

BIOMASS PERFORMANCE

Monitoring and Control in
Bio-pharmaceutical
Production



Ronald Neeleman

Biomass Performance

Monitoring and Control in Bio-pharmaceutical Production

Ronald Neeleman

Promotor: prof. dr. ir. G. van Straten
hoogleraar in de meet-, regel- en systeemtechniek

Co-promotoren: dr. ir. A.J.B. van Boxtel
universitair docent, departement Agrotechnologie en Voeding

dr. E.C. Beuvery
hoofd laboratorium product- en procesontwikkeling, RIVM, Bilthoven

Samenstelling promotiecommissie:

prof. dr. J. van Impe (Katholieke Universiteit Leuven, België)
prof. dr. ir. A.C.P.M. Backx (Technische Universiteit Eindhoven)
prof. dr. ir. R.M. Boom (Wageningen Universiteit)
dr. ir. R.H. Wijffels (Wageningen Universiteit)
dr. ir. P. van der Meijden (AKZO-Nobel Diosynth, Oss)

Biomass Performance
Monitoring and Control in Bio-pharmaceutical Production

Ronald Neeleman

Proefschrift

ter verkrijging van de graad van doctor
op gezag van de rector magnificus
van Wageningen Universiteit,
prof. dr. ir. L. Speelman
in het openbaar te verdedigen
op maandag 25 november 2002
des namiddags te vier uur in de Aula

Neeleman, R.

Biomass Performance: Monitoring and Control in Bio-pharmaceutical Production

Thesis Wageningen University.-With ref.-With summary in Dutch

ISBN 90-5808-733-6

PREFACE

It has been more than four years since I decided to become a Ph.D. student in the Systems and Control Group of Wageningen University. In retrospect, I think that by far most of the ideas you will find in this thesis were thought of on the train or on my bike going home after a long fruitless day. In more than fifty percent of the cases, the ideas turned out to be useful, even the next day...

After four years, I have indeed managed to complete this book. Of course, I have not done it all alone and this should be the place to thank all those who helped me to make this thesis as thick as it is (although none of them were present on the train or on my bike). Without them this thesis would have been only half as thick. However, giving proper thanks to all those involved would let this preface take over half of the thesis, thus quadrupling its size.

Therefore, I will refrain from summing up a huge thank-you list. I will stick to only mentioning my promotor and co-promotors for their enthusiasm and encouragement. They gave me the motivation to complete this study. And all the students for their major contributions to the work involved in this thesis. Finally, none of this work would have been possible without the permission and support from the RIVM.

Ronald Neeleman

CONTENTS

1	BIOMASS PERFORMANCE	1
1.1	BIO-PHARMACEUTICALS AND THEIR PRODUCTION.....	2
	Production Process.....	2
	Regulations	3
	The Case of <i>Bordetella pertussis</i>	3
1.2	MODELLING WITH A VIEW TO MONITORING	4
	Why Should Everybody Model?	4
	Modelling Bioprocesses.....	4
1.3	SOFTWARE-SENSORS FOR BIOREACTOR MONITORING	5
	Need for Monitoring	5
	What Is a Software-Sensor?.....	6
1.4	APPLICATION OF MONITORING	6
	Controlled Fed-batch	6
	Prediction of Future Trajectory	7
	Product Release.....	8
1.5	THESIS OUTLINE.....	8
	Aim and Outline.....	8
	Monitoring	8
	Application of Monitoring.....	9
1.6	REFERENCES.....	9

2	PRINCIPLES AND DEFINITIONS OF MODELLING, ESTIMATION AND CONTROL	13
2.1	INTRODUCTION.....	14
2.2	MODELLING THE BIOREACTOR.....	14
	Modelling Biotechnological Processes.....	14
	General Model Structure.....	16
	Unstructured, Nonsegregated Modelling.....	20
2.3	OBSERVING THE BIOREACTOR.....	24
	Indirect Measurements and Correlations.....	25
	Software sensors.....	25
	State estimation.....	26
	Parameter Estimation.....	27
	Kalman Filtering.....	27
	Tuning of software sensors.....	30
2.4	CONTROL.....	31
	Feedback control.....	31
	Feed-Forward Control.....	31
	Feed-forward-Feedback Control.....	32
2.5	NOMENCLATURE.....	33
2.6	REFERENCES.....	35
3	RESPIRATION QUOTIENT: MEASUREMENT AND ESTIMATION DURING BATCH CULTIVATION IN BICARBONATE BUFFERED MEDIA	37
3.1	ABSTRACT.....	38
3.2	INTRODUCTION.....	38
3.3	MATERIALS AND METHODS.....	40
	Cell line and culture medium.....	40
	Bioreactor.....	40
	Analysis.....	40
3.4	GAS CONCENTRATIONS IN BATCH-WISE CELL CULTURES.....	41
	Oxygen Uptake Rate (OUR).....	41
	Carbon dioxide equilibrium in the gas phase.....	41
	Carbon dioxide equilibrium in the liquid phase.....	42

3.5	SOFTWARE SENSOR DESIGN.....	44
	Dynamic model.....	44
	The Kalman Filter algorithm	46
3.6	APPLICATION OF THE SOFTWARE SENSOR	48
	Validation of the software sensor	48
	Application to cell cultivation	51
	Robustness of the software sensor.....	52
3.7	CONCLUDING REMARKS.....	53
3.8	NOMENCLATURE	54
3.9	REFERENCE	55
4	ESTIMATION OF SPECIFIC GROWTH RATE FROM CELL DENSITY MEASUREMENTS	57
4.1	ABSTRACT.....	58
4.2	INTRODUCTION.....	58
4.3	MATERIALS AND METHODS	59
	Strains and culture media.....	59
	Bioreactor conditions.....	59
	Analysis	59
	Software.....	60
4.4	METHODS FOR GROWTH RATE ESTIMATION	60
	Traditional methods	60
	Recursive estimation.....	61
	Forward Extended Kalman Filter (EKF)	62
	Backward Extended Kalman Filter.....	63
	Smoothing estimation	64
4.5	RESULTS AND DISCUSSION.....	66
	Application to batch cultivation of <i>B. pertussis</i>	66
	Application to batch cultivation of <i>N. meningitidis</i>	67
	Application to fed-batch cultivation.....	68
4.6	CONCLUSIONS	70
4.7	NOMENCLATURE	71
4.8	APPENDIX.....	72

	Linearisation and discretisation of Eq. (4).....	72
	Forward EKF algorithm.....	72
	Backward EKF algorithm.....	72
	Smoothing algorithm.....	73
4.9	REFERENCES.....	73
5	DUAL-SUBSTRATE UTILISATION BY <i>BORDETELLA PERTUSSIS</i>.....	75
5.1	ABSTRACT.....	76
5.2	INTRODUCTION.....	76
5.3	MATERIALS AND METHODS.....	77
	Strains and Culture Media.....	77
	Bioreactor Conditions.....	77
	Analysis.....	77
5.4	RESULTS.....	78
	Dual-Substrate Limitation.....	78
	Growth Rate.....	79
	Substrate Utilisation.....	80
	Step-wise Experimentation for Parameter Estimation.....	81
	Model Validation.....	82
5.5	DISCUSSION.....	83
5.6	NOMENCLATURE.....	83
5.7	APPENDIX.....	84
	Lactate Yield.....	84
	Glutamate Yield.....	84
5.8	REFERENCES.....	85
6	SEQUENTIAL ESTIMATION OF BIOMASS AND SPECIFIC GROWTH RATE.....	87
6.1	ABSTRACT.....	88
6.2	INTRODUCTION.....	88
6.3	MATERIALS AND METHODS.....	89
	Strain, Medium and Bioreactor Conditions.....	89
	Analysis.....	90
	Hard- and Software Set-up.....	90

6.4	SEQUENTIAL ESTIMATION.....	90
6.5	ESTIMATION OF THE SPECIFIC GROWTH RATE.....	92
	Model.....	92
	Observer.....	93
	Stability Analysis.....	94
6.6	ESTIMATION OF BIOMASS CONCENTRATION.....	95
	Model.....	95
	Observer.....	95
	Stability Analysis.....	96
6.7	RESULTS.....	98
	Sequential Estimation.....	98
	Tuning.....	98
	Experimental Validation.....	100
6.8	CONCLUSIONS.....	102
6.9	NOMENCLATURE.....	103
6.10	REFERENCES.....	103
7	ONLINE HARVEST PREDICTION FOR THE PRODUCTION OF BIOLOGICALS.....	107
7.1	ABSTRACT.....	108
7.2	INTRODUCTION.....	108
7.3	MATERIALS AND METHODS.....	109
	Strain and Culture Medium.....	109
	Bioreactor Operation.....	109
	Analysis.....	109
	Software.....	110
7.4	DUAL-SUBSTRATE KINETIC MODEL OF <i>B. PERTUSSIS</i>	110
7.5	OBSERVER FOR SPECIFIC GROWTH RATE.....	112
7.6	OBSERVER FOR BIOMASS.....	114
7.7	OBSERVER BASED CALCULATION OF SUBSTRATES.....	116
7.8	OBSERVER BASED PREDICTION OF HARVEST TIME.....	117
7.9	CONCLUDING REMARKS.....	118
7.10	NOMENCLATURE.....	119

7.11	REFERENCES.....	120
8	DUAL-SUBSTRATE FEEDBACK CONTROL OF THE SPECIFIC GROWTH RATE.....	121
8.1	ABSTRACT.....	122
8.2	INTRODUCTION.....	122
8.3	MATERIALS AND METHODS.....	123
	Strain, Medium and Bioreactor Conditions.....	123
	Analysis.....	124
	Hard- and Software Set-up.....	124
8.4	SEQUENTIAL ESTIMATION.....	125
	Estimation of Specific Growth Rate.....	125
	Estimation of Biomass Concentration.....	127
	Tuning.....	129
8.5	FEED-FORWARD CONTROLLER.....	129
	Dual-Substrate Model.....	129
	Calculations of Feed-forward Controller.....	130
8.6	FEEDBACK CONTROLLER.....	132
	Feedback Control.....	133
8.7	RESULTS.....	134
	Feed-Forward Control Only.....	134
	Feed-Forward-Feedback Controller.....	135
8.8	CONCLUSIONS.....	137
8.9	NOMENCLATURE.....	138
8.10	REFERENCES.....	139
9	ADVANCED PROCESS MONITORING: THE ULTIMATE ALTERNATIVE.....	143
9.1	ABSTRACT.....	144
9.2	INTRODUCTION.....	144
9.3	PARAMETRIC RELEASE.....	146
	Definition.....	146
	Terminally sterilised biologics.....	146
	In-process controls versus finished product controls.....	147

	<i>B. pertussis</i> cultivation.....	147
9.4	ADVANCED PROCESS MONITORING	149
	Modelling.....	149
	Monitoring	150
	Robustness	151
	Predictability.....	153
9.5	DISCUSSION.....	153
	Critical process parameters.....	153
	Prerequisites for parametric release.....	153
	Parametric release: the ultimate alternative.....	154
9.6	REFERENCES.....	155
10	CONCLUSIONS	157
10.1	INTRODUCTION.....	158
10.2	MONITORING BIOMASS PERFORMANCE	158
10.3	KNOWLEDGE OF BIOMASS PERFORMANCE	160
10.4	PREDICTING BIOMASS PERFORMANCE.....	160
10.5	MANIPULATING BIOMASS PERFORMANCE	161
10.6	ASSURANCE OF BIOMASS PERFORMANCE.....	162
10.7	THE MODULAR APPROACH	162
10.8	PROSPECTS	163
	Managing Biomass Performance.....	163
	In-depth Biomass Performance.....	164
	Beyond Biomass Performance.....	165
	From Biomass Performance to Company Performance	165
10.9	REFERENCES.....	165
	SUMMARY	167
	SAMENVATTING	171
	BIBLIOGRAPHY.....	175
	CURRICULUM VITAE	177

1 BIOMASS PERFORMANCE

1.1 Bio-Pharmaceuticals and their Production

The primary concern in the pharmaceutical industry is not the reduction of manufacturing cost (although that is still a very desirable goal), but the production of a product of consistently high quality in amounts to satisfy the medical needs of the population, while insuring safety and efficacy. This has resulted in ‘process monitoring’ becoming an integral part of process operation. The aim of process monitoring is to achieve early indicators of changes in process operation that may result in low product yield or low quality, or the observation of a disturbance that may result in the batch being terminated early as the process is out-of-specification.

Process monitoring as it is generally applied in the pharmaceutical industry comprises of monitoring the physical and chemical state of the process (e.g. pH, temperature, dissolved oxygen, etc.). However, since the quality of the product and the consistency of the process are determined by the physiological state of the living cells inside the bioreactor [1], the ‘performance of biomass’ should be monitored. Therefore this thesis will discuss the use of monitoring methodologies for the acquisition of information directly related to process performance and especially to *biomass performance*.

Production Process

The production process of most bio-pharmaceuticals can be divided into two major steps; upstream processing where the product is made and downstream processing where the product is purified and treated for final presentation. Upstream processing involves the preparation of media, cultivation of the seed lot on inoculum scale, and the final cultivation of biomass in the bioreactor. During upstream processing the main goal is to maximise product yield, which can be biomass, proteins, DNA or a combination. When the final cultivation is finished the process stops, the biological material is *harvested* and downstream processing starts. Downstream processing is usually more versatile and may involve purification, inactivation, detoxification, blending, adjuvation, and filling. In between and during the various steps of bio-pharmaceutical production samples - called in-process controls - are taken, measured and compared to specifications.

This thesis is mainly focused on the cultivation step of upstream processing.

Regulations

Regulatory approval of a pharmaceutical product is for the product *and* process together. Thus, process modifications may require new clinical trials to test the safety of the resulting product. Since clinical trials are very expensive, process improvements are made under a limited set of circumstances. Even during clinical trials it is difficult to make major process modifications.

Pharmaceuticals sold on the market or used in clinical trials must come from facilities that are certified as GMP. GMP stands for Good Manufacturing Practice and concerns the actual manufacturing facility design and layout, the equipment and procedures, training of production personnel, control of process inputs (e.g., starting materials and cultures), processing, and handling of product. The GMP guidelines stress the need for documented procedures to validate performance. “Process validation is establishing documented evidence which provides a high degree of assurance that a specific process will consistently produce a product meeting its predetermined specifications and quality characteristics” and “There shall be written procedures for production and process-control to assure that products have the identity, strength, quality, and purity they purport or are represented to possess.”

The key point is that process modifications cannot be made without considering their considerable regulatory impact.

The Case of *Bordetella pertussis*

All theory discussed in this thesis is tested experimentally with various organisms; *Bordetella pertussis*, *Neisseria meningitidis*, insect cells, and hybridoma's. However, most research is performed using *B. pertussis*, which is the causative organism of whooping cough, a disease of the respiratory tract that may be life threatening in infants and young children. The disease can effectively be controlled by immunization with a vaccine consisting of inactivated whole cells. Whole cell vaccines comprising of heat-killed virulent *B. pertussis* cells were introduced in most European countries in the 1950s. In the Netherlands this vaccine is administered together with diphtheria, tetanus and polio vaccines to infants aged 2, 3, 4, and 11 months in four subsequent intramuscular injections.

Because whole cell *B. pertussis* vaccines have been associated with adverse reactions, several acellular vaccines have been developed to replace the whole cell vaccine. These vaccines are comprised of purified antigens from *B. pertussis* such as toxin, filamentous hemagglutinin, pertactin, and fimbriae [2]. The concentration of these virulence-factors in one dose of acellular vaccine is generally higher than the corresponding concentrations in one dose of whole-cell vaccine. Using data from literature it is estimated that the number of doses acellular vaccine obtained per bioreactor volume is at best fivefold lower than for whole-cell vaccines [3, 4]. So, to make the acellular vaccine profitable, the volumetric productivity of *B. pertussis* has to increase.

1.2 Modelling With a View to Monitoring

Why Should Everybody Model?

The ability to describe an object, material, process or action in mathematical terms is central to our paradigms for analysis and control. Detailed mathematical models are used by engineers in academia and industry to gain competitive advantage through applications of model-based process design, monitoring, control and optimisation. Thus, building high quality and validated mechanistic models is a key activity in process engineering.

When developing models that use *a priori* knowledge of physical, chemical or biological *laws* to propose (one or more) possible models, these *laws* dictate the model structure and invariably contain adjustable parameters. Typically, it is desired to establish the most appropriate model structure and the best values for the parameters, thus providing the best fit to experimental data.

Modelling Bioprocesses

The specific growth rate is one of the most important parameters in biotechnological cultivation processes, and the relationships between the rate of growth, substrate consumption and product formation are crucial for monitoring, controlling, and optimising these processes. Therefore, it is of importance to be able to properly model the specific growth rate as a function of the states, e.g. biomass and substrates [5].

A common and classic approach for modelling bioprocesses is the assumption that only one substrate is limiting. In the case of more than one limiting

substrate, like with *B. pertussis*, non-interactive models are used that assumes growth of micro-organisms to be limited by the most limiting substrate [6], or an interactive limitation model is used where both substrates are simultaneously essential and limiting. Although *B. pertussis* is clearly dual-substrate limited, neither interactive nor non-interactive models seem appropriate [7].

Biological models usually involve two kinds of parameters: the pseudo-stoichiometric coefficients that rely on the underlying metabolic network and the kinetic coefficients that indicate the relative speed of that network. The difference of nature between these two kinds of parameters should be taken into account, thus stoichiometry and kinetics should not be treated as a whole in the identification exercise. A problem with dual-substrate limited growth models is the estimation of these parameters, such as yields and maximal specific growth rate. Contrary to single-substrate models it is not possible to determine the yields independently from the growth rate, since the origin of the formed biomass can not be reduced to one of the substrates. In this thesis it will be shown how to implement a two-step procedure for identifying separately the stoichiometry and the kinetics for the case of *B. pertussis* [7].

1.3 Software-Sensors for Bioreactor Monitoring

Need for Monitoring

The performance of the cultivation process depends heavily on the environment suitable for efficient production of both biomass and product. Local controllers maintain such an environment and in turn depend heavily on the accurate and regular measurements. Typical online measurements include temperature, dissolved oxygen, pH, agitation, and foam, (all used for control), and off-gas analysis (used to calculate respiration data). With the increasing move to high cell density cultivation requiring fed-batch growth strategy to avoid substrate inhibition, as well as to provide an increasing measure of control, the ability to access more crucial data online, such as biomass concentration and growth rate, is becoming increasingly desirable.

What Is a Software-Sensor?

Monitoring bioprocesses by on-line (in-situ) techniques is highly desirable since it has the potential to produce significant improvements in process control. However, an issue with pharmaceutical batch and fed-batch processes is the relatively low level of sophistication of sensor-based instrumentation. Whilst direct on-line measurement of the states may not be possible, the influence of their variation can be observed in available online measurements. It is therefore in certain instances possible to obtain an online inference of the process states – such an approach is termed a software-sensor [8].

In other words, a software-sensor is a software algorithm giving an online estimation for state variables in a cultivation, e.g. biomass, substrate and/or product, whose analyses are normally time consuming, labour intensive and costly. A software-sensor calculates this prediction from the available online measurements using a model and proper mathematical inference. Recent developments in the area of advanced bioprocess control have demonstrated the applicability of inferential estimation of bioprocess state variables from secondary variables monitored frequently and regularly online [9-16].

1.4 Application of Monitoring

The information gathered by one or more software-sensors can be used to control the process, improve the accuracy and reproducibility of the feeding strategy and reduce the labour-intensive offline analysis required. Moreover, the ability of online monitoring the physiological conditions of biomass, the *biomass performance*, potentially yields a better measure of quality than current animal tests.

Controlled Fed-batch

Batch processing is synonymous with flexible manufacturing and has advantage over continuous production in that the system can be more easily modified to take account of changes in product specifications resulting from changes in or evolving customer requirements. However, apart from varying inoculum conditions and maintaining environmental conditions there is no effective way to direct a batch process into a certain region of interest. Thus, using a fed-batch process where substrates are fed to the bioreactor increases the ability to control the process.

Furthermore, fed-batch operation has been found particularly effective for processes in which effects such as substrate inhibition, catabolite repression, and product inhibition are important. These phenomena lead to growth rate expressions that exhibit a maximum with respect to one or more substrate concentrations. A living cell possesses a complex internal control system involving phenomena such as activation, inhibition, induction, and repression, which lead to such non-monotonic growth rate expressions. With optimal control theory it can be shown that whenever the growth rate expression is a non-monotonic function of a substrate, a fed-batch process is likely to outperform both a continuous process and a batch process. The problem is then the determination of the *best* feed-rate of substrate as a function of time, where the meaning of *best* varies from problem to problem [17-27].

Fed-batch bioreactors may be operated in a variety of ways by regulating the feed rate in a predetermined manner (feed-forward control) or by using a feedback control. The most commonly used are constantly fed, exponentially fed, and extended fed-batch. Where in extended fed-batch cultivation, the feed rate is regulated to maintain the substrate concentration constant until the bioreactor is full. However, the application of extended fed-batch is hindered by the lack of online sensors for substrate. As a result a model-based controller should be developed, which uses the information provided by software-sensors.

Prediction of Future Trajectory

While planning and scheduling is a common industrial problem, the manufacture of pharmaceuticals by biotechnological processes has certain unique planning and scheduling aspects. Since batch sizes are often fixed due to process specifications and its corresponding validation requirements, the cultivation time as important variable is left. Logically one should not harvest too early in the process to avoid poor yields. And since the quality of the product usually deteriorates within several hours after substrate depletion, it is hence of great importance to harvest in time.

For the prediction of the moment to harvest it is important to determine what has happened, is happening, and probably will happen in the bioreactor, based on all available knowledge about the bioreactor and the process. Using a proper model in combination with advanced monitoring techniques like

software-sensors it becomes possible to reconstruct what happened in the past and what the current state of the process is. From that the future trajectory of a cultivation can be predicted, and thus the moment of harvest can be predicted. In such case a simulation is started on operator request or on a regular schedule, with the current reconstructed state to calculate the moment of harvest.

Product Release

Besides using process monitoring for control purposes, it can be applied to indicate process consistency and consequently to provide assurance of product quality. When the physical state of the process is tightly controlled and the physiological state, the *biomass performance*, can thoroughly be evaluated, the opportunity is offered for batch release and regulatory approval based on such data, without conducting a test on the final product. This technique is called parametric release and it has the potency to be the ultimate alternative for the use of animal tests.

1.5 Thesis Outline

Aim and Outline

The aim of this research is to examine the performance and reliability of software-sensors for online monitoring the production of biologicals and to inquire into the applicability of these sensors. Besides covering practical applications for prediction, control and consistency checking, the perspectives of these techniques in a pharmaceutical environment are inquired as well.

Monitoring

Chapter 2 gives some general background in biotechnology, process engineering and systems theory for modelling and estimation as it is applied in this thesis. This single chapter is not intended to cover these subjects as a whole but as a cursory introduction to the theory used in the following chapters. Chapter 3 describes an algorithm that estimates the dissolved carbon dioxide concentration during a process. With this observed environmental variable the respiration quotient of an organism could be determined. Thus, it gives information on the pathways that the organism is using. In Chapter 4, a growth rate observer is designed which proves to be a good technique for off-

line calculation of the specific growth rate from available cell concentration measurements. Both these chapters (3 and 4) present techniques that are independent of the organism used and give valuable information about the state of the organism. As a result the growth rate estimation algorithm was used as a tool to understand the behaviour of *B. pertussis* in such a way that a model could be developed in Chapter 5. Chapter 6 describes the development of a generic on-line sequential observer for the specific growth rate and biomass concentration based on common oxygen measurements. This technique is again independent of the organism used. Besides an improvement in the medium composition for *B. pertussis*, the techniques depicted in chapters 5 and 6 showed to be valuable when applied for on-line monitoring process performance.

Application of Monitoring

Chapter 7 combines the model of *B. pertussis* (Chapter 5) with the sequential estimator (Chapter 6) to reconstruct the complete state of the process, e.g. biomass and substrate concentrations, online. Then it uses these reconstructed and estimated values to simulate the future trajectory of the cultivation in order to achieve a prediction of the right moment of harvesting. The sequential estimator and model are then used to construct a feed-forward-feedback controller of the specific growth rate in Chapter 8. Chapter 9 describes its application for indicating process robustness and consistency in a pharmaceutical environment in order to considerably simplify validation and ultimately reduce the number of animal tests used for the release of product. Finally, Chapter 10 is a combination of concluding and reflecting the achievements in this thesis and some future perspectives.

1.6 References

1. Rani, K.Y. and V.S.R. Rao (1999). Control of fermenters - a review. *Bioprocess Engineering* 21: 77-88.
2. Willems, R.J.L. and F.R. Mooi (1997). From whole cell to acellular whooping cough vaccines. *Reviews in Medical Microbiology* 7: 0.
3. Ibsen, P.H., J.W. Petersen and I. Heron (1993). Quantification of pertussis toxin, filamentous haemagglutinin, 69 kDa outer-membrane

- protein, agglutinogens 2 and 3 and lipopolysaccharide in the Danish whole-cell pertussis vaccine. *Vaccine* 11: 318-322.
4. Thalen, M., J. Van De IJssel, W. Jiskoot, B. Zomer, P. Roholl, C. De Gooijer, C. Beuvery and J. Tramper (1999). Rational medium design for *Bordetella pertussis*: basic metabolism. *Journal of Biotechnology* 75(2-3): 147-159.
 5. Neeleman, R. and A.J.B. Van Boxtel (2001). Estimation of specific growth rate from cell density measurements. *Bioprocess Engineering* 24(3): 179-185.
 6. Ryder, D.N. and C.G. Synclair (1972). Model for the growth of aerobic microorganisms under oxygen limiting conditions. *Biotechnology and Bioengineering* 14: 787-798.
 7. Neeleman, R., M. Joerink and A.J.B. Van Boxtel (2001). Dual-Substrate Utilisation By *Bordetella pertussis*. *Applied Microbiology and Biotechnology* 57: 489-493.
 8. Dorresteyjn, R.C. (1997). Software sensors as a tool for optimization of animal-cell cultures. PhD Thesis, Process Engineering, Wageningen University, The Netherlands.
 9. Chéryy, A. (1997). Software sensors in bioprocess engineering. *Journal of Biotechnology* 52: 193-199.
 10. Claes, J.E. and J.F. Van Impe (2000). Combining yield coefficients and exit-gas analysis for monitoring of the baker's yeast fed-batch fermentation. *Bioprocess Engineering* 22: 195-200.
 11. Estler, M.U. (1995). Recursive on-line estimation of the specific growth rate from off-gas analysis for the adaptive control of fed-batch processes. *Bioprocess Engineering* 12: 205-207.
 12. Lombardi, M., K. Fiaty and P. Laurent (1999). Implementation of observer for on-line estimation of concentration in continuous-stirred membrane bioreactor: Application to the fermentation of lactose. *Chemical Engineering Science* 54: 2689-2696.
 13. Náhlik, J. and Z. Burianec (1988). On-line parameter and state estimation of continuous cultivation by extended Kalman filter. *Applied Microbiology and Biotechnology* 28: 128-134.

14. San, K.-Y. and G. Stephanopoulos (1984). Studies on On-Line Bioreactor Identification. II. Numerical and Experimental Results. *Biotechnology and Bioengineering* 26: 1189-1197.
15. Stephanopoulos, G. and K.Y. San (1984). Studies on On-Line Bioreactor Identification. I. Theory. *Biotechnology and Bioengineering* 26: 1176-1188.
16. Tatiraju, S., M. Soroush and R. Mutharasan (1999). Multi-rate nonlinear state and parameter estimation in a bioreactor. *Biotechnology and Bioengineering* 63: 22-32.
17. Alves, T.L.M., A.C. Costa, A.W.S. Henriques and E.L. Lima (1998). Adaptive optimal control of fed-batch alcoholic fermentation. *applied Biochemistry and Biotechnology* 70-72: 463-478.
18. Berber, R., C. Pertev and M. Türker (1999). Optimization of feeding profile for baker's yeast production by dynamic programming. *Bioprocess Engineering* 20: 263-269.
19. Lee, J. and W.F. Ramirez (1996). On-line optimal control of induced foreign protein production by recombinant bacteria in fed-batch reactors. *Chemical Engineering Science* 51(4): 521-534.
20. Lee, J.-H., H.C. Lim, Y.J. Yoo and Y.H. Park (1999). Optimization of feed rate profile for the monoclonal antibody production. *Bioprocess Engineering* 20: 137-146.
21. Levisauskas, D., V. Galvanauskas, R. Simutis and A. Lübbert (1999). Model based calculation of substrate/inducer feed-rate profiles in fed-batch processes for recombinant protein production. *Biotechnology Techniques* 13: 37-42.
22. Pyun, Y.R., J.M. Modak, Y.K. Chang and H.C. Lin (1989). Optimization of Biphasic Growth of *Saccharomyces carlsbergensis* in Fed-batch Culture. *Biotechnology and Bioengineering* 33: 1-10.
23. Riesenber, D., V. Schulz, W.A. Knorre, H.D. Pohl, D. Korz, E.A. Sanders and A. Roß (1991). High cell density cultivation of *Escherichia coli* at controlled specific growth rate. *Journal of Biotechnology* 20: 17-28.
24. Suzuki, T., T. Yamane and S. Shimizu (1986). Mass production of poly-B-hydroxybutyric acid by fed-batch culture with controlled

carbon/nitrogen feeding. *Applied Microbiology and Biotechnology* 24: 370-374.

25. Tartakovsky, B., B.E. Dainson, D.R. Lewin and M. Sheintuch (1996). Observer-based non-linear control of a fed-batch autoinductive fermentation process. *The Chemical Engineering Journal* 61: 139-148.
26. Wang, F.S. and C.H. Shyu (1997). Optimal feed policy for fed-batch fermentation of ethanol production by *Zymomous mobilis*. *Bioprocess Engineering* 17(2): 63-68.
27. Zafiriou, E. and J. Zhu (1989). Optimal feed rate profile determination for fed-batch fermentations in the presence of model-plant mismatch, American Control Conference, Pittsburgh, PA, 2006-2009.

2 PRINCIPLES AND DEFINITIONS OF MODELLING, ESTIMATION AND CONTROL

2.1 Introduction

Modelling has become a prerequisite for the development of optimisation and control methods. The usefulness of simulation models in connection with experimental planning is also evident. Thus modelling and simulation are valuable tools in basic biological research by directing investigators towards a search for quantifiable results. Furthermore, for the practical implementation of control, estimation of non-measurable state variables is necessary as well. This chapter is written to introduce the techniques of modelling and estimating biotechnological processes. The subjects will not be covered extensively but briefly as an introduction to the theory used in the following chapters.

2.2 Modelling the Bioreactor

Modelling Biotechnological Processes

The modelling of biotechnological processes began with the equations of Blackman in 1905 and Monod in 1942, which related the concentration of the limiting substrate to the growth rate of the microorganism. Ever since, a great number of models have been developed for a wide variety of processes using different microorganisms. Most of them still contain kinetics of the Monod type, which shows the fundamental nature of that equation.

Within the framework of this thesis it would be excessive to provide a complete survey of all models or model structures. Therefore, emphasis will be put on introducing the ideas and structure of models for biotechnological processes together with some interesting examples.

A typical time course of batch cultivation is shown in Figure 1. With respect to cell mass, the process can be divided into an initial lag phase, a growth or exponential phase, a stationary phase, and a declining or death phase. The initial lag phase is due to regulatory phenomena of the microorganisms as an adaptation from the inoculum to the conditions in the bioreactor. During the lag and growth phases substrate is in excess. In a later stage of the process, substrate limitation, product inhibition, or other phenomena lead to the stationary phase with a zero growth rate. This can be followed by a declining phase in which the cell mass decreases due to lysis or endogenous metabolism.

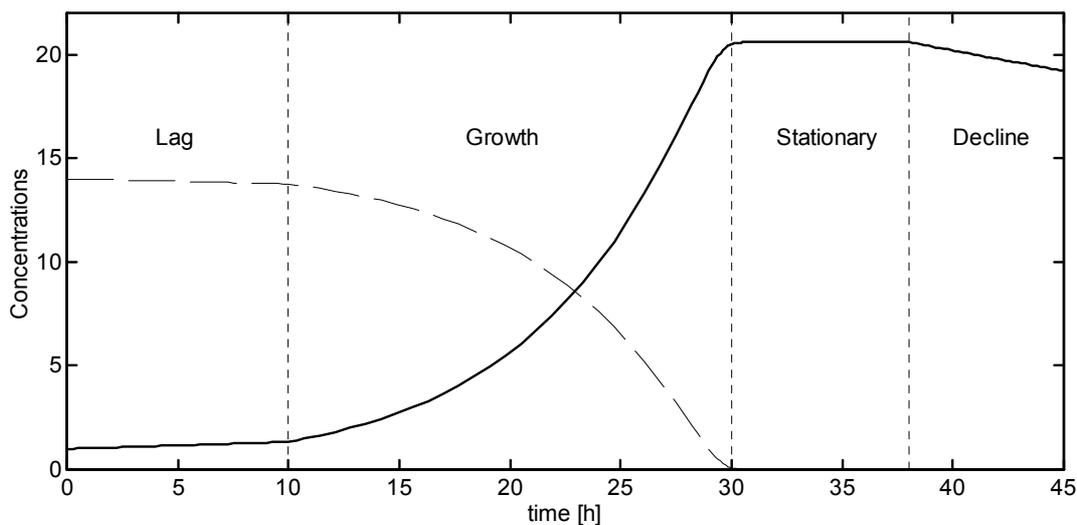


Figure 1: Typical time course of a batch process with biomass (solid) and limiting substrate (dashed) concentrations. Showing a lag-phase, exponential- or growth-phase, stationary-phase, and decline- or death-phase.

The behaviour is more complicated when product concentration is considered, some products are fully growth associated and increase proportionally to the growth rate. More complicated kinetics will be necessary, when separate growth and production phases or secondary metabolites are involved.

The function of a biological model is to describe the metabolic reaction rates and their stoichiometry on the basis of present and past bioreactor conditions. Biological models or growth models can be classified using the conceptual framework first suggested by Tsuchiya et al [1], as either structured or unstructured and either segregated or nonsegregated.

Structured versus Unstructured Models

Unstructured models take the cell mass as a uniform quantity without internal dynamics. The reaction rates depend only upon the macroscopic conditions in the liquid phase of the bioreactor. Therefore the models only contain kinetics of growth, substrate uptake, and product formation. This is a good approximation if the response time to changes in the environment of the cell is either negligibly small or very long compared to the duration of the cultivation process.

In contrast to these unstructured models, the internal state of the cells can also be considered which leads to so-called structured models. Such models are structured on the basis of biomass components such as concentrations of

metabolites, enzymes, DNA and/or RNA. With these models it becomes possible to describe a lag phase or transient phase.

Segregated versus Nonsegregated Models

Nonsegregated models treat the culture as a collection of average cells, all with the same characteristics at any given time. Usually unstructured models are also nonsegregated models. Segregated models treat each cell as independent, and a population as a collection of such distinct cells. They describe different morphological types of cells or cell ageing and sometimes describe the interactions between different cells.

General Model Structure

Biotechnological processes generally have the following structural elements: the liquid phase, gas phase, and the biotic phase, the latter consists of the cells or enzymes. All the reactions catalysed by microorganisms take place in the liquid phase. The relevant components are:

- Cell mass or biomass, C_X , autocatalytically synthesised from the substrates.
- Substrates, C_S , as energy and nutrient suppliers.
- Products, C_P , inside or outside of the cells.
- Dissolved gases, mainly oxygen, C_O , and carbon dioxide, C_C , which are connected to the gas phase by mass exchange.

In Figure 2 the corresponding structure of a model of a biotechnological process is shown. It consists of the liquid phase model, a gas phase model, and

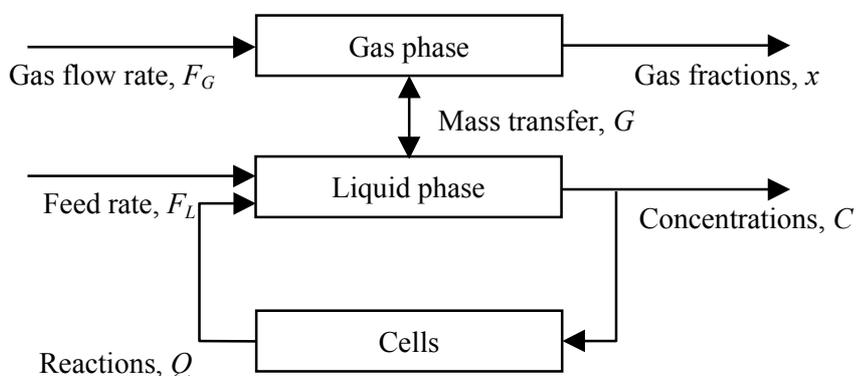


Figure 2: Structure of models for biotechnological processes.

a cell model. Taking this biological model as a local microkinetic model, it should be independent of the type of reactor and mode of operation.

Liquid Phase

The liquid phase of a mixed tank bioreactor can be modelled using input, output and consumption mass balances:

$$\begin{aligned}\frac{dV_L C}{dt} &= F_{L,in} C_{in} - F_{L,out} C + Q V_L + G V_L \\ \frac{dV_L}{dt} &= F_{L,in} - F_{L,out}\end{aligned}\quad (1)$$

Where C is the vector of concentrations of reactants (C_X, C_S, C_P, C_O, C_C), $F_{L,in}$ and $F_{L,out}$ are the inflow and outflow rate of the reactor, V_L is the liquid phase volume, C_{in} is the vector of concentrations in the inflow, Q is the vector of reaction rates in the liquid phase, and G is the mass exchange vector with the gas phase. Rewriting these equations gives:

$$\begin{aligned}\frac{dC}{dt} &= \frac{F_{L,in}}{V_L} \cdot (C_{in} - C) + Q + G \\ \frac{dV_L}{dt} &= F_{L,in} - F_{L,out}\end{aligned}\quad (2)$$

With the above model equations different kinds of processes can be described:

- Batch cultivation: $F_{L,in}=F_{L,out}=0$, $V_L=\text{constant}$
- Fedbatch cultivation: $F_{L,in}\neq 0$, $F_{L,out}=0$, $V_L\neq\text{constant}$
- Semi-continuous cultivation: $F_{L,in}\neq F_{L,out}\neq 0$, $V_L\neq\text{constant}$
- Continuous cultivation: $F_{L,in}=F_{L,out}\neq 0$, $V_L=\text{constant}$

Besides this physical model, a kinetic model has to be established, describing the biological reactions, Q , by the cells. A general characteristic of microbially catalysed reactions is that the reaction rates are proportional to the amount of active biomass:

$$Q = q \cdot C_X \quad (3)$$

Where q is defined as the specific reaction rate expressed in $\text{mol}\cdot\text{g}^{-1}\cdot\text{h}^{-1}$ and is related by laws of conservation and stoichiometry to represent the kinetic equations of the cell model. For unstructured cell models these equations are algebraic ones, while structured models also include further differential equations for biological compounds [2].

Gas Phase

The main components of the gas phase are oxygen, carbon dioxide, nitrogen and water vapour. Sometimes a biological product may also evaporate into the gas phase, like ethanol. The mass balances for the gas phase are similar to those for the liquid phase, except they are expressed by their mole fractions, x_O , x_C , x_N , x_W , and x_P , for oxygen, carbon dioxide, nitrogen, water vapour, and evaporated product, respectively:

$$\frac{dx}{dt} \cdot \frac{p \cdot V_G}{T} = \frac{F_{G,in}}{T_{in}} \cdot p_{in} \cdot x_{in} - \frac{F_{G,out}}{T} \cdot p \cdot x - R \cdot V_L \cdot G \quad (4)$$

Where, x and x_{in} are vectors with mole fractions of the components in the gas phase and the inlet gas flow, V_G is the volume of the gas phase and $F_{G,in}$ and $F_{G,out}$ are the rates of the ingoing and outgoing gas flows. G is a vector of the transfer rates of the components from the gas phase to the liquid phase. p and p_{in} (Pa) are the pressures of the gas phase and inlet gas flow, T and T_{in} are the temperatures of the gas phase and inlet gas flow, and R is the gas constant. It is assumed that no nitrogen is exchanged between gas and liquid phases, and that no product is present in the gas flow at the inlet:

$$x_{in} = [x_{O,in} \quad x_{C,in} \quad x_{N,in} \quad x_{W,in} \quad 0]^T \quad (5)$$

And

$$G = [OTR \quad CTR \quad 0 \quad WTR \quad PTR]^T \quad (6)$$

Where OTR , CTR , WTR and PTR are the transfer rates from gas to liquid phase of oxygen, carbon dioxide, water vapour, and evaporated product respectively.

Mass Transfer Between Gas Phase and Liquid Phase

The mass transfer rate between gas phase and liquid phase is proportional to the concentration gradient in the interfacial area and to the volumetric mass transfer coefficient, $k_L a$ (h^{-1}). It can be calculated by the film model, here given for oxygen:

$$OTR = k_L^O a (C_O^* - C_O) \quad (7)$$

Where * indicates the saturation concentration in the gas-liquid interface. The saturation concentration for the dissolved gases depends on the mole fractions in the gas phase. Following Henry's law the saturation concentration for dissolved oxygen, with H_O in as the Henry constant for oxygen, can be calculated by:

$$C_O^* = \frac{p_O}{H_O} = \frac{p}{H_O} x_O \quad (8)$$

This oxygen concentration in the liquid phase at saturation is proportional to the partial pressure in the gas phase. The mass transfer coefficient, $k_L a$, depends on the hydrodynamics in the reactor and the physical properties of the cultivation medium. It can be determined experimentally, e.g. by a dynamic [3].

When the oxygen concentration in the liquid is kept constant by a controller, the accumulation of the dissolved oxygen concentration will be very small and makes the oxygen uptake rate by the cells, OUR , practically equal to the OTR . As a consequence the OUR can be deduced directly from gas and liquid analyses only. First it is assumed that the OTR is approximately equal to $-CTR$ and the PTR and WTR are negligible, so the flow rate of the inlet gas equals the flow rate of the outlet gas. Second the difference between the pressure of the inlet gas and the pressure of the gas phase is assumed negligible. Now the oxygen fraction in the gas phase can be described by:

$$\frac{dx_O}{dt} = \frac{F_G}{V_G} \cdot \left(\frac{T}{T_{in}} x_{O,in} - x_O \right) - \frac{R \cdot T \cdot V_L}{p \cdot V_G} \cdot OUR \quad (9)$$

Where

$$OUR = k_L^O a \cdot \left(\frac{p}{H_O} \cdot x_O - C_O \right) \quad (10)$$

When $x_O(t=0)$ is known, the only unknown variable in Eqs. (9) and (10) is OUR .

Unstructured, Nonsegregated Modelling

In unstructured, nonsegregated models, the biological reactions depend solely on macroscopic variables that describe the conditions in the reactor and the only biological state variable is the cell mass concentration, C_X . Nevertheless, this type of model can cover many phenomena in biotechnological processes and as such is used widely. Unstructured models only fail when intracellular dynamics must be considered. This thesis will solely deal with unstructured, nonsegregated modelling.

Simple Growth and Single Substrate Kinetics

During growth, new cell mass is formed autocatalytically with specific growth rate μ , which is generally defined as a function of the concentration of the limiting substrate. Microbial reactions usually show saturation at high substrate concentrations, that is, the reaction rate approaches a maximum value, μ_{max} . On the other hand, the reaction rate equals zero if no substrate is available. So a general representation of the specific growth rate is:

$$\mu(C_S) = \mu_{max} \cdot f(C_S) \quad (11)$$

Where f is a normalised function (between 0 and 1) of the growth kinetics. Table 1 summarises several basic types of normalised kinetics for the specific growth rate, as been proposed in literature. The ratio of cell mass formed to substrate consumed, which is called the yield coefficient Y_{XS} , is usually assumed constant. For a simple process with one substrate Eq. (3) becomes:

$$q(C_S) = \begin{bmatrix} q_X \\ q_S \end{bmatrix} = \begin{bmatrix} 1 \\ -Y_{XS} & -1 \end{bmatrix} \cdot \mu(C_S) \quad (12)$$

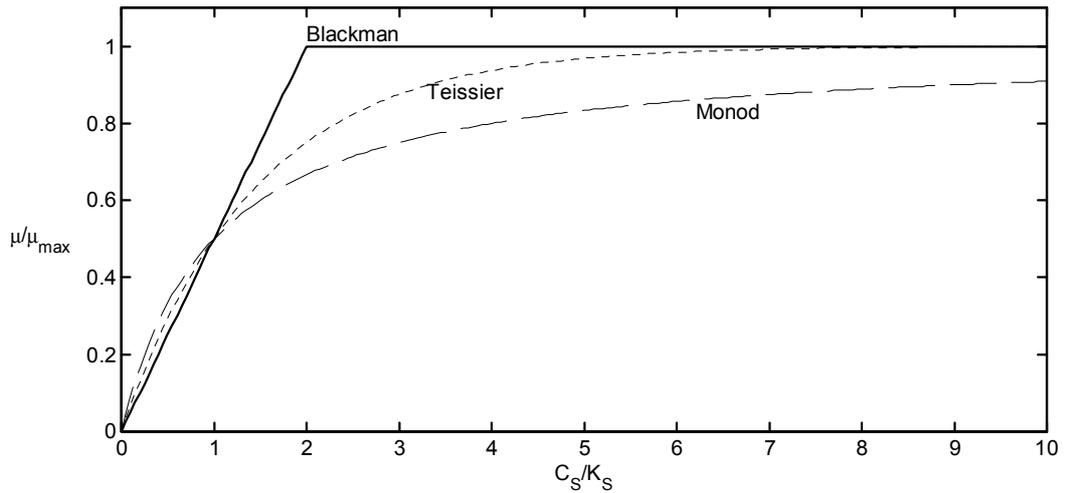


Figure 3: Normalised Blackman, Monod and Teissier kinetics, versus normalised substrate concentration.

Since substrate uptake is proportional to the growth rate, the kinetics of Table 1 can also be taken as normalised functions of substrate uptake kinetics. Although the equations in Table 1 should be taken as kinetics for the globally observed behaviour of the culture, which need not have a close relation to microkinetics of the biological reaction, an interpretation of certain kinetics can be given. The Monod equation is analogous to the Michaelis-Menten enzyme kinetics and can be seen as a growth rate limiting, carrier-mediated transport system for the substrate. The Blackman equation can be interpreted such that, at high substrate concentrations not the substrate uptake, but another metabolic reaction, is rate limiting and the Teissier model is very similar to Monod, however, with a more rapid saturation [4].

A plot of the $\mu(C_S)$ characteristics of these kinetics is given in Figure 3. This figure shows that when considering measurement errors, it is difficult to discriminate between the different kinetics, especially when using batch culture data only. Therefore, the Monod kinetics is generally a good choice for a single substrate model [4].

Table 1: Growth kinetics for a single substrate

model name	normalised kinetics, f
Blackman (1905)	$\min(1, K_B \cdot C_S)$
Monod (1942)	$\frac{C_S}{K_M + C_S}$
Teissier (1942)	$1 - e^{-K_T \cdot C_S}$

Inhibition

Besides substrate limitation, inhibition by substrates, products or biomass is quite often found in biotechnological processes. Han and Levenspiel [5] give an extensive review of inhibition kinetics. Most of the kinetics are extensions of the Monod equation and have been derived from enzyme inhibition kinetics. The most cited inhibition kinetics is proposed by (and called) Haldane:

$$f(C_S) = \frac{C_S}{K_M + C_S + \frac{C_S^2}{K_I}} \quad (13)$$

Product, substrate, or cell inhibition can be obtained by choosing the variable C_I as C_P , C_S or C_X , respectively.

Multiple Essential Substrates

The above kinetics are valid for a single limiting compound. However, when multiple substrates limit the growth of the organism the elementary kinetics of Table 1 and equation (13) have to be extended. When all substrates are essential, meaning there is no growth if only one of the substrates is lacking, there are two ways of modelling the kinetics; interacting and non-interacting. In the interacting model all substrates together determine the growth rate, for N substrates, this can be expressed by a product of N normalised kinetics:

$$\mu(C_{S1}, C_{S2}, \dots, C_{SN}) = \mu_{max} \cdot f_1(C_{S1}) \cdot f_2(C_{S2}) \cdot \dots \cdot f_N(C_{SN}) \quad (14)$$

In the non-interacting model, it is assumed that only the substrate with the greatest limitation determines the growth rate. This leads to a selection of the minimum growth rate allowed among all substrates:

$$\mu(C_{S1}, C_{S2}, \dots, C_{SN}) = \min\{\mu_{S1}(C_{S1}), \mu_{S2}(C_{S2}), \dots, \mu_{SN}(C_{SN})\} \quad (15)$$

Where all terms μ_{Si} have the form of Eq. (11). As an example, both are represented below for a situation with two limiting substrates:

Interacting model,

$$\mu(C_{S1}, C_{S2}) = \mu_{max} \cdot \frac{C_{S1}}{K_{M1} + C_{S1}} \cdot \frac{C_{S2}}{K_{M2} + C_{S2}} \quad (16)$$

Non-interacting model,

$$\mu(C_{S1}, C_{S2}) = \mu_{max} \cdot \min \left\{ \frac{C_{S1}}{K_{M1} + C_{S1}}, \frac{C_{S2}}{K_{M2} + C_{S2}} \right\} \quad (17)$$

Growth Enhancing by Alternative Substrates

Another case where multiple substrates have to be considered is when one of the substrates is sufficient for growth and others are used up in parallel, and so increase the growth rate. For modelling these kinds of phenomena the entire kinetics can be derived from the sum of elementary kinetics, Eq. (11):

$$\mu(C_{S1}, C_{S2}, \dots, C_{SN}) = \mu_1(C_{S1}) + \mu_2(C_{S2}) + \dots + \mu_N(C_{SN}) \quad (18)$$

The biological meaning of additive kinetics is that the bottleneck for growth is not associated with substrate uptake steps. While the multi-substrate kinetics Eqs (14) and (15) for essential substrates is used regularly in literature and seem to give a good description of observed phenomena, the case for enhancing substrates is not often described. Experimental data that support Eq. (18) are very rare. The reason might be a too-simple approximation of the growth kinetics due to the neglect of the regulator response of the microorganisms, which can only be described by structured models.

Cell Maintenance

At low growth rates one often can observe a decrease in cell yield. This can be explained by an additional growth-independent substrate consumption for maintaining the cell structure, or by lysis processes. Herbert (1959) took this effect into account by discriminating between observed net specific growth rate, μ , and true specific growth rate μ_G . The difference is the specific rate of endogenous metabolism, m_X .

$$\mu = \mu_G - m_X \quad (19)$$

Compared with Eq. (12) the observed growth dependent yield becomes:

$$Y_{XS}(\mu) = Y_{XSG} \cdot \frac{\mu}{\mu + m_X} \quad (20)$$

Here, μ_G may be expressed as any of the previously given kinetics. During strong growth limitation the observed growth rate, μ , can become less than zero. In this case the rate m_X also can be viewed as a specific rate of cell lysis or the degradation rate of intracellular storage material.

Pirt [6] and Ierusalimsky [7] described the phenomena of the decreasing yield as an additional need for substrate consumption for maintenance of the cell structure, expressed by the specific rate m_S . Thus, if the uptake rate of substrate used only for growth is q_{SG} , then the total substrate uptake q_S is:

$$q_S = q_{SG} - m_S \quad (21)$$

Note that production is defined as positive rate, and thus q_{SG} is negative. Now the specific growth rate is given by

$$\mu = -q_{SG} \cdot Y_{XSG} \quad (22)$$

And the yield becomes, comparing Eq. (8),

$$Y_{XS}(\mu) = Y_{XSG} \cdot \frac{\mu}{\mu + m_S \cdot Y_{XSG}} \quad (23)$$

By comparing Eqs. (20) and (23) it can be seen that both are equivalent in the regions of normal growth, if

$$m_S = \frac{m_X}{Y_{XSG}} \quad (24)$$

The main drawback of Eq. (21) is that in dynamic simulations the substrate concentration can become less than zero because of the zero-order maintenance reaction. This situation cannot occur with the model in Eq (19), where with strong limitation the growth rate becomes less than zero.

2.3 Observing the bioreactor

In this paragraph, some general concepts on state and parameter estimation for biological processes are briefly discussed. This introductory paragraph is based on the works of Bastin and Dochain [8], Stephanopoulos and San [9], Jazwinski [10], and Lewis [11].

Indirect Measurements and Correlations

Certain data collected from bioreactors can be used directly (e.g. pH, dissolved oxygen concentration, carbon dioxide fraction, etc.), while others have very little physical significance by themselves and must therefore be combined with other measurements to provide physically significant information. Such a procedure is called an indirect measurement and provides information on cellular metabolism and cultivation conditions. The rates of consumption or production of many species can be calculated by taking a simple species balance around the bioreactor. Among the quantities most easily acquired on-line and most frequently calculated are the oxygen uptake rate (*OUR*), the carbon dioxide evolution rate (*CER*), and the respiratory quotient (*RQ*). These provide very useful information and are very good indicators of cellular respiratory activities.

Software sensors

During cultivation, information on the biological components and well-being of the micro-organisms (biomass performance) is of vital importance. For better control and optimisation of cultivation processes it is essential to know on-line these physiological parameters. Measuring is difficult, if not impossible, for many of these parameters. Therefore, any control or optimisation based on physiological parameters cannot be implemented unless values can be measured on-line to provide the necessary information required by the controller or optimiser. Although efforts to develop such sensors are underway, the availability and reliability of instruments are very limited. There is thus an urgent need to provide the necessary information by estimating these parameters from others that are simpler and easier to measure. Such an estimation technique is called an observer or a *software-sensor*.

A software-sensor is a systematic method to utilise information of varying

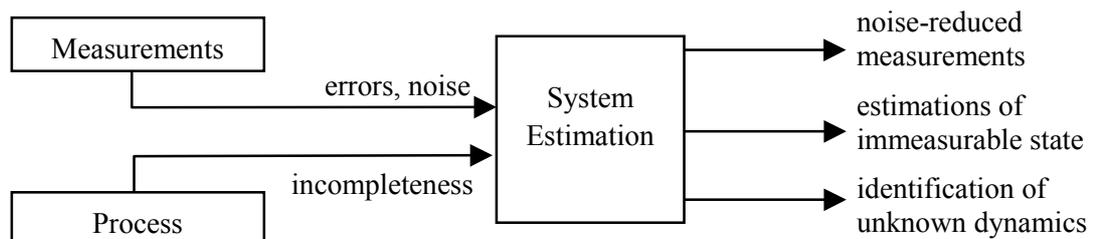


Figure 4: Schematic description of system estimation with inputs and outputs.

form and accuracy for the purpose of identifying the state of cultivation. As indicated in Figure 4, the methodology utilises general process knowledge (with its incompleteness) as well as information contained in the process measurements (with its noise and errors) in order to produce noise-reduced, corrected measurements, to estimate immeasurable states, and to identify unknown or uncertain elements of system dynamics. This output can be an instrument for control or optimisation.

State estimation

Consider the general model for the biological process introduced in paragraph 2.2:

$$\frac{dC}{dt} = Q(C) + D(C_{in} - C) + G(C) \quad (25)$$

where the development through time of the state (concentrations in the liquid) C , depend on the reaction rates $Q(C)$, the dilution rate D , the incoming concentrations C_{in} , and the mass exchange with the gas phase $G(C)$. From this model, a general class of state observers can be written as follows:

$$\frac{d\hat{C}}{dt} = Q(\hat{C}) + D(C_{in} - \hat{C}) + G(\hat{C}) + p(\hat{C})[C_I^m - \hat{C}_I] \quad (26)$$

in which \hat{C} denote the estimated concentrations, C_I^m a measurable subset 1 of the concentrations, \hat{C}_I the estimated subset 1 of the concentrations, and $p(\hat{C})$ the observer gain. Note that the observer consists of two parts, a copy of the model and an error term. In the case of perfect estimation, the latter equals zero [8, 10, 11].

Not all combinations of model equations and measurements yield a suitable estimate of the desired state. The concept of *observability* is introduced to test whether the model structure and subset of measured concentrations contain the right information for estimating the immeasurable concentrations. If the convergence speed of the implemented observer can be fixed arbitrarily, then this is called *exponential* observability. If the dynamics are exclusively determined by the experimental conditions, this is called *asymptotic* observability [8].

The design of an observer entails finding a suitable expression for $p(\hat{C})$. To this end, the dynamics of the observation error, $e = C - \hat{C}$, is investigated:

$$\frac{de}{dt} = Q(\hat{C} + e) - Q(\hat{C}) - De + G(\hat{C} + e) - G(\hat{C}) + p(\hat{C})[C_I^m - \hat{C}_I] \quad (27)$$

In the design rule for a *Luenberger* observer, which is described above, the goal is to make $e = 0$ an asymptotically stable equilibrium point of Eq (27). Therefore, $p(\hat{C})$ is selected such that the eigenvalues have a strictly negative real part [8].

Parameter Estimation

Biological systems are characterised by complex, non-linear relationships involving poorly identified parameters. This makes them likely candidates for applying adaptive algorithms. For instance, to avoid explicitly modelling the specific rates, these can be treated as parameters, and estimated along with the state variables. The dynamics of the latter can then be described by simple mass-balance equations.

Stephanopoulos and San [9] followed this approach by implementing an adaptive Extended Kalman Filter (EKF). Specific rates and state variables were both satisfactorily estimated under steady-state or transient operation from measurements provided by the off-gas analysis of a cultivation.

Kalman Filtering

The *Luenberger* observer as described above is applicable in many situations, however, when knowledge of measurement noise and model uncertainty is available the use of a Kalman filter would be preferred. A Kalman Filter estimates the state of a system based on the knowledge of the system input (with its uncertainty), the measurement of the system output (with its uncertainty), and a model of the relation between the input and output (with its uncertainty). Kalman filters are well known and widely applied but are usually treated in a very concise and formal way, concentrating on the mathematical derivation, thus making bioengineers reluctant in applying the technique. In this paragraph the Kalman filter will be introduced briefly and intuitively.

The Kalman filter provides an estimate of the state that tries to minimise the inconsistencies with all pieces of information in the least squares sense. When

doing so, the inconsistencies are weighted with a measure of the certainty of the information. Uncertain information is given low weight, whereas highly certain information is given a very high weight. When all information is available at once, it can be processed in batch, resulting in a classical weighted least squares problem. If, however, the information becomes available incrementally, as is the case for an on-line Kalman filter, a recursive formulation of the estimation problem is necessary.

Linear Kalman Filter Algorithm

Suppose the state of a system evolves according to the linear discrete-time state equation:

$$x_{k+1} = Ax_k + Bu_k + w_k \tag{28}$$

Where A is the state matrix, B the input matrix, u_k the input vector at time step k , and w_k represents the process uncertainty at time step k with covariance matrix Q . A linear system has a linear output (measurement) equation:

$$y_k = Cx_k + v_k \tag{29}$$

Where y_k is the measurement vector, C the measurement matrix, and v_k represents Gaussian measurement inaccuracy with zero mean and covariance R . Depending upon the matrix C , a (multiple) of all states, some of the states,

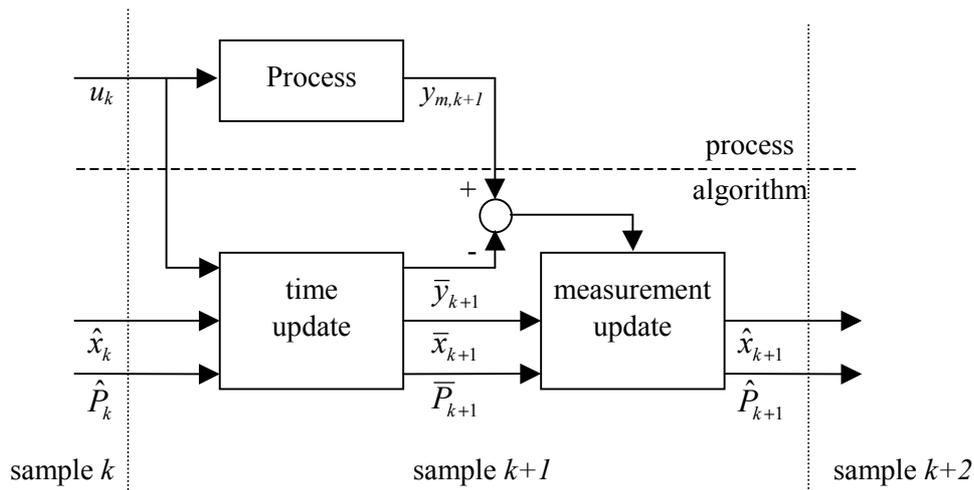


Figure 5: Schematic representation of the Kalman filter calculations. It shows the different steps to be taken at each time instant k . The dashed horizontal line makes distinction between the algorithm and process, the dotted vertical lines symbolise the successive sample moments.

or a linear combination of the states are measured.

The linear Kalman filtering algorithm is depicted in Figure 5. In this figure, a distinction is made between the cultivation process that gives the data and the algorithm that processes the data. The algorithm has two steps, respectively the time update and the measurement update. In the time update a one sample ahead prediction is made for the state and output variables (respectively $\bar{x}_{k+1}, \bar{y}_{k+1}$) and the prediction variance of the states (\bar{P}_{k+1}). The actual values of the input variables (u_k) and the available estimated results of the Kalman filter (\hat{x}_k, \hat{P}_k) at sample moment k are used for this prediction.

The next step is the measurement update, which takes place as soon as new data becomes available and where the prediction of the output variable (\bar{y}_{k+1}) is corrected. The correction is proportional to the *innovation*, which is the difference between the measured process output (y_{k+1}) and the predicted output (\bar{y}_{k+1}). The magnitude of the correction, called the *Kalman gain* K_k , varies for successive samples and aims to minimise the error covariance (\hat{P}_k) for the state and output variables. Now, the measurement update gives the best estimate (\hat{x}_{k+1}) of the states and its variance (\hat{P}_{k+1}). The values for the state estimate (\hat{x}_{k+1}) are used as the software sensor output. The entire set of these

Table 2: Discrete time linear Kalman filter algorithm

System and measurement model
$x_{k+1} = A_k x_k + B_k u_k + w_k \quad w_k \sim (0, Q_k)$ $y_k = C_k x_k + v_k \quad v_k \sim (0, R_k)$
Initialisation
$P_0 = P_{x_0}, \quad \hat{x}_0 = \bar{x}_0$
Time update (one sample-ahead prediction)
$\bar{x}_{k+1} = A_k \hat{x}_k + B_k u_k$ $\bar{y}_{k+1} = C_k \bar{x}_{k+1}$ $\bar{P}_{k+1} = A_k \hat{P}_k A_k^T + Q_k$
Measurement update (after including the measurement)
$\hat{P}_{k+1} = \bar{P}_{k+1} - \bar{P}_{k+1} C_{k+1}^T R_{k+1}^{-1} C_{k+1} \bar{P}_{k+1}$ $K_k = \hat{P}_{k+1} C_{k+1}^T R_{k+1}^{-1}$ $\hat{x}_{k+1} = \bar{x}_{k+1} + K_k (y_{k+1} - \bar{y}_{k+1})$

equations comprises what is called the discrete time Kalman filter [12, 11] and is summarised in Table 2.

A few remarkable advantageous features of this linear Kalman filtering algorithm can easily be observed. First, each calculation step only requires the last estimate and a new set of measurement data. The essential advantage of such simple "step-by-step" structure of the computational scheme is that there is no need to store all old results and measurement data for each up-dating state estimate, and this saves computer memory and processor time, especially in real-time (on-line) applications. Second, all recursive formulas of the algorithm are straightforward and linear, consisting of only matrix multiplication and addition, and a single matrix inversion in the calculation of the Kalman gain.

Non-linear Kalman Filter

The state equation (28) and measurement equation (29) calculate respectively the measurement and the predicted state as a linear function of the state x . Unfortunately, in practice these functions are often non-linear:

$$\begin{aligned} x_k &= f(x_{k-1}, u_{k-1}) + w \\ z_k &= h(x_k) + v \end{aligned} \tag{30}$$

In order to apply the Kalman filter these equations are linearised in the most recent estimate to obtain the so-called Extended Kalman filter. Jazwinski [10] and Lewis [11] wrote classical books on Kalman filters which contain, among other things, a study of stability, sensitivity to modelling errors for linear Kalman filters, and various approaches to non-linear filtering.

Tuning of software sensors

The first prerequisite for tuning software sensors is the availability of performance criteria such as convergence speed and noise sensitivity. The tuning of software sensors can then be studied in various ways. By finding the relation between the tuning parameters and the performance criteria a performance analysis can be made. From such an analysis, one can gain insight in the tuning procedure. However, this usually does not yield values for the tuning parameters that are suitable for the actual process. Another approach is defining an objective function from the performance criteria

mentioned above. By using simulation data or experimental data from previous runs, the optimal values for the tuning parameters can be found by using an optimisation routine. Although this method will not give insight in the tuning problem, it will result in an observer that is tuned properly for practical implementation [13].

2.4 Control

The main goals in applying control methods to microbiological systems are to improve operational stability and production efficiency and profit, and to handle dynamic changes during start-up and shutdown. The number of ways to control a biotechnological process and the number of reports about such control is vast, therefore, this paragraph will briefly describe only a few well-known concepts of general control.

Feedback control

A classical automatic control system in the form of feedback control is shown in Figure 6. Since classical control techniques are covered extensively in standard references [14], the treatment given here is at most cursory. Almost all feedback control systems use negative feedback. The controller generates an error signal, ε , by subtracting the measured process output (y , the controlled variable) from a desired value (y_{sp} , set-point), and calculates the control signal, u , by applying a certain algorithm to the error signal. This control signal then manipulates the process input (control element) to reduce the error. The most common controller is called a PID controller and uses proportional, integral, and derivative actions [14].

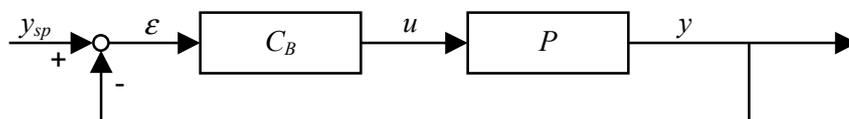


Figure 6. Feedback controller C_B for process P .

Feed-Forward Control

In most situations simple feedback control is adequate. However, there are situations in which simple feedback control is inadequate, and more advanced control is required. For instance with feed-forward control where, unlike with normal feedback control, one does not wait until a disturbance actually affects

the output. Instead the disturbance is measured and corrective control is applied in anticipation of the expected effect. This is a better control for eliminating the effect of disturbance [14].

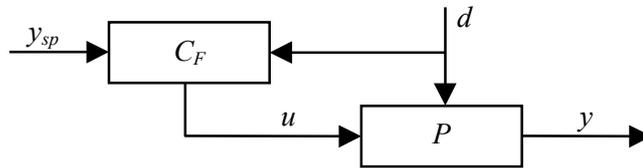


Figure 7. Feed-forward controller C_F eliminates disturbance d .

The disadvantages of a feed-forward controller as depicted by Figure 7, are the need for a good process model, its sensitivity to process parameter variation, the need to know the disturbance a priori, and its inability to cope with unmeasured disturbances.

Feed-forward-Feedback Control

A feed-forward control scheme is rarely used alone for the reasons mentioned above. Generally it is combined with feedback control to correct errors caused by imperfect knowledge in predicting the effect of disturbances on the output. A combination of feedback and feed-forward control is shown in Figure 8. This permits the major effect of the disturbance to be corrected by the feed-forward controller leaving the fine trimming to the feedback loop. Table 3 shows the advantages and disadvantages of feed-forward and feedback control [14].

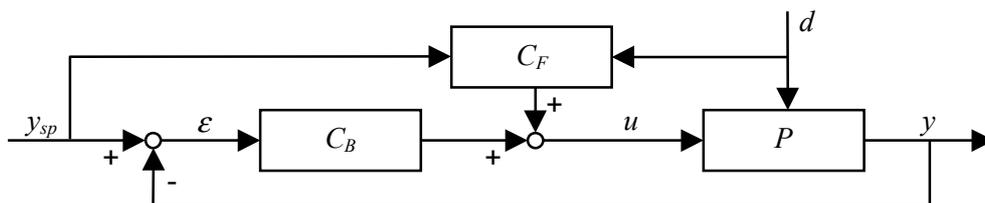


Figure 8. Combined feed-forward C_F and feedback C_B control.

Table 3: Advantages and disadvantages of feed-forward and feedback control.

Advantages	Disadvantages
<i>Feed-forward</i>	
Acts before the effect of a possible disturbance Good for slow systems Does not introduce instability	Requires an accurate process model Cannot cope with unmeasured disturbances Sensitive to process parameter variation
<i>Feedback</i>	
Does not require a good model insensitive to modelling errors insensitive to parameter changes	Waits until the effect of a disturbance is felt unsatisfactory for slow systems might create instability

2.5 Nomenclature

Paragraph 2.2

C_C	$\text{mol}\cdot\text{m}^{-3}$	carbon dioxide concentration in the liquid phase
C_{in}	$\text{mol}\cdot\text{m}^{-3}$	vector of concentrations in the liquid inflow
C_O	$\text{mol}\cdot\text{m}^{-3}$	oxygen concentration in the liquid phase
C_O^*	$\text{mol}\cdot\text{m}^{-3}$	oxygen saturation concentration
C_P	$\text{mol}\cdot\text{m}^{-3}$	product concentration
C_S	$\text{mol}\cdot\text{m}^{-3}$	substrate concentration
C_X	$\text{g}\cdot\text{m}^3$	biomass concentration
CTR	$\text{mol}\cdot\text{m}^{-3}\cdot\text{h}^{-1}$	carbon dioxide transfer rate from gas to liquid phase
$F_{G,in}$	$\text{m}^3\cdot\text{h}^{-1}$	inlet gas flow rate
$F_{G,out}$	$\text{m}^3\cdot\text{h}^{-1}$	outlet gas flow rate
$F_{L,in}$	$\text{m}^3\cdot\text{h}^{-1}$	liquid inflow rate
$F_{L,out}$	$\text{m}^3\cdot\text{h}^{-1}$	liquid outflow rate
G	$\text{mol}\cdot\text{m}^{-3}\cdot\text{h}^{-1}$	vector of mass exchange between the liquid and gas phase
f	-	normalised function of growth kinetics
H_O	$\text{Pa}\cdot\text{m}^3\cdot\text{mol}^{-1}$	Henry constant for oxygen
K_B	$\text{m}^3\cdot\text{mol}^{-1}$	saturation constant of Blackman kinetics
K_I	$\text{mol}^2\cdot\text{m}^{-6}$	inhibition constant
K_M	$\text{mol}\cdot\text{m}^{-3}$	Monod saturation constant

CHAPTER: 2

K_T	$\text{m}^3 \cdot \text{mol}^{-1}$	saturation constant of Teissier kinetics
$k_L a$	h^{-1}	volumetric mass transfer coefficient between gas and liquid
$k_L^O a$	h^{-1}	volumetric mass transfer rate for oxygen
m_S	$\text{mol} \cdot \text{g}^{-1} \cdot \text{h}^{-1}$	specific substrate consumption rate for maintenance
m_X	h^{-1}	specific rate of endogenous metabolism (maintenance)
OTR	$\text{mol} \cdot \text{m}^{-3} \cdot \text{h}^{-1}$	oxygen transfer rate from gas to liquid phase
OUR	$\text{mol} \cdot \text{m}^{-3} \cdot \text{h}^{-1}$	oxygen uptake rate
p	Pa	pressure of the gas phase
p_{in}	Pa	pressure of the inlet gas flow
PTR	$\text{mol} \cdot \text{m}^{-3} \cdot \text{h}^{-1}$	evaporated product transfer rate from gas to liquid phase
Q	$\text{mol} \cdot \text{m}^{-3} \cdot \text{h}^{-1}$	vector of reaction rates in liquid phase
q	$\text{mol} \cdot \text{g}^{-1} \cdot \text{h}^{-1}$	vector of specific reaction rates
q_S	$\text{mol} \cdot \text{g}^{-1} \cdot \text{h}^{-1}$	specific reaction rate for substrate
q_X	$\text{mol} \cdot \text{g}^{-1} \cdot \text{h}^{-1}$	specific reaction rate for biomass
R	$\text{J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$	gas constant ($8.314 \text{ J} \cdot \text{mol}^{-1} \cdot \text{K}^{-1}$)
T	K	temperature of the gas phase
T_{in}	K	temperature of the inlet gas flow
V_L	m^3	volume of the liquid phase
V_G	m^3	volume of the gas phase
WTR	$\text{mol} \cdot \text{m}^{-3} \cdot \text{h}^{-1}$	water vapour transfer rate from gas to liquid phase
x	-	vector of fractions in the headspace
x_C	-	carbon dioxide fraction in the headspace
x_{in}	-	vector of fractions in the inlet gas flow
x_N	-	nitrogen fraction in the headspace
x_O	-	oxygen fraction in the headspace
x_P	-	product fraction in the headspace
x_W	-	water vapour fraction in the headspace
Y_{XS}	$\text{g} \cdot \text{mol}^{-1} \cdot \text{h}^{-1}$	yield coefficient of biomass on substrate
μ	h^{-1}	specific growth rate
μ_{max}	h^{-1}	maximal specific growth rate

Paragraph 2.3

A	state matrix
B	input matrix
C	measurement matrix
K_k	Kalman gain at time step k
k	moment in time
\bar{P}_k	predicted predicted covariance matrix at time step k
\hat{P}_k	estimated covariance matrix at time step k
Q_k	covariance matrix of w_k
R_k	covariance of v_k
u_k	input vector at time step k
v	Gaussian measurement uncertainty
w	process uncertainty / noise
x_k	state vector at time step k
\hat{x}_k	estimated state vector at time step k
\bar{x}_k	predicted state vector at time step k
y_k	measurement vector at time step k
\hat{y}_k	estimated measurement vector at time step k

2.6 References

1. Tsuchiya, H.M., A.G. Fredrickson and R. Aris (1966). Dynamics of Microbial Cell Populations. *Advance Chemical Engineering* 6: 125-132.
2. Roels, J.A. and N.W.F. Kossen (1978). On the modelling of microbial metabolism. *Prog. Ind. Microbiol.* 14: 95.
3. Van 't Riet, K. and J. Tramper (1991). *Basic Bioreactor Design*. New York, Marcel Dekker.
4. Bellgardt, K.H. (1991). *Cell Models. Measuring, Modelling and Control*. K. Schügerl. Weinheim, Verlagsgesellschaft mbH. 4: 267-298.
5. Han, K. and O. Levenspiel (1988). Extended Monod-equation for substrate, product and cell inhibition. *Biotechnology and Bioengineering* 32: 430.

6. Pirt, S.J. (1965). The maintenance energy of bacteria in growing cultures. Proc. R. Soc. Ser. B 163: 224.
7. Ierusalimsky, N.D. (1967). Bottle-necks in metabolism as growth rate controlling factors, Microbial Physiology and Continuous Culture, 3rd International Symposium, London, 23.
8. Bastin, G. and D. Dochain (1990). On-line Estimation and Adaptive Control of Bioreactors. Amsterdam, Elsevier Science Publishers.
9. Stephanopoulos, G. and K.Y. San (1984). Studies on On-Line Bioreactor Identification. I. Theory. Biotechnology and Bioengineering 26: 1176-1188.
10. Jazwinski, A.H. (1970). Stochastic processes and filtering theory. New York, Academic Press.
11. Lewis, F.L. (1986). Optimal Estimation. New York, USA, John Wiley and Sons.
12. Chen, G. (1994). Approximate Kalman Filtering. Singapore, World Scientific.
13. Rykaert, V.G. (1998). Model based optimization and control of bioprocesses - From theory to practice. PhD thesis, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium.
14. Stephanopoulos, G. (1984). Chemical Process Control: An Introduction to Theory and Practice. London, UK, Prentice-Hall International Inc.

3 RESPIRATION QUOTIENT: MEASUREMENT AND ESTIMATION DURING BATCH CULTIVATION IN BICARBONATE BUFFERED MEDIA

This chapter is based on:

R. Neeleman, E.J. van den End, and A.J.B. van Boxtel (2000). Estimation of the Respiration Quotient in a Bicarbonate Buffered Batch Cell Cultivation. Journal of Biotechnology 80: 85-95.

R. Neeleman (2001). Respiration Quotient: Measurement and estimation during batch cultivation in bicarbonate buffered media. In: Focus on Biotechnology. Volume IV, pp. 203-216. Kluwer Academic Publishers, The Netherlands.

3.1 Abstract

The Respiration Quotient (RQ) is a key metabolic parameter for cell cultures and is usually determined from gas analysis only. In bicarbonate buffered media the carbon dioxide balance is affected by accumulation and hence the RQ can not directly be calculated from gas measurements. A Kalman Filter as software sensor that estimates the carbon dioxide evolution rate can cope with these buffering capacities and thus is used for determining the RQ . The model used by the Kalman Filter lumps all carbonate in the liquid to one term in order to eliminate the role of a priori knowledge of cell and medium kinetics without affecting the performance. A reference experiment verified the performance of the software sensor and subsequent experiments with insect cells showed the progress of the RQ during cultivation.

3.2 Introduction

Animal cells, yeast cells and aerobic microbial cells oxidise organic compounds into water, carbon dioxide and other organic compounds to gain energy for their maintenance and growth. As such these organisms consume oxygen and produce carbon dioxide with rates called the oxygen uptake rate (OUR) and the carbon dioxide evolution rate (CER). These rates are direct indicators of metabolic activity. Their ratio called the respiration quotient

$$RQ = \frac{CER}{OUR} \quad (1)$$

varies with the nature of the substrates and products of the organism. Bonarius et al. [1] and Royce [2] argue that for cell cultures the RQ can be considered as a key metabolic parameter making it possible to detect on what medium substrate the organism grows.

Furthermore, stoichiometric coefficients, e.g. yields, are usually determined using elemental balancing (conservation of chemical elements). This set of balance equations can mostly not be solved since the number of unknown coefficients exceeds the number of balance equations. Additional information is required and with the right set of extra measurements the equations become solvable. Measuring the RQ introduces such information to solve the balance equations and calculate the stoichiometric coefficients. By means of RQ measurements it has been possible to close mass balances and to determine

metabolic flux distributions for yeast [3]. RQ data has also been used for on-line bioreactor control, for example to minimise glucose effects [4] or to optimise substrate consumption in yeast [5]. Its accurate evaluation is hence of great importance.

The RQ -value can be calculated using the oxygen and carbon dioxide concentrations in the gas stream into and from the reactor headspace [1]. However, the effect of pH control action on carbon dioxide evolved from the medium troubles the measurement of the CER . Royce [2] describes this phenomenon together with a solution to compensate for these so-called pH-effects. Besides these pH-effects, buffering capacities (e.g. bicarbonate) of the medium also trouble on-line determination of the CER . When the reactor is not in steady state (e.g. batch cultivation), direct measurement of carbon dioxide concentrations in the inlet and outlet of the reactor headspace does not satisfy. Dissociation and accumulation of carbon dioxide in the medium and headspace disturbs the steady state balance and results in an incorrect CER - and thus RQ -calculation.

To manage such problems software-sensors based on standard measurements are combined with mathematical observers to derive the internal states of the system. Stephanopoulos and San [6] gave an extensive discussion on such observer-based software-sensors for the reconstruction of the process states. For example the achieved amount of biomass and substrate concentrations are derived from oxygen and carbon dioxide measurements. Since that time the importance of observer based software- sensors increased. Other important work in this area was produced by Bastin and Dochain [7] and concerned observer-based software-sensors for adaptive control. All the work in this area was focussed on biomass and substrate estimation and requires more or less detailed a priori knowledge of the system. However, for monitoring and control purposes it is not always necessary to know the states, in a lot of applications only an indicator is needed on what substrate is being used or is necessary to be added.

This chapter describes and discusses a Kalman Filtering algorithm as software-sensor for estimating the RQ -value for batch cell cultivation from carbon dioxide measurements. The sensor will take account for buffering capacities of the media for carbon dioxide.

3.3 Materials and methods

Cell line and culture medium

The insect cell culture used was Se-IZD 2109, subclone G9, developed from *Spodoptera exigua* (Se) cells (German Institute for Zoology of the Technical University Darmstadt). For cultivation of the insect cells HyQ-CCM3 (Hyclone, Utah, USA) medium was used with 2.5% serum. It contained 0.1% Pluronic to protect the cells against shear forces. To avoid contamination, 50 mg.l⁻¹ gentamicin, 50 mg.l⁻¹ streptomycin and 10 mg.l⁻¹ penicillin were added.

Bioreactor

The cells were grown in a 1.8-litre flat bottom reactor containing 1 litre medium. A three-bladed marine impeller was used to agitate the medium and a four bladed turbine impeller was situated in the headspace. Both stirrers were mounted on the same axis and baffles in the reactor prevented vortex formation. Temperature, pH, and stirrer speed were controlled at 28°C, 6.3, and 400 rpm respectively. Oxygen was transported through the headspace only and the dissolved oxygen concentration was controlled at 30% air saturation by changing the oxygen fraction in a nitrogen/oxygen gas mixture. The total gas flow was kept constant at 100 ml.min⁻¹.

Analysis

The dissolved oxygen in the medium was measured by a polarographic electrode (Ingold, Urdorf, Switzerland). The pH was measured by a pH electrode (Ingold, Urdorf, Switzerland). The temperature was measured with a Pt100 temperature sensor. The CO₂ fraction in the outflow was measured by a Servomex 1400 series CO₂-Analyser (Servomex, Zoetermeer, The Netherlands). Data were logged and saved by an application in Matlab on a PC connected to the Bio Controller. Bicarbonate concentrations were calculated using the pCO₂ pressure, measured by a Radiometer ABL 500 (Radiometer, Zoetermeer, The Netherlands), and the acid dissociation constant for the CO₂-HCO₃⁻ equilibrium. Glucose and glutamine concentrations were measured with a YSI 2750 Select (Yellow Springs Instruments, Ohio, USA).

3.4 Gas concentrations in batch-wise cell cultures

For the on-line calculation of the RQ , the OUR and CER have to be known (measured or estimated) on-line. Two different methods are applied in the determination of these rates. First the OUR will be deduced directly from gas analyses and liquid measurements. Since there are no satisfactory sensors for dissolved carbon dioxide and the accumulation of carbon dioxide can not be discarded, an estimator is developed for the on-line estimation of the CER . Finally, the ratio of these rates will give the RQ -value.

Oxygen Uptake Rate (OUR)

In their general form, the mass balances for oxygen in the gas and liquid phases include accumulation terms for oxygen. However, oxygen is sparingly soluble in aqueous solutions and usually kept constant by a controller. Therefore accumulation of the dissolved oxygen concentration will be very small and makes the OUR practically equal to the transfer rate of oxygen over the gas-liquid interface. As a consequence the OUR can be deduced directly from gas and liquid analyses only:

$$OUR = k_L^O a \left(\frac{p - p_w}{H_O} x_O - C_O \right) \quad (2)$$

For this expression it is assumed that the gas and the liquid phase are sufficiently mixed and that the gas flow rate is high so that $x_O^{in} \approx x_O^{out}$. x_O represents the fraction of oxygen in the headspace of the bioreactor, C_O the oxygen concentration in the liquid phase, p the headspace pressure and p_w the partial pressure of water vapour, which has to be taken into account because the oxygen fraction is expressed on a dry basis. H_O is Henry's constant for oxygen. The mass transfer coefficient for oxygen transfer, $k_L^O a$, can be experimentally determined using a dynamic method as described by Van 't Riet [8].

Carbon dioxide equilibrium in the gas phase

The transfer of carbon dioxide across the gas-liquid interface is a physical process (as for oxygen), which is liquid-film limited:

$$CTR = -k_L^C a \left(\frac{p - p_w}{H_C} x_C - C_C \right) \quad (3)$$

Where x_C is the fraction of carbon dioxide in the headspace, C_C is the carbon dioxide concentration in the liquid phase, and H_C the Henry's constant for carbon dioxide. In this expression, it is assumed that both the gas-phase and the liquid-phase are well mixed. The mass transfer coefficient for carbon dioxide, $k_L^C a$, is difficult to determine by the dynamic method. Transfer kinetics could be influenced by pH control and pCO₂ electrodes suffer from poor response times. However, the ratio of $k_L a$ values for CO₂ and O₂ is proportional to the ratio of their liquid phase diffusivities, and thus of the square root of their mole mass:

$$\frac{k_L^O a}{k_L^C a} \cong \frac{D_O}{D_C} \cong \sqrt{\frac{M_O}{M_C}} = 0.89 \quad (4)$$

Since the $k_L^O a$ can be determined experimentally as mentioned in the previous paragraph and the mole masses are known, the $k_L^C a$ can be calculated. The total mass balance for the CO₂ concentration in the headspace is given by

$$\frac{dx_C}{dt} = -\frac{F_G R T x_C}{V_H (p - p_w)} - k_L^C a \left(\frac{p - p_w}{H_C} x_C - C_C \right) \frac{V_L R T}{V_H (p - p_w)} \quad (5)$$

With F_G as the gas flow, assuming absence of carbon dioxide in the incoming gas. And R is the gas constant, T the temperature and V_H the headspace volume.

Carbon dioxide equilibrium in the liquid phase

In contrast to the assumptions for the *OUR* there is a discrepancy between the carbon dioxide transfer rate (*CTR*) across the gas-liquid interface, available from gas analyses, and the carbon dioxide evolution rate (*CER*) of the biomass in the bioreactor. The *CER* cannot be measured directly as carbon dioxide has a much higher solubility than oxygen that is enhanced by its hydrolysis to bicarbonate. Changes in the concentrations of dissolved carbon dioxide and bicarbonate results in differences between the *CTR* and *CER* [2], which increase with the bicarbonate buffering capacity of the medium. By measuring carbon dioxide in the off-gas of the bioreactor, a substantial amount of carbon

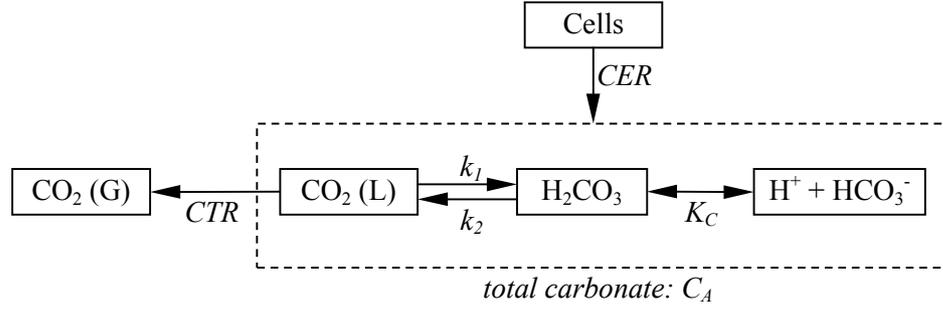


Figure 1: Diagram of the CO₂ balance in a cell culture. Showing the rate constants k_1 and k_2 and the dissociation constant K_C . All components in the dashed area are lumped to C_A , which is called the 'total carbonate concentration'.

dioxide coming from the buffer system will inevitably be measured, and not only the carbon dioxide produced by the cells [1]. Figure 1 shows the diffusion of carbon dioxide through the cell membrane and the subsequently following reactions with k_1 and k_2 as the rate constants for the indicated reactions and K_C is the carbonic acid dissociation constant. Further dissociation, of bicarbonate into carbonate is negligible for the pH range used during standard cell cultures (pH < 7.8). The concentration of carbonic acid is always very small in comparison to that of dissolved carbon dioxide, C_C . The mass balance for carbon dioxide and bicarbonate is:

$$\left(\frac{dC_C}{dt} + \frac{dC_H}{dt} \right) = CER - CTR \quad (6)$$

with C_C the concentration of carbon dioxide and C_H the concentration bicarbonate, $[HCO_3^-]$. As there is no sink for bicarbonate ions other than by dehydration, the time scale of changes in cell culture is long enough to ensure that the reactions involved in the dehydration of bicarbonate are close to equilibrium. The rate constants associated with carbonic acid dissociation are so large that this reaction can be considered to be at equilibrium.

During batch cultivations, the pH is controlled closely to the desired set-point. Therefore it is not necessary to model both the carbon dioxide and the bicarbonate concentration separately, because carbon dioxide and bicarbonate are in equilibrium; the dissociation constant for hydrolysis of bicarbonate is given by:

$$K_A = \frac{C_H \times 10^{-pH}}{C_C} \quad (7)$$

Now, a lumped variable C_A , called ‘total carbonate concentration’, can be introduced:

$$C_A = C_H + C_C \quad (8)$$

and thus the carbon dioxide concentration in the liquid is given by:

$$C_C = \frac{C_A}{1 + \frac{K_A}{10^{-pH}}} \quad (9)$$

The dashed area in Figure 1 represents the lumped concentration. Combining Equations (3), (6), (8) and (9) gives the mass balance for the ‘total carbonate concentration’:

$$\frac{dC_A}{dt} = CER + k_L^C a \left(\frac{p - p_w}{H_C} x_C - \frac{C_A}{1 + \frac{K_A}{10^{-pH}}} \right) - \frac{F_S}{V_L} C_A \quad (10)$$

The last term in this equation, with sample flow rate F_S , is inserted to take the loss of C_A due to sampling in account. Equations (5) and (10) represent the complete carbon dioxide mass balance for liquid and gas phase. It must be noticed that information about cell kinetics is not necessary and for the medium only known physical constants like H_C , K_A and R , are used.

3.5 Software sensor design

The software sensor is not based upon an equilibrium model ($CTR = CER$) but on a dynamic model for the carbon dioxide concentrations in headspace and medium, Equations (5) and (10). This dynamic model is used to reconstruct the actual CER from the measured carbon dioxide in the headspace.

Dynamic model

In systems theory there are two concepts, which must be clearly distinguished from each other. The plant is the actual physical system that needs to be observed. The model is the mathematical description of the physical system, which is used for the filter design stage. Within the control-engineering field, it is common to rewrite a model to its state-space representation. In such a

presentation all relations are linear or linearised and written in matrix-notation. The input, state, output, and noise of the model are given as vectors. The input vector usually contains the manipulable variables of the model, the state vector consists of those variables that develop through time dependant of each other and of the input, and the output usually contains all the measured variables. The total system can be written as an *ABCD*-system:

$$\begin{aligned}\frac{dx}{dt} &= Ax + Bu \\ y &= Cx + Du\end{aligned}\tag{11}$$

Where u is the input vector with m inputs, x the state vector with n states and y the output vector with k outputs. The matrices A and B contain information about the way the states develop through time and with the matrices C and D the output can be calculated. A is a $n \cdot n$, B a $n \cdot m$, C a $k \cdot n$ and D a $k \cdot m$ matrix.

In this situation the model, consisting of Equations 4 and 9, are extended with an extra differential equation used for estimating the *CER*. The prediction of the *CER* is a zero mean random walk process, i.e. its derivative is zero:

$$\frac{dCER}{dt} = 0\tag{12}$$

For use in a discrete Kalman Filter algorithm the model is discretised and rewritten to a state-space representation by defining the input, state and output vectors. The state vector consists of the *CER*, the lumped concentration of both dissolved carbon dioxide and bicarbonate (C_A) and the molar fraction of carbon dioxide in the headspace (x_C). There is no input vector and the output vector is constructed from the measured variables, which is only the molar fraction of carbon dioxide in the off-gas.

This model is linear in the state variables for the equations so the *ABCD*-matrices can be developed straightforward. A consequence of computer application for process monitoring is data sampling. Only process values at discrete time moments are available. Therefore, instead of continuous time models, the process behaviour is given by the discrete time equivalents. Moreover, since there is no input, there are no B and D matrices. Now the state vector, A matrix and C vector are:

$$\begin{aligned}
 x &= [C_A \quad x_C \quad CER]^T \\
 A &= \begin{bmatrix} 1 - \frac{k_L^C a \Delta t}{1 + \frac{K_A}{10^{-pH}}} - \frac{F_S \Delta t}{V_L} & \frac{k_L^C a (p - p_w) \Delta t}{H_C} & \Delta t \\ \frac{k_L^C a V_L R T \Delta t}{(p - p_w) V_H (1 + \frac{K_A}{10^{-pH}})} & 1 - \frac{F_G R T \Delta t}{(p - p_w) V_H} - \frac{k_L^C a V_L R T \Delta t}{H_C V_H} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
 C &= [0 \quad 1 \quad 0]
 \end{aligned} \tag{13}$$

The Kalman Filter algorithm

Both the plant and the model use the same known input. The state of the plant is, of course, unknown and its output can be measured. The state and output of the model is based on the predictions of the model, as can be seen in Table 1, where w_k and v_k represent process noise and measurement noise vectors, respectively, they are assumed to be random with mean values zero and variances Q and R .

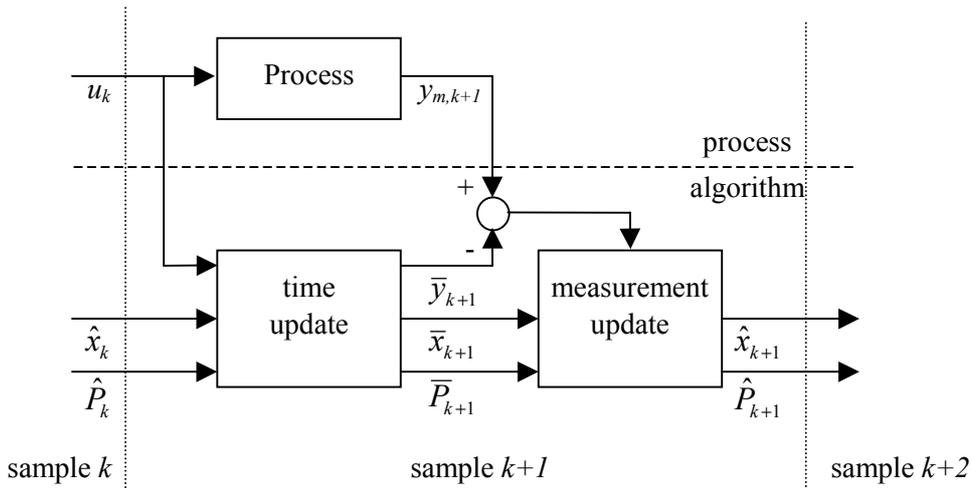


Figure 2: Schematic representation of the Kalman filter calculations. It shows the different steps to be taken at each time instant k . The dashed horizontal line makes distinction between the algorithm and process, the dotted vertical lines symbolise the successive sample moments.

The Kalman Filtering algorithm is depicted in Figure 2. In this figure, a distinction is made between the cultivation process that gives the data and the algorithm that processes the data. The algorithm has two steps, respectively the time update and the measurement update. In the time update a one sample ahead prediction is made for the state and output variables (respectively $\bar{x}_{k+1}, \bar{y}_{k+1}$) and the prediction variance of the states (\bar{P}_{k+1}). The actual values of the input variables (u_k) and the available estimated results of the Kalman filter (\hat{x}_k, \hat{P}_k) at sample moment k are used for this prediction.

The next step is the measurement update, which takes place as soon as new data becomes available and where the prediction of the output variable (\bar{y}_{k+1}) is corrected. The correction is proportional to the error between measured process output ($y_{m,k+1}$) and predicted output (\bar{y}_{k+1}). The magnitude of the correction gain varies for successive samples and aims to minimise the error covariance (\hat{P}_k) for the state and output variables. Now, the measurement update gives the best estimate (\hat{x}_{k+1}) of the states and its variance (\hat{P}_{k+1}). The values for the state estimate (\hat{x}_{k+1}) are used as the software sensor output. The entire set of these equations comprises what is called the discrete time Kalman Filter [9, 10] and is summarised in Table 1.

Table 1: Discrete time Kalman Filter algorithm

System and measurement model	
$x_{k+1} = A_k x_k + B_k u_k + w_k$	$w_k \sim (0, Q_k)$
$y_k = C_k x_k + v_k$	$v_k \sim (0, R_k)$
Initialisation	
$P_0 = P_{x_0}$	$\hat{x}_0 = \bar{x}_0$
Time update (one sample-ahead prediction)	
$\bar{x}_{k+1} = A_k \hat{x}_k + B_k u_k$	
$\bar{y}_{k+1} = C_k \bar{x}_{k+1}$	
$\bar{P}_{k+1} = A_k \hat{P}_k A_k^T + Q_k$	
Measurement update (after including the measurement)	
$\hat{P}_{k+1} = \bar{P}_{k+1} - \bar{P}_{k+1} C_{k+1}^T R_{k+1}^{-1} C_{k+1} \bar{P}_{k+1}$	
$K_k = \hat{P}_{k+1} C_{k+1}^T R_{k+1}^{-1}$	
$\hat{x}_{k+1} = \bar{x}_{k+1} + K_k (y_{m,k+1} - \bar{y}_{k+1})$	

A few remarkable advantageous features of the Kalman filtering algorithm can easily be observed. First, each calculation step only requires the last estimate and a new set of measurement data. The essential advantage of such simple "step-by-step" structure of the computational scheme is that there is no need to store all old results and measurement data for each up-dating state estimate, and this saves computer memory and processor time, especially in real-time (on-line) applications. Second, all recursive formulas of the algorithm are straightforward and linear, consisting of only matrix multiplication and addition, and a single matrix inversion in the calculation of the Kalman gain.

3.6 Application of the software sensor

The Kalman Filtering algorithm is used as software sensor for the *CER*. First the performance was validated by an experiment, then the application and use was proven by a series of experiments and finally two experiments show the robustness of the software sensor in coping with disturbances. Table 2 shows the values for the experimental conditions and used parameters.

Validation of the software sensor

Before applying the software-sensor, its performance needs to be validated. The bottleneck in the *RQ*-estimation procedure is that the software sensor must be able to detect the *CER* correctly. To validate the sensor at this point medium, to which a known amount of bicarbonate was added, was pumped

Table 2: Experimental conditions and parameter values

Parameter	value
F_G	0.004 mol·min ⁻¹ ·cells ⁻¹ 0.010 mol·min ⁻¹ (sensor validation)
H_C	3.137×10 ⁶ Pa·l·mol ⁻¹ at 28°C
K_A	4.42×10 ⁷ mol·l ⁻¹
$k_L^C a$	0.0206 min ⁻¹
$k_L^O a$	0.0231 min ⁻¹
$p-p_w$	101325 Pa
R	8.3143 J·mol ⁻¹ ·K ⁻¹
T	28 and 35°C
V_H	0.825 l
V_L	1 l

from a storage vessel into the reactor. So the carbon production by cells was imitated by pumping an exactly known amount of bicarbonate into the reactor. If the software sensor operates well, it should be able to detect this amount of bicarbonate correctly. Moreover, as the software sensor gives also the value of the total carbonate content (C_A), the performance can also be checked by analysis of the total carbonate concentration in the solution.

During the experiment, there was a minimum of headspace in the storage vessel, which limits the loss of bicarbonate from the liquid phase. Consequently, the total carbonate concentration in the feed to the reactor remained constant. Feeding of the bicarbonate solution to the fixed volume reactor by a continuous pumping system imitated the CER by cells. From the measurements of the total carbonate concentration (C_A) in the feed and reactor together with the feed rate, the actual CER was calculated. During the experiment the feed flow rate was manipulated in order to obtain several levels for the CER and C_A .

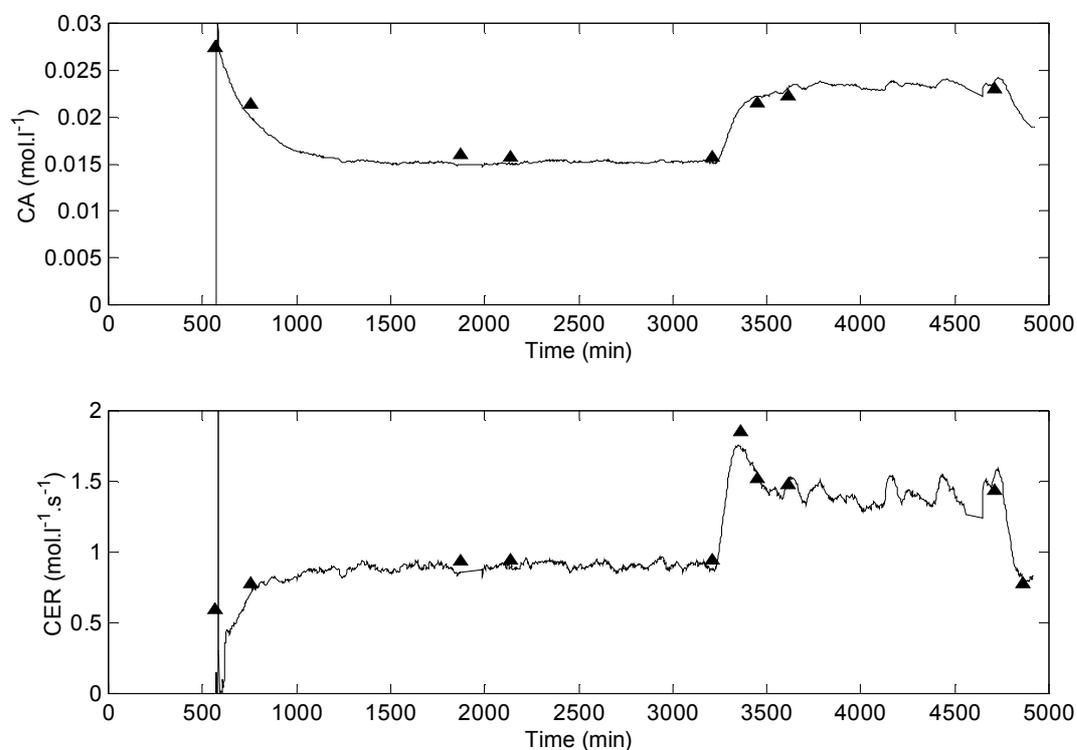


Figure 3: Validation of the software sensor, imitation of the CER of cells by feeding a given amount of bicarbonate to the reactor. A: Measured (\blacktriangle) and estimated (solid line) total carbonate concentrations, C_A . B: Actual (\blacktriangle) and estimated (solid line) CER .

Figure 3 shows the C_A and CER obtained by the software sensor compared with the measured C_A and the “imitated” CER (obtained by pumping the bicarbonate medium in the reactor). At 600 minutes the reactor contained an amount of bicarbonate and the feed flow and software sensor were started. The sensor results of the first five minutes are affected by the start-up of the sensor. The total carbonate concentration, C_A , in the reactor decreased during the first 500 minutes due to its high start concentration. At 3200 minutes, together with the feed flow rate, the CER was increased and at 4700 minutes, it was decreased again. During the whole experiment, there is a good match between the C_A and CER values obtained from the software sensor and the laboratory measurements. The CER signal from the software sensor in the second part of the experiment reveals some variations, which are a consequence of variations in the measured carbon dioxide concentrations in the off-gas. The good match between the obtained results confirms the use of the software sensor for C_A and CER measurement.

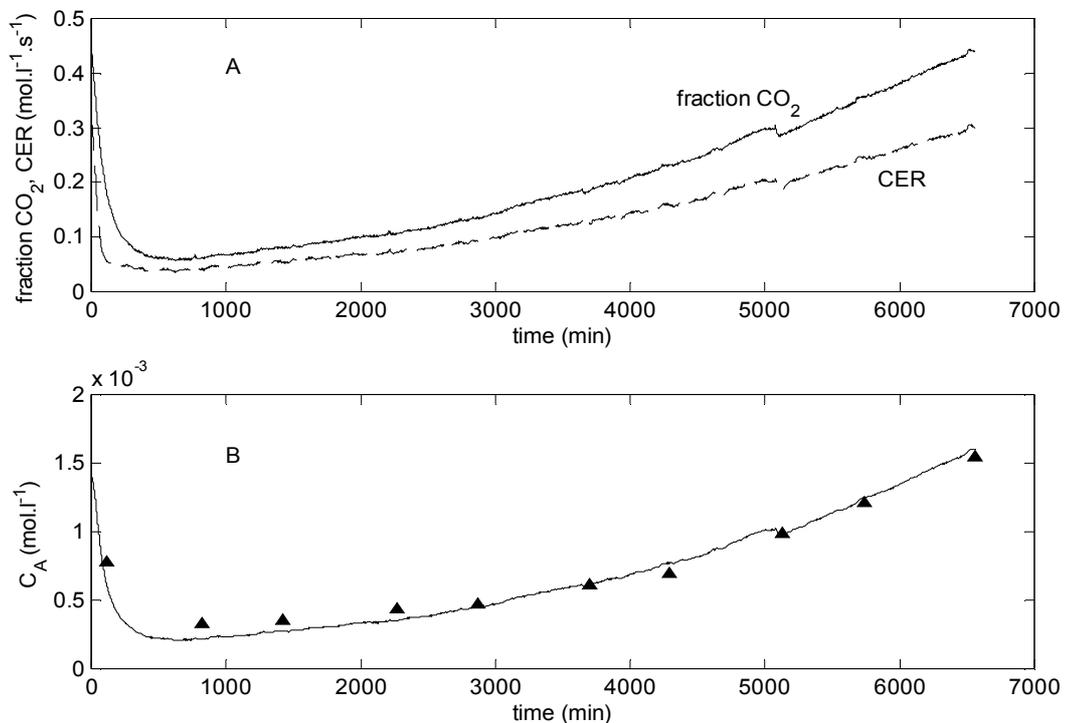


Figure 4: Batch-wise insect cell cultivation. A: Measured molar fraction of carbon dioxide in the off-gas (solid line) and the estimated CER (dashed line). B: Measured total carbonate C_A (▲) and the estimated total carbonate (solid line).

Application to cell cultivation

The on-line application of the *CER* software sensor was evaluated for 18 batch cultivations of insect cells and 2 mammalian cell cultivations. Figure 4A shows standard results for the measured carbon dioxide in the off-gas and the *CER*-values from the software sensor during one of the insect cell cultivations. The *CER*-values are in the range that can be expected for the used cells. Due to a re-calibration of the pH sensor at 5100 minutes, the pH controller reacted suddenly, which in turn affected the measured carbon dioxide and the estimated *CER*. In this experiment, together with the estimation of the *CER*, the C_A is estimated. Figure 4B shows the estimated C_A and the laboratory measurements, which correspond well. The mean error between the estimated and measured 'total carbonate concentration' is $0.014 \text{ mmol.l}^{-1}$ with a standard deviation of $0.066 \text{ mmol.l}^{-1}$. With the estimated *CER* and calculated *OUR* the *RQ* could be calculated.

Figure 5 shows the estimated *RQ* for a part of a batch together with the measured substrate concentrations: glutamine and glucose. The peaks in the *RQ* signal fall together with the sampling moments when some C_A containing medium is replaced by the same amount of C_A free medium. Figure 5 illustrates that when glucose and glutamine (till 4500 minutes) are in

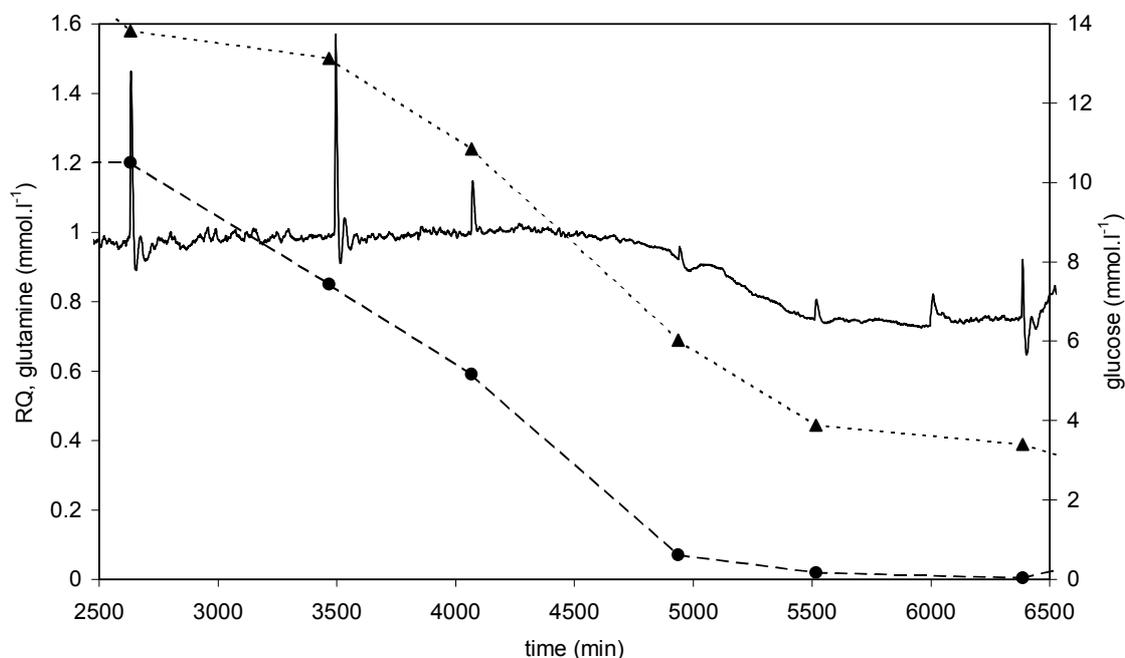


Figure 5: The estimated *RQ* (solid line) and the concentrations of glutamine (▲, left axis) and glucose (●, right axis) during batch wise cultivation of insect cells.

sufficiently high concentration available the organism grows well with a RQ around 1.0. After 5000 minutes, the glucose is depleted and now glutamine is the main substrate for the organism. The RQ -value that falls from 1.0 to 0.8 reflects the change in the physiological state of the organism due to the lack of glucose. Thus, with the ability to estimate the RQ on-line the physiological state of the cells can be monitored on-line and be used as an indicator for adding extra substrates. This finding confirms the statement of Royce [2] and Bonarius et al. [1] that the RQ can be used as parameter to detect the metabolic activity of the cell.

Robustness of the software sensor

Examples of RQ estimates for other cultures are given in figure 6. Figure 6A shows the estimated RQ during a batch where at 2700 minutes the communication between the computer and the PLC failed for approximately 1 hour. The Kalman filter waited 1 hour before a new measurement update could be calculated. After this disturbance, the software sensor quickly recovered the same RQ value. Figure 6B shows the estimated RQ during another batch where at 1800 and 2600 minutes huge samples were taken and

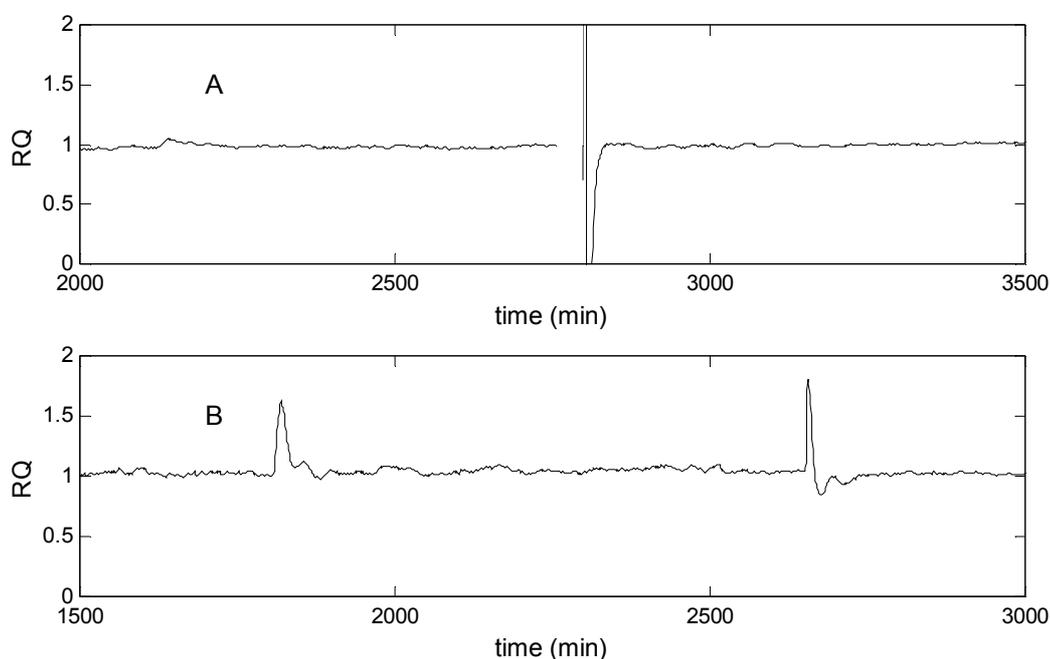


Figure 6: Two examples of estimated RQ values during batches where the Kalman filter had to cope with disturbances. A: The estimated RQ during a batch where at 2700 minutes the communication failed for 1 hour. B: The estimated RQ during a batch where at 1800 and 2600 minutes huge samples were replaced with the same amount of fresh medium.

replaced with fresh medium. Samples were taken by pressurising the headspace to withdraw the liquid. This way of sample taking gives a major impact on the headspace pressure and CO₂ fraction. Furthermore, the fresh medium instantly changes the 'total carbonate concentration' in the medium. Again it can be seen that the software sensor quickly recovers from these disturbances.

3.7 Concluding Remarks

For batch cell cultivation in media which buffer carbon dioxide, the carbon dioxide evolution rate (*CER*) and respiration quotient (*RQ*) cannot be obtained from an equilibrium model only. As an alternative, a software sensor is used. Main features of the sensor are:

- The *CER* detection is based on a dynamic model that covers the transient phases for bicarbonate buffering media and batch wise operations. In the model all carbon dioxide and bicarbonate in the liquid phase are lumped to a single component "total carbonate" and as a result no detailed knowledge of the reaction kinetics is necessary; only physical constants are required.
- The sensor uses a discrete time Kalman Filter algorithm.
- In a validation experiment where cell carbon dioxide production was imitated, the software sensor proved to be successful for estimation of the *CER* from on-line carbon dioxide measurements in the off-gas.
- Results obtained from experiments with insect cells showed that the *RQ*-value was close to 1.0, a value that is common for cells growing on glucose and glutamine as main substrates. As the physiological state of the organism changes due to substrate limitations, the *RQ*-value reflects this change.
- Results of experiments where disturbances occurred showed that the Kalman Filter could cope efficiently with communication failures and sudden medium or headspace changes.

Important difference with the work of Stephanopoulos and San [6] and Bastin and Dochain [7] is that a minimum of a priori system knowledge is needed for the estimation of the *RQ* value. In contrast to a limited amount of information

that can be obtained from laboratory samples the software sensor produces a continuous stream of RQ and CER values to monitor the state of the organism. Therefore, it is a powerful tool to be used for various cell cultivations.

3.8 Nomenclature

Model

C_A	$\text{mol}\cdot\text{l}^{-1}$	‘total carbonate concentration’
C_C	$\text{mol}\cdot\text{l}^{-1}$	carbon dioxide concentration in the liquid
C_H	$\text{mol}\cdot\text{l}^{-1}$	bicarbonate concentration
C_O	$\text{mol}\cdot\text{l}^{-1}$	oxygen concentration in the liquid
CER	$\text{mol}\cdot\text{l}^{-1}\cdot\text{min}^{-1}$	carbon dioxide evolution rate
CTR	$\text{mol}\cdot\text{l}^{-1}\cdot\text{min}^{-1}$	carbon dioxide transfer rate
F_G	$\text{mol}\cdot\text{min}^{-1}$	gas flow rate
F_S	$\text{l}\cdot\text{min}^{-1}$	liquid sample flow rate
H_C	$\text{Pa}\cdot\text{l}\cdot\text{mol}^{-1}$	Henry's constant for carbon dioxide
H_O	$\text{Pa}\cdot\text{l}\cdot\text{mol}^{-1}$	Henry's constant for oxygen
K_A	$\text{mol}\cdot\text{l}^{-1}$	dissociation constant for bicarbonate hydrolysis
K_C	$\text{mol}\cdot\text{l}^{-1}$	carbonic acid dissociation constant
$k_{\#}$	min^{-1}	rate constant of reaction #
$k_L^C a$	min^{-1}	mass transfer coefficient for carbon dioxide
$k_L^O a$	min^{-1}	mass transfer coefficient for oxygen
OUR	$\text{mol}\cdot\text{l}^{-1}\cdot\text{min}^{-1}$	oxygen uptake rate
p	Pa	pressure
p_w	Pa	partial pressure of water vapour
pH	-	pH in the bioreactor
R	$\text{kJ}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$	molar gas constant
RQ	-	respiration quotient
T	K	temperature
V	l	volume
x_O	-	fraction of oxygen in the headspace
x_C	-	fraction of carbon dioxide in the headspace

Kalman filter algorithm

A, B, C	system matrices
G	selection matrix for state variance
K	Kalman gain
k	sample moment
P	state variance matrix
Q	variance matrix of state variables
R	variance matrix output variables
u	input variable vector
v	output uncertainty
w	process uncertainty
x	state variable vector
y	output variable vector
y_m	measured process output

3.9 Reference

1. Bonarius, H.P.J., C.D. de Gooijer, J. Tramper and G. Schmid (1995). Determination of the Respiration Quotient in Mammalian Cell Culture in Bicarbonate Buffered Media. *Biotechnology and Bioengineering* 45: 524-535.
2. Royce, P.N. (1992). Effects of Changes in the pH and Carbon Dioxide Evolution Rate on the Measured Respiratory Quotient of Fermentations. *Biotechnology and Bioengineering* 40: 1129-1138.
3. Vallino, J.J. and G. Stephanopoulos (1993). Metabolic flux distributions in *Corynebacterium glytamicum* during growth and lysine overproduction. *Biotechnology and Bioengineering* 41: 633-646.
4. Wang, H.Y., C.L. Cooney and D.I.C. Wang (1979). Computer control of bakers' yeast production. *Biotechnology and Bioengineering* 21: 975-995.
5. Wu, W.-T. and P.-M. Wang (1993). On-line optimal control for ethanol production. *Journal of Biotechnology* 29: 257-266.

6. Stephanopoulos, G. and K.Y. San (1984). Studies on On-Line Bioreactor Identification. I. Theory. *Biotechnology and Bioengineering* 26: 1176-1188.
7. Bastin, G. and D. Dochain (1990). *On-line Estimation and Adaptive Control of Bioreactors*. Amsterdam, Elsevier Science Publishers.
8. Van 't Riet, K. and J. Tramper (1991). *Basic Bioreactor Design*. New York, Marcel Dekker.
9. Chen, G. (1994). *Approximate Kalman Filtering*. Singapore, World Scientific.
10. Lewis, F.L. (1986). *Optimal Estimation*. New York, USA, John Wiley and Sons.

4 ESTIMATION OF SPECIFIC GROWTH RATE FROM CELL DENSITY MEASUREMENTS

This chapter is based on:

R. Neeleman and A.J.B. van Boxtel (2001). Estimation of specific growth rate from cell density measurements. *Bioprocess and Biosystems Engineering*, 24(3), 179-185.

4.1 Abstract

For modelling purposes it is of great importance to derive the specific growth rate as a function of time from biomass measurements. Traditional methods such as exponential or polynomial fitting of biomass measurements do not give satisfactory results nor do these methods take the noise characteristics of the biomass measurements into account. Standard recursive techniques, such as Kalman filtering, use only the data up to the time under consideration and are dependent of a good initial estimation. This paper describes a technique based on combining subsequent backward and forward extended Kalman filtering to give a smoothing estimator for the specific growth rate. The estimator does not need an initial value and is shown to have a single tuning parameter. The applicability of the estimator is demonstrated on batch and fed-batch cultivations of two organisms: *Bordetella pertussis* and *Neisseria meningitidis*.

4.2 Introduction

The specific growth rate is one of the most important parameters in biotechnological processes and the relationships between the rates of growth, substrate consumption, and product formation are crucial for monitoring and controlling these processes. Common practice in specific growth rate determination is fitting exponential or polynomial functions through the biomass measurement results. Apart from the unknown order of the polynomial function, the main disadvantage of such an approach is the assumption that the specific growth rate remains constant or follows a polynomial pattern. In batch and fed-batch processes, substrate and biomass concentrations change and as a consequence the specific growth rate is subject to changes. Therefore, it is of importance to be able to monitor the specific growth rate as a function of time without a priori knowledge of the order in which it changes. Especially when developing a structured model and for microorganisms in a multi-substrate or multi-limiting environment such an impartial specific growth rate determination is of great importance.

A great number of applications of recursive estimation of biological processes are reported [1-9]. Several different techniques are used amongst which are the extended Kalman filter and the asymptotic observer. These techniques have in common that they are intended as software sensors for online use. Each

estimate is based only on the measurements up to the time under consideration and as a result the estimators react too late to changes.

For model development where the estimation of the specific growth rate is done offline, these recursive techniques use only measurements from the past and thus do not exploit all available data. This paper describes a smoothing technique based on the extended Kalman filter [10, 11], where all data, both past and future, are taken into account for estimating the specific growth rate. A typical feature of this smoother is that the initial-value problem of the general Kalman filter is cancelled. Furthermore, for estimation of the specific growth rate, the problem of tuning is reduced to a single scalar parameter.

4.3 Materials and Methods

Strains and culture media

Bordetella pertussis strain 509, which is one of the two strains included in the DPT-Polio vaccine applied in the Netherlands, was used with the chemically defined medium containing glutamate and lactate as main carbon substrates [12]. For the experiments using *Neisseria meningitidis* strain JB16215 on, glucose and glutamate were the main carbon substrates in the chemically defined medium.

Bioreactor conditions

The cells were grown in a 5-litre round-bottom reactor containing 3-litre medium. An eight-bladed marine impeller was used to agitate the medium. Temperature, pH, dissolved oxygen and stirrer speed were controlled at 34°C, 7.2, 20% and 400 rpm, respectively, for *B. pertussis* and 35°C, 7.0, 30% and 600 rpm, respectively, for *N. meningitidis*. Oxygen was transported through the headspace only and controlled by changing the oxygen fraction in the gas flow. The total gas flow was kept constant at 1-l·min⁻¹.

Analysis

A polarographic electrode (Ingold, Urdorf, Switzerland) was used to measure the dissolved oxygen in the medium. A pH electrode (Ingold) was used to measure the pH and the temperature was measured with a Pt100 temperature sensor. Biomass was measured by optical density using a Vitalab 10 (Vital

Scientific, Dieren, the Netherlands) at 590 nm. The optical densities of *B. pertussis* were recalculated to $\text{mmol}\cdot\text{l}^{-1}$ according to a calibration curve. Glutamate and L-lactate were determined with a YSI 2750 Select analyser (Yellow Springs Instruments, Yellow Springs, USA).

Software

Matlab and Simulink (Mathworks, Mass., USA) were used for the development of the forward and backward extended Kalman filter algorithms, the smoothing algorithm, the simulation of microbial growth, and for the visualisation of the results. All the used algorithms can be downloaded freely at <http://www.aenf.wau.nl/mrs/staff/neeleman/estmu>.

4.4 Methods for growth rate estimation

Traditional methods

For model development one of the main issues is to reconstruct the time function of the specific growth rate, $\mu(t)$ from biomass measurements, $C_X(t)$. Besides dividing the total growth rate by the average biomass concentration between two measurements, biotechnology textbooks, e.g. [13], present the following frequently used function to calculate the specific growth rate between two consecutive biomass measurements:

$$\mu = \frac{\ln\left(\frac{C_{X,k+1}}{C_{X,k}}\right)}{t_{k+1} - t_k} \quad (1)$$

With this approach, the specific growth rate is calculated over each time interval; however, with a decreasing time interval, noise on the biomass measurements plays an increasing role. Other approaches seen regularly are fitting exponential curves or sigmoids. Main disadvantage of fitting such equations through data is the assumption that the specific growth rate is constant and so these methods are not appropriate to determine changes. Polynomials or splines solve that problem, but fix the specific growth rate to change according to a polynomial of a lower degree. The order of this polynomial function is a crucial and unknown parameter. The main disadvantage of all these traditional techniques is that they consider each

measurement to be exact by disregarding the accuracy and noise characteristics of the biomass measurements.

Recursive estimation

Several authors described recursive estimators; Stephanopoulos and San reported on the subject [8, 9] and Claes and Van Impe [3] validated the performance of an observer online. Such a recursive estimator uses a model to reconstruct the state when only part of the state is measured. The model necessary for specific growth rate estimation describes changes of biomass as directly related to the biomass concentration, net specific growth rate and the dilution rate, D :

$$\frac{dC_X}{dt} = \mu C_X - D C_X \quad (2)$$

The behaviour of the specific growth rate during cultivation is not known a priori, and so the change of the specific growth rate is assumed to be a random walk process:

$$\frac{d\mu}{dt} = 0 + z \quad z \sim (0, \gamma) \quad (3)$$

With z as a white-noise process with zero mean and spectral density γ . The expected variability of the specific growth rate in time is reflected by the choice of the variance of γ in the algorithm. Transferring this model, Eq. (2) and Eq. (3), into state representation for the estimator, using the state vector $x = [C_X \ \mu]^T$, input $u = [D]$ and output $y = C_X$:

$$\begin{aligned} \frac{dx}{dt} &= f(x, u) + w & w &\sim (0, Q^s) \\ y_k &= Hx_k + v_k & v_k &\sim (0, R_k) \end{aligned} \quad (4)$$

With f , the non-linear function of x and u . w is zero mean white noise with spectral density Q^s , y_k is the sampled output with H as the output matrix and v_k as zero mean white noise with covariance R_k . Biomass is determined by measuring the optical density of the cells and due to the limited range of the analyser, samples with high biomass concentration are diluted. This results in a standard deviation of the biomass measurements, which is proportional to biomass: $\sigma = \lambda \cdot C_X$. Thus, the covariance R_k of these measurements equals σ^2 . The parameter λ can easily be determined from actual measurements. Since all

unpredicted changes of the biomass have to be reflected by the specific growth rate, the model description for biomass growth is assumed perfect. Thus, the spectral density of the biomass prediction equals zero. The remaining single tuning parameter is γ .

$$Q^s = \begin{bmatrix} 0 & 0 \\ 0 & \gamma \end{bmatrix} \quad (5)$$

Forward Extended Kalman Filter (EKF)

The EKF algorithm, a frequently used recursive estimator [8, 9], is a two-step algorithm using a time-update and a measurement-update. The algorithm is given in detail in the appendix. Given the current estimate, \hat{x}_k , the time update predicts the state value at the next sample, \bar{x}_{k+1} , (one-step-ahead prediction) using Runge-Kutta integration. The prediction of the corresponding error covariance, \bar{P}_{k+1} , is calculated using the linearised model. The next step in the EKF algorithm, the measurements-update, consists of a Joseph stabilised approximation of the error covariance, \hat{P}_{k+1} , and the correction of the state based on the new measurement, y_{k+1} . The correction term is a function of the innovation, that is, the discrepancy between the measured and predicted values. The innovation gain, K , is chosen to minimise the covariance of the estimation error given the covariances Q_k and R_k .

So, in contrast to the traditional techniques, the Kalman filter algorithm takes the measurement error characteristics into account. The drawback of this algorithm is the unknown initial state \hat{x}_0 and the corresponding error covariance \hat{P}_0 . Furthermore, due to the restriction that the specific growth rate changes as a random walk process, it reacts too late to changes, as can be seen in Figure 1. This figure presents a simulated pattern together with the estimated specific growth rate and error covariance. For the simulation, a sample time of 1 hour was used and random Gaussian distributed noise was added to the biomass ‘measurements’. It shows that, although the initial state was far from the actual value, the EKF converged to the true state within two steps. This filter can be implemented online since it is based on the measurements up to and including the moment of estimation.

The performance of the filter depends on the single tuning parameter γ . High values result in a specific growth rate with high variability, while low values

ESTIMATION OF SPECIFIC GROWTH RATE FROM CELL DENSITY MEASUREMENTS

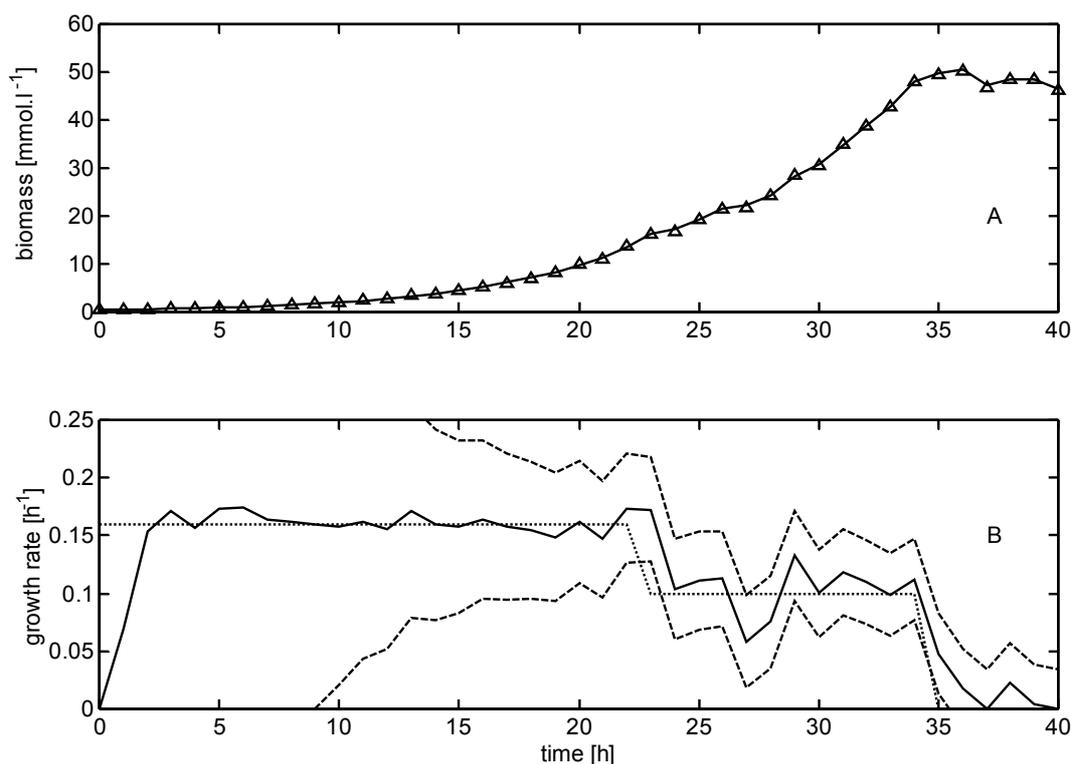


Figure 1: Demonstration of estimator results on simulated biomass measurements. A: simulated biomass ‘measurements’ (triangles) and estimated biomass (solid line). B: corresponding estimated specific growth rate (solid line) and variance (dashed line). Data were generated with a simulated specific growth rate pattern (dotted lines) and noise was added to the biomass ‘measurements’. In the period 22 to 23 hours the simulated specific growth rate changed from 0.16 to 0.10 h⁻¹.

result in a delayed response to changes. An intermediate between noisy and delayed estimation has to be determined by tuning.

Backward Extended Kalman Filter

For the above-mentioned forward EKF, the initial specific growth rate is unknown. As a result, during the first hours of the batch, the estimation is inaccurate and the corresponding error covariances are high. Since the final specific growth rate of a cultivation is usually known, this problem can also be approached in an opposite way. Standard batch and fed-batch processes are terminated when the substrates are depleted, the specific growth rate approaches zero and biomass no longer increases. At this moment, the state, x_{tf} , is most accurately known. Therefore, the estimator can be implemented to run backwards starting with a specific growth rate of zero (the algorithm is presented in detail in the appendix). Figure 2 shows the result of the backward

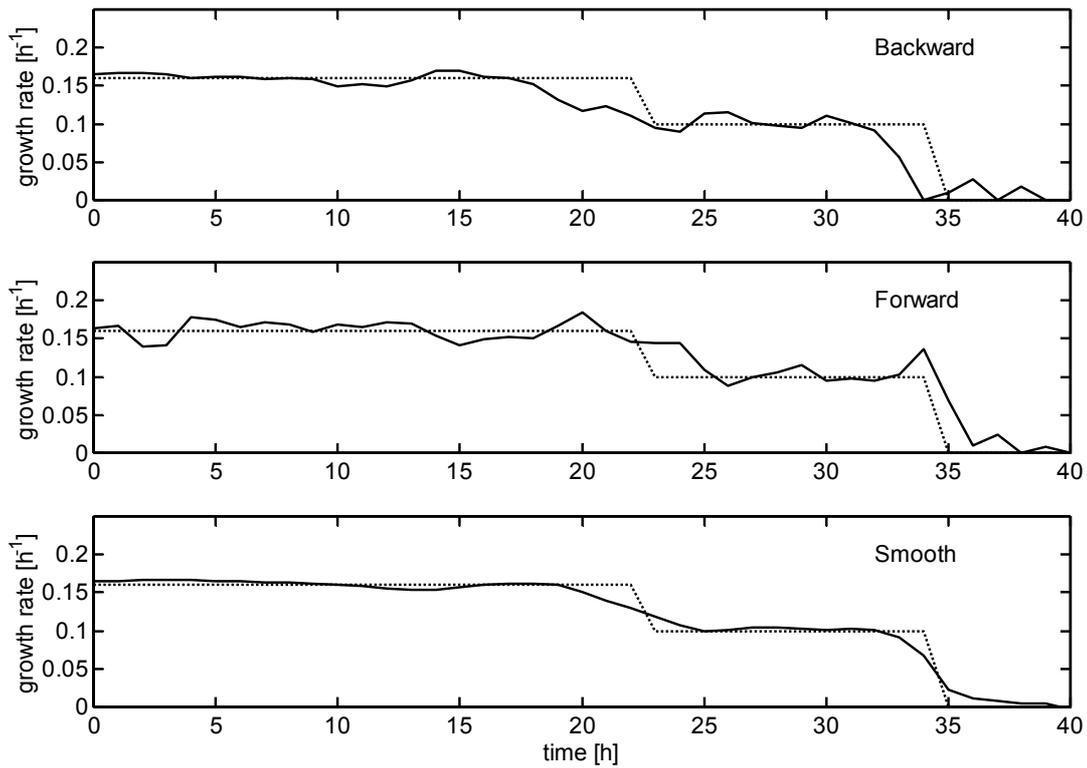


Figure 2: Specific growth rate estimation results (solid) with variances (dashed) of the backward (top), subsequent forward (middle) and smoothing (bottom) estimators. Data were generated with a simulated specific growth rate pattern (dotted) and noise was added to the biomass ‘measurements’.

estimation together with the simulated specific growth rate. Now, the initial state is chosen properly and the specific growth rate during the first hours of the batch is estimated correctly. The drawback of the backward filter is that it tends to overshoot in the other direction; seen in normal time direction, the changes in specific growth rate are detected too early.

Smoothing estimation

With the final estimation of the backward filter, $\hat{x}_b(t_0)$, a good estimation of the initial specific growth rate is made. This value is used by a subsequent forward EKF, avoiding the earlier mentioned initial-value problem. Figure 2 shows the estimates of the subsequent forward filter.

Combining the better of the two approaches gives a smoothing estimator [10, 11]. The smooth estimate $\hat{x}_{s,k}$ is obtained by averaging the backward estimates, $\hat{x}_{b,k}$, with the subsequent forward estimates, $\hat{x}_{f,k}$, weighted by their respective error covariances. The algorithm is described in detail in the

appendix. Now, each estimation incorporates all available data. Figure 2 shows the smooth estimate of the simulated specific growth rate. It can be seen that the variability of the smooth estimate is less than the variability of the forward and backward estimates. Furthermore, the timing of the smooth estimate is more accurate in detecting growth rate changes compared to the forward and backward filters.

Besides averaging states, the error covariances are also smoothed. Figure 3 shows the error covariances of the forward, backward and smooth estimations. Since each measurement increases the accuracy of the estimation, the error covariance of the smooth estimate, using all data, is smaller than or equal to the smallest error covariance of the other two estimates.

Both the forward and backward filters need an initial guess of the error covariance matrix, respectively P_0 and P_{tf} . Since high error covariances have a minor effect on the smoothed estimation, while low error covariances have a major effect, the initialisations are done based on rules of thumb and have small influence on the final smoothed estimate.

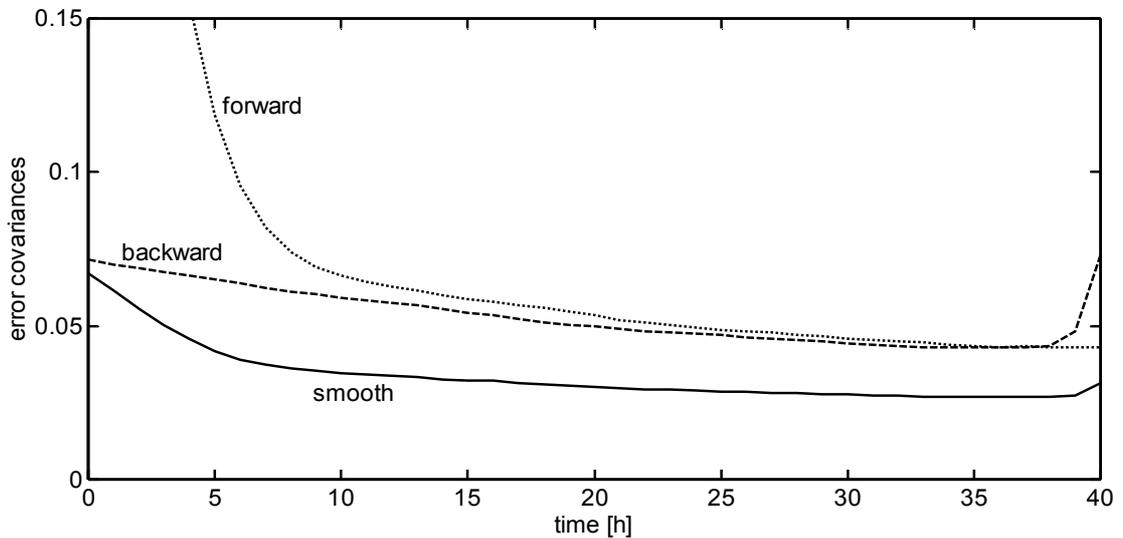


Figure 3: Forward (dotted), backward (dashed) and smooth (solid) error covariances for the specific growth rate of the smoothing estimator for a batch simulation.

4.5 Results and Discussion

Application to batch cultivation of *B. pertussis*

The smoothing estimator was used to estimate the specific growth rate of several experiments. Figure 4 shows the estimated specific growth rate during a cultivation of *B. pertussis*. The two substrates are depleted at 25 h and the specific growth rate during the unlimited part of the cultivation remained constant around 0.15 h^{-1} . Note that the substrate concentrations are not used by the smoothing estimator, but are shown to illustrate the specific growth rate dependency on the substrate concentrations.

Figure 5 shows a second cultivation of *B. pertussis*, where different starting concentrations of glutamate and lactate were used. Clearly, the specific growth rate of *B. pertussis* equals 0.15 h^{-1} when all substrates are present in excess.

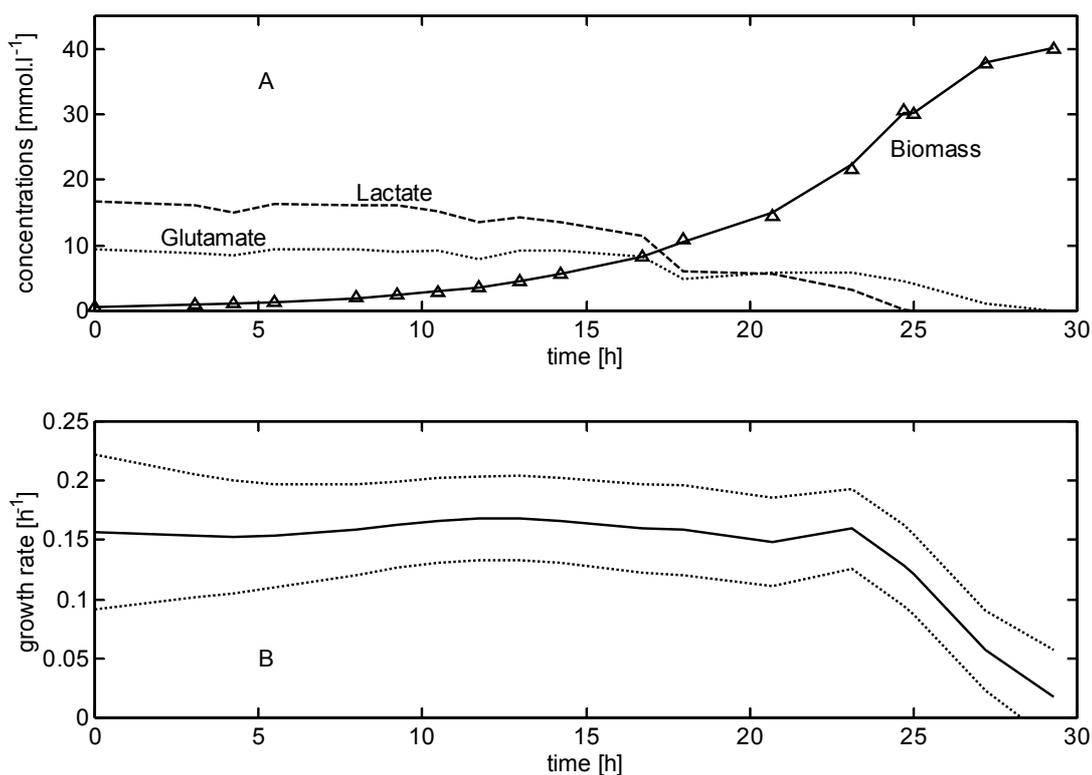


Figure 4: Estimation of the specific growth rate during a cultivation of *B. pertussis* where both substrates are depleted simultaneously. A: Biomass measurements (triangles), estimations (solid) and concentrations of glutamate (dotted) and lactate (dashed). B: Smooth estimation of the specific growth rate (solid) and variance (dotted).

ESTIMATION OF SPECIFIC GROWTH RATE FROM CELL DENSITY MEASUREMENTS

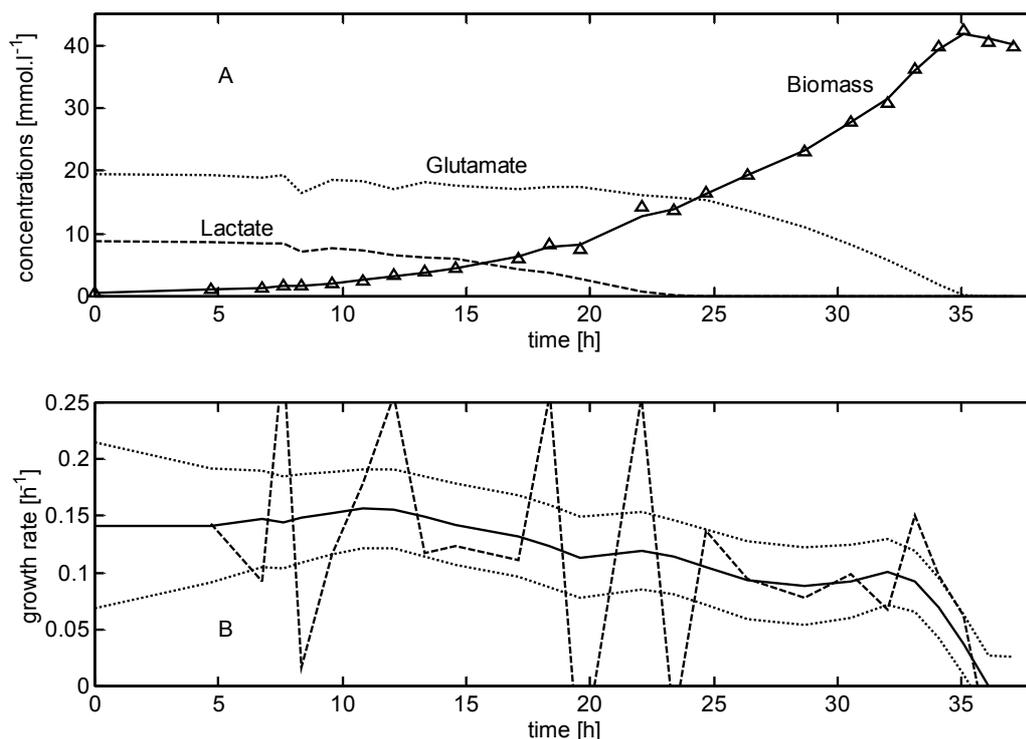


Figure 5 Estimation of the specific growth rate during a batch cultivation of *B. pertussis*. A: biomass measurements (triangles) and estimations (solid) and concentrations of glutamate (dotted) and lactate (dashed). B: smooth estimation of the specific growth rate (solid) and corresponding variance (dotted) and calculated by Eq. (1) of specific growth rate (dashed).

Also, when after 22 h lactate is depleted, the specific growth rate is at a lower level of 0.10 h^{-1} . Furthermore, the specific growth rate decreases towards zero at 35 h when the glutamate becomes depleted. Besides the estimated specific growth rate, the specific growth rate according to Eq. (1) is shown as well. Clearly, the traditional technique is noisier and not a good indicator for changes of the specific growth rate. Furthermore, experiments with lactate only (data not shown here) showed no growth at all. From these experiments, it is deduced that the specific growth rate of *B. pertussis* is enhanced by the presence of lactate. In this way, the smoothing specific growth rate estimator elucidated the model structure for *B. pertussis*.

Application to batch cultivation of *N. meningitidis*

Since the method is not specific to *B. pertussis*, the smoothing estimator can be used for various microorganisms, such as *N. meningitidis*. In this case, as there is a linear relationship between optical density and dry weight of the

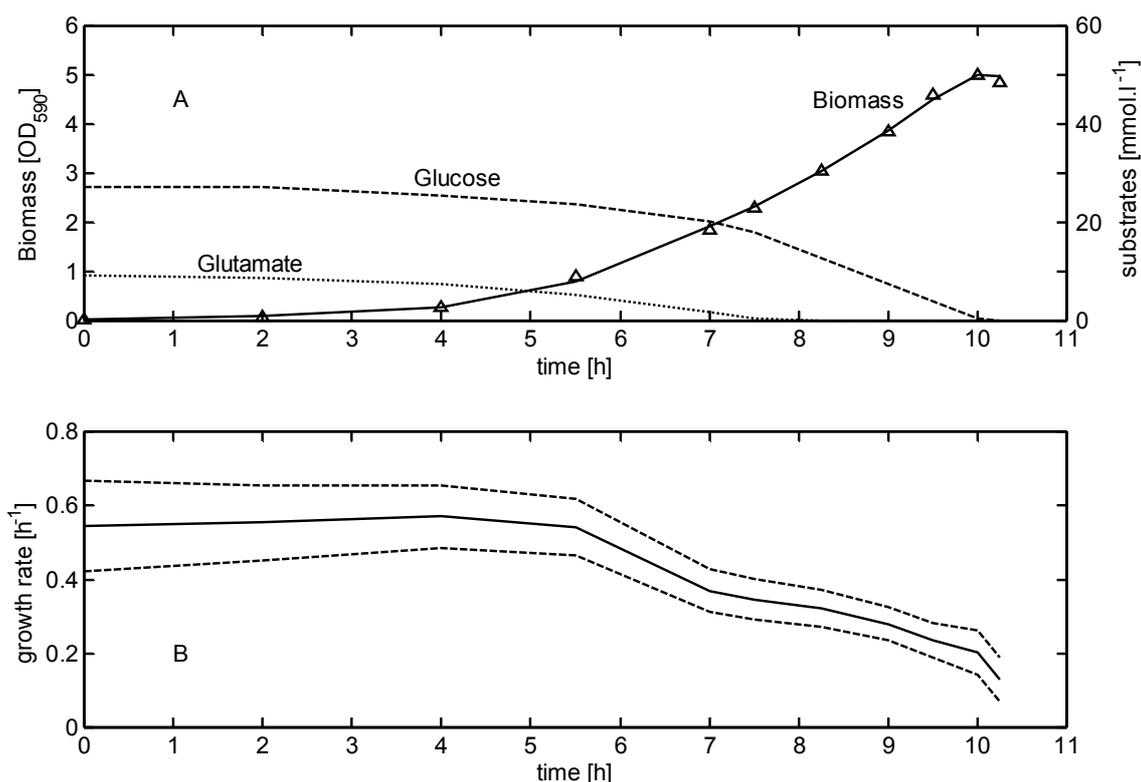


Figure 6: Estimation of the specific growth rate during a batch of *N. meningitidis*. A: optical density measurements (triangles) and estimations (solid) and concentrations of glutamate (dotted) and glucose (dashed). B: smooth estimation of the specific growth rate (solid) and variance (dotted).

cells, the optical density is used directly instead of the biomass concentration. Figure 6 shows the estimated specific growth rate during batch cultivation. For this cultivation, a chemically defined experimental medium was used with a lower concentration of glutamate. After 7 h, the glutamate became depleted and it can be seen that the specific growth rate decreased from 0.55 h^{-1} to 0.35 h^{-1} . The final specific growth rate could not be estimated properly due to the limited number of samples at the end of the growth phase. Clearly, the smoothing estimator can be applied to various microorganisms and biomass need not be converted to $\text{g}\cdot\text{l}^{-1}$ or $\text{mmol}\cdot\text{l}^{-1}$ when a linear relationship between optical density and dry weight is found.

Application to fed-batch cultivation

Besides estimation during batch cultivation, it is also of great interest to estimate the specific growth rate during fed-batch. Since the smoothing estimator is based on a model describing the effect of the dilution rate, it can be used for batch, fed-batch and continuous cultures. Figure 7 shows the

ESTIMATION OF SPECIFIC GROWTH RATE FROM CELL DENSITY MEASUREMENTS

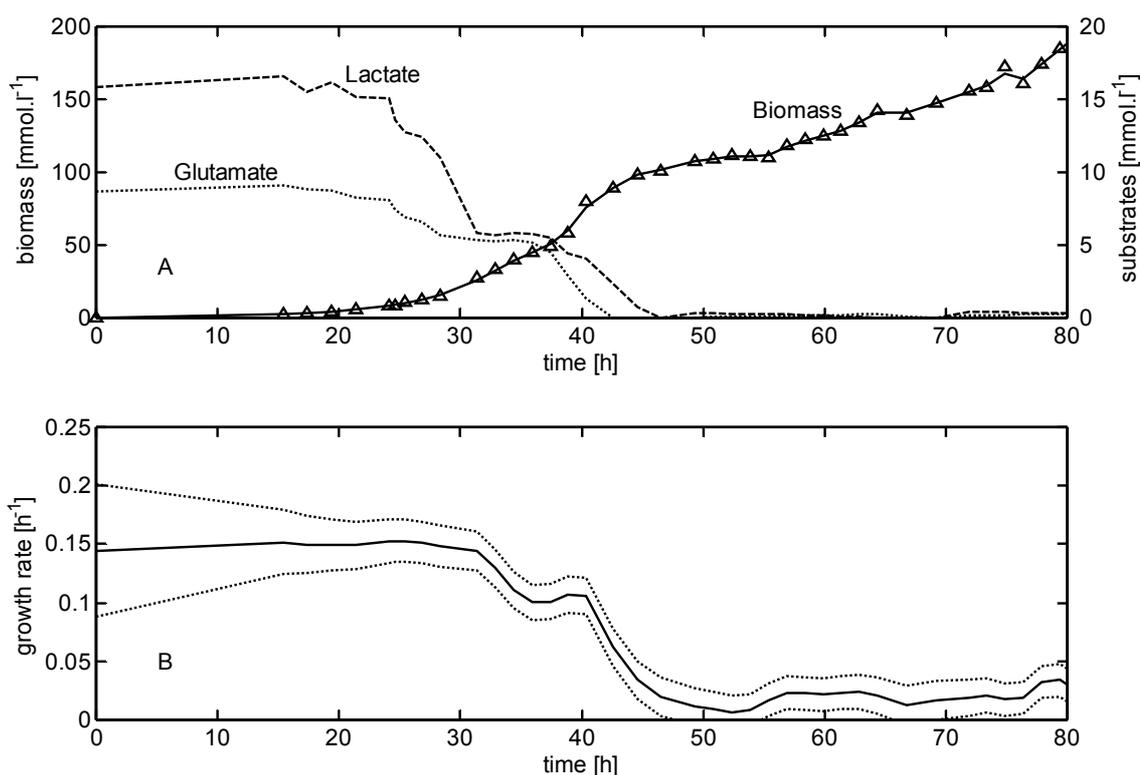


Figure 7 Estimation of the specific growth rate during a fed-batch cultivation of *B. pertussis*. A: biomass measurements (triangles) and estimations (solid) and concentrations of glutamate (dotted) and lactate (dashed) on the right axis. B: smooth estimation of the specific growth rate (solid) and variance (dotted).

estimated specific growth rate of a semi-fed-batch cultivation of *B. pertussis*. The first 32 h of the cultivation were batch, after which the feed flow rate was set constant at 12-ml·h⁻¹ containing 250 mmol·l⁻¹ glutamate and 415 mmol·l⁻¹ lactate. The volume was kept constant by sample taking and with a level pipe. The specific growth rate started at 0.15 h⁻¹ and became dependent on the substrate concentrations after 30 hours. The estimated specific growth rate reacted to this phenomenon by a decrease. At 32 h, the feed was started and the supply of substrates temporarily prevented further limitation. This resulted in a short period where the specific growth rate stabilised. After 40 h, the substrate consumption by the incremented biomass surmounts the feed and the substrate concentrations further decreased. From 40 h, the substrates became further limiting and the specific growth rate dropped to a lower nearly constant level after 50 h. Sudden changes in biomass at high cell densities are reflected by slight variations in the specific growth rate towards the end.

4.6 Conclusions

For offline model development, traditional methods such as fitting exponential or polynomial functions are not satisfactory, since they do not cope with dilution or with the characteristics of biomass measurement errors. Furthermore, it is shown that the specific growth rate of organisms is subject to change, while fitting an exponential curve over the complete cultivation interval assumes it to be constant. Calculating the specific growth rate in between consecutive biomass measurements results in noisy estimations at high sample rates.

There are many reports about recursive techniques, which are designed for online use and take into account the characteristics of measurement errors. These recursive techniques depend on a good guess of initial values and corresponding high error covariances and, when applied offline, not all data are used for each estimation.

In this paper, an offline smooth estimation algorithm for the specific growth rate is presented based on consecutive backward and forward extended Kalman filtering. The main features of this algorithm are the following:

- All data, both past and future, are taken into account for each estimation and, since the algorithm starts with the last known state, the estimations do not depend on an initial value.
- The model does not use organism-specific parameters and thus can be applied to various microorganisms.
- Contrary to the traditional techniques, biomass measurement error characteristics are not neglected.
- A single tuning parameter describes the trade-off between smoothness and adaptation. The initial guess of the error covariance matrix has a negligible effect on the final result.
- When a linear relationship between optical density and dry weight of the cells exists, the optical density can be used directly by the algorithm.

Application of the algorithm showed:

- When the specific growth rate in a multi-substrate environment changes, this technique accurately traces the changes in specific growth rate. Thus, making the smoothing estimator a powerful tool for model structure identification. It must be noted that the filter smoothes fast changes in growth rate unless a large number of samples are taken during the change.
- Two different micro-organisms were used to show the micro-organism independency of the algorithm.
- Experimental data showed the applicability of the algorithm for both batch and fed-batch operation.

Important differences with established recursive techniques are the lack of necessity for an initial estimation and the use of all data over the cultivation time interval.

4.7 Nomenclature

C_x	biomass concentration ($\text{mmol}\cdot\text{l}^{-1}$)
D	dilution rate (h^{-1})
F	linearised and discretised model f
G	selection matrix for state variance
H	output matrix
K	Kalman gain
k	sample moment
P	state variance matrix / error covariance
Q	spectral density / covariance of w
R	variance matrix of v
u	input variable vector
v	output uncertainty
w	process uncertainty
x	state variable vector
y	measured process output
μ	specific growth rate (h^{-1})

t	time (h)
z	white-noise
γ	spectral density of z

4.8 Appendix

Linearisation and discretisation of Eq. (4)

For both EKF's, the model $f(x,u)$ is linearised and discretised (using a first-order approximation) at sample time k to the matrix F_k :

$$F_k = I + \frac{df(x_k, u_k)}{dx_k} \Delta t$$

The covariance matrix Q_k equals the discretised spectral-density matrix Q^s evaluated at sample time k .

Forward EKF algorithm

Time update:

$$\begin{aligned}\bar{x}_{k+1} &= \hat{x}_k + \int_{t_k}^{t_{k+1}} f(\hat{x}, u) dt \\ \bar{P}_{k+1} &= F_k \hat{P}_k F_k^T + Q_k\end{aligned}$$

Measurement update:

$$\begin{aligned}K_{k+1} &= \bar{P}_{k+1} H^T (H \bar{P}_{k+1} H^T + R_{k+1})^{-1} \\ \hat{P}_{k+1} &= (I - K_{k+1} H) \bar{P}_{k+1} (I - K_{k+1} H)^T + K_{k+1} R_{k+1} K_{k+1}^T \\ \hat{x}_{k+1} &= \bar{x}_{k+1} + K_{k+1} (y_{k+1} - H \bar{x}_{k+1})\end{aligned}$$

Backward EKF algorithm

Time update:

$$\begin{aligned}\bar{x}_{k-1} &= \hat{x}_k - \int_{t_{k-1}}^{t_k} f(\hat{x}, u) dt \\ \bar{P}_{k-1} &= F_k \hat{P}_k F_k^T + Q_k\end{aligned}$$

Measurement update:

$$\begin{aligned} K_{k-1} &= \bar{P}_{k-1} H^T (H \bar{P}_{k-1} H^T + R_{k-1})^{-1} \\ \hat{P}_{k-1} &= (I - K_{k-1} H) \bar{P}_{k-1} (I - K_{k-1} H)^T + K_{k-1} R_{k-1} K_{k-1}^T \\ \hat{x}_{k-1} &= \bar{x}_{k-1} + K_{k-1} (y_{k-1} - H \bar{x}_{k-1}) \end{aligned}$$

Smoothing algorithm

$$\begin{aligned} K_{s,k} &= \hat{P}_{b,k} (\hat{P}_{f,k} + \hat{P}_{b,k})^{-1} \\ \hat{P}_{s,k} &= (I - K_{s,k}) \hat{P}_{b,k} \\ \hat{x}_{s,k} &= K_{s,k} \hat{x}_{f,k} + (I - K_{s,k}) \hat{x}_{b,k} \end{aligned}$$

4.9 References

1. Bastin, G. and D. Dochain (1990). On-line Estimation and Adaptive Control of Bioreactors. Amsterdam, Elsevier Science Publishers.
2. Cazzador, L. and V. Lubenova (1995). Nonlinear Estimation of specific Growth Rate for Aerobic Fermentation Processes. *Biotechnology and Bioengineering* 47: 626-632.
3. Claes, J.E. and J.F. Van Impe (1999). On-line estimation of the specific growth rate based on viable biomass measurements: experimental validation. *Bioprocess Engineering* 21: 389-395.
4. Estler, M.U. (1995). Recursive on-line estimation of the specific growth rate from off-gas analysis for the adaptive control of fed-batch processes. *Bioprocess Engineering* 12: 205-207.
5. Farza, M., H. Hammouri, S. Othman and K. Busawon (1997). Nonlinear observers for parameter estimation in bioprocesses. *Chemical Engineering Science* 52(23): 4251-4267.
6. Gee, D.A. and W.F. Ramirez (1996). On-Line State Estimation and Parameter Identification for Batch Fermentation. *Biotechnology Progress* 12: 132-140.
7. Popova, S. and T. Patarinska (1998). Biomass estimation on the basis of yeast cultivation model with morphophysiological parameters. *Bioprocess Engineering* 19(4): 313-315.

8. San, K.-Y. and G. Stephanopoulos (1984). Studies on On-Line Bioreactor Identification. II. Numerical and Experimental Results. *Biotechnology and Bioengineering* 26: 1189-1197.
9. Stephanopoulos, G. and K.Y. San (1984). Studies on On-Line Bioreactor Identification. I. Theory. *Biotechnology and Bioengineering* 26: 1176-1188.
10. Gelb, A. (1974). *Applied Optimal Estimation*. Cambridge, Massachusetts, MIT Press.
11. Lewis, F.L. (1986). *Optimal Estimation*. New York, USA, John Wiley and Sons.
12. Thalen, M., J. Van De IJssel, W. Jiskoot, B. Zomer, P. Roholl, C. De Gooijer, C. Beuvery and J. Tramper (1999). Rational medium design for *Bordetella pertussis*: basic metabolism. *Journal of Biotechnology* 75(2-3): 147-159.
13. Van 't Riet, K. and J. Tramper (1991). *Basic Bioreactor Design*. New York, Marcel Dekker.

5 DUAL-SUBSTRATE UTILISATION BY *BORDETELLA PERTUSSIS*

This chapter is based on:

R. Neeleman, M. Joerink, and A.J.B. van Boxtel (2001). Dual-Substrate Utilisation by *Bordetella pertussis*. *Applied Microbiology and Biotechnology*, 57, 489-493.

5.1 Abstract

To improve the cultivation of *Bordetella pertussis* and take advantage of the newest techniques in monitoring and control, a quantitative description of substrate utilisation is necessary. Growth of the organism is limited by two main substrates. However, neither interactive nor non-interactive modelling seems appropriate. A model that combines essential and enhanced kinetics was developed based on experimental observation. Instead of fitting all model parameters at once, a step-wise experimentation procedure was used. Finally, two cultivations showed the accuracy of the model.

5.2 Introduction

Bordetella pertussis is the bacterium responsible for whooping cough, a disease of the respiratory tract that may be life threatening in infants and young children. The disease can effectively be controlled by immunisation with a vaccine consisting of inactivated whole cells. Thalen et al. [1] described the metabolism of *B. pertussis* in detail and noticed a deficiency in the citric acid cycle. Experimental studies showed that the organism is not capable of utilising lactate as a main carbon source. However, when grown on a medium with glutamate as the main carbon source, addition of lactate increased the growth rate [2]. This clearly indicated dual-substrate growth limitation and, although the complete metabolic pathway is known, the kinetics have not been quantitatively described.

Pharmaceutical production processes are tightly controlled, and online quality monitoring is of increasing interest. To develop such advanced monitoring tools and to enhance the production efficiency of a vaccine against whooping cough, a quantitative model of the process is needed. A common approach for modelling bioprocesses with multiple-substrate limitation is the use of a non-interactive model that assumes growth of micro-organisms to be limited by the most-limiting substrate [3], or an interactive limitation where both substrates are essential and limiting simultaneously. *B. pertussis* is clearly dual-substrate limited, but neither interactive nor non-interactive models seem appropriate. Another problem with dual-substrate limited growth models is the estimation of the parameters, such as yields and maximal specific growth rate. Contrary to single-substrate models, it is not possible to determine the yields

independently from the growth rate, since the origin of the formed biomass can not be reduced to one of the substrates.

Hence, the main purpose of this study was to develop a model of *B. pertussis* in a dual-substrate-limited environment and to find a solution for the estimation of the parameters other than simply fitting them all at once.

5.3 Materials and Methods

Strains and Culture Media

B. pertussis strain 509 (available from the RIVM collection, Bilthoven, The Netherlands), one of the two strains included in the DPT-Polio vaccine used in the Netherlands, was used with the chemically defined Thijs-medium (RIVM, Bilthoven, The Netherlands) containing glutamate and lactate as main carbon substrates.

Bioreactor Conditions

The cells were grown in a 5-l round-bottomed reactor containing 3-l medium. An eight-bladed marine impeller was used to agitate the medium. Temperature, pH, dissolved oxygen, and stirrer speed were controlled at 34°C, 7.2, 20%, and 400 rpm, respectively. Oxygen was transported through the headspace only and controlled by changing the oxygen fraction in the gas flow. The total gas flow was maintained at 1 l.min⁻¹.

Analysis

A polarographic electrode (Ingold, Urdorf, Switzerland) measured the dissolved oxygen in the medium. A pH electrode (Ingold) measured the pH. The temperature was measured with a Pt100 temperature sensor. Biomass was measured by optical density using a Vitalab 10 (Vital Scientific, Dieren, The Netherlands), at 590 nm and by dry weight. The optical densities of *B. pertussis* were recalculated to g·l⁻¹ dry weight according to a calibration curve. Glutamate and L-lactate were determined with a YSI 2750 Select analyser (Yellow Springs Instruments, Yellow Springs, USA).

Table 1: Three main types of kinetics when growth is limited by two substrates.

Model	Kinetics
Interactive / Multiplicative	$\mu = \mu_1(C_{S1}) \cdot \mu_2(C_{S2})$
Non-interactive	$\mu = \min\{\mu_1(C_{S1}), \mu_2(C_{S2})\}$
Additive / Enhanced	$\mu = \mu_1(C_{S1}) + \mu_2(C_{S2})$

5.4 Results

Dual-Substrate Limitation

Generally, unstructured non-segregated models describe microbial growth as an autocatalytic process; the change of biomass in batch cultivation is directly related to the product of biomass concentration C_X and growth rate μ . For a medium with more than one main substrate, the assumption is made that the growth rate is environmentally limited by the most-limiting substrate [3]. Although dated, this concept, called ‘non-interactive modelling’, still holds today for many cases. However, the number of reports of microorganisms that are simultaneously limited by two or more substrates is increasing [4-7].

Unstructured dual-substrate models are characterised according to whether they are ‘interactive’ or ‘non-interactive’. ‘Interactive’ models assume that growth is limited by both substrates simultaneously. For these models both substrates are essential and thus have to be present for growth to take place. Tsao and Hanson [8] extended these models to include the effect of growth-enhancing substrates. These substrates are assumed to increase the growth rate and at the same time are not essential for growth to take place. For this third ‘additive’ case, the growth rate is the sum of the two individual growth rates, which are measured separately. Moreover, the growth rate would change when one of the substrates becomes depleted while the other is still saturating. The three main types of dual-substrate models are shown in Table 1. All of these models mainly concern biomass growth, and although the consumption rate of the substrates is of importance, they have not been discussed previously.

Growth Rate

From previous experiments it was noticed that the growth rate of *B. pertussis* on glutamate only is lower than the growth rate on medium containing both glutamate and lactate. Furthermore, a dual-substrate-limited approach seems appropriate, since the specific growth rate changes during cultivation when one substrate becomes depleted (Figure 1). In Figure 1, the growth rate is estimated from cell density measurements by a model-free smoothing algorithm [2]. When both substrates are present during the first 20 h, the growth rate is maximal at approximately 0.15 h^{-1} , when lactate becomes depleted after 23 h, the growth rate drops to approximately 0.10 h^{-1} .

This contradicts the first two ‘essential’ cases of dual limited growth. One possible explanation for this situation is additive growth, the third case, which means that *B. pertussis* is capable of utilising lactate in parallel with glutamate and enhancing its growth rate. Growth on glutamate would then be independent of growth on lactate, and the total observable growth rate would be the sum of the two growth rates. However, this does not reflect reality since

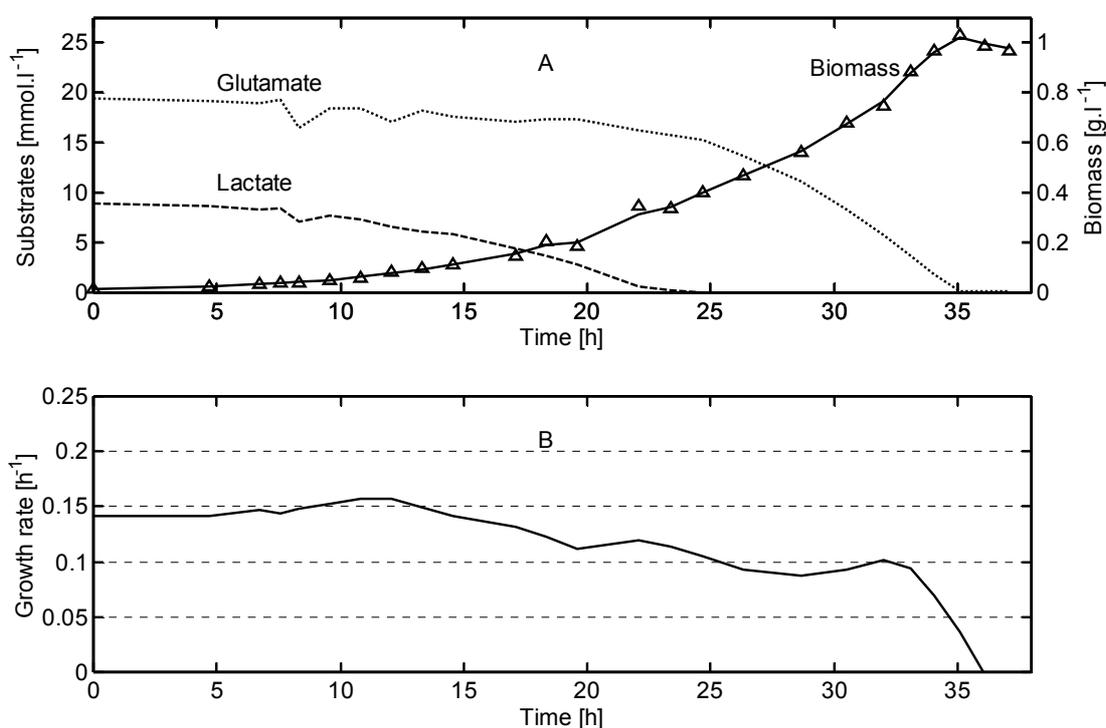


Figure 1: Estimated specific growth rate during a batch cultivation of *B. pertussis*. A: Biomass measurements (triangles) and estimations (solid) and concentrations of glutamate (dotted) and lactate (dashed). B: Smooth estimation of the specific growth rate.

it assumes that the organism can grow on lactate only, which is not the case [1]. Hence, using the current medium, a combination of essential and additive modelling has been used, where glutamate is an essential, and lactate is an enhancing substrate.

Substrate Utilisation

Glutamate is used by *B. pertussis* both as a carbon and nitrogen source for building material for growth and as an energy source. The metabolism of *B. pertussis* has been described in detail by Thalen et al. [1] and can be generalised to the formation of biomass from glutamate and lactate by two major pathways. The first pathway describes growth on glutamate only and the second pathway describes enhanced growth on both substrates. Growth via these two pathways is assumed parallel, and thus the individual growth rates can be added (enhanced growth). For both pathways, glutamate is essential. Assuming Monod kinetics for the pathways, the complete model comprises:

$$\begin{aligned}
 \frac{dC_X}{dt} &= (\mu_1 + \mu_2) C_X \\
 \frac{dC_G}{dt} &= - \left(\frac{\mu_1}{Y_{G1}} - \frac{\mu_2}{Y_{G2}} \right) C_X \\
 \frac{dC_L}{dt} &= - \frac{\mu_2}{Y_{L2}} C_X
 \end{aligned}
 \quad
 \begin{aligned}
 \mu_1 &= \mu_{\max} \frac{C_G}{K_G + C_G} \\
 \mu_2 &= \mu_{enh} \frac{C_G}{K_G + C_G} \frac{C_L}{K_L + C_L}
 \end{aligned}
 \tag{1}$$

Where C_G and C_L are the concentrations of respectively, glutamate and lactate, Y_{G1} , Y_{G2} and Y_{L2} are respectively, the yield of biomass on glutamate over pathway 1, the yield of biomass on glutamate over pathway 2, and the yield of biomass on lactate over pathway 2. The maximal growth rate over pathway 1 is μ_{\max} and the ‘enhancing’ growth rate over pathway 2 is μ_{enh} . K_G and K_L are the half-saturation constants for the Monod kinetics.

When both substrates are present the overall yield of biomass on glutamate, Y_G^{ov} , which can be calculated from measurements, will be higher than for a cultivation with glutamate only. This is due to a part of the biomass that is also formed from lactate in the second pathway and thus explains the second negative sign in the above derivation of the glutamate concentration.

Step-wise Experimentation for Parameter Estimation

For single-substrate models the yield can easily be calculated independent of the growth rate by linear regression of biomass and substrate measurements. However, for dual-substrate models it is not possible to separate which amount of biomass is formed from which substrate, and consequently the yields cannot be determined independently of the growth rate. To avoid fitting all parameters at once, a step-wise procedure is used. Firstly, a cultivation with only glutamate is performed. Since this is a single-substrate cultivation the maximal growth rate, μ_{max} , and the yield over the first pathway, Y_{G1} , are determined independently. Figure 2a shows the measured biomass concentrations against the measured glutamate concentrations for linear regression to determine Y_{G1} , the yield of biomass on glutamate over pathway 1. The maximal growth rate, μ_{max} , is determined by fitting an exponential curve through the biomass measurements during the saturating part of the batch.

Secondly, a cultivation is performed with both substrates present during the complete cultivation. With the results of this cultivation the overall growth rate is determined and thus the enhancing growth rate, μ_{enh} , is calculated. Figure 2b shows the biomass concentrations against the glutamate and lactate concentrations. From these measurements the overall yield of biomass on

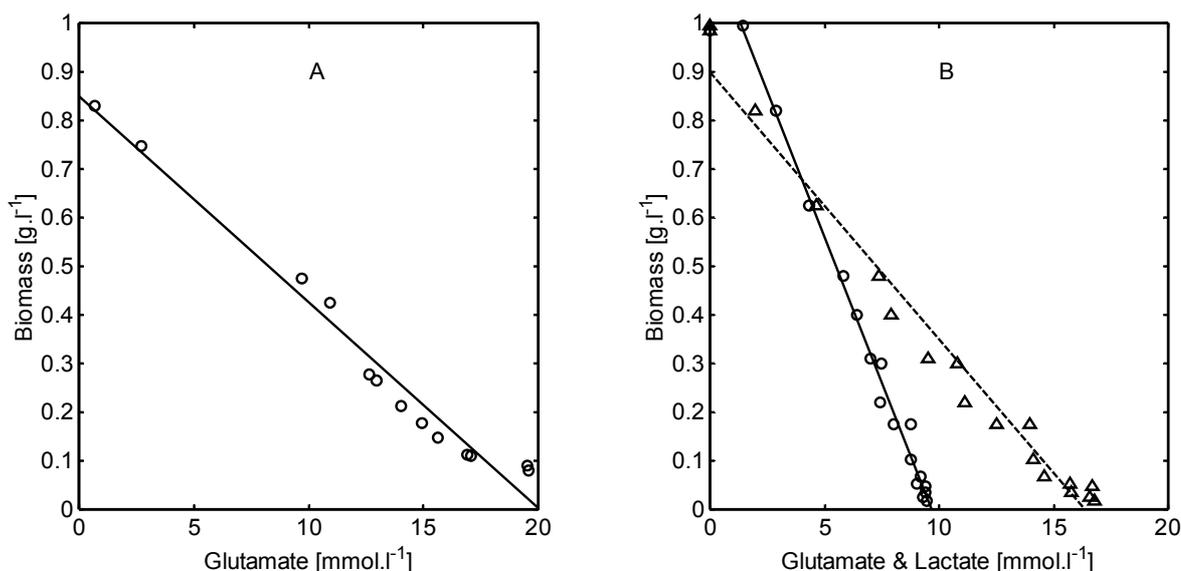


Figure 2: Biomass measurements against substrate measurements and the results of linear regression (solid) to determine the yields. A: Cultivation with glutamate only. B: Cultivation with glutamate (circles) and lactate (triangles).

Table 2: Parameters of the dual-substrate-limited growth model of *B. pertussis*.

Path 1	Value	Path 2	Value
μ_{max}	0.128 h ⁻¹	μ_{enh}	0.032 h ⁻¹
Y_{G1}	42.35 g·mol ⁻¹	Y_{L2}	10.92 g·mol ⁻¹
Y_{G2}	18.69 g·mol ⁻¹	Y_L^{ov}	55.04 g·mol ⁻¹
Y_G^{ov}	120.29 g·mol ⁻¹	K_L	0.50 mmol·l ⁻¹
K_G	0.50 mmol·l ⁻¹		

lactate, Y_L^{ov} and the overall yield of biomass on glutamate, Y_G^{ov} , are determined by linear regression. Subsequently the yield of biomass on lactate over pathway 2, Y_{L2} , is derived from the overall yield, Y_L^{ov} , according to the Appendix. The yield of biomass on glutamate over pathway 2, Y_{G2} , is derived from the overall yield, Y_G^{ov} , and the yield over pathway 1, Y_{G1} , according to the Appendix. Finally the half-saturation constants, K_G and K_L , are chosen to be small, since they only influence the very low substrate concentrations at the end of the batch. All derived parameter values are shown in Table 2.

Model Validation

These parameter values were used to validate the model with a set of independent experiments. Different starting concentrations were used for both the simulations and actual experiments. Figure 3a shows the results of a cultivation and a simulation of a standard cultivation where there is a balanced

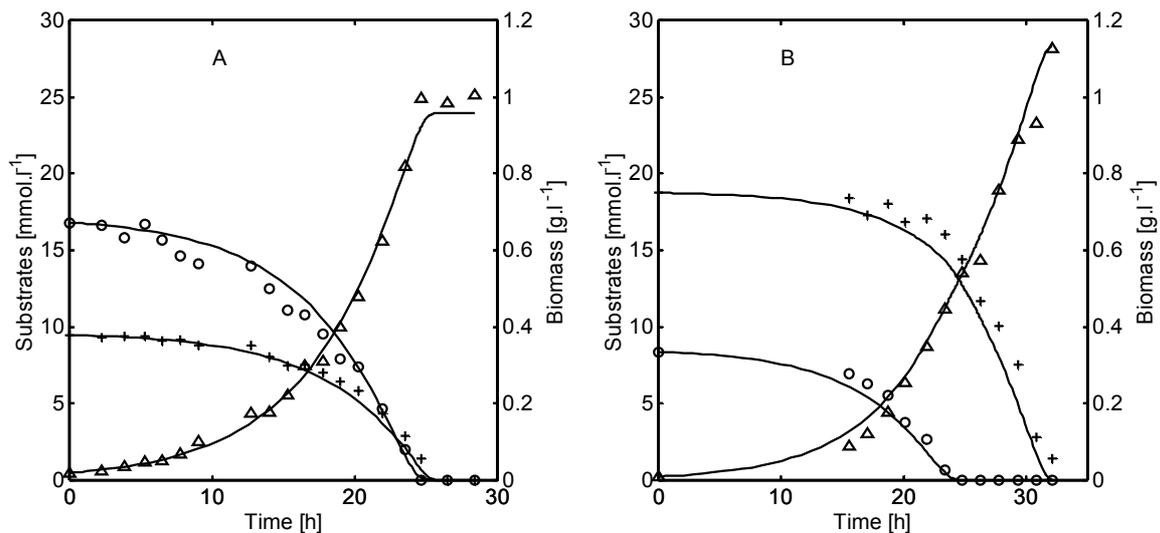


Figure 3: Measured and simulated (solid) concentrations of biomass (triangles), glutamate (+), and lactate (o) during two cultivations. A: Cultivation with standard medium (balanced amounts of glutamate and lactate). B: Cultivation with more glutamate and less lactate.

amount of substrates, such that both are depleted at approximately the same time. Consequently, during this complete cultivation both pathways are active. The model shows good accuracy in predicting the behaviour of *B. pertussis* when grown using both pathways. Figure 3b shows the results of a simulation and a cultivation with higher glutamate concentration and lower lactate concentration. Halfway through the cultivation, at 23 h, the lactate is depleted and growth continues on glutamate only. At that time, the organism switches from both pathways to pathway 1 only. Again, the model and its parameters show good accuracy in predicting growth on both pathways, as well as growth on pathway 1.

5.5 Discussion

The experiments showed that the model is capable of predicting the behaviour of *B. pertussis* when grown on a mixture of glutamate and lactate. Furthermore, they showed that halfway through cultivation when lactate becomes depleting, the switch in metabolism is described accurately. At present, this dual-substrate model coupled with a state estimator is used to satisfactorily predict the future trajectory of a cultivation on-line [9].

Besides the perspective of advanced monitoring, the development of this model also yielded an improvement of the used medium. From the model and its parameters, the optimum ratio between glutamate and lactate was determined to make sure that both substrates are simultaneously depleted.

5.6 Nomenclature

C_G	$\text{mmol}\cdot\text{l}^{-1}$	glutamate concentration
C_L	$\text{mmol}\cdot\text{l}^{-1}$	lactate concentration
C_S	$\text{mmol}\cdot\text{l}^{-1}$	substrate concentration
C_X	$\text{g}\cdot\text{l}^{-1}$	biomass concentration
K_G	$\text{mmol}\cdot\text{l}^{-1}$	glutamate saturation constant
K_L	$\text{mmol}\cdot\text{l}^{-1}$	lactate saturation constant
Y_{G1}	$\text{g}\cdot\text{mmol}^{-1}$	yield of biomass on glutamate for pathway 1
Y_{G2}	$\text{g}\cdot\text{mmol}^{-1}$	yield of biomass on glutamate for pathway 2
Y_G^{ov}	$\text{g}\cdot\text{mmol}^{-1}$	overall yield of biomass on glutamate

Y_{L2}	$\text{g}\cdot\text{mmol}^{-1}$	yield of biomass on lactate for pathway 2
Y_L^{ov}	$\text{g}\cdot\text{mmol}^{-1}$	overall yield of biomass on lactate
μ	h^{-1}	specific growth rate
μ_1	h^{-1}	specific growth rate for pathway 1
μ_2	h^{-1}	specific growth rate for pathway 2
μ_{max}	h^{-1}	maximal specific growth rate (pathway 1)
μ_{enh}	h^{-1}	enhancing specific growth rate (pathway 2)

5.7 Appendix

Lactate Yield

The measured overall yield of biomass on lactate, Y_L^{ov} , is calculated by linear regression from the measurements as follows:

$$\frac{dC_X}{dC_L} = -Y_L^{ov} \quad (\text{A1})$$

In the model, the yield is split over two pathways and since lactate is used only by the second pathway, the yield of biomass on lactate over pathway 2, Y_{L2} , is reconstructed from the overall yield and the model:

$$\frac{dC_X}{dC_L} = \frac{\frac{dC_X}{dt}}{\frac{dC_L}{dt}} = \frac{(\mu_1 + \mu_2)C_X}{-\frac{\mu_2}{Y_{L2}}C_X} = -Y_L^{ov} \quad (\text{A2})$$

Rewriting this equation, assuming saturating conditions and thus using the maximal and ‘enhanced’ growth rates, gives:

$$Y_{L2} = \frac{\mu_{enh}}{\mu_{max} + \mu_{enh}} Y_L^{ov} \quad (\text{A3})$$

Glutamate Yield

The measured overall yield of biomass on glutamate, Y_G^{ov} , is calculated by linear regression from the measurements as follows:

$$\frac{dC_X}{dC_G} = -Y_G^{ov} \quad (A4)$$

In the model, the yield is split in growth over pathway 1 and over pathway 2. Since the yield over pathway 1, Y_{G1} , is known and the overall yield is measured, the yield over pathway 2, Y_{G2} , can be derived:

$$\frac{dC_X}{dC_G} = \frac{\frac{dC_X}{dt}}{\frac{dC_G}{dt}} = \frac{(\mu_1 + \mu_2)C_X}{-\left(\frac{\mu_1}{Y_{G1}} - \frac{\mu_2}{Y_{G2}}\right)C_X} = -Y_G^{ov} \quad (A5)$$

Rewriting this equation, assuming saturating conditions and thus using the maximal growth rate and the ‘enhancing’ growth rate, gives:

$$Y_{G2} = \frac{\mu_{enh} Y_{G1} Y_G^{ov}}{\mu_{max} Y_G^{ov} - (\mu_{max} + \mu_{enh}) Y_{G1}} \quad (A6)$$

5.8 References

1. Thalen, M., J. Van De IJssel, W. Jiskoot, B. Zomer, P. Roholl, C. De Gooijer, C. Beuvery and J. Tramper (1999). Rational medium design for Bordetella pertussis: basic metabolism. *Journal of Biotechnology* 75(2-3): 147-159.
2. Neeleman, R. and A.J.B. Van Boxtel (2001). Estimation of specific growth rate from cell density measurements. *Bioprocess Engineering* 24(3): 179-185.
3. Ryder, D.N. and C.G. Synclair (1972). Model for the growth of aerobic microorganisms under oxygen limiting conditions. *Biotechnology and Bioengineering* 14: 787-798.
4. Beyenal, H. and A. Tanyolac (1997). A combined growth model of Zoogloea ramigera including multisubstrate, pH, and agitation effects. *Enzyme and Microbial Technology* 21(2): 74-78.
5. Lee, A.L., M.M. Ataa and M.L. Shuler (1984). Double-substrate-limited growth of Escherichia coli. *Biotechnology and Bioengineering* 26: 1398-1401.

CHAPTER 5

6. Machado, R. and C.P.L. Grady (1989). Dual substrate removal by an axenic bacterial culture. *Biotechnology and Bioengineering* 33: 327-337.
7. Sambanis, A., S. Pavlou and A.G. Fredrickson (1986). Analysis of the dynamics of ciliate-bacterial interactions in a CST. *Chemical Engineering Science* 41(6): 1455-1469.
8. Tsao, G.T. and T.P. Hanson (1975). Extended monod equation for batch cultures with multiple exponential phases. *Biotechnology and Bioengineering* 17: 1591-1598.
9. Neeleman, R., M. Joerink, A.J.B. Van Boxtel, C. Beuvery and G. Van Straten (2001). On-line harvest prediction for the production of biologicals, *Computer Applications in Biotechnology*, Quebec, Canada, 430-435.

6 SEQUENTIAL ESTIMATION OF BIOMASS AND SPECIFIC GROWTH RATE

6.1 Abstract

The biomass concentration is a very important state variable and its specific growth rate is a crucial parameter of almost every bioprocess. Due to the lack of reliable and cheap sensors, they cannot often be measured directly or estimated from related variables, such as the concentrations of substrates or products. In this chapter, a sequence of two stable observers is used to estimate the specific growth rate and the biomass concentration for processes where the measurement of oxygen uptake rate is available on-line. A good agreement between the sequentially estimated values and the measured values is shown for the cultivation process of *Bordetella pertussis*, thus proving the applicability of the algorithm.

6.2 Introduction

For better control and optimisation of bioreactor processes, it is essential to measure the key physiological parameters on-line. However, such measurements are rarely available for most of these parameters. Therefore, any control or optimisation based on key physiological parameters cannot be implemented unless values can be measured on-line to provide the necessary information required by the controller or optimiser. Although efforts to develop such sensors are underway, the availability and reliability of these instruments are still very limited. There is thus an urgent need to provide the necessary information by estimating these parameters from other values that are simpler and easier to measure. In this chapter, parameter and state estimation algorithms that can be used to estimate a key parameter or a physiological state of a bioreactor will be presented.

Many works in this area have focused on the estimation of the specific growth rate in simple cultures (with one substrate, one biomass, and in some cases one product) [1-3]. In most cases, the proposed estimators are based on an extended Kalman Filter approach, which generally leads to complex algorithms. Moreover, despite some apparently good results, there is no a-priori guarantee of the convergence and stability of these algorithms. Another approach for estimating the kinetic rates in bioreactors is that developed by Bastin and Dochain [4]. That approach needs a measurement of part of the state of the process [5, 6]. A good example of an estimator based on exit-gas analysis is proposed by Estler [7]. The use of extended Kalman Filter for the

on-line estimation of both biological states and kinetic parameters is studied in detail by Stephanopoulos and San [8] and Pons [9], where experimental applications are reported.

In this chapter, a different type of algorithm is presented. Since the dynamics of the state and the parameters are different in nature, their estimation will be separated. This has been suggested as a natural way of implementing an adaptive state estimator by Goodwin and Sin [10]. The potential advantages of this approach are:

- It makes use of system's linearity (the mass-balance equations become linear).
- Partial convergence can be proven; fewer factors have to be tuned than i.e. an extended state Kalman filter.
- Since the specific rates are obtained separately, the algorithm can be tuned separately.

6.3 Materials and Methods

Strain, Medium and Bioreactor Conditions

Bordetella pertussis strain 509 (RIVM, Bilthoven, The Netherlands), one of the two strains included in the DPT-Polio vaccine used in the Netherlands, was used with the chemically defined THIJS-medium containing glutamate and lactate as main carbon sources (RIVM). The cells were grown in a 5-l round-bottomed glass reactor containing 3-l medium. An eight-bladed marine impeller was used to agitate the medium. Temperature, pH, dissolved oxygen, and stirrer speed were controlled at 34°C, 7.2, 20%, and 400 to 500 rpm, respectively. Oxygen was transported through the headspace only and controlled by changing the oxygen fraction in the gas flow. The total gas flow was maintained at 1 l.min⁻¹.

For the fed-batch experiment two 500-ml concentrated stock solutions of glutamate and lactate were used for the two separate feeds and placed on balances to monitor the feed rate. The feeds were added to the bioreactor by two pumps (101U/R 32 rpm, Watson Marlow ltd., Cornwall UK) connected to the bioreactor control system.

Analysis

A polarographic electrode (Ingold, Urdorf, Switzerland) measured the dissolved oxygen in the medium. A pH electrode (Ingold) measured the pH. The temperature was measured with a Pt100 temperature sensor. Biomass was measured by optical density using a Vitalab 10 (Vital Scientific, Dieren, The Netherlands), at 590 nm and by dry weight. The optical densities of *B. pertussis* were recalculated to $\text{g}\cdot\text{l}^{-1}$ dry weight according to a calibration curve. Glutamate and L-lactate were determined with a YSI 2750 Select analyser (Yellow Springs Instruments, Yellow Springs, USA).

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). All standard control-loops (dissolved oxygen, pH, temperature, etc.) were performed in the BCSV-software and logged to a disc on the UNIX machine. Using FTP, a Windows NT machine downloaded new data every minute. These data were processed by the sequential estimator, which was implemented in Matlab (Mathworks, Massachusetts USA) to estimate biomass concentration and specific growth rate.

6.4 Sequential Estimation

A significant part of the overall bioreactor identification is the estimation of unknown or uncertain model parameters. Usually, this is accomplished by treating the parameters as additional state variables. In this chapter, another state/parameter estimation method is presented. It is based on the observation that, in an adequately formulated model, the model parameters are distinguished from the state variables in that they change markedly slower than the state variables, and sometimes they are, in fact, time invariant. In order to illustrate this point, consider a model in which the parameters are truly time invariant. If these parameter values are accurately known, then the state estimation and its future projection using the model equations will result in small prediction errors at first. That results in slower convergence of the estimation of the parameters and consequently causes divergence of estimates,

especially when the true states of the model depend strongly on the parameters.

Therefore, the estimation of the states will be separated from the estimation of the parameters. Such an algorithm is shown in Figure 1 and can be summarised as follows. First, the parameters, p , are estimated from the available measurements, y , at a certain moment in time, t . This leads to an update of the system dynamic model. Then, the estimation of the state vector, x , is made with the measurements and the updated model. The implicit hypothesis of such an approach is that the dynamics of the unknown parameters are slower than the dynamics of the state variables. This approach is called sequential state/parameter estimation, SSPE [11-13]. This algorithm is applicable to systems with adequate models for which the parameters are not expected to vary significantly with time.

Estler [7] reports the use of a sequential algorithm for adaptive control of the ammonium concentration in a fed-batch process. Park and Ramirez [12] report the use of the SSPE algorithm for regulating the glucose concentration in yeast fermentation. Both estimate the states separately from the parameters. With such applications it should be noted that the underlying basis for the SSPE algorithm is the distinctively lower rate of parameter variation compared to that of the state variables. Since μ and Y can be strongly dependent upon culture conditions such as substrate concentration, dissolved oxygen concentration, temperature, and pH, the parameterisation of such metabolic variables may not be appropriate for SSPE application. However, if the culture conditions are regulated at constant levels, it is reasonable to expect that these

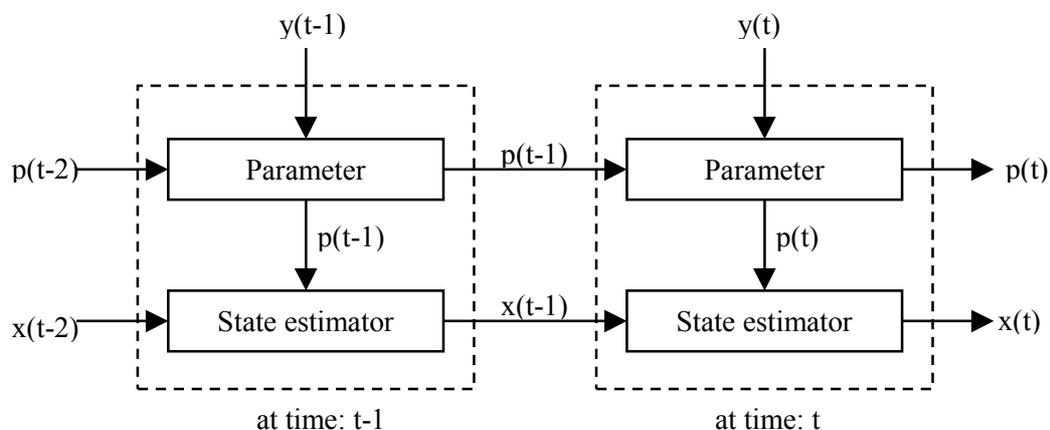


Figure 1: Scheme of a sequential state/parameter estimator. With state x , parameters p , measurements y , at time $t-1$ and t .

metabolic variables do not show considerable variation. Both Estler [7] and Park and Ramirez [12] used SSPE coupled with regulation of culture dynamic conditions.

6.5 Estimation of the Specific Growth Rate

Model

A suitable asymptotic observer of the specific growth rate based on the measurements of the oxygen uptake rate OUR , is developed by Lubenova et al. [14] for batch processes. This observer uses a generic model for biomass growth and oxygen consumption that describes oxygen consumption as the sum of oxygen used for growth and oxygen used for maintenance. Representing this generic model to accommodate a fed-batch process leads to Equation (1):

$$\begin{aligned} \frac{dC_X}{dt} &= (\mu - D)C_X \\ OUR &= \left(\frac{\mu}{Y_O} + m_O \right) C_X \end{aligned} \quad (1)$$

Where C_X is the concentration of biomass, μ the specific growth rate, D the dilution rate, Y_O the yield of oxygen on biomass growth, and m_O the maintenance coefficient for oxygen. The parameters for the case of *B. pertussis* were determined from experimental data [15] and are shown in Table 1. The derivative of the oxygen uptake rate, OUR , can be calculated as follows:

$$\begin{aligned} \frac{dOUR}{dt} &= \left(\frac{\mu}{Y_O} + m_O \right) \frac{dC_X}{dt} + \frac{d \left(\frac{\mu}{Y_O} + m_O \right)}{dt} C_X \\ &= \left(\frac{\mu}{Y_O} + m_O \right) (\mu - D) C_X + \frac{d\mu}{dt} \frac{C_X}{Y_O} \end{aligned} \quad (2)$$

Rewriting this equation with Eq. (1) gives:

$$\begin{aligned}\frac{dOUR}{dt} &= \left(\mu - D + \frac{\frac{d\mu}{dt}}{(\mu + Y_O m_O)} \right) OUR \\ &= \alpha \cdot OUR\end{aligned}\quad (3)$$

Where α is called the ‘apparent specific growth rate’ and depicts the apparent specific growth rate of OUR . Rewriting the equation for α gives:

$$\frac{d\mu}{dt} = (\alpha - \mu + D) (\mu + Y_O m_O) \quad (4)$$

Note that when the specific growth rate is constant, α equals $(\mu - D)$ and the OUR increases with the same exponential rate as the biomass concentration.

Observer

To identify the imperfectly known and possibly slowly varying parameter α , an observer is constructed. For that purpose the system is extended with the differential equation of α with a zero right-hand part added for the unknown parameter ‘dynamics’. In other words, α is modelled as a random walk process. The μ -observer as described by Lubenova et al. [14] was developed for batch cultivation only. Adjusting the algorithm for fed-batch cultivation results in:

$$\begin{aligned}\frac{dO\hat{U}R}{dt} &= \hat{\alpha} OUR_m + p_1 OUR_m (OUR_m - O\hat{U}R) \\ \frac{d\hat{\alpha}}{dt} &= 0 + p_2 OUR_m (OUR_m - O\hat{U}R) \\ \frac{d\hat{\mu}}{dt} &= (\hat{\alpha} - \hat{\mu} + D) (\hat{\mu} + Y_O m_O)\end{aligned}\quad (5)$$

where a hat (^) depicts the estimations and p_1 is a tuning parameter describing the trade-off between the noise of the measurements, OUR_m , and the ‘noise’ of the model. Tuning parameter p_2 determines the characteristics of the random walk process of α . The tuning parameters have to be chosen such that the observer works properly and is stable.

Stability Analysis

The stability and convergence of this specific growth rate estimator are proven by analysis of the system error (real value minus estimations):

$$\begin{aligned}\varepsilon_{OUR} &= OUR - O\hat{U}R \\ \varepsilon_{\alpha} &= \alpha - \hat{\alpha}\end{aligned}\tag{6}$$

The dynamics of the estimation errors are:

$$\begin{aligned}\frac{d\varepsilon_{OUR}}{dt} &= \frac{dOUR}{dt} - \frac{dO\hat{U}R}{dt} \\ &= \alpha OUR - \hat{\alpha} OUR_m - p_1 OUR_m (OUR_m - O\hat{U}R) \\ \frac{d\varepsilon_{\alpha}}{dt} &= \frac{d\alpha}{dt} - \frac{d\hat{\alpha}}{dt} \\ &= \frac{d\alpha}{dt} - p_2 OUR_m (OUR_m - O\hat{U}R)\end{aligned}\tag{7}$$

Now assume that the real OUR data are corrupted by the additive measurement noise ε_m :

$$OUR_m = OUR + \varepsilon_m\tag{8}$$

Combining equations (6), (7), and (8) gives:

$$\begin{aligned}\frac{d\varepsilon_{OUR}}{dt} &= \alpha (OUR_m - \varepsilon_m) - \hat{\alpha} OUR_m - p_1 OUR_m (\varepsilon_{OUR} + \varepsilon_m) \\ &= \varepsilon_{\alpha} OUR_m - \alpha \varepsilon_m - p_1 OUR_m \varepsilon_{OUR} - p_1 OUR_m \varepsilon_m \\ \frac{d\varepsilon_{\alpha}}{dt} &= \frac{d\alpha}{dt} - p_2 OUR_m (\varepsilon_{OUR} + \varepsilon_m) \\ &= \frac{d\alpha}{dt} - p_2 OUR_m \varepsilon_{OUR} - p_2 OUR_m \varepsilon_m\end{aligned}\tag{9}$$

Now the dynamics of the estimation error $E = [\varepsilon_{OUR} \ \varepsilon_{\alpha}]^T$ is expressed by the following vector equation:

$$\frac{dE}{dt} = A \cdot E + B\tag{10}$$

With

$$\frac{dE}{dt} = \begin{bmatrix} -p_1 & 1 \\ -p_2 & 0 \end{bmatrix} \cdot OUR_m \cdot \begin{bmatrix} \varepsilon_{OUR} \\ \varepsilon_\alpha \end{bmatrix} + \begin{bmatrix} (-\alpha - p_1 OUR_m) \varepsilon_m \\ -p_2 OUR_m \varepsilon_m + \frac{d\alpha}{dt} \end{bmatrix} \quad (11)$$

Since the OUR_m can be taken out of the observer matrix A , the eigenvalues of this matrix are independent of the measured OUR_m . The system is called stable when the real parts of the eigenvalues of the observer matrix are negative or zero [16]. This is guaranteed when:

$$\begin{aligned} p_1 &\geq 2 \cdot \sqrt{p_2} \\ p_2 &\geq 0 \end{aligned} \quad (12)$$

6.6 Estimation of Biomass Concentration

Model

The separately estimated specific growth rate and its derivative are used in combination with the measured oxygen uptake rate OUR_m to construct an observer for the biomass concentration according to Ignatova and Lubenova [17]. First a model of the development through time of OUR and the biomass growth rate R_X is used:

$$\begin{aligned} \frac{dOUR}{dt} &= \left(\frac{\mu}{Y_O} + m_O \right) (\mu - D) C_X + \frac{d\mu}{dt} \frac{C_X}{Y_O} \\ \frac{dC_X}{dt} &= (\mu - D) C_X = R_X \end{aligned} \quad (13)$$

Observer

This model, equation (13), is used to derive the following biomass observer. Note that C_X is modelled as a random walk process with coloured noise by describing its derivative, R_X , as a random walk process with zero mean noise [18]. This means that the dynamics of C_X are affected by a random and a systematic part. The latter, R_X , is also subject to random noise:

$$\begin{aligned}
 \frac{dO\bar{U}R}{dt} &= \left(\frac{\hat{\mu}}{Y_O} + m_O \right) \bar{R}_X + \frac{d\hat{\mu}}{dt} \frac{\bar{C}_X}{Y_O} + p_3 (OUR_m - O\bar{U}R) \\
 \frac{d\bar{C}_X}{dt} &= \bar{R}_X + p_4 (OUR_m - O\bar{U}R) \\
 \frac{d\bar{R}_X}{dt} &= p_5 (OUR_m - O\bar{U}R)
 \end{aligned} \tag{14}$$

where a hat (^) indicates the estimations of the first observer and a dash (-) the estimations of the second observer. p_3 indicates the trade-off between noise from the OUR model and its measurement, p_4 indicates the correction of the error prediction to the biomass model and p_5 depicts the random walk characteristics of R_X . The used values of the tuning parameters are mentioned in Table 1.

Stability Analysis

The stability and convergence of this biomass estimator are proven by analysis of the system error, (real values minus estimations):

$$\begin{aligned}
 \varepsilon_{OUR} &= OUR - O\bar{U}R \\
 \varepsilon_R &= R_X - \bar{R}_X \\
 \varepsilon_X &= C_X - \bar{C}_X
 \end{aligned} \tag{15}$$

The dynamics of the estimation errors are:

$$\begin{aligned}
 \frac{d\varepsilon_{OUR}}{dt} &= \frac{dOUR}{dt} - \frac{dO\bar{U}R}{dt} \\
 &= \left(\frac{\hat{\mu}}{Y_O} + m_O \right) \varepsilon_R + \frac{d\hat{\mu}}{Y_O \cdot dt} \frac{1}{Y_O} \varepsilon_X - p_3 (OUR_m - O\bar{U}R) \\
 \frac{d\varepsilon_R}{dt} &= \frac{dR_X}{dt} - \frac{d\bar{R}_X}{dt} \\
 &= \frac{dR_X}{dt} - p_5 (OUR_m - O\bar{U}R) \\
 \frac{d\varepsilon_X}{dt} &= \frac{dC_X}{dt} - \frac{d\bar{C}_X}{dt} = R_X - \bar{R}_X - p_4 (OUR_m - O\bar{U}R) \\
 &= \varepsilon_R - p_4 (OUR_m - O\bar{U}R)
 \end{aligned} \tag{16}$$

Assume that the real OUR data are corrupted by the measurement noise ε_m :

$$OUR_m = OUR + \varepsilon_m \quad (17)$$

Combining equations (15), (16), and (17) gives:

$$\begin{aligned} \frac{d\varepsilon_{OUR}}{dt} &= -p_3\varepsilon_{OUR} + \left(\frac{\hat{\mu}}{Y_O} + m_O \right) \varepsilon_R + \frac{d\hat{\mu}}{Y_O \cdot dt} \varepsilon_X - p_3\varepsilon_m \\ \frac{d\varepsilon_R}{dt} &= -p_5\varepsilon_{OUR} - p_5\varepsilon_m + \frac{dR_X}{dt} \\ \frac{d\varepsilon_X}{dt} &= -p_4\varepsilon_{OUR} + \varepsilon_R - p_4\varepsilon_m \end{aligned} \quad (18)$$

Now the dynamics of the estimation error $E = [\varepsilon_{OUR} \ \varepsilon_R \ \varepsilon_X]^T$ is expressed by the following vector equation:

$$\frac{dE}{dt} = A \cdot E + B \quad (19)$$

With

$$\frac{dE}{dt} = \begin{bmatrix} -p_3 & \frac{\hat{\mu}}{Y_O} + m_O & \frac{d\hat{\mu}}{Y_O \cdot dt} \\ -p_5 & 0 & 0 \\ -p_4 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \varepsilon_{OUR} \\ \varepsilon_R \\ \varepsilon_X \end{bmatrix} + \begin{bmatrix} -p_3\varepsilon_m \\ -p_5\varepsilon_m + \frac{dR_X}{dt} \\ -p_4\varepsilon_m \end{bmatrix} \quad (20)$$

If the observer matrix A has real distinct eigenvalues smaller or equal to zero then the observer is called stable [16]. Consequently, in order to be stable the parameters p_3 , p_4 , and p_5 need to be bounded by:

$$\begin{aligned} p_3 &\geq 2 \\ p_4 &\geq \frac{\mu}{Y_O} + m_O + D + 1 \\ p_5 &\geq \frac{1}{Y_O} \cdot \frac{d\mu}{dt} + 4 \end{aligned} \quad (21)$$

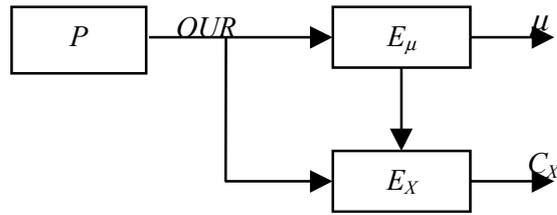


Figure 2: Diagram of the complete sequential estimation algorithm with process P , growth rate observer E_μ and biomass observer E_X .

6.7 Results

Sequential Estimation

These two observers, equations (5) and (14), are combined in sequence as illustrated by Figure 2, to estimate the specific growth rate and biomass concentration on-line from OUR measurements only. First the specific growth rate and its derivative are estimated from the measured OUR . These estimations are used to update the dynamics of the model used for estimating the biomass concentration in the second estimator. The observers are initialised with the maximum specific growth rate known from data and the inoculation biomass concentration.

Tuning

The problem of tuning the estimators is solved by finding the proper values for p_1 to p_5 . For this purpose, performance criteria such as the estimation error and the sensitivity to noise are evaluated and related to the values of the tuning parameters. For the growth rate observer those are the estimation error of OUR and the sensitivity to noise for the specific growth, and for the biomass observer those are the estimation errors of the OUR and the biomass

Table 1: Parameters used by the observers and dual-substrate model of *B. pertussis*.

parameter	value	parameter	value
p_1	16.76	Y_O	$0.6 \text{ g}\cdot\text{mmol}^{-1}$
p_2	7.81	m_O	0.01
p_3	2.65		
p_4	12.78		
p_5	15.65		

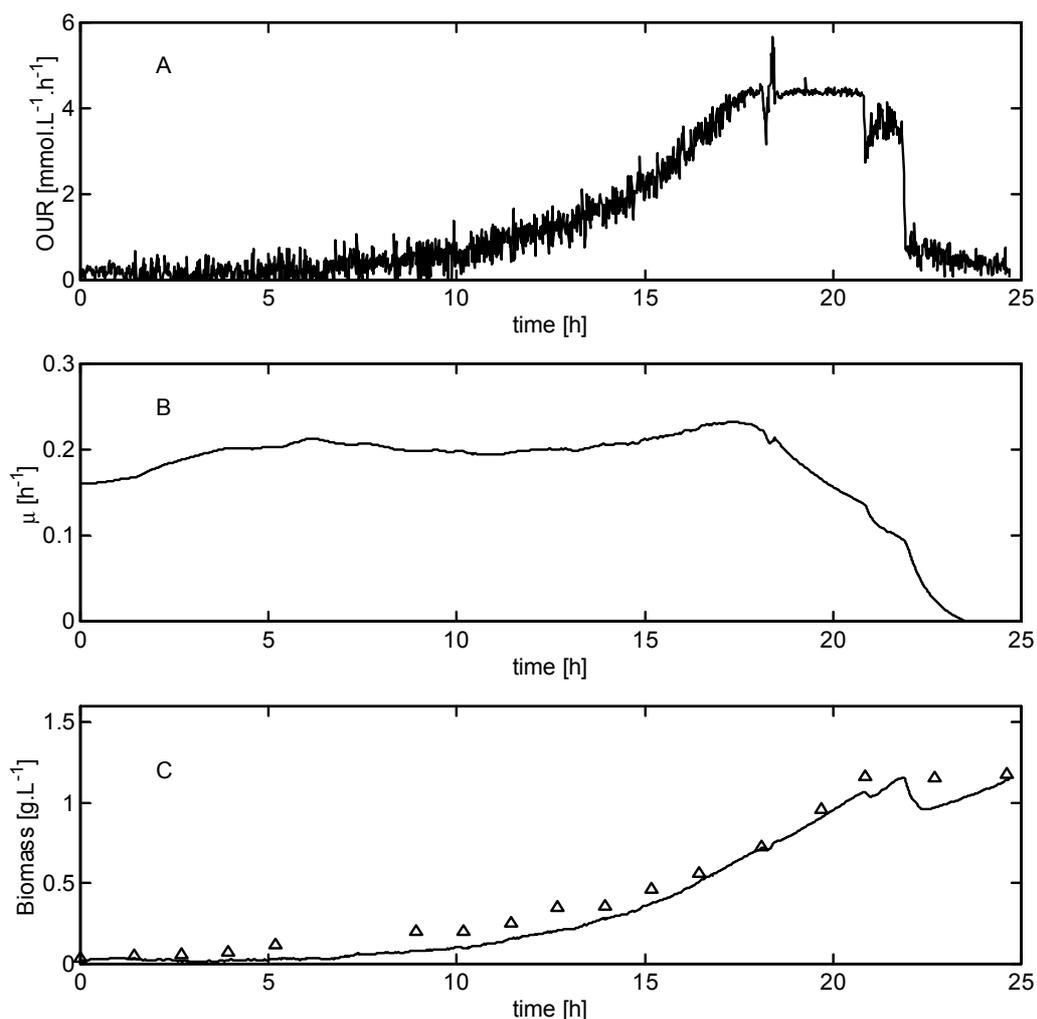


Figure 3: Sequentially estimated batch cultivation of *B. pertussis*. Top graph (A): The measured OUR during cultivation. Middle graph (B): estimated specific growth rate. Bottom graph (C): The estimated (solid) and measured (triangles) biomass concentrations.

concentration and the sensitivity to noise for the biomass concentration. Such an analysis provides insight in the tuning procedure, but does not yield values for the tuning parameters that are suitable for the actual process.

A more suitable approach is to define an objective function from these performance criteria. By using simulation data and experimental data from previous runs, the optimal values for the tuning parameters were found by using an optimisation routine. The result is a trade off between the estimation errors and the sensitivities to noise that stays within the boundaries mentioned in equations (12) and (21).

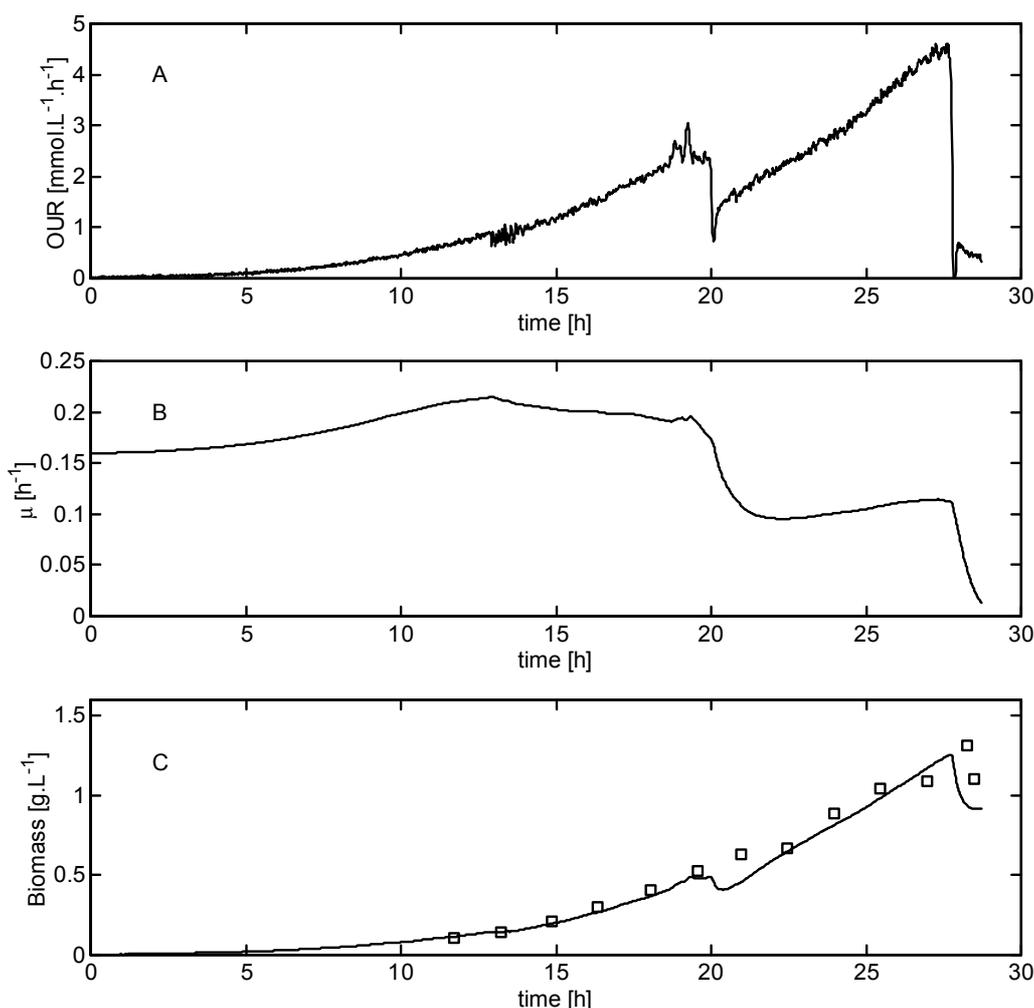


Figure 4: Sequentially estimated batch cultivation of *B. pertussis*. Top graph (A): The measured OUR during cultivation. Middle graph (B): estimated specific growth rate. Bottom graph (C): The estimated (solid) and measured (triangles) biomass concentrations.

Experimental Validation

Several experiments with *B. pertussis* were performed in order to test the performance of the sequential estimator. Figure 3 shows the results for batch cultivation. During the first part of this batch, the relative noise on the *OUR*-signal is high and consequently the biomass concentration is under-estimated during the first part of the batch. Clearly, the quality of the *OUR*-data is crucial for the performance of the sequential estimator. Towards the end of the batch cultivation, the relative noise on the *OUR*-signal decreases and the biomass estimations converge to their measured values. *B. pertussis* is interactively limited by two substrates (glutamate and lactate) as reported by

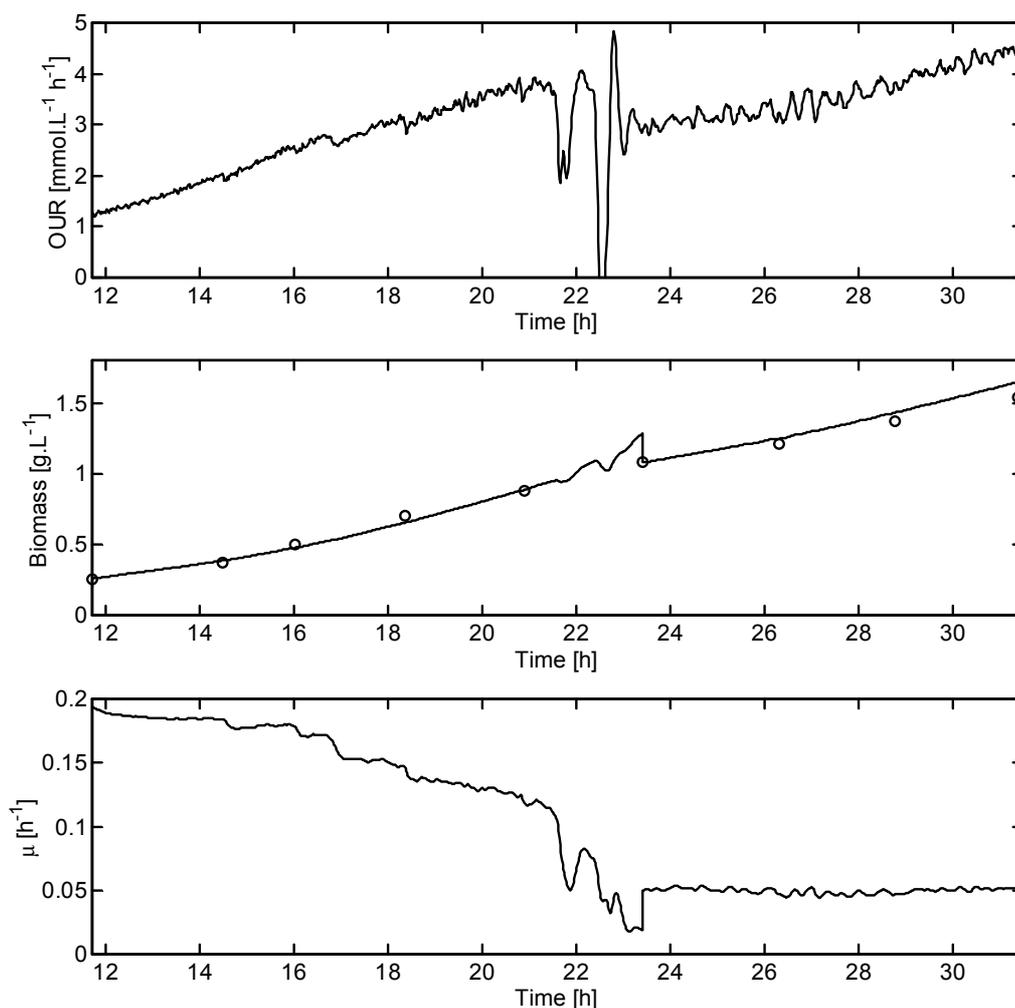


Figure 5: Sequentially estimated batch cultivation of *B. pertussis*. Top graph (A): The measured OUR during cultivation. Middle graph (B): The estimated (solid) and measured (triangles) biomass concentrations. Bottom graph (C): The estimated specific growth rate.

Neeleman et al. [19]. After approximately 21 hours, the first substrate (lactate) is depleted and 1 hour later, the second and final substrate (glutamate) is depleted. The *OUR*-signal clearly indicates these two moments of substrate depletion.

Figure 4 shows the results for another batch cultivation where after 20 hours the first substrate (lactate) is depleted. Since the organism is interactively limited by two substrates, the specific growth rate decreases with the depletion of lactate. Figure 4b shows the change of the specific growth rate around 20 hours, indicating that the sequential estimator is capable of recognising a changing specific growth rate. Figure 4c shows a good agreement between the estimated and measured biomass concentration. Clearly the decreased level of

noise (due to a more accurate dissolved oxygen sensor and better tuning of the dissolved oxygen controller) improved the performance of the sequential estimator.

To illustrate the applicability of the sequential estimator for fed-batch cultivation a final experiment was conducted where glutamate and lactate were fed into the bioreactor. Figure 5a shows the measured *OUR* during this fed-batch cultivation. Until 23 hours, the cultivation was performed in batch. Around 21.5 hours the first substrate, lactate was depleted, and growth continued for just one hour on the remaining glutamate. Then around 23 hours the feed was started. Since these events happened in rapid succession, the control actions of the dissolved oxygen controller fluctuated strongly. Consequently, the *OUR*-signal fluctuated strongly as well, which subverted the performance of the sequential estimator. Therefore after 23.5 hours, shortly after the feed started, the state of the sequential estimator was reset with the measured biomass concentration. Figure 5b shows the estimated and measured biomass concentrations and figure 5c shows the estimated specific growth rate. With this experiment, the performance of the sequential estimator is proven for fed-batch cultivation.

6.8 Conclusions

This chapter has considered the problems of designing a sequential estimator for the specific growth rate and biomass concentration for a class of aerobic microbial growth processes. The results are relevant to the implementation of control laws.

The generally good agreement between the on-line estimates and the off-line measurements of biomass concentration indicated that the proposed combination of algorithms is useful in the estimation of the biomass concentration and its specific growth rate. Furthermore, the flexibility and the general adaptive features of these algorithms suggest that they could be used in the estimation of time-varying parameters of other situations as well.

By modelling biomass in the second observer as a random walk process with coloured noise, the response is quicker but a tendency to overshoot the true value might be present. Stephanopoulos and San [18] suggested the use of a damping force in the description of R_X to correct a possible overshoot.

The sequential estimator turned out to be stable, reliable in operation, and suitable for application in a system for prediction and/or adaptive control.

6.9 Nomenclature

C_X	$\text{g}\cdot\text{L}^{-1}$	biomass concentration
D	h^{-1}	dilution rate
E	-	vector of estimation errors
m_O	$\text{mmol}\cdot\text{g}^{-1}\cdot\text{h}^{-1}$	maintenance coefficient for oxygen
OUR	$\text{mmol}\cdot\text{L}^{-1}\cdot\text{h}^{-1}$	oxygen uptake rate
OUR_m	$\text{mmol}\cdot\text{L}^{-1}\cdot\text{h}^{-1}$	measured oxygen uptake rate
$p_{\#}$	-	tuning parameter #
R_X	$\text{g}\cdot\text{L}^{-1}\cdot\text{h}^{-1}$	biomass growth rate
t	h	time
Y_O	$\text{g}\cdot\text{mmol}^{-1}$	yield of biomass on oxygen
<i>Greek symbols</i>		
α	h^{-1}	apparent specific growth rate
ε_{α}	h^{-1}	estimation error of α
ε_m	$\text{mmol}\cdot\text{L}^{-1}\cdot\text{h}^{-1}$	OUR measurement error
ε_{OUR}	$\text{mmol}\cdot\text{L}^{-1}\cdot\text{h}^{-1}$	estimation error of OUR
μ	h^{-1}	specific growth rate

6.10 References

1. Cazzador, L. and V. Lubenova (1995). Nonlinear Estimation of specific Growth Rate for Aerobic Fermentation Processes. *Biotechnology and Bioengineering* 47: 626-632.
2. Holmberg, A. and J. Ranta (1982). Procedures for Parameter and State Estimation of Microbial Growth Process Models. *Automatica* 18(2): 181-193.
3. Shimizu, H., T. Takamatsu, S. Shioya and K.I. Suga (1989). An algorithmic approach to constructing the on-line estimation system for the specific growth rate. *Biotechnology and Bioengineering* 33: 354-364.

4. Bastin, G. and D. Dochain (1986). On-line Estimation of Microbial Specific Growth Rates. *Automatica* 22(6): 705-709.
5. Claes, J.E. and J.F. Van Impe (1999). On-line estimation of the specific growth rate based on viable biomass measurements: experimental validation. *Bioprocess Engineering* 21: 389-395.
6. Farza, M., S. Othman, H. Hammouri and J. Biston (1997). A nonlinear approach for the on-line estimation of the kinetic rates in bioreactors - Application to a lactic acid production process. *Bioprocess Engineering* 17(3): 143-150.
7. Estler, M.U. (1995). Recursive on-line estimation of the specific growth rate from off-gas analysis for the adaptive control of fed-batch processes. *Bioprocess Engineering* 12: 205-207.
8. Stephanopoulos, G. and K.Y. San (1984). Studies on On-Line Bioreactor Identification. I. Theory. *Biotechnology and Bioengineering* 26: 1176-1188.
9. Pons, M.N. (1992). *Bioprocess monitoring and control*. Oxford UK, Hanser Publishers, Oxford University Press.
10. Goodwin, G.C. and K.S. Sin (1984). *Adaptive filtering prediction and control*. Englewood Cliffs, NJ, Prentice-Hall.
11. Chattaway, T. and G. Stephanopoulos (1989). An adaptive state estimator for detecting contaminants in bioreactors. *Biotechnology and Bioengineering* 34: 647.
12. Park, S. and W.F. Ramirez (1989). Optimal regulatory control of bioreactor nutrient concentration incorporating system identification. *Chemical Engineering Science*.
13. Stephanopoulos, G. and S. Park (1990). Bioreactor State Estimation. *Biotechnology, Measuring Modelling and Control*. H.-J. Rehm and G. Reed. Weinheim Germany, VCH. 4: 225-250.
14. Lubenova, V., M. Ignatova and S. Tsonkov (1993). On-Line Estimation of Specific Growth Rate for a Class of Aerobic Batch Processes. *Chemical and Biochemical Engineering Quarterly* 7(2): 101-106.
15. Neeleman, R., M. Joerink, A.J.B. Van Boxtel, C. Beuvery and G. Van Straten (2001). On-line harvest prediction for the production of biologicals, *Computer Applications in Biotechnology*, Quebec, Canada, 430-435.

16. Stephanopoulos, G. (1984). *Chemical Process Control: An Introduction to Theory and Practice*. London, UK, Prentice-Hall International Inc.
17. Ignatova, M. and V. Lubenova (1995). An approach for estimation of biotechnological processes variables. *Chemical and Biochemical Engineering Quarterly* 9(1): 19-25.
18. Stephanopoulos, G. and K.Y. San (1982). On-Line Estimation of Time-Varying Parameters Application to Biochemical Reactors. IFAC, *Modelling and Control of Biotechnological Processes*, Helsinki Finland, 195-199.
19. Neeleman, R., M. Joerink and A.J.B. Van Boxtel (2001). Dual-Substrate Utilisation By *Bordetella pertussis*. *Applied Microbiology and Biotechnology* 57: 489-493.

7 ONLINE HARVEST PREDICTION FOR THE PRODUCTION OF BIOLOGICALS

This chapter is based on:

R. Neeleman, M. Joerink, A.J.B. van Boxtel, C. Beuvery, and G. van Straten (2001). On-line harvest prediction for the production of biologicals. *Computer Applications in Biotechnology*, Quebec, Canada, pp. 403-435.

7.1 Abstract

To achieve good product quality and consistency the bioreactor has to be harvested in time. A prediction algorithm is developed which aids the operator in determining and scheduling the right moment of harvest. An observer was build for biomass concentration and specific growth rate from oxygen consumption measurements only. With these estimations and a dual-substrate kinetic model the substrate concentrations are calculated. With the estimated biomass and calculated substrate concentrations the future trajectory of the cultivation is computed by forward simulation to successfully determine the harvest time on-line. Several experiments with a benchscale bioreactor to cultivate *Bordetella pertussis* were carried out.

7.2 Introduction

A major goal in pharmaceutical industry, also the subject of local regulations, is to ensure a high level of consistency. When process consistency is mainly determined by batch cultivation, the only manipulatable operation is the moment of harvest. Logically one should not harvest too early in the process to avoid poor yields. And since the quality of the product usually deteriorates within several hours after substrate depletion it is hence of great importance to harvest in time.

The production process of a vaccine against whooping cough is used to investigate the possibility of harvest prediction. Oxygen consumption proves to be a good indicator for the time to harvest. As soon as substrates are depleted the oxygen consumption decreases rapidly. However, *B. pertussis*, the organism responsible for whooping cough, grows on two major substrates, lactate and glutamate. When lactate is depleted a short decrease in oxygen consumption is observed. If there is still glutamate available the organism continues growth until this component is depleted. To properly determine whether both substrates are depleted, a reaction time of at least half an hour is necessary. Besides this long reaction time, also the ability to schedule downstream processing favours the ability to predict the harvest time well in time.

To predict the harvest time a model is needed that describes biomass growth and consumption of glutamate and lactate. A regularly scheduled or operator

requested forward simulation of this model has to predict the harvest time. However, the lack of on-line in situ measurements of the biomass and substrate concentrations makes it impossible to determine the current state of the cultivation, which is needed by the simulation routine.

A second model relating biomass and growth rate to oxygen consumption has to be derived and used to build an observer for specific growth rate and biomass concentration from oxygen uptake rate (*OUR*) measurements only. The *OUR* is on-line calculated from dissolved oxygen and off-gas measurements. Combining such biomass estimation with a multi-substrate kinetic model the substrates can be calculated on-line. Based on the estimated current state of the process a forward simulation to predict the moment of substrate depletion can be carried out at any time during the cultivation.

7.3 Materials and Methods

Strain and Culture Medium

B. pertussis strain 509 which is one of the two strains included in the DPT-Polio vaccine applied in the Netherlands, was used with a chemically defined medium containing glutamate and L-lactate as main carbon substrates [1].

Bioreactor Operation

The cells were grown in a 5-litre round bottom reactor containing 3-litre medium. An eight-bladed marine impeller was used to agitate the medium. Temperature, pH and dissolved oxygen were controlled at 34°C, 7.2 and 20% air saturation respectively. Oxygen was transported through the headspace only and controlled by changing the oxygen fraction in the gas flow. The total gas flow was kept constant at 1·l·min⁻¹.

Analysis

A polarographic electrode (Ingold, Urdorf, Switzerland) measured the dissolved oxygen in the medium. A pH electrode (Ingold, Urdorf, Switzerland) measured the pH. The temperature was measured with a Pt100 temperature sensor. Biomass was measured off-line by optical density using a Vitalab 10 (Vital Scientific, Dieren, The Netherlands), at 590 nm. The optical densities of *B. pertussis* were recalculated to mmol·l⁻¹ according to a

calibration curve. Glutamate and L-lactate were determined off-line with a YSI 2750 Select analyser (Yellow Springs Instruments, Yellow Springs, USA).

Software

Matlab (Mathworks, Massachusetts, USA) was used for the development of the model, the observers, the simulation routines, and for visualisation of the results.

7.4 Dual-Substrate Kinetic Model of *B. pertussis*

The medium used for *B. pertussis* cultivation contains glutamate and lactate as main carbon sources. *B. pertussis* is not capable of growing on lactate only, which makes lactate not an essential substrate. Contrarily growth on glutamate only is possible and so one would conclude growth to be limited by glutamate only. However, since the growth rate decreases when lactate is depleted and moreover the growth rate on glutamate only is lower than on both substrates, a combined essential and additive model [2] is derived for *B. pertussis* and shown by equation 1.

$$\begin{aligned}\frac{dC_X}{dt} &= (\mu_G + \mu_L)C_X \\ \frac{dC_G}{dt} &= -\left(\frac{\mu_G}{Y_{G1}} + \frac{\mu_L}{Y_{G2}}\right)C_X \\ \frac{dC_L}{dt} &= -\frac{\mu_L}{Y_L}C_X\end{aligned}\tag{1}$$

Where C_X , C_G and C_L are the concentrations of biomass, glutamate and lactate respectively. μ_G denotes the specific growth rate on glutamate only and μ_L the enhanced growth rate in the presence of lactate:

$$\begin{aligned}\mu_G &= \mu_G^{max} \frac{C_G}{K_G + C_G} \\ \mu_L &= \mu_L^{max} \frac{C_G}{K_G + C_G} \frac{C_L}{K_L + C_L}\end{aligned}\tag{2}$$

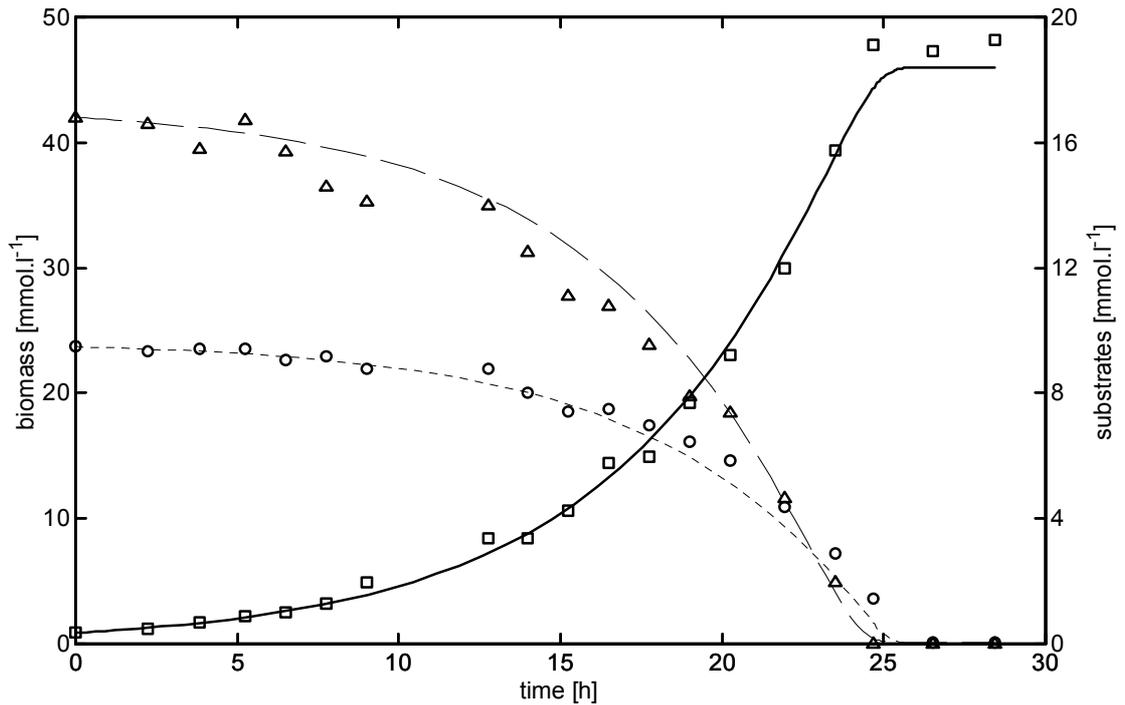


Figure 1: Measured biomass (squares), glutamate (circles) and lactate (triangles) concentrations during batch cultivation together with the simulation results (solid).

The parameters are determined by a step-wise experimentation procedure. First a cultivation experiment was carried out with glutamate as the only main substrate. Since the organism is capable of growing on glutamate only, the parameters linked to glutamate consumption could be calculated separately. The yield, Y_{G1} , was determined by linear regression of glutamate against biomass concentrations and the maximum specific growth rate, μ_G^{max} , could be determined by fitting the growth curve to an exponential function.

A second experiment with both glutamate and lactate was used to determine the additive effect of lactate to the specific growth rate, μ_L^{max} . And the final yields, Y_{G2} and Y_L , could again be determined by linear regression. Table 1 shows the derived parameters. The negative value of the glutamate yield in the lactate pathway, Y_{G2} , indicates a decreased maintenance on glutamate in the presence of lactate.

The performance of the dual-substrate model is demonstrated by a third experiment, Figure 1 shows the measured biomass, glutamate and lactate concentrations together with the simulated concentrations using the multi-substrate model. Clearly this figures shows that the model is capable of

Table 1: Parameter values of the dual-substrate kinetic model

parameter	value
μ_G^{max}	0.12 h ⁻¹
K_G	0.50 mmol.l ⁻¹
Y_{G1}	2.243 mol.mol ⁻¹
Y_{G2}	-2.5 mol.mol ⁻¹
μ_L^{max}	0.055 h ⁻¹
K_L	0.50 mmol.l ⁻¹
Y_L	0.75 mol.mol ⁻¹

accurately predicting the development through time of the biomass and substrate concentrations and thus of the harvest time.

To predict the future trajectory of the cultivation, the current state of the biomass and substrate concentrations is needed. Since these are not measured on-line, they have to be estimated from e.g. the measured oxygen consumption. However, adding oxygen consumption to this model does not result in an observable system, because it is not possible to determine from which substrate the biomass was formed. Therefore a generic biomass observer will be used to estimate the biomass concentration on-line. The development through time of the substrate concentrations will then be calculated by simulation with the previous mentioned model and the estimated biomass concentration.

7.5 Observer for Specific Growth Rate

A generic model for biomass growth and oxygen consumption was used to construct the observer. This model, equation 3, describes oxygen consumption as the sum of oxygen used for growth and oxygen used for maintenance.

$$\frac{dC_X}{dt} = \mu C_X$$

$$OUR = \frac{\mu}{Y_O} C_X + m_O C_X \quad (3)$$

Where Y_O is the yield of oxygen on biomass growth and m_O the maintenance coefficient of oxygen. These parameters were determined from experimental data and are shown in Table 1. The derivative of the OUR can be calculated as follows:

$$\begin{aligned}\frac{dOUR}{dt} &= \left(\frac{\mu}{Y_O} + m_O \right) \frac{dC_X}{dt} + \frac{d\left(\frac{\mu}{Y_O} + m_O \right)}{dt} C_X \\ &= \left(\frac{\mu}{Y_O} + m_O \right) \mu C_X + \frac{d\mu}{dt} \frac{C_X}{Y_O}\end{aligned}\quad (4)$$

Rewriting this equation with equation 3 gives:

$$\frac{dOUR}{dt} = \left(\mu + \frac{\frac{d\mu}{dt}}{\mu + Y_O m_O} \right) OUR = \alpha \cdot OUR \quad (5)$$

Where the auxiliary parameter α , is the 'apparent specific growth rate', such that:

$$\frac{d\mu}{dt} = (\alpha - \mu) (\mu + Y_O m_O) \quad (6)$$

Since, the development through time of the apparent specific growth rate, α , is unknown it is modelled as a random walk process. The final growth rate observer as described by Ignatova and Lubenova [3, 4] is described by equation 7:

$$\begin{aligned}\frac{dO\hat{U}R}{dt} &= \hat{\alpha} OUR_m + p_1 OUR_m (OUR_m - O\hat{U}R) \\ \frac{d\hat{\mu}}{dt} &= (\hat{\alpha} - \hat{\mu}) (\hat{\mu} + Y_O m_O) \\ \frac{d\hat{\alpha}}{dt} &= p_2 OUR_m (OUR_m - O\hat{U}R)\end{aligned}\quad (7)$$

Where p_1 is a tuning parameter describing the trade-off between the noise of the measurements, OUR_m , and the 'noise' of the model. The random walk process of the apparent specific growth rate, α , is depicted by the tuning parameter p_2 . The tuning parameters are chosen such that the observer works properly and is stable (Table 2). The system is called stable when the real parts of the eigenvalues of the linearised observer matrix are negative or zero [5]. Both eigenvalues of this system can be made negative or zero by the tuning parameters:

$$\begin{aligned} p_1 &\geq 2 \cdot \sqrt{p_2} \\ p_2 &\geq 0 \end{aligned} \quad (8)$$

7.6 Observer for Biomass

The estimated specific growth rate, its derivative and the measured *OUR* are used to construct an observer for the biomass concentration according Ignatova and Lubenova [3, 4]. Equation 9 shows the model used by this observer:

$$\begin{aligned} \frac{dOUR}{dt} &= \left(\frac{\mu}{Y_O} + m_O \right) \mu C_X + \frac{d\mu}{dt} \frac{C_X}{Y_O} \\ R_X &= \mu C_X \end{aligned} \quad (9)$$

Where R_X equals the growth rate of biomass, which is modelled as a random walk process. Equation 10 shows the derived biomass observer:

$$\begin{aligned} \frac{dO\bar{U}R}{dt} &= \left(\frac{\hat{\mu}}{Y_O} + m_O \right) \bar{R}_X + \frac{d\hat{\mu}}{dt} \frac{\bar{C}_X}{Y_O} + p_3 (OUR_m - O\bar{U}R) \\ \frac{d\bar{C}_X}{dt} &= \bar{R}_X + p_4 (OUR_m - O\bar{U}R) \\ \frac{d\bar{R}_X}{dt} &= p_5 (OUR_m - O\bar{U}R) \end{aligned} \quad (10)$$

Where p_3 indicates the trade-off between noise from the *OUR*-model and its measurement, p_4 indicates the correction of the error prediction to the biomass model and p_5 depicts the random walk characteristics of R_X (Table 2).

The parameters are again chosen such that the observer works properly and is stable [3]. All three eigenvalues can be manipulated by the tuning parameters to be negative or zero, and thus to make the system stable:

$$\begin{aligned} p_3 &\geq 2 \\ p_4 &\geq \frac{\mu}{Y_O} + m_O + 1 \\ p_5 &\geq \frac{1}{Y_O} \cdot \frac{d\mu}{dt} + 4 \end{aligned} \quad (11)$$

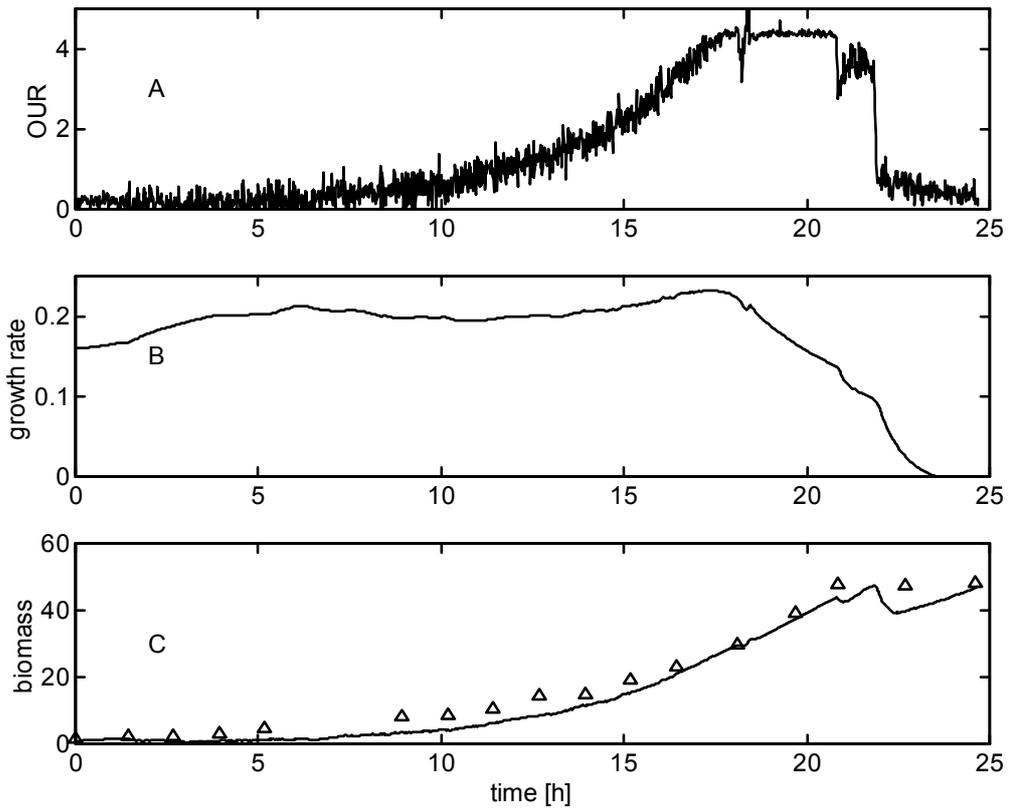


Figure 2: A: measured OUR. B: estimated growth rate. C: estimated (solid) and measured (triangles) biomass concentration.

These two observers, equations 7 and 10, are used to estimate biomass concentrations on-line for *B. pertussis* cultivations from *OUR* measurements only. The two observers are initialised with the maximal specific growth rate known from the model and with the inoculation biomass concentration.

Results are shown by Figure 2. Figure 2a shows the measured *OUR* that is used to estimate the specific growth rate and biomass concentration. Figure 2b shows the estimated specific growth rate. The estimated specific growth rate is

Table. 2: Parameter values of the oxygen model and the two observers

parameter	value
Y_O	$1.67 \text{ mmol.mmol}^{-1}$
m_O	0.01 h^{-1}
p_1	16.76
p_2	7.81
p_3	2.65
p_4	12.78
p_5	15.65

nearly constant during the first 18 hours, then it decreases slowly until 23 hours when the growth rate is zero. Figure 2c shows the measured and estimated biomass concentration. At the beginning of the batch the biomass concentration is estimated too low, but after 15 hours the estimation follows the measurements more accurately. Clearly this two-step observer is capable of on-line biomass estimation.

7.7 Observer based calculation of substrates

Using the estimated biomass concentration and the measured start-concentrations of the substrates the development through time of the substrate concentrations can be calculated with the dual-substrate model of equations 1 and 2.

The observers are initialised without taking a sample. The maximal specific growth rate as known from the model is used and the starting concentrations

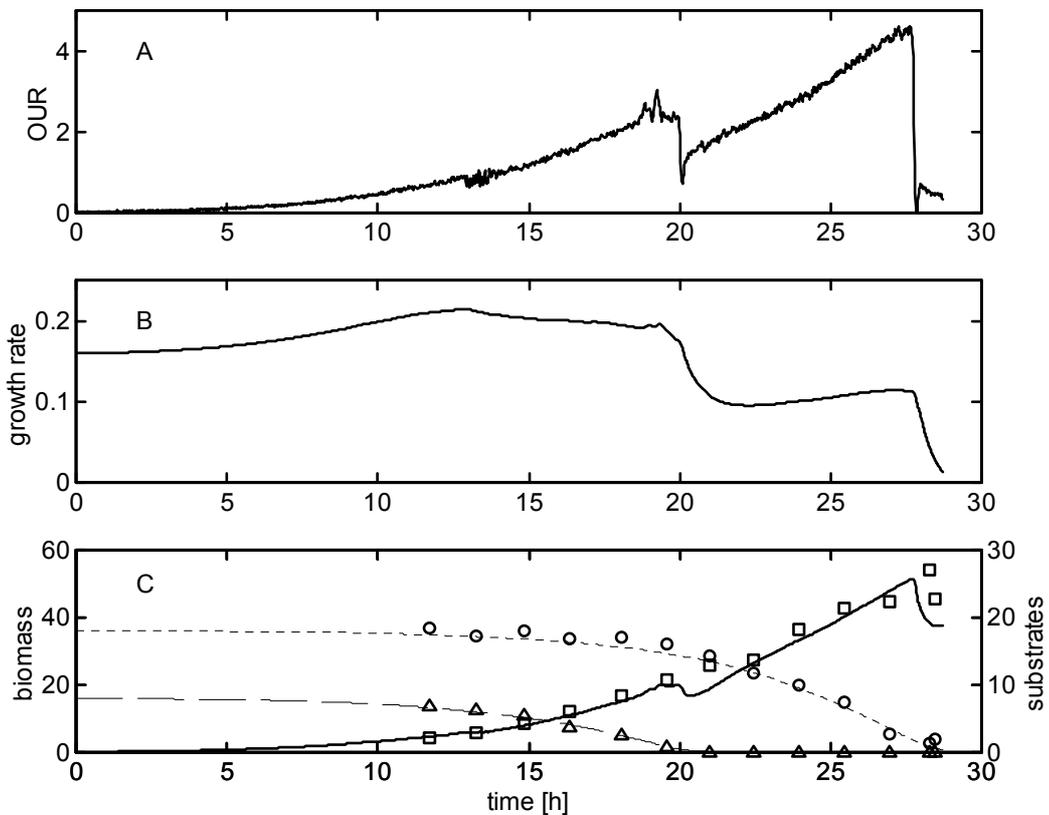


Figure 3: A: measured *OUR*. B: estimated growth rate. C: estimated concentrations (lines) and measurements of biomass (squares), lactate (triangles) and glutamate (circles).

of the substrates as they were added to the medium. The starting concentration of biomass was set to zero. Figure 3a shows the measured *OUR* during a cultivation with extra glutamate. Halfway during this cultivation, at 19 hours, lactate is depleted as shown by the sudden decrease in *OUR*. This lactate depletion is also shown by the estimated growth rate in Figure 3b; it drops to the maximum specific growth rate of glutamate only (0.1 h^{-1}). The estimated biomass concentration and the calculated substrate concentrations are shown in Figure 3c.

Clearly this combination of observers and model based substrates calculation is capable of giving good estimates of the current state of the cultivation without taking a sample for initialisation. Subsequently these results can be used to predict the future development of the concentrations and thus the harvest time. Besides harvest time prediction this combination of observers has a high potential to be used for control purposes.

7.8 Observer based prediction of harvest time

To predict the harvest time the current estimated biomass concentration and calculated substrate concentrations are used for a simulation to determine when the substrates are depleted. This simulation runs when requested by the operator or on a scheduled interval. Figure 4 shows a diagram of the complete harvest prediction procedure together with the used equations. Figure 5 shows the error of the predicted harvest time for two cultivations as it is calculated every hour.

This figure shows a late harvest time prediction at the start of the cultivation due to an estimated lag phase. Following the estimated lag phase the biomass

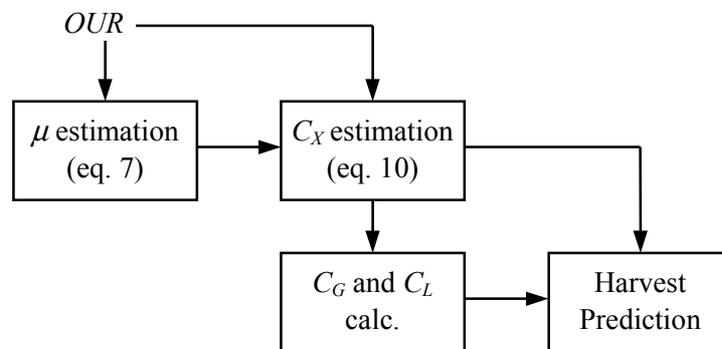


Figure 4: Schematic representation of the step-wise harvest prediction procedure.

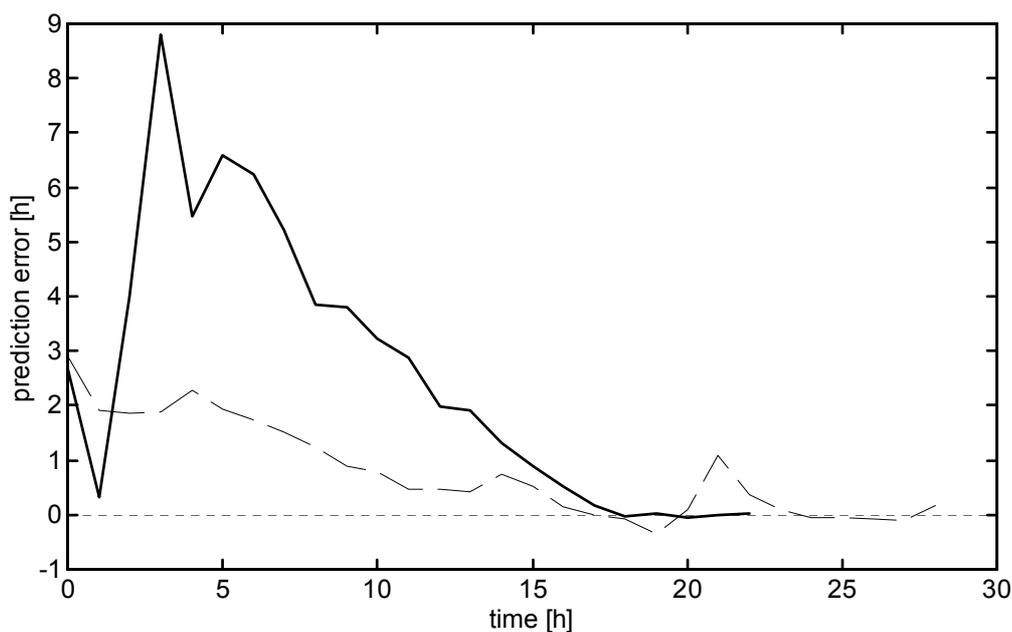


Figure 5: Error of the predicted harvest time at the moment of prediction for two batches.

accumulates faster than predicted by the multi-substrate model. As a result the harvest time prediction decreases. For both cultivations, 8 hours before harvest time, the error in prediction is less than 30 minutes.

7.9 Concluding remarks

To achieve good product quality and consistency the bioreactor has to be harvested in time. For processes with varying cultivation time a regular or requested prediction of the time to harvest aids the operator in determining the right moment of harvest and scheduling the following steps of downstream processing.

Based on theoretical inference and a step-wise experimentation procedure a dual-substrate kinetic model was developed to describe the growth of *B. pertussis* and the consumption of the two substrates. This model was shown to give a good prediction of a cultivation with different initial substrate concentrations.

A second model was derived which related biomass, growth and oxygen consumption. This model was used to build a two-step observer that was shown to give a good estimation of the growth rate and biomass concentration on-line based on *OUR* measurements only.

A combination of the dual-substrate kinetic model and the two-step observer was shown to give a good prediction of the current state, e.g. biomass, and substrate concentrations, of the cultivation without taking a sample for initialisation. This current state was used as starting point for a simulation with the dual-substrate kinetic model to determine the harvest time. Determination of harvest time can be done on a scheduled interval or on an operator request. It was shown that the harvest prediction adjusts to lag phases and varying growth rates and is a powerful indicator in the production process of biologicals.

7.10 Nomenclature

C_G	$\text{mmol}\cdot\text{l}^{-1}$	glutamate concentration
C_L	$\text{mmol}\cdot\text{l}^{-1}$	lactate concentration
C_X	$\text{mmol}\cdot\text{l}^{-1}$	biomass concentration
\bar{C}_X	$\text{mmol}\cdot\text{l}^{-1}$	estimated biomass concentration
K_G	$\text{mmol}\cdot\text{l}^{-1}$	glutamate saturation concentration
K_L	$\text{mmol}\cdot\text{l}^{-1}$	lactate saturation concentration
m_O	h^{-1}	maintenance of biomass on oxygen
OUR	$\text{mmol}\cdot\text{l}^{-1}\cdot\text{h}^{-1}$	oxygen uptake rate
OUR_m	$\text{mmol}\cdot\text{l}^{-1}\cdot\text{h}^{-1}$	measured oxygen uptake rate
$\hat{O}UR$	$\text{mmol}\cdot\text{l}^{-1}\cdot\text{h}^{-1}$	oxygen uptake rate estimated by the growth observer
$\bar{O}UR$	$\text{mmol}\cdot\text{l}^{-1}\cdot\text{h}^{-1}$	oxygen uptake rate estimated by the biomass observer
$p_{\#}$	-	observer tuning parameter #
R_X	$\text{mmol}\cdot\text{l}^{-1}\cdot\text{h}^{-1}$	biomass growth rate
\bar{R}_X	$\text{mmol}\cdot\text{l}^{-1}\cdot\text{h}^{-1}$	biomass growth rate
Y_{G1}	$\text{mol}\cdot\text{mol}^{-1}$	yield of biomass on glutamate for pathway 1
Y_{G2}	$\text{mol}\cdot\text{mol}^{-1}$	yield of biomass on glutamate for pathway 2
Y_L	$\text{mol}\cdot\text{mol}^{-1}$	yield of biomass on lactate
Y_O	$\text{mol}\cdot\text{mol}^{-1}$	yield of biomass on oxygen
α	h^{-1}	apparent specific growth rate
$\hat{\mu}$	h^{-1}	estimated specific growth rate
μ_G	h^{-1}	specific growth rate on glutamate

μ_G^{max}	h^{-1}	maximal specific growth rate on glutamate
μ_L	h^{-1}	specific growth rate on lactate
μ_L^{max}	h^{-1}	enhancing specific growth rate on lactate

7.11 References

1. Thalen, M., J. Van De IJssel, W. Jiskoot, B. Zomer, P. Roholl, C. De Gooijer, C. Beuvery and J. Tramper (1999). Rational medium design for *Bordetella pertussis*: basic metabolism. *Journal of Biotechnology* 75(2-3): 147-159.
2. Mankad, T. and E.B. Nauman (1992). Modeling of microbial growth under dual limitations. *The Chemical Engineering Journal* 48: B9-B11.
3. Ignatova, M. and V. Lubenova (1995). An approach for estimation of biotechnological processes variables. *Chemical and Biochemical Engineering Quarterly* 9(1): 19-25.
4. Lubenova, V., M. Ignatova and S. Tsonkov (1993). On-Line Estimation of Specific Growth Rate for a Class of Aerobic Batch Processes. *Chemical and Biochemical Engineering Quarterly* 7(2): 101-106.
5. Stephanopoulos, G. (1984). *Chemical Process Control: An Introduction to Theory and Practice*. London, UK, Prentice-Hall International Inc.

8 DUAL-SUBSTRATE FEEDBACK CONTROL OF THE SPECIFIC GROWTH RATE

8.1 Abstract

For pharmaceutical production, consistency is a major goal. To achieve this goal during the cultivation process, the metabolic state of the micro-organisms in the bioreactor has to be kept constant. By controlling the specific growth rate of biomass such metabolic stability is guaranteed. This chapter describes the combined use of a sequential estimator and feed-forward-feedback-controller to keep the specific growth rate constant for a dual-substrate utilising organism. The performance of the combined algorithm is demonstrated by several lab-scale experiments with *Bordetella pertussis*.

8.2 Introduction

The pharmaceutical industry is specially interested in consistency in production. With standard cultivation techniques the metabolic activity of the organisms may change, e.g. due to the varying amount of substrates. To achieve more consistent production, methods to enforce metabolic stability are needed. The specific growth rate characterises the metabolic state of the organisms. Thus proper control of the specific growth rate would increase the consistency of the production process.

Such a controller (μ -controller) should keep the specific growth rate at a desired set-point by adjusting the feed rate of nutrients on-line during fed-batch or continuous cultivation. However, there are no direct sensors for the specific growth rate and for many applications there is a lack of reliable on-line instruments to measure biomass concentration in order to deduce the specific growth rate. Therefore to meet this problem a model-supported estimation technique, software sensor [1], is needed to determine the specific growth rate from other available measurements such as oxygen consumption rate, *OUR* [2]. Fig 1 depicts a scheme of such a set-up with the process, the estimator and the controller.

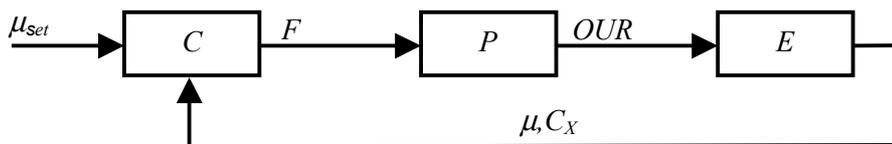


Figure 1: Diagram of a specific growth rate controller consisting of the process P , an estimator E , and a controller C . The biomass concentration C_x and the specific growth rate μ are estimated based on the oxygen uptake rate OUR to calculate the control feed rate F in order to keep the specific growth rate at set-

For fed-batch cultivation the output of the μ -controller in figure 1 must generate an exponentially increasing feed rate in order to keep the specific growth rate constant. Thus with a near constant input of the error between the set-point and the estimated specific growth rate the controller should require permanent adaptation of the settings if it is a standard feedback controller.

A strategy for μ -control frequently discussed in biotechnology is a feed-forward controller based on a pre-determined trajectory or on estimations of the biomass concentrations [3-5]. However such a feed-forward controller needs a very accurate model and does not react to mismatches between the predetermined cell growth and the actual cell growth or unexpected disturbances on the system.

Therefore to compensate for the effects of such mismatches robustness is improved by adding a feedback-controller to the feed-forward controller. Such a feedback controller can be a simple and direct PI-controller. Several reports have been made of combined feed-forward-feedback-control of the specific growth rate based on estimations [6-11]. However, these cases use on-line measurements of states (like biomass or substrate concentrations) and/or discuss single-substrate-limiting situations. While in many cases such on-line measurements are not available and multiple-substrates-limiting situations apply.

This study concerns the development of a feed-forward-feedback- μ -controller for the cultivation of a micro-organism interactively limited by two substrates. This complicates the development of a μ -controller, since it involves the control of two separate feeds. The performance of the controller is shown by actual experimentation with *B. pertussis*, the organism responsible for whooping cough. A dual-substrate model [12] and a sequential estimator for specific growth rate and biomass concentration based on measurements of *OUR* [13, 2] were developed and are integrated in the μ -controller.

8.3 Materials and Methods

Strain, Medium and Bioreactor Conditions

B. pertussis strain 509 (RIVM collection, Bilthoven, The Netherlands), one of the two strains included in the DPT-Polio vaccine used in the Netherlands, was used with the chemically defined THIJS-medium containing glutamate

and lactate as main carbon sources (RIVM). The cells were grown in a 5-l round-bottomed glass reactor containing 3-l medium. An eight-bladed marine impeller was used to agitate the medium. Temperature, pH, dissolved oxygen, and stirrer speed were controlled at 34°C, 7.2, 20%, and 400 to 500 rpm, respectively. Oxygen was transported through the headspace only and controlled by changing the oxygen fraction in the gas flow. The total gas flow was maintained at 1-l.min⁻¹.

Two 500-ml concentrated stock solutions of glutamate and lactate were used for the two separate feeds and placed on balances to monitor the feed rate. The feeds were added to the bioreactor by two pumps (101U/R 32 rpm, Watson Marlow ltd., Cornwall UK) connected to the bioreactor control system.

Analysis

A polarographic electrode (Ingold, Urdorf, Switzerland) measured the dissolved oxygen in the medium. A pH electrode (Ingold) measured the pH. The temperature was measured with a Pt100 temperature sensor. Biomass was measured by optical density using a Vitalab 10 (Vital Scientific, Dieren, The Netherlands), at 590 nm and by dry weight. The optical densities of *B. pertussis* were recalculated to g·l⁻¹ dry weight according to a calibration curve. Glutamate and L-lactate were determined with a YSI 2750 Select analyser (Yellow Springs Instruments, Yellow Springs, USA).

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). All standard control-loops (dissolved oxygen, pH, temperature, etc.) were performed in the BCSV-software and logged to a disc on the UNIX machine. Using FTP, a Windows NT machine downloaded new data every minute. Data were processed by a routine written in Matlab (Mathworks, Massachusetts USA) to estimate biomass concentration and specific growth rate, and to calculate new set-points for the two pumps. Data were visualised in graphs on the Windows NT machine as well. The new set-points for the two pumps were then uploaded to the UNIX machine and to the pumps by BCSV.

8.4 Sequential Estimation

A significant part of the overall bioreactor identification is the estimation of unknown or uncertain model parameters. Usually in system identification this is accomplished by treating the parameters as additional state variables. In this study, a different state/parameter identification method is presented. It is based on the observation that, in an adequately formulated model, the model parameters are distinguished from the state variables in that they change markedly slower than the state variables, and sometimes they are, in fact, time invariant.

As a result a suitable estimation algorithm can be summarised as follows. First the parameters are estimated from the available measurements at a certain moment in time. This leads to an update of the system dynamic model. Then, the estimation of the state vector is made with the measurements and the updated model. This is called sequential state and parameter estimation, SSPE [14-16]. It should be noted that the underlying basis for the SSPE algorithm is the distinctively lower rate of parameter variation compared to that of the state variables. The parameterisation of μ might not be appropriate for SSPE application, since it can be strongly dependent upon culture conditions such as substrate concentrations. However, if the culture conditions are regulated at constant levels, it is reasonable to expect that this metabolic variable does not show considerable variation.

Estimation of Specific Growth Rate

A suitable asymptotic observer of the specific growth rate based on the measurements of the oxygen uptake rate OUR , is developed by Lubenova et al. [2] for batch processes. This observer uses a generic model for biomass growth and oxygen consumption that describes oxygen consumption as the sum of oxygen used for growth and oxygen used for maintenance. Eq. (1) shows this generic model after adjustment for fed-batch processes:

$$\begin{aligned} \frac{dC_X}{dt} &= (\mu - D)C_X \\ OUR &= \left(\frac{\mu}{Y_O} + m_O \right) C_X \end{aligned} \tag{1}$$

Where C_X is the concentration of biomass, μ the specific growth rate, D the dilution rate, Y_O the yield of oxygen on biomass growth, and m_O the maintenance coefficient for oxygen. The parameters for *B. pertussis* were determined from experimental data [17] and are shown in Table 1. The derivative of the oxygen uptake rate, OUR , can be calculated as follows:

$$\begin{aligned}\frac{dOUR}{dt} &= \left(\frac{\mu}{Y_O} + m_O \right) \frac{dC_X}{dt} + \frac{d\left(\frac{\mu}{Y_O} + m_O \right)}{dt} C_X \\ &= \left(\frac{\mu}{Y_O} + m_O \right) (\mu - D) C_X + \frac{d\mu}{dt} \frac{C_X}{Y_O}\end{aligned}\quad (2)$$

Rewriting this equation with Eq. (1) gives:

$$\begin{aligned}\frac{dOUR}{dt} &= \left(\mu - D + \frac{\frac{d\mu}{dt}}{(\mu + Y_O m_O)} \right) OUR \\ &= \alpha \cdot OUR\end{aligned}\quad (3)$$

Where α depicts the ‘apparent specific growth rate’ of OUR . Rewriting the equation for α gives:

$$\frac{d\mu}{dt} = (\alpha - \mu + D) (\mu + Y_O m_O) \quad (4)$$

Note that when the specific growth rate is constant, α equals $(\mu - D)$. An observer will now be used to identify the imperfectly known and possibly slowly varying parameter α . For this purpose the state of the system is extended to contain the differential equation of α with a zero right-hand part added for the unknown parameter ‘dynamics’. In other words, α is modelled as a random walk process. The μ -observer as described by Lubenova et al. [2] was developed for batch cultivation only. Adjusting that algorithm for fed-batch results in:

$$\begin{aligned}
 \frac{dO\hat{U}R}{dt} &= \hat{\alpha}OUR_m + p_1 OUR_m (OUR_m - O\hat{U}R) \\
 \frac{d\hat{\alpha}}{dt} &= 0 + p_2 OUR_m (OUR_m - O\hat{U}R) \\
 \frac{d\hat{\mu}}{dt} &= (\hat{\alpha} - \hat{\mu} + D)(\hat{\mu} + Y_O m_O)
 \end{aligned} \tag{5}$$

Where p_1 is a tuning parameter describing the trade-off between the noise of the measurements, OUR_m , and the 'noise' of the model. Tuning parameter p_2 depicts the characteristics of the random walk process of α . The tuning parameters are chosen (Table 1) such that the observer works properly and is stable. The system is called stable when the real parts of the eigenvalues of the linearised observer matrix are negative or zero [18]. As shown in detail in Chapter 6, this is guaranteed when:

$$\begin{aligned}
 p_1 &\geq 2 \cdot \sqrt{p_2} \\
 p_2 &\geq 0
 \end{aligned} \tag{6}$$

Estimation of Biomass Concentration

The estimated specific growth rate, its derivative, and the measured oxygen uptake rate OUR_m are used to construct an observer for the biomass concentration according to Ignatova and Lubenova [13]. First a model of the development through time of OUR and the biomass growth rate R_X is used:

$$\begin{aligned}
 \frac{dOUR}{dt} &= \left(\frac{\mu}{Y_O} + m_O \right) (\mu - D) C_X + \frac{d\mu}{dt} \frac{C_X}{Y_O} \\
 \frac{dC_X}{dt} &= (\mu - D) C_X = R_X
 \end{aligned} \tag{7}$$

From this model the following biomass observer is derived, note that C_X is modelled as a random walk process with coloured noise by modelling R_X , its derivative, as a random walk process:

$$\begin{aligned}
 \frac{dO\bar{U}R}{dt} &= \left(\frac{\hat{\mu}}{Y_O} + m_O \right) \bar{R}_X + \frac{d\hat{\mu}}{dt} \frac{\bar{C}_X}{Y_O} + p_3 (OUR_m - O\bar{U}R) \\
 \frac{d\bar{C}_X}{dt} &= \bar{R}_X + p_4 (OUR_m - O\bar{U}R) \\
 \frac{d\bar{R}_X}{dt} &= p_5 (OUR_m - O\bar{U}R)
 \end{aligned} \tag{8}$$

Where p_3 indicates the trade-off between noise from the *OUR* model and its measurement, p_4 indicates the correction of the error prediction to the biomass model and p_5 depicts the random walk characteristics of R_X (Table 1).

The parameters are chosen such that the observer works properly and is stable [13]. As shown in detail in Chapter 6, all three eigenvalues can be manipulated by the tuning parameters to be negative or zero, and thus to make the system stable:

$$\begin{aligned}
 p_3 &\geq 2 \\
 p_4 &\geq \frac{\mu}{Y_O} + m_O + D + 1 \\
 p_5 &\geq \frac{1}{Y_O} \cdot \frac{d\mu}{dt} + 4
 \end{aligned} \tag{9}$$

These two observers, equations (5) and (8), are combined sequentially to estimate the specific growth rate and biomass concentration on-line from *OUR* measurements. The observers are initialised with the maximal specific growth rate known from data, since it is assumed that the cells will start growing at this maximal growth rate, and the inoculation biomass concentration.

Table 1: Parameters used by the observers and dual-substrate model of *B. pertussis*.

parameter	value	parameter	value
Y_O	0.6 g·mmol ⁻¹	μ_{max}	0.12 h ⁻¹
m_O	0.01 h ⁻¹	K_G	0.50 mmol·l ⁻¹
p_1	16.76	Y_{G1}	42.35 g·mmol ⁻¹
p_2	7.81	Y_{G2}	120.29 g·mmol ⁻¹
p_3	2.65	μ_{enh}	0.032 h ⁻¹
p_4	12.78	K_L	0.50 mmol·l ⁻¹
p_5	15.65	Y_{L2}	10.92 g·mmol ⁻¹

Tuning

The problem of tuning the estimators is reduced to finding the proper values for p_1 to p_5 . For this purpose, the trade off between convergence and noise sensitivity is investigated in simulation and on data sets while staying within the boundaries as mentioned in equations (6) and (9).

8.5 Feed-forward Controller

Dual-Substrate Model

To construct a feed-forward controller a model is necessary which describes the substrate usage and growth of *B. pertussis*. Experimental studies showed that *B. pertussis* is not capable of utilising lactate as main carbon source. However, Chapter 4 showed that when grown on a medium with glutamate as main carbon source, addition of lactate increases the growth rate [19]. Based on these observations, in Chapter 5, a dual-substrate model for *B. pertussis* was developed [12]. The complete metabolism is lumped to the formation of biomass from glutamate and lactate from two main pathways. The first pathway concerns growth on glutamate only and the second pathway concerns growth on both substrates. Growth via these two pathways is assumed to be parallel, thus the individual specific growth rates can be added, furthermore, glutamate is essential for both pathways and Monod kinetics are assumed for the pathways. The model as described in Chapter 5 is modified to incorporate substrate feeding:

$$\begin{aligned}
 \frac{dC_X}{dt} &= (\mu_{max}f_G + \mu_{enh}f_Gf_L - D)C_X \\
 \frac{dC_G}{dt} &= \frac{F_G^{in}}{V}C_G^{in} - \frac{F_G^{in} + F_L^{in}}{V}C_G - \left(\frac{\mu_{max}f_G}{Y_{G1}} - \frac{\mu_{enh}f_Gf_L}{Y_{G2}} \right) C_X \\
 \frac{dC_L}{dt} &= \frac{F_L^{in}}{V}C_L^{in} - \frac{F_G^{in} + F_L^{in}}{V}C_L - \frac{\mu_{enh}f_Gf_L}{Y_{L2}} C_X \\
 \frac{dV}{dt} &= F_G^{in} + F_L^{in} - F_{sample}
 \end{aligned} \tag{10}$$

Where C_G and C_L are the concentrations of glutamate and lactate respectively, C_G^{in} and C_L^{in} are the incoming feed concentrations of glutamate and lactate

respectively, and Y_{G1} , Y_{G2} and Y_{L2} are the yields of biomass on glutamate over pathway 1, glutamate over pathway 2, and lactate over pathway 2 respectively. μ_{max} is the maximal specific growth rate over pathway 1, μ_{enh} is the maximal enhancing specific growth rate over pathway 2, and V is the liquid volume. All parameters are given in Table 1.

The kinetics are determined by f_G and f_L , which are normalised Monod kinetics for glutamate and lactate and are determined by the Monod saturation constants for glutamate, K_G , and lactate, K_L :

$$\begin{aligned} f_G &= \frac{C_G}{K_G + C_G} \\ f_L &= \frac{C_L}{K_L + C_L} \end{aligned} \tag{11}$$

Calculations of Feed-forward Controller

The goal of the feed-forward controller is to maintain a reference specific growth rate, μ^{ref} . This is achieved by adding exactly the amount of substrate that is needed by the organisms to grow with the reference specific growth rate.

Since metabolic stability is the underlying goal of the controller, both normalised Monod kinetics have to be kept constant at f_G^{ref} and f_L^{ref} . Choosing both kinetics smaller than one will make C_G and C_L negligible compared to C_G^{in} and C_L^{in} .

Based on the first measurement of the biomass concentration a feed profile can be determined for the rest of the fed-batch cultivation assuming exponential growth of biomass. Korz et al. [20] applied such an approach for a single-substrate case. Major disadvantage of this technique is the inability to adapt the feed rate to unexpected disturbances on the process. A better technique would be to use the estimated biomass concentration to calculate the correct feed rate on-line:

$$F_G^{in}(t) = \left(\frac{\mu_{max} f_G^{ref}}{Y_{G1}} - \frac{\mu_{enh} f_G^{ref} f_L^{ref}}{Y_{G2}} \right) \frac{\bar{C}_X(t) \cdot V(t)}{C_G^{in}} \quad (12)$$

$$F_L^{in}(t) = \frac{\mu_{enh} f_G^{ref} f_L^{ref}}{Y_{L2}} \frac{\bar{C}_X(t) \cdot V(t)}{C_L^{in}}$$

In these equations the unknowns are the kinetics f_G^{ref} and f_L^{ref} . In principle, there is an unlimited number of possible feed rates, because:

$$\mu^{ref} = \mu_{max} f_G + \mu_{enh} f_G f_L \quad (13)$$

However, since metabolic stability is the underlying goal, the ratio between the normalised kinetics should be fixed:

$$\beta = \frac{f_L}{f_G} \quad (14)$$

Figure 2 shows the values of f_G and f_L at various β and μ^{ref} . The white parts of the figure show the admissible regions of the normalised kinetics.

$$\max \left(0, \frac{\mu^{ref} - \mu_{max}}{\mu_{enh}} \right) < \beta < \frac{\mu_{max} + \mu_{enh}}{\mu^{ref}} \quad (15)$$

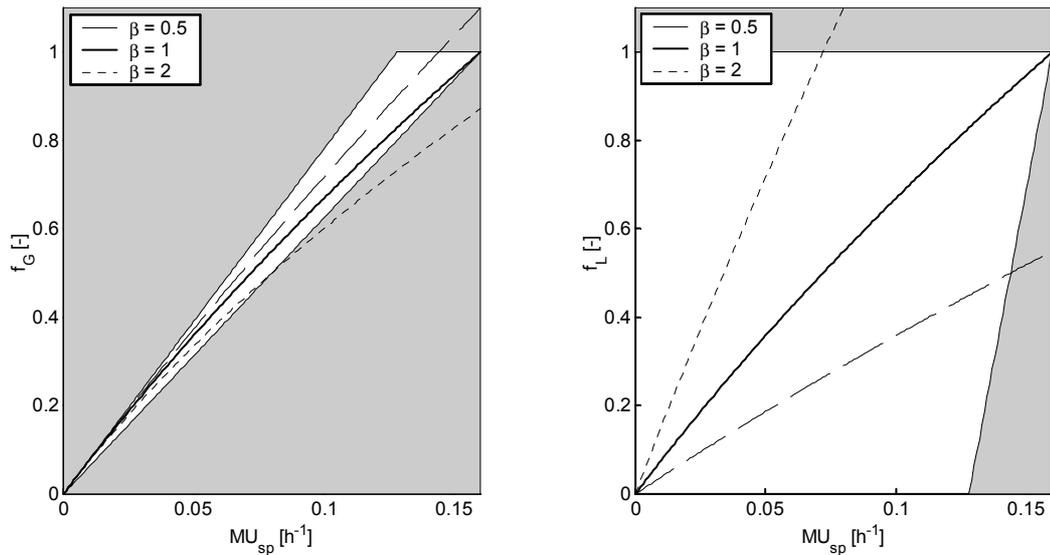


Figure 2: Diagram of the complete algorithm with process P , growth rate observer E_{μ} , biomass observer E_X , feed-forward controller C_{FF} , feedback controller C_{FB} , and supervisory boundary controller C_{SB} .

With a fixed ratio, β , the two feed rates for glutamate and lactate can be calculated. For the experiments in this study a ratio of 1 was used. Although the performance of a feed-forward controller is promising, a disadvantage of this feed-forward controller is the necessity to know all model-parameters exactly and these parameters are not allowed to change during cultivation. This dependency on an accurate model makes development time-consuming and costly and above all this controller would not be very robust. Since robustness is a major goal in pharmaceutical production using only a feed-forward controller is not sufficient.

8.6 Feedback Controller

A direct feedback controller for the feed rates could not be used since the process is strongly non-linear. In other words, the controller has to exponentially increase the feed rates based on a (hopefully) small error between the estimated specific growth rate and the set-point for the specific growth rate. Tartakovsky et al. [21] reported an input-output-linearising controller to solve this problem. Moreover, in the dual-substrate case, such a direct feedback controller has to generate two varying feed rates based on a single input signal.

When the feed rates of glutamate and lactate are calculated by the feed-forward controller, the feed-forward controller, process, and estimator can be seen as one combined system. That system is input-output linear and scalar with the reference specific growth rate as input and the estimated specific growth rate as output. Now a direct feedback controller would suffice to eliminate the error which is introduced when the calculations in the feed-forward controller do not satisfy due to mismatch between the used model and the actual process characteristics or due to unknown disturbances.

The feedback controller adjusts the reference specific growth rate, μ^{ref} , which is the input of the feed-forward controller in order to let the estimated specific growth rate approach the set-point, μ^{set} . In other words, by putting the feedback controller in sequence with the feed-forward controller instead of in parallel the gain of the feedback controller does not have to be adapted during cultivation.

Feedback Control

The feedback controller should only adjust the reference value slightly and not react to noise on the estimation signal. A slow/smooth controller is therefore desired. For this case the velocity form of a PI-controller [18] is implemented to adjust μ^{ref} such that the estimated specific growth rate will converge to the desired μ^{set} :

$$\begin{aligned} \varepsilon_k &= \mu_{set} - \hat{\mu}_k \\ \mu_{ref,k} &= \mu_{ref,k-1} + K_P \cdot (\varepsilon_k - \varepsilon_{k-1}) + K_I \cdot \varepsilon_k \cdot (t_k - t_{k-1}) \end{aligned} \quad (16)$$

Where K_P and K_I are the two parameters that define the controllers (dynamic) response to errors between the estimated and set-point specific growth rate. Tuning these parameters was conducted by performing several simulations of open-loop step-responses at various cell densities. Subsequently closed-loop simulations were performed with varying levels of noise and drift in order to let the error converge properly to zero with the right slow dynamics. Final tuning is done based on experimental data.

Figure 3 shows the complete implementation of the sequential estimator, the feed-forward controller, and the feedback controller.

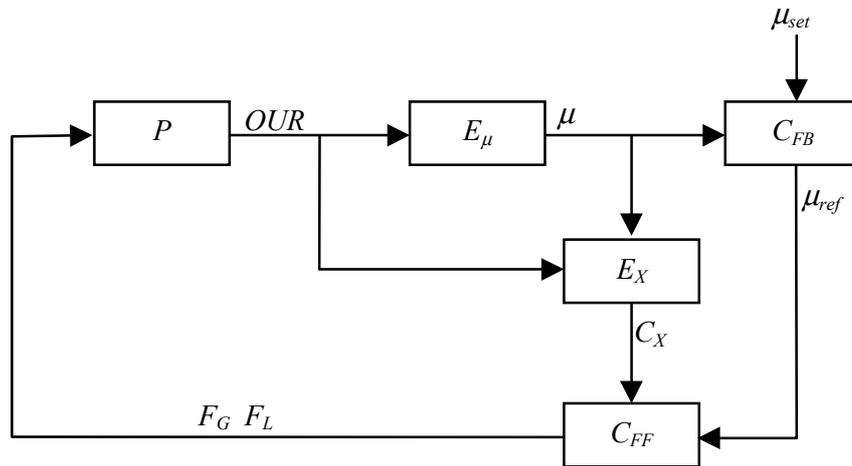


Figure 3: Diagram of the complete algorithm with process P , growth rate observer E_μ , biomass observer E_X , feed-forward controller C_{FF} , feedback controller C_{FB} , and supervisory boundary controller C_{SB} .

8.7 Results

Feed-Forward Control Only

Several experiments with *B. pertussis* were conducted with the sequential estimator and feed-forward controller only. For these experiments no feedback control was applied. Figure 4 shows an experiment where at 36 hours the batch phase of the cultivation was ended (all substrates were depleted) and the feed-forward controller was started. Although the calculated feed rates were correct at the start of the fed-batch, the pumps were erroneously set twice too

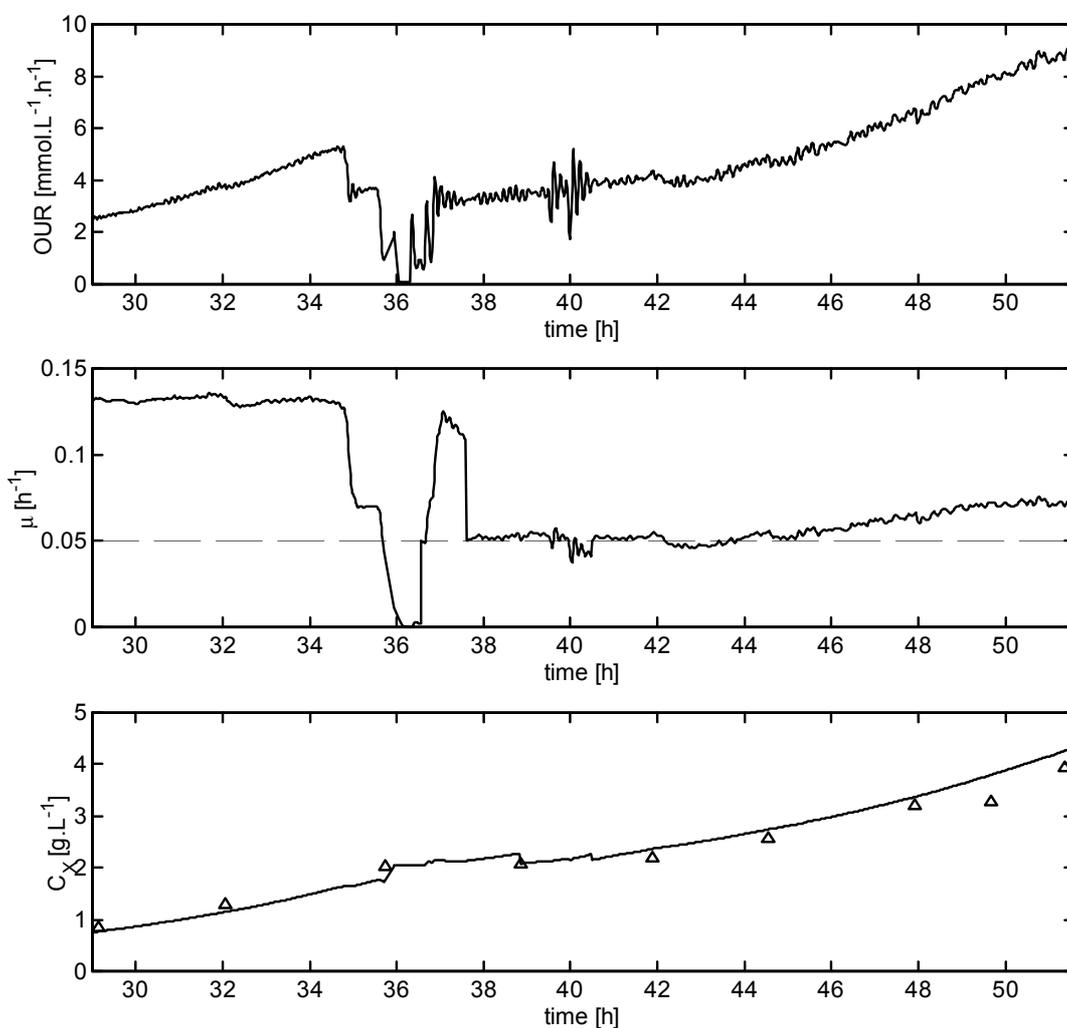


Figure 4. Fed-batch cultivation of *B. pertussis* with the feed-forward controller only, to maintain the specific growth rate at 0.05 h⁻¹. Top graph (A): The *OUR* during cultivation. Middle graph (B): The set-point (dashed) and estimated specific growth rate (solid). Bottom graph (C): The estimated (solid) and measured (triangles) biomass concentrations.

high. This resulted in a too high specific growth rate, which was evidently noticed by the sequential estimator. After 37.5 hours the pumps were correctly set and the specific growth rate was maintained at set-point ($\mu^{set} = 0.05 \text{ h}^{-1}$) for approximately 7 hours. The off-line biomass samples correspond well with the estimated biomass concentrations. Around 40 hours inaccuracy in the *OUR* signal resulted in increased noise on the specific growth rate estimations. After 45 hours the biomass concentration was slightly over-estimated resulting in a slow drift of the specific growth rate from set-point. At the end the specific growth rate had drifted to 0.07 h^{-1} . Clearly the feed-forward controller strongly relies on accurate model parameters and a good estimation of the biomass concentration. Errors in these parameters or estimation result in undesirable behaviour, emphasising that feed-forward control only is not satisfactory.

Feed-Forward-Feedback Controller

Adding the feedback controller to the feed-forward controller according to the diagram in figure 3 should prevent a slow drift from set-point as shown by the

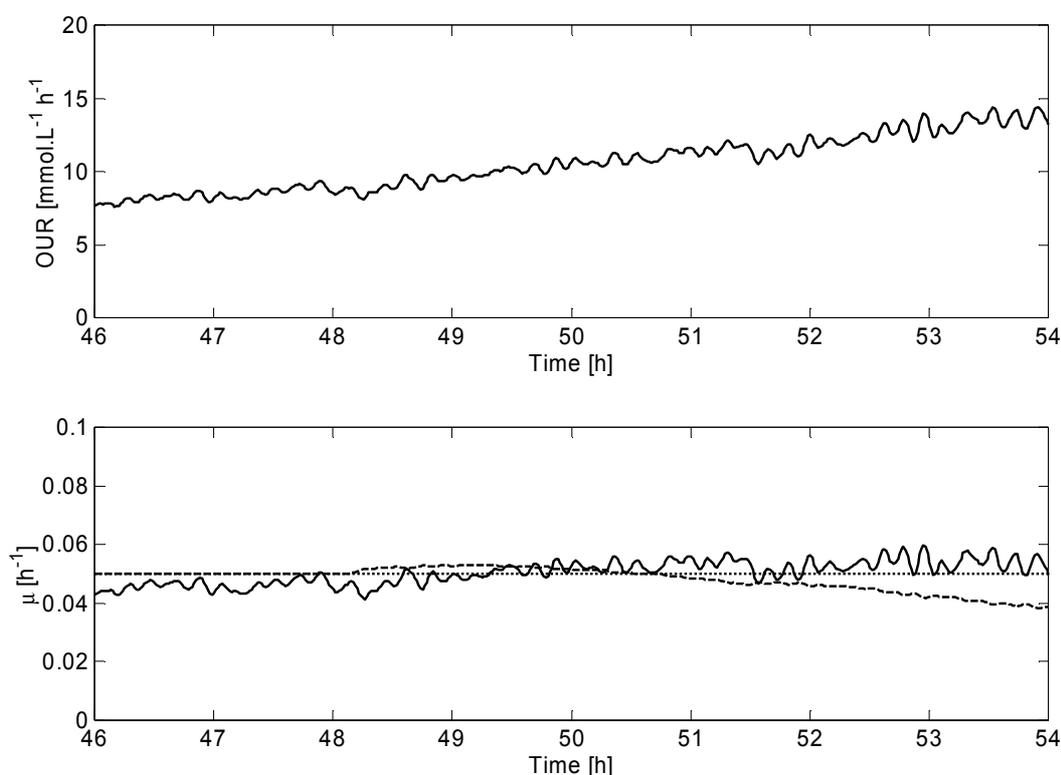


Figure 5. Fed-batch cultivation of *B. pertussis* with the feed-forward-feedback controller with set-point 0.05 h^{-1} . Top graph (A): The *OUR* during cultivation. Bottom graph (B): The estimated specific growth rate (solid) and the controller output μ_{ref} (dashed).

experiment with feed-forward only. Figure 5 shows the results for an experiment with the combined feed-forward-feedback controller, figure 5a shows the measured *OUR* and figure 5b shows the estimated specific growth rate together with the set-point and output signal of the controller. Until 48 hours feed-forward only was applied. As can be seen in figure 5b, the specific growth rate was slightly off set-point. After 48 hours the feedback controller was started and the specific growth rate was brought back to set-point by the feedback controller. Figure 5b also shows that the controller output, μ^{ref} , slowly decreased to avoid a further increasing specific growth rate. Without the feedback controller the feed would have been too high and the specific growth rate would have increased too much. Looking at the dynamics of the noise on the estimated specific growth rate, this noise was clearly not induced by the controller but by the noise on the *OUR*-signal. Fine-tuning the dissolved oxygen controller or pre-filtering the *OUR*-signal might reduce this noise. Tuning the estimator might reduce the noise as well, but would decrease the error-convergence. Therefore this behaviour is considered desirable as long as the controller is sluggish enough not to respond to this noise.

Figure 6 shows the results for another experiment where the specific growth rate was maintained at set-point for over 15 hours and the total amount of biomass was nearly quadrupled (increased final concentration and volume). At approximately 24 hours the feed-forward-feedback controller was started with set-point 0.05 h^{-1} . Until 32 hours the specific growth rate was maintained accurately around this set-point. At 32 hours the pumps reached their maximal flow rates and accordingly their ranges were adjusted, as a consequence the controllers had to be reset. However, the estimated specific growth rate was maintained satisfactorily around set-point emphasising the robustness of the algorithm.

Then, at 38 hours, the set-point was increased to 0.07 h^{-1} and maintained for another 5 hours. This higher specific growth rate resulted in a heavily increased noise level on the *OUR*, and consequently on the estimated specific growth rate. Despite the increased noise the controller kept the specific growth rate at this new higher set-point. During the complete course of this experiment the biomass concentration was estimated accurately as can be seen by figure 6b.

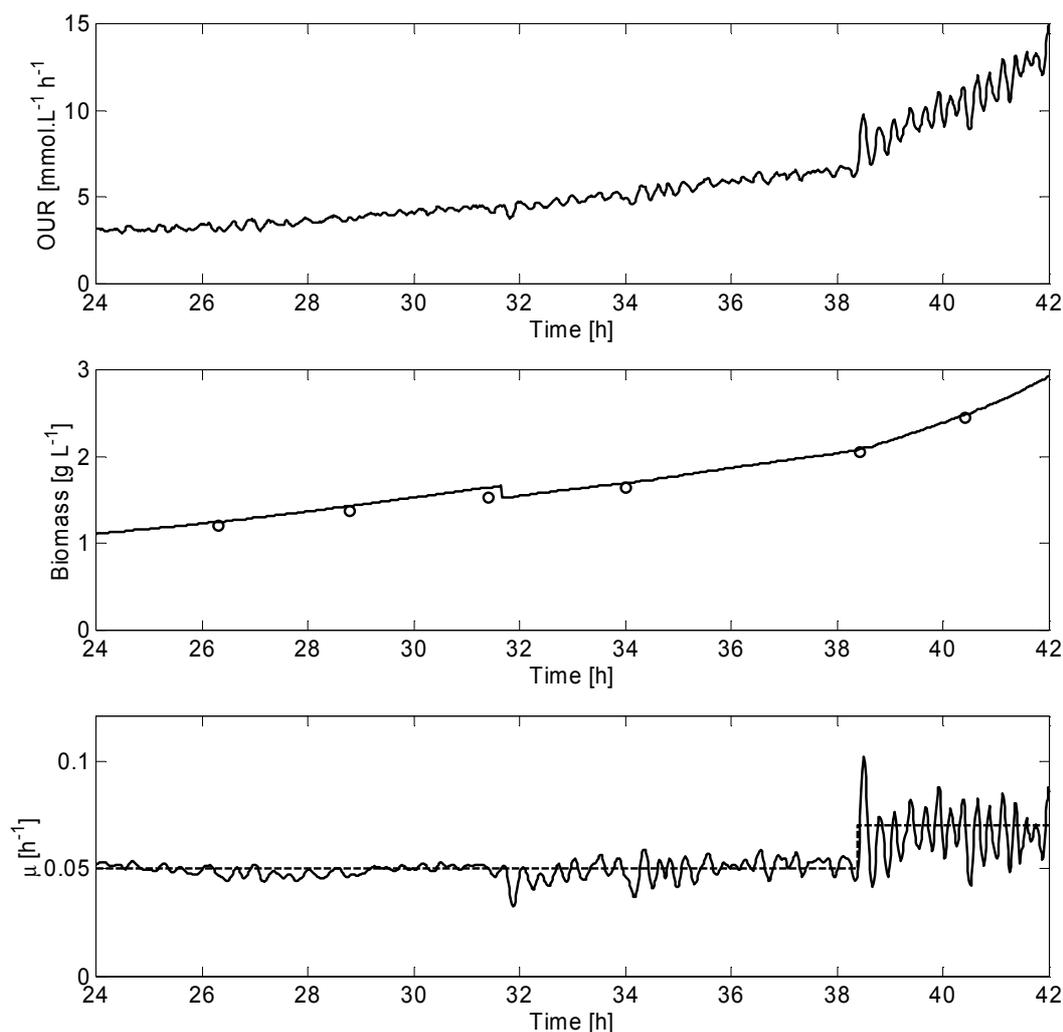


Figure 6. Feed-forward-feedback controlled cultivation of *B. pertussis* where at 38 hours the set-point was increased from 0.05 h⁻¹ to 0.07 h⁻¹. Top graph (A): The *OUR* during cultivation. Middle graph (B): The estimated (solid) and measured (circles) biomass concentrations. Bottom graph (C): The estimated specific growth rate (solid) and the set-point (dashed).

8.8 Conclusions

In this chapter a complete scheme for the control of the specific growth rate is presented (figure 3) and applied to the cultivation process of *B. pertussis*. The main features of this μ -controller are:

- The applied sequential estimator does not need complex or expensive measuring devices. A simple gas analyses is sufficient to provide the required on-line data. No complex models are involved; just little a priori

knowledge and a few experiments are necessary to determine the parameters.

- By separating the controller in a feedback and feed-forward part the system becomes intuitive and easily adjustable. The feed-forward controller makes the system input-output linear and scalar such that a simple direct feedback controller, e.g. PID-controller, can be used.
- In order to apply the feed-forward controller to other processes it is necessary to conduct a relatively small number of experiments providing enough data to determine and tune the parameters mentioned in Table 1. For single-substrate limited systems such an operation is rather straightforward. For multi-substrate limited systems this exercise could become more cumbersome as described in Chapter 5 [12].
- The controller stability was first tested in simulated cultivations and after that, successful tests have been performed in real experiments at 3-L scale with *B. pertussis*. Growth was controlled at reduced specific growth rate (0.05 h^{-1}) over a period of more than 15 hours. All tests showed a sufficient robust behaviour. The controller did not need time consuming tuning procedures.

This study focused on keeping the specific growth rate constant at 0.05 h^{-1} , which was a suitable value to demonstrate the controller performance. The used level of 0.05 h^{-1} was enough to almost quadruple the total amount of biomass. However, there is a possibility that another level of the specific growth rate yields higher production capacity or product quality. The presented μ -controller gives the perspective for a systematic variation of the specific growth rate to find the optimal specific growth rate.

8.9 Nomenclature

C_G	$\text{mmol}\cdot\text{l}^{-1}$	glutamate concentration
C_G^{in}	$\text{mmol}\cdot\text{l}^{-1}$	glutamate concentration in glutamate feed
C_L	$\text{mmol}\cdot\text{l}^{-1}$	lactate concentration
C_L^{in}	$\text{mmol}\cdot\text{l}^{-1}$	lactate concentration in lactate feed
C_X	$\text{g}\cdot\text{l}^{-1}$	biomass concentration
D	h^{-1}	dilution rate

DUAL-SUBSTRATE FEEDBACK CONTROL OF THE SPECIFIC GROWTH RATE

F_G^{in}	$l \cdot h^{-1}$	incoming glutamate feed rate
F_L^{in}	$l \cdot h^{-1}$	incoming lactate feed rate
F_{sample}	$l \cdot h^{-1}$	feed rate of sampletaking
K_G	$mmol \cdot l^{-1}$	Monod saturation constant for glutamate
K_L	$mmol \cdot l^{-1}$	Monod saturation constant for lactate
m_O	$mmol \cdot g^{-1} \cdot h^{-1}$	maintenance coefficient for oxygen
OUR	$mmol \cdot l^{-1} \cdot h^{-1}$	oxygen uptake rate
OUR_m	$mmol \cdot l^{-1} \cdot h^{-1}$	measured oxygen uptake rate
$p\#$	-	tuning parameters
R_X	$g \cdot l^{-1} \cdot h^{-1}$	biomass growth rate
t	h	time
V	l	volume
Y_{G1}	$g \cdot mmol^{-1}$	yield of biomass on glutamate over pathway 1
Y_{G2}	$g \cdot mmol^{-1}$	yield of biomass on glutamate over pathway 2
Y_{L2}	$g \cdot mmol^{-1}$	yield of biomass on lactate over pathway 2
Y_O	$g \cdot mmol^{-1}$	yield of biomass on oxygen
α	h^{-1}	apparent specific growth rate
μ	h^{-1}	overall specific growth rate
μ_1	h^{-1}	specific growth rate over pathway 1
μ_2	h^{-1}	specific growth rate over pathway 2
μ_{max}	h^{-1}	maximal specific growth rate for pathway 1
μ_{enh}	h^{-1}	maximal enhancing specific growth rate for pathway 2

8.10 References

1. Bastin, G. and D. Dochain (1990). On-line Estimation and Adaptive Control of Bioreactors. Amsterdam, Elsevier Science Publishers.
2. Lubenova, V., M. Ignatova and S. Tsonkov (1993). On-Line Estimation of Specific Growth Rate for a Class of Aerobic Batch Processes. Chemical and Biochemical Engineering Quarterly 7(2): 101-106.
3. Lee, J., S.Y. Lee, S. Park and A.P.J. Middelberg (1999). Control of fed-batch fermentations. Biotechnology Advances 17: 29-48.

4. Paalme, T., K. Tiisma, A. Kahru, K. Vanatulu and R. Vilu (1990). Glucose-limited fed-batch cultivation of *Escherichia coli* with computer-controlled fixed growth rate. *Biotechnology and Bioengineering* 35: 312-319.
5. Zabriskie, D.W., D.A. Wareheim and M.J. Polansky (1987). Effects of Fermentation Feeding Strategies prior to Induction of Expression of a Recombinant Malaria Antigen in *Escherichia coli*. *Journal of Industrial Microbiology* 2: 87-95.
6. Gudi, R.D., S.L. Shah, M.R. Gray and P.K. Yegneswaran (1997). Adaptive multirate estimation and control of nutrient levels in a fed-batch fermentation using off-line and on-line measurements. *The Canadian Journal of Chemical Engineering* 75: 562-573.
7. Hitzmann, B., O. Broxtermann, Y.-L. Cha, O. Sobieh, E. Stärk and T. Scheper (2000). The control of glucose concentration during yeast fed-batch cultivation using a fast measurement complemented by an extended Kalman Filter. *Bioprocess Engineering* 23: 337-341.
8. Levisauskas, D. (2001). Inferential control of the specific growth rate in fed-batch cultivation processes. *Biotechnology Letters* 23: 1189-1195.
9. Levisauskas, D., R. Simutis, D. Borvitz and A. Lübbert (1996). Automatic control of the specific growth rate in fed-batch cultivation processes based on an exhaust gas analysis. *Bioprocess Engineering* 15: 145-150.
10. Massimo, C.D., M.J. Willis, G.A. Montague, M.T. Tham and A.J. Morris (1991). Bioprocess model building using artificial neural networks. *Bioprocess Engineering* 7: 77-82.
11. Riesenber, D., V. Schulz, W.A. Knorre, H.D. Pohl, D. Korz, E.A. Sanders and A. Roß (1991). High cell density cultivation of *Escherichia coli* at controlled specific growth rate. *Journal of Biotechnology* 20: 17-28.
12. Neeleman, R., M. Joerink and A.J.B. Van Boxtel (2001). Dual-Substrate Utilisation By *Bordetella pertussis*. *Applied Microbiology and Biotechnology* 57: 489-493.
13. Ignatova, M. and V. Lubenova (1995). An approach for estimation of biotechnological processes variables. *Chemical and Biochemical Engineering Quarterly* 9(1): 19-25.

14. Chattaway, T. and G. Stephanopoulos (1989). An adaptive state estimator for detecting contaminants in bioreactors. *Biotechnology and Bioengineering* 34: 647.
15. Park, S. and W.F. Ramirez (1989). Optimal regulatory control of bioreactor nutrient concentration incorporating system identification. *Chemical Engineering Science*.
16. Stephanopoulos, G. and S. Park (1990). Bioreactor State Estimation. *Biotechnology, Measuring Modelling and Control*. H.-J. Rehm and G. Reed. Weinheim Germany, VCH. 4: 225-250.
17. Neeleman, R., M. Joerink, A.J.B. Van Boxtel, C. Beuvery and G. Van Straten (2001). On-line harvest prediction for the production of biologicals, *Computer Applications in Biotechnology*, Quebec, Canada, 430-435.
18. Stephanopoulos, G. (1984). *Chemical Process Control: An Introduction to Theory and Practice*. London, UK, Prentice-Hall International Inc.
19. Neeleman, R. and A.J.B. Van Boxtel (2001). Estimation of specific growth rate from cell density measurements. *Bioprocess Engineering* 24(3): 179-185.
20. Korz, D.J., U. Rinas, K. Hellmuth, E.A. Sanders and W.D. Deckwer (1995). Simple fed-batch technique for high cell density cultivation of *Escherichia coli*. *Journal of Biotechnology* 39: 59-65.
21. Tartakovsky, B., B.E. Dainson, D.R. Lewin and M. Sheintuch (1996). Observer-based non-linear control of a fed-batch autoinductive fermentation process. *The Chemical Engineering Journal* 61: 139-148.

9 ADVANCED PROCESS MONITORING: THE ULTIMATE ALTERNATIVE

This chapter is based on:

R. Neeleman and E.C. Beuvery (2002). Advanced Process Monitoring: The Ultimate Alternative. In: Developments in Biologicals, Volume 111, p. 275-284. IABs, Basel, Karger, Switzerland

9.1 Abstract

Parametric release has the potency to be the ultimate alternative for the use of animal tests. It involves the release of product using evaluation of tightly controlled physical parameters, without the performance of any test on the final product. This practice is strengthened if there is a lack of validity for the final biological or animal test to indicate quality or consistency. That pleads for in-process monitoring replacing finished product controls when amongst others batch-to-batch consistency has been formerly validated. Currently parametric release is a common practice when it comes to terminally sterilised and aseptically filled biologicals. It is illustrated for the production of inactivated cells of *Bordetella pertussis* that new techniques are able to monitor the process in detail. Furthermore, it is shown that the production of *B. pertussis* is a well defined and highly predictable. In parallel with the tests for terminally sterilised biologicals, mouse potency tests are statistically hardly capable of indicating batch-to-batch variation. Monitoring on-line process parameters provides a better assurance of consistency and quality, thus giving the opportunity for batch release and regulatory approval without the use of animal tests.

9.2 Introduction

The object of the regulations in the European Community in the field of biologicals is to ensure a high level of safety and efficacy; unfortunately, they tend towards the increased use of animals for testing purposes. For ethical, economic and practical reasons, the number of animals used must be reduced which makes the quest for an alternative highly relevant.

An important issue for the reduction of animal tests is parametric release. Currently parametric release is a common practice when it comes to terminally sterilised biologics [1] and is the object of efforts for aseptically filled biologics [2]. It involves the release of product using evaluation of tightly controlled physical parameters, without the performance of a sterility test. The lack of validity for the sterility test to indicate product sterility strengthens this practice.

Soulebot [3] took parametric release one step further when he pleaded for the replacement of animal tests of the final product with in vitro tests before

adsorption and blending. Thus pleading that in-process controls replace finished product controls when amongst others batch-to-batch consistency has been formerly validated.

For the current combination vaccine (diphtheria, tetanus, pertussis, and inactivated polio vaccine), as applied in The Netherlands, animal tests are performed both on the individual bulk products and on the final product. One of these bulk products is a vaccine consisting of inactivated whole cells of *B. pertussis*. The antigenic composition i.e. the quality of this vaccine is determined during the cultivation step. By rigorously defining the production process and using a chemically defined medium, the operating margins of the cultivation are narrowed down. Consequently, the consistency of the production has increased. Furthermore, the metabolism of *B. pertussis* is known extensively [4], thus yielding the possibility of using process data to monitor the behaviour of the organism during cultivation.

In order to apply a parametric release system to the production of inactivated whole cells of *B. pertussis* the performance of biomass should be monitored on-line. However, sensors for biomass are not existent, expensive and/or not sterilisable. Alternatively, sampling for off-line measurements has to be limited for operability, and sterility reasons. A model describing biomass, substrates, specific growth rate and oxygen consumption is derived and used to build a 'software-sensor' for specific growth rate and biomass concentration. This software-sensor provides a good on-line estimation of the specific growth rate and biomass concentration based on standard oxygen measurements only. Finally, the same model is used to show that the production process of *B. pertussis* is highly predictable [5].

Parallel to terminally sterilised biologics, the process is now well defined and highly predictable. Furthermore, alike the sterility tests, mouse potency tests are statistically hardly capable of indicating neither batch-to-batch variation nor product quality. Monitoring on-line process characteristics provides a better assurance of consistency and quality. Supported by in vitro tests such process-monitoring yields the opportunity for regulatory approval and batch release, thus eliminating mouse potency tests [6].

9.3 Parametric Release

Definition

Within this framework, the term parametric release will be used for the release of pharmaceuticals without conducting a test on the final or bulk product. The quality of the product is ascertained solely by review of the parameters achieved during the process. Such a procedure is only relevant when the process is validated to be consistent and the product test is unreliable or statistically not precise enough. In other words, no test on the final product has to be done if and only if the process is fully defined, validated, consistent and its deviations are more accurately monitored than possible with product testing. Such an approach would be consistent with the general notices of the European Pharmacopoeia [7]:

“With the agreement of the competent authority, alternative methods of analysis may be used for control purposes, provided that the methods used enable an unequivocal decision to be made as to whether compliance with the standards of the monographs would be achieved if the official methods were used.”

Several cases of parametric release have been approved or proposed in the last few years [1-3] and will be reviewed shortly.

Terminally sterilised biologics

Parametric release is a common practise for the release of terminally sterilised pharmaceuticals without conducting a sterility test. The sterility of a lot is examined solely by monitoring the sterilisation cycle [1]. Typically this requires sterilisation of a product in the final container by means of a validated sterilisation procedure achieving a SAL of 10^{-6} or better. Products sterilised in their final container by validated, well-controlled moist heat sterilisation processes may be released for distribution based on a review of the recorded sterilisation process parameters, provided the recorded parameters accord with the validated process, without reliance on a sterility test [8]. This is acknowledged in pharmacopoeias by statements to the effect that release of sterile products may not be solely based on sterility testing, but has to rely on validated procedures. According to the general texts on sterility of the European Pharmacopoeia [9]:

“When a fully validated terminal sterilisation method by steam, dry heat or ionising radiation is used, parametric release, that is the release of a batch of sterilised items based on process data rather than on the basis of submitting a sample of the items to sterility testing, may be carried out, subject to the approval of the competent authority.”

Besides being common practice for terminally sterilised biologicals, parametric release is the object of efforts for aseptically filled biologicals according to the same concept.

In-process controls versus finished product controls

In vitro testing of the final product is complicated since the active components are often adsorbed onto alum or presented as an emulsion, and tests are easier to carry out prior to final blending operations. Soulebot [3] suggested that in-process controls could replace finished product controls when amongst others batch-to-batch consistency has been formerly validated. Since the adsorption and blending process is a fully defined and validated process and the test on the final product is only possible with an inaccurate in vivo test. A parallel could be made with the terminally sterilised biologicals. The process is a fully defined physical process with a good test on the incoming material. Thus, the outcome of the process (the quality of the product) can be estimated with more assurance by monitoring the process and establishing proper input conditions than by testing the final product. In other words, Soulebot pleaded for parametric release.

***B. pertussis* cultivation**

For the current combination vaccine (diphtheria, tetanus, pertussis, and inactivated polio), as applied in The Netherlands, the bulk pertussis vaccines consist of inactivated whole cells of two strains of *B. pertussis*. The required assays for the release of whole cell pertussis vaccine are the mouse weight gain test and the mouse protection test (‘Kendrick test’) and are performed both on the individual bulk products and on the final blended product. These tests involve at least 156 mice per strain and are criticised for the use of many animals, poor precision and insensitivity [10]. Therefore, such tests give poor assurance of process consistency and whether the product is within specifications. Since the quality of the final product is determined during the

cultivation process, it is essential to verify whether parametric release could be applied to this process.

For parametric release, specifications of the input conditions should be kept and the process should be monitored such that the consistency and product quality is estimated with more accuracy and assurance than the mouse tests. Input conditions involve the medium composition, which is fully defined and can easily be prepared consistently and the biomass concentration in the inoculum, which is measured. The cultivation process is controlled severally, not only by maintaining constant environmental conditions like pH, temperature, and dissolved oxygen concentrations, but also by using computer recipes preventing anomalies during the process. As a result of using the defined medium and applying such rigorous control systems the process has become very consistent.

Monitoring the process to verify whether crucial process parameters remain within specifications is crucial for parametric release. However, monitoring the process should not be done by looking at environmental conditions only, such as whether the pH is maintained at set-point, or by looking at derived conditions like the amount of oxygen consumed. Preferably, the state of the process, the biomass concentration, should be monitored. And ultimately, it would give even more assurance if not the state of the process, but the performance of the biomass is monitored by for instance measuring the growth rate or metabolic fluxes on-line. However, sensors for biomass concentration are at present not mature enough for pharmaceutical application and sensors for measuring the performance of biomass in a bioreactor are not existent.

Therefore, advanced process monitoring where the state of biomass can be estimated using available measurements and model based inference seems promising. These algorithms have been developed in the 1970s and have been adopted by the biotechnology industry in the 1990s. Many applications have proven its power and robustness in a production environment and thus proven its maturity for pharmaceutical application.

9.4 Advanced Process Monitoring

Modelling

Since the medium used for cultivating *B. pertussis* is a chemically defined medium and the process is rigorously defined by software recipes, the behaviour of the organism is fully dependent of its known environment. Consequently, a model was made which relates the states of the organism, the growth rate, pathway distribution, and respiration to its environmental conditions. Growth of the organism can be described using two main pathways for utilising the two main substrates, glutamate, and lactate. The used ‘dual-substrate model’ is described by Neeleman et al. [11]. Figure 1, shows the prediction of the dual-substrate model compared with the measurements taken during cultivation. The figure shows that this dual-substrate model gives a good prediction of the actual behaviour of *B. pertussis*. Furthermore, it gives crucial inside information on the cellular state, growth rate and pathway distribution.

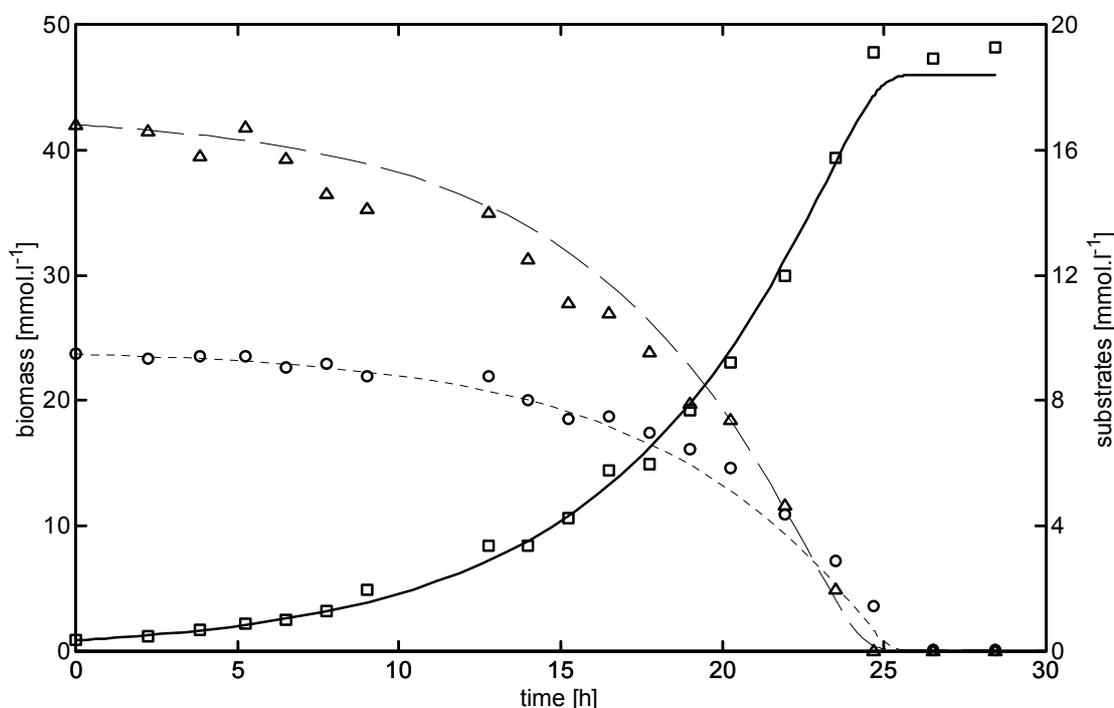


Figure 1: Measured biomass (squares), glutamate (circles) and lactate (triangles) concentrations during batch cultivation together with the simulation results (solid).

Monitoring

Whilst direct online measurement of the states may not be possible. The influence of their variation can be observed in available online measurements. It is therefore in certain instances possible to obtain an online inference of the process states – such an approach is termed a ‘software-sensor’. In other words, a software-sensor is a software algorithm giving an online prediction for parameters during a cultivation, in this case biomass and growth rate, whose analyses are time consuming, labour intensive and costly. A software-sensor calculates this prediction from the available online measurements using a model and proper mathematical inference.

Based on the dual-substrate model such a software-sensor is derived which estimates the current state of the process and biomass based on the oxygen uptake rate (OUR). The OUR is calculated from standard available measurements of the dissolved and gaseous oxygen concentrations. The OUR is used to estimate the specific growth rate and the biomass concentration. Figure 2 shows a schematic representation of the software-sensor algorithm.

This software-sensor currently operates on-line during cultivation. Figure 3 shows the results for a standard cultivation of *B. pertussis*. Figure 3A shows the ‘measured’ OUR that is used to estimate the specific growth rate and biomass concentration. Figure 3B shows the estimated specific growth rate. The estimated specific growth rate is nearly constant during the first 18 hours, then decreases slowly until 23 hours when the growth rate is zero. Figure 3C shows the measured and estimated biomass concentration. At the beginning of the batch, the biomass concentration is estimated too low, but after 15 hours, the estimation follows the measurements more accurately. Clearly, this software-sensor is capable of on-line biomass and growth rate estimation.

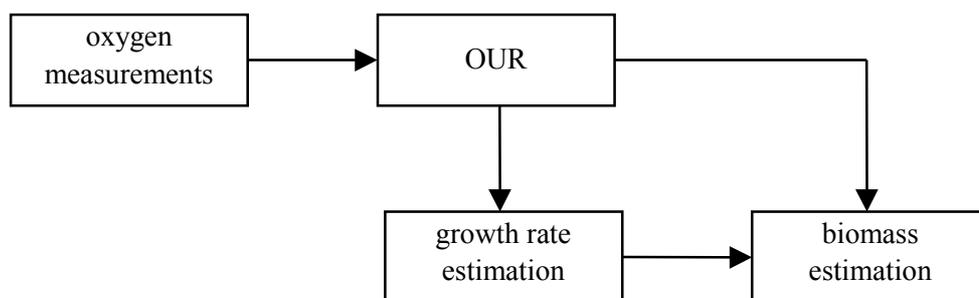


Figure 2: Schematic representation of the software-sensor.

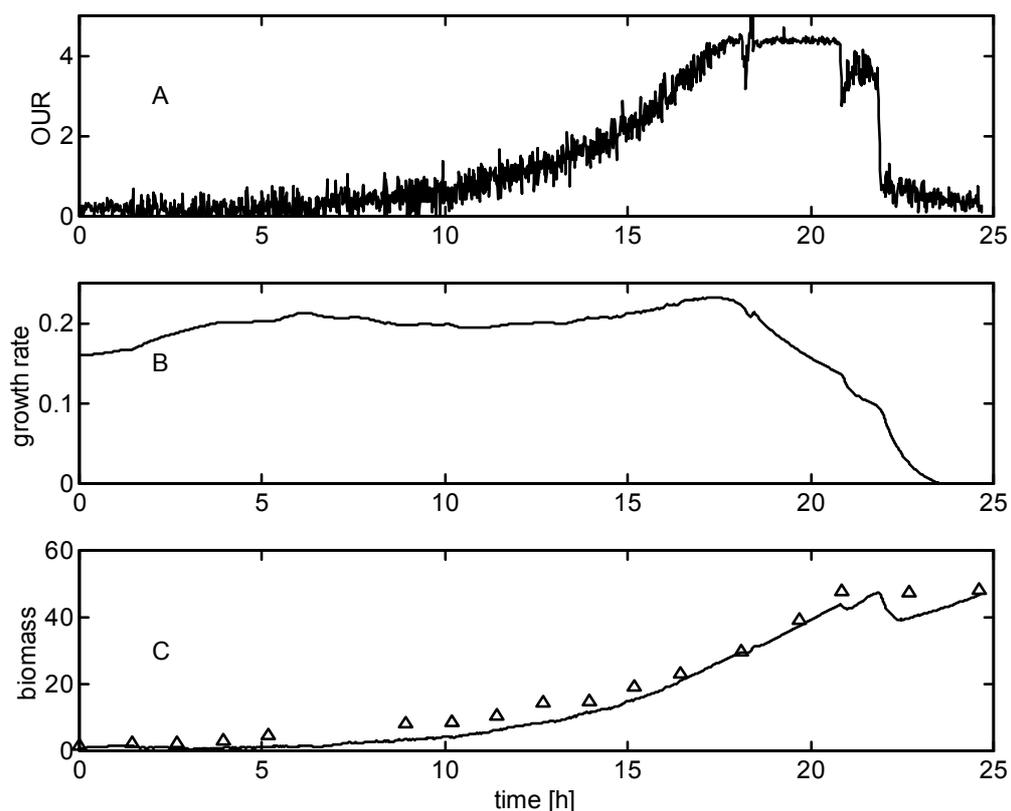


Figure 3: A: measured OUR. B: estimated growth rate. C: estimated (solid) and measured (triangles) biomass concentration.

Robustness

Important for application in the pharmaceutical industry is to test the proposed algorithm for its robustness. Here, robustness of this algorithm is defined as the ability to converge to the ‘true’ state after a malicious start and the ability to cope with anomalies during the process.

Normally the software-sensor is initialised by taking a sample of the cultivation broth. This measurement is then used as starting condition for the algorithm. Convergence of the software-sensor is tested by initialising without that sample. The starting specific growth rate is set to maximal and the starting biomass concentration is set to zero. Figure 4A shows the measured OUR and figure 4C shows the estimated and measured biomass concentration. The estimated biomass concentration, which was set to zero at initialisation, converges rapidly to its ‘true’ measured value indicating proper convergence of the software-sensor.

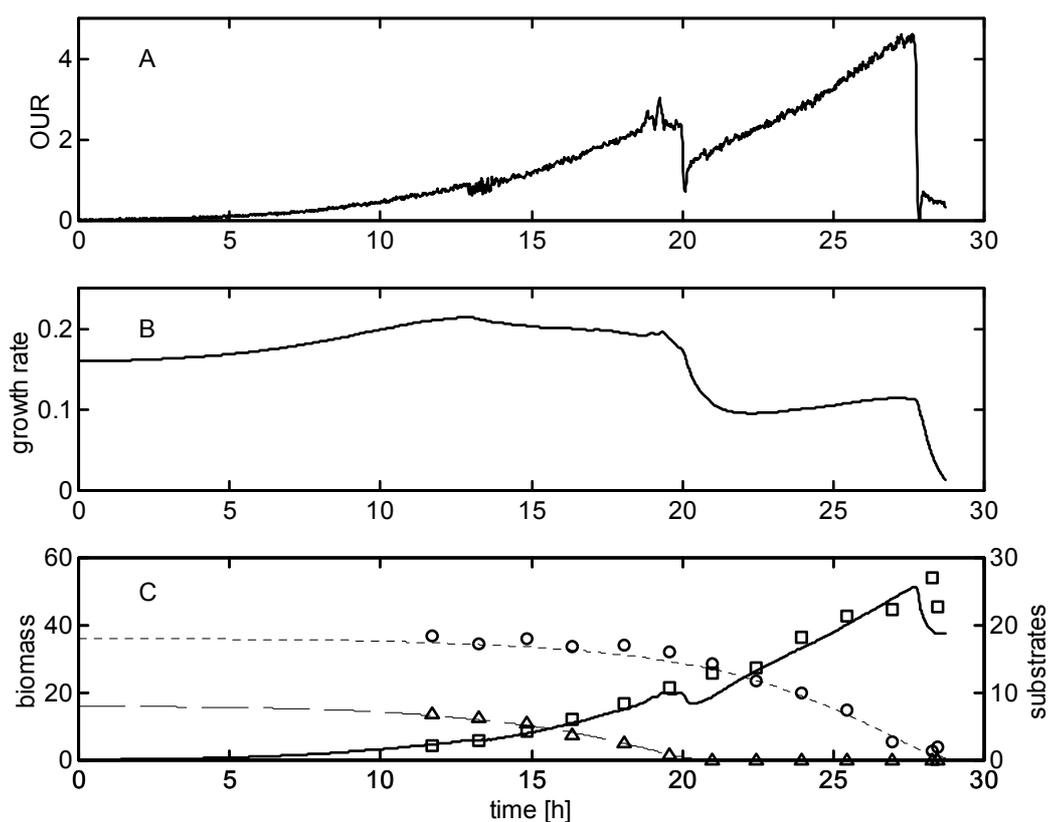


Figure 4: A: measured *OUR*. B: estimated growth rate. C: estimated concentrations (lines) and measurements of biomass (squares), lactate (triangles) and glutamate (circles).

Furthermore the tracking capabilities of the software-sensor are tested by increasing the amount of glutamate added to the medium such that halfway through cultivation, lactate will be depleted and the organism has to shift from growing on both substrates towards growth on glutamate only. Figure 4A shows that halfway during the cultivation, at 20 hours when lactate is depleted, the *OUR* suddenly decreases indicating the decrease in oxygen consumption by the organism. This lactate depletion is also shown by the estimated growth rate in Figure 4B; it drops from the maximum overall specific growth rate to the maximum specific growth rate possible for growth on glutamate only (0.1 h^{-1}). The estimated and measured biomass and substrate concentrations are shown in Figure 4C. Clearly, the software-sensor illustrated to be able to track changes in growth rate due to substrate depletion. Now the software-sensor proved to be robust enough for regulatory compliance and thus pharmaceutical application.

Predictability

The results of the software-sensor can subsequently be used to predict the future development of the concentrations and thus the harvest time by simulation with the dual-substrate model. This simulation runs when requested by the operator or on a scheduled interval and has shown to give a prediction of the harvest time with an accuracy of 15 minutes when run more than 12 hours in advance [5]. Concatenating, the process is fully defined and predictable and can be monitored in detail with the 'software-sensor' resulting in a more confident quality measure of the product than the animal tests.

9.5 Discussion

Parametric release is possible if and only if the following conditions are met: critical input parameters and critical process parameters are kept within their previously set limits and the monitoring system is more accurate in establishing deviations or proving consistency than the final test. The use of a software-sensor enables in-process monitoring of the performance of biomass in the bioreactor, thus verifies whether the specific growth rate, a critical process parameter, remains within specifications.

Critical process parameters

Besides using the specific growth rate as a critical process parameter, a list of other critical process parameters and their respective operating limits should be defined and established. Since these process parameters depend on the strain, medium, and specific equipment used it will be beyond the scope of this paper to elaborate on this list in detail. Such process parameters might include, for example: dissolved oxygen concentration, temperature, pH, and more importantly biomass concentration. When critical process parameters defined in the standard operating procedures have not been achieved, the product must not be released to the market. Critical process parameters must be recorded.

Prerequisites for parametric release

A program must be established describing in detail the release steps and the documentation required. The decision to release a product is primarily based on the acceptance criteria defined in the validation records. The following

points are prerequisites for parametric release and thus form the basis for the assessment of quality and subsequent release of product.

- Procedures are established that ensure critical process input parameters are met before every run.
- Critical process parameters are defined together with their respective operating limits prior to release of the product and procedures are established that ensure these operating limits are met during every part of the process.
- Procedures are established that ensure all relevant data are documented and reviewed during the process of releasing the product.
- A document (batch record) is retained that contains all relevant data including the recorded critical process parameters reviewed for the release decision of the product. The document must be available for review by the controlling regulatory (competent) authority.
- Procedures for maintenance, periodic re-qualification, with acceptance criteria and schedules are established and followed.
- Evidence of successful validation of the used equipment and the monitoring program must exist prior to processing and release of the product. Additionally, the process should be re-examined periodically to ensure it continues to function as intended.
- A deviation and change control procedure must be in place describing the reporting and actions to be taken in the event of any deviation from the acceptance criteria of the process and of the monitoring program. This includes an assessment of the significance of the deviation, and a decision if revalidation is necessary or not.
- Once parametric release has been decided upon, a release or non-release decision cannot be overruled by use of another method.

Parametric release: the ultimate alternative

This work showed that a cultivation process as part of producing a pharmaceutical could be considered a well-defined and accurately predictable process. Furthermore, by using advanced techniques the biomass

concentration and more importantly the performance of the biomass could be monitored giving crucial information about the quality and consistency of the final bulk product.

The used software-sensor is not organism-specific and thus applicable for various other organisms. Moreover, the complete procedure and prerequisites proposed in this work can be applied to a wide variety of organisms and processes.

This work is an attempt to set up a framework for releasing product based on advanced monitoring techniques. This technique has a strong potential for eliminating the number of animals used for testing and thus becomes a novel and promising alternative. Such an approach would not only save animal lives, an ethical requirement, but also save costs (animal tests are particularly costly and often of long duration) and solve some practical difficulties (certain tests can adequately be performed in SPF animals or animals free of specific antibodies or that have never previously encountered the specific organism). In many cases, such animals are plainly not available.

Thus the pharmaceutical industry and regulatory agencies should jointly facilitate progress in this field to lead to parametric release systems based on process characteristics which can accurately be measured using softsensing techniques.

9.6 References

1. Moldenhauer, J.E. and R. Madsen (2000). Parametric Release - Much ado about nothing. *Journal of Pharmaceutical Science and Technology* 54(1): 32.
2. Olson, W.P. (1998). Opinion: Aseptic Processing and Parametric Release. *Journal of Pharmaceutical Science and Technology* 52(3): 81.
3. Soulebot, J.P., F. Milward and P. Prevost (1993). In vitro potency testing of inactivated biologics: current situation in the E.E.C. *Veterinary Microbiology* 37(3-4): 241-251.
4. Thalen, M., J. Van De IJssel, W. Jiskoot, B. Zomer, P. Roholl, C. De Gooijer, C. Beuvery and J. Tramper (1999). Rational medium design for *Bordetella pertussis*: basic metabolism. *Journal of Biotechnology* 75(2-3): 147-159.

5. Neeleman, R., M. Joerink, A.J.B. Van Boxtel, C. Beuvery and G. Van Straten (2001). On-line harvest prediction for the production of biologicals, *Computer Applications in Biotechnology*, Quebec, Canada, 430-435.
6. Metz, B., C.F.M. Hendriksen, W. Jiskoot and G.F.A. Kersten (2002). Reduction of animal use in human vaccine quality control: opportunities and problems. *Vaccine* 20: 2411-2430.
7. Anonymous (1997). General Statements (1.1.). *European Pharmacopoeia*.
8. Anonymous (1999). PDA Technical Report No. 30: Parametric Release of Pharmaceuticals Terminally Sterilized by Moist Heat. *Journal of Pharmaceutical Science and Technology* 53(4): 217-222.
9. Anonymous (1997). Methods of preparation of sterile products (5.1.1.). *European Pharmacopoeia*: 283-285.
10. Van Der Ark, A.A.J., I. Van Straaten, A.M. Akkermans, C.F.M. Hendriksen and H.J.M. Van De Donk (1994). Development of Pertussis Serological Potency Test: Serological assessment of antibody response induced by whole cell vaccine as an alternative to mouse protection in an intracerebral challenge model. *Biologicals* 22: 233-244.
11. Neeleman, R., M. Joerink and A.J.B. Van Boxtel (2001). Dual-Substrate Utilisation By *Bordetella pertussis*. *Applied Microbiology and Biotechnology* 57: 489-493.

10 CONCLUSIONS

10.1 Introduction

The central theme in this thesis, as well as its title, is *biomass performance*. With biomass performance the well being of biomass in a process is intended. Amongst others, that involves the concentration of biomass and its growth rate. In this thesis the use of process monitoring methodologies for the acquisition of information related to biomass performance is discussed. Obtaining this information yielded improved knowledge of the process and an accurate on-line monitoring tool.

Robust and reliable on-line technology enables the development of prediction and control strategies such as the right moment to harvest and substrate feeding. These techniques in turn enable the process operation to move away from predetermined schedules to physiologically based process decisions allowing manipulation of the engineering environment to control product formation and minimise unwanted impurities, by-products, or low yields. Furthermore, such monitoring tools result in increased assurance of consistency in process and product quality, ultimately leading to the prospects of parametric release.

10.2 Monitoring Biomass Performance

Successfully monitoring cultivation processes requires a good understanding of the interactions between the cell physiology and the physical and chemical environment in the bioreactor. Subsequently, the combination of off-line and on-line process data can be used to define the process, demonstrate process consistency, and maintain regulatory compliance. In this thesis several monitoring techniques are developed for which no or only little *a priori* information is needed. Two of these are organism-independent and can therefore be used for a variety of organisms and/or processes.

The respiration quotient is considered a good indicator of metabolic activity and is calculated from the oxygen uptake rate and carbon dioxide production rate. Measuring the carbon dioxide production rate on-line is difficult, especially when bicarbonate buffered media are applied. Chapter 3 describes an on-line Kalman filter that has no knowledge of the used organism and still gives a good estimate of the carbon dioxide consumption rate regardless of an existing bicarbonate buffer. The organism-independence and on-line

CONCLUSIONS

characteristics of this algorithm make it generally applicable in many processes. Subsequently, the on-line available respiration quotient can be used to monitor the process and/or control the process such that a specific respiration quotient is maintained.

The next step in monitoring biomass performance is the acquisition of specific knowledge about the micro-organism used, in order to develop a model that describes its growth and substrate usage. In particular, knowing the development through time of the specific growth rate improves the knowledge of the performance of the micro-organisms. Traditionally the specific growth rate is calculated off-line from biomass concentration measurements with methods such as exponential or polynomial fitting or walking averages. However, these methods do not give satisfactory results nor do these methods take the noise characteristics of the biomass measurements into account. In Chapter 4 a smooth estimation technique for off-line determination of the specific growth rate is developed. This technique does not need knowledge of the organism, its intercellular behaviour, or the process. With cultivation data of *B. pertussis* and *N. meningitidis* this technique showed to be useful for determining the growth characteristics of micro-organisms based on measurements of biomass concentration only.

Where off-line estimation of the specific growth rate is a powerful tool for model development and process consistency checking, on-line estimation finds its applications in monitoring, control, and optimisation. If at each moment in time estimates of the specific growth rate and biomass concentration are available, they can be exerted to calculate a feed rate or predict the future behaviour of the process.

An on-line sequential estimator is developed in Chapter 6, where only little organism-specific information is needed to give accurate estimations of the biomass concentration and its specific growth rate. These estimations are based on generally available measurements of the oxygen concentration only. Although in Chapter 6 the estimator is developed for the cultivation of *B. pertussis*, it is based on a generic growth and oxygen consumption model, and can thus be applied to various bioprocesses. With only a few data sets, minor parameter estimation and tuning, the parameters of the sequential estimator are adjusted for another micro-organism, making this a strong and universal tool.

10.3 Knowledge of Biomass Performance

Modelling has become a prerequisite for the development, optimisation, and control of bioprocesses. With the aid of an accurate model it is feasible to perform simulations in order to improve process knowledge and investigate/test applications of monitoring and control. The usefulness of simulation models in connection with experimental planning is also evident.

An important step in modelling micro-organisms in a bioreactor is correlating the development through time of the specific growth rate with substrate consumption. The tool presented in Chapter 4 gives such useful information by objectively estimating the specific growth rate from biomass samples. For *B. pertussis* this resulted in the development of a dual-substrate model where the two substrates simultaneously limit the growth of the organism in an exceptional way, as opposed to the commonly expected proposition that growth is limited by the most limiting substrate. Chapter 5 describes this model that combines essential and enhanced kinetics and which was validated with experiments.

With the improved understanding of the metabolism of the organism and its substrate utilisation the process and the used medium are fine-tuned. Thus modelling and simulation are valuable tools in basic biological research. Moreover, the quest for a proper model direct investigators towards a search for quantifiable results.

10.4 Predicting Biomass Performance

To achieve good product quality and consistency the bioreactor has to be harvested in time. Since downstream processing needs to be prepared and operators need to be ready for harvest, there is a demand for an algorithm that is capable of predicting the right moment of harvesting. Furthermore, a prediction algorithm gains the advantage to let the moment of harvesting be dictated by the state of the organisms instead of by a predetermined time-schedule.

In order to predict the future trajectory of a bioprocess it is inevitable to know the past and current status of the bioprocess. Combining the dual-substrate model of Chapter 5 with the monitoring system of Chapter 6 resulted in the on-line reconstruction of the complete state of the process, i.e. concentrations

of biomass and substrates. With this known current status, a simulation of the future trajectory is made in order to predict the moment of substrate depletion and thus the right moment of harvesting. Chapter 7 describes such a prediction algorithm which can be performed on operator request or fully automatic on a regular interval.

Besides using a prediction algorithm for the moment of harvesting it can be used for other events that have to take place during a production process, e.g. virus inoculation, feeding, inducing expression, etc..

10.5 Manipulating Biomass Performance

A very popular way to influence or manipulate a bioprocess is operating in fed-batch mode where the substrate is slowly fed to the reactor but no product is drawn until the end. A fed-batch culture has the advantage of avoiding substrate overfeeding, which can inhibit the growth of micro-organisms. Moreover, the fact that no substance is withdrawn from the reactor helps the process to work in good sterilised conditions. On the other hand, from the control engineer's viewpoint, it is the fed-batch process that presents the greatest challenge: the process variables are difficult to measure, the 'quality' of the product is difficult to define yet very important, the process model usually contains strongly time-varying parameters, etc. But above all, the challenge arises mainly because the control of the process by adjusting the feed rate is an on-line, dynamical problem.

In spite of its importance, control of the specific growth rate is not widespread in the pharmaceutical industry. One reason is that typical industrial cultivations are not sufficiently instrumented for applications of the control algorithms proposed so far. Another reason is that many of the control algorithms are too complex and often not robust and reliable enough. Chapter 8 describes the use of the on-line monitoring system presented in Chapter 6 to create a simple and robust automatic control system for the specific growth rate. The algorithm uses only oxygen measurements and was shown to maintain a constant specific growth rate for more than 15 hours. Moreover the complete system has a modular approach and each individual module can be implemented, tested, tuned, and validated separately. That results in a less complex, easy to understand, and off-the-shelf application for the pharmaceutical industry.

10.6 Assurance of Biomass Performance

The interest in alternatives, reductions, and refinement methods for the use of animals in the development and control of biological products has increased considerably in the last few years. The issues raised are of concern to the scientific community as well as to manufacturers, regulatory authorities, and those involved in animal ethics and welfare. For this purpose, a large number of physicochemical, immunochemical, biochemical, and cell biological methods have been proposed [1].

For the release of product a monitoring system should be in place that gives the highest possible assurance of consistency in product quality. Here the keyword 'assurance' should be emphasised. Measuring the actual product quality is not always feasible and does not necessarily have to give the highest level of assurance. For instance, mouse potency tests are statistically hardly capable of indicating neither batch-to-batch variation nor product quality. Since the monitoring technique described in Chapter 6 does give accurate information about batch-to-batch variation, the release of product should be based on these on-line process data. Such a procedure would be called parametric release.

The basis for parametric release is intrinsic process quality and extensive product and process knowledge, leading to clear, thorough, robust, and consistent processing. Although the implementation of such a procedure was demonstrated, and validation is easily attainable, regulatory acceptance is still at a very early stage. Therefore, efforts should be made in strengthening the communication between research scientists, manufacturers and regulators to promote the exchange of scientific knowledge and to increase awareness of the importance of parametric release.

10.7 The Modular Approach

All techniques in this thesis have been developed such that they can be applied in a modular approach. Each chapter describes one or two separate tools or modules to be used and combined to reach one of the final goals. Each module or tool is implemented in a separate and individual software routine. The strength of a modular approach is the possibility to implement, tune, test, and validate each module or tool separately and the opportunity to replace a single

module with a new or improved algorithm without harming the combined application.

When connecting all modules to a combined application, the input and output specifications of each module need to be defined precisely. Table 1 shows briefly all modules, tools, and combinations with their inputs (measurements or outputs of other modules or tools) and outputs.

Table 1: Modules and tools discussed in the various chapters of this thesis with their inputs and outputs.

Tool / Module	Ch.	Input meas./ Output of	Output
A RQ-estimation	3	CO_2 and O_2	RQ
B off-line μ estimation	4	C_X	μ
C simulation model	5	$C_X(t)$, $C_S(t)$, tool B	$C_X(t+1)$, $C_S(t+1)$
D on-line μ estimation	6	OUR	μ , $d\mu/dt$
E on-line C_X estimation	6	OUR , module D	C_X
F state reconstruction	7	module D and E	C_X , C_G , C_L
G harvest prediction	7	module F	time to harvest
H feedback μ -control	8	μ_{sp} , module D	μ_{ref}
I feed-forward μ -control	8	μ_{ref} , module C, E, and H	F_G , F_L
J parametric release	9	module D and E	approval
K safety control		App $C_X(off-line)$, module I	F_G , F_L

10.8 Prospects

Managing Biomass Performance

In the past, computers in bioprocessing generally have been used for process monitoring, data acquisition, data storage, low-level control, and error detection. Computer capabilities are now further exploited by the implementation of advanced control and optimisation strategies, which potentially provide significant improvement in product quality, yield, consistency, and productivity [2].

The extent of computer application depends on the development of better on-line instrumentation that can monitor cellular physiology, as well as on reliable mathematical models that can describe various cellular dynamics. When the availability of on-line sensors is limited, mathematical models and estimation techniques will play very important roles in optimising bioreactors as shown in this thesis.

Traditionally, in operator controlled processes, the operator has been responsible for monitoring the current status of the bioreactor, initiating proper control commands (for example emptying the vessel for harvesting or starting the feed in a fed-batch culture), and for specifying control parameters (set-points, flow rates). The next logical challenge in the production of biopharmaceuticals is a broadly applicable algorithm that is capable of plant-wide control. Such an expert system will detect off-specifications, will give advice about starting the feed or initiating harvest, and will monitor the performance. In other words computer technology can now be used to perform tasks usually carried out manually in order to automate process decision making. Moreover, for direct process control, e.g. control of the growth rate, advanced computer control is imperative since such control can not be done manually. Chapter 7 and 8 present applications that make process decisions based upon an on-line monitoring system.

In-depth Biomass Performance

One of the obstacles hampering the further development of bioprocesses is the lack of understanding microbial metabolic pathways and cellular control mechanisms. Such knowledge is needed in order to model the process and formulate meaningful process control algorithms so that the final objective of process optimisation can be achieved.

There have been few reports in the literature on extending metabolic flux analysis to on-line state recognition of cultivation processes, using only measured values acquired on-line. On-line estimation of metabolic reaction rates has been successfully performed by macroscopic elemental and heat balance methods. However, the estimated reaction rates were limited to extracellular metabolites only. Takiguchi et al. [3] present an on-line method for estimating the flux distribution of intracellular metabolites from on-line measurements only.

Beyond Biomass Performance

Besides monitoring biomass performance during the cultivation step, such advanced monitoring techniques are also promising for application at other bio-pharmaceutical process steps. Filter fouling and column ageing are only two examples of downstream processing steps that will benefit from advanced monitoring techniques. The ability to recognise filter fouling from pressure changes or column ageing from the changes in elution pattern would greatly contribute process quality. Furthermore, it would shift process decisions being made on a pre-determined schedule to being dictated by the state of the filter or column.

From Biomass Performance to Company Performance

Accelerating the drug development process while controlling costs is of critical importance in the biopharmaceutical industry, now faced with shortened product life cycles due to increased competition from generic drugs. The time span over which companies have to recoup their investment is shrinking while the cost of developing drugs is rising. Hence, to achieve an acceptable return on this investment, biopharmaceutical companies need to focus on cutting down the cost of drug development and improving the overall time-to-market.

Biopharmaceutical companies typically have a portfolio of drug candidates to manufacture for clinical trials, but with finite response resources, budget and capacity. Therefore, it is imperative that process development is improved and R&D portfolio management is made more effective. However, managing an R&D portfolio is complicated by the fact that each project is subject to technical and market uncertainties. Incorporating the effects of risk analysis helps enhance the quality of decision-making within a company. Consequently, the need for computer-aided data analysis and simulation tools, capable of capturing both the technical and business aspects of manufacturing processes, is critical for such decision-making.

10.9 References

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

2. Dorresteyn, R.C., G. Wieten, P.T.E. Van Santen, M.C. Philippi, C.D. De Gooijer, J. Tramper and E.C. Beuvery (1997). Current good manufacturing practice in plant automation of biological production processes. *Cytotechnology* 23(1-3): 19-28.
3. Takiguchi, N., H. Shimizu and S. Shioya (1997). An on-line physiological state recognition system for the lysine fermentation process based on a metabolic reaction model. *Biotechnology and Bioengineering* 55(1): 170-181.

SUMMARY

The primary concern in the pharmaceutical industry is not the optimisation of product yield or the reduction of manufacturing cost, but the production of a product of consistently high quality. This has resulted in 'process monitoring' becoming an integral part of process operation. In this thesis process monitoring is one of the central themes, from monitoring the environment of the micro-organisms to monitoring the micro-organisms themselves. The latter is called monitoring *biomass performance*. Subsequently, in this thesis, monitoring is applied to predict the future trajectory of the cultivation and thus the time to harvest. Furthermore, it is applied to indicate process consistency and thus to give assurance about product quality. Finally, monitoring is applied to control the specific growth rate in order to improve process consistency and thus to improve assurance of product quality.

One of the key metabolic indicators suitable for monitoring the performance of cell cultures is the respiration quotient. Usually gas analysis is used to monitor this variable. However, in bicarbonate buffered media the carbon dioxide balance is affected by accumulation and hence the respiration quotient can not directly be calculated from gas measurements. A Kalman filter is introduced that estimates the carbon dioxide evolution rate and copes with these buffering capacities. The model used by the Kalman filter lumps all carbonate in the liquid to one term in order to eliminate the role of a priori knowledge of cell and medium kinetics without affecting the performance. A reference experiment verified the performance of the software sensor and subsequent experiments with insect cells showed the progress of the respiration quotient during cultivation.

To take advantage of the newest techniques in monitoring and control, more knowledge of the micro-organism involved is necessary. The specific growth

rate is an important indicator of *biomass performance*. An off-line algorithm is developed that is capable of determining the specific growth rate from off-line biomass samples. The technique is based on combining subsequent backward and forward extended Kalman filtering to give a smoothed and optimal estimation. It can be used as a powerful tool for various organisms since again no a priori knowledge of the micro-organism is needed. This estimator gained improved knowledge of the growth characteristics of *Bordetella pertussis* and *Neisseria meningitidis*. Both organisms seemed to be simultaneously limited by two substrates. However, for the case of *B. pertussis*, neither interactive nor non-interactive modelling seemed appropriate and a model that combines essential and enhanced kinetics was developed based on experimental observation. Instead of fitting all model parameters at once, a step-wise experimentation procedure was used and the accuracy of the dual-substrate model was shown by two cultivations.

The next step in monitoring the cultivation step in pharmaceutical production is monitoring the biomass concentration and its growth rate on-line. Due to the lack of reliable and cheap sensors, they cannot often be measured directly or estimated from related variables, such as the concentrations of substrates or products. A sequential observer is introduced to estimate the specific growth rate and the biomass concentration for processes where the measurement of oxygen uptake rate is available on-line. The applicability of the algorithm is proven by the good agreement between the sequentially estimated values and the measured values for the cultivation process of *B. pertussis*.

One of the applications of this monitoring system is the combination of the current estimated state of the cultivation with the dual-substrate model of *B. pertussis* in order to reconstruct on-line, the total current state of the cultivation (e.g. biomass and substrate concentrations). With the estimated current state, the future trajectory of the cultivation is computed by forward simulation. This prediction algorithm successfully aids the operator in determining and scheduling the right moment of harvest on-line. Thus, this prediction algorithm gains the advantage of letting the moment of harvest being dictated by the state of the organisms instead of by a predetermined time-schedule, which ultimately leads to improved process consistency.

Another application of the monitoring system is the use of the available on-line estimates to track whether the process remains within specifications. The

SUMMARY

sequential estimator gains more information than regular process monitoring. Now *biomass performance* is available on-line, and can thus be used to indicate consistency at a metabolic level. For the case of *B. pertussis*, the final product test is statistically hardly capable of indicating batch-to-batch variation. Monitoring on-line process parameters provides a better assurance of consistency and quality, thus giving the opportunity for batch release and regulatory approval without the use of the final animal tests.

To further improve process consistency during cultivation, the metabolic state of the micro-organisms in the bioreactor has to be kept constant. By controlling the specific growth rate of biomass such metabolic stability is guaranteed. This thesis demonstrates the combined use of the sequential estimator, the dual-substrate model, and a feed-forward-feedback-controller to keep the specific growth rate constant for *B. pertussis*. Successful tests were performed in real experiments where the specific growth rate was controlled at reduced specific growth rate over a period of more than 15 hours, quadrupling the total amount of biomass. In order to apply this controller to other processes it is necessary to conduct a relatively small number of experiments providing enough data to determine and tune the parameters.

All techniques in this thesis are proposed in a modular approach where each tool or module is implemented in a separate software routine, such that they can be tuned, tested, improved and validated separately. Connecting the modules results in one of the mentioned combined applications.

SAMENVATTING

Het grootste belang in de farmaceutische industrie is niet de optimalisatie van product opbrengst of de reductie van productie kosten, maar de productie van een product van consistente en hoge kwaliteit. Dat heeft erin geresulteerd dat ‘proces monitoring’ een integraal deel uitmaakt van de procesvoering. In deze thesis is ‘proces monitoring’ één van de centrale thema’s. Van het monitoren van de omgeving van de micro-organismen tot het monitoren van de micro-organismen zelf. Het laatste wordt in deze thesis het monitoren van *biomassa prestatie* genoemd. Vervolgens wordt in deze thesis ‘monitoring’ gebruikt om het toekomstig verloop van een cultivatie te voorspellen en dus het moment van oogsten. Verder wordt het toegepast om proces consistentie aan te tonen en dus om een vorm van betrouwbaarheid te geven over product kwaliteit. Tenslotte wordt ‘monitoring’ toegepast om de proces consistentie en dus de product kwaliteit te verhogen.

Één van de sleutel variabelen geschikt voor het monitoren van de prestatie van cel culturen is the respiratie quotiënt. Gebruikelijk wordt gas analyse gebruikt om deze variabele te monitoren. Echter, in bicarbonaat gebufferde media wordt de koolstofdioxide balans aangetast door accumulatie en dus kan de respiratie quotiënt niet direct van de gasbalans worden berekend. Een Kalman filter die rekening houdt met deze bufferende werking is gebruikt om de koolstofdioxide productie snelheid te schatten. Het model dat door het Kalman filter wordt gebruikt voegt alle carbonaat termen in de vloeistof samen tot één term. Een referentie experiment verifieerde de prestatie van de software sensor. Enkele daaropvolgende experimenten met insectencellen demonstreerde het verloop van de respiratie quotiënt gedurende cultivatie

Om gebruik te maken van de nieuwste technieken op het gebied van meet- en regeltechniek is er meer kennis van de micro-organismen nodig. De specifieke

groeisnelheid is een belangrijke indicator van *biomassa prestatie*. Een off-line algoritme is ontwikkeld dat in staat is om de specifieke groeisnelheid te schatten vanuit off-line biomassa monsters. Deze techniek is gebaseerd op de combinatie van een voorwaarts en een terugwaarts ‘extended Kalman filter’ om zo een vloeiende en optimale schatting te krijgen. Het kan worden gebruikt als een krachtig gereedschap voor diverse organismen omdat geen voorkennis over het micro-organisme nodig is. De schatter zorgde voor verbeterde kennis omtrent de groeikarakteristieken van *Bordetella pertussis* en *Neisseria meningitidis*. Beide organismen lijken door twee substraten tegelijkertijd gelimiteerd te worden. Echter, voor *B. pertussis*, is noch een interactief noch een niet-interactief model van toepassing. Zodoende is op basis van experimentele observaties een model ontwikkeld dat essentiële en additieve kinetiek combineert. In plaats van alle model parameters tegelijkertijd te schatten is er gebruik gemaakt van een stapsgewijze experimentele procedure. De nauwkeurigheid van dit dubbelsubstraat model is gedemonstreerd met behulp van twee extra experimenten.

De volgende stap in ‘monitoring’ de cultivatie is het schatten van de biomassa concentratie en de specifieke groeisnelheid. Wegens een gebrek aan betrouwbare en goedkope sensoren, kunnen deze gegevens niet worden gemeten of direct worden geschat vanuit bijvoorbeeld substraat metingen. Een sequentiële schatter is ontwikkeld welke de specifieke groeisnelheid en biomassa concentratie schat vanuit on-line metingen van de zuurstof opname snelheid. De toepasbaarheid van het algoritme is bewezen door de goede overeenkomst tussen de sequentieel geschatte waarden en de gemeten waarden voor de cultivatie van *B. pertussis*.

Één van de toepassingen van dit ‘monitoring’ systeem is de combinatie van de huidige schatting met het dubbelsubstraat model van *B. pertussis* om zo de complete huidige toestand (biomassa en substraat concentraties) on-line te reconstrueren. Met deze geschatte toestand kan middels simulatie het toekomstig verloop van de cultivatie worden berekend. Dit predictie algoritme helpt de operator met het bepalen van het juiste moment van oogsten. Dus, dit predictie algoritme zorgt ervoor dat het oogstmoment kan worden bepaald door de toestand van de organismen zelf, in plaats van op een voorafgestelde tijd, wat uiteindelijk leidt tot verbeterde proces kwaliteit en consistentie.

Een andere toepassing van het ‘monitoring’ systeem is om de schattingen te gebruiken om on-line te bepalen of het proces nog binnen de opgestelde limieten blijft. De sequentiële schatter verwerft meer belangrijke informatie dan standaard sensoren. Hiermee is de *biomassa prestatie* dus toegankelijk on-line en kan worden gebruikt om consistentie op metabool niveau aan te tonen. Voor *B. pertussis* is de laatste product test statistisch moeilijk in staat om batch-op-batch variatie aan te tonen. Daartegenover geeft dit ‘monitoring’ systeem meer zekerheid over consistentie en kwaliteit en dus de gelegenheid voor batch vrijgifte zonder het gebruik van diertests.

Om de proces consistentie verder te verbeteren moet de metabole toestand van de organismen in de bioreactor constant worden gehouden. Door de specifieke groeisnelheid te regelen kan zo’n metabool stabiele situatie worden bereikt. In deze thesis wordt een combinatie van de sequentiële schatter, het dubbelsubstraat model en een ‘feed-forward-feedback’ regelaar toegepast om de specifieke groeisnelheid van *B. pertussis* constant te houden. Succesvolle experimenten tonen aan dat de specifieke groeisnelheid constant kan worden gehouden over een periode van meer dan 15 uur, onderwijl de totale biomassa hoeveelheid verviervoudigend. Om deze regelaar toe te passen bij andere organismen is een relatief klein aantal experimenten noodzakelijk om genoeg data te krijgen voor het bepalen en ‘tunen’ van de parameters.

Alle technieken in deze thesis zijn modulair van opbouw en ieder stuk gereedschap of module zit in een aparte software routine zodat ze afzonderlijk kunnen worden geïmplementeerd, getuned, getest, verbeterd en gevalideerd. Het verbinden van de modules resulteert in één van de genoemde toepassingen.

BIBLIOGRAPHY

- R. Neeleman (1995). Process-Integration and Plant-Controllability for Continuous Affinity Recycle Extraction. Masters Thesis, Wageningen University, The Netherlands.
- R. Neeleman (1996). On-line Estimation of the Respiration Quotient of Mammalian Cells during Batch Cultivation. Masters Thesis, Wageningen University, The Netherlands.
- R. Neeleman, J.E. Claes and J.F. Van Impe (1996). Modeling, Monitoring and Control of Bakers' Yeast Fed-Batch Fermentation. Technical Report LIMB 1996-08, BioTec Group, Catholic University Leuven, Belgium
- R. Neeleman, E.J. van den End and A.J.B. van Boxtel (1999). Estimation of Respiration Quotient in Bicarbonate Buffered Media. Proceedings of 9th European Congress on Biotechnology, Brussels, Belgium, p. 2346.
- R. Neeleman (2000) Modelling of enhanced substrate utilisation with the aid of extended Kalman filtering. Proceedings of 19th Benelux Meeting on Systems and Control, Mierlo The Netherlands, p. 28.
- R. Neeleman, E.J. van den End and A.J.B. van Boxtel (2000). Estimation of the Respiration Quotient in a Bicarbonate Buffered Batch Cell Cultivation. Journal of Biotechnology, Vol 80, pp. 85-95.
- R. Neeleman and A.J.B. van Boxtel (2001). Observer-based prediction of Time to Harvest for Bioreactor Cultivation. Proceedings of 20th Benelux Meeting on Systems and Control, Houffalize Belgium, p. 131.

BIBLIOGRAPHY

R. Neeleman, M. Joerink, A.J.B. van Boxtel, E.C. Beuvery and G. van Straten (2001). On-line Harvest Prediction for the Production of Biologicals. Proceedings of 8th Int. Conf. on Computer Applications in Biotechnology, Quebec Canada, pp. 430-435.

R. Neeleman (2001). Respiration Quotient: Estimation during batch cultivation in bicarbonate buffered media. In: Focus on Biotechnology, Volume IV, pp. 203-216. Kluwer Academic Publishers, The Netherlands, ISBN 0-7923-6927-0.

R. Neeleman, A.J.B. van Boxtel (2001). Estimation of Specific Growth Rate from Cell Density Measurements. Bioprocess and BioSystems Engineering, Vol 24, pp. 179-185.

R. Neeleman, M. Joerink, E.C. Beuvery and A.J.B. van Boxtel (2001). Dual-Substrate Utilisation by *Bordetella pertussis*. Applied Microbiology and Biotechnology, Vol57(4), pp. 489-493.

R. Neeleman and E.C. Beuvery (2002). Advanced Process Monitoring: The Ultimate Alternative. In: Developments in Biologicals, Volume 111, pp. 275-284. IABs, Basel, Karger, Switzerland.

R. Neeleman, M. Joerink, D. Vries and A.J.B. van Boxtel. Sequential Estimation of Biomass and Specific Growth Rate. In preparation.

R. Neeleman, D. Vries, Z. Soons and A.J.B. van Boxtel. Dual-Substrate Feedback Control of the Specific Growth Rate. In preparation.

CURRICULUM VITAE

Ronald Neeleman was born in July 1971 in Rotterdam, The Netherlands. He received his Master of Science degree in biotechnology at Wageningen University in 1996. From 1996 to 1997, he worked as bioprocess engineer at the laboratory for product and process development of the Dutch National Institute of Public Health and the Environment (RIVM) in Bilthoven, The Netherlands. In February 1998 he started his Ph.D. study at the Systems and Control Group of Wageningen University and received the best presentation award during the Benelux meeting on Systems and Control in March 2001. Besides his study and current work he was chairman of UNIT-DISC, the Ph.D. department of the Dutch Institute of systems and Control from 1999 to 2001, is an active IDD scuba dive instructor since March 2001 and started his own pharmaceutical software company called Nephaso in November 2001. Currently he is again employed at the laboratory for product and process development at RIVM.