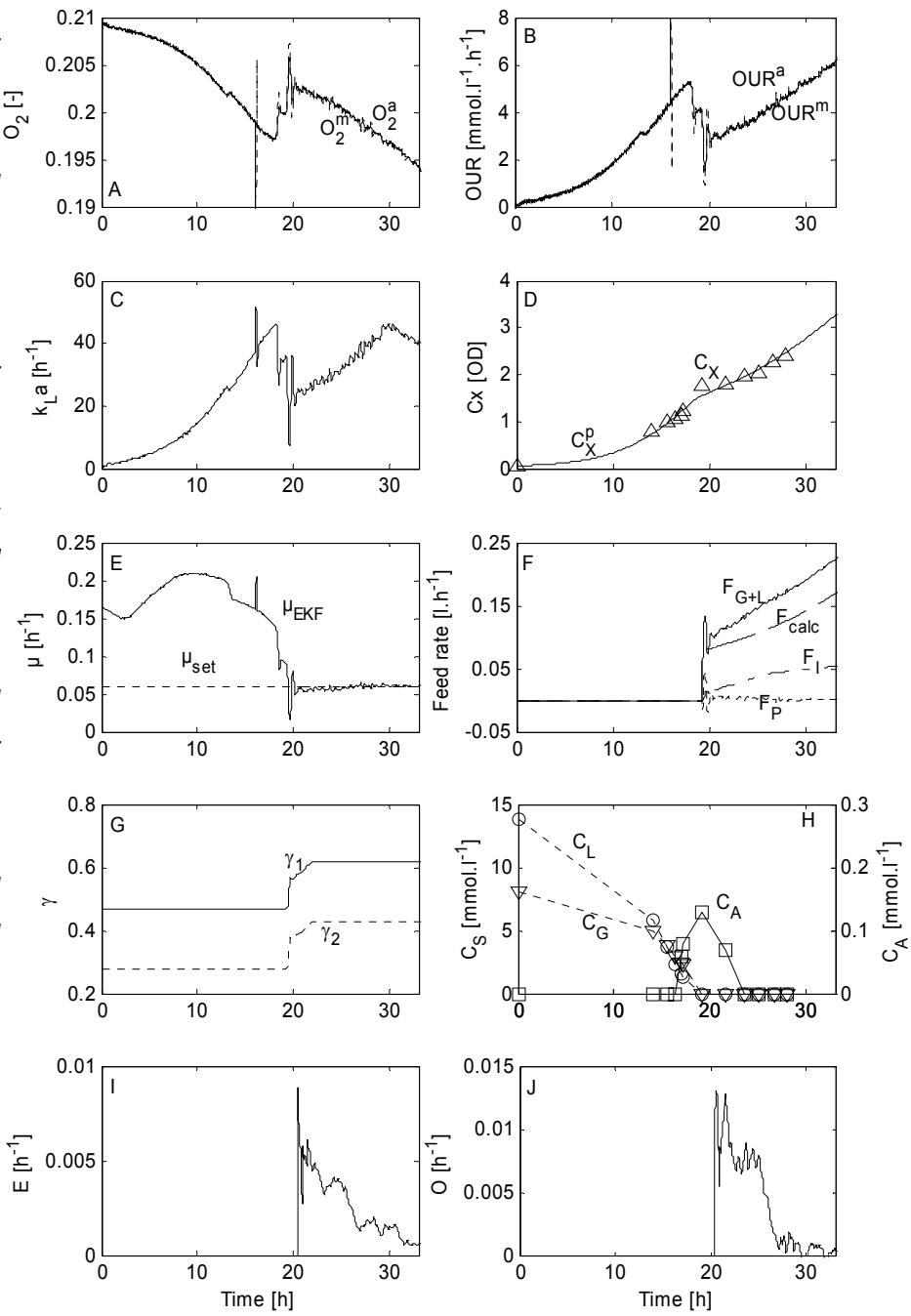


Figure 7. *Experimental data for B. pertussis with adaptive control of the specific growth rate.* ***A.*** *Measured (O_2^m) and auxiliary (O_2^a) oxygen fraction* ***B.*** *Steady state method (OUR^m) and auxiliary (OUR^a) oxygen uptake rate.* ***C.*** *Oxygen transfer coefficient.* ***D.*** *Offline (ΔC_X) and predicted biomass concentration (- C_X^p)* ***E.*** *Specific growth rate* ***F.*** *Substrate feed rate. Glutamate and lactate (F_{G+L}), prior calculation (F_{calc}), proportional action (F_P), and integral action (F_I)* ***G.*** *Controller tuning parameters* ***H.*** *Substrate concentrations: glutamate (C_G left axis) and lactate (C_L left axis); and acetoacetate (C_A right axis).* ***I.*** *Mean absolute error* ***J.*** *Oscillation measure.*

Two *B. pertussis* cultivations were performed using headspace aeration. The oxygen transfer coefficient was smaller for headspace aeration compared with sparger aeration of the *N. meningitidis* cultivations (compare Fig. 6E with 7C), resulting in smaller differences in inlet and outlet gas concentrations and more noise on the *OUR*.

The gas flow rate was kept constant (5 $l.min^{-1}$) and consequently the mixing effects and tubing delays were approximately constant and small (50 seconds) compared to the initial phase of the *N. meningitidis* cultivation (4.5 minutes). As a consequence, the mixing effects and delays in the gas phase are negligible in the experiment with *B. pertussis* due to the higher gas flows applied for headspace aeration.

Figure 7 shows the results for the fed-batch cultivation with *B. pertussis*. First, the cells were grown in batch cultivation until the substrates almost depleted. At t=18.5h the enhancing substrate, lactate, was depleted resulting in a lower specific growth rate; at 19.5h the main substrate, glutamate, was depleted and the specific growth rate dropped.

At this point, the feed with limiting substrates and the auto-tuning controller started automatically to control the specific growth rate at the set-point of $0.06h^{-1}$. At the start of the fed-batch, a few minutes were needed to fill the tubes with the feed and to actually start the substrates entering the bioreactor. After that, as a result of the controller action an exponentially increasing feed was added into the bioreactor to cope with the exponentially increasing biomass. The controller maintained the specific growth rate close to the set-point.

After some hours of fed-batch, the mean absolute error and the oscillation measure (see appendix A for the definition) decreased and eventually became approximately constant indicating that controller performance was satisfactory and that further auto-tuning would have no effect. The performance monitor qualified these measures to be small and decided to continue the cultivation with the current settings for γ_1 and γ_2 and to deactivate the automatic tuning. This means that the correct tuning has been “learned” after one run. Subsequently, the following cultivations can be performed with the obtained settings.

The observed biomass concentrations coincide well with the offline measurements; and also the specific growth rate was accurately observed throughout the whole fed-batch cultivation.

Figure 7H (left axis) shows that the substrates glutamate and lactate decreased during the batch cultivation and were kept constant at very low concentrations during the fed-batch part of the cultivation. *B. pertussis* is able to form and excrete acetoacetate in certain conditions, e.g. when glutamate and lactate are depleted, thereby reducing the biomass

yield (Thalen et al., 1999). The undesired component acetoacetate was only formed in small amounts at the start of fed-batch (Fig. 7H right axis, note the small values on the scale), when neither glutamate nor lactate were available in the medium. At t=23h, glutamate and lactate were again present; the acetoacetate was consumed and remained zero in the following hours of the fed-batch.

Figure 7C shows the time-varying course of the oxygen transfer coefficient k_La for *B. pertussis*. The stirrer speed, and so k_La, is increased throughout the cultivation to cope with the increasing oxygen demand. At t=29h the stirrer speed reached its maximum value and remained constant until the end of the cultivation and the dissolved oxygen controller was not capable anymore to track the set-point. The oxygen transfer coefficient, however, decreased due to blanketing effects (caused by e.g. viscosity, formation of proteins, and lysis).

The applications show that the observer and controller are highly applicable: from small-scale to pilot-scale cultivations; from batch to fed-batch cultivations, from low biomass concentrations to high biomass concentrations, from *OUR* calculated using the off- gas composition to *OUR* calculated using k_La, dissolved oxygen and inlet concentrations; and from cultivation with *B. pertussis* to *N. meningitidis*.

Conclusions

In this work two aspects concerning the scale-up from laboratory to pilot-scale cultivation have been considered and successfully implemented:

- The first is an improved computation of oxygen uptake rate taking into account gas-phase dynamics that is important for large-scale equipment. If the volume of the gas phase is large compared to the gas flows, mixing effects and tubing delays tend to filter out the dynamics of the oxygen uptake rate. Incorporating the gas phase dynamics in online monitoring reduces the error observations of the oxygen consumption and biomass growth up to 30%. Using models which contain system-relevant parameters enhances the applicability of the method.
- The second aspect is the monitoring and control of the biomass growth. The M&C system –designed for small-scale cultivations- accurately monitored and controlled the specific growth rate and biomass for pilot-scale cultivations in real-time. It appeared that the M&C system could be applied directly to the larger-scale cultivation without modifications like redesigning or re-tuning the observer or

controller. So, this strongly suggests that the M&C system is robust with respect to scale and micro-organism.

On large-scale, vaccines are currently produced in batch cultivation. In traditional batch cultivation with *B. pertussis*, the cells grow until the substrates are depleted up to approximately 1.6 OD. The methods developed in this paper show that fed-batch cultivation has potential for production on (near) production-scale. In the fed-batch phase of the cultivation the biomass more than doubles up to 3.5 OD by adding a feed with limiting substrates to control the specific growth rate. Higher biomass concentrations can be obtained, like in small-scale cultivations (Soons et al., 2006), if the headspace aeration were replaced by sparger aeration to obtain proper oxygen transfer.

So, the combination of enhancing an existing process with controlled fed-batch cultivation and the achievement of at least doubled concentration of bulk vaccine on pilot-scale yields a favourable upstream production method for bulk whooping cough vaccine.

Acknowledgments

Wageningen University and the Netherlands Vaccine Institute (NVI) work together in a project with Applikon Biotechnology BV and Siemens NV in order to improve the production process by release of biopharmaceuticals on the basis of new online techniques. The Dutch Ministry of Economical Affairs (TSGE3067) funds the project. Laboratory experiments were performed at NVI, The Netherlands. We like to thank Martin Hamzink for performing the NMR measurements and Gino Baart for providing the experimental data with *N. meningitides*.

Appendix A: Monitoring and control

The monitoring and control system consists of an observer, controller for specific growth rate, auto-tuning of the controller, and a performance monitor.

- *Observer.* An Extended Kalman Filter (EKF) observes specific growth rate (μ^m) and biomass (C_X^m) using the oxygen uptake rate (*OUR*) every minute (Soons et al., 2006). The EKF is based on the following model for biomass growth and oxygen consumption:

$$\frac{dC_X}{dt} = \mu C_X$$
$$OUR = \left(\frac{\mu}{Y_O} + m_O\right) C_X \qquad \text{A1}$$

- *Specific growth rate control.* The control purpose is to regulate the specific growth rate (μ) to a desired value by adding a feed with limiting substrates using the controller of (Soons et al., 2006):

$$F_{G+L} = F_{calc} + F_P + F_I$$
$$F_{calc} = \frac{ac + bd}{aC_G^{in} + bC_L^{in}} C_X^p V$$
$$F_P = K_P \cdot \varepsilon \qquad \text{A2a}$$
$$F_I = \int_0^t K_I \cdot \varepsilon \, dt$$

where the error ε is the set point minus the observed specific growth rate (μ_{set}-μ_p) and where the controller gains K_P and K_I are adjusted online to the changing volume:

$$K_P = \frac{V}{\gamma_1 (aC_G^{in} + bC_L^{in})}, K_I = \frac{V}{\gamma_2 (aC_G^{in} + bC_L^{in})} \qquad \text{A2b}$$

γ_1 and γ_2 are the tuning parameters for the controller and *a*, *b*, *c*, and *d* are constants depending on the micro-organism (in the work *Bordetella pertussis*) and the set-point for specific growth rate.

- *Automatic tuning* was performed as in (Soons et al., 2007a):

$$\frac{d\gamma_1}{dt} = \beta_1 \cdot \frac{\varepsilon}{\mu_{set}} + \beta_2 \cdot \frac{\varepsilon^2}{\mu_{set}^2}$$
$$\frac{d\gamma_2}{dt} = \beta_1 \cdot \frac{\varepsilon}{\mu_{set}} + \beta_2 \cdot \frac{\varepsilon^2}{\mu_{set}^2} \qquad \text{A3}$$

Where the tuning parameters γ_1 and γ_2 are adapted proportional to the error ε and proportional to the quadratic error ε^2, which is large if the error is large. β is the adaptation rate, the tuning parameters for auto-tuning.

- *Performance monitor.* The performance monitor has two tasks in the system. The first function of the performance monitor is to qualify controller performance using the criteria given in Eqs. A4-A5 (Soons et al., 2007a). Second, based on the qualification, the performance monitor decides to retune the process using automatic tuning or to stop the auto-tuning activity and continue with the current

settings. The first measure is the mean absolute error over the past *N* points of a moving window at time point *k*:

$$E(k) = \frac{1}{N} \cdot \sum_{k-N}^{k} \left| \varepsilon(k) \right| \qquad \text{A4}$$

The second is the oscillation measure and is defined by

$$O(k) = \frac{1}{N} \cdot \left(\sum_{k-N}^{k} \left| \varepsilon(k) \right| - \sum_{k-N}^{k} \varepsilon(k) \right) \qquad \text{A5}$$

Appendix B: parameter values

Table 1. *parameter values*

Parameter		Value
Volume	headspace	20-30 *l*
	centrifuge	1.34 *l*
	tubing	2.0 *l*
	incinerator	2.15 *l*
	cooler	0.66 *l*
	cultivation	30-40 *l*
	broth	
B. pertussis	Y_O	0.033 $OD.mmol^{-1}$
	m_O	0.69 $mmol.OD^{-1}.h^{-1}$
N. meningitidis	Y_O	0.077 $OD.mmol^{-1}$
	m_O	0 $mmol.OD^{-1}.h^{-1}$
C_L^{in}		835 $mmol.l^{-1}$
C_G^{in}		500 $mmol.l^{-1}$
$\tau_{analyzer}$		13 *s*
μ_{set}		0.06 h^{-1}
β_1		0.25
β_2		0.50
N		90

CHAPTER 7

Modelling the cultivation of Bordetella pertussis: the effect of specific growth rate on expression of virulence factors on gene and protein level

Submitted to Journal of Biotechnology as: Zita I. T. A. Soons Nicole S. D. Larmonie, Bas van de Waterbeemd, Mathieu Streefland, Dirk E. Martens, Gerrit van Straten, and Anton J. B. van Boxtel

Abstract

The quality of whole cell whooping cough vaccine is based on the presence of outer-membrane proteins that are important for inducing a protective immune response. This work investigates the effect of specific growth rate on gene and protein expression of *Bordetella pertussis*. While modelling of the synthesis of complex products and processes is uncommon and difficult, an attempt was made to model the formation of virulence factors by enhancing an existing model for the cultivation of *B. pertussis* for the production of whole-cell vaccine against whooping. Four chemostat cultivations were performed to model the biomass growth and formation of virulence factors. Although this study is based on a limited number of measurements, it indicates the potential of using models for improved vaccine manufacturing processes. The data and model suggest that vaccine quality is not deteriorating with increasing growth rate over the investigated range, and thus a high growth rate can be selected to achieve a high productivity.

Keywords

Model, specific growth rate, *Bordetella pertussis*, virulence factors, gene and protein expression, chemostat cultivation.

Nomenclature

a_{FHA}, a_{PT}	non-growth associated specific production rate of FHA and PT [$\mu g.OD^{-1}.h^{-1}$]
a_{LPS}	non-growth associated specific production rate of LPS [$mmol.OD^{-1}.h^{-1}$]
b_{FHA}, b_{PT}	growth associated production of FHA and PT [$\mu g.OD^{-1}$]
b_{LPS}	growth associated production of LPS [$nmol.OD^{-1}$]
Bvg	Bordetella virulence genes
c_{FHA}	filamentous haemagglutinin concentration [$mg.l^{-1}$]
C_G, C_G^{in}	glutamate concentration in the medium, respectively in the feed [$mmol.l^{-1}$]
C_L, C_L^{in}	lactate concentration in the medium, respectively in the feed [$mmol.l^{-1}$]
c_{LPS}	lipopolysaccharide concentration [$\mu mol.l^{-1}$]
c_{PT}	pertussis toxin concentration [$mg.l^{-1}$]
C_X	biomass concentration [*OD*]
D	dilution rate [$l.h^{-1}$]
E	performance criterion
FHA	filamentous haemagglutinin
F_S	substrate feed rate [$l.h^{-1}$]
GE	gene Expression
I	emission intensity
K_G, K_L	Monod constant on glutamate, respectively lactate [$mmol.l^{-1}$]
LPS	lipopolysaccharide
MB	membrane bound
m_G, m_L	maintenance coefficient on glutamate and lactate [$mmol.OD^{-1}.h^{-1}$]
PT	pertussis toxin
q_{FHA}, q_{PT},	specific FHA, PT, and LPS production rate [$\mu g.OD^{-1}.h^{-1}$]
q_{LPS}	specific FHA, PT, and LPS production rate [$nmol.OD^{-1}.h^{-1}$]
q_G, q_L	specific glutamate and lactate consumption rate [$mmol.OD^{-1}.h^{-1}$]
SUP	supernatant
t	cultivation time [*h*]
V	liquid volume [*l*]
vag, vrg	virulence activated gene, virulence repressed gene
VR	virtual Reference
WCS	whole-cell suspension
Y_{G1}, Y_{G2}	biomass yield on glutamate over pathway 1 and pathway 2 [$OD.mmol^{-1}$]
Y_L	biomass yield on lactate [$OD.mmol^{-1}$]

Greek letters

μ	specific growth rate [h^{-1}]
μ_{max}, μ_{enh}	maximum, respectively enhanced specific growth rate [h^{-1}]

Superscripts and subscripts

m	measured values

Introduction

Prediction and optimization of vaccine production processes have not reached the state of development that can be found in other industries. One reason is the accuracy and reliability by which the states of the complex cultivation processes can be monitored online. To stimulate improved (bio)pharmaceutical manufacturing, the US Food and Drug Administration (FDA) launched the PAT initiative (FDA, 2004). It requires a science-based approach to understand the process and its critical variables that have to be monitored and controlled online.

Currently vaccines are released on the basis of offline tests on the final product. In most cases animal models are required for potency assays. Such assays are inherently inaccurate and do not comply with possibility for feedback during the process. A model to predict the end-quality of the vaccine in real-time could be the first step towards better understanding, monitoring, and control of the cultivation. This work aims to understand the expected vaccine quality in response to the specific growth rate during the cultivation step for the production of a vaccine against whooping cough disease, keeping in mind that specific growth rate is accessible for control (Soons et al., 2006).

Bordetella pertussis is a pathogenic bacterium that causes the disease whooping cough. Whooping cough is highly contagious and can be dangerous for children. Consequently, an effective and safe vaccine is needed against this disease. An effective vaccine is based on virulence factors, which are essential for the survival and pathogenicity of *B. pertussis.* The virulence factors have different roles depending on the stage of infection, from entry of the pathogen into the host and attachment to the target tissues to dissemination of the pathogen or its products (Weiss and Hewlett *et* al., 1986; Kerr and Matthews, 2000).

The virulence factors of *B. pertussis* are among others filamentous haemagglutinin (FHA) and pertussis toxin (PT). FHA is involved in the attachment of the cells and enables colonisation of the respiratory tract. PT is unique for *B. pertussis* and is the main toxin as it is involved in

all stages of infection. Another component, lipopolysaccharide (LPS) is involved in the last two stages of the infection in causing local damage at the site of infection and in dissemination.

The virulence factors are important aspects for the quality of *B. pertussis* suspensions used for vaccine production against infection with whooping cough (Westdijk et al., 1997; Rodriguez et al., 1994). Westdijk et al. (1997) demonstrated that during batch cultivation the growth phase determines the antigen production and release. No conclusion, however, is given about the relation between virulence factor production and specific growth rate or nutrient composition. Rodriguez et al. (1994) and Licari et al. (1991) show that the production of pertussis toxin and LPS is strongly growth-associated and that a high specific growth rate is an effective way for producing PT. High amounts of LPS, however, are regarded as a drawback, because LPS is suspect of causing adverse reactions. So, in order to obtain a high quality-vaccine, it is important to restrict LPS content of the vaccine, while still providing efficient other virulence factors.

The two-component system Bordetella virulence genes BvgA/S regulates the expression of virulence factors of *B. pertussis* (Scarlato et al., 1990). BvgS is a sensor that triggers BvgA to activate the virulence activated genes (vags) or deactivate virulence repressed genes (vrgs). In order to obtain a high-quality vaccine, this BvgA/S system needs to be in the *bvg+* mode, in which the bacteria express virulence factors. DNA micro-arrays can be used to monitor the state of the BvgA/S system and the activation and repression of genes and thus to monitor the quality of the final vaccine (Streefland et al., 2007). Certain laboratory conditions - such as low temperature or the presence of $MgSO_4$ or nicotinic acid in the growth medium - modulate the expression of *B. pertussis* vags and vrgs (Lacey, 1960, Hot et al., 2003, Streefland et al. 2007). Nakamura et al. (2006) showed that the growth phase determines the expressions of virulence factors. Out of 81 virulence factors, 51% had significantly lower transcript abundance in stationary phase than in log phase and 2% significantly higher abundance. These authors suggest that nutrient limitation is one of the several interacting factors affecting virulence. Although end-of-batch effects have been investigated, the effect of specific growth rate on gene expression is yet unexplored.

Modelling the production of complex products and processes is uncommon (Junker and Wang, 2006) and is difficult because of the large number of complex reactions in a cell, the nonlinearity, and the time-dependency (Karim et al., 2003). Furthermore, the data in cell biology often contain a high degree of uncertainty. This uncertainty translates into the difficulty of measuring model parameters and hence the impossibility of obtaining a precise model (Sontag, 2007). Nevertheless, models do contain valuable information and can be

used to gain better understanding, monitoring & control and/ or optimization of cultivation strategies.

Jenzsch et al. (2007), Levisauskas et al. (2003), Dhir et al. (2000) and few others (Junker and Wang, 2006) optimised the production of single proteins in real-time. However, if the final product is a heterogeneous mixture, which is the case in cultivation of *B. pertussis* for a whole-cell vaccine for whooping cough, it is more difficult to detect the products and monitor these cultivations (Junker and Wang, 2006).

This work gives an indication of the effect of specific growth rate (μ) on vaccine quality by combining data on protein levels for virulence factors and gene expression on mRNA level. It also uses metabolic information for substrate consumption and biomass formation. In addition a model for growth and antigen formation of *B. pertussis* is proposed. To this end, a series of four substrate-limited continuous-flow stirred-tank reactor (CSTR) cultivations are performed at different dilution rates and hence different specific growth rates. The next sections describe the materials and methods for the laboratory experiments; postulate the model for virulence factor formation; and discuss the results of the experiments and the model. The model for biomass growth and the calculation of its parameters are given in Appendix A.

Materials & methods

Bioreactor conditions

CSTR cultivations with the dual substrate consuming bacterium *Bordetella pertussis* were performed in a 5-liter bioreactor containing 2.87 litres medium with glutamate and L-lactate as the main carbon sources (Thalen et al. (1999). A six-bladed turbine stirrer was used to agitate the medium. Temperature was controlled at 34ºC, agitation speed at 470 rpm, and dissolved oxygen at 30% air saturation by headspace aeration only (by changing the incoming oxygen fraction in an oxygen/nitrogen mixture). The total gas flow was kept constant at 1 l/min. Four dilution rates were applied: 0.03, 0.05, 0.08, and 0.11 h^{-1}. The bacteria washed out in a cultivation at a dilution rate of 0.14 h^{-1}, hypothetically due to limitation of a component in the medium. The CSTR phase was started when the biomass reached an optical density of 1. The cultivation was considered in steady-state after at least five dilutions.

Analysis

Online measurements

A polarographic electrode (Applikon, the Netherlands) was used to measure dissolved oxygen in the medium. A pH electrode (Mettler Toledo, Udorf, Switzerland) was used to measure pH, and temperature was measured with a Pt100 temperature sensor.

Offline measurements

Glutamate and L-lactate were measured offline with a YSI 2750 select analyser (Yellow Springs Instruments, Yellow Springs, USA). *Biomass* was measured offline by measuring optical density (OD) at 590 nm of 1 ml suspension using a Vitalab 10 (Vital Scientific, the Netherlands). Nuclear magnetic resonance NMR (OXFORD NMR AS400, BoveBid, Amsterdam) was applied to detect accumulation of *waste products* or *metabolites*.

Gene expression was analysed for a set of 3582 *B. pertussis* genes by means of two-colour micro-arrays as in Streefland et al. (2007). Preparation and analysis of the micro-arrays was done in four steps:

- First, RNA was isolated and pre-treated to fix the expression profile. The samples were corrected for the amount of biomass per cultivation.
- Second, the two samples –to compare on one array- were labelled with two different fluorescent dyes (Cy3/Cy5), which have emissions at two different wavelengths. Each specific growth rate sample was compared with each other sample in six arrays (0.03-0.05, 003-0.08, 0.03-0.11, 0.05-0.08, 0.05-0.11, and 0.08-0.11 h^{-1}).

The differences in emission intensities of each fluorescent dye, or gene expression, were used in a ration-based analysis to identify up- and down regulated genes, the third and fourth step.

- Third, the expression data were natural-log transformed and quantile normalized to correct for differences between arrays without affecting the measurement of the biological variation. A virtual reference (*VR*) was calculated by taking the average of each gene for all slides and for both dyes together.
- Fourth, the gene expression for each gene was calculated in a semi-quantitative way by comparing the emission intensity *I* of each specific growth rate sample with each

other intensity of the sample on the same arrays. For instance, the expression *GE* for each gene *i* at 0.03 h^{-1} reads:

$$GE_{0.03}^{i} = \frac{1}{4} \cdot \left(I_{0.03}^{0.03-0.08,i} - I_{0.08}^{0.03-0.08,i} + I_{0.03}^{0.03-0.05,i} - I_{0.05}^{0.03-0.05,i} + I_{0.03}^{0.03-0.11,i} - I_{0.11}^{0.03-0.11,i} + VR^{i} \right) \quad 1$$

where the superscript indicates the micro-array and the subscript the specific growth rate sample on that micro-array.

We refer to Streefland et al. (2007) for more details on the preparation of the micro-arrays.

Quantification of cell-associated *proteins* (FHA and PT) was done using ELISA developed by Westdijk et al. (1997). Each sample was first compounded with 0.5% Na-azide and incubated for 10 minutes ate 56°C and next divided in whole-cell suspensions (WCS), membrane bound proteins (MB) and supernatant (SUP) as follows. Half of a 20-ml sample was directly used for the quantification of the virulence factors in the whole cell suspension. The other half was centrifuged at 6500 x g for 10 minutes at 4°C. The resulting supernatant was used for quantification of the cell-free virulence factors, the resulting pellet for the membrane-bound virulence factors. Ideally if measurement errors do not occur, the sum of the MB and SUP virulence factors is identical to the virulence factors in the whole-cell suspensions.

Membrane bound LPS was measured using gas chromatography (Welch, 1991).

Model for production of virulence factors and LPS

Biomass growth on glutamate and L-lactate was modelled earlier (Neeleman et al., 2001). The model is summarized in Appendix A. The steady state conditions in the CSTR experiments allow estimation of parameters that are not easy to obtain in batch. As shown in the appendix, the model yields good fits on observed biomass and substrate concentrations at various specific growth rates, and can thus serve as a basis for modelling of virulence factors. To this end, the model is with an empirical model for formation of the virulence factors (VF) FHA, PT, and of LPS. The mass balances for their production the bioreactor are given by:

$$\frac{dC_{VF}}{dt} = q_{VF} \cdot C_X - D \cdot C_{VF} \quad 2$$

where q_{VF} is the specific virulence factor production rate, which is postulated to be proportional to the specific growth rate according:

$$q_{VF} = a_{VF} \cdot \mu + b_{VF} \quad 3$$

with a_{VF} a constant for growth associated virulence factor formation, and b_{VF} a constant for non-growth associated specific virulence factor production rate.

The measured values for the specific production rates of the virulence factors and its model predictions are shown in Fig. 1. The slopes correspond to the growth associated formation a_{VF}, the intercepts to the non-growth associated parts b_{VF}.

Although it would be possible to obtain a better fit using more fit parameters, the data are not that indicative to do this with confidence. A straight line almost through the origin is in line with the hypothesis that the content of a particular factor per unit cell is invariant with specific growth rate. Moreover, as shown in the next section a linear relationship gives a good prediction of FHA observed in batch at a much higher specific growth rate. The values of the parameters in Eq. 3 obtained from the experiments are given in Table A1

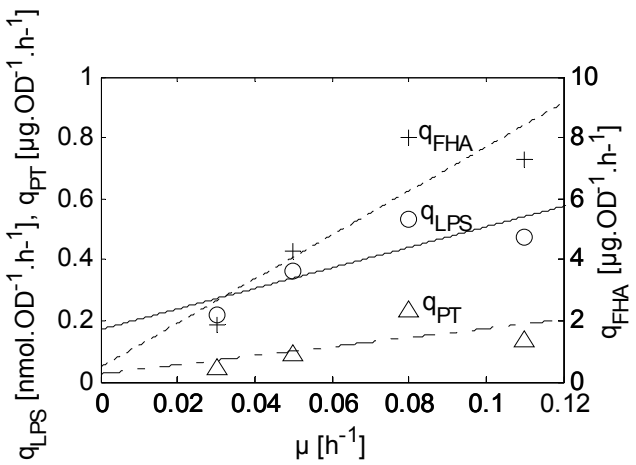


Figure 1. *Specific production rates for the virulence factors. Left axis: q_{LPS} (o) and q_{PT} (Δ). Right axis: q_{FHA} (+).*

Experimental and modelling results

Virulence factors on gene level

Most literature shows a switch in expression of virulence factors between the *bvg*+ state when virulence factors are abundantly expressed and *bvg*- state when virulence factor expression is absent. Cummings et al. (2006) mentioned that gene regulation studies have historically focused on the endpoints of the spectrum. Recently, Cummings et al. (2006) and Stenson and Peppler (2007) suggested that *bvg* regulation is more smooth rather than an on/off switch by identifying at least two intermediate phases between the *bvg*+ and *bvg*- extremes. Cummings et al. (2006) also stated that a higher-resolution examination of the *bvg* regulation spectrum may reveal further distinct phenotypic phases.

This work investigates potential differences in virulence factor expression in response to specific growth rate. The virulence activated genes are taken from Streefland et al., (2007). The average expression of each vag, the virtual reference VR_i, was between 6.4 and 10.9 indicating that the bacteria were in the *bvg*+ phase for all specific growth rates. Figure 2 shows that gene expression of the virulence activated genes was approximately constant throughout the different specific growth rates.

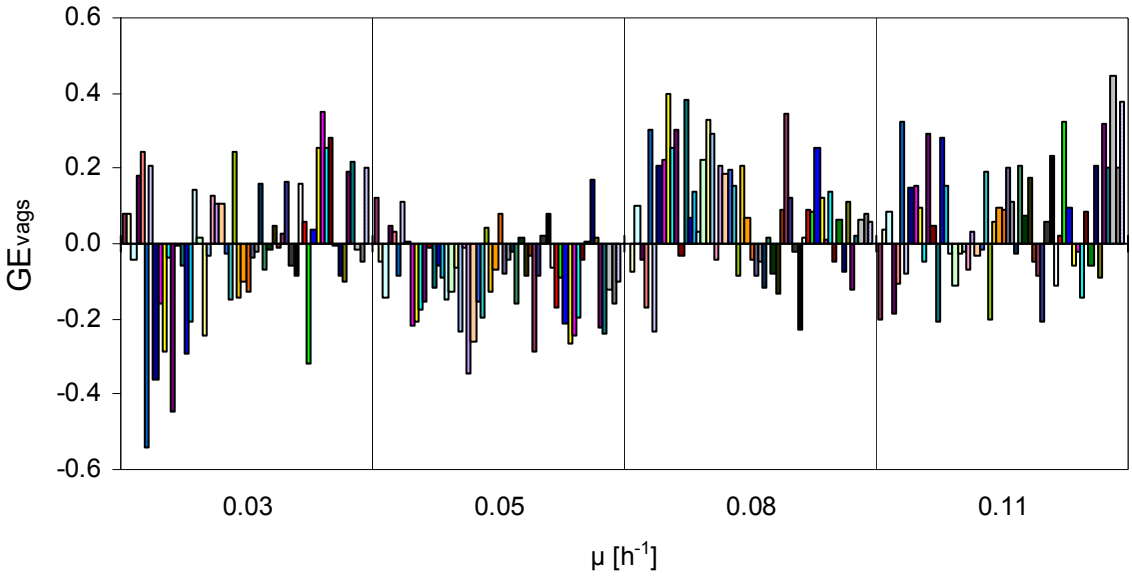


Figure 2. *Effect of specific growth rate on gene expression for the 56 virulence activated genes compared to their virtual reference (VR_i: the average expression of each gene).*

Growth rate dependent genes

In the previous section, only genes associated with virulence of *B. pertussis* were considered. In addition all other genes were examined for growth rate dependency. Out of 3600 genes the gene expression of 159 *B. pertussis* genes turned out to be growth rate dependent. The largest groups are the ribosomal protein genes and genes related to ABC transporters. Ribosomes are involved in the protein biosynthesis by translating mRNA to protein. Naturally, the expression of ribosomal genes increases with the specific growth rate (and so the formation of biomass and proteins). 95% out of 22 genes was positively correlated with the specific growth rate, 5% negatively correlated.

The ABC transporters are transmembrane proteins that function in the transport of a wide variety of compounds across extra- and intracellular membranes, including metabolic

products, lipids, and drugs. 63% Out of 11 genes for the expression of ABC transporters was negatively correlated with the specific growth rate, 27% positively correlated. The reason for a negative correlation may be the fact that the substrate concentrations are lower with lower specific growth rates, which may increase the effort to transport the substrates from the environment into the cells.

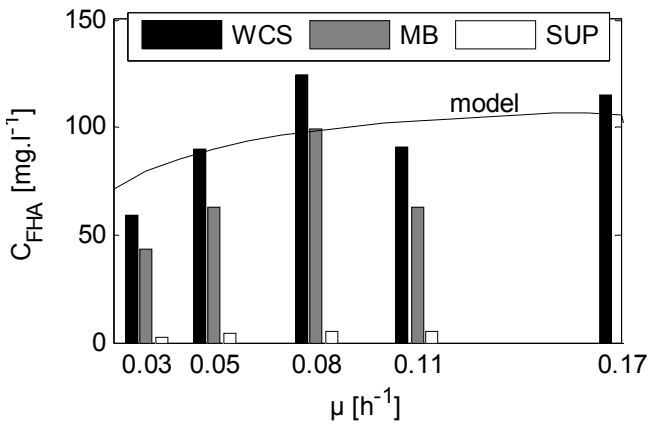
Virulence factors and LPS on product level

High cell surface densities of protective proteins are likely to augment the potency of *pertussis* whole cell vaccine. Therefore, it is important to examine the content of outer membrane components like FHA and PT per cell besides protein activity per volume (Westdijk et al., 1997). The presented model fits are obtained by simulation of the model Eqs. A1-A7 and 1-2. This section focuses on the results of the experiments and the model.

FHA

Jacob-Dubuisson et al. (2000) and van den Berg et al. (1999) state that FHA is loosely associated with the outer membrane and is secreted into the extra cellular environment during growth. Westdijk et al. (1995) measured relatively small amounts of FHA in the supernatant. Figure 3 shows that FHA is predominately membrane bound (MB) in this work, which is in agreement with literature. The production of membrane-bound as well as excreted FHA is growth associated (Figs. 1, 3). The FHA content, which can be observed by dividing the FHA concentration by the biomass concentration, appeared to be almost invariant.

Figure 3. *Experimental results and model fit for the effect of specific growth rate on the steady-state FHA concentration and one batch measurement. WCS is whole-cell suspension, MB membrane bound, and SUP supernatant.*



The predictions for the concentrations of FHA are the result of biomass growth and FHA production proportional to the specific growth rate (Eqs. A1-A7, 1-2). Given the complexity and uncertainty of the protein data, the model gives a fair fit of the FHA concentrations of the whole cell suspensions for the CSTR cultivations.

The model derived from CSTR experiments offers new prospects to predict or even optimize the quality of the production stage for the whole cell vaccine of *B. pertussis*. As an example, it is used to predict FHA during batch cultivations. The specific growth rate was about 0.17 h^{-1} in the main part of the batch. The suspension was harvested before the substrates were depleted and the specific growth rate dropped. The predicted FHA concentration was 108 $\mu g.l^{-1}$; the measured concentration at the end of the exponential phase of the batch was 123 ± 15 $\mu g.l^{-1}$ (for five experiments, see Fig. 3).

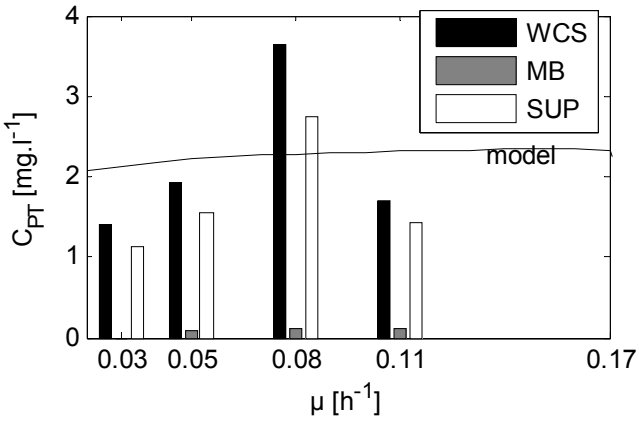
Figure 3 shows that the FHA concentrations at the end of the exponential phase of the batch are close to the concentrations measured in CSTR cultivations. FHA expressed per biomass, however, is slightly lower at the end of the batch, because biomass concentrations are slightly higher.

Note that the model is obtained using steady-state experiments; so quasi-stationary responses can be predicted. The effect of disturbances or abrupt changes in specific growth rate on mRNA and protein formation could be in the order of minutes to hours and is not incorporated in the model. These effects could be inferred using time series of gene and protein expression in batch or perturbed CSTR or fed-batch experiments in future.

PT

Figure 4 shows that PT is predominately secreted. Thalen et al. (2006) measured that PT is also predominately secreted during batch cultivation. The expression of PT is growth associated (Fig. 4), which is in agreement with Licari et al. (1991) and Rodriguez et al. (1994) who measured that PT production growth associated up to a maximum applied specific growth rate of 0.07 h^{-1}. The model (Eqs. A1-A7, 1-2) predicts similar μ-profiles for PT (Fig. 4) as shown in the previous paragraph for FHA. PT was not measured during batch cultivations.

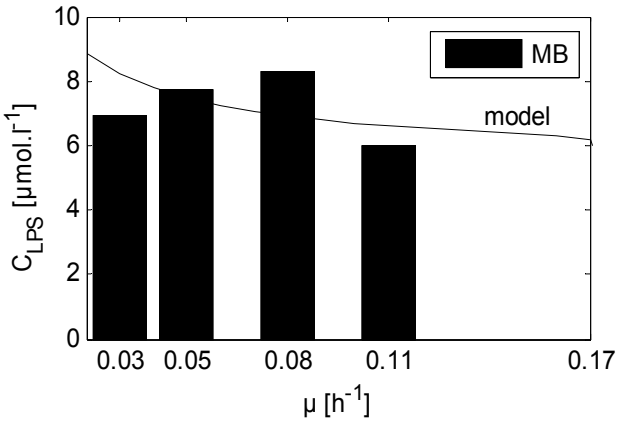
Figure 4. *Experimental results and model fit for the effect of specific growth rate on the steady-state PT concentration. WCS is whole-cell suspension, MB membrane bound, and SUP supernatant.*



LPS

Rodriguez et al. (1994) measured that LPS production was always growth associated up to a maximum applied specific growth rate of 0.07 h^{-1}. Figure 5 indicates that the formation of LPS per cell is growth associated, which is in agreement with Rodriguez et al. (1994). The differences in the measured LPS concentrations are small (Fig. 5). In our CSTR cultivations, the minimum LPS concentration was obtained at 0.11 h^{-1} (Fig. 5). The model (Eqs. A1-A7, 1-2) predicts slightly lower LPS concentrations with increasing specific growth rate.

Figure 5. *Experimental results and model fit for the effect of specific growth rate on the steady-state concentration of membrane bound (MB) LPS.*



Virulence factors and LPS on gene-product level

On protein level, the content of a particular virulence factor per unit cell is almost invariant with specific growth rate; on gene level the expression of the virulence activated genes was approximately constant throughout the different specific growth rates. So based on the

correspondence between the gene and protein patterns the gene–protein data pairs may be classified as regulated.

The genes for FHA and PT are positioned on the same operon and therefore the gene expression is expected to be coupled, which is confirmed by the data. The genes for LPS expression are on a different operon and can therefore exhibit different expression profiles. The differences in gene expression for LPS are smaller than for the virulence factors. On product-level, the specific production rates for LPS contain a larger non-growth associated dependency (and a smaller growth associated dependency) than for FHA and PT.

Discussion and conclusions

Traditionally, most vaccines are produced in batch cultivation, where cells grow until the main nutrients are depleted. The product is released on the basis of offline tests at the end of the cultivation step (e.g. sterility, potency, homogeneity, and animal tests), involving large variability, delayed release, and high costs. In order to improve consistency and quality (see also Soons et al 2006), to save time, and to reduce costs, it is attractive to replace the traditional batch cultivation by specific growth rate controlled fed-batch or continuous cultivation. Although the model postulated in this work is based on a limited number of measurements and it is not that precise, the encouraging results suggest that modelling can be used to predict formation of virulence factors as function of specific growth rate. In this way product quality could be designed and known beforehand instead of being only tested in the end product.

In literature, the mechanism for expression of virulence factors is classified as an on/off switch, in which the expression of virulence factors is activated or deactivated. The results in this work indicate that *B. pertussis* is in the activated phase (*bvg*+) by abundantly expressing the virulence factors.

A high specific growth rate ($\mu \geq 0.08\ h^{-1}$) favours high virulence factor formation, fast production of biomass and low LPS content per cell. The choice of specific growth rate to produce the best quality vaccine depends on multiple objectives like: maximize virulence factors per cell and/or the total amount of virulence factors to induce protection, minimize LPS content to minimize side effects, maximize biomass in minimum time to safe time and costs, etc. The model for formation of virulence factors presented here is a very first step towards more insight. It must be further scrutinized by more experimentation, modification and adjustment on the basis of additional experimental evidence. However, as has been

demonstrated here, once this has been done, the availability of such a model opens perspectives to design feed strategies during the cultivation and to optimize the quality of the vaccine using an objective function to weight the different quality aspects. In this way, a more science-based method for producing vaccines can be developed, meeting the demands of the FDA's PAT initiative to improving the production process by online monitoring and controlling product quality.

Acknowledgments

We like to thank Alex de Haan for measuring LPS, Jeroen Pennings for analyzing the micro-arrays, Joeri Kint for performing the batch cultivations, and Bert Zomer for performing the NMR measurements. The Dutch Ministry of Economical Affairs (TSGE3067) funds the project. Laboratory experiments were performed at NVI, The Netherlands.

Appendix A: Model and parameter determination for dual substrate model of biomass growth

Growth of *B. pertussis* is limited by two substrates. The metabolism is described in detail by Thalen et al. (1999) and can be generalised to the formation of biomass from glutamate and lactate by two major pathways: glutamate alone (pathway 1) or glutamate and lactate (pathway 2). The organism can grow on glutamate only, but growth on lactate alone is not possible. In the current medium, glutamate is an essential, and lactate is an enhancing substrate. Growth via these two pathways is assumed to be parallel, and thus the individual growth rates can be added. Neeleman et al. (2001) proposed the following dual substrate model (Eqs. 1-6), which applies for batch, fed-batch, and continuous-flow stirred-tank (CSTR) cultivations. Monod kinetics and oxygen excess are assumed.

$$\mu(C_G, C_L) = \mu_{\max} \cdot \frac{C_G}{K_G + C_G} + \mu_{enh} \cdot \frac{C_G}{K_G + C_G} \cdot \frac{C_L}{K_L + C_L} \qquad \text{A1}$$

$$\frac{dV}{dt} = F_S - F_{out} \qquad \text{A2}$$

$$\frac{dC_X}{dt} = \left(\mu(C_G, C_L) - \frac{F_S}{V} \right) C_X \qquad \text{A3}$$

$$\frac{dC_G}{dt} = \frac{F_S}{V}\left(C_G^{in} - C_G \right) - q_G^{ov} C_X \qquad \text{A4}$$

$$q_G^{ov} = \frac{\mu_{\max} \frac{C_G}{K_G + C_G}}{Y_{G1}} - \frac{\mu_{enh} \frac{C_G}{K_G + C_G} \frac{C_L}{K_L + C_L}}{Y_{G2}} + m_G \qquad \text{A5}$$

$$\frac{dC_L}{dt} = \frac{F_S}{V}\left(C_L^{in} - C_L\right) - q_L^{ov} C_X \qquad \text{A6}$$

$$q_L^{ov} = \frac{\mu_{enh} \frac{C_G}{K_G + C_G} \frac{C_L}{K_L + C_L}}{Y_L} + m_L \qquad \text{A7}$$

where C_G and C_L are the glutamate respectively lactate concentration, q_G^{ov} and q_L^{ov} the overall specific consumption rates for glutamate and lactate, Y_{G1}, Y_{G2}, and Y_{L2} the biomass yields over the different pathways. μ_{max} is the maximum specific growth rate over pathway 1, μ_{enh} the "enhancing" specific growth rate over pathway 2. K_G and K_L are Monod constants. The biomass growth rate is directly related to biomass (C_X), specific growth rate (μ), and dilution rate (the incoming substrate feed rate (F_S) divided by the volume of the broth (V)).

Neeleman et al. (2001) calculated the yield using a series of batch cultivations in a step-wise procedure. First, the yield on glutamate only Y_{G1} is calculated by linear regression using data from cultivations with glutamate only; second the overall yield on lactate Y_L^{ov} and on glutamate Y_G^{ov} are calculated on cultivations containing lactate and glutamate; third, Y_{G2} and Y_L are calculated using the overall yields:

$$Y_L = \frac{\mu_{enh}}{\mu_{enh} + \mu_{enh} Y_L^{ov}} \qquad \text{A8}$$

$$Y_{G2} = \frac{\mu_{enh} Y_{G1} Y_G^{ov}}{\mu_{\max} Y_G^{ov} - \left(\mu_{\max} + \mu_{enh}\right) Y_{G1}} \qquad \text{A9}$$

Although several parameters, e.g. μ_{max}, μ_{enh}, can be properly calculated from batch cultivations, maintenance and Monod constants can only be approximated, because their contribution is small during major part of the batch ($\mu \cdot Y^{-1} >> m$).

CSTR cultivations allow better calculation of these parameters. The overall yields and maintenance coefficients can be calculated from the "measured" overall consumption rates, which are calculated from the steady state balances Eqs. A3-A7 (where the subscript $_m$ indicates the measurements from experiments) for the substrates:

$$q_{G,m}^{ov} = \frac{\left(C_{G,m}^{in} - C_{G,m}\right) D}{C_{X,m}}, q_{L,m}^{ov} = \frac{\left(C_{L,m}^{in} - C_{L,m}\right) D}{C_{X,m}} \qquad \text{A10}$$

where the dilution rate is equal to the specific growth rate in steady-state:

$$D = \frac{F_S}{V} = \mu \qquad \text{A11}$$

Figure A1 shows that the yield and maintenance coefficients for substrate consumption are constant and independent of the specific growth rate.

The Monod constants are estimated by minimisation of the normalized squared model error *E* for a steady-state situation:

$$\min_{K_G,K_L} E = \min_{K_G,K_L} \frac{\sqrt{\sum_{k=1}^{N}\left[C_G^m(k) - C_G^{\text{mod}}(k)\right]^2}}{\bar{C}_G^m(N-1)} + \frac{\sqrt{\sum_{k=1}^{N}\left[C_L^m(k) - C_L^{\text{mod}}(k)\right]^2}}{\bar{C}_L^m \cdot (N-1)} \qquad \text{A12}$$

Figure A2 shows that the dual substrate model fits well the measurements. The estimated parameters are summarized in Table 1.

The presented model Eqs. A1-A9 is based on primary substrates. Accumulating metabolites or waste products (e.g. acetoacetate), however, may affect growth and should then be part of the model. NMR measurements, performed to detect accumulation of metabolites or waste products in the supernatant of the cultivation broth, showed that such components were absent at any specific growth rate.

Figure A1. *Specific glutamate and lactate consumption rates as function of specific growth rate for calculation of the overall yields as in Pirt, 1982 (Y_G^{ov} = 0.168 and Y_L^{ov} = 0.101).*

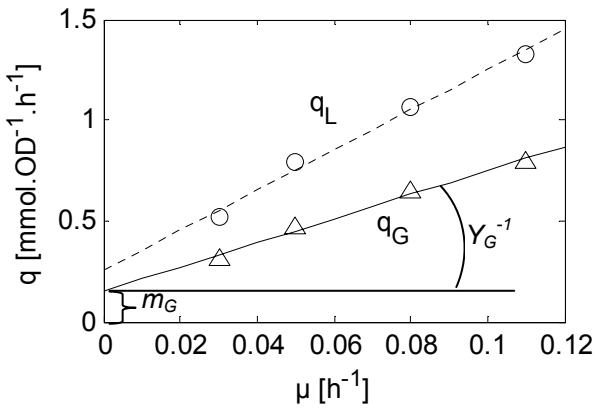


Figure A2. *Measured biomass (+), glutamate (Δ), and lactate (○) concentrations at four specific growth rates, and model fit using Eqs. A1-A7 (lines).*

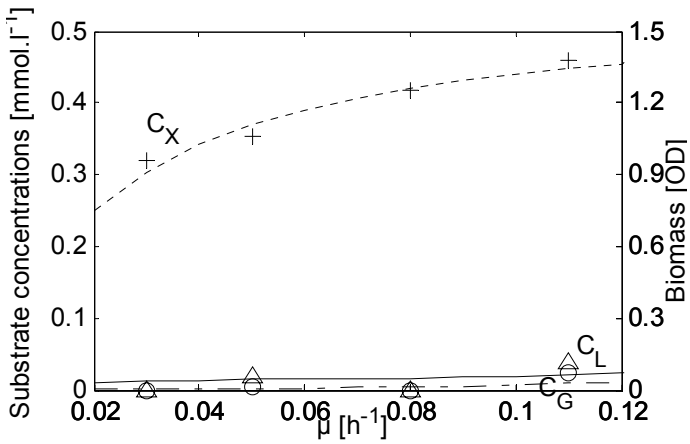


Table A1. *Yields and maintenance coefficients for the dual substrate model of B. pertussis*

	This work	Literature (Neeleman et al. 2002)
Y_{G1}	-	0.055 $OD.mmol^{-1}$
Y_{G2}	0.048 $OD.mmol^{-1}$	0.061 $OD.mmol^{-1}$
Y_L	0.032 $OD.mmol^{-1}$	0.018 $OD.mmol^{-1}$
m_G	0.15 $mmol.OD^{-1}.h^{-1}$	0 $mmol.OD^{-1}.h^{-1}$
m_L	0.26 $mmol.OD^{-1}.h^{-1}$	0 $mmol.OD^{-1}.h^{-1}$
K_G	5.10^{-3} $mmol.l^{-1}$	0.5 $mmol.l^{-1}$
K_L	1.10^{-5} $mmol.l^{-1}$	0.5 $mmol.l^{-1}$
Y_O	0.033 $OD.mmol^{-1}$	0.041 $OD.mmol^{-1}$
m_O	0.69 $mmol.OD^{-1}.h^{-1}$	0.41 $mmol.OD^{-1}.h^{-1}$
a_{FHA}	72 $\mu g.OD^{-1}$	-
b_{FHA}	0.46 $\mu g.OD^{-1}.h^{-1}$	-
a_{PT}	1.49 $\mu g.OD^{-1}$	-
b_{PT}	0.025 $\mu g.OD^{-1}.h^{-1}$	-
a_{LPS}	3.37 $nmol.OD^{-1}$	-
b_{LPS}	0.17 $nmol.OD^{-1}.h^{-1}$	-

CHAPTER 8

Conclusions & Perspectives

Conclusions

In biopharmaceutical manufacturing quality of the product is currently tested at the end of the production process only. Bioprocesses naturally contain variability due to the raw materials, initial conditions, human intervention, and varying properties of the micro-organism. By setting fixed protocols the current GMP regulations do not encompass the reduction of these deviations. This leads to products containing a considerable amount of variability, and possibly out of specifications. Recently the FDA released guidelines on process analytical technology (PAT) (FDA, 2004) - a framework for innovative pharmaceutical development, manufacturing, and quality assurance - in which they stress the need for reducing variability and aim at improving manufacturing processes by introducing timely measurement and control of quality variables.

In different parts of science the word “control” has different meanings (Wold, 2006). In the world of process analytical technology or statistical process control “control” often means “check” whether the process is still within the desired specifications by means of monitoring or classifying the process. In the engineering sense “control” means ensure that the process stays close to the predefined set-points or trajectories. Wold (2006) also states that “PAT process control” in the engineering sense is still a challenge of the future.

Although such control technologies may still be a challenge using spectroscopic instrumentation often associated with PAT, control in the engineering sense is already available for PAT (Soons et al., 2006; Gnoth et al., 2007).

This thesis is a step forward to real-time feedback control for PAT by modelling, monitoring, and controlling biomass and product formation of the cultivation step of *Bordetella pertussis* for the production of bulk whole cell whooping cough vaccine towards desired quality specifications and enhanced batch-to-batch consistency. This answers to the five research questions formulated in chapter 1 are summarized below. They were obtained by combining different techniques from a multidisciplinary field of systems and control theory, biology, and chemometrics. The research questions and solutions are summarized in the following sections.

How to improve monitoring of complex bioprocesses?

An advanced technique has been developed and evaluated to enhance bioprocess monitoring by reconstructing biomass and specific growth rate. The software sensor is sensor based on a standard and cheap dissolved oxygen probe and is evaluated in comparison to another advanced technology, spectroscopy based on a dedicated near infrared sensor.

- *Software sensor*. The software sensor monitors the specific growth rate and biomass in real-time, which can not be measured online due to the lack of suitable sensors. The software sensor is based on an Extended Kalman Filter that combines standard measurements of the oxygen uptake rate with a generic model for biomass growth to observe the otherwise unknown biomass growth while reducing the effect of measurement noise. Tuning is based on simulations with a worst case scenario for measurement noise and guided by stability criteria, such that the software sensor accurate, robust, and is widely applicable. Biomass and specific growth rate are accurately estimated over the whole operating range from low to high biomass concentrations during fed-batch cultivations. Offline biomass measurements are incorporated to improve and safeguard the biomass estimation and to observe the effect of biomass growth on the oxygen transfer coefficient. This ability offers potential for high cell density cultivations, in which k_La will drastically decrease due to blanketing effects.
- *Near infrared spectroscopy*. Near infrared spectroscopy is a technique often associated with PAT to enhance process monitoring. The interpretation of near infrared spectra is challenging due to the large number of wavelengths recorded and the overlapping absorbance features of near infrared spectroscopy. A controlled random search procedure selects an optimal window of wavelengths, which is used in the partial least squared routine to regress on the biomass measurements. The proposed wavelengths selection procedure outperforms the traditional calibration procedures by enhancing the prediction accuracy of biomass in real-time. The near infrared predictions depend on the quality of the training dataset, which needs to encompass all possible sources of temporal disturbances like pH and dissolved oxygen.

The option to use both sensors is attractive in practice to safeguard against sensor failure. Considering the current limitations of near infrared monitoring and the challenges to be solved, such as a fixed path length, a linear PLS model, and a limited number of batches, at

present the software sensor is the preferred choice for monitoring and feedback control of biomass and specific growth rate.

How can the cultivation process be controlled at a desired level while coping with the nonlinear and time-varying characteristics of bioprocesses?

The control aim was to track the set-point for specific growth rate by adding a feed with limiting substrates. Using the laws of model-reference control in combination with a dual substrate model for biomass growth, a stable adaptive "PI" controller was derived. The controller settings are adapted to the changing process dynamics (volume, biomass, and set-point for specific growth rate) in order to cope with the time-varying characteristics of fed-batch processes. Tuning based on a combination of deviations in specific growth rate and dissolved oxygen leads to bioreactor control without interactions between specific growth rate and dissolved oxygen controller.

Experiments showed that the controller performance was good for its ability to cope with time-varying kinetics and states, noise, and external disturbances. Enhancing the standard batch cultivation with a fed-batch phase yielded a six times higher biomass yield and a constant and high antigen activity of the bulk vaccine. The controller, therefore, contributes to enhanced batch-to-batch consistency and quality. So being a step forward to "PAT process control", "PAT process control" is less a challenge of the future than commonly believed.

How to automate controller tuning?

The challenge was to upgrade performance of poorly acting controllers automatically and thereby reducing the tuning effort. Automating controller tuning may facilitate non-professionals to work with advanced controllers. Three automatic tuning methods were applied to upgrade control performance for specific growth rate control during fed-batch cultivation. We expect that a cultivation controlled at the desired specific growth rate will result in smaller variations in end quality (vaccine titer) and thus yield a better product (vaccine). The best two methods were qualified by laboratory experiments for *B. pertussis*. The methods do not require online identification, thus avoiding the need for process perturbation and complex implementation. Control performance is evaluated online on the basis of the current mean absolute error and oscillation measure. If control performance is poor, application of automatic tuning yields good performance within five hours by adapting

the controller parameters such that the mean absolute error and oscillation measure decrease at least ten-fold. The closed loop with automatic tuning is stable at any point along the trajectory of fed-batch cultivation.

What set-points are needed to obtain the best-quality vaccine?

The quality of whole cell whooping cough vaccine is based on the presence of outer-membrane proteins that are important for inducing a protective immune response. While modelling of the synthesis of complex products and processes is uncommon and difficult, an attempt was made to model the expression of virulence factors by enhancing an existing model for the cultivation of *B. pertussis* for the production of whole-cell vaccine against whooping cough. The data and model suggest that the formation rate of virulence factors is proportional to the specific growth rate; and thus that a high specific growth rate ($\mu \geq 0.08\ h^{-1}$) is favourable for required vaccine quality, in the sense of high contents of virulence factors per cell and relatively low content of the undesired lipopolysaccharide component.

The choice of the required specific growth rate to produce the best quality vaccine depends on multiple objectives like: maximize virulence factors per cell and/or the total amount of virulence factors to induce protection, minimize lipopolysaccharide content to minimize side effects, maximize biomass in minimum time to safe time and costs, etc. Modelling the formation of virulence factors opens perspectives for online or offline use to design feed strategies during the cultivation and to optimize the quality of the vaccine using an objective function that weight different quality aspects. In this way, a more science-based method for producing vaccines can be developed, meeting the demands of the FDA's PAT initiative to improve the production process by online monitoring and control of product quality.

Can the developed monitoring & control tools be applied to cultivation on production-scale?

Monitoring and control systems in biotechnology are usually designed and tested in laboratory-scale experiments. Application on production-scale is limited in literature. In scaling-up from laboratory to pilot-scale and beyond two problems are encountered for which solutions have been worked out. The presence of a headspace in the bioreactor and tubing to the analyzer cause mixing effects and delays. Incorporating the gas phase dynamics in the monitoring system gives up to 30% more accurate observations of the oxygen consumption and biomass growth. Using models which contain system-relevant parameters

that can be easily obtained from the equipment specifications enhances the applicability of the method. The other issue is whether the application of the observer and controller for biomass growth developed for small-scale cultivation processes is applicable to pilot-scale cultivation. It appeared that the observer and controller could be applied directly to the larger-scale cultivation without any modifications. So, the observer and controller are independent of scale and micro-organism. The monitoring and control system for (fed-)batch cultivation shows good performance and enables an at least doubled concentration of bulk vaccine on pilot-scale compared to the standard batch production process in a controlled manner. So a more favourable upstream production method for vaccines.

Perspectives

The potency of vaccines is mainly determined in the cultivation step of the micro-organisms, in which the bulk product is formed. Currently vaccines are released on the basis of potency and safety tests on the final vaccine lots. In Europe, approximately 1.5 million laboratory animals are being used for routine quality control of vaccines (Metz et al., 2002). Manufacturers, regulatory authorities, and ethical commissions wish to reduce animal use, because of drawbacks like costs, imprecision, and ethical concerns. Metz (2005) proposes to minimize animal tests by developing *in vitro* tests using physiochemical and immunochemical techniques as an alternative for the *in vivo* potency test; Streefland et al. (2007) by developing DNA arrays to asses quality indicators at the end of the cultivation. Another development to reduce animal tests in the future is parametric release, or real-time release on the basis of online monitored and controlled key variables. Critical variables must be defined that correlate well with product quality and that can be accurately monitored and controlled online, such that quality of the product at the end of the process is guaranteed. Subsequently the bulk product can be released for the next step. This thesis is a first step in real-time monitoring and control towards high quality and consistently produced vaccines. Some further opportunities for improvements in biopharmaceutical manufacturing are discussed below.

Modelling towards control of vaccine quality

The model for formation of virulence factors developed in this thesis opens perspectives for online or offline use to design feed strategies during the cultivation and to optimize the quality of the bulk vaccine. Specific growth rate may be one critical variable for release, but

surely there are more critical variables determining vaccine quality. Other variables that can be monitored online, which influence the formation of virulence factors are e.g. dissolved oxygen, pH, and osmolarity. The effect of these variables on the expression of the virulence factors on gene and protein level is valuable for the vaccine quality model. The extended model has the potential to be the basis for an enhanced software sensor and feedback control towards higher consistency and quality. Moreover, the integration of data on gene expression, protein formation, and metabolites will allow insight into the systems properties, prediction of perturbation effects, prediction of responses on certain set-points, and gains insight in preferred feedback control strategies.

Enhanced monitoring

Release on the basis of online monitored variables requires a high accuracy and robustness in monitoring in order to prevent release of unsatisfactory products or decline of good products. In view of this purpose and in view of feedback control, the accuracy of biomass estimation from near infrared spectroscopy must be enhanced. Potential improvements of accuracy and robustness may be obtained by incorporating dissolved oxygen and pH as independent factors in the regression, perhaps in combination with nonlinear regression methods based on a larger number of batches that contain the natural variability of biological processes.

Near infrared spectroscopy is based on molecular overtone and combination vibrations. Although an advantage of near infrared is that the near infrared light typically penetrates much further into a sample, the absorbance of specific chemical components at a given wavelength is weak. Due to the complexity of the near infrared spectra it can be difficult to calibrate specific components. Mid infrared spectroscopy is based on fundamental vibrations associated with the rotational-vibrational structure. The absorbance of specific components at a given wavelength is stronger for mid infrared spectroscopy and may therefore be promising for biomass monitoring.

The near infrared sensor is often mentioned as the ideal sensor to monitor quality performance of the process. None of the currently available sensors, however, can directly measure product quality in real-time. Vaidyanathan et al. (2001) observed that low concentrations of protein can only be successfully modelled by using the filtrates and not whole broth measurements due to the dominance of the biomass absorbance that overlaps the protein absorbance. We doubt whether near infrared is suited for directly monitoring of virulence factors (attached to the biomass), essential for quality of vaccines. But comparing

the course of the process with known qualities can indirectly predict whether the process is within specifications. This could be used together with an enhanced software sensor that predicts formation of virulence factors on the basis of measurements from standard sensors.

Pilot-scale fed-batch vaccine production

Approximately $1x10^7$ doses of vaccine are necessary to vaccinate the Dutch child population in one year. So far, most vaccines are produced in large-scale bioreactors (>1000L) by means of batch cultivation. According to the draft European guidelines for whole cell pertussis vaccine, bacterial concentration should not exceed 1 OD per single human dose (Anonymous, 2007). This means that one production run at large-scale (method 1 in table 1) would yield sufficient bulk whooping cough doses to vaccinate all young children in the Netherlands in one year. This is a highly inefficient way of producing vaccines: the down-time is high compared to the running time. Moreover, this production method may cause problems to meet the GMP guidelines, because difficulties may arise to maintain skills and experience of the production personnel to properly perform the cultivations if so few runs are performed every year; and because the equipment has to be reconstructed and carefully cleaned to be able to produce different products throughout the year with potentially a higher risk of cross-contamination. Moreover, large-scale production may lead to increased inhomogeneity due to large mixing times.

Alternatively, nine controlled pilot-scale fed-batch cultivations, as developed in this thesis, would be necessary (method 2), or three runs using an improved process, where the oxygen transfer coefficient is increased by using sparger instead of headspace aeration (method 3). In view of the efficient use of the bioreactor, method 2 and 3 are more profitable than the traditional batch cultivation on production-scale, still allowing possibilities for vaccine production for larger markets if more pilot-scale runs are performed.

So, the combination of enhancing an existing process with controlled fed-batch cultivation with the achievement of higher yields of bulk vaccine on pilot-scale in a controlled way gives a favourable upstream production method for vaccines. Whooping cough vaccine was shown as an example in the foregoing. Many more biopharmaceuticals may profit from the cultivation methods proposed in this thesis.

Table 1. *Cultivations strategies. Data for method 1 and 2 are obtained from real cultivations, data for method 3 are extrapolated from small-scale experiments.*

Reactor (cultivation broth) volume [l]	End biomass [OD]	Doses of vaccine	Number of required runs	Method
1200 (800)	1.6	$1.3x10^7$	1	1. Batch
60 (35)	3.5	$1.2x10^6$	9	2. Fed-batch (headspace aeration)
60 (40)	10	$4.0x10^6$	3	3. Improved fed-batch (sparger aeration)

REFERENCES

Abrahamsson C, Johansson J, Sparén A, Lindgren F (2003) Comparison of different variable selection methods conducted on NIR transmission measurements on intact tablets. Chemometrics Intelligent Lab Syst 69: 3-12

Alford JS (2006) Bioprocess control: Advances and challenges. Comput Chem Eng 30: 1464-1475

Akay B, Ertunç S, Kahvecioğlu A, Hapoğlu H, Alpbaz M (2002) Adaptive control of *S. cerevisae* production. Trans IchemE 80: 28-38

Alvarez, A, Simutis R (2004) Application of Kalman filter algorithm in GMC control strategy for fed-batch cultivation process. Informacinės Technologijos Ir Valdymas 1: 7-12

Anonymous (2007) Pertussis vaccine (whole cell, absorbed). Reference: PA/PH/Exp 15/T(07) 16 ANP. Pharmeuropa 19: 644-646

Åström KJ, Hägglund T (1988) Automatic PID tuning. United States of America: Instrument Society of America

Åström KJ, Hägglund T, Hang CC, Ho WK (1993) Automatic tuning and adaptation for PID controllers –A survey. Control Eng Prac 1: 699-714

Åström KJ, Wittenmark B (1995) Adaptive control. Addison-Wesley Publishing Company, United States of America

Baart GJE, De Jong G, Philippi M, Van't Riet K, Van Der Pol LA, Beuvery EC, Tramper J, Martens DE (2007) Scale-up for bulk production of vaccine against meningococcal disease. Vaccine 25: 6399-6408

Babuška R, Damen MR, Hellinga C, Maarleveld H (2003) Intelligent adaptive control of bioreactors. J Intell Manuf 14: 255-265

Bakshi BR, Locher G, Stephanopoulos Gr, Stephanopoulos Ge (1994) Analysis of operating data for evaluation, diagnosis and control of batch operations. J Proc Cont 4: 179-194

Bastin, G, Dochain D (1990) On-line estimation and adaptive control of bioreactors. Elsevier Science Publishers, Amsterdam, The Netherlands

Berber R, Kravaris C, Niemiec M, Brosilow CB (1998) Nonlinear model based process control. Kluwer Academic Publishers, Dordrecht, The Netherlands, 115-142

Bloemen HHJ, Wu L, Gulik WM, Heijnen JJ, Verhaegen MHG (2003), Reconstruction of the O_2 uptake rate and CO_2 evolution on a time scale of seconds. Bioeng Food Natural Products 49: 1985-1908

Casas López JL, Rodríguez Porcel EM, Oller Alberola I, Ballesteros Martin MM, Sánchez Pérez JA, Fernández Sevilla JM, Chisti Y (2006) Simultaneous determination of oxygen consumption rate and volumetric oxygen transfer coefficient in pneumatically agitated bioreactors. Ind Eng Chem Res 45: 1167-1171

Chang D (2003) The snowball effect in fed-batch bioreactors. Biotechnol Prog 19: 1064-1070

Chen L, Bastin G, Van Breusegem V (1995) A case study of adaptive nonlinear regulation of fed-batch biological reactors. Automatica 31: 55-65

Cimander C, Mandenius C-F (2004) Bioprocess control from a multivariate process trajectory, Bioprocess Biosyst Eng 26: 401-411

Cummings CA, Bootsma HJ, Relman DA, Miller JF (2006) Species- and strain-specific control of complex, flexible regulon by *Bordetella* BvgAS. J Bacteriol 188: 1775-1785

Dagci OH, Efe MO, Kaynak O, Yu X (2001) Variable structure systems of biochemical processes. Proceedings of ISIE, Pusan, Korea, 1690-1695

Dekkers RM (1982) State estimation of a fed-batch baker's yeast fermentation, Modelling and Control of Biotechnological Processes, Helsinki, Finland, published by Perfamom, Oxford, U. K., 201-211

Deming WE (1986) Out of the crisis. The MIT press, Massachusetts.

Desbourough L, Miller R (2002) Increasing customer value of industrial control performance monitoring—Honeywell's experience, Sixth International Conference on Chemical Process Control, AIChE Symposium Series Number 326, 98

Dhir S, Morrow Jr KJ, Rhinehart RR, Wiesner T (2000) Dynamic optimization of hybridoma growth in a fed-batch bioreactor. Biotechnol Bioeng 67: 197-205

Dochain D (2003) State and parameter estimation in chemical and biochemical processes: a tutorial. J Proc Cont 13: 801-818

Dondo R, Marqués D (2003) Simulation results for on-line optimization of a batch bioreactor using nonlinear filtering and optimal control. ISA Transactions 42: 289-303.

Doran PM (1995) Bioprocess engineering principles. Academic Press, San Diego

Estler MU (1995) Recursive on-line estimation of the specific growth rate from off-gas analysis for the adaptive control of fed-batch processes. Bioprocess Eng. 12: 205-207

Ferreira AP, Cardoso de Menezes J (2007) Determination of protein and fat content in fermentation raw materials with NIR reflectance spectroscopy, Preprints Vol. I. Computer Applications in Biotechnology 10: 347-351

Gnoth S, Jenzsch M, Simutis R, Lübbert A (2007) Process Analytical Technology (PAT): Batch-to-batch reproducibility for fermentation processes by robust process operational design and control J Biotechnol 132: 180-186

Food and Drug Administration (FDA) (2004) Guidance for Industry PAT – A framework for Innovative Pharmaceutical Development, Manufacturing and Quality Assurance, Guidance Document. US Department of Health and Human Services

Frahm B, Lane P, Atzert H, Munack A, Hoffmann M, Hass VC, Pörtner R (2002) Adaptive, model-based control by the Open-Loop-Feedback-Optimal (OLFO) controller for the effective fed-batch cultivation of hybridoma cells. Biotechnol Prog 18: 1095-1103

Galaction AI, Cascaval D, Oniscu C, Turnea M (2004) Prediction of oxygen mass transfer coefficients in stirred bioreactors for bacteria, yeasts and fungus broths. Biochem Eng J 20: 85-94

Gee DA, Ramirez WF (1996) On-line state estimation and parameter identification for batch fermentation. Biotechnol Prog 12: 132-140

Ghoul M, Boudrant J, Engasser JM (1991) A comparison of different techniques for the control of the growth of *Candida utilis* CBS 621. Proc Biochemistry 26: 135-142

Gnoth S, Jenzsch M, Simutis R, Lübbert A (2007) Process Analytical Technology (PAT): Batch-to-batch reproducibility for fermentation processes by robust process operational design and control. J Biotechnol 132: 180-186

Gorry PA (1990) General least-squares smoothing and differentiation by the convolution (Savitzky-Golay) method. Anal Chem 62: 570-573

Gudi RD, Shah SL, Gray MR, Aran PKY (1997) Adaptive multirate estimation and control of nutrient levels in a fed-batch fermentation using off-line and on-line measurements. Canadian J Chem Eng 75: 562-573

Gudi RD, Shah SL, Gray MR (1995) Adaptive multirate state and parameter estimation strategies with application to a bioreactor. AIChE J 41: 2451-2464

Haaland DM, Thomas EV (1988) Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. Anal Chem 60: 1193-1202

Helland IS (1988) On the structure of partial least squares regression. Commun Stat – Simul 17: 581-607

Hot D, Antoine R, Renauld-Mongénie G, Caro V, Hennuy B, Levillain E, Huot L, Wittmann G Poncet D, Jacob-Dubuisson F, Guyard C, Rimlinger F, Aujame L, Godfroid E, Guiso N, Quentin-Millet M-J, Lemoine Y, Locht C (2003) Differential modulation of *Bordetella pertussis* as evidenced by DNA microarray analysis. Mol. Gen. Genomics 269: 475–486

Ignatova M, Lubenova V, Georgieva P (2000) MIMO adaptive linearizing control of fed-batch amino acids simultaneous production. Bioprocess Eng 22: 79-84

Ignatova MN (2003) Adaptive biomass observer on the basis of measurements of oxygen. 11th Mediterranean conference on control and automations, Rhodes, Greece

Ioannou PA, Sun J (1996) Robust adaptive control. Prentice Hall Inc, New Jersey, USA

Jacob-Dubuisson F, Kehoe B, Willery E, Reveneau N, Locht C, Relman DA (2000) Molecular characterization of *Bordetella bronchiseptica* filamentous haemagglutinin and its secretion machinery. Microbiol 146: 1211–1221

Jenzsch M, Simutis R, Eisbrenner G, Stückrath I, Lübbert A (2006) Estimation of biomass concentrations in fermentation processes for recombinant protein production. Bioproc Biosyst Eng 29: 19-27

Jenzsch M, Gnoth S, Kleinschmidt M, Simutis R, Lubbert A (2007) Improving the batch-to-batch reproducibility of microbial cultures during recombinant protein production by regulation of the total carbon dioxide production. J Biotechnol 128: 858-867

Junker BH, Wang HY (2006) Bioprocess monitoring and computer control: key roots of the current PAT initiative. Biotechnol Bioeng 95: 226-261

Karim MN, Hodge D, Simon L (2003), Data-based modeling and analysis of bioprocesses: some real experiences. Biotechnol Prog19: 1591:1605

Lopes JA, Costa PF, Alves TP, Menezes JC. 2004. Chemometrics in bioprocess engineering: process analytical technology (PAT) applications. Chemometrics Intelligent Lab Syst 74:269– 275.

Keesman KJ (2002) State and parameter estimation in biotechnical batch reactors. Control Eng Practice 10: 219-225

Kerr JR, Matthews RC (2000) *Bordetella pertussis* infection: pathogenesis, diagnosis, management, and the role of protective immunity. Eur J Clin Microbiol Infect Dis 19: 77-88

Krastanov, MI, Dimitrova NS (2003) Stabilizing feedback of a nonlinear process involving uncertain data. Bioprocess Biosyst Eng 25: 217-220

Lacey BW (1960) Antigenic modulation of *Bordetella pertussis*. J Hyg 58: 57–93

Levisauskas D, Simutis R, Borvitz D, Lübbert, A (1996) Automatic control of the specific growth rate in fed-batch cultivation processes based on exhaust gas analysis. Bioprocess Eng 15: 145-150

Levisauskas D (2001) Inferential control of the specific growth rate in fed-batch cultivation processes. Biotechnol Letters 23: 1189-1195

Levisauskas D, Galvanauskas V, Henrich S, Wilhelm K, Volk N, Lubbert A (2003) Model-based optimization of viral capsid protein productionin fed-batch culture of recombinant *Escherichia coli*. Bioproc Biosyst Eng 25: 255–262

Lewis FJ (1986) Optimal estimation. John Wiley and Sons, New York.

Licari P, Siber GR, Swartz R (1991) Production of cell mass and Pertussis toxin by *Bordetella pertussis*, Journal of Biotechnology 20, 117-130

Ljung L (1979) Asymptotic behaviour of the Extended Kalman Filter as a parameter estimator for linear systems. IEEE transaction on Automatic Control 24: 36-50

Lübbert A, Jørgensen SB (2001) Bioreactor performance: a more scientific approach for practice. J Biotechnol 85: 187-212

Lubenova VN (1999) Stable adaptive algorithm for simultaneous estimation of time-varying parameters and state variables in aerobic bioprocesses. Bioprocess Biosyst Eng 21: 219-226

Lubenova, V, Rocha I, Ferreira EC (2003) Estimation of multiple biomass growth rates and biomass concentration in a class of bioprocesses. Bioprocess Biosyst Eng 25, 395-406.

Metz B, Hendriksen C, Jiskoot W, Kersten G (2002) Reduction of animal use in human vaccine quality control: opportunities and problems. Vaccine 20: 2411-2430

Metz B, Structural characterization of diphtheria toxoid, Thesis, Utrecht University, The Netherlands

Montague GA, Martin EB (2007) Forecasting for fermentation operational decision making, Preprints Vol. I. Computer Applications in Biotechnology 10: 141-146

Mutha RK, Cluett WR, Penlidis A (1997) A new multirate-measurement-based estimator: emulsion copolymerization batch reactor case study. Ind Eng Chem Res 36: 1036-1047

Myers MA, Kang S, Luecke RH (1995) State estimation and control for systems with delayed off-line measurements. Comput Chem Eng 20: 585-588

Nakamura MM, Liew S-Y, Cummings CA, Bring MM, Dieterich C, Relman DA (2006) Growth phase- and nutrient limitation-associated transcript abundance regulation in *Bordetella pertussis.* Infect Immun 74: 5537-5548

Neeleman R, Joerink M, Van Boxtel AJB (2001) Dual substrate utilisation by *Bordetella pertussis*. Appl Microbiol Biotechnol 57: 489-493

Neeleman R, van Boxtel AJB (2001a) Estimation of specific growth rate from cell density measurements. Bioprocess Biosyst Eng 24: 179-185

Neeleman R (2002) Biomass performance – Monitoring and control in bio-pharmaceutical production, Thesis, Wageningen University, The Netherlands

Neeleman R, Beuvery C, Vries D, Van Straten G, Van Boxtel AJB (2004) Dual-substrate feedback control of the specific growth-rate in vaccine production. Proceedings CAB 9, Nancy, France

Nise NS (2000) Control Systems engineering. John Wiley & Sons, USA

November EJ, Van Impe JF (2002)The tuning of a model-based estimator for the specific growth rate of *Candida utilis*. Bioprocess Biosyst Eng 25: 1-12

Perrier M, Feyo de Azevedo S, Ferreira EC, Dochain D (2000) Tuning of observer-based observers: theory and application to the on-line estimation of kinetic parameters. Control Eng Practice 8: 377-388

Picó-Marco E, Picó J, De Battista H, Navarro JL (2004) Robust adaptive control in biotechnological fed-batch processes. Proceedings CAB9, Nancy, France

Picó-Marco E, Picó J, De Battista H (2005). Sliding mode scheme for adaptive specific growth rate control in biotechnological fed-batch processes. Int J Control 78: 128-141

Pirt SJ (1982) Maintenance energy: a general model for energy-limited and energy-sufficient growth. Arch Microbiol 133: 300- 302

Pörtner R, Schwabe J, Frahm B (2004) Evaluation of selected strategies for fed-batch cultures of a hybridoma cell line. Biotechnol Appl Biochem 40: 47-50

Price WL (1977) A controlled random search procedure for global optimization. Computer J 20: Price WL. 1977. A controlled random search procedure for global optimisation, The Computer J 20:367-370.

Robertsson G (2001) Simple and efficient multivariate calibration of nonlinear spectroscopic data. Appl Spectrosc 55: 98-104

Rocha I, Ferreira EC (2002) Model-based adaptive control of acetate concentration during the production of recombinant proteins with *E. coli*. 15th Triennial World Congress, Barcelona, Spain

Rodriguez ME, Hozbor DF, Samo AL, Ertola R, Yantorno OM (1994) Effect of dilution rate on the release of pertussis toxin and lipopolysaccharide of *Bordetella pertussis.* J Ind Microbiol 13: 273-278

Roggo Y, Chalus P, Maurer L, Lema-Martinez C, Edmond A, Jent N (2007) A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. J Pharm Biomed Anal 44: 683–700

Sabra W, Kim E, Zeng A (2002) Physiological responses of *Pseudomonas aeruginosa* PAO1 to oxidative stress in controlled microaerobic and aerobic cultures. Microbiol 148: 3195-3202

Scarff M, Arnold SA, Harvey LM, McNeil B (2006) Near infrared spectroscopy for bioprocess monitoring and control: current status and future trends. CRC Crit Rev Biotechnol 26: 17-39

Scarlato V, Prugnol A, Arico B, Rappuoli R (1990) Positive transcriptional feedback at the bvg locus controls expression of virulence factors in *Bordetella pertussis.* Proc Natl Acad Sci USA 87: 6753–6757

Siegell SD and Gaden JR EL (1962) Automatic control of dissolved oxygen levels in fermentation. Biotechnol Bioeng 4: 354-356

Smets IY, Bastin GP, Van Impe JF (2002) Feedback stabilization of fed-batch bioreactors: non-monotonic growth kinetics. Biotechnol Prog 18: 116-125

Smets IY, Claes JE, November EJ, Bastin GP, Van Impe JF (2004) Optimal adaptive control of (bio)chemical reactors; past, present and future. J Process Contr 14: 795-805

Sontag ED (2001) Input to state stability and related notions. Proceedings Chemical Process Control 6, Tueson, 109-120

Sontag ED (2006) Some remarks on input choices for biochemical systems, arXiv:math/0606261v1

Sontag ED (2007) Monotone and near-monotone biochemical networks. Syst Synth Biol 1: 59-87

Soons ZITA, Voogt JA, Van Straten G, Van Boxtel AJB (2006) Constant specific growth rate in fed-batch cultivation of *Bordetella pertussis* using adaptive control. J Biotechnol 125: 252-268

Soons ZITA, Van Straten G, Van Boxtel AJB (2006a) Automatic tuning and adaptation for specific growth rate control of fed-batch cultivation. Proceedings Chemical Process Control 7, Lake Louise, Canada

Soons ZITA, Shi J, Van Der Pol LA, Van Straten G, Van Boxtel AJB (2007) Biomass growth and k_La estimation using online and offline measurements. Preprints Vol. I. Computer Applications in Biotechnology, Cancun, Mexico, 81-86

Soons ZITA, Voogt JA, Van Straten G, L.A. Van der Pol, Van Boxtel AJB (2007a) Online automatic tuning and control for fed-batch cultivation. Bioprocess Biosyst Eng, online first

Soons ZITA, Shi J, Stigter JD, Van der Pol LA, Van Straten G, Van Boxtel AJB (2008) Observer design and tuning for biomass growth and k_La estimation using online and offline measurements. J Process Contr, in press

Spriet JA, Botterman J, De Buyser DR, De Visser PL, Vandamme EJ (1982) A computer-aided noninterfering on-line technique for monitoring oxygen-transfer characteristics during fermentation processes. Biotechnol Bioeng 24: 1605-1621

Stenson TH, Peppler MS (2007) Osmolarity affects Bvg-mediated virulence regulation by *Bordetella pertussis*. Can J Microbiol 53: 1053-1061

Stephanopoulos G, San KY (1984) Studies on On-Line Bioreactor Identification. I. Theory. Biotechnol Bioeng 26: 1176-1188

Streefland M, Van De Waterbeemd B, Happé H, Van Der Pol LA, Beuvery EC, Tramper J, Martens DE (2007) PAT for vaccines: The first stage of PAT implementation for development of a well-defined whole-cell vaccine against whooping cough disease. Vaccine 25: 2994-3000

Tatiraju S, Soroush M, Mutharasan R (1999) Multi-rate nonlinear state and parameters estimation in a bioreactor. Biotechnol Bioeng 63: 22-32

Thalen M, Van Den IJssel J, Jiskoot W, Zomer B, Roholl P, De Gooijer C, Beuvery C. Tramper H (1999) Rational medium design for *Bordetella pertussis*; basic metabolism. J Biotechnol 75: 147-159

Thalen M, Venema M, Van Den IJssel J, Berwald L, Beuvery C, Martens D, Tramper J (2006) Effect of relevant culture parameters on pertussis toxin expression by *Bordetella pertussis*. Biologicals 34L 213-220

Triadaphillou S, Martin EB, Montague GA, Nordon A, Jeffkins P, Stimpson S (2007) Fermentation process tracking through enhanced spectral calibration modeling. Biotechnol Bioeng 97: 554-567

Vaidyanathan S, Harvey LM, McNeil B (2001) Deconvolution of near-infrared spectral information for monitoring mycelial biomass and other key analytes in a submerged fungal bioprocess. Anal Chim Acta 428: 41-59

Vaidyanathan S, Macaloney G, McNeil B (1999) Fundamental investigations on the near-infrared spectra of microbial biomass as applicable to bioprocess monitoring. Analyst 124: 157-162

Van Den Berg B, Beekhuizen H, Willems RJL (1999) Role of *Bordetella pertussis* virulence factors in adherence to epithelial cell lines derived from the human respiratory tract. Infect Immun 3: 1056-1062

Van Impe JF, Bastin G (1995) Optimal adaptive control of fed-batch fermentation processes. Control Eng Practice 3: 939-954

Van Sprang ENM, Streefland M, van der Pol LA, Beuvery EC, Ramaker H-J, Smilde AK (2007) Manufacturing vaccines: an illustration of using PAT tools for controlling the cultivation of *Bordetella pertussis*. Quality Eng 19: 373-384

Van't Riet, K, Tramper J (1991) Basis Bioreactor Design. Marcel Dekker, New York

Verboven S, Hubert M (2005) LIBRA: a MATLAB library for robust analysis, Chemometrics Intelligent Lab Syst 75: 127-136

Wang NS, Stephanopoulos GN (1984) Computer applications to fermentation processes. CRC Crit Rev Biotechnol 2: 1-90

Weiss AA, Hewlett EL (1986) Virulence factors of *Bordetella pertussis.* Ann Rev Microbiol 40: 661-686

Welch DF (1991) Applications of Cellular Fatty Acid Analysis. Clin Microbiol Rev 4: 422-438

Westdijk J, Van Den IJssel J, Thalen M, Beuvery C, Jiskoot W (1997) Quantification of cell-associated and free antigens in *Bordetella pertussis* suspensions by antigen binding ELISA. J Immunoassay 18: 267-284

Wieten G, Dorresteijn RC, Habben-Janssen M, De Clercq G, Beuvery EC (1995) Batch control system vaccines (BCSV): a new man machine interface for bioreactors. Cytotechnology 18: 57–66.

Wilson DI, Agarwal M, Rippin DWT (1998) Experiences implementing the extended Kalman filter on an industrial batch reactor. Comput Chem Eng 22: 1653-1672

Wold S, Cheney J, Kettaneh N, McCready C (2006) The chemometric ananlysis of point and dynamic data in pharmaceutical and biotech production (PAT) – some objectives and approaches. Chemometrics Intelligent Lab Syst 84: 159-163

Wu L, Lange HC, Gulik WM, Heijnen JJ (2003) Determination of in vivo oxygen uptake and carbon dioxide evolution rates from off-gas measurements under highly dynamic conditions. Biotechnol Bioeng 81: 448-458

Zlateva P (1997) Sliding-mode control of fermentation processes. Bioprocess Eng 16: 383-387

SUMMARY

In traditional biopharmaceutical production quality of the product is currently tested at the end of the production process only. In Europe, approximately 1.5 million laboratory animals are being used every year for routine quality control of vaccines. Manufacturers, regulatory authorities, and ethical commissions wish to reduce animal use, because the high costs, imprecision of the tests, and ethical concerns.

Bioprocesses are characterised by natural variability in the raw materials, initial conditions, human intervention, and varying properties of the micro-organism. By setting fixed protocols the current "Good Manufacturing Practice" (GMP) regulations do not encompass the reduction of these deviations. This leads to products that have a variable quality, and are even sometimes out of specifications. Recently the Food and Drug Administration (FDA) released guidelines on process analytical technology (PAT), a framework for innovative pharmaceutical development, manufacturing, and quality assurance, in which they stress the need for reducing variability in the end-products and aim to improve manufacturing processes by introducing timely measurements and control of quality variables.

This thesis intended to be a step forward to real-time feedback control for PAT by modelling, monitoring, and controlling biomass and product formation towards the desired specifications, enhanced batch-to-batch consistency, and ultimately release on the basis of online monitored and controlled key variables. As an example the cultivation of *Bordetella pertussis* is chosen for the production of bulk whole cell whooping cough vaccine.

This thesis addresses five research questions. The answers are obtained by combining different techniques from a multidisciplinary field of systems and control theory, biology, and chemometrics.

How to improve monitoring of bioprocesses?

The specific growth rate and biomass formation are important variables that we would like to control during the cultivation of bacteria for vaccine production. This is particularly important when the produced bacteria form the basis of a whole cell vaccine like whooping cough. Biomass growth, however, can not be measured online, because there are no suitable sensors. In this thesis, two advanced techniques have been developed and evaluated for the whooping cough application to enhance bioprocess monitoring:

- *Software sensor*. The software sensor estimates biomass growth on the basis of a generic model and standard measurements of oxygen consumption. This software sensor accurately observes biomass growth over the whole range from low to high biomass concentrations during different types of cultivation (batch, fed-batch, and continuous cultivations). Offline biomass measurements can be incorporated to improve the accuracy of the biomass estimation, to safeguard the estimation, and to observe the effect of biomass growth on the oxygen transfer coefficient. This developed software sensor offers potential for high cell density cultivations, in which the oxygen transfer coefficient is drastically affected.
- *Near infrared spectroscopy*. The near infrared sensor generates spectra at a high sampling rate during the cultivation that contain wavelengths between 833 and 2500 nm. The evolution of these spectra contains information on the evolution of the cultivation process of the bacteria that are the basis of the final vaccine. The interpretation of these spectra is challenging due to the large number of wavelengths and due to the overlapping absorbance bands of specific components. In this thesis an advanced method is proposed that allows automatic selection of wavelengths that contain relevant information on the biomass concentration. The automatic wavelengths selection outperforms the traditional procedures for near infrared monitoring, thus enhancing the accuracy. A drawback of near infrared monitoring is the sensitivity for temporary disturbances in the process conditions like pH and dissolved oxygen. In order to enclose these in the calibration, the training dataset should encompass this kind of variability.

The option to use both sensors is attractive in practice to safeguard against sensor failure. Considering the current limitations of near infrared monitoring and the challenges to be solved, the software sensor is the preferred choice for monitoring and feedback control of biomass growth.

How can the cultivation process be controlled at a desired level?

Control of bioprocesses is a difficult task due to the time-varying characteristics of cultivation processes, from low to high biomass concentrations and from low to high oxygen and substrate consumption. The controller actions need to be adjusted to these changes to successfully control the process. Hence the challenge was to control the specific growth rate at a constant set-point by adding substrates in such a way that the feed rate is low at the

start and high at the end of the cultivation to cope with the increasing demands for substrates.

By combining a dual substrate model for biomass growth with specifications on the controller behaviour (a reference model), a controller is derived that adapts its settings in response to time-varying conditions. Consequently it is possible to extend the standard batch cultivation with a controlled fed-batch phase, what yields an up to six times higher biomass concentration and a constant and high antigen activity of the bulk vaccine, as is shown in experiments The controller, therefore, contributes to enhanced batch-to-batch consistency and quality.

How to automate controller tuning?

The challenge was to upgrade performance of poorly acting controllers automatically thereby eliminating the manual tuning effort. Automating controller tuning may facilitate non-professionals to work with advanced controllers. For this purpose, three automatic tuning methods are tested in simulations. The best two methods are qualified by laboratory experiments for *B. pertussis*. The methods do not require identification procedures, which are often required to automate tuning, thus avoiding the need for process perturbation. The settings of the controller are automatically adapted such that controller performance improves up to ten times for a process that was not or badly tuned in advance. We expect that a cultivation controlled at the desired specific growth rate will result in smaller variations in end quality (vaccine titer) and thus yield a better product (vaccine).

What set-points are needed to obtain the best-quality vaccine?

The quality of whole cell whooping cough vaccine is based on the presence of virulence factors, outer-membrane proteins that are important for inducing a protective immune response. To obtain an indication how the formation of these virulence factors depend on biomass growth a series of experiments has been done in which the set-point for specific growth rate was varied. While modelling of the synthesis of complex products and processes is uncommon and difficult, we attempted to model the expression of three virulence factors by enhancing an existing model for biomass formation. The data and model suggest that vaccine quality is not deteriorating with increasing growth rate over the investigated range, and thus a high growth rate can be selected to achieve a high productivity.

The final choice of specific growth rate to produce the best quality vaccine depends on multiple objectives like: maximize virulence factors per cell and or per volume to induce protection, minimize lipopolysaccharide to minimize side effects, maximize biomass in minimum time to safe time and costs, etcetera. The model for formation of virulence factors presented here is a very first step towards more insight. It must be further scrutinized by more experimentation, modification and adjustment on the basis of additional experimental evidence. However, as has been demonstrated in this thesis, once this has been done, the availability of such a model opens perspectives to design feed strategies during the cultivation and to optimize the quality of the vaccine using an objective function to weight the different quality aspects.

Can the developed monitoring & control tools be applied on production-scale?

Monitoring and control systems in biotechnology are usually designed and tested in laboratory-scale experiments. Application on production-scale is limited in literature and also the preceding aspects (1-4) in this thesis have been done at small-scale. To make the applicability to production-scale plausible, two issues encountered in the implementation of the monitoring and control system to larger scale have been worked out. The first issue is the adjusted calculation for oxygen consumption on large-scale, the starting point for online monitoring and control. This issue needs special attention, because the headspace and tubing cause deviations from the real oxygen consumption. The standard method, which does not take the headspace and tubing into account, gives an erroneous result which may go up to 30% of the value. The new method, which corrects for the dynamic effects, gives correct results. The second issue is the actual application of the observer and controller for biomass growth to pilot-scale cultivation. Enhancing the traditional batch cultivation with a controlled fed-batch phase enables an at least doubled concentration of bulk vaccine on pilot-scale compared to the standard batch production process. So a more favourable upstream production method for vaccines.

SAMENVATTING

In traditionele biopharmaceutische productie wordt productkwaliteit tot nu toe alleen getest in het eindproduct. Jaarlijks worden voor een aantal testen ongeveer 1,5 miljoen proefdieren gebruikt ten einde de veiligheid en werkzaamheid vast te stellen. Veel mensen en instanties zien graag een vermindering in het gebruik van dierproeven vanwege de hoge kosten, onnauwkeurigheid van de testen en ethische bezwaren.

Bioprocessen worden gekenmerkt door natuurlijke variatie in de grondstoffen, initiële condities, menselijke interventie en variërende eigenschappen van het micro-organisme. De huidige "Good Manufacturing Practice" (GMP) voorschriften proberen deze variatie binnen de perken te houden door vaste protocollen te gebruiken, maar anticiperen aldus niet op mogelijkheden de effecten van de afwijkingen te verminderen of te elimineren. Dit leidt in de praktijk tot producten met een aanzienlijke variatie in kwaliteit en producten die mogelijk niet binnen de specificaties vallen. Onlangs heeft de Food and Drug Administration (FDA) richtlijnen gepubliceerd, een raamwerk voor innovatieve farmaceutische ontwikkeling, productie en kwaliteitswaarborging, dat algemeen wordt aangeduid met PAT (Proces Analytische Technologie). Hierin beklemtonen zij de behoefte om variatie in de eindproducten te verminderen en pleiten zij voor verbeterde productieprocessen door het introduceren van vroegtijdige metingen en controle van kwaliteitsvariabelen.

Dit proefschrift beoogt een stap voorwaarts te zijn in de richting van online procesbeheersing via modelleren, monitoren en controleren van biomassa- en productvorming, teneinde de gewenste kwaliteitsspecificaties en verbeterde consistentie te bereiken en uiteindelijk vrijgave op basis van sleutelvariabelen die tijdens het proces waargenomen en geregeld worden. Als voorbeeld is de cultivatie van *Bordetella pertussis* gekozen voor de productie van bulk cellulair kinkhoest vaccin.

Dit proefschrift gaat in op vijf onderzoekvragen. De antwoorden zijn verkregen door verschillende technieken te combineren uit een multidisciplinair gebied van systeem- en regeltheorie, biologie, en chemometrics.

Hoe kan het monitoren van bioprocessen verbeterd worden?

De specifieke groeisnelheid en biomassavorming zijn belangrijke variabelen die we graag willen regelen tijdens het kweken van bacteriën voor vaccinproductie. Dit is vooral van belang wanneer de geproduceerde bacteriën de basis zijn van een cellulair vaccin zoals

kinkhoest. Biomassagroei kan echter niet online gemeten worden, omdat er geen geschikte sensoren zijn. In dit proefschrift zijn voor de kinkhoest applicatie twee technieken ontwikkeld en geëvalueerd om het monitoren van biomassagroei te verbeteren:

- *Software sensor*. De software sensor schat de biomassagroei op basis van een generiek model en standaard metingen van het zuurstofverbruik. Deze software sensor observeert de biomassagroei nauwkeurig over het hele bereik van lage naar hoge biomassaconcentratie tijdens verschillende types cultivatie (batch, fed-batch en continu-kweken). Om de schatting te verbeteren, als waarborg en om het effect van biomassagroei op de zuurstofoverdracht tijdens de kweek te weten te komen, kunnen offline biomassametingen toegevoegd worden. De ontwikkelde software sensor biedt perspectief voor kweken met hoge celdichtheden met variabele zuurstofoverdracht.
- *Nabij-infrarood spectroscopie*. Met behulp van een nabij-infrarood sensor kunnen tijdens de kweek met een hoge frequentie spectra gegenereerd worden die bestaan uit golflengten tussen 833 en 2500 nm. Het verloop van deze spectra bevat informatie over het verloop van het kweekproces van de bacteriën die uiteindelijk het vaccin vormen. De interpretatie van de spectra is uitdagend door het grote aantal golflengten en doordat er overlap is van de absorptiebanden van specifieke componenten. In dit proefschrift is een geavanceerde methode uit de literatuur voorgesteld om automatisch golflengten te selecteren die relevante informatie bevatten. Automatische selectie verbetert de standaard procedures voor nabij-infrarood monitoren door een verhoogde nauwkeurigheid. Een nadeel van nabij-infrarood monitoren is de gevoeligheid voor tijdelijke verstoringen in de procescondities pH en opgeloste zuurstof. Om deze in de callibratie te kunnen meenemen, moeten in de training dataset al dit soort variaties aanwezig zijn.

Het gebruik van twee verschillende sensoren kan complementaire informatie opleveren en daarmee de kweek waarborgen als één sensor faalt. Gezien de huidige beperkingen van nabij-infrarood monitoren en de nog niet aangegane uitdagingen om monitoren verder te verbeteren, is op dit moment voor monitoren en regelen van biomassagroei de software sensor beter geschikt.

Hoe kan het kweekproces geregeld worden op een gewenst niveau?

Het regelen van bioprocessen is een lastige taak, omdat de omstandigheden veranderen tijdens het kweekproces, van weinig naar veel biomassa en van lage naar een hoge zuurstof-

en substraatconsumptie, en de regelaaracties hierop dienen te worden aangepast om het proces succesvol te regelen. De uitdaging was dan ook de specifieke groeisnelheid van de bacteriën te regelen door het toevoegen van substraten zodanig dat de groeisnelheid constant is. Deze toevoersnelheid van substraten dient laag te zijn aan het begin en hoog aan het eind van de kweek om het hoofd bieden aan de toenemende biomassa en substraatbehoefte.

Door een model voor biomassagroei op twee substraten te combineren met specificaties over het regelaargedrag (een referentiemodel) is een regelaar afgeleid die in staat is zijn instellingen aan te passen aan de variabele omstandigheden. Daardoor is het mogelijk het standaard batch proces te verlengen met een geregelde fed-batch fase, wat een tot zes keer zo hoge opbrengst oplevert en een constante en hoge antigeen activiteit van het bulkvaccin, zoals experimenten hebben laten zien. De adaptieve regelaar draagt daarmee bij aan een verhoogde consistentie en kwaliteit.

Hoe kan een regelaar ontworpen worden zodat de instelling gedaan kan worden door non-professionals?

De uitdaging was om de prestatie van slechte regelaars of niet-ingestelde regelaars automatisch in te stellen en daarbij de inspanning van het handmatig instellen van regelaars te sparen. Hiertoe zijn drie methoden uitgetest in simulaties. De twee beste methoden zijn gevalideerd in labexperimenten met *B. pertussis*. De methoden vereisen geen identificatieprocedures, zoals vaak nodig bij automatisch instellen, en vermijden daarmee de noodzaak tot verstoring van het proces. De instellingen van de regelaar worden automatisch aangepast zodanig dat de regelprestaties tot wel tien maal verbeteren bij een vooraf slecht of niet ingesteld proces. Op basis van de online geëvalueerde regelaarprestaties kan beslist worden of er (extra) auto-tuning nodig is of dat het instellen klaar is en het proces met de huidige instellingen uitgevoerd kan worden in een volgende run. We verwachten dat een cultivatie die automatisch en goed geregeld is op het gewenste setpoint zal resulteren in kleinere variaties in vaccinkwaliteit en een beter product.

Welke setpoints leiden tot de beste vaccinkwaliteit?

De kwaliteit van cellulair kinkhoest vaccin is gebaseerd op de aanwezigheid van virulentiefactoren, eiwitten op het buitenmembraan van de *B. pertussis* bacterie, die belangrijk zijn voor het opwekken van een beschermende immuunrespons. Om een indicatie

te verkrijgen hoe de vorming van deze virulentiefactoren afhangt van de biomassagroei zijn een aantal experimenten uitgevoerd waarin het setpoint voor de specifieke groeisnelheid is gevarieerd. Hoewel het modelleren de vorming van complexe producten en processen zeldzaam en moeilijk is, hebben wij een poging gedaan de vorming van een drietal virulentiefactoren te modelleren door een bestaand model voor biomassavorming uit te breiden. De data en het model suggereren dat een vaccinkwaliteit niet verslechtert met een toenemende specifieke groeisnelheid in het onderzochte bereik; en dus dat een hoge specifieke groeisnelheid gekozen kan worden om hoge productiviteit te bereiken.

De uiteindelijke keuze van het (verloop van het) setpoint hangt af van verschillende doelen als: maximaliseren van virulentiefactoren per cel of per volume om een goede immuunrespons te induceren, minimaliseren lipopolysaccharide tegen bijwerkingen, maximaliseren van de biomassa om tijd en geld te besparen, etc. Het model voor vorming van virulentiefactoren kan perspectieven openen om de setpoints zo in te stellen en dat de kwaliteit van het vaccin geoptimaliseerd wordt door gebruik te maken van een doelfunctie die de verschillende kwaliteitsaspecten meet.

Kunnen de ontwikkelde monitor- & controletechnieken toegepast worden op productieschaal?

Over het algemeen worden de technieken om bioprocessen te monitoren en te beheersen ontworpen en uitgetest op laboratoriumschaal. De toepassing op productieschaal is gelimiteerd in de literatuur en ook de voorgaande aspecten (1-4) in dit proefschrift zijn uitgevoerd op kleine schaal. Om toepasbaarheid op productieschaal aannemelijk te maken, zijn twee kwesties onderzocht die zich voordoen bij de implementatie op grotere schaal van een monitor- en controlesysteem. De eerste kwestie is een aangepaste berekening van het zuurstofverbruik op grote schaal, het uitgangspunt voor online monitoren en controle. Deze kwestie krijgt speciale aandacht, omdat de headspace en andere apparatuur zorgt voor verstoringen in de meting van het echte zuurstofverbruik. De nieuwe methode reduceert de fout in de berekening van het zuurstofverbruik tot 30%. De tweede kwestie is het daadwerkelijk toepassen van het online monitoren en regelen van het productieproces voor kinkhoest op pilotschaal. Tijdens de experimenten is een verdubbelde concentratie van bulk vaccin op pilotschaal bereikt vergeleken met het traditionele proces op productieschaal, wat leidt tot een gunstigere methode voor vaccinproductie.

CURRICULUM VITAE

Zita Soons was born on October 15th, 1979 in Maastricht, The Netherlands. After completion of high school at “Jeanne d’Arc College” in Maastricht, she studied bioprocess engineering at Wageningen University from 1998-2003. Her first MSc thesis concerned “Observer based growth rate control for vaccines” and the second thesis “Modelling growth and product formation in solid-state fermentation”. From 2003-2004 she was a research assistant at Quest (ICI) in Naarden, the Netherlands. In 2004-2008 she was working as a PhD-student in a project called “PaRel” or “Parametric Release”. Her work was a collaboration between the Systems and Control Group of Wageningen University and the Process Development group at the Netherlands Vaccine Institute. The results of the research are presented in this thesis.

LIST OF PUBLICATIONS

Soons ZITA, Shi J, Stigter HD, Van Der Pol LA, Van Straten G, Van Boxtel AJB (2007) Observer design and tuning for biomass growth and k_La using online and offline measurements. Journal of Process Control, in press

Soons ZITA, JA Voogt, Van Straten G, Van Boxtel AJB (2006) Constant specific growth rate in fed-batch cultivation of Bordetella pertussis using adaptive control. Journal of Biotechnology 125: 252-268

Soons, ZITA, Van Straten G, Van Der Pol LA, Van Boxtel AJB (2007) Online automatic tuning and control for fed-batch cultivation. Bioprocess and Biosystems Engineering, online first

Soons ZITA, Shi J, Van Der Pol LA, Van Straten G, Van Boxtel AJB (2007) Biomass growth and k_La estimation using online and offline measurements. Preprints Vol. I. Computer Applications in Biotechnology, Cancun, Mexico, 81-86

Soons ZITA, Van Straten G, Van Boxtel AJB (2006a) Automatic tuning and adaptation for specific growth rate control of fed-batch cultivation. Proceedings Chemical Process Control 7, Lake Louise, Canada

Soons ZITA, Streefland M, Van Straten G, Van Boxtel AJB, Towards PAT monitoring and control: near infrared and software sensor. Submitted to Chemometrics and Intelligent Laboratory Systems.

Soons ZITA, Van Den IJssel J, Van der Pol LA, Van Boxtel AJB, Scaling-up vaccine production: implementation aspects of a biomass growth observer and controller. Submitted to Bioprocess and Biosystems Engineering

Soons ZITA, Larmonie NSD, Van De Waterbeemd B, Streefland M, Beuvery EC, Martens DE, Van Straten G, Van Boxtel AJB. Modelling the cultivation of *Bordetella pertussis*: the effect of specific growth rate on expression of virulence factors on gene and protein level. Submitted to Journal of Biotechnology.

TRAINING ACTIVITIES

Discipline specific activities	**Year**	**Credits (ECTS)**
Systems and control theory of non-linear systems	2004	4.2
Multivariate analysis	2005	1.4
Model predictive control	2005	2.8
Practical training performing Bioreactor cultivations	2004	2.8
Symposium enhanced whooping cough vaccine	2005	0.3
Seminar A Turn-Key Solutionfor PAT, Copenhagen	2005	0.3
Bioprocess engineering course (Croatia)	2006	2.0
Benelux meeting	2005, 2006	3.0
Computer Applications in Biotechnology, Mexico	2007	1.5
Chemical Process Control 7, Canada	2006	1.5
Control theory for systems biology	2007	0.9
Netherlands Biotechnology congress	2006	0.6
Poster presentation	2004	0.6
DISC day Delft		
General courses		
Techniques for writing and presenting a scientific paper	2005	1.4
Organizing and supervising thesis work	2005	0.7
Personal efficacy	2006	0.6
“Veilige Microbiologische Technieken”	2006	0.6
Good Manufacturing Practice	2005	0.3
Optional		
Research proposal	2004	6.0
Beurs Industrial processing	2004	0.3
The Instrument, Upscaling in food, farma & biotech	2004	0.3
Biomedical Symposium, Wageningen	2006	0.3
Food quality management	2004	6
Total		**38.4**

MONITORING + CONTROL

ISBN 978-90-8504-933-3