Identification and control of a fed-batch process : application to culture of Saccharomyces cerevisiae

Citation for published version (APA):
Keulers, M. L. B. (1993). *Identification and control of a fed-batch process : application to culture of Saccharomyces cerevisiae*. [Phd Thesis 1 (Research TU/e / Graduation TU/e), Electrical Engineering]. Technische Universiteit Eindhoven. https://doi.org/10.6100/IR400351

DOI:
10.6100/IR400351

Document status and date:
Published: 01/01/1993

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

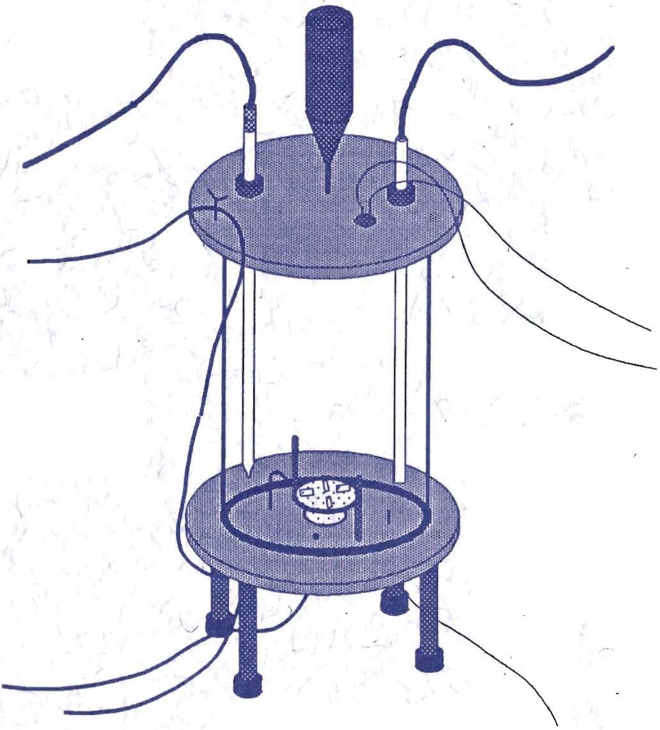
Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 04. Oct. 2023

Identification and Control of a Fed-Batch Process; Application to Culture of *Saccharomyces cerevisiae*



Marc Keulers

Identification and Control of a Fed-Batch Process; Application to Culture of *Saccharomyces cerevisiae*

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de
Technische Universiteit Eindhoven, op gezag van
de Rector Magnificus, prof. dr. J.H. van Lint,
voor een commissie aangewezen door het College
van Dekanen in het openbaar te verdedigen op
donderdag 2 september 1993 om 16.00 uur

door

Marcus Laurentius Bernardine Keulers

Geboren te Voerendaal

Druk: Boek- en Offsetdrukkerij Letru, Helmond, 04920-37797

Dit proefschrift is goedgekeurd
door de promotoren

prof. dr. ir. P. Eykhoff
en
prof. dr. ir. A.C.P.M. Backx

Copromotor : dr.ir. A.J.W van den Boom

Abstract

The production of bakers' yeast is not only important for bakers. From an industrial point of view, (bakers') yeast is becoming more and more important as it sometimes, after modification, supplies enzymes, flavours, essences, proteins, etc. The control engineer has to ensure that the yeast can grow under favourable conditions to produce these products. Yet, most of the production techniques used nowadays are based on feed recipes of important nutrients. In order to develop new production techniques process modelling, identification, and control are applied to the bakers' yeast process.

Before one is able to do any modelling, identification, or control, the process itself must be studied. A comprehensive study has been done on the process characteristics, the sensors, and the actuators. The information obtained sets the limitations for any identification or control application.

The next step is to make a simulation model of the process that can be used for testing identification and control algorithms. Both a batch and fed-batch phase simulation model of the growth of *Saccharomyces cerevisiae* (bakers' yeast) have been derived. These models are based on physiological/physical laws and equations. The parameters of the simulation models have been estimated using experimental laboratory data. The simulation models have been validated with data from other experimental laboratory experiments.

The goal of this work is to control the growth of *Saccharomyces cerevisiae*. To be able to do so, control models need to be identified. The identification techniques used to obtain control models included both linear and non-linear algorithms. A linear model of the oxidative growth on glucose is identified. This is done as follows. Firstly the process inputs are compensated with a non-linear compensation such that the resulting input-output behaviour is more or less linear. Secondly, well-known techniques such as Auto Regressive and Output Error technique are used to make a model of the linearized process.

The non-linear identification techniques used are block-oriented models based on Volterra kernels, non-linear models linear-in-the-parameters, and the Radial Basis Function network. The Radial Basis Function network is a sort of neural network. The models found using these non-linear identification techniques are not very suitable for control. Nevertheless, they can be used for one step ahead prediction and to correct sensor failure.

Despite using different identification techniques, no model of the specific growth rate could be made. The specific growth rate describes the growth of the culture and is thus a key variable. To overcome the problem of the non-measurable specific growth rate a simple observer has been developed on the basis of on-line measurements. This observer is able to estimate the specific growth rates and the cell concentration of *Saccharomyces cerevisiae*. Furthermore, the observer can distinguish between the different growth contributions of the culture. This observer is inherently stable and converges.

Based on the observer and using the non-linear compensation found with the linear models, a control strategy involving the aerobic and fermentative growth of *Saccharomyces cerevisiae* has been developed. The controller includes a PI block and an exponential feed-forward block. The controller has first been extensively tested and tuned on the simulation models. Good experimental results have been obtained with this controller. After using the controller for regulation of the specific growth rate, some new, unexpected dynamic behaviour of the culture could be reported.

Contents

Glossary

1 **Introduction** 1
- 1.1 Motivation 3
- 1.2 Problem Statement 5
- 1.3 Outline of this Thesis 8
- References 11

2 **Process Description** 13
- 2.1 The Yeast Cell 14
- 2.2 The Laboratory Model Process 16
- 2.3 Microorganisms and Culturing Techniques 17
- 2.4 Fermenter Operation 18
- 2.5 The Experimental Site 19
- 2.6 Conclusions 27
- References 30

3 **Simulation Models** 31
- 3.1 Bio-Process Modelling 32
- 3.2 Quantitative Process Analysis 35
- 3.3 The Batch Model 43
- 3.4 The Fed-Batch Model 59

3.5 Conclusions 68
References 68

4 Linear Identification 73
4.1 Process Analysis 74
4.2 Non-Linear Compensation 77
4.3 Experiments 78
4.4 Estimation of a Linear Model 82
4.5 Validation of the Model 86
4.6 Conclusions 87
References 88

5 Non-Linear Identification 91
5.1 Identification Methods 92
5.2 Simulation Data 98
5.3 Results 99
5.4 Conclusions 127
References 129

6 Observers for Control 131
6.1 Bio-Process Observers 132
6.2 Model Based Observer 134
6.3 Results 138
6.4 Conclusions 143
References 144

7 Control 147
7.1 Literature Review 148
7.2 Control Structure 151
7.3 Results 171
7.4 Conclusions 184
References 185

8 Conclusions 191

List of Examples 197

List of Figures 199

List of Tables 203

Samenvatting 205

Curriculum Vitae 207

Glossary

Abbreviations

ADP	Adenosine-5'-DiPhosphate
ARMAX	Auto-Regressive Moving Average auXilary input
ATP	Adenosine-5'-TriPhosphate
BJ	Box-Jenkins
CDAS	Control and Data Acquisition Software
FAD	Flavine Adenine Dinucleotide
NAD	β-Nicotinamide-Adenine Dinucleotide
NADH	β-Nicotinamide-Adenine Dinucleotide, reduced
NLC	Non-Linear Compensation
OE	Output Error
PI	Proportional Integral controller
PRBS	Pseudo Random Binary Sequence
P_c	Product costs
P_s	Price per unit product
Q_{crit}	Quality criterion
TC	Time Constant
TCA	TriCarboxylic Acid
VAX	Computer
YPD	Yeast extract, Peptone, and Dextrose
dcw	dry cell weight

f_s	sample frequency

Variables

A,B,C,D,F	State equation matrices	[-]
AF	Air Flow	[dm^3.h]
CPR	Carbon-dioxide Production Rate	[mole.h^{-1}.dm^{-3}]
DOT	Dissolved Oxygen Tension	[mole]
E	Ethanol concentration	[mole.dm^{-3}]
GF	Glucose Flow	[dm^3.min^{-1}]
K_La	Oxygen transfer coefficient	[h^{-1}]
OUR	Oxygen Uptake Rate	[mole.h^{-1}.dm^{-3}]
Q_{emax}	Regulatory ethanol enzyme	[mole.h^{-1}.dm^{-3}]
Q_{omax}	Regulatory oxygen enzyme	[mole.h^{-1}.dm^{-3}]
Q_{smax}	Regulatory glucose enzyme	[mole.h^{-1}.dm^{-3}]
RQ	Respiratory Quotient	[-]
S	Glucose concentration	[mole.dm^{-3}]
SS	Stirrer Speed	[rpm]
T	Time	[h]
U	Dimensionless variable	[-]
V	Volume	[dm^3]
X	Biomass concentration	[mole.dm-^{3}]
$Y_{x\|s}$	Biomass yield on glucose	[mole.mole^{-1}]
e	Noise	[-]
$h_{0,1,2}$	Volterra kernels	[-]
u	Input	[-]
y	Output	[-]

Constants

K_e	Saturation parameter for ethanol uptake	[mole.dm^{-3}]
K_i	Inhibition parameter: free glucose inhibits ethanol uptake	[mole.dm^{-3}]
K_m	Substrate saturation constant for the induction of the production of oxidation capacity	[mole.dm^{-3}]
K_n	Substrate saturation constant for the induction of the production of oxidation capacity	[mole.dm^{-3}]
K_o	Saturation parameter for oxygen uptake	[mole.dm^{-3}]
K_s	Saturation parameter for glucose uptake	[mole.dm^{-3}]
Q_{em}	Maximum regulatory ethanol enzyme	[mole.dm^{-3}]
Q_{om}	Maximum regulatory oxygen enzyme	[mole.dm^{-3}]
Q_{sm}	Maximum regulatory glucose enzyme	[mole.dm^{-3}]
a-j	Stoichiometric constants	[-]
ms	Maintenance	[h^{-1}]
nk	Delay	[-]
q_s	Glucose consumption rate	[h^{-1}]

Greek Symbols

μ	Specific growth rate	$[h^{-1}]$
τ	Time constant	[h]

Superscripts

oxid	Oxidative
red	Reductive

Subscripts

e	Ethanol
f	Final time
o	Oxygen
s	Glucose
hx	Hydrogen content in yeast
nx	Nitrogen content in yeast
ox	Oxygen content in yeast
max	Maximum
0	Initial time, t=0 [h]

1

Introduction

Working with yeasts is a fascinating experience as they tend to be like human beings; at the moment you think you do understand them they will surprise you with something completely different. It has been a challenge for many researchers to try to understand this living cell and its chemical reactions, especially the dynamic responses. Due to their complex and ever changing structure (mutations) this challenge will remain, possibly forever. Not only at cell level is there a challenge. From an industrial point of view, yeasts are becoming more and more important as they, sometimes after modification are suppliers of enzymes, flavours, etc. For production of any of these products the conditions for the biotechnical process have to be beneficial. So, maintaining favourable conditions for these workhorses is one of the main targets of the control engineer. Another working field for the control engineer is the development of laboratory tools

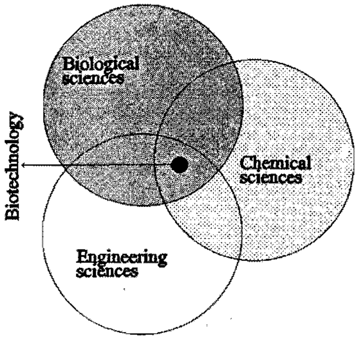


Figure 1.1 *Multidisciplinary nature of biotechnology*

to be able to expose the working of this magnificent creature in order to help to understand its behaviour better. Later on in this thesis we will elucidate these two working fields of the control engineer, but first we will show aspects of the role which control science plays in biotechnology.

As depicted in Figure 1.1, it is recognized that biotechnology is a coalescent of different sciences, i.e. biological, chemical, and engineering sciences. From this coalescence of sciences we will derive models. These models are the core of this thesis. All the models are, somehow, related to, or based on the biotechnological and chemical reactions taking place in a fermenter. However, the reader is reminded that the research in this thesis is carried out with a control engineer's viewpoint, as many of the theories and practical algorithms emerge from the systems control field. Before we get to the motivation and problem statement, we will give some brief history on biotechnology and a short explanation of a yeast cell.

History

Today biotechnology is often presented as something relatively new, perhaps the latest in the technical development, but biotechnology is rather old. The origins of biotechnology are found at the beginning of human history. The history of biotechnology can be divided into four phases (Enari, 1982).

The first biotechnical processes were alcoholic fermentations (wine and beer) and lactic acid fermentation of milk products. The Sumerians discovered alcoholic fermentation perhaps ten thousand years ago. About five thousand years later the Egyptians learned to use yeast for baking.

The second period in the history of biotechnology was the development of microbiology, about a hundred years ago. In 1867 Louis Pasteur published one of his books, "Étude sur la bière", in which he showed that yeast was responsible for fermentation. He also discovered that lactic acid bacteria were spoiling wine and beer by causing lactic acid fermentation. In 1882 Robert Koch developed solid media, which made it possible to isolate single cell cultures. Just one year later pure cultures were applied for the first time to industrial fermentations by Emil Christian Hansen at the Carlsberg brewery.

The third step was the development of fermenter technology. This period started in the 1940s with the production of penicillin and lead to the development of large submerged fermenters. The largest of these units operating today has a capacity of 1500 m^3. Further steps in fermenter technology were the introduction of continuous fermentation ten years later, and the development of immobilized enzymes and cells in the 1970s.

The fourth phase in the history of biotechnology can be called the period of recombinant DNA technology. This technology is concerned with the biological information stored in the hereditary material of the cells. Recombinant DNA technology, which emerged from research in the 1970s, introduced the possibility of manipulating the genetic properties of microorganisms. At the same time computer technology was introduced, which made process control feasible in the 1970s.

Yeast

We will give a general overview of the yeast cell and define some relevant process parameters. A more extensive explanation of how the yeast cell works is given in the next Chapter. A graphical representation of the yeast cell is shown in Figure 1.2. As can be seen from the figure, the yeast cell multiplies through budding. In order to be able to bud the yeast cell needs building material and energy. The yeast cell uses sugars, e.g. glucose or maltose, to derive building material and energy. Beside the sugars (carbon sources) there are a few more items, e.g. vitamins, minerals, etc., the yeast cell uses when it grows. One of the main items is oxygen (see Figure 1.2). The yeast cell can use the oxygen together with the sugar to grow without ethanol production. If no oxygen is available the yeast can grow as well, but ethanol will be produced (beer and wine production). Yeast cell growth on oxygen and sugar is more efficient than yeast cell growth on sugar with ethanol production. Furthermore, high ethanol concentrations are toxic to the yeast cells. A detailed description of the ethanol production is given in Section 2.1. As ethanol is a carbon source itself the yeast cell can use the ethanol as well for growth (see Figure 1.2). However, the yeast cell has preference of sugars over ethanol. During growth the yeast cell produces carbon-dioxide.

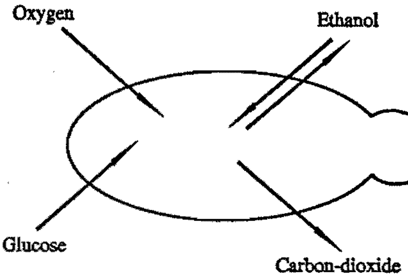


Figure 1.2 *A simple representation of a yeast cell and its main inputs and outputs*

If we look at fermenter level there are numerous yeast cells present in the fermenter. At fermenter level the carbon-dioxide production rate (*CPR*) is a measure for the rate of carbon-dioxide production of all yeast cells. The rate of oxygen utilisation of all yeast cells is the oxygen uptake rate (*OUR*). These two rates can be measured on-line by a mass-spectrometer, see Section 2.6. Together with the ethanol measurement (*E*) the rates give a good description of the condition of the yeast cells, when the yeast cells produce ethanol or not. As the yeast cell does not use any oxygen when it produces ethanol another reliable variable to determine the yeast condition would be the ratio of *CPR* and *OUR*. This ratio is called the respiratory quotient (*RQ*) and is defined as *CPR* divided by *OUR*.

1.1 Motivation

This research done on *Saccharomyces cerevisiae* is motivated by the fact that in biotechnology the role the control engineer plays is still underestimated. Many researchers have spent time and effort in dealing with various aspects of the combination of control theory and biotechnology, e.g. Axellson, 1989; Flaus, 1990; Gee, 1990; Ramseier, 1991; Chen, 1992; van Impe, 1993.

The production of *Saccharomyces cerevisiae* can be influenced at several levels, see Figure 1.3. We will explain Figure 1.3 and give at every level a brief example. At the cell level, or microorganism level, the molecular biologists and physiologist are the ones that motivate the production. The yeast strain will be chosen and may even be genetically modified. Next, at the process level, control and biotechnical engineers try to regulate the process. For example, the control engineer will try to maintain favourable conditions for the yeast cells and the biotechnical engineer tries to optimize the feed composition. Then the product is governed by the customer and the possible applications of the product. Yeast for baking bread needs to produce a lot of gas (CO_2) during the first minutes when the bread is in the oven, otherwise the bread will not rise. Finally, at plant level, down stream processing and scheduling are the most important control items. In the large production vessel (e.g. 1600 m^3) the broth will not be mixed homogeneous causing all sorts of effects. These effects have to be controlled All these levels interact with each other, constrain each other. The place of the control engineer is now clearly defined, but is restricted by the limitations that are set at microorganism and product level.

Man has actually been able to produce alcoholic beverages for some ten thousand years without measuring anything quantitatively, only qualitative determinations were done. It can be said that based on a few simple measurements the fermentation process can be controlled. So, is there a need to control any of the levels depicted in Figure 1.3? The answer to the question is 'yes'.

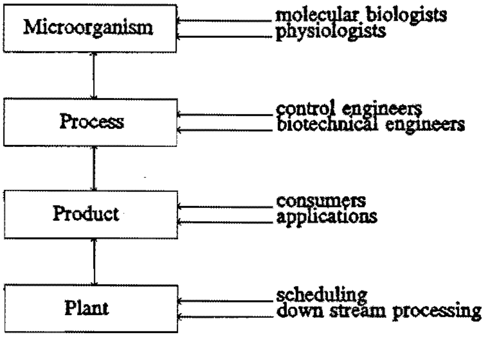


Figure 1.3 *Different levels for controlling Saccharomyces cerevisiae*

In this thesis, we try to control the process such that it can be run efficiently and smoothly at process level (see Figure 1.3). We will also try to regulate the process such that the influence on the microorganism and product level can be enhanced. We expect the control engineer to develop control systems which in future enable us to regulate the enzyme activity (Alverez and Meraz, 1993) or even the intracellular events, see Example 1.1.

Example 1.1 *Control of intracellular events; a future approach*

It is not simple and perhaps not necessary to measure directly reactions inside the cell during a fermentation. A way of indirect measurement is the use of a model of the process or event, and to use this model to control the events inside the cell. By combining the information collected by microbiologists, biochemists, and molecular biologists, models can be constructed which reveal the important reaction, the trigger mechanisms in the cell. When the mechanisms that regulate events in the cell are really understood it should thus be possible to influence the intracellular events by relatively simple measurements.

For our research only the overall effects of the culture could be measured on-line, i.e. oxygen uptake, carbon dioxide production, dissolved oxygen, temperature, and pH. By sampling and off-line analysis, some compounds in the culture medium could be measured, such as biomass and glucose. With the aid of modern (software) sensors non measurable states (the specific growth rate and its metabolic parts) were reconstructed from the on-line measurements.

From the analysis of (the metabolic parts of) the specific growth rate it is postulated that trehalose degradation can be induced by a 'high' extracellular glucose concentration and at certain specific growth rates (Keulers *et al.*, 1993). The specific growth rate can be controlled at a set-point by manipulating the glucose flow. In this case the trehalose storage/degradation, an intracellular event, could be regulated by manipulating both the specific growth rate and the extracellular glucose concentration by glucose flow control. ■

With Example 1.1 we try to elucidate the advanced possibilities control can offer to biotechnology, not only at process level, but also at the microorganism and product level. But, firstly we have to control the process at process level in order to be able to advance to the other levels. In this thesis we will deal with control of the process level only.

Yeast cells are grown preferably without ethanol production as this is the most efficient way. A way to control the growth of the yeast cells such that no ethanol production occurs is to keep the RQ, as defined above, at a predefined value. A severe drawback of this way of controlling the growth of the yeast cells is that one can not regulate the specific growth rate (growth on a carbon source), it is fixed at a certain, unknown value. The specific growth is unknown as the RQ remains constant for a wide and interesting range of the specific growth rate, i.e. growth on oxygen without production of ethanol. In the region of growth without ethanol production certain growth profiles need to be followed in order to control the quality of the yeast cells. An obvious solution to the problem is the regulation of the specific growth rate itself instead of keeping RQ constant.

The 'normal' way of regulating the specific growth rate is carried out by open loop dosage schemes. The dosage schemes are precalculated nutrient additions during a production. The additions are based on experience and theoretical knowledge. To obtain more reproducible production and to account for variations in the sugar feed and theoretical values, e.g. the yield, a feedback strategy seems most appropriate. In the feedback strategy the specific growth rate is controlled directly. However, the specific growth rate can not be measured as no industrial sensor is available. To overcome this problem appropriate identification and control based techniques are used; see next Chapters.

1.2 Problem Statement

In order to be able to control a process the first task of a control engineer is to maintain favourable process conditions. Fermentation processes, especially in industry, require efficient control systems because of material costs and tight product specifications. Any change in operational conditions can drastically reduce product quality and productivity; it could even result in complete failure of the overall process. We will discuss some of the requirements to be met; firstly the biotechnological requirements and secondly the industrial production requirements will be given. The requirements for industrial production are not all applicable during the laboratory experiments, so only relevant terms will be selected and investigated.

Biotechnological Requirements

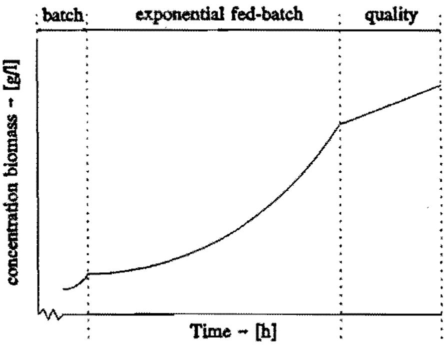


Figure 1.4 *The concentration of biomass as a function of time during the three phases of the laboratory process*

Yeast can be produced in various ways, subject to its purpose. In this thesis the focus is on fed-batch bakers' yeast (*Saccharomyces cerevisiae*) production. The industrial production is 'translated' to a laboratory process which has three distinct phases (see Figure 1.4): a *batch phase* (*seed stage* in production terms) in which the start volume of the yeast is prepared, a *fed-batch phase* in which production of the yeast is the main item, and a *quality phase* where the yeast is cultivated according to certain quality constraints. The last two phases form the *propagation stage* in industrial production. The 'translation' from industrial production phases to laboratory process phases has been done because of the reproducibility (batch phase) of the experiments and because the fed-batch and quality phase need different control objectives.

Note that little is known quantitatively about the quality phase. Only a few research groups have dealt with it so far. Dairaku *et al.*, (1982) describe that the characteristics of an industrially produced cell depend on its life cycle, and the average values of the distributed quantities of these cells are important for industrial use. They state that the age distribution of microorganisms has an essential relationship to the quality of bakers' yeast, e.g. the ability of fermentation in the making of bread.

Industrial Production

In industrial processes, at plant level, not only the product itself is an important item to control but also the energy consumption, the production time, the consumption of nutrients, the labour, and the use of the plant. They all effect the economic performance of the process. One can express this in a profit equation, e.g.:

$$Profit = \frac{P_s \cdot [X(T_f) - X(T_0)] \cdot [1 - Q_{crit}] - P_c}{T_f - T_0} \tag{1.1}$$

where

P_s = price per unit product
X = biomass
Q_{crit} = quality criterion
P_c = costs of product components
T_f = final time
T_0 = begin time.

The costs of product components, P_c, consists of raw material (80%), energy costs (9%), capital costs (4%), daily costs (5%), and others (2%). Saving on nutrients (raw material) and energy could bring a large profit return. This also holds for cutting down on the production time, $T_f - T_0$. The quality criterion is a bit misleading as normally a higher quality would lead to a higher price and thus more profit. However, this increase in profit is already incorporated in P_s. The quality criterion takes into account the additional effort (more difficult feeding schemes, more products out of range, special precautions, etc.), not covered by the other parameters. However, a higher quality normally requires greater production cost. Any gain in profit is incorporated in the price.

Considering the bakers' yeast laboratory process, the main energy consumption is by the cooling/heating section of the fermenter and by the stirrer motor which mixes the broth so that homogeneous reactions take place. Glucose or molasses are the main energy source for the growth of the yeast cells. Some of these industrial requirements can be 'translated' into biotechnical requirements for the laboratory process as defined in the previous subsection. The biotechnical requirement for the different laboratory process phases can be formulated as:

batch phase	*At the end of the batch phase a prespecified amount of yeast has to be produced.*
fed-batch phase	*Complete the fed-batch phase as soon as possible using as little energy as possible to produce a prespecified amount of yeast. Beside this, use as little nutrients as possible and, if applicable, follow any prespecified growth pattern.*
quality phase	*Complete the quality phase obeying prespecified growth patterns and user defined limitations on nitrogen, oxygen and nutrients.*

Problem Statement

After translating the biotechnical requirements into control 'language', the control engineer is confronted with a *non-linear*, *time-varying*, sometimes *poorly understood* process. Nevertheless, a control strategy has to be developed in order to satisfy all the above requirements. The general lack of reliable sensors is another complicating factor, so many key variables, e.g. biomass concentration, can not be measured on-line.

Depending on the type of product, one is interested in either optimizing the fermenter's productivity (fed-batch phase), or in controlling a variable, e.g. overall growth rate, to meet certain product specifications (quality phase). There are many ways to accomplish these objectives, e.g. quality control by means of the ratio of budding cells (Dairaku *et al.*, 1982). For our laboratory process a suitable control mechanism is required, based on linear, non-linear, or other modelling techniques, capable of regulating both fed-batch as well as quality phase. This leads to the following problem statement:

Problem Statement

Given a yeast culture, Saccharomyces cerevisiae SU32, grown at laboratory scale, under optimal (favourable) process conditions, develop a controller that can regulate the specific growth rate of the culture according to any prespecified pattern.

An evaluation of different methods, their applicability, limitations, and effectiveness to attain the above goal is demonstrated in the next Chapters.

1.3 Outline of this Thesis

This thesis consists of three parts: the laboratory process and its *simulation models* (used to test identification and control algorithms), identification methods for *control based models* (used to control the process), and *observers* (used to substitute for not (yet) existing or available sensors), and specific growth rate control (see Figure 1.5).

In Chapter 2 some relevant background in yeast biology is given together with a characterisation of the equipment used. The focus is on relevant aspects for the understanding of the cell and on limitations of the process. This knowledge is expanded in Chapter 3 with both a theoretical and an experimental analysis of the process. This is done in order to be able to develop the simulation models presented in Chapter 3. These simulation models will be used to test various identification and control strategies.

The novelty of the simulation models presented in Chapter 3 is the use of both experimental batch and fed-batch data for parameter estimation. Most models presented in literature used data from shake flask or continuous culture experiments to estimate most of their parameters. Another novelty is the use of validation data to check the proposed simulation model

The next three Chapters will focus on structure identification and parameter estimation of the model process. For those readers not familiar with identification and parameter estimation, Beck and Young (1988) wrote a good introduction to 'system identification, parameter, and state estimation' with a pictorial presentation of the most important terms used.

In Chapter 4, a linear dynamic model of the oxidative metabolism of yeast will be identified. This linear model is derived by first compensating one of the inputs of the process with a 'non-linear compensation'; the resulting process can be considered linear. This model can be used as a simulation and prediction model in a controller. As the process is highly non-linear, some non-linear identification and structure determination algorithms are evaluated on the process, as described in Chapter 5. The identification algorithms are tested on both simulation models and experimental data. Due to the lack of appropriate sensors one of the important process variable can not be measured on-line: the growth rate. This process variable can, however, be observed and estimated using on-line exhaust-gas measurements as is presented in Chapter 6. Beside the specific growth rate and its major metabolic components, the biomass concentration is estimated.

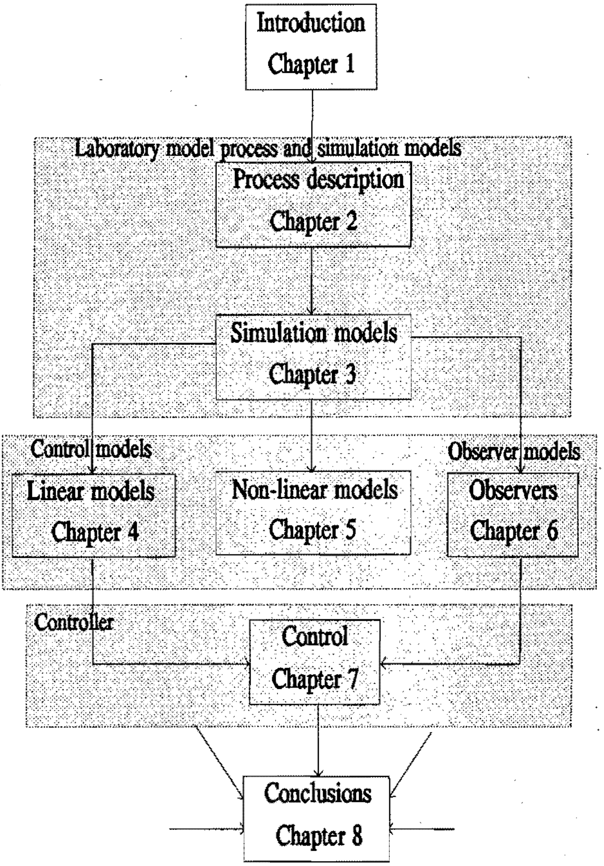

Introduction
Chapter 1
Laboratory model process and simulation models
Process description
Chapter 2
Simulation models
Chapter 3
Control models
Observer models
Linear models
Chapter 4
Non-linear models
Chapter 5
Observers
Chapter 6
Controller
Control
Chapter 7
Conclusions
Chapter 8

Figure 1.5 *Outline of the thesis by Chapters*

Although there have been some papers and reports dealing with linear identification of the culture of *Saccharomyces cerevisiae*, none of those works involved oxidative growth on glucose only, combined with a non-linear compensation. Furthermore, most of the models obtained were ARMAX models with a one-step ahead prediction whereas the models developed here are output error models having an extended horizon. Concerning the non-linear structure identification algorithms some applications of Volterra kernels and Radial Basis Function networks are known. No applications of the non-linear model linear in-the-parameters have been reported. The presented observer is unique as it can estimate all metabolic parts of the specific growth rate and the biomass concentration from on-line measurements. The biggest drawback is that the proposed observer is not generic, i.e. the observer can only work for bakers' yeast type of cultures after adaptation of the culture specific stoichiometry.

Chapter 7 covers specific growth rate control of the fed-batch process. The control scheme chosen is discussed. It incorporates the non-linear compensation used in Chapter 4 and a PI block. Furthermore experimental limitations are discussed and some recommendations are presented to overcome these limitations. Finally controller results are presented. The presented structure for the specific growth rate control is unique: an observer, feed-forward controller, and a PI-controller block. Due to the 'high' set-point control of the specific growth rate unknown dynamic effects could be measured. These dynamic effects on the specific growth rate have not yet been reported or investigated. We try to control the dynamic effects by adding an additional dynamic compensation in the controller scheme.

Finally, conclusions and suggestions for further research are given in Chapter 8. The main conclusion is that it is possible to control the specific growth rate of *Saccharomyces cerevisiae* at set-point or along a pattern of small set-point changes.

Models

In order to avoid any confusion about models, their names and application a diagram of the models used is presented in Figure 1.6.

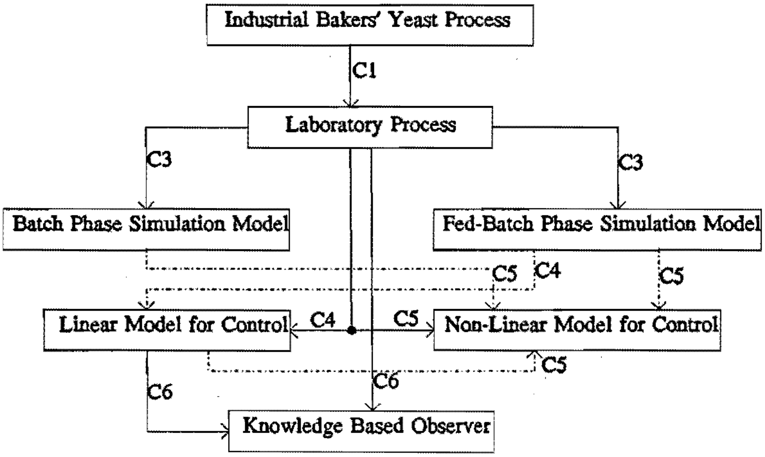

Industrial Bakers' Yeast Process
C1
Laboratory Process
C3
C3
Batch Phase Simulation Model
Fed-Batch Phase Simulation Model
C5
C4
C5
Linear Model for Control
C4
C5
Non-Linear Model for Control
C5
C6
C6
Knowledge Based Observer

Figure 1.6 *A schematic overview of the models used in this thesis, the solid lines between the boxes represent models based on experimental data, the dashed lines between the boxes represent models based on simulation data*

In this Chapter (C1) the industrial bakers yeast process is modelled as laboratory process with three different stages, as explained before in one of the previous sections. Of two of these phases, the batch phase and the fed-batch phase, a simulation model will be made. This will be carried out in Chapter 3 (C3). The simulation models are used to test identification and control models. This is done in Chapter 4 and 5, (C4, C5). These identification and control models are thus models of the simulation models of Chapter 3. The models of simulation models are dashed lines in Figure 1.6, whereas the solid line represent models made using experimental data. As can be seen in Figure 1.6, beside the models of simulation models also experimental data form the laboratory model process is used to make some identification and control models (C4, C5). The last model that is made is the observer model which is constructed from the model found in Chapter 4 and the experimental data from the laboratory model process (C6).

References

Alverez, J. and M. Meraz (1993). Control of Enzyme Activity in Stirred Bioreactors. *The Chemical Engineering Journal*, Vol. 51, pp. B11-B16.

Axellson, J.P. (1989). Modelling and Control of Fermentation Processes. *Ph.D. Thesis*, Lund Institute of Technology, Sweden.

Beck, M.B. and P.C. Young (1988). An Introduction to System Identification and Parameter and State Estimation. *Computer Applications in Fermentation Technology: Modelling and Control of Biotechnological Processes*, (Eds. Fish, N.M. *et al.*) London-New York: Elsevier Applied Science, London - New York, pp. 129-158.

Chen, L. (1992). Modelling, Identifiability and Control of Complex Biotechnological Systems. *Ph.D. Thesis*, Université Catholique de Louvain, Belgium.

Dairaku, K., Izumoto, E., Morikawa, H., Shioya, S. and T. Takamatsu (1982) Simulation and Quality Control of Bakers' Yeast Fed-Batch Culture Using Population Models. *Biotechnology and Bioengineering*, Vol. 24, pp. 2661-2666.

Enari, T.M. (1982) Biotechnology and Control Engineering. *IFAC symposium, Modelling and Control of Biotechnological Processes*, Helsinki, Finland, pp. 1-3.

Flaus, J-M. (1990) Estimation de l'État de Bioprocédés à Partir de Measures Indirectes. *Ph.D. Thesis*, L'Institut National Polytechnique de Grenoble, France.

Gee, D.A. (1990). Modelling, Optimal Control, State Estimation, and Parameter Identification Applied to a Batch Fermentation Process. *Ph.D. Thesis*, Ann Arbor University, Michigan, USA.

Keulers, M., Ariaans, L. and M.L.F. Giuseppin (1993). The Dynamic Step Response of a Fed-Batch Bakers' Yeast Process. *Preprints of the 6th European Congress on Biotechnology*, Florence, Italy.

Ramseier, M. (1991). Nonlinear Adaptive Control of Fermentation Processes Utilizing *a priori* Process Knowledge. *Ph.D. Thesis*, University of California, Santa Barbara, USA.

van Impe, J. (1993). Modelling and Optimal Adaptive Control of Biotechnological Processes. *Ph.D. Thesis*, Katholieke Universiteit Leuven, Belgium.

2

Process Description

Some of the process dynamics can only be understood by looking at the heart of the matter, the yeast cell. Hence we will focus on some relevant mechanisms. This elementary focus on the cell is not extensive, but a help to understand some of the choices made later on. In the cell thousands of biochemical reactions take place, integrated into a functioning whole. Some of these reactions interact with the environment as the cell needs materials for growth and maintenance. There is an exchange of compounds between the cell and the environment. The environment can change, not in the least due to the (re)actions of the cell; the yeast cell is capable of accumulating and storing reserve material. The cell essentially has the capability of adapting to new conditions after a certain adjustment phase, also known as the lag phase.

In industrial fermentation (see Section 1.2) the main interest of the plant engineer is the maintaining of a favourable environment for the yeast. These favourable conditions may differ according to the type of microorganism. They may change during the process itself. Changes in environmental conditions should be optimal, i.e. as smoothly as possible. Not only the yeast cell dictates our way of handling the process; sensors, actuators, and the process control system do as well. The process control system not only supplies a favourable ambient condition for the cells by regulating some of the physical process variables, it also influences the dynamics of the process. This interaction can result in extended lag phases, loss of control, or formation of undesired (by-)products. A comprehensive investigation of all process equipment, of the way the yeast is cultivated, and of the technical limitations is necessary in order to define operational boundaries for the process under study.

The yeast cell is excellently described by various authors (e.g. Rose and Harrison, 1971). In Section 2.1, we will focus on the cell itself, the glycolysis (sugar metabolism), and the ethanol formation. Section 2.2 will focus on the laboratory process. The kind of microorganism used and how it is cultured is discussed in Section 2.3. How the fermenter is operated during all phases is discussed in Section 2.4. The fermenter, its actuators and sensors, and the process control system and some off-line cell concentration measurement techniques are given in Section 2.5. Finally the conclusions are drawn in Section 2.6.

2.1 The Yeast Cell

Yeast is a large and complex microorganism that belongs to the class of Fungi. It has a volume a few times larger than a bacteria and (like all higher organisms) it has a cell nucleus. The yeast cell itself has a diameter of 5-8 [μm]. Generally it reproduces by budding and cell division. The yeast cell is a robust organism in several aspects. The cell wall has a rigid structure that constitutes 25% of the cell dry weight. Not only does it protect the cell but it also contains several enzymes.

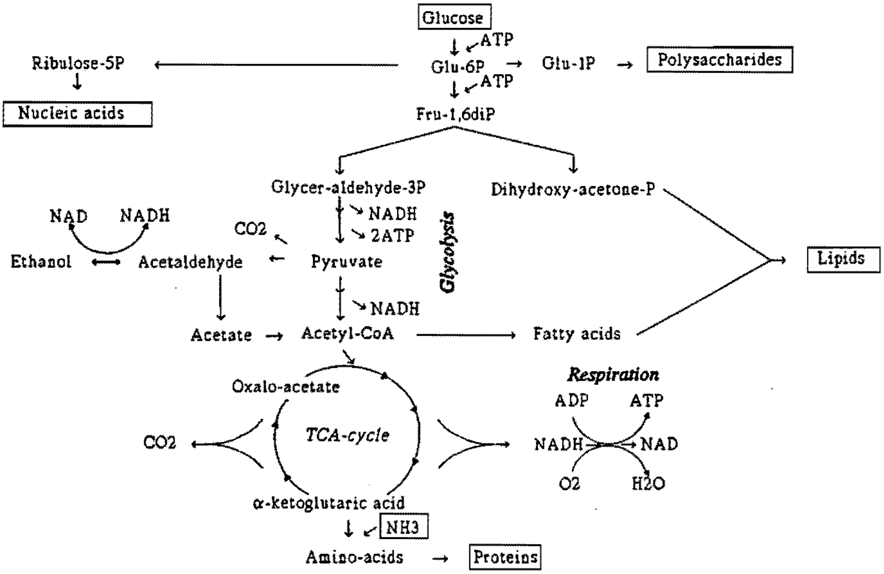

Glucose
ATP
Ribulose-5P
Glu-6P
Glu-1P
Polysaccharides
ATP
Nucleic acids
Fru-1,6diP
Glycer-aldehyde-3P
Dihydroxy-acetone-P
NAD
NADH
NADH
CO2
2ATP
Glycolysis
Ethanol
Acetaldehyde
Pyruvate
Lipids
NADH
Acetate
Acetyl-CoA
Fatty acids
Oxalo-acetate
Respiration
ADP
ATP
CO2
TCA-cycle
NADH
NAD
O2
H2O
α-ketoglutaric acid
NH3
Amino-acids
Proteins

Figure 2.1 *A simplified metabolic map involving the glycolysis of sugar to pyruvate, its connection with the ethanol production of the TriCarboxylic Acid (TCA) cycle and connected respiration (after Enfors et al. 1990)*

The inside of the cell contains a cell nucleus where the DNA is located. There is a number of different functional units, organelles, e.g. up to 50 mitochondria, where an important part of the cell energy metabolism and generation takes place. Other parts of the metabolism take place in the cytosol. Sugars and lipids are stored in storage granules. A large part of the cell interior is made up of vacuoles. These have various functions, e.g. serving as a storage reservoir for nutrients and certain enzymes. There is also a hypothesis that wastes and toxic products accumulate in a vacuole.

Glycolysis

The yeast needs sugar as a carbon source and as an energy source. The sugar metabolism consists of many steps and is spread out over large parts of the cell. There are many internal control mechanisms that integrate the different parts. Figure 2.1 presents a summary of the sugar metabolism in *Saccharomyces cerevisiae.* Glucose enters the glycolysis and is oxidized to pyruvate by co-enzymes (β-Nicotinamide-Adenine Dinucleotide, NAD) containing enzymes. Some Adenoside-5'-TriPhosphate (ATP) is produced in the glycolysis. From this branch point ethanol may be formed under decarboxylation of pyruvate and re-oxidation of reduced co-enzyme (β-Nicotinamide-Adenine Dinucleotide, reduced, NADH). This enables energy yielding oxidation of the glucose in the absence of oxygen. Alternatively, pyruvate is decarboxylated by pyruvate decarboxylase into acetyl-coenzyme A. The acetyl-coenzyme A enters the TCA-cycle in which the carbon is oxidized by different co-enzymes (NAD, Flavine Adenine Dinucleotide) containing enzymes to carbon-dioxide. Comparatively much more energy is obtained in the subsequent re-oxidation of the co-enzymes in the respiration chain, which is coupled to phosphorylation of Adenoside-5'-DiPhosphate (ADP) to ATP. While glycolysis is performed in the cytoplasm, the TCA-cycle and respiration is located in the mitochondria.

Ethanol Formation

The ethanol formation can be seen as an overflow reaction. If the glycolysis in the cytosol produces more pyruvate than the respiratory system and the TCA-cycle in the mitochondria can metabolize, the excess pyruvate is reduced to ethanol. On the other hand, if there is ethanol and too little glucose available to the cell, ethanol can be converted and metabolized by the mitochondria. In case of excess glucose and ethanol in the broth, glucose is metabolized preferably.

Ethanol is converted to acetate in the cytosol and further oxidized in the citric acid cycle and glyoxylate cycle in the mitochondria and the glyoxsomes respectively. The growth on ethanol requires the production of many precursors for synthesis that glycolysis otherwise provides when growth is on glucose. In order to produce these precursors the glycolysis is essentially run backwards. It is referred to as the gluconeogenic pathway.

There are several theories for locating the rate limitation step of this process. For our purpose, it suffices to regard it as a limited respiratory capacity (Sonnleitner and Käppeli, 1986). Under certain circumstances, glucose is stored within the cell in the form of the polysaccharides glycogen and trehalose. At low growth rate, there is a pronounced interaction with this storage. Another storage area is the lipid fraction of the cell that may play a role in glucose metabolism. Altogether these storage units can account for up to 30 % of the cell mass.

During growth, the individual yeast cells change their metabolism depending on their cell cycle stage. Two main phases can be distinguished, the single cell phase and the budding phase. The single cell phase is characterized by respiration and volumetric growth at a rate depending on the substrate concentration. The budding phase has a constant time length, the metabolism is fast, ethanol is produced and excreted from the cell (von Meyenburg, 1969).

Under anaerobic conditions a yeast cell culture partially metabolizes glucose to ethanol under various growth conditions. During this process the cell extracts energy. Such microbiological reactions where energy is extracted without consumption of oxygen are called fermentations. The yeast cell has a large capacity to break down sugar to ethanol. However, this way the cell gets only about 5-10% of the energy available in sugar.

At high sugar concentrations the yeast cell growth rate increases far beyond the capacity of complete oxidation of glucose, and large amounts of ethanol are produced despite aerobic conditions (Crabtree effect).

Some of this knowledge about the yeast cell will be used in the next Chapters on modelling and identification. The knowledge will either be used to model certain events (simulation model) or to limit the purpose of a model (control model). The macroscopic behaviour of the culture will be addressed in Chapter 3, as this Chapter deals with the simulation model. The next Section will continue with a brief discussion on the laboratory process.

2.2 The Laboratory process

The laboratory process is defined in Section 1.2 as a sequence of batch, fed-batch, and quality phases each with their own requirements. Together they represent the *seed* and *propagation* stages of an industrial process. Due to the distinct definitions, each phase of the laboratory process can be seen as a separate process. A detailed description of the requirements of each phase is given in Section 2.5, the experimental site.

This laboratory process is carried out in an aerated, continuously stirred fermenter. This 10 litre fermenter (see Figure 2.2) is the core of the fermentation system. The glasswall vessel is fitted with an air-tight rubber ring on both bottom and top. The top plate contains a number of ports that are used to support a pH probe, a dissolved oxygen probe, an anti-foam sensor, an outlet air-cooling device, and several ports to add or withdraw chemicals. The bottom plate contains ports for a temperature probe, a heating and a cooling element, a combined air and feed inlet, a sample needle and a waste outlet. Agitation is provided by a magnetically-coupled drive motor with adjustable speed. An impeller and four baffles provide good mixing. Air is blown into the fermenter through a small pipe ending just below the impeller, providing instant mixing of the air. The fermenter is held sterile during the fermentation in order to avoid contamination.

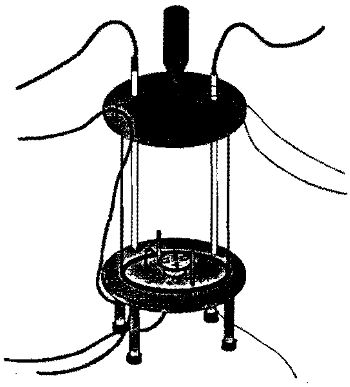


Figure 2.2 *The fermenter*

The laboratory process is performed at favourable environmental conditions, i.e. temperature at 30 [°C] and pH at 5. Further, by adding anti-foam agents, foam production in the head space of the fermenter is minimized. Depending on the phase of the laboratory process, the nutrients are added either beforehand or during the process. Again care is taken that this carried out in a sterile manner. During the processes, on-line measurements are made of the pH, temperature, dissolved oxygen tension, and the effluent gas composition, i.e. oxygen, carbon-dioxide, and ethanol are taken. A more detailed description of these sensors can be found in Section 2.5 and Table 2.8.

The laboratory process is a good representation of the *seed* and *propagation* stages of an industrial process.

2.3 Microorganisms and Culturing Techniques

The microorganism used in the experiments in this study is bakers' yeast, *Saccharomyces cerevisiae* (strain SU 32).

Culture Maintenance

The strain is kept in a refrigerator at -80 [°C]. Inocula are prepared by transferring a colony to 2 shake-flasks each containing 0.2 [dm^3] YPD (= Yeast extract, Pepton, and Dextrose). The culture is kept in the shaker incubator at 30 [°C] and 250-300 [rpm]. After developing for 24 [h] the culture can be used as inoculum.

Table 2.1 *Medium composition for batch and fed-batch*

Component	Batch	Fed-batch	Units
$C_6H_{12}O_6.H_2O$	33	440	[g.dm^{-3}]
$(NH_4)_2SO_4$	2	10	[g.dm^{-3}]
KH_2PO_4	2	4	[g.dm^{-3}]
K_2HPO_4	2	4	[g.dm^{-3}]
NaCl	0.5	-	[g.dm^{-3}]
$CaCl_2$	0.2	0.8(2)	[g.dm^{-3}]
$MgSO_4$. $7H_2O$	0.6	6	[g.dm^{-3}]
Trace-Metals Solution (Egli)	10	16	[cm^3.dm^{-3}]
Vitamin Cocktail (Egli)	1(1)	4(1)	[cm^3.dm^{-3}]

(1) *Filter sterilization.*

(2) *Separate sterilization.*

The Synthetic Medium

For both batch and fed-batch experiments we use a so called synthetic medium. This is an aqueous solution of a mixture of salts and organic compounds with glucose as the main energy and carbon source. When the yeast grows on this medium, growth is carbon limited. The exact medium compositions for both batch and fed-batch are listed in Table 2.1. The medium is prepared freshly each time it is needed. It is sterilized at 120 [°C] for 45 and 20 [min] for the batch and fed-batch respectively. The batch medium is sterilized inside the fermenter. Vitamins are sterilized separately by a filter, so is the $CaCl_2$ of the fed-batch medium. Both media are brought to pH 3 before sterilization. The weight of broth is 4000 [g]. The concentration of the glucose in the fed-batch medium is 440 [$g.dm^{-3}$] or 2.47 [$mole.dm^{-3}$].

For every experiment two new inocula should be prepared from the freezer (-80 [°C]) as practice showed that keeping a strain in the refrigerator (+5 [°C]) for more than a week could change the reaction of the culture. In this way the quality of the yeast at the beginning of a fed-batch experiment is constant.

2.4 Fermenter Operation

One day before start-up, two inocula are prepared and placed in the shaker incubator. The pH-probe is calibrated with buffer-solutions of pH 4.01 and 7.00 before sterilization. On the day of start-up the fermenter, including all tanks, the pH probe, the Dissolve Oxygen Tension (DOT)-probe, the batch medium, the anti-foam, and all tubes are sterilized at 120 [°C] for 45 [min]. The fermenter is then allowed to settle at the set temperature with agitation and air flow turned on. During the settling of the fermenter the fed-batch medium and the $CaCl_2$ are sterilized for 20 [min]. After the settling of the fermenter the DOT-probe is calibrated at pH 5 and 30 [°C].

Table 2.2 *Operating conditions during batch and fed-batch phase of important variables*

Parameter	Batch	Fed-batch	units
Glucose feed	-	0-200	[$cm^3.h^{-1}$]
Cell mass	1-10.5	10.5-70	[$g.dm^{-3}$]
Glucose	0-33	<0.200	[$g.dm^{-3}$]
Ethanol	0-15	0-1	[$g.dm^{-3}$]
Agitation	750	400-1250	[rpm]
Air flow	3	3-8	[$dm^3.min^{-1}$]
Temperature	30	30	[°C]
pH	5.0-5.5	5.0	[-]
Liquid volume	4.4-4.45	4.45-7	[dm^3]
Anti-foaming agent	10	40	[cm^3]

Next a computer program is started to control the pH, temperature, anti-foam, and to log data. At this point the system is ready for inoculation. To ensure its purity, the inoculum is checked with a microscope. The total amount, 400 [cm^3], is carefully transferred to the fermenter, minimizing the risk of contamination by heating the air around the inoculum port. Since the initial amount of yeast cells is very low, the yeast grows first on the batch medium. After about 16 [h] the batch phase ends with a cellmass concentration of about 10.5 [g.dm^{-3}]. All glucose and ethanol present during the batch phase have been consumed at this stage. The end of the batch stage is detected by total consumption of the ethanol and the *DOT*-level rising back to 90% (see Figure 3.1). Table 2.2 lists the main conditions as controlled or set during both batch and fed-batch phase. The fermenter is now ready for further experiments.

Before starting an experiment the glucose pump is calibrated. The pump tubing is renewed before every experiment. During an experiment liquid samples are taken automatically for off-line analysis. The mass-spectrometer is calibrated twice a day. Normally an experiment can not last for more than about 24 [h] because, due to the high cell concentrations, the viscosity of the medium starts to play an important role and the oxygen concentration in the broth becomes limiting (less than 10%). If the growth rate is low, around 0.1 [h^{-1}], the experiment can last for more than 24 [h]. The fermenter is switched to low growth rates during the night.

2.5 The Experimental Site

A complete schematic overview of the experimental site of the laboratory process is given in Figure 2.4. Below each function is described, such as actuators, sensors, and the data acquisition system and its rates. This includes the time or acquisition protocol, a study of the sensors and actuators and finally a description of the optical density and dry-weight measurements.

Monitoring

The experimental site has a number of independent measurement devices. Between each of these devices and the μVAX-computer is a communication link with an acquisition time protocol. The μVAX can dictate the acquisition time protocol to the Applikon unit. However the mass-spectrometer is a master to the μVAX and thus dictates the acquisition time protocol to the μVAX. Inside the μVAX the CDAS data logging system is master of all results stored in the data base regardless of which is slave or master. At certain time intervals, with a minimum of 1 [min], the data is logged. Note that the logging time has no link

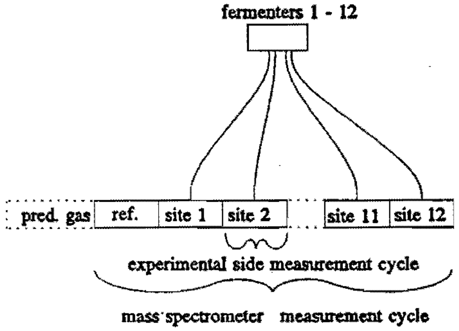

fermenters 1 - 12
pred. gas
ref.
site 1
site 2
site 11
site 12
experimental side measurement cycle
mass spectrometer measurement cycle

Figure 2.3 *The mass-spectrometer measurement cycle*

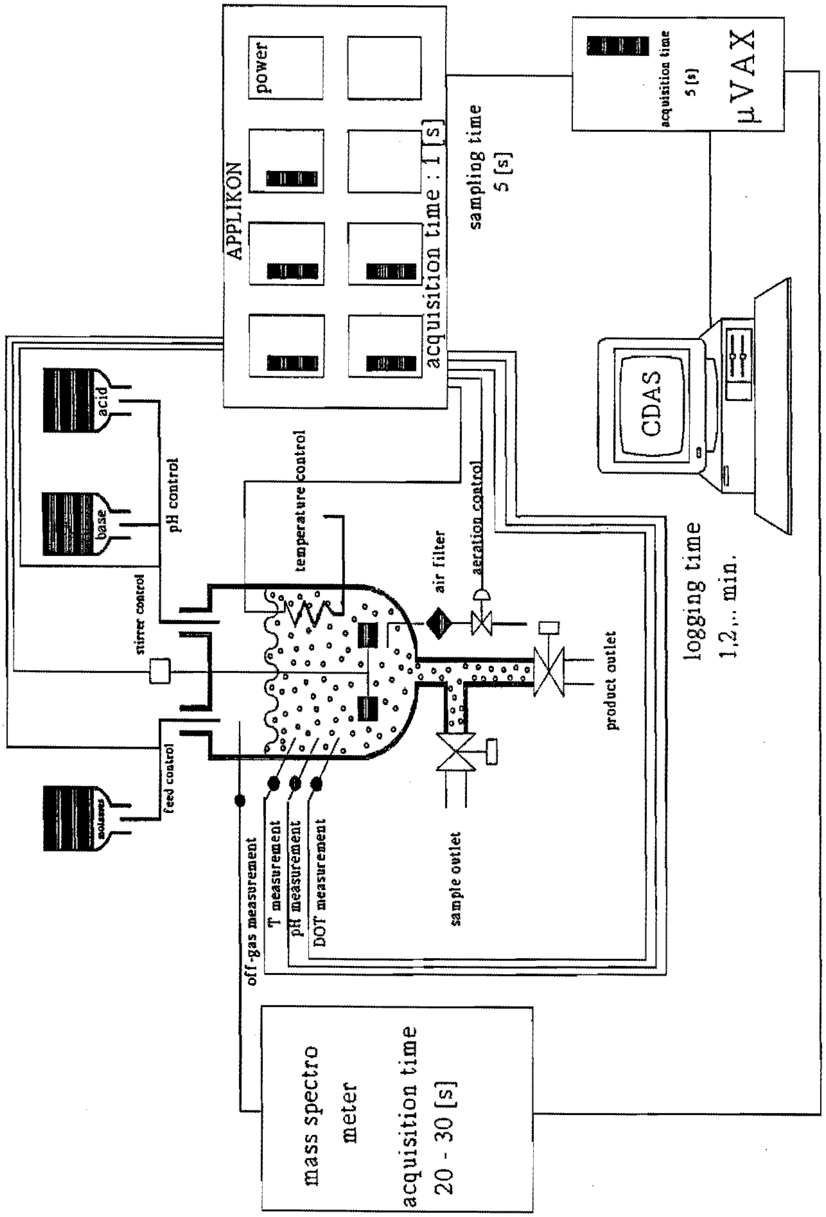

molasses
feed control
stirrer control
base
acid
pH control
APPLIKON
power
acquisition time : 1 [s]
off-gas measurement
T measurement
pH measurement
DOT measurement
temperature control
air filter
aeration control
sample outlet
product outlet
mass spectro
meter
acquisition time
20 - 30 [s]
sampling time
5 [s]
acquisition time
5 [s]
µVAX
CDAS
logging time
1,2,.. min.
sampling time 30 [s] - 6 [min]

Figure 2.4 *Schematic overview of the experimental site with the fermenter, the CDAS unit, the µVAX, and the mass-spectrometer*

with the actual acquisition time. The acquisition time of the Applikon unit is 1 [s] (for balance, T, pH, *DOT*, *Stirrer Speed (SS)*, and *Air Flow (AF)*). These values can be sampled every 5 [s] by the μVAX. With dedicated software this high sample rate can be logged within CDAS directly to disk.

The mass spectrometer has a measurement cycle depicted in Figure 2.3. Hourly a predefined gas (pred. gas) is inserted and a calibration is performed. This is done before the normal reference gas (ref.). Every measurement cycle takes about 3 [min]. To speed up this cycle the experimental sites that are not used are skipped, retaining a shorter cycle with a minimum of a calibration gas and one experimental site measurement. This minimal cycle takes about 30 [s]. This is thus the minimal sample frequency of the mass spectrometer. Note that CDAS has a minimal sample frequency of 1 [min]. Thus in order to use the 30 [s] sample frequency more than one channel has to measure the outlet gas of one experimental site. Later on the results have to be combined.

Sensors and Actuators

The main purpose of this subsection is the analysis of the actuators and sensors. The analysis should be carried out accurately to be able to determine the limits of identification and control. The characteristics analyzed are: the standard deviation (σ) of the output signals, the bandwidth of the output signals, the signal to noise ratio (S/N) of the output signals, the delay time of the sensors/actuators, and the accuracy of the logged data within the Computer and Data Acquisition Software (CDAS)-data base. Table 2.8 (page 28 and 29) lists the available sensors, actuators and their specifications. Next, some more detailed comments will be given on the primary variables, the process variables, and the input variables.

Primary Variables

There are three primary variables to be controlled: temperature, pH and foam. The primary variables are controlled by local controllers. The local controllers are situated in the Applikon unit and have a sampling frequency of 1 [Hz]. The control inputs of the primary variables are listed below together with some requirements.

The temperature should be constant (= 30 [°C]) during all three phases of the laboratory process. The maximum deviation from the required temperature is 0.1 [°C]. The operation region for the temperature is from 29 to 31 [°C]. The temperature is kept at the set-point via heating and cooling water valves, a PT100 sensor, and a PID-controller. Note that during both fed-batch and quality phase no hot water is used as the system only needs cooling (exothermical process). The primary controller has to be adjusted because of this change.

The pH should be constant (= 5) during batch, fed-batch and quality phase. The maximum deviation from the pH is 0.05. The pH is kept at the set-point via acid and base roller pumps, an Ingold pH-electrode, and a PID-controller. The operation region for the pH ranges from 4 to 6 for all phases. The pH measurement is temperature dependent (see Table 2.4), but this dependence will be neglected because it is smaller than the logging accuracy within the maximum deviation from the set-point. Beside this dependence the measurement is also anti-foam dependent. However the dependence is even smaller than the temperature dependence, thus it can also be neglected.

Table 2.3 *Set-point, maximum deviation from the set-point and operation region of temperature and pH for the laboratory process*

	Temperature [°C]	pH [-]
Set-point (maximum deviation)	30 (0.1)	5 (0.05)
Operation region	29 - 31	4 - 6

The foam production can occur during all three phases and should be minimized. Foam can be detected via a conductivity measurement and regulated with a roller pump adding anti-foaming agent. This feature has not been used during the experiments as a sufficient anti-foaming agent is added at the beginning of both the batch and fed-batch phase.

Table 2.4 *Temperature, pH and anti-foaming agent dependence of the pH and DOT-measurement*

	Temperature	Anti-foaming agent
DOT	1.9 % /°C	-
pH	-0.02 pH /°C	-0.0016 pH / cm^3

Process Variables

The process variables can not be controlled directly as most of them are not measurable. Adequate control of the process variables on the basis of the input variables has yet to be designed.

Both biomass and glucose concentration can not be measured on-line. An off-line measurement test is performed for the biomass. The glucose concentration is not measured off-line as these measurements normally are not done during the industrial scale production. For more detailed information on the off-line biomass measurements see the subsection on dry-weight analysis.

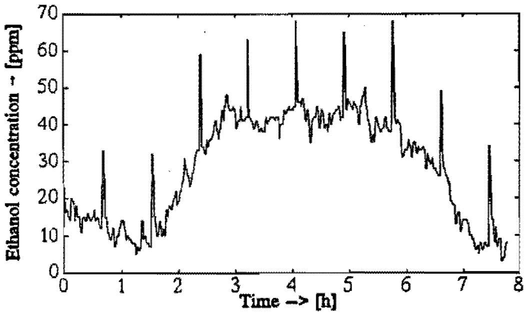


Figure 2.5 *Ethanol measurement disturbed with hourly inserted reference gas*

The ethanol concentration in the effluent-gas is measured by the mass-spectrometer. The ethanol measurements with the mass-spectrometer can be performed in two ways: with the Faraday detector (not very sensitive but stable) or the electron multiplier (1000 times more sensitive than the Faraday detector but not stable). The ethanol measurements are influenced by other users, if their processes produce ethanol as some of the ethanol remains in the measurement device. A reference gas that is injected automatically every hour, does influence the measurements tremendously (Figure 2.5). Further, due to daily calibration, trends are visible in the ethanol measurements. These trends are very small (about 10 [ppm]) and can be neglected.

The *DOT* is measured inside the broth via an Ingold DOT-probe. The measurement shows a temperature dependence that is rather strong (see Table 2.4). This dependence has a first order transfer function which can not be measured because of the relatively low sampling rate. It is assumed that the temperature changes are very small and very smooth thus that this effect can be neglected. Note that the *DOT* measurement is not pH or anti-foam dependent.

Table 2.5 *Set-points, maximum deviation from the set-points, settling times for the set-points and operation regions of the outputs for the three phases of the laboratory process*

	DOT [%]	*Ethanol* [ppm]	*RQ* [-]
physical limitations	0 - 100	0 - 6000	0 - 20
Set-points (maximum deviation)			
batch	15 (1.5)	(3)	(3)
fed-batch	15 (1.5)	0 (< 100)	1.08 (0.05)
quality	15 (1.5)	0 (< 100)	(1)
Settling time of new set-point values			
quality	(2)	(2)	15 min
Operation regions			
batch	10 - 100	< 3000	(3)
fed-batch	10 - 30	< 200	1.00 - 1.15
quality	10 - 30	< 100	(1,3)

(1) depends on the requirements set by the operator.
(2) no constraints.
(3) may have any physical value.

The *OUR* and *CPR* are calculated from the effluent-gas analysis, the $O_{2,out}(t)$ and $CO_{2,out}(t)$ mass-spectrometer measurements. These measurements can be disturbed by changes in the stirrer speed and air flow (new equilibrium of the fluid-gas transfer in the fermenter). Due to malfunctioning of the mass-spectrometer or to human interaction, spikes can occur in the mass-spectrometer measurement.

For most process variables there are some control objectives during the different phases of the process. Of interest is in what way these variables have to be controlled, what are the set-points, what is the permitted deviation from these set-points and what are the operating regions. For every variable a short description of these requirements will be given and the requirements are listed in Table 2.5.

The *DOT* should be constant (e.g. 15%) during all three phases. During the batch phase a higher amount is permitted if the lower limits of both stirrer speed and air flow are used. The lower limit is chosen in order to reduce energy consumption to a minimum. Note that during an uncontrolled batch phase the changes in the *DOT* are very slow and smooth and the control of the *DOT* can be very simple (e.g. PID). During the next two phases the control of the *DOT* has to be quicker in order to minimize the error due to changes in the glucose flow.

Unlimited ethanol production is permitted during the batch phase. But during both fed-batch and quality phase no ethanol may be produced. Due to the fact that the accuracy of the mass-spectrometer is 100 [ppm], the operation region and the maximum deviation from the set-point for both the fed-batch and quality phase are not set to zero but to 100 or 200 [ppm]. Within this region one may assume no ethanol is produced.

No limitation is set on the *RQ* during the batch phase. During the exponential fed-batch phase the *RQ* is set to 1.08 and kept at this value. Deviations from this set-point are allowed as long as optimal growth is achieved and no ethanol is produced. During the quality phase all kinds of trajectories are possible. The deviation level from these trajectories is normally 1% (which is extremely low). The operating region for the batch phase ranges from 0.4 to 10. For both the fed-batch and quality phase it ranges from 1.00 to 1.15.

Input Variables

The input variables in Table 2.8 (actuators) marked with an (*) have primary controllers to minimize the set-point error. The base pump has no primary controller, the flow is measured through the number of rotations the pump makes. Calibration for different tubes is necessary and can be done by means of a balance. The air flow is measured by a flow meter that corrects the flow if problems should occur. The nutrient flow is measured by a balance (decrease of weight). These weight changes are fed into the controller to give the desired flow. If one wants to calculate the flow from off-line data, the data of a balance measuring the amount of glucose feed left have to be used. The stirrer speed has no primary controller.

Input manipulations are necessary in order to keep the outputs at their set-points or to keep track of the trajectories. Of interest are the physical limitations of the inputs, the operation regions, and the maximum rate of change allowed for these inputs. These items will be discussed for each input separately and summarized in Table 2.6.

The airflow is kept low (3.0 [$dm^3.min^{-1}$]) during the batch phase in order to ensure minimum energy consumption. During the fed-batch phase an increase (up to

8.0 [$dm^3.min^{-1}$]) of the air flow is necessary in order to preserve enough oxygen in the broth. During the quality phase the air flow is set to maximum in order to ensure quick regulation of the *DOT*.

The stirrer set-point is kept low, 300 [rpm], during the batch phase. If necessary it is increased, to ensure a *DOT*-set-point. During the exponential fed-batch and quality phase regulation of the *SS* is necessary in order to ensure the *DOT*-set-point level.

Table 2.6 *Physical limitations, operation regions and maximum rate of change for the three phases of the laboratory process*

	Air flow [$dm^3.min^{-1}$]	Stirrer speed [rpm]	Glucose flow [$dm^3.min^{-1}$]	NH_4OH [$dm^3.min^{-1}$]
physical limitations	0 - 10	0 - 1500	(3)	(3)
		Operation regions		
batch	3	300	0	(1)
fed-batch	3 - 8	400 - 1000	(1)	(1)
quality	3 - 8	500 - 1000	(2)	0
		Maximum rate of change		
batch	0	0	0	(1)
fed-batch	0.1	100	(1)	(1)
quality	0	100	(2)	0

(1) depends on the roller pump, tube and desired end concentration of the biomass
(2) depends on desired RQ trajectory
(3) depends on the roller pump and tube

The glucose/molasses[1] flow is zero during the batch phase and increases exponentially during the fed-batch phase. The maximum deviation has to be calculated each time a certain roller pump, tube and the desired end concentration of the biomass is required. The same holds for the operating region. No information is available about the quality phase. The physical limitations depend on the tube and roller pump used.

The NH_4OH dosing during the fed-batch phase and quality phase can be subject to limited nitrogen control.

---

[1] Molasses are not used in these experiments because both the composition and quality of batches of molasses vary too much. Thus the experiments would not be reproducible.

Disturbances

The sensor measurements are disturbed by noise. The characteristics are listed in Table 2.8. Beside the noise, other disturbances act upon the measurements. The disturbances will be listed below according to the measurement device they act upon.

The mass spectrometer measurement is disturbed by regular (once an hour) calibration measurements. The disturbance is two-fold: first, due to the calibration measurement, the time necessary to complete one mass-spectrometer cycle is enlarged (see Figure 2.3); secondly the excessive amount of ethanol in the calibration gas influences the ethanol measurement (see Figure 2.5). The DOT-probe measurement is disturbed by addition of anti-foaming agent if the foam level on the broth is high. By addition of anti-foaming agent the DOT-level drops several percent due to the formation of a new liquid-gas equilibrium. Temperature fluctuations are also visible in the DOT-measurements. The pH measurement is similarly disturbed by temperature fluctuations. The CO_2 and O_2 measurements of the mass-spectrometer are influenced by variations in the stirrer speed and air flow. If the stirrer speed is increased, the CO_2 and O_2 increase respectively due to changes in pressure in the bioreactor. Increases in air flow yield similar results. Due to the fact that the balance is placed on a table which is not damped, any (human) disturbance action which acts upon the table will cause a disturbance in the balance measurements.

Optical Density Measurements and Dry-Weight Analysis

In order to get an indication of the biomass growth one can perform dry weight measurements or optical density measurements. Both give an indication of the concentration of the biomass. Normally the dry-weight measurements are only used for final weights analysis, at the end of both batch and fed-batch phase. This is because the analysis requires about 24 [h] whereas the optical density measurement requires only 10 [min]. Dry-weight analysis gives a reliable value, the optical density measurement has to be correlated with the dry-weight ones.

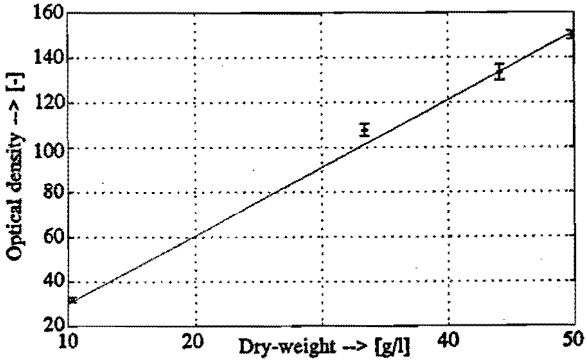


Figure 2.6 *Correlation between cell dry-weight and optical density; error bars indicate standard deviations*

For the dry-weight measurements a sample of 1 [cm^3] is taken from the fermenter and is put in a centrifuge for 10 [min] at a speed of 5000 [rpm]. The centrifuge stops due to air friction. Effluent liquid is taken away with a 'pipet' and the remaining yeast is washed with 5 [cm^3] water. The yeast is diluted again with water, put in the centrifuge and effluent liquid is taken away. Then the remaining yeast, together with adhering water, is put in a pre-weighed cylinder and put in an oven at 95

[°C] until all the liquid is evaporated. Next the cylinder is cooled at room temperature and re-weighted.

For the optical density measurements a sample (about 1 [cm^3]) is taken from the fermenter. The sample is analyzed in a Vitalab 20 spectrophotometer which has an optical filter of 610 [nm] and a light source of 500-650 [nm]. The sample is measured and, if necessary, diluted such that the reading of the spectrophotometer is in the range of 0.20 to 0.40. The dilution is done in as few steps as possible. Next ten samples are made through dilution and are measured, the mean as well as the standard deviation is calculated.

A few measurements were done in order to determine the relation between the optical density measurements and the dry-weight results, see Table 2.7. The resulting numerical correlation (2.1) was used for all experiments

$$\textit{Dry-weight cell concentration} = 0.33 \cdot \textit{Optical density} \tag{2.1}$$

which has an accuracy of 3 %. Figure 2.6 gives the graphical representation of some of the points from Table 2.7.

Table 2.7 *Some dry-weight and optical density measurements*

dry-weight [g.dm^{-3}]	mean opt. [-]	std opt. [-]	dw/opt. [g.dm^{-3}]
44.09	133.40	3.54	0.33
10.25	31.32 32.09	0.93 1.15	0.33 0.32
10.75	32.10	0.83	0.335
10.69	30.94	0.86	0.35
33.43	107.92	2.77	0.31
10.47	31.60	0.59	0.33
49.79	149.88	1.82	0.33
10.45 10.64	33.01	0.95	0.32 0.31

2.6 Conclusions

Much is known about the yeast itself in terms of intracellular reactions, some of which contribute to the model building, both for simulation and control models. Care has to be taken not to incorporate too much knowledge into a controller model as this could make the model too complex for its intended use.

Table 2.8 *Available actuators and sensors and their specifications and characteristics*

Variable Symbol [unit]	Sensor	Actuator	Standard deviation [σ]	Noise band-width [Hz]	Delay [s]	Accuracy
Primary variables						
Temperature* T [°C]	PT100	Heating and cooling water valves (ON/OFF)	0.0246	n.m	5	0.01 [°C]
pH* pH [-]	Ingold pH-electrode	Base and acid roller pumps (PID controller)	0.0032	n.m	10	0.01 [pH]
Foam* - [-]	Conductive foam -detector	Anti-foam roller pump (ON/OFF)	n.m.	n.m.	5	0.01 [cm^3]
Process variables						
Biomass X [g.dm^{-3}]	On-line	None	n.m.	n.m.	n.m.	n.m.
	Off-line		n.m.	n.m.	+/- 600	dilution factor
Glucose G [g.dm^{-3}]	On-line	None	n.m.	n.m.	n.m.	n.m.
	Off-line		n.m.	n.m.	n.m.	n.m.

n.m. = not measurable

Table 2.8 *Available actuators and sensors and their specifications and characteristics. Continued*

Variable Symbol [unit]	Sensor	Actuator	Standard deviation [σ]	Noise band-width [Hz]	Delay [s]	Accuracy
			Process variables			
Ethanol E [ppm]	Mass-spectrometer	None	5.2458	1/16	+/- 30	1 [ppm]
DOT DOT [%]	Ingold DOT-probe	None	0.04 / 0.0279	1/1.6	5	0.1 [%]
O_2 [mole] CO_2 [mole]	Mass-spectrometer	None	0.0074 0.0006745	1/4 1/4	+/- 30 +/- 30	0.001 0.001
			Input variables			
NH_4 NH_4 [$cm^3.min^{-1}$]	NH_4 roller pump	NH_4 roller pump	0.0051	0	5	depends on tube used
Air flow* AF [$dm^3.min^{-1}$]	Flow meter	Air valve	0.060	0	5	0.01 [$dm^3.min^{-1}$]
Glucose flow* GF [$dm^3.min^{-1}$]	Balance	Nutrient roller pump	0.0176	0	5	0.01 [$dm^3.min^{-1}$]
Stirrer speed SS [rpm]	Not applied yet	Stirrer motor	n.m.	n.m.	n.m.	n.m.

n.m. = not measurable

The laboratory process, as defined in Chapter 1, is a good approximation of the real, industrial process through the definition of the three phases: batch, fed-batch, and quality. By applying a fresh inoculum from the freezer a constant quality of experiments is provided. Following the correct procedure to operate the fermenter and following the operational conditions yields a constant amount of biomass, i.e. 10.5 [g.dm^{-3}], at the end of the batch phase. With this constant amount the fed-batch and quality phase can be operated smoothly.

The sensors and actuators could be described quite accurately, e.g. in terms of standard deviation of the noise and bandwidth. Some of the measurements show a dependence on temperature, but not in a severe and deteriorating way. The most important conclusion from the analysis of the sensors and actuators is that key process variables such as biomass and glucose concentration can not be measured on-line. On the other hand primary variables like pH and temperature are controlled adequately to provide the laboratory process with favourable environmental conditions. To get an indication of the biomass growth, dry-weight measurements and related off-line optical density measurements are carried out.

References

Enfors, S.O., Hedenberg, J. and K. Olsson (1990). Simulation of the Dynamics in the Bakers' Yeast Process. *Bioprocess Engineering*, Vol. 5, pp. 191-198.

Rose, A.H. and J.S. Harrison (1971). Physiology and Biochemistry of Yeasts. *Yeasts*, Vol. 2, Academic Press, London, UK.

Sonnleitner, B. and O. Käppeli (1986). Growth of *Saccharomyces cerevisiae* is Controlled by its Limited Respiratory Capacity: Formulation and Verification of a Hypothesis. *Biotechnology and Bioengineering*, Vol. 28, pp. 927-937.

von Meyenburg, K. (1969). Katabolic-Repression und der Sprossungszyclus von *Saccharomyces cerevisiae*. *Ph.D. Thesis*, ETH, Zürich, Zwitserland.

3

Simulation Models

Modelling is found everywhere in science and engineering. For the control engineer modelling is one of the basics. Models used within control science are developed for a wide variety of purposes: to understand processes better, to predict a process behaviour, to simulate part of a process, or to serve as an internal model for a controller. The models made for different purposes need not to be comparable. A simulation model of a process may be based on physical laws (physical state equations), whereas a control model of the same process is based on mathematical (no physical laws) equations (input-output data). So the first question one asks oneself is what has to be modelled and why a process has to be modelled, i.e. the intended use of the model. If we know the purpose of the model we can decide upon the inputs and outputs that have to be modelled, which are the system variables and the parameters describing them. The purpose of the model decides whether parameters and states should have a physical meaning (diagnostic and simulation model) or that the model should only represent process input-output relations.

Next we have to make some assumptions about what kind of approximation is appropriate, as all models are abstractions of reality. The approximation is a trade-off between the accuracy and the complexity of a model, and it is normally determined in an iterative way. This trade-off can be particularly critical when the model is used for on-line process control. The model has to be simple, regarding computer time and response time, as well as robust; the model should be able to keep the controlled process within limits.

The following modelling step involves the actual description of the process in terms of physiological, physical or chemical differential equations, input-output relations, or mathematical equations. The type of model is related to the purpose, e.g. a physical model in order to understand the process better, a black-box model to serve as model for control, etc. For physical models most of these parameters relate to a physical property and can thus be obtained from literature or determined using experiments to estimate the physical parameter. This is not the case for the black-box model parameters. The parameters need not to be related in a direct way to any physical property at all. The black-box model parameters are identified using especially designed experiments.

The last step of modelling is the most important one. The model is checked against experimental data. These data should cover all possible process conditions the model is designed for. These process conditions depend on the purpose of the model. If the model is used as an internal reference model for a controller it should describe the process well in terms of the L_2-norm of the model output minus the process output. If a simulation model is needed, it is of importance that the model can describe the process response for any relevant process condition.

The goal of this Chapter is to come up with two simulation models of the laboratory process of *Saccharomyces cerevisiae*: one for the batch and one for the fed-batch phase. These simulation models should be suitable for testing identification as well as control strategies. These simulation models are based on physiological knowledge and are not intended to serve as control models. Models suitable for control will be derived in the next Chapters.

The contents of this Chapter is as follows: first we deal with a classification and a general overview of bio-process models. We will restrict ourselves to mathematical models; although at present the development of neural-network models is very interesting, it lies beyond the scope of this Chapter. Next we will analyze the fed-batch phase of *Saccharomyces cerevisiae* theoretically and experimentally. Here we look at aspects that are essential for simulation model design. The actual fed-batch phase simulation model is based on a batch phase model discussed in Section 3.3. Parameters are taken from literature and from experiments executed in the laboratory. In Section 3.4, the fed-batch simulation model is presented and evaluated. To end this Chapter conclusions are given.

3.1 Bio-Process Modelling

Many models of the yeast culture *Saccharomyces cerevisiae* have been published. These models may be classified, using the conceptual framework first suggested by Tsuchiya *et al.* (1966), as either *structured* or *nonstructured*, and either *segregated* or *nonsegregated* models. A different classification is used by Schügerl (1982), who also investigates enzyme synthesis models, kinetic models, and reactor models.

If no difference is made between cells, the model is called nonsegregated. Nonsegregated models treat the culture as a collection of average cells, all with the same characteristics at any given time. Segregated models treat each cell as an independent entity, and a population as a collection of such distinct cells.

If one describes the processes happening in a macroscopic way, the description is called nonstructured. Structured models have a detailed description of what is occurring inside the cell. The distinction structured versus nonstructured is hard to make. In this thesis we will use the pragmatic distinction that modelling more than biomass, substrate(s), products, carbon-dioxide, oxygen, and volume is considered structured. Furthermore, if there is more than one description of the condition of the biomass the model is considered segregated. This is somewhat different from the definition used by Tsuchiya *et al.* (1966). Note that all these models give, to a certain extent, more insight into the process or could be used to simulate (part of) the process.

Nonstructured, Nonsegregated Models

These models have cells that are indistinguishable from one another; the cell is a 'black-box' with respect to physiological and metabolic processes. The models come close to models used in controller design. The presented models are simple but can not be validated with experimental data as they have fixed parameters which are determined by physical experiments that do not relate to the process range the models where built for. The models are based on ordinary differential equations of state variables like biomass, glucose, and ethanol. Along with these some expressions for output such as carbon-dioxide production rate, oxygen uptake rate, and dissolved oxygen are presented. The difference between the models reported lies in the algebraic equations used to calculate different rates.

One of the best studied models in this category is the one by Sonnleitner and Käppeli (1986), where the growth of *Saccharomyces cerevisiae* is limited by its oxidative capacity. The model is based on the fact that glucose degradation proceeds via two pathways under conditions of aerobic ethanol formation. Part of the glucose is metabolized oxidatively and part reductively, with ethanol being a product of the reductive energy metabolism. Furthermore ethanol can only be used oxidatively. Enfors *et al.* (1990) described a simulation model for fed-batch bakers' yeast production based on the theory of Sonnleitner and Käppeli (1986). The theory is further supported by investigations and experimental results. The model involves the concept of maximum respiratory capacity of the cell. If the sugar concentration is increased above a certain critical value, corresponding to a critical rate of glycolysis and a maximum rate of respiration, then all the excess sugar consumed is converted into ethanol. This critical sugar concentration is dependent on the type of sugar present. Some experimental evidence was presented for the model but no numerical model validation was carried out. Szewczyk (1989) derived a simple model describing the basic phenomona of bakers' yeast growth. From the biochemical point of view the model is similar to the one presented by Sonnleitner and Käppeli (1986).

Some other models are given in the work of Bajpai and Reuß (1980); e.g. they presented a model for the production of penicillin. They incorporated oxygen limitation on growth and product formation. Various experimental penicillin trends have been simulated qualitatively. A way to model the ethanol production/consumption rate in a fed-batch culture of bakers' yeast is presented by Okada *et al.* (1981). This model has an empirical equation for the ethanol production/consumption rate. A simple model is found by Esener *et al.* (1981) that describes the experiments quite well, except under big transients. A complete different approach has been reported by Savageau and Voit (1982) who used the power-law formalism to model the growth. A comprehensive treatment of this theory is beyond the scope of this thesis.

Nonstructured, Segregated Models

Normally the segregated models are structured as well because they have to take into account the energy and mass flows between the different cell descriptions. In this case the cell is a 'black-box' with respect to physiological and metabolic processes. One of the examples of an nonstructured but segregated model is the model designed by Dantigny *et al.* (1992) where the biomass is divided into two parts, one accounting for the growth on glucose and one accounting for the growth on ethanol. There is a flow from one biomass state to the other depending on the concentration of the glucose and ethanol. A nice example of a simple model is the work of Bailey and Nicholson (1989). Their model not only takes into account the cell mass but also the fresh dry weight and the viability. Cazzador and Mariani (1988; 1989) made a segregated population model for yeast budding. The model is based on the dependence of the critical mass level for budding from the genealogical age and the substrate concentration. In the paper of Takamatsu *et al.*, (1985) three different models describing the population balance of microorganisms are presented. The population balance is based on the assumption of symmetric cell division. The models are used to explain the dynamic behaviour of the ratio of budding cells to the total cells.

Structured, Nonsegregated Models

A vast number of structured, nonsegregated models has been built due to the fact that these models are mostly used for understanding of the process or for testing some hypothesis. The biomass cell is divided into different parts, ranging from a simple nonstructured model with an additional differential equation accounting for a lag phase (Bergter and Knorre, 1972) to very complex ones (Steinmeyer and Shuler, 1989). The range of complexity is given by the trade-off between accuracy and utility of a model and the goal of the model.

One of the earliest simulation models of *Saccharomyces cerevisiae* is described by Bergter and Knorre (1972). Both batch and continuous phase models are presented describing aerobic growth in a glucose limited medium. The models are structured in the sense that they incorporate the diauxic lag phase (lag phase in the batch phase when the glucose has been consumed and the yeast changes to ethanol consumption), modelled as an extra differential equation. Due to its simplicity, the model is very suited for fast simulations of biomass, glucose, and ethanol behaviour.

Steinmeyer and Shuler (1989) presented a model to predict batch phase growth of the yeast on a variety of nitrogen sources and at variable dissolved oxygen concentrations. Furthermore the slowing of growth through the combined effect of nutrient exhaustion and product inhibition in the later stages of the fermentation was modelled. The model presented is extremely complex, 26 rate expressions and 18 mass balances, and is poorly validated. Barford and coworkers (Barford, 1989a, 1989b; Barford and Hall, 1981; Hall and Barford, 1981) presented several models for both batch and continuous phase. The models are based on the experimental observations that the respiratory capacity of the organism may become saturated and may exhibit a maximum specific oxygen uptake. They concluded that this observation could lead to the possibility that transport in and out of mitochondria was of major importance in the overall metabolism of *Saccharomyces cerevisiae.* Based on this a simulation model has been derived that allows precise quantification of the changes in rate and stoichiometry of energy production. The model included an ATP balance. A rather complex model, including gas-liquid phase modelling has been presented by Bellgardt *et al.* (1982). The model is structured but not segregated as it is found in recent work of Bellgardt and coworkers. The mathematical model for the growth of *Saccharomyces cerevisiae* describes growth on glucose and ethanol in aerobic and anaerobic circumstances.

By means of a structured but nonsegregated model describing the accumulation of carbohydrates Heinzle *et al.* (1982) simulated sustained oscillations in continuous cultures of *Saccharomyces cerevisiae*. A model for recombinant yeast which consists of 19 differential equations, 24 analytical reactions, and 48 parameters has been proposed by Coppella and Dhurjati (1990). Most of the parameters have been identified using experimental data. Heijnen and Roels (1981) used a nonsegregated, structured model to verify the hypothesis on yield factors and maintenance in aerobic fermentation processes. A model describing the growth of *Saccharomyces cerevisiae* subject to inhibition by any known product is described by Fukuda *et al.* (1978). The inhibitive substance is modelled as an extra differential equation in both fed-batch and continuous models. Scheiding and Hecht (1982) developed a mathematical model describing the biological activities of the cell plus all balanced equations for the liquid and gas phase of the stirred tank reactor.

Structured, Segregated Models

These models take everything into account and are thus regarded as the ones closest to real process behaviour. Due to the fact that the models tend to become very complex as they are already segregated, the structure level is very low, such that they can still be utilized. Bellgardt and coworkers (Bellgardt *et al.*, 1991; Bellgardt and Yuan, 1992; Yuan *et al.*, 1992) developed a structured, segregated model of the growth of *Saccharomyces cerevisiae*. They built a cell cyclic model for unequal budding yeast and verified it with experimental data available in literature. The idea behind the model is that the quality of the yeast is strongly influenced by the cyclic process and the model can be used to optimize and control the quality of the harvested yeast cells. The model could also be called nonstructured as it is only the ATP that is modelled in excess of all other normal nonstructured parameters.

3.2 Quantitative Process Analysis

Saccharomyces cerevisiae is, together with *Saccharomyces carlsbergenius* used for beer brewing, perhaps one of the best studied yeast cultures; much is known about the metabolic flows in the cell and how they relate to growth. Most of this knowledge is from continuous culture experiments where yeast cells are cultivated under practically steady state conditions. This knowledge is expressed as stoichiometric relations and specific fluxes of components. From an identification and control point of view it is, however, much more important to understand the dynamics and (inherent) non-linearities of the process. This knowledge can help to design a suitable simulation model, experiments to identify/validate a model, and a correct control strategy.

In this Section we will look more closely at the various aspects of the laboratory process. In order to design a simulation model we have to adopt one of the theories accounting for yeast growth; here we will follow the bottleneck theory of

Sonnleitner and Käppeli (1986). Together with this theory some non-linearities and time constants of the culture can be explored. Having classified the type of yeast culture model we want to build, we can do experiments in order to find the dynamical characteristics of the yeast culture (Backx and Damen, 1989). With the information gained we can design appropriate experiments for identification and validation of the simulation model, and we are able to use the information for other model design, e.g. control.

The Theory of Cell Growth

Glucose is used by the cell both as an energy source and the sole carbon source for building material for growth. Ammonia is used as the nitrogen source. The glucose metabolism and the cell growth can be described as in Section 2.1. The simulation model can be of a lower complexity as intracellular components are not measured on-line. The batch phase model will be built around the measurable output variables and will account for some of the lag phases occurring during the batch phase.

The Stoichiometry
Depending on the glucose supply rate the culture can grow under different conditions. Glucose can be consumed both oxidatively and reductively. In the former the glucose is, together with oxygen and nitrogen, converted to energy, cellmass, carbon dioxide, and water. In the latter no oxygen is present and ethanol is produced as a by-product. This ethanol can then be consumed together with oxygen and nitrogen to yield energy, cellmass, carbon dioxide, and water. We assume that during all reactions taking place the cellmass composition remains the same. The metabolic pathways can be given as follows:

Oxidative growth on glucose

$$C_6H_{12}O_6 + aO_2 + bnxNH_3 \rightarrow bCH_{hx}O_{ox}N_{nx} + cCO_2 + dH_2O \qquad (3.1)$$

Reductive growth on glucose

$$C_6H_{12}O_6 + gnxNH_3 \rightarrow gCH_{hx}O_{ox}N_{nx} + hCO_2 + iH_2O + jC_2H_6O \qquad (3.2)$$

Oxidative growth on ethanol

$$C_2H_6O + kO_2 + lnxNH_3 \rightarrow lCH_{hx}O_{ox}N_{nx} + mCO_2 + nH_2O \qquad (3.3)$$

Maintenance reaction

$$\alpha O_2 + CH_{hx}O_{ox}N_{nx} \rightarrow CO_2 + \beta H_2O + nxNH_3 \qquad (3.4)$$

In all these pathways $C_6H_{12}O_6$ stands for glucose, C_2H_6O for ethanol and $CH_{hx}O_{ox}N_{nx}$ for the normalized elemental composition of yeast. The parameters hx (=1.8), ox (=0.5), and nx (=0.2) represent the composition of the yeast (Roels, 1983). Next we want to determine the stoichiometric coefficients a to j, α and β. Elemental balances for C, H, O, and N give rise to a system of equations. There are three independent equations and four unknowns; hence a measurement must be made.

Table 3.1 *Stoichiometric constants based on the cell yields*

O_2	$CH_{hx}O_{ox}N_{nx}$	CO_2	H_2O	C_2H_6O
a = 2.1675	b = 3.65	c = 2.35	d = 3.81	
	g = 0.36	h = 1.892	i = 0.1642	j = 1.874
k = 1.614	l = 1.32	m = 0.68	n =2.208	
α = 1.05			β = 0.6	

This can be:

the cell yield (Axelsson, 1989),
the respiratory quotient (Stephanopoulos *et al.*, 1984),
the ethanol production rate divided by the oxygen uptake rate together with the respiratory quotient (Shimizu *et al.*, 1989), or
the carbon dioxide production rate and oxygen uptake rate divided by the ammonia consumption rate (Shioya *et al.*, 1985).

In the last case the ammonia is not directly related to the biomass. Table 3.1 lists the used stoichiometric coefficients based on the cell yield. Another possibility is to compute the stoichiometric constants using singular value decomposition, Saner *et al.* (1992).

The maintenance reaction, Eq. (3.4), can be seen as composed of two parts: firstly, if no glucose is available, Eq. (3.4) represents the decay of the culture; secondly, if glucose is available, a part of the growth on glucose or ethanol is 'consumed', Eq. (3.1) - (3.3), for maintenance.

The Bottle-Neck Principle.

The different pathways are described as separate reactions; in reality, however, the total reaction is a combination of several of these pathways depending on the availability of glucose, ethanol and oxygen. We assume that the maintenance reaction is always active. If glucose is present, the proportion assigned to the oxidative reaction and the reductive reaction depends on the amount of glucose present compared to the respiratory capacity, which is dependent on the amount of oxygen present. If the glucose concentration is lower than the respiratory capacity, the growth will be purely oxidative. If the glucose concentration exceeds the respiratory capacity, part of the glucose will be metabolised into ethanol. Another case exists if both ethanol and glucose are present and the glucose concentration is less than the respiratory capacity. In this case the glucose will be metabolised preferably. The bottleneck principle does not support simultaneous production and consumption of ethanol. This holds if the medium is mixed homogeneously. A schematic representation of the bottle-neck principle is given in Figure 3.1.

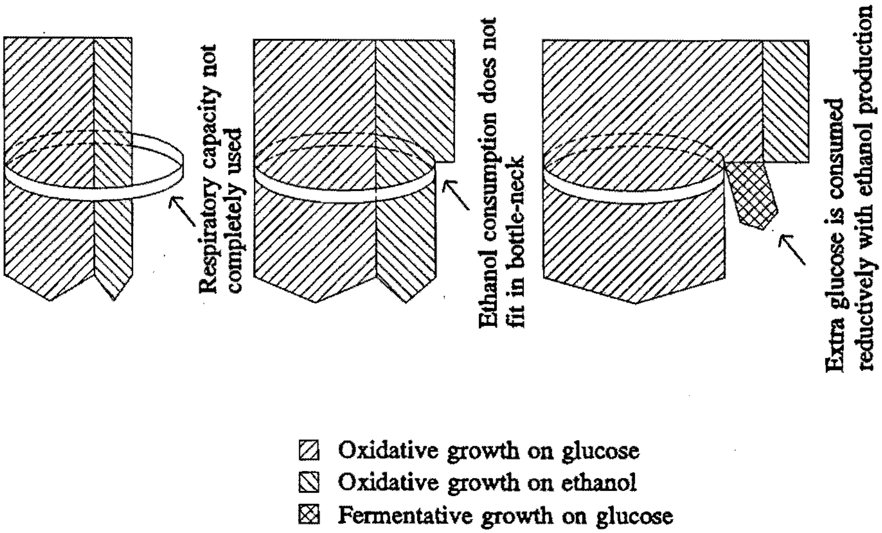

Respiratory capacity not completely used
Ethanol consumption does not fit in bottle-neck
Extra glucose is consumed reductively with ethanol production
▨ Oxidative growth on glucose
▧ Oxidative growth on ethanol
▩ Fermentative growth on glucose

Figure 3.1 *Bottle-neck principle*
Left picture, respiratory capacity (open band) exceeds available glucose and ethanol. Both glucose and ethanol are metabolized oxidatively.
Middle picture, glucose does not, but glucose and ethanol do exceed the respiratory capacity. Part of the ethanol is not metabolized.
Right picture, glucose exceeds the respiratory bottleneck and excess glucose is metabolized reductively to ethanol. Ethanol is not consumed in this case.

System Non-Linearities

The non-linearities of the process can be divided into two groups: nearly static non-linearities and dynamic non-linearities. Nearly static non-linearities are the exponential growth, the Monod type of kinetics and the oxygen transfer coefficient. Dynamic non-linearities are found in the switch between different pathways, the oxygen limitation, and the influence of both stirrer and air flow on the equilibrium between gas and liquid phase in the head space of the fermenter.

Static Non-Linearities

Most of the nearly static non-linearities are found in the physiological part of the fermentation process. These static non-linearities normally do not change during the process. Their time constants are very small in comparison to the other dynamics of the system. Because of their very small time constants they are called 'nearly' static non-linearities. We can distinguish the following nearly static non-linearities:

- *Exponential growth.* The yeast cells are growing through cell budding. This can be modelled as an exponentially growing system. The dynamics are very slow but are sensitive to environmental parameters and to events in the past (the memory effect of yeast).

- *Monod kinetics.* Monod kinetics are used to describe steady-state systems, here we also use them to describe dynamic events. The algebraic description of the Monod kinetics is:

$$U = \frac{[\ ]}{k_{[\ ]} + [\ ]} \qquad 0 \leq U \leq 1 \tag{3.5}$$

with: $[\ ]$ = e.g. concentration of ethanol/glucose/.... and U a dimensionless variable.

- *Oxygen transfer coefficient.* The relation between the dissolved oxygen tension, the stirrer speed and the air flow is expressed as transfer coefficient of oxygen into the medium:

$$K_L a(t) = C_{DOT} SS(t)^{C_{SS}} AF(t)^{C_{AF}} \tag{3.6}$$

The non-linearities are the multiplications of the different terms and the power coefficients C_{SS} and C_{AF}. This non-linearity is static and it is believed that the parameters only vary very slowly during the laboratory process, due to the changing viscosity of the medium.

Dynamic Non-Linearities

The nearly static non-linearities can be approximated by some linear functions or compensated by appropriate choices of the input signals. However, the dynamic disturbances require a more careful treatment. The following dynamic non-linearities occur:

- *Oxygen limitation.* Oxygen limitation is a shortage of oxygen in the broth due to insufficient oxygen supply. During the fed-batch phase oxygen limitation will probably not occur. If the effect takes place it results in an excessive production of ethanol. The process dynamics change rapidly under this situation.
- *Ethanol production/consumption.* The switch between ethanol production and consumption is a kind of lag phase. This lag phase is dependent on the enzymes available at that moment. These enzymes are dependent on the history of the fed-batch cultivation, e.g. the amount of ethanol consumed.
- *Stirrer influences.* During the fed-batch phase an equilibrium between liquid and gas exists in the head space of the fermenter. If the stirrer speed is changed there is a pressure change in the liquid leading to a new equilibrium between gas and liquid. During the settling of this equilibrium the measurements are disturbed. To avoid these non-linearities the rate of change of the stirrer is limited. Note that the same holds for the air flow.

Transitions

The transitions from one phase to another are characteristic for the laboratory process. The biological behaviour of the process changes during a transition. An example of such a transition is the change from ethanol production to ethanol consumption during the first part of the batch phase of the model process. Another, more important, example is given in Example 3.1

Example 3.1 *Transition from batch phase to fed-batch phase of the laboratory process*

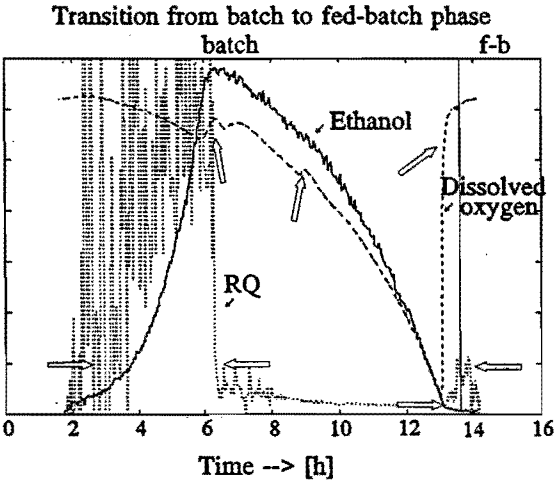


Figure 3.2 *Batch and begin of fed-batch phase; big arrows denote possible detection points; the plot is to illustrate the transition, no values are given*

The transition from batch phase to fed-batch phase is important in order to start any control actions in time and in order to prevent the yeast from starvation and quality loss. The change in phase is detectable by the following items (see Figure 3.2):

- a noticeable rise of the dissolved oxygen tension although stirrer speed and air flow are kept constant.
- a decrease to almost zero of the ethanol concentration.
- a noticeable rise of the pH if not corrected by acid addition, otherwise a noticeable acid addition.
- less cooling of the fermenter.
- RQ is at 1.08.

Note that all these effects occur simultaneously and that detection by using only one of these effects could lead to the wrong estimated time point of the transition. In Figure 3.2, it can be seen that at several time instants one, or even two, of the above mentioned effects occur. The effects are marked with big arrows. However, the batch phase is not yet completed at those time instants but only around 14 [h] when all effects are present simultaneously. ■

Dynamic Characteristics of the Process

Although a great number of characteristics have been found through theoretical analysis of the process, we still need more information in order to be able to design experiments. The set up of the analysis of dynamical characteristics is based on the work of Backx and Damen (1989). Here we mention the steps taken and the results obtained from these steps.

Table 3.2 *Free run settings of the inputs*

Glucose flow (*GF*)	$\frac{\mu}{Y_{x\mid s}\cdot S_0}X_0\cdot e^{\mu\cdot t}$	[mole.h^{-1}]
Air flow (*AF*)	4.0	[dm^3.min^{-1}]
Stirrer speed (*SS*)	750	[rpm]

To begin we interviewed operators about 'normal' process operations, actuator and sensor behaviour; this provides a first indication of the dominant time constants of the process. Most of this information is denoted in Tables 2.5 to 2.8. With this information we designed so called 'free run' experiments to determine the noise characteristics of the process. In our case the process was operated in open loop with exponential feeding of glucose, constant air flow and constant stirrer speed; see Table 3.2. The results of the noise characteristics can be found in Table 2.8. No information is available regarding the process noise as the noise generated by the process was lower than the measurement noise of the sensors.

Next we designed 'staircase' signals for all inputs, *GF*, *AF*, and *SS*, to check for the steady state gains and the largest time constants of the process. Due to the fact that we know a-priorily that the system is non-linear, a number of steps is taken and for the glucose flow these steps are multiplied with the open-loop feed curve (see Table 3.2). For the other inputs the same default values (Table 3.2) are taken as for the free run experiment. A step of 10% from the nominal value is made for each of the inputs. The results of the staircase experiments are combined in Table 3.3. Note that the gains are normalized by the inputs.

There are two distinct entries in the *SS-DOT* field of Table 3.3. This is because the dissolved oxygen tension reacts differently to stirrer speed changes if the *DOT* is 'high' (>30%), 'normal' (30%<*DOT*<10%), or 'low' (<10%). The value of 0.5 [%.rpm^{-1}] is for the high *DOT* values, the 1.6 [%.rpm^{-1}] is for the normal values. For the low *DOT* values no gain could be estimated.

Table 3.3 *The estimated gains of the staircase experiments*

	OUR [mmole.h^{-1}]	*CPR* [mmole.h^{-1}]	*DOT* [%]
GF [% X_0]	5.8	5.3	0.73
AF [dm^3.min^{-1}]	134	142	16.5
SS [rpm]	1.8	1.3	0.5 / 1.6

The smallest process time constant is smaller than the equivalent time constant of the sample frequency of 4/60 [Hz], see Table 3.4. The results of Table 3.4 are a only rough indication of the real time constants because of the low sample frequency. Only the *DOT* could be sampled at a higher frequency, high enough for reliable results. There are again three cases to be distinguished. These three cases differ from the ones presented discussing the gains:

a. The *DOT* reacts quickly, the time constant τ is calculated to be 0.006 [h] (= 20 [s]), if the step is positive, caused by a positive *SS* change, and no oxygen limitation occurs. This is a purely physical reaction.
b. The *DOT* reacts slowly, τ is 0.04 [h] (2.4 [min]), if the step is negative, caused by a positive *GF* change, and no oxygen limitations occurs. Now the decrease of the *DOT* is caused by the physiological process of the biomass.
c. If oxygen limitation occurs, the changes in the *DOT* are so small that no time constant can be estimated.

The next step in the quantitative process analysis is to apply fast Pseudo Random Binary Sequences to the laboratory process inputs. From the laboratory process response we can calculate the bandwidth of the process, the delay times, and we can distinguish between input influences (Table 3.4). The PRBS sequences have to be mutually independent and are applied symmetrically around the normal operating points (Table 3.2). The clock frequency of the PRBS equals the sample frequency such that a flat frequency spectrum is used on the bandwidth $f_s/2$. The duration of the sequence should be about 5 to 10 times the largest time constant. Normally the sizes of the input variations are chosen in relation to their signal to noise ratios but, as the process is non-linear, this is not possible; instead visual inspection of the staircase experiments sets the excitation.

Table 3.4 *Time constants, bandwidth and delay times of the process; all times are in minutes*

	Ethanol	*DOT*	*OUR*	*CPR*
Bandwidth	1/16	1/1.6	1/4	1/4
Smallest TC	*	2.1	4	4
Delay Time	*	2	3	3

Normally the delay times are estimated from the PRBS data using correlation techniques. If a cross correlation between an input and output is calculated the reaction of the output due to another input is averaged. This is possible if the inputs are mutually independent and if the process is linear. However, if during the process the gain of one of the input-output combinations changes, the reactions of the changed gain on other inputs-output combinations are not averaged, and the results of the cross correlation of the different input-output reactions are corrupted. In this case it is not possible to use these techniques. Step responses have to be used to estimate the delay times. Visual inspection gives the time the output starts to react on the step applied at the input; this is the delay time (Table 3.4). Not all outputs react on the inputs and these delay times can not be estimated; for example the ethanol concentration does not react on all inputs.

Before identifying the batch phase of the laboratory process some considerations would enhance the identification results. From the analysis of the smallest time constant it is known that the *DOT* shows a 'fast' reaction to stirrer speed changes and a 'slow' reaction to glucose flow changes. This difference can be explained by the difference in speed for physiological (cell related) events which are 'slow' and physical (broth related) events which are 'fast'. The fast dynamics need a sample frequency, 0.2 [Hz], that is ten times higher than the slow dynamics, 0.02 [Hz]. The fastest sample frequency available is 1/60 [Hz], as presented in Section 2.5. This is just enough to cover the slow dynamics; for the fast dynamics some other solution has to be found if one wants to identify the dynamics. For control purposes we are not interested in the physical, fast dynamics and consequently will not identify the related parameters of the physical dynamics for any of the models.

3.3 The Batch Model

This Section discusses the estimation of the parameters of the physiological simulation model of the batch phase of the laboratory process. The batch phase simulation model is based on physiological laws. The choice for a simulation model based on physiological laws is explained at the beginning of this Chapter. The simulation model that will be built is a structured, nonsegregated one. This Section will have three subsections: the simulation model, the experimental data, and the actual parameter estimation.

The Simulation Model

We built a physiological model of the batch phase of the model process. The simulation model is capable of describing both the steady states and the transients of the phase. The model is based on the oxygen limited capacity of the yeast cell (Sonnleitner and Käppeli, 1986) combined with a regulated enzyme production of the cell (Sweere, 1988) and a maintenance term. Furthermore the dissolved oxygen tension, oxygen uptake rate, and carbon dioxide production rate equations are imbedded in the model. Glucose is the only limiting substrate. Vitamins, trace elements and salts are considered to be not limiting. The physiological description has been transformed in mathematical terms. It is based on first order ordinary differential equations combined with empirical rules and algebraic expressions.

A schematic diagram of the differential and algebraic equations is given in Figure 3.3. In the figure only the basic equations are given; maintenance, dilution terms, the oxygen related equations, the inputs and outputs are not shown. Showing these items as well would obstruct the insight in the model structure. Three blocks can be distinguished in Figure 3.3. Firstly, there are the four balances, biomass, oxygen, glucose, and ethanol. These balances, except for the oxygen one, contain integrators. Next, to the right of the Figure is the oxygen bottle-neck. The grey shaded area represents the same function as the one explained with Figure 3.1, and the same as the theory introduced by Sonnleitner and Käppeli, 1986. In this block diagram the maintenance and carbon-dioxide balances are not included.

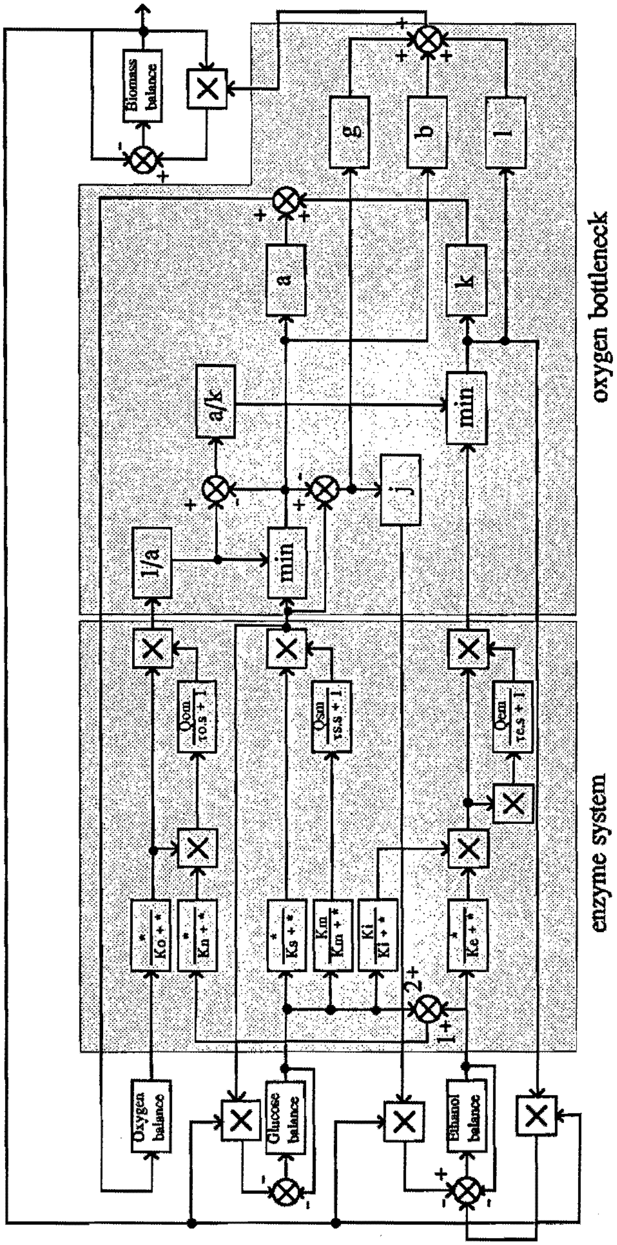

Biomass balance
Oxygen balance
Glucose balance
Ethanol balance
oxygen bottleneck
enzyme system
g
b
l
a
k
a/k
1/a
min
min
j
Qom / τo.s + 1
Qsm / τs.s + 1
Qem / τe.s + 1
* / Ko + *
* / Kn + *
* / Ks + *
Km / Km + *
Ki / Ki + *
* / Ke + *

Figure 3.3 *A schematic overview of part of the batch phase simulation model*

At the left of Figure 3.3 is the regulatory enzyme system (Sweere, 1988) which influences the diauxic phase of the batch phase of the model process. The diauxic phase is the change from reductive growth on glucose to oxidative growth on ethanol. At the right hand side of the enzyme system block area are the integrators representing the regulatory enzymes. At the left hand side are several Monod blocks representing the influence of the states on the enzymes.

The physiological model is extended with the physical environment (see Section 2.5) valid for our laboratory model process:

- The fermentation takes place in a 10 litre continuously stirred tank reactor.
- The continuously stirred tank reactor is perfectly mixed; no gradients occur within the broth.
- The temperature and pH are kept constant at 30 [°C] and 5.0 respectively.
- Foam production is negligible.
- Addition of anti-foam and NH_4OH does not influence the process in any other way than keeping foam negligible and the pH at 5.
- There are no correction terms for time constants, time lags, etc. of actuators and sensors.

The model can be described by a set of differential and algebraic equations. The differential equations concern the biomass concentration (X), the glucose concentration (S), the ethanol concentration (E), the regulatory oxygen enzyme (Q_{omax}), the regulatory glucose enzyme (Q_{smax}), and the regulatory ethanol enzyme (Q_{emax}). These differential equations are represented in Eq. (3.7) - (3.12):

$$\frac{dX(t)}{dt} = \left[g \cdot Q_s^{red} + b \cdot Q_s^{oxid} + l \cdot Q_e^{oxid} - ms \right] \cdot X(t) \tag{3.7}$$

$$\frac{dS(t)}{dt} = -Q_{smax}(t) \cdot \frac{S}{K_s + S} \cdot X(t) \tag{3.8}$$

$$\frac{dE(t)}{dt} = \left[Q_e^{oxid} - j \cdot Q_s^{red} \right] \cdot X(t) \tag{3.9}$$

$$\frac{dQ_{smax}(t)}{dt} = \left[Q_{sm} \cdot \frac{S}{K_n + S} - Q_{smax}(t) \right] / \tau_s \tag{3.10}$$

$$\frac{dQ_{omax}(t)}{dt} = \left[Q_{om} \cdot \frac{DOT}{K_o + DOT} \cdot \frac{2S + E}{K_m + 2S + E} - Q_{omax}(t) \right] / \tau_o \tag{3.11}$$

$$\frac{dQ_{emax}(t)}{dt} = \left[Q_{em} \cdot \frac{E}{K_e + E} \cdot \frac{K_i}{K_i + S} \cdot \frac{DOT}{K_o + DOT} - Q_{emax}(t) \right] / \tau_e \tag{3.12}$$

The above relations describe the state of the batch phase of the model process. The regulatory enzyme production parameters and the oxygen bottle-neck related parameters are described by the following algebraic equations:

$$K_La(t) = C_{DOT}SS(t)^{C_{ss}}AF(t)^{C_{af}} \tag{3.13}$$

$$OUR(t) = \left[a \cdot Q_s^{oxid} + k \cdot Q_e^{oxid} + \frac{ms}{\alpha}\right] \cdot X(t) \tag{3.14}$$

$$CPR(t) = \left[c \cdot Q_s^{oxid} + h \cdot Q_s^{red} + j \cdot Q_e^{oxid} + ms\right] \cdot X(t) \tag{3.15}$$

$$DOT(t) = DOT_{\max} - \frac{OUR(t)}{K_La(t)} \tag{3.16}$$

$$Q_s^{oxid} = \min \begin{cases} Q_{smax} \cdot \dfrac{S}{K_s + S} \\ \dfrac{1}{a} \cdot Q_{omax} \cdot \dfrac{DOT}{K_o + DOT} \end{cases} \tag{3.17}$$

$$Q_s^{red} = Q_{smax} \cdot \frac{S}{K_s + S} - Q_s^{oxid} \tag{3.18}$$

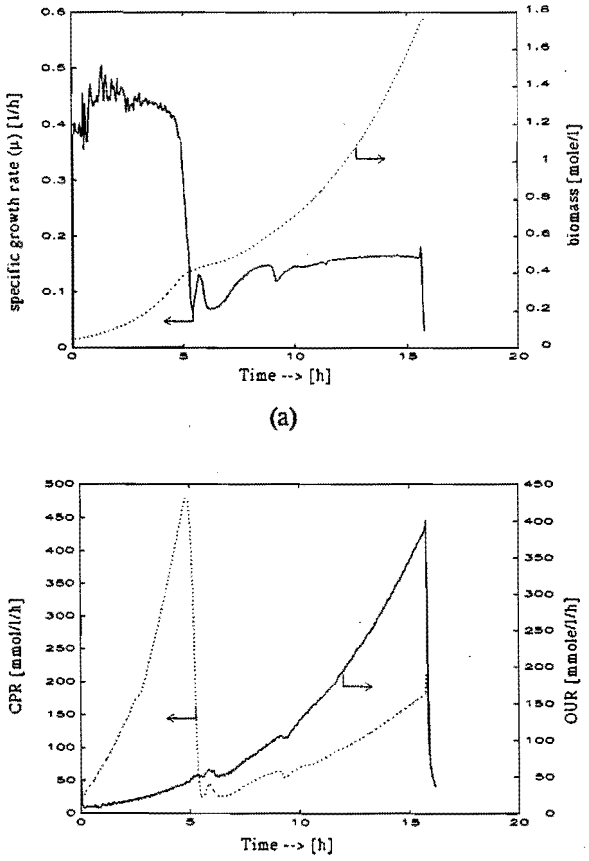
$$Q_e^{oxid} = \min \begin{cases} Q_{emax} \cdot \dfrac{E}{K_e + E} \\ \dfrac{1}{k} \cdot Q_{omax} \cdot \dfrac{DOT}{K_o + DOT} - \dfrac{a}{k} \cdot Q_s^{oxid} \end{cases} \tag{3.19}$$

Relevant process inputs and outputs are chosen. There are two inputs: the stirrer speed (*SS*) and the airflow (*AF*) and four outputs: the ethanol concentration (*E*), the dissolved oxygen tension (*DOT*), the oxygen uptake rate (*OUR*), and the carbon-dioxide production rate (*CPR*). Implemented in a simulation environment, the simulation model contains 16 parameters, 6 of which are related to Monod kinetics (K_j, $j = o,e,s,n,i,m$), a maintenance term (*ms*), 6 parameters that are related to specific regulatory enzyme productivity (Q_{jm}, τ_j, $j = o,e,s$), and 3 that are related to the oxygen transfer from gas to liquid (C_j, $j = DOT,SS,AF$). These parameters have to be estimated and their estimated values have to be validated. In order to do so, suitable experiments are designed.

Experimental Data

The batch phase simulation model will serve as an initial model for the fed-batch phase simulation model. Normal operation, as described in Chapter 2, is followed in order to get a specific predefined amount of yeast for the fed-batch phase. This incorporates a high glucose level at the beginning of the batch and high ethanol levels as a result of this. Due to these specific batch phase conditions only batch data can serve as estimation/validation data for the batch phase simulation model. There is however one unfavourable characteristic about the batch phase; none of the inputs as described above is excited. We try to overcome this problem by varying the initial amount of glucose at the beginning of the batch phase. Two experiments were executed, one with S_0 = 30 [g.dm^{-3}], Example 3.2, and one with S_0 = 5 [g.dm^{-3}], Example 3.3.

Example 3.2 *Experimental batch fermentation with S_0 = 30 [g.dm^{-3}]*



(a)

(b)

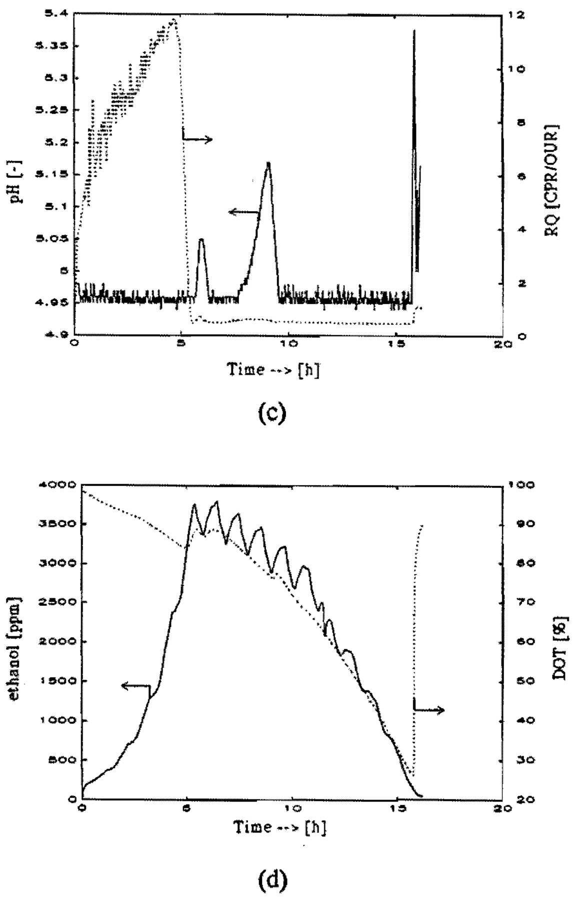


Figure 3.4 *Experimental batch phase results of Saccharomyces cerevisiae and a synthetic medium with S_0 = 30 [g.dm^{-3}], part (a) is the estimated specific growth rate and biomass concentration versus time, part (b) is the oxygen uptake rate and carbon-dioxide production rate versus time, part (c) is the pH and respiratory quotient versus time, and part (d) is the ethanol concentration and dissolved oxygen tension versus time*

An experimental aerobic batch phase is depicted in Figure 3.4. Note that the specific growth rate and biomass concentration are estimated values, not measured values. The estimated biomass concentration and specific growth rate in Figure 3.4(a) show the diauxic phase of yeast as described by amongst others Sonnleitner and Käppeli, (1986). From the same diagram we can see that the specific growth rate is about 0.45 [h^{-1}] during reductive growth on glucose, and 0.18 [h^{-1}] during oxidative growth on ethanol. During the diauxic phase, the change from reductive growth on glucose to oxidative growth on ethanol, the specific growth rate varies between 0.07

[h^{-1}] and 0.16 [h^{-1}]. This phenomenon can be related to the change of extracellular components such as pyruvate, acetate and acetaldehyde which are consumed before or together with ethanol (Dantigny *et al.*, 1989). Furthermore, enzymes related to the uptake of ethanol, glucose and oxygen change during the diauxic phase.

The change during the diauxic phase is also seen from the exhaust-gas plots in Figure 3.4(b) and 3.4(c) where high *CPR* and low *OUR* are seen in the first phase of the growth when the glucose is fermented.

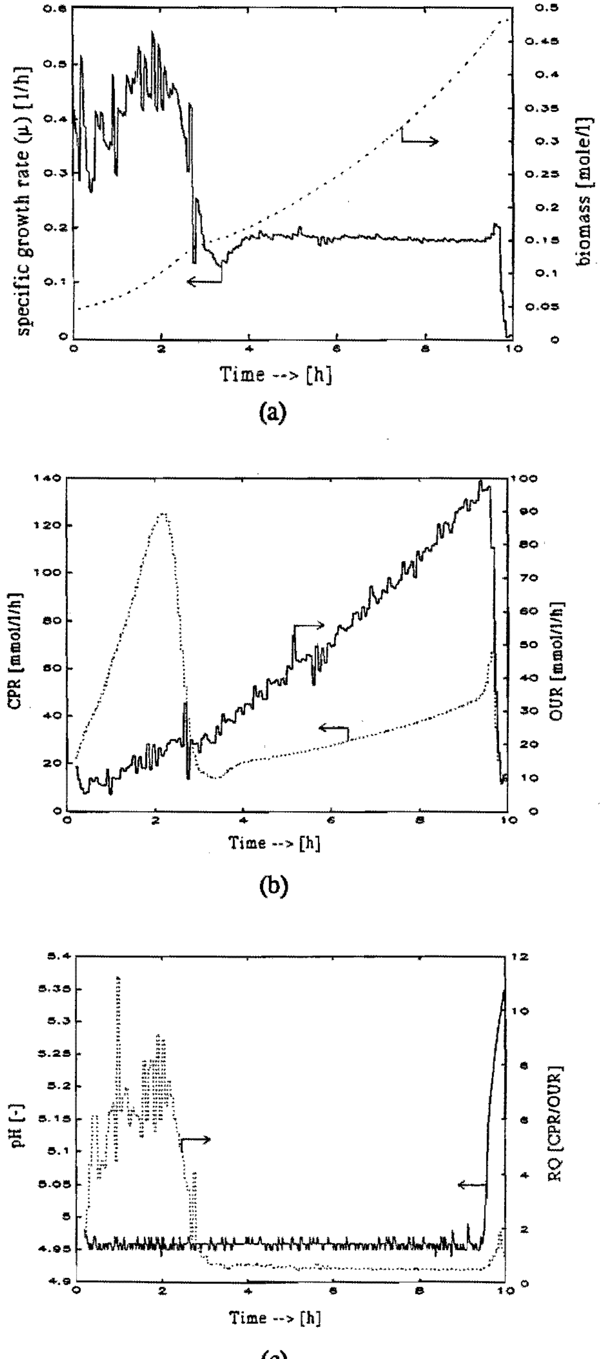
The *RQ* rises to a value above 10 with an abrupt change as the glucose is totally consumed and the cell has to adapt to oxidative growth on ethanol. Figure 3.4(c) shows the pH changes to follow the yeast growth curve. Note that NH_4OH is added to keep the pH at 5 but no acid is added if the broth turns basic. Three (small) deviations from the set-point are registered; first after the fermentation phase, there is a very small increase probably due to consumption of pyruvate; then about two hours later there is an increase probably due to the consumption of acetate. At the end of the batch phase the pH increases again after all nutrients have been consumed. The ethanol concentration is shown in Figure 3.4(d) together with the *DOT*. One can see that preference is given to glucose over ethanol consumption in the first part of the batch phase, after which ethanol is consumed. The oscillations in the ethanol curve are due to unknown cyclic events. ■

We don't know if the extracelluar components pyruvate, acetate, and acetaldehyde would also be visible in the diagrams if a lower amount of ethanol would be produced. Therefore the S_0 = 5 [$g.dm^{-3}$] experiment is carried out to see if the same events occur as with the S_0 = 30 [$g.dm^{-3}$] experiment. From the glycolysis of sugar (Figure 2.1) it is believed that due to the lower concentrations of glucose and ethanol no, or few, extracellular effects would be seen in the diagrams.

Example 3.3 *Experimental batch fermentation with* $S_0 = 5$ [$g.dm^{-3}$]

The results of this experiment with a low glucose concentration are shown in Figure 3.5. If we compare the figures with the batch phase for 30 [$g.dm^{-3}$] some distinct changes are noticeable; the pH is smooth over the whole range of the batch phase, so no base excretion has been detected.

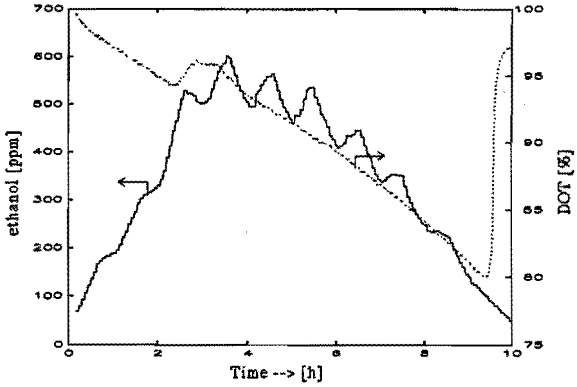
In addition to this, the changes in *CPR*, *OUR*, and *DOT*, noticeable around the diauxic phase, are not present except for the change of metabolism. It is thus believed that extracellular components such as acetate, acetal-dehyde, and pyruvate are present in smaller concentrations than in the batch phase with S_0 = 30 [$g.dm^{-3}$] and that they do not influence the batch phase behaviour whatsoever. The parameters found with the 5 [$g.dm^{-3}$] experiment will be more applicable to the fed-batch phase, as high glucose and ethanol concentrations are not intended in the laboratory model fed-batch phase.


specific growth rate (μ) [1/h]
biomass [mole/l]
Time --> [h]
0
0.1
0.2
0.3
0.4
0.5
0.6
0.05
0.15
0.25
0.35
0.45
2
4
6
8
10
CPR [mmol/l/h]
OUR [mmol/l/h]
20
40
60
80
100
120
140
30
50
70
90
pH [-]
RQ [CPR/OUR]
4.9
4.95
5
5.05
5.1
5.15
5.2
5.25
5.3
5.35
5.4
12

(a)

(b)

(c)



(d)

Figure 3.5 *Experimental batch phase results of Saccharomyces cerevisiae and a synthetic medium with S_0 = 5 [g.dm^{-3}], part (a) is the estimated specific growth rate and biomass concentration versus time, part (b) is the oxygen uptake rate and carbon-dioxide production rate versus time, part (c) is the pH and respiratory quotient versus time, and part (d) is the ethanol concentration and dissolved oxygen tension versus time* ■

From the above two Examples we learn that, although the initial amount of glucose differs by a factor of 6, the shape of the batch phase remains the same; excessive ethanol production in the first part of the batch phase followed by ethanol consumption in the second part. Both parts of the batch phase are linked through the diauxic phase. In the diauxic phase the biggest difference between the two Examples is seen due to the postulated extracellular concentrations of pyruvate, acetate, and acetaldehyde. If present, the extracellular components give rise to some events during the diauxic phase, otherwise the effects are not noticeable.

Estimation of the Parameters

The 15 model parameters are estimated using common sense and least squares algorithms. By common sense, it is meant that we try to minimize the number of parameters to be estimated at a time. This is done by looking at the batch phase and determining the sensitivity of the parameters at a certain stage. Similar approaches using sensitivity functions are reported by Kishimoto *et al.*, 1976; Choi and Park, 1981; Vialas *et al.*, 1985; Young *et al.*, 1986; Espie and Macchietto, 1989; Munack, 1989; Posten and Munack, 1990; Menawat and Balachander, 1992; Schneider *et al.*, 1992. The parameters are estimated using data were the sensitivity of the parameters is high. The least squares algorithms are used to minimize the output error of both the *OUR* and the *CPR*, i.e. the error = the process *OUR* (*CPR*) - the simulation model *OUR* (*CPR*). The *OUR* and *CPR* are used for the physiologically related parameters whereas the *DOT* is used for the physically related parameters. Firstly we will deal with parameters associated with the dissolved oxygen tension as they can be reformulated depending on stirrer speed, air flow and oxygen uptake rate only, and all these can be measured. Next we split the data into areas where ethanol is produced and areas where ethanol is consumed, and estimate the parameters belonging to either of these two pathways.

A normal mass balance on dissolved oxygen (*DOT*) over the fermenter yields (*Lee et al.*, 1991):

$$\frac{dDOT(t)}{dt} = K_L a(t) \cdot \left(DOT_{saturation} - DOT(t)\right) - OUR(t) \tag{3.21}$$

A pseudo steady state approximation for the *DOT* can be used since the mass transfer rate of O_2 from gas to liquid and the rate of O_2 utilization are much larger than the time rate of change in *DOT* (about 300 times). Therefore the dissolved oxygen parameters can be seen as part from the *OUR* and the K_La:

$$DOT(t) = DOT_{saturation} - \frac{OUR(t)}{K_L a(t)} \tag{3.22}$$

where K_La is as in Eq. (3.6).

An additional dependency on the biomass due to changing viscosity can be modelled as (Coppella and Dhurjati, 1990):

$$DOT^{*}_{saturation}(t) = DOT_{saturation} \cdot \left[1 - \frac{X(t)}{X_{visc}}\right] \tag{3.23}$$

with X_{visc} = 35 [g dcw.dm^{-3}], according to Coppella and Dhurjati (1990). The X_{visc} is rather low as during our experiments much higher concentrations of biomass were registered without much influence on the *DOT*. We assume for our process that $X_{visc} \gg X$. As there are no stirrer speed or air flow changes, the constants related to the K_La have to be determined from fed-batch experiments. The values found are denoted in Table 3.5 with the standard deviations. These values are valid for both experiments as there was no difference in aeration or stirrer speed.

Table 3.5 *Estimated K_La parameters*

parameter	value	standard deviation	dimension
C_{DOT}	15.1	$1.32 \cdot 10^{-2}$	[h.dm^{-3}]
C_{SS}	1.50	-	[-]
C_{AF}	0.46	0.014	[-]

The oxygen related to the oxidative pathway Eq. (3.1) can be calculated using the stoichiometric coefficients (Table 3.1) and the measured oxygen uptake rate (*OUR*). The relation between *OUR* and oxygen used is:

$$OUR(t) = \left[Q_o(t) + \frac{ms}{\alpha}\right] \cdot X(t) \tag{3.24}$$

From this equation we can calculate $Q_o(t)$ if we know *ms* and *X(t)*. The maintenance *ms* has been given and with an estimation algorithm (see Chapter 6) we can estimate *X(t)*. Using these figures as starting values, we can estimate $Q_o(t)$ for the oxidative growth on glucose and ethanol. This $Q_o(t)$ is now directly associated to the enzyme system related to oxygen consumption:

$$\frac{dQ_o(t)}{dt} = \frac{Q_{om} - Q_o(t)}{\tau_o} \tag{3.25}$$

such that the two related parameters Q_{om} and τ_o can be estimated as well, together with the Monod parameters K_o and K_m (see Table 3.6).

Table 3.6 *The estimated parameters of the batch phase simulation model*

Parameter	S_0 = 30 [g.dm^{-3}]	S_0 = 5 [g.dm^{-3}]
K_s	$2.4 \cdot 10^{-2}$	$1.1 \cdot 10^{-2}$
K_o	$1.2 \cdot 10^{-6}$	$1.2 \cdot 10^{-6}$
K_e	$1.1 \cdot 10^{-3}$	$1.1 \cdot 10^{-3}$
K_n	$6.43 \cdot 10^{-5}$	$6.43 \cdot 10^{-5}$
K_i	$2.4 \cdot 10^{-3}$	$2.4 \cdot 10^{-3}$
K_m	0.138	0.164
Q_{sm}	1.03	1.20
Q_{om}	0.2244	0.24
Q_{em}	0.17	0.17
τ_s	2.65	1.90
τ_o	3.07	1.90
τ_e	0.67	0.67

The same can be done with the parameters related to the Q_e, where we know that the yeast grows on ethanol only Eq. (3.3), that is the change from the reductive growth on glucose to oxidative growth on ethanol, the diauxic phase. During this period holds (from the *OUR* and *CPR* equations):

$$OUR(t) = \left[k \cdot Q_e(t) + \frac{ms}{\alpha}\right] X(t) \tag{3.26}$$

and

$$CPR(t) = \left(m \cdot Q_e(t) + ms\right) VX(t) \tag{3.27}$$

such that with the knowledge obtained from the previous estimation we calculate Q_e(t) and then estimate the ethanol related parameters Q_{em}, τ_e, K_e, and K_i.

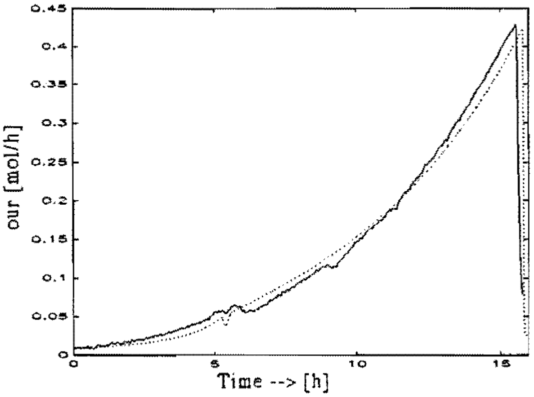
As we have only ethanol production during the reductive growth on glucose, we can use the *CPR* to calculate the Q_s(t) using (3.24) and:

$$CPR(t) = \left[\left[\frac{c}{a} - \frac{h}{a}\right] \cdot Q_o(t) + h \cdot Q_s(t) + ms\right] \cdot X(t) \tag{3.28}$$

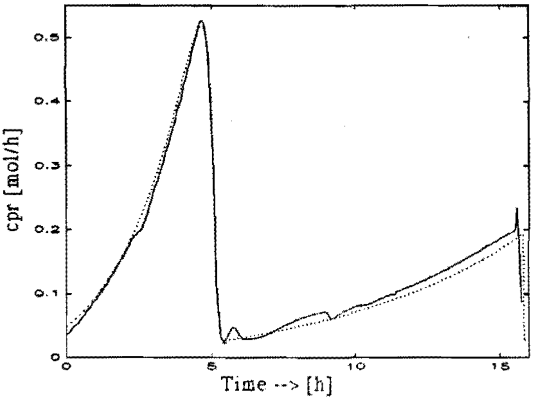
Once again we can estimate the parameters related to Q_s(t); Q_{sm}, τ_s, and K_s. Furthermore, some Monod related parameters, K_j, j=e,s,n are best estimated when transients occur in the batch phase. This is around the diauxic phase and at the end of the batch.

All these estimates are combined and the outputs of the simulation model are compared with the process outputs. This is done in Example 3.4 for the 5 [g.dm^{-3}] experiment and the 30 [g.dm^{-3}] experiment.

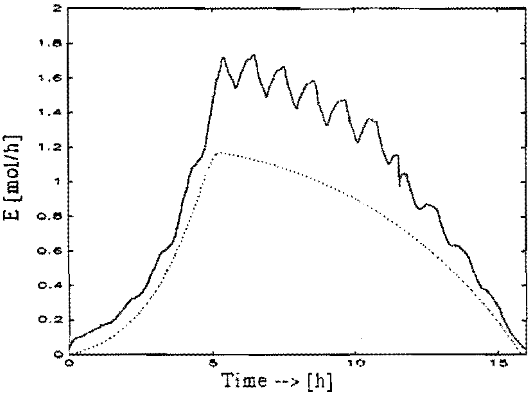
Example 3.4 *Experimental batch fermentations and the simulation model responses*



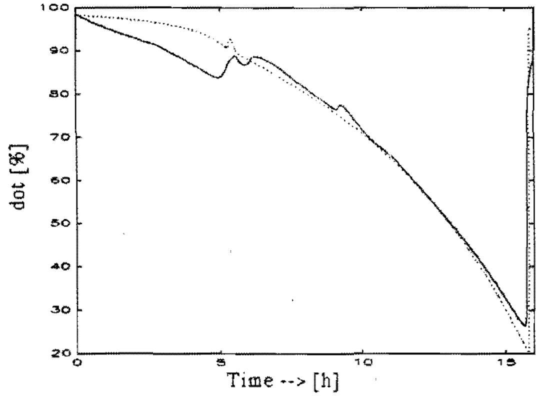
(a)


cpr [mol/h]
0.5
0.4
0.3
0.2
0.1
0
0
5
10
15
Time --> [h]

(b)


E [mol/h]
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
5
10
15
Time --> [h]

(c)

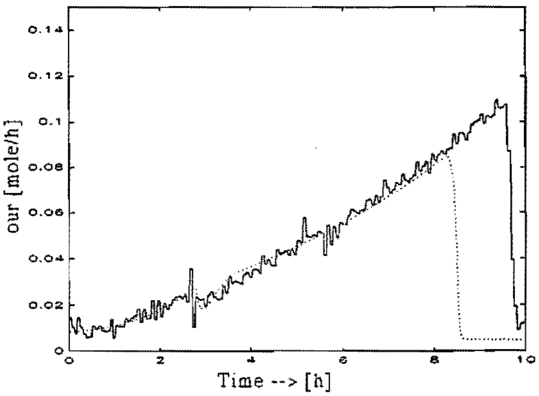

dot [%]
100
90
80
70
60
50
40
30
20
0
5
10
15
Time --> [h]

(d)

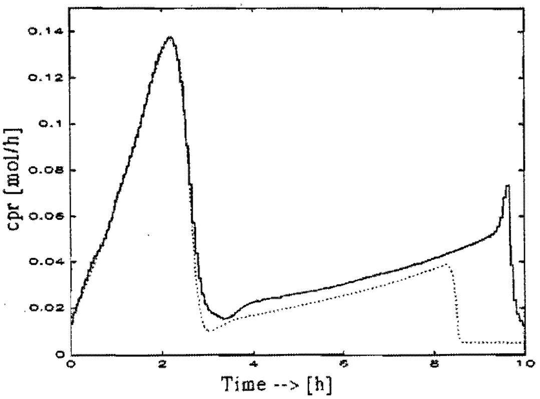
Figure 3.6 *Experimental batch fermentation results of Saccharomyces cerevisiae in synthetic medium with S_0 = 30 [g.dm^{-3}]; the solid line is the measured output, the dotted line is the model output, part (a) is the oxygen uptake rate versus time, part (b) is the carbon-dioxide production rate versus time, part (c) is the ethanol concentration versus time, and part (d) is the dissolved oxygen tension versus time*

Figure 3.6 depicts the four outputs for the S_0 = 30 [g.dm^{-3}] experiment, measured and estimated. The agreement for the *OUR* and *CPR* is quite good, which is not surprising as these two outputs have been used for minimization of the least squares algorithm. The agreement for the *DOT* is less, especially in the first part of the batch phase where the simulation model predicts less oxygen consumption than in the experiment. The second part of the *DOT* curve is as good as the *OUR* and *CPR* curves. The ethanol curve is a factor smaller than the process output; this factor could be due to erroneous calibration.

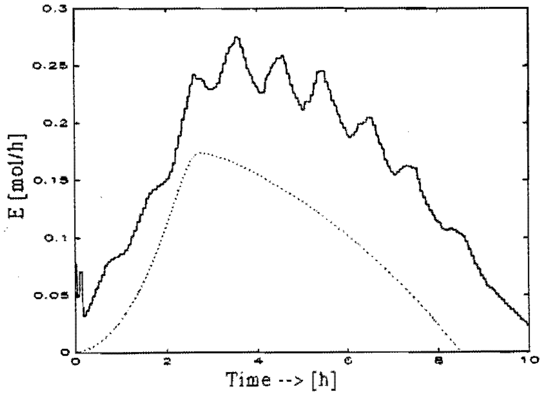
The simulation model performs poorly around the diauxic phase, seen for all outputs. The diauxic phase is enlarged due to the presence of the extracellular components and as these are not modelled by equations, the simulation model is not capable to simulate the events caused by the extracellular components. Most of the comments given on the S_0 = 30 [g.dm^{-3}] experiments are also valid for the S_0 = 5 [g.dm^{-3}] experiment. Here we see, Figure 3.7, an additional disagreement; around 8.5 [h] the ethanol has been consumed and the culture stops growing. This erroneous results could be due to a wrong yield for ethanol production on glucose or an incorrect yield for oxidative ethanol consumption.



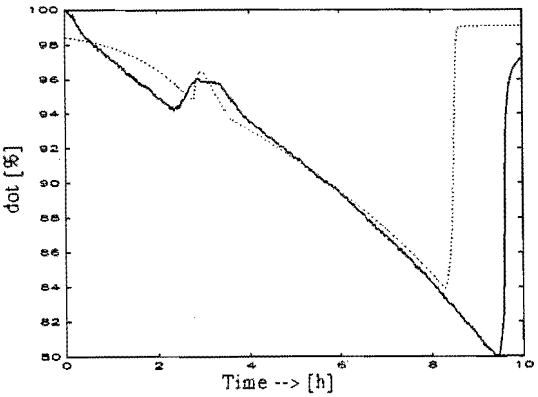
(a)


cpr [mol/h]
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0
2
4
6
8
10
Time --> [h]

(b)


E [mol/h]
0.3
0.25
0.2
0.15
0.1
0.05
0
0
2
4
6
8
10
Time --> [h]

(c)


dot [%]
100
98
96
94
92
90
88
86
84
82
80
0
2
4
6
8
10
Time --> [h]

(d)

Figure 3.7 *Experimental batch fermentation results of Saccharomyces cerevisiae in synthetic medium with* $S_0 = 5$ *[g.dm*$^{-3}$*]. The solid line is the measured output, the dotted line is the model output, part (a) is the oxygen uptake rate versus time, part (b) is the carbon-dioxide production rate versus time, part (c) is the ethanol concentration versus time, and part (d) is the dissolved oxygen tension versus time*

The oxidative yield of biomass on glucose is correct as the models enters the diauxic phase at the same time as the process. As it is believed that no or few extracellular components are present during this batch phase, the diauxic phase should be modelled by the differential equations related to the regulatory enzymes. This agreement is fairly good as is seen from the *DOT* and *CPR* plots. ■

For both simulation models the relative sum of squared errors is calculated and denoted in Table 3.7. The relative error for the ethanol is large, but can be postulated, as mentioned in Example 3.4, as a calibration error. The rest is acceptable, except for the *OUR* for the $S_0 = 5$ [g.dm^{-3}] experiment. If we compare the parameter values found in Table 3.6, one can see that the regulatory enzyme related parameters differ, especially the time constants. The difference is due to the fact that the extracellular components in the $S_0 = 30$ [g.dm^{-3}] experiment influence the diauxic phase, making it last longer. Consequently the related time constants are bigger.

Table 3.7 *Relative sum of squared errors for both simulation models*

	$S_0 = 30$ [g.dm^{-3}]	$S_0 = 5$ [g.dm^{-3}]$^{1)}$
OUR	1.51 %	25.78 % $^{2)}$
CPR	4.39 %	1.06 %
E	9.34 %	20.34 % $^{2)}$
DOT	0.45 %	$7 \cdot 10^{-3}$ %

1) The relative error is only calculated over the period up till where ethanol has been consumed. The total relative error is worse, but not comparable any more with the $S_0 = 30$ [g.dm^{-3}] experiment.
2) The high relative error is due to the high noise level on the *OUR* and ethanol measurements.

Due to the fact that the diauxic phase is modelled as the interplay of three regulatory enzyme systems, associated with three differential equations in the model (Sweere, 1988), the number of initial conditions for the batch phase simulation model is enlarged from 4 to 7. Therefore the initial conditions of the enzyme related differential equations have to be estimated as well (Table 3.8). This is done along with the other 12 parameters.

The initial value of Q_e is not important as $Q_e(t)$ stays (or goes to) zero during the fermentative part of the batch phase and can be set to 0. The additional enzyme initial conditions make it very difficult to evaluate this model on other batch data, because the initial conditions are data dependent, as can be seen from the two batch experiments used so far.

Table 3.8 *Initial conditions for the batch phase simulations models*

Initial states	S_0 = 30 [g.dm^{-3}]	S_0 = 5 [g.dm^{-3}]
X(0)	0.044	0.044
S(0)	132/198	22/198
E(0)	0	0
DOT(0)	$23.85 \cdot 10^{-5}$	$23.85 \cdot 10^{-5}$
$Q_s(0)$	0.563	0.1625
$Q_o(0)$	0.189	0.1755
$Q_e(0)$	0	0

3.4 The Fed-Batch Model

This Section will extend the batch phase simulation model described in Section 3.3 to a fed-batch phase simulation model. The kind of alterations that are suggested can be found in the first subsection. The simulation model that will be extracted from the batch phase simulation model is a nonstructured, nonsegregated one. The following subsection will, as with the batch phase simulation model, describe the experiments and the estimation/validation of the simulation model.

The Simulation Model

The assumptions for the batch phase simulation model based on the theory of Sonnleitner and Käppeli (1986) are still valid in the fed-batch phase simulation model. The main difference between the batch phase and the fed-batch phase simulation model is the glucose feed rate *GF(t)*. In the case of the batch phase simulation model the volume did not change, in the fed-batch phase simulation model it does, such that the model given by Eq. (3.7) - (3.16) has to be changed. Inclusion of the volume is not the only difference; the diauxic phase, clearly visible in the S_0 = 30 [g.dm^{-3}] batch phase experiment, and to a lesser extent in the S_0 = 5 [g.dm^{-3}] batch phase experiment, will not occur unless extreme overfeeding occurs.

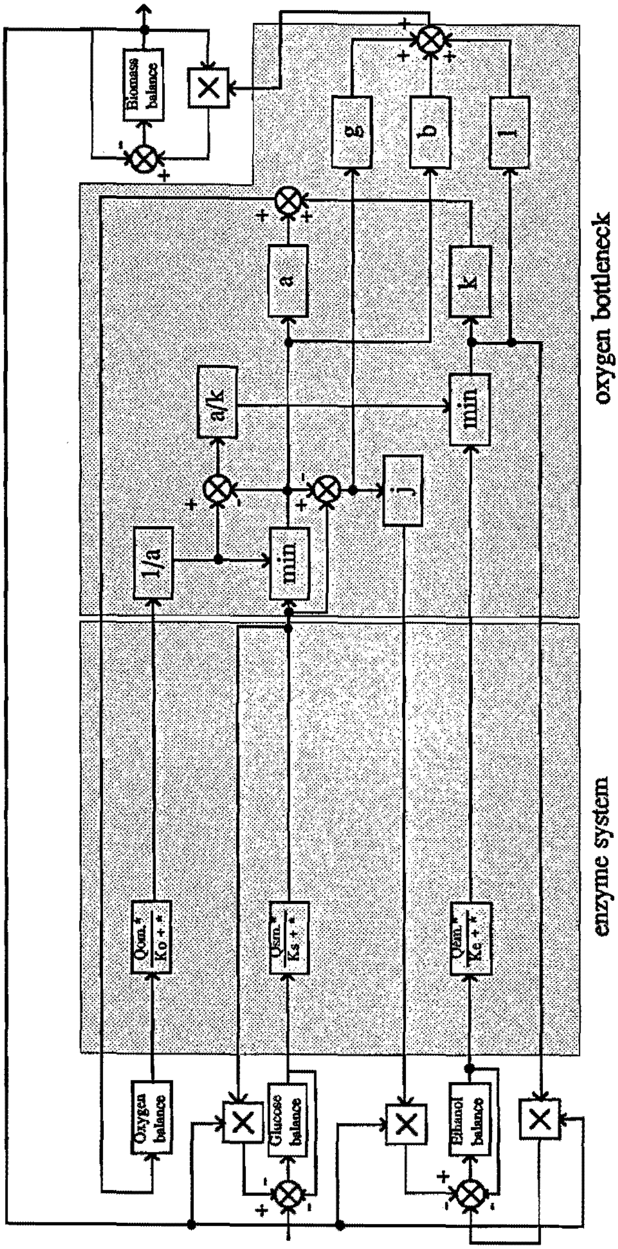

Biomass balance
g
b
l
a
k
a/k
1/a
min
j
min
Qom·*/(Ko + *)
Qsm·*/(Ks + *)
Qem·*/(Ke + *)
Oxygen balance
Glucose balance
Ethanol balance
oxygen bottleneck
enzyme system
+
-

Figure 3.8 *A schematic overview of part of the fed-batch phase simulation model*

The adaptation of Sweere (1988), concerning the regulated enzyme production, will be changed for the fed-batch phase simulation model. The parameters related to the lag phase, τ_s, τ_o, and τ_e, together with the differential equations associated with them can be discarded. The three maximum specific uptake rates Q_{sm}, Q_{om}, and Q_{em} are now constant during each simulation. All environmental settings discussed in Section 3.3 are valid for the fed-batch phase as well. The settings do change when high concentrations of yeast are reached at the end of the fed-batch phase, as viscosity starts to play a role and homogeneous mixing can not be assured any more. Then temperature fluctuations are larger, due to the biological oxidation and the higher stirrer speed, but are kept within the limits as denoted in Section 2.2.

For sake of simplicity these effects are noticed, but not taken into account for the fed-batch phase model. Furthermore, it is noticed that during the first two hours of a normal fed-batch phase run the yeast has to adapt to the new substrate (change from ethanol to glucose) and does not behave the way it does during the rest of the time. A possible explanation for this could be the reactivation or production of certain enzymes related to oxidative growth. In this model, these phenomenon are not taken into account as the model has to serve as a simulation model for control and identification design only.

A schematic view of the differential equation is given in Figure 3.8. Like in Figure 3.3, the batch phase simulation model, only the basic equations are given. With the discussion of the Figure we will only comment on the changes made compared to the batch phase simulation model. The big change is found in the regulatory enzyme system where all differential equations are substituted by simple Monod equations. Furthermore, the glucose balance has an extra input, the glucose feed (*GF*), which makes the diagram a fed-batch one.

The model can, as with the batch phase simulation model, be described by a few differential equations and algebraic expressions. As the diauxic phase does not occur, the set of differential equations is limited to four: the biomass concentration (X), the glucose concentration (S), the ethanol concentration (E), and the volume (V). These differential equations are denoted in Eq. (3.28) - (3.31).

$$\frac{dX(t)}{dt} = \left[g \cdot Q_s^{red} + b \cdot Q_s^{oxid} + l \cdot Q_e^{oxid} - ms \right] \cdot X(t) - \frac{F(t)}{V(t)} \cdot X(t) \tag{3.28}$$

$$\frac{dS(t)}{dt} = -Q_{smax}(t) \cdot \frac{S}{K_s + S} \cdot X(t) - \frac{F(t)}{V(t)} \{S(t) - GF\} \tag{3.29}$$

$$\frac{dE(t)}{dt} = \left[Q_e^{oxid} - j \cdot Q_s^{red} \right] \cdot X(t) - \frac{F(t)}{V(t)} \cdot E(t) \tag{3.30}$$

$$\frac{dV(t)}{dt} = GF(t) \tag{3.31}$$

To go with this set of differential equations the same set of algebraic equations as reported with the batch phase simulation model are valid. That is Eq. (3.13) - (3.19).

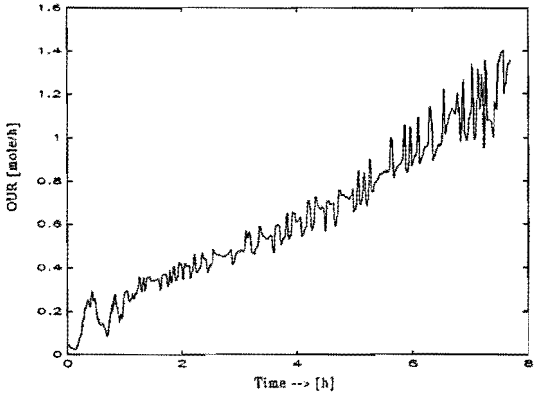
Relevant process inputs and outputs are chosen again. There are three inputs: the ethanol concentration (*E*), the dissolved oxygen tension (*DOT*), the oxygen uptake rate (*OUR*), and the carbon-dioxide production rate (*CPR*). Implemented in a simulation environment, the simulation model contains 6 parameters, 3 of which are related to Monod kinetics (K_j, $j = o,e,s$), a maintenance term (*ms*), and 3 that are related to the oxygen transfer from gas to liquid (C_j, j=*DOT,SS,AF*). These parameters have to be estimated and their estimated values have to be validated. In order to do so, suitable experiments are designed.

Experiment Design

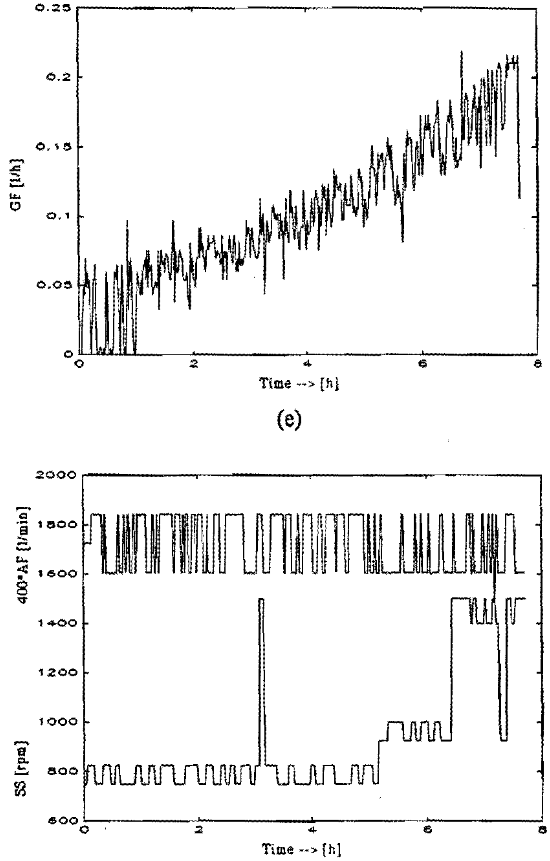
The experiments to identify parameters for the fed-batch phase simulation model are different from the batch phase simulation model experiments as presented in Example 3.3 and 3.4. In the case of the fed-batch phase we are able to excite all inputs. Furthermore, it is noted that during the fed-batch phase the process changes are quick. These process changes are: different growth rate and ethanol production ⇔ ethanol consumption. The next Example will show an identification set.

Example 3.5 *Experimental fed-batch phase data using glucose as feed*

From experiment-design and modelling theory (Section 3.2) we know that PRBS sequences with a specified frequency and pulse width will excite our system over the bandwidth specified by the process to be modelled. Normally simultaneous excitation of all inputs with mutually independent PRBS-sequences will provide us with a rich data-set, suitable for minimization (see Figure 3.6). However, from the qualitative process analysis, we learned that the total process exists of a "fast" physical part, associated with the stirrer speed, air flow, and dissolved oxygen tension, and a "slow" physiological part associated with biomass growth, etc.



(a)



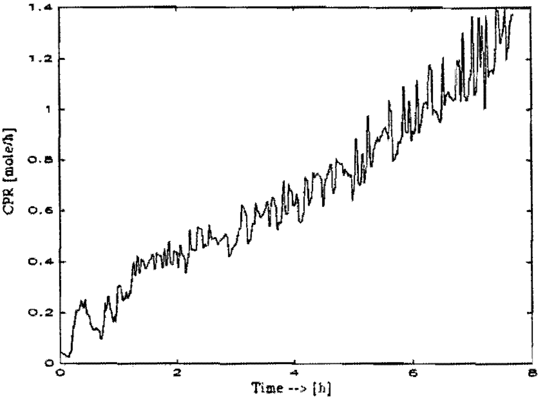
(e)

(f)

Figure 3.9 *Experimental fed-batch phase data of Saccharomyces cerevisiae with synthetic media as feed, part (a) is the oxygen uptake rate, part (b) is the carbon-dioxide production rate, part (c) is the dissolved oxygen tension, part (d) is the ethanol, part (e) is the glucose flow and part (f) the air flow and stirrer speed*

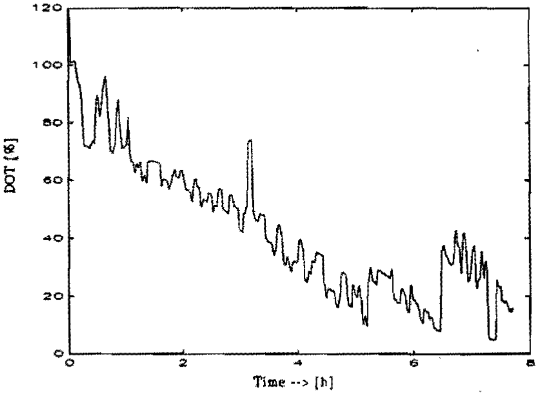
In order to carry out relevant experiments we have to distinguish between those two parts and design appropriate input sequences. The parameters related to the physical part, associated with the K_La, can be identified as in the batch phase simulation model (see Section 3.3). During the identification of the physiological parameters, the input sequences of the stirrer speed and air flow should either be constant, but high enough to ensure high dissolved oxygen levels, or increasing in accordance to the amount of biomass. ■

The physiological part of the yeast process is only influenced by glucose flow changes. Only the oxygen suppliers, *SS* and *AF* influence the biomass growth and thus the physiological system. In order to excite the process in the best way PRBS signals are



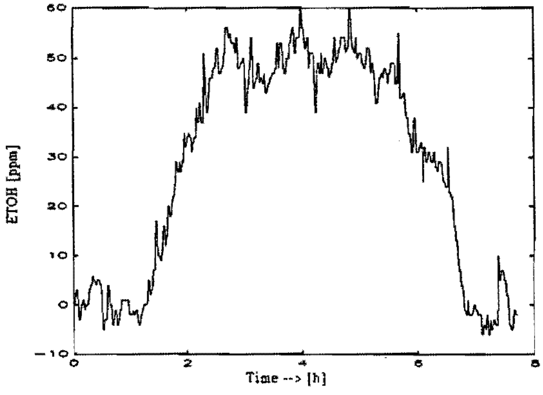
CPR [mole/h]
1.4
1.2
1
0.8
0.6
0.4
0.2
0
2
4
6
8
Time --> [h]

(b)



DOT [%]
120
100
80
60
40
20
0
2
4
6
8
Time --> [h]

(c)



ETOH [ppm]
60
50
40
30
20
10
0
−10
2
4
6
8
Time --> [h]

(d)

used. The PRBS signals settings are described in Table 3.9.

Table 3.9 *The glucose feed rate PRBS-settings*

T_s	1	[min]
T_{PRBS}	4	[min]
T_f	9	[h]
μ_{nom}	0.15	[h^{-1}]

Estimation of the Parameters

The parameters of the fed-batch phase simulation model are identified using least-squares methods. Again minimization is done using the *OUR* and *CPR* outputs. In the estimation of the parameters of the batch phase simulation model we used fed-batch phase data to come up with the K_La related parameters. The parameters values are still valid for the fed-batch phase simulation model and we use the parameters listed in Table 3.5. The rest of the parameters have to be identified, yet we can distinguish between ethanol free phases and ethanol consumption or production phases, such that the number of parameters to be estimated simultaneously is reduced.

Firstly, we take as the initial set of parameter values the parameters found with the 5 [g.dm^{-3}] experiment. Then we estimate the parameters as two sets, one when ethanol production and consumption occurs, and one when there is only oxidative growth on glucose. Then the final minimization is performed on the complete data set. The results of this minimization are found in Table 3.10.

Table 3.10 *The parameters of the fed-batch phase simulation model*

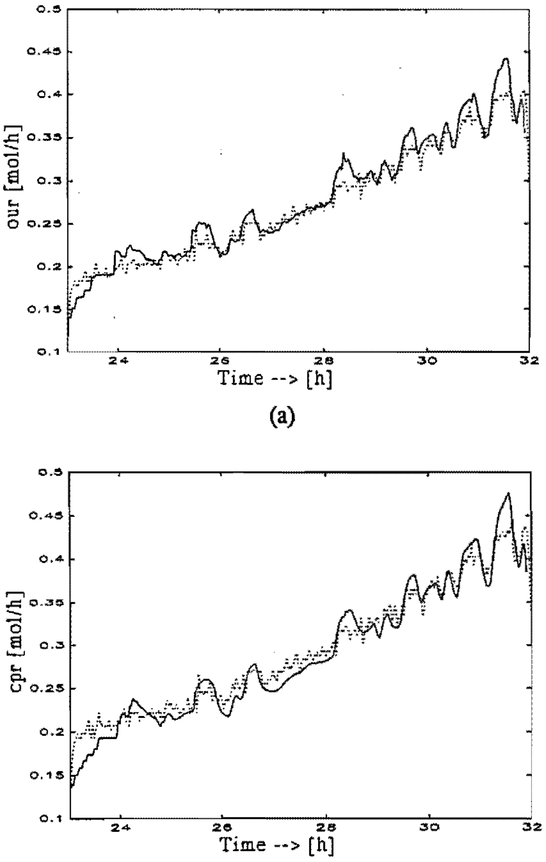
Parameter	Fed-batch phase
K_s	$5\cdot10^{-4}$
K_o	$3.4\cdot10^{-4}$
K_e	$2.2\cdot10^{-3}$
Q_{sm}	0.64
Q_{om}	0.1485
Q_{em}	0.13

Validation Set

The simulation model is then tested on experimental fed-batch phase data used as validation set. The next Example reports on this validation.

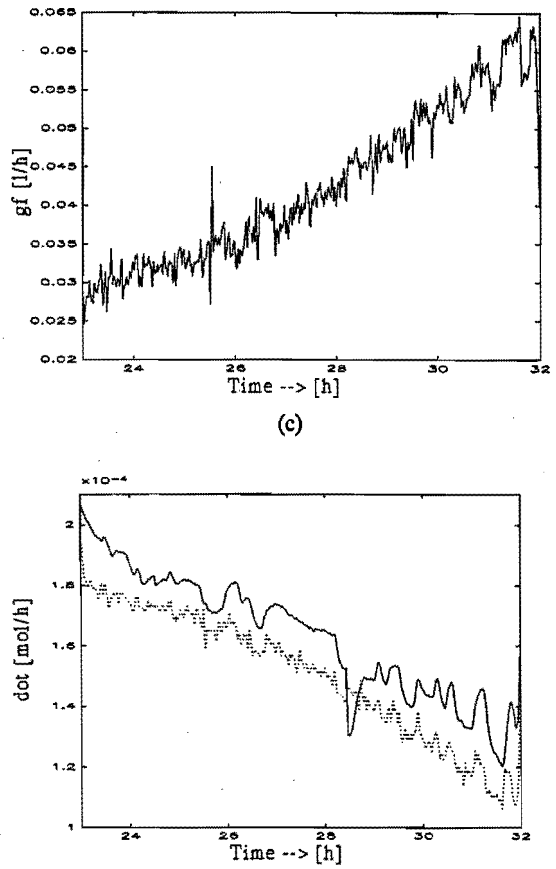
Example 3.6 *Experimental fed-batch phase data used as validation set for the fed-batch phase simulation model*

The most used situation for the fed-batch simulation model will be the oxidative growth on glucose with eventually a little ethanol production. An experiment has been carried out in which there is no ethanol production, so only pure oxidative growth on glucose occurred. The model response together with the real data are depicted in Figure 3.10. Together with the outputs the glucose flow is given. The glucose flow, an input, is the reason why all simulation model outputs seem to have a high 'noise' level. The glucose flow is reconstructed from the weight measurements of a balance, thus introducing quantization noise and measurement noise not seen by the real process. The fact that the laboratory process effluent gas-measurements *OUR* and *CPR* are rather noise free despite the noise on the *GF* input is due to the fact that the head space of the fermenter acts as a low pass filter.



(a)

(b)



(c)

(d)

Figure 3.10 *Experimental fed-batch phase data of Saccharomyces cerevisiae with synthetic feed; this is a validation set, where the solid line represent the measured data and the dotted lines the simulation model response; note that the delay has been corrected for, part (a) is the oxygen uptake rate, part (b) is the carbon-dioxide production rate, part (c) is the glucose flow, and part (d) is the dissolved oxygen tension*

The 'noise' due to the glucose flow is completely filtered. As one can see, the response of the *DOT* is adequate except for an offset. This is due to the fact that the geometry of the fermenter used is different from the K_La estimation data such that the K_La value is different. We do not account for this as the model is a only a (rough) generic simulation model, not intended to present one specific fermenter geometry only. The agreement for the *OUR* and *CPR* is good. In the beginning there is small discrepancy, due to the start up of the fed-batch phase and the fact that the culture has to adapt. At the end some of the positive steps are not followed correctly. This

> disagreement is due to not modelled dynamical effects in the culture which can not be accounted for by this fed-batch phase simulation model. The dynamical effects are probably due to degradation of trehalose (Keulers, *et al.*, 1993). ■

The resulting fed-batch phase simulation model is of order 4 (S,X,E,V), has 3 inputs (*AF,SS,GF*), and 4 outputs (*OUR,CPR,DOT,E*).

3.5 Conclusions

The laboratory process can be divided into a physical part (*DOT, SS*, and *AF*) and a physiological part (X, S, E, *GF, OUR*, and *CPR*). The physical part is 'fast' compared to the physiological part which is 'slow'. During any identification, for a simulation or a control model, this knowledge should be exploited when designing experiments.

The batch phase simulation model is based on physiological laws and the physiological parameters associated are identified using experimental batch phase data. The agreement between the experimental batch phase data and the simulation model data is quite good for both the *OUR* and *CPR*. The agreement between measured and simulated ethanol is less due to incorrect parameters and noisy measurements. The *DOT* is modelled reasonably but shows some deviations during reductive growth on glucose. Both a S_0 = 5 and 30 [g.dm^{-3}] batch phase experiment are modelled. It is noticed that the diauxic phase is modelled reasonably by the regulatory enzyme system. Due to non-modelled extracellular components the agreement of the diauxic phase for the 30 [g.dm^{-3}] experiment is not good.

The fed-batch simulation model is a stripped version of the batch phase model as the need for the regulatory enzyme system is no longer required. The regulatory enzyme system governs the diauxic phase which does not occur during the fed-batch phase. The agreement of the outputs *OUR* and *CPR* is remarkably good; only at the beginning of the fed-batch phase is there a disagreement due to a non-modelled adjustment phase. Some non-modelled events like the degradation of trehalose during a positive step on the glucose give an erroneous simulation model response. The *DOT* is also in good agreement, except for an offset due to (changing) fermenter geometry.

References

Backx, A.C.P.M., and A.A.H. Damen (1989). Identification of Industrial MIMO Processes for Fixed Controllers. Part 1 General Theory and Practice. *Journal A*, Vol. 1, pp. 3-12.

Bailey C.M. and H. Nicholson (1989). A New Structured Model for Plant Cell Culture. *Biotechnology and Bioengineering*, Vol. 34, pp. 1331-1336.

Bajpai R.K. and M. Reuß (1980). A Mechanistic Model for Penicillin Production. *Journal of Chemical Technology and Biotechnology*, Vol. 30, pp. 332-344.

Barford J.P. (1989a). A General Model for the Aerobic Yeast Growth: Batch Growth. *Biotechnology and Bioengineering*, Vol. 35, pp. 907-920.

Barford J.P. (1989b). A General Model for the Aerobic Yeast Growth: Continuous Culture. *Biotechnology and Bioengineering*, Vol. 35, pp. 921-927.

Barford J.P. and R.J. Hall (1981). A Mathematical Model for the Aerobic Growth of

Saccharomyces cerevisiae with a Saturated Respiratory Capacity. *Biotechnology and Bioengineering*, Vol. 23, pp. 1735-1762.

Bellgardt, K.H. and J.Q. Yuan (1992). Process Models: Optimization of Yeast Production. - A Case Study. Chapter 12. *Biotechnology*, Vol. 4, Rehm and Reed.

Bellgardt, K.H., Kuhlmann, W. and H.D. Meyer (1982). Deterministic Growth Model of *Saccharomyces cerevisiae*, Parameter Identification and Simulation. *IFAC Symposium, Modelling and Control of Biotechnical Processes*, Helsinki, Finland, pp. 67-74.

Bellgardt, K.H., Yuan, J.Q., Jiang, W.S. and W.D. Deckwer (1991). Optimum Quality Control of a Bakers' Yeast Production Process. *Proceedings of the ECC*, Grenoble, France, pp. 230-235.

Bergter, F. and W.A. Knorre (1972). Computersimulation von Wachstum und Produktbildung bei *Saccharomyces cerevisiae. Zeitschrift für Allgemeine Mikrobiologie*. Vol. 12, No. 8, pp. 613-629.

Cazzador, L. and L. Mariani (1988). A Simulation Program Based on a Structured Population Model for Biotechnological Yeast Processes. *Applied microbiology and biotechnology*, Vol. 29, pp. 198-202.

Cazzador, L. and L. Mariani (1989). Structured Modelling and Parameter Identification of Budding Yeast Populations. In: *Computer Applications in Fermentation Technology: Modelling and Control of Biotechnological Processes*, (Eds. Fish, N.M. *et al.*) London-New York: Elsevier Applied Science, London - New York, pp. 211-216.

Choi, C.Y. and S.Y. Park (1981). The Parametric Sensitivity of the Optimal Fed-Batch Fermentation Policy. *Journal of Fermentation Technology*, Vol. 59, No. 1, pp. 65-71.

Coppella, S.J. and P. Dhurjati (1990). A Mathematical Description of Recombinant Yeast. *Biotechnology and Bioengineering*, Vol. 35, pp. 356-374.

Dantigny, P., Ninow, J.L., Marc, I., and J.M. Engasser (1989). Representation of Changes in the Metabolic Pattern of Bakers' Yeast from Measurements of Extracellular Pyruvate, Acetate, Acetaldehyde and Ethanol. *Biotechnology letters*, Vol. 11, No. 7, pp. 515-520.

Dantigny, P., Ziouras, K. and J.A. Howell (1992). A Structured Model of Bakers' Yeast Fed-Batch Growth. *IFAC Symposium, Modelling and Control of Biotechnical Processes*, Keystone, Colorado, USA, pp. 223-226.

Dhurjati, P.S. and R.J. Leipold (1990). Biological Modelling, Chapter 8 of: *Computer Control of Fermentation Processes*, CRC Press, D.R. Omstead (ed.), pp. 207-220.

Enfors, S.O., Hedenberg, J. and K. Olsson (1990). Simulation of the Dynamics in the Bakers' Yeast Process. *Bioprocess Engineering*, Vol. 5, pp. 191-198.

Esener, A.A., Roels, J.A. and N.W.F. Kossen (1981). Fed-Batch Culture: Modeling and Application in the Study of Microbial Energetics. *Biotechnology and Bioengineering*, Vol. 23, pp. 1851-1871.

Espie, D. and S. Macchietto (1989). The Optimal Design of Dynamic Experiments. *AIChE Journal*, Vol.35, No. 2, pp. 223-229.

Fukuda, H., Shiotani, T., Okada, W. and H. Moriskawa (1978). A Mathematical Model of the Growth of Bakers' Yeast Subject to Product Inhibition. *Journal* *of Fermentation Technology*, Vol. 56, No. 4, pp. 361-368.

Hall, R.J. and J.P. Barford (1981). Simulation of the Integration of the Internal Energy Metabolism and the Cell Cycle of *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering*, Vol. 23, pp. 1763-1795.

Heijnen, J.J. and J.A. Roels (1981). A Macroscopic Model Describing Yield and Maintenance Relationships in Aerobic Fermentation Processes. *Biotechnology and Bioengineering*, Vol. 23, pp. 739-763.

Heinzle, E., Dunn, I.J., Furukawa, K. and R.D. Tanner (1982). Modelling of Sustained Oscillations Observed in Continuous Culture of *Saccharomyces cerevisiae*. *IFAC Symposium, Modelling and Control of Biotechnical Processes*, Helsinki, Finland, pp. 57-65.

Keulers, M., Ariaans, L. and M.L.F. Giuseppin (1993). The Dynamic Step Response of a Fed-Batch Bakers' Yeast Process. *Preprints of the 6th European Congress on Biotechnology*, Florence, Italy.

Kishimoto, M., Yamane, T. and F. Yoshida (1976). Sensitivity Analysis for Exponential Fed-Batch Culture. *Journal of Fermentation Technology*, Vol. 54, No. 12, pp. 891-901.

Lee, S.C., and Hwang, Y.B., Chang, H.N. and Y.K. Chang (1991). Adaptive Control of Dissolved Oxygen in a Bioreactor. *Biotechnology and Bioengineering*, Vol 37, pp. 597-607.

Menawat, A.S. and J. Balachander (1992). Alternate Control-Structures and Plant-Model Mismatch for Fed-Batch Bioreactors. *Journal of Chemical Technology and Biotechnology*, Vol. 55, pp. 177-184.

Munack, A. (1989). Optimal Feeding Strategy for Identification of Monod-Type Models by Fed-Batch Experiments. In: *Computer Applications in Fermentation Technology: Modelling and Control of Biotechnological Processes*, (Eds. Fish, N.M. *et al.*) London-New York: Elsevier Applied Science, London - New York, pp. 195-204.

Posten, C. and A. Munack (1990). On-Line Application of Parameter Estimation Accuracy to Biotechnical Processes. *Proceedings of the ACC*, Boston, USA, pp. 2181-2186.

Okada, T., Fukuda, H. and H. Morikawa (1981). Kinetic Expressions of Ethanol Production Rate and Ethanol Consumption Rate in Bakers' Yeast Cultivation. *Journal of Fermentation Technology*, Vol. 59, No. 2, pp. 103-109.

Roels, J.A. (1983). *Energetics and Kinetics in Biotechnology*. Elsevier Biomedical Press, Amsterdam - New York - Oxford.

Saner, U., Heinzle, E. and D. Bonvin (1992). Computation of Stoichiometric Models for Bioprocess. *IFAC Symposium, Modelling and Control of Biotechnical Processes*, Colorado, USA, pp. 327-330.

Savageau, M.A. and E.O. Voit (1982). Powe-Law Approach to Modelling Biological Systems, I, II, and III. *Journal of Fermentation Technology*, Vol. 60, No. 3, pp. 221-241.

Schneider, R., Posten, C. and A. Munack (1992). Application of Linear Balance Equations in an On-Line Observation System for Fermentation Processes. *IFAC Symposium, Modelling and Control of Biotechnical Processes*, Colorado, USA, pp. 319-322.

Schügerl, K. (1982). Modelling of Biotechnical Processes. *IFAC Symposium,* *Modelling and Control of Biotechnical Processes*. Helsinki, Finland, pp. 13-31.

Shimizu, H., Takamatsu, T., Shioya, S. and K.-I. Suga (1989). An Algorithm Approach to Constructing the On-Line Estimation System for the Specific Growth Rate. *Biotechnology and Bioengineering*, Vol. 33, pp. 354-364.

Shioya, S., Shimizu, H., Ogata, M. and T. Takamatsu (1985). Simulation and Experimental Studies of the Profile Control of the Specific Growth Rate in a Fed-Batch Culture. *IFAC Symposium, Modelling and Control of Biotechnological Processes*, Noordwijkerhout, The Netherlands, pp. 79-84.

Sonnleitner B. and O. Käppeli (1986). Growth of *Saccharomyces cerevisiae* is Controlled by its Limited Respiratory Capacity: Formulation and Verification of a Hypothesis. *Biotechnology and Bioengineering*, Vol. 28, pp. 927-937.

Steinmeyer D.E. and M.L. Shuler (1989). Structered Model for *Saccharomyces ccerevisiae*. *Chemical Engineering Science*. Vol. 44, No. 9, pp. 2017-2030.

Stephanopoulos, G. and K. San (1984). Studies of On-Line Bioreactor Identification, I, II, III and IV. *Biotechnology and Bioengineering*, Vol. 26, pp. 1176-1218.

Sweere, A. (1988). Response of Bakers' Yeast to Transient Environmental Conditions Relevant to Large-Scale Fermentation Processes. *Ph.D.-Thesis*, Delft University of Technology, The Netherlands.

Szewczyk, K.W. (1989). A Model for Bakers' Yeast Growth. *Bioprocess Engineering*, Vol. 4, pp. 261-264.

Takamatsu, T., Shioya, S., Chikatani, H. and K. Dairaku (1985). Comparison of Simple Population Models in a Bakers' Yeast Fed-Batch Culture. *Chemical Engineering Science*, Vol. 40, No. 3, pp. 499-507.

Tsuchiya, H.M., Fredrickson, A.G. and R. Aris (1966). Dynamics of Microbial Cell Populations. *Advance Chemical Engineering*, Vol. 6, pp.125-132.

Vialis, C., Cheruy, A. and S. Gentil (1985). An Experimental Approach to Improve the Monod Model Identification. *IFAC Symposium, Modelling and Control of Biotechnological Processes*, Noordwijkerhout, The Netherlands, pp. 175-179.

Young, J.Y., Marino-Galarraga, M., Hong, J. and R.T. Hatch (1986). Experimental Design for Parameter Estimation from Batch Culture. *Biotechnology and Bioengineering*, Vol. 28, pp. 836-841.

Yuan, J.Q., Bellgardt, K.H., Posten, C., Jiang, W.S. and W.D. Deckwer (1992). A Combined Cell-Cycle and Metabolic Model of Bakers' Yeast and its Application in Process Optimization. *Proceedings of the ACC*, Chicago, USA, pp. 2667-2672.

4

Linear Identification

In the previous Chapter we developed simulation models for the batch and fed-batch phase of the process. These simulation models are based on physical/physiological assumptions. The models are suited for simulation purposes but are less adequate for control design, especially model based control. This is because the simulation models are non-linear (switches) and time variant (regulatory enzyme system) whereas most of the control algorithms are based on linear(ized) process models around a working point. We will build a model for control design that is linear and time invariant. This model will be built from data obtained from the culture of *Saccharomyces cerevisiae* grown in fed-batch phase conditions. This linear model can be used to design linear or estimation based control algorithms to keep the process at its set-point during the fed-batch phase. We restrict ourselves to modelling of the oxidative growth on glucose, as all other processes are not within the scope of the control strategy of maximizing the productivity during the fed-batch phase.

There are various outstanding books and articles on (practical) identification of linear models from input-output data, e.g. Eykhoff 1974; Fasol and Jörgl 1980; Isermann 1980; Ljung 1987. Most of them relate to either electrical or mechanical engineering when the authors explain theory. Wang and Stephanopoulos (1984), Stephanopoulos and coworkers (1984) and Bastin and Dochain (1990) introduced some specific identification techniques for biotechnological processes. In this thesis we will apply the 'standard' techniques from Eykhoff (1974) and Ljung (1987) together with some *a-priori* knowledge. In literature some articles describe the identification using these 'standard' techniques. We will give a short description of some of these articles.

Holmberg and Ranta (1982) discuss some practical problems concerning parameter and state estimation of microbial growth processes, both continuous and batch processes. Andersen and Jørgensen (1989) identified, based on data of a simulated continuous yeast fermentation process consisting of eight state variables, a MIMO ARMAX-model using a multivariable pseudolinear recursive extended least squares method with variable forgetting factor. This model was identified using frequency base singular value analyses methods, by linearizing the process around this working point. In another paper Andersen *et al.* (1991) used the same techniques to experimental data obtaining only a third order model of the continuous system.

Before doing any experiment design for identification, one should know that a good understanding of disturbances and process dynamics facilitates the identification and experiment design. Most of this has already been carried out in Chapter 2 and 3. Here we will analyze the simulation model to get some feeling about the process dynamics. This inquiry will be done both analytically, by looking at the ordinary differential equations of a fed-batch phase system, and by analyzing the fed-batch phase simulation model identified in the previous Chapter. There will be a difference between the process dynamics and the simulation model dynamics as the simulation model is only a description of the reality. The results of the analysis are presented in the first Section. As the process is non-linear we propose a 'non-linear compensation' to overcome one of the basic non-linearities of the system: exponential biomass growth. This growth occurs by budding (Chapter 2) and a simple scheme is given in Section 4.2 to compensate for this. The process will now, more or less, be linearized, such that an appropriate experiment design can be carried out to come up with a rich data set. Dynamic experiments based on PRBS-sequences are performed in order to identify a linear input-output model of the system around a working point. It is assumed that this working point is valid for a class of working points all in the same metabolism of the yeast cell. Special attention is paid to dynamic properties of the actuators and sensors of the system with respect to the modelling error. The experimental design is denoted in Section 4.3. Having gathered the data, the estimation of the linear input-output model can be performed. We will identify three outputs (*DOT*, *OUR*, and *CPR*) related to one input (glucose feed rate) in a SISO way. This estimation is treated in Section 4.4 together with the validation. The Chapter ends with the conclusions.

4.1 Process Analysis

The yeast culture is assumed to grow in a well-mixed reactor under substrate limited conditions. The growth rate will be near the switching point for ethanol production/consumption. Further we assume that oxygen will not be limiting. Referring to the differential equations of the fed-batch simulation model of Chapter 3.4, we simplify the model by discarding

- any volume change due to dilution
- ethanol production or consumption
- any influence of the oxygen

Furthermore, the specific growth rate will not be separated into different contributions, but treated as one, $\mu(t,S)$:

$$\frac{dVX(t)}{dt} = \mu(t,S)\cdot VX(t) \tag{4.1}$$

$$\frac{dVS(t)}{dt} = -\; q_s(t,S)\cdot VX(t) + GF(t)\cdot S_{in} \tag{4.2}$$

$$\frac{dV(t)}{dt} = GF(t) \tag{4.3}$$

We assume that the process will be operated below the critical growth rate (no ethanol production) such that only glucose consumption is taken into consideration. Furthermore, only small deviations from the glucose concentration will be applied:

$$\mu(t,S) = \mu(t) \leq \mu_{crit} \tag{4.4}$$

$$q_s(t,S) = q_s^o + \alpha\cdot S \tag{4.5}$$

We can write the differential equation for glucose as a linear equation:

$$\begin{aligned} \frac{dVS(t)}{dt} &= -q_s^o\cdot VX(t) \;-\alpha X(t)\cdot VS(t) + S_{in}\cdot GF(t) \\ \frac{dVX(t)}{dt} &= \mu(t)\cdot VX(t) \end{aligned} \tag{4.6}$$

The feed rate *GF(t)* to obtain a stationary *VS(t)* is related to the biomass growth:

$$S_{in}\cdot GF(t) = q_s(t)\cdot VX(t) = q_s^o\cdot VX(t) \tag{4.7}$$

The actual metabolism of the yeast culture will react in a few seconds whereas the cell growth and metabolic adaptation have a far larger time-constant. We can approximate the glucose dynamics by an appropriate time invariant model:

$$\tau\frac{d\Delta S}{dt} + \Delta S = 0 \tag{4.8}$$

where $\tau = 1/\alpha X$. This differential equation is elementary. Note that the glucose dynamics change during a cultivation as the biomass tends to grow. There are some factors not modelled in Eq. (4.8), such as oxygen limitation, by-products, etc. We assume that these factors are not important for the oxidative growth on glucose.

The fed-batch phase simulation model of Chapter 3 will now be examined. The model is excited with the data from the validation experiment of Chapter 3. Here, during 9 hours there are several operating points, different values of *XV(t)*, and one can linearize the model at every point. Only the point of t = 27 [h] will be linearized, as this is somewhere in the middle of the experiment. As states we select X and S, as input the glucose feed only and as output the *OUR*, *CPR* and *DOT*. This yields the following state space description:

$$\begin{aligned}\begin{bmatrix}\dot{X}\\ \dot{S}\end{bmatrix} &= A(t)\cdot\begin{bmatrix}X\\ S\end{bmatrix}+B(t)\cdot[GF]\\ \begin{bmatrix}OUR\\ CPR\\ DOT\end{bmatrix} &= C(t)\cdot\begin{bmatrix}X\\ S\end{bmatrix}+D(t)\cdot[GF]\end{aligned} \tag{4.9}$$

Note that the matrices A,B,C, and D are time variant. Figure 4.1, the Bode plot, represents the frequency response of the process. Over the frequency range of interest, 10^{-4} [rad.s^{-1}] until 1 [rad.s^{-1}], the gain as well as the phase change little. The limits on the frequency range are caused by the sampling frequency (4/60 [Hz]) and the slowest dynamical time constant, yeast growth (τ = 2 [h]). The little change in both phase and gain is confirmed by spectral analysis of different input-output data. The fact that both gain and phase do not change means that the linearized process has a transfer function equal to one, i.e. it is called static.

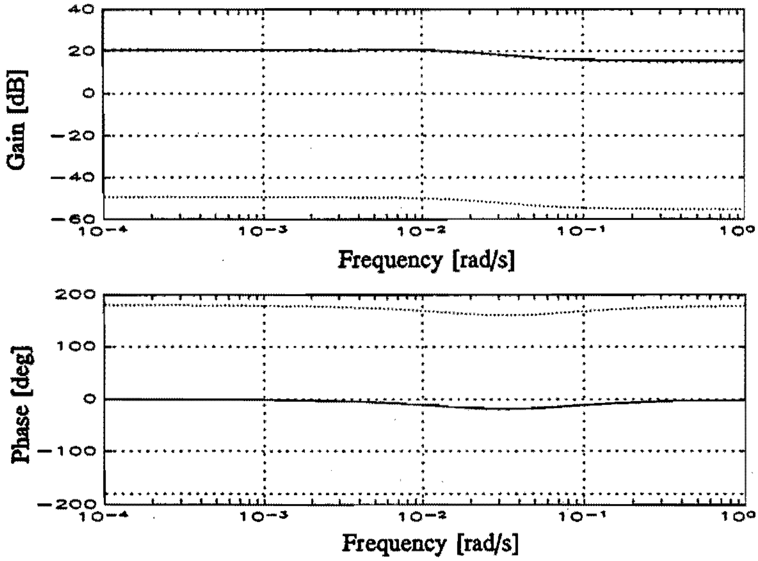


Figure 4.1 *The Bode plot of the simulation model taken at t = 27 [h] of the validation data presented in Section 3.4.; the solid and dashed lines are the OUR and CPR gains and phases, and the dotted line is the DOT gain and phase*

The phase of the *DOT* is different from the phase of the *OUR* and *CPR* due to the fact that the *DOT* has an inverse response to the *GF* compared with the *OUR* and *CPR* responses. The fact that both gain and phase differ little over the frequency range seems to contradict the findings of Axellson (1989) who found an amplitude decrease of 10^{+3} and a phase decrease of 200 over a frequency range from 10^{-3} [rad.s^{-1}] to 0.1 [rad.s^{-1}]. However his metabolism was ethanol production. From this 'static' Bode plot one can assume that the process can be modelled by low order or static models.

Preference will be given to a low order model over a static model. This is because we know from the fed-batch phase simulation model (Section 3.3) that some dynamic terms are not covered by the simulation model. Furthermore, due to the fact that the process analysis is based on fed-batch phase simulation data everything is known a-priori and the non-linear compensation (see next Section) could be calculated 'exactly'. In practice one can expect more deviations from the non-linear compensation such that building a model becomes more difficult.

The state space description is transformed to 3 SISO transfer functions, such that both states, which can not measure, are eliminated:

$$\begin{aligned}
\left(1+89.74\cdot z^{-1}+2.11\cdot z^{-2}\right)\cdot OUR &= \left(4.81\cdot z^{-1}+0.22\cdot z^{-2}\right)\cdot GF\\
\left(1+89.74\cdot z^{-1}+2.11\cdot z^{-2}\right)\cdot CPR &= \left(5.21\cdot z^{-1}+0.24\cdot z^{-2}\right)\cdot GF\\
\left(1+89.74\cdot z^{-1}+2.11\cdot z^{-2}\right)\cdot DOT &= -\left(0.16\cdot z^{-1}+7.11\cdot z^{-2}\right)\cdot GF
\end{aligned} \tag{4.10}$$

If one looks at the poles and zero, see Table 4.1, there is the same poles and zero for all transfer functions; only the factor K is different for each output. The pole and zero near the origin have almost the same value, such that the effect on the gain and phase margin is almost negligible (see Figure 4.1). The pole in 89 will be neglected for the frequency range of interest. The linearized model is a stable minimum-phase process.

Table 4.1 *Poles, zero and K-factors of the SISO transfer functions*

Output	Poles	Zero	K-factor
OUR	-89 -0.024	-0.045	481
CPR			521
DOT			-0.15

It can thus be expected that, with the proper 'non-linear compensation', the actual linear identification will yield a simple low order model.

4.2 Non-Linear Compensation

As mentioned before, we will only consider the oxidative growth on glucose. If no ethanol is present, this part of the process can be regarded as linear, as only maintenance and oxidative growth on glucose can occur. However the process itself remains highly non-linear due to the exponential character of the biomass growth. The growth is an important variable of the fed-batch process; despite its importance it can not be measured on-line. At present only off-line measurements can be done (Chapter 2) or the biomass concentration and overall growth rate can be estimated using software sensors (Chapter 6). These measurements techniques will be combined in order to compensate for the non-linearity associated with the biomass growth. This will be done in the following sense that, assuming one knows the biomass concentration and overall growth rate, one can compensate all input signals with the estimated growth of the biomass, an exponential function. This is called the 'non-linear compensation' (NLC); Axellson (1989) refers to it as the 'basic dosage scheme'.

The biomass concentration is modelled as Eq. (4.11), a solution of Eq. (4.1). The initial biomass and the expected growth rate are obtained from information supplied by estimators, dry-weight, and optical density measurements. The 'non-linear compensation' is simply related to the biomass concentration by Eq. (4.11) as an exponentially increasing function depending upon the growth rate (μ) of the biomass.

$$VX(t) = VX(0) \cdot e^{\mu(t)t} \tag{4.11}$$

Considering the input signals we also relate the glucose flow (*GF*) directly to the modelled growth Eq. (4.7). The stirrer speed and the air flow are also directly correlated with the modelled growth in order to ensure that the oxygen concentration remains high enough. The initial stirrer speed and air flow values are given in Table 2.5. With the proper 'non-linear compensation' the expected *DOT* is at a constant level and is large enough to guarantee that no oxygen limitation occurs. The carbon-dioxide rate and the oxygen output rate are expected to increase with the glucose flow as they are directly related to the amount of biomass present Eq. (3.2). Although no ethanol is assumed, there will, in practice, be a negligible production/consumption of ethanol due to some local non-homogeneties in the culture. A possible way to model these non-homogeneties is to form a multi-compartment model made from several simulation models.

4.3 Experiments

The basic idea for identifying a linear model is to add persistently exciting test signals containing enough information to the inputs of the process such that the system is excited over its whole bandwidth. Normally those test signals are added to the input signals. During the experiment the ratio of the amplitude of the persistently exciting test signal compared to the amplitude of the normal input signal remains constant. In the fed-batch phase the normal input signal is the non-linear compensation which is exponentially increasing, Eq. (4.11). The ratio would become very small at the end of the experiment. Thus the influence of the test signals on the process will become very small and consequently, the output signal response will contain little information. In order to keep the ratio and thus the influence equal during the experiment the test signals are not added but multiplied with the normal input signals. An example of an input signal applied to the process is denoted in Eq. (4.12).

$$Input(t) = NLC \cdot (1 + a \cdot testsignal(t)) \tag{4.12}$$

As test signals PRBS sequences are used. The factor a, which determines the deviation from the 'non-linear compensation', has to be small enough not to deviate from its nominal growth, as has been explained in the previous Section.

Figure 4.2 shows the model to be estimated together with the NLC. The small blocks represent the models of the actuators and sensors. In most cases the transfer function of the actuators is one within the bandwidth of the experiment. This is because of the low sampling rate of 0.0167 [h], all of the actuators and sensors have come to a stationary value in between two sample points. Only if we want to determine the dynamics of the physical part of the process, the H_{DOT} should be taken as:

$$H_{DOT}(s) = \frac{1}{\tau_{DOT}s + 1} \tag{4.13}$$

with $\tau_{DOT} = 1.57 \cdot 10^{-3}$ [h]. The τ_{DOT} is determined using a high sampling frequency, 0.2 [Hz], and sodiumdithionite ($Na_2S_2O_4$) to register the *DOT* response. The applied PRBS-signals have an amplitude of 5% of the nominal input signal. The constant *a* is 0.1, it is chosen to subtract the PRBS-signals because this ensures that no ethanol will be produced during such an experiment as $\mu < \mu_{crit}$.

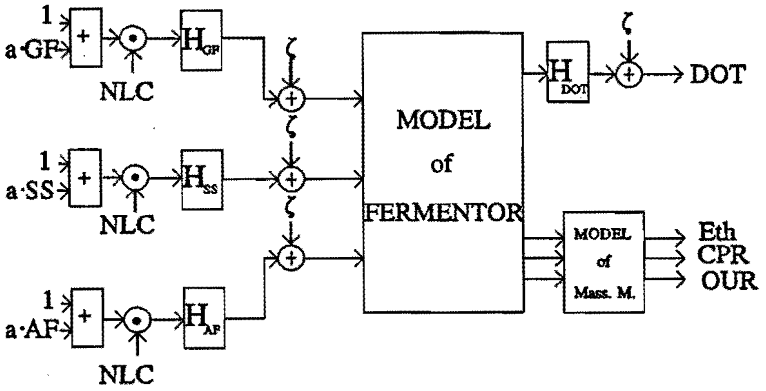


Figure 4.2 *A schematic overview of the experiment set-up*

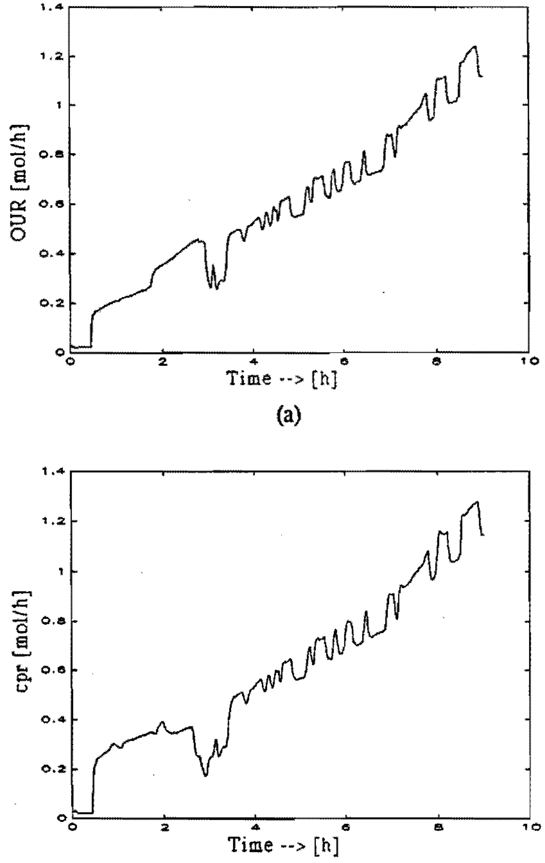
A set of experiments has been carried out to generate data for the identification of the linear model. All the experiments had the same set-up, see Example 4.1. Part of the experimental data will be used for estimation, part will be used for validation of the linear model for control. During all experiments a small amount of ethanol was produced at the beginning of the fed-batch phase. This indicates that the growth rate and/or the initial biomass of the basic dosage scheme is not correct. This is due to an erroneous setting of the estimated biomass in the non-linear compensation, Eq. (4.11). Furthermore, the culture has to adjust to new conditions, changing from growth on ethanol to growth on glucose, which causes some ethanol production during the growth on glucose. During the rest of the experiments the glucose was metabolized oxidatively. The process remained in this pathway for the rest of the experiments.

In Example 4.1 we show a experimental data set obtained using the non-linear compensation and a PRBS signal.

Example 4.1 *Experimental data using the non-linear compensation and a PRBS as input signals for the glucose flow*

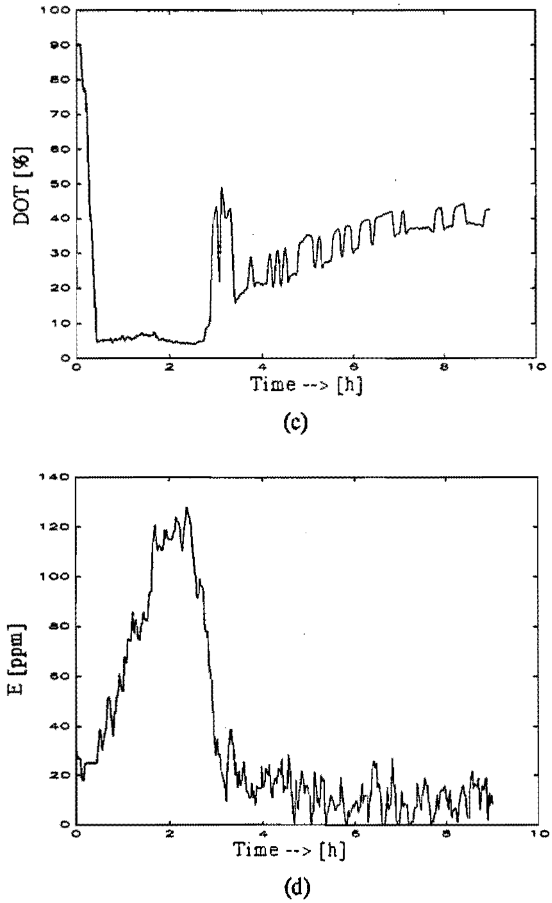
As one can see not the whole data-set of the experiment is suited for identification. As mentioned a small amount of ethanol was produced in the beginning from 0-3 [h], Figure 4.3, indicating that another pathway was followed. The rest of the experiment, from 3-9 [h], appears to contain usable data. If the yeast grows purely oxidative on the glucose, the dynamics of the off-gas outputs *OUR* and *CPR*, except for a constant, appeared to be identical. Ethanol is not produced in this pathway; the small amount of measured ethanol is seen as measurement noise.

Despite the fact that the glucose is metabolized oxidatively, the output data-sets do not have the expected characteristics. We expected a variation of the *DOT* around a steady state; instead the *DOT* mean value is increasing and the gain is decreasing. This could be due to the fact that with higher *DOT* values the gain decreases, but it seems more likely that the non-linear compensation with respect to the expected growth was incorrect.



(a)

(b)



(c)

(d)

Figure 4.3 *Fed-batch fermentation data of Saccharomyces cerevisiae with synthetic feed; part (a) is the oxygen uptake rate, part (b) is the carbon-dioxide production rate, part (c) is the dissolved oxygen tension, and part (d) is the ethanol* ■

Signal Processing

The obtained output signals are not directly suited for identification. First of all the non-linear compensation has to be removed from the input signals. This is done by simply dividing the applied input signals by the non-linear compensation, or by removing the linear trend of the logarithm of the input signals. Next the output signals have to be corrected. This is carried out in the same way as with the input signals for the off-gas measurements, and the *DOT* is corrected such that a normal signal remains.

Signal processing is performed after non-linear corrections, peaks are removed and replaced by a linear interpolation of two surrounding data-points and to prevent aliasing in the data-set, the data has to be filtered with an analog low-pass filter. As the bandwidth of the process is close to the sample Nyquist rate, a low-pass filter with a high cut-off frequency is necessary and the order of the filter has to be high in order to get a proper filtering. An eighth order Chebychev filter is chosen because of its sharp decay. The last step of the pre-processing phase is to correct the obtained data for their time-delays. The delay is estimated using correlation functions of input and output data. Note that estimation algorithms are able to estimate delays as well. It is found that due to high noise levels it is more time efficient to correct the delay before doing any estimation.

4.4 Estimation of a Linear Model

For all input-output combinations transfer functions mentioned in Section 4.1, Eq. (4.10), a linear model will be estimated by minimizing the error between the process response and the linear model response. Ethanol will not be considered as a relevant output as, during oxidative growth on glucose, no ethanol is expected to be produced or consumed. From the fed-batch simulation process analysis we know that the transfer functions of the *OUR* and *CPR* output to changes in the *GF* input are equal, except for a gain. So identification of just one transfer function, e.g. *GF* ↔ *OUR*, would be sufficient. But as it does not require much extra effort we will identify both models.

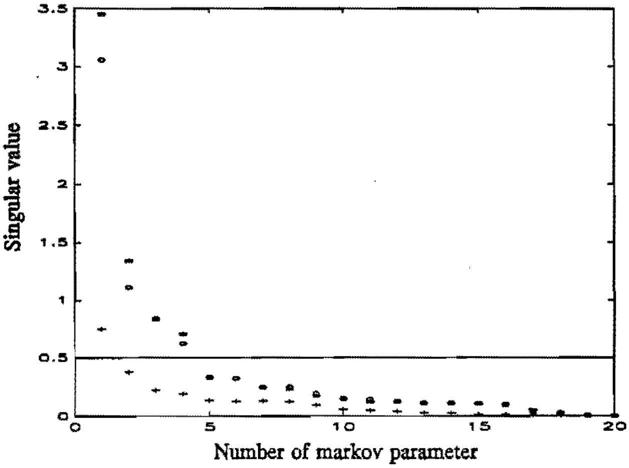


Figure 4.4 *Singular values of the estimated Markov parameters; (*) are the OUR, (o) the CPR and (+) the DOT related values*

Even the dynamical difference between the *OUR/CPR* and the *DOT* is small, but as they are measured with two different sensors a comparison could be useful. We start the estimation with the determination of the McMillan degree for all input-output data sets. The McMillan degree is based on the singular values of a finite block Hankel matrix of estimated FIR Markov parameters from input-output data (Backx and Damen, 1989). This McMillan degree represents the estimated number of states in a state space description of the process (the order of the process). The outcome is used to get an impression of the numbers of poles of the process; the zeros are determined afterwards. Figure 4.4 shows the plot of the singular values of all input-output data sets. We can see that the order of the processes (the McMillan degree) should be 4 for the *GF* ↔ *OUR* and *GF* ↔ *CPR* model and 1 for the *GF* ↔ *DOT* model. Next we continue the estimation with the determination of an initial model order for an output error estimation by choosing an ARX-model Eq. (4.14) as starting point.

$$A(z^{-1}) \cdot y(t) = B(z^{-1}) \cdot u(t-nk) + e(t) \tag{4.14}$$

The ARX-structure is chosen because of the simplicity of mathematical calculations. The model order of the resulting ARX model is chosen by means of the Akaike criterion. Note that the model order chosen did not necessarily give the lowest relative output error. The results of these initial model orders are given in Table 4.2. The second order set for the *GF*↔*DOT* denoted in Table 4.2 are the values for the uncorrected *DOT* signal.

Table 4.2 *Model orders found with ARX-structure determination*

model	n_A	n_B	rel. outp. error
GF↔*OUR*	4	4	0.1116
GF↔*CPR*	4	4	0.1261
GF↔*DOT*	4/6	4/7	0.1856

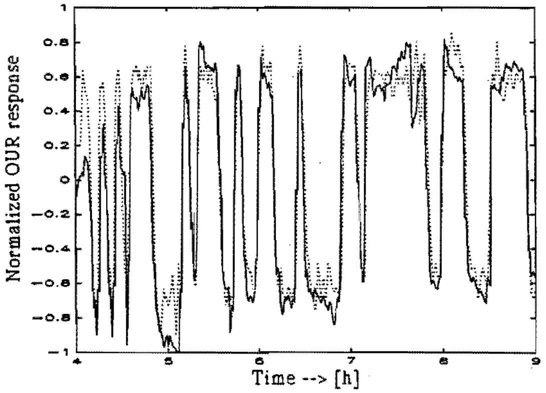
With these initial model orders output-error models Eq. (4.15) are determined.

$$y(t) = \frac{B(z^{-1})}{F(z^{-1})} \cdot u(t-nk) + e(t) \tag{4.15}$$

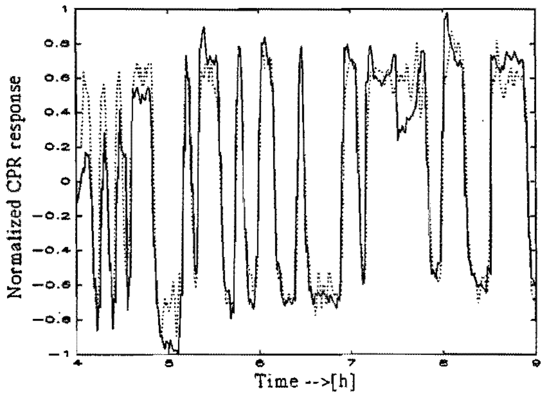
The preference is given to output error models over prediction (equation) error models as the output error model has a longer prediction horizon than the one step ahead prediction of the equation error model. The model order is varied from 1 to 4, the choice of the final model is based on the relative quadratic error of the real process minus the model output.

Example 4.2 *Estimation of linear models for control of three input-output data sets, GF ↔ OUR, GF ↔ CPR, and GF ↔ DOT*

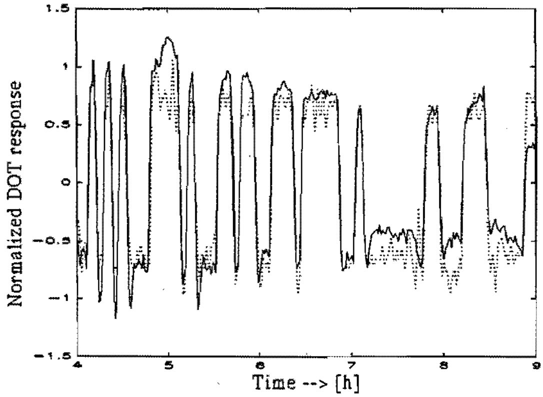
It is found that the relative output error hardly differs for the different model structures (see Table 4.3). This is in concordance with the expectations formulated with the linearized model of Section 4.1. If we look at the table, first or second order models should be enough to model the process. A set of second order model responses is given in Figure 4.5.


1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
Normalized OUR response
4
5
6
7
8
9
Time --> [h]

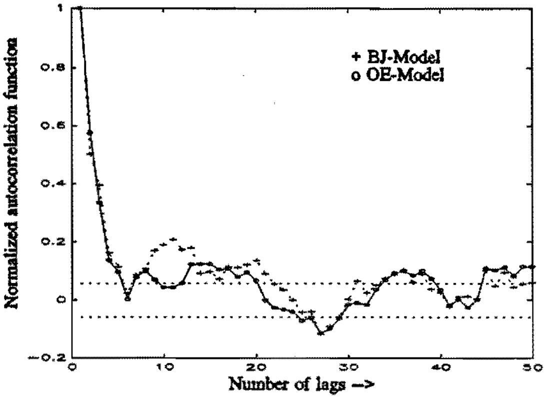
(a)


1
0.8
0.6
0.4
0.2
0
−0.2
−0.4
−0.6
−0.8
−1
Normalized CPR response
4
5
6
7
8
9
Time -->[h]

(b)


1.5
1
0.5
0
−0.5
−1
−1.5
Normalized DOT response
4
5
6
7
8
9
Time --> [h]

(c)



(d)

Figure 4.5 *Linearized and normalized experimental data and second order model response of a fed-batch phase of Saccharomyces cerevisiae with synthetic feed, part (a) is the GF ↔ OUR data and model response, part (b) is the GF ↔ CPR data and model response, part (c) is the GF ↔ DOT data and model response, part (d) the correlation of the output error of the GF ↔ OUR for OE and BJ model; the solid line (a-c) is the process response, the dotted line the model response; for (d) yields that the solid line is the OE autocorrelation of the error, the dotted line the BJ autocorrelation of the error*

The resulting output error is not white, as can be seen from Figure 4.5(d). We could apply Box Jenkins (BJ) models Eq. (4.16) to model the error as well, but they do not yield an error that is better than the one obtained with the output error method.

$$y(t) = \frac{B(z^{-1})}{F(z^{-1})} \cdot u(t-nk) + \frac{C(z^{-1})}{D(z^{-1})} \cdot e(t) \qquad (4.16)$$

As one can see from Table 4.2 the 4th order models obtained with the ARX structure have a relative output error that is less than the output error models. Note that the identification was restricted by the sample rate which could not be higher than once per minute. It is felt that an optimal sample rate for constructing good oxidative models should be once per 30-40 [s]. Furthermore, the models resulting from this identification are not suited for any other pathway, ethanol production/consumption, or oxygen limitation. ■

Table 4.3 *Relative output error of different model order for all outputs for output error (OE) and Box Jenkins (BJ) models. n_C and n_D only for BJ-models*

	OUR		*CPR*		*DOT*	
[n_F n_B n_C n_D]	OE	BJ	OE	BJ	OE	BJ
1 1 1 1	0.1348	0.2203	0.1302	0.1962	0.1456	0.2876
2 1 2 1	0.1344	0.2177	0.1304	0.1900	0.1478	0.2867
1 2 1 2	0.1328	0.4600	0.1302	0.4054	0.1437	0.4634
2 2 2 2	0.1344	0.4477	0.1260	0.4010	0.1466	0.4645
3 3 - -	0.1541	-	0.1340	-	0.1564	-
4 4 - -	0.1318	-	0.1312	-	0.1435	-

4.5 Validation of the Linear Model

A data set is used to validate the model obtained. Although the model is of low order and the results are somewhat predictable, the validation is still carried out. In order to obtain the true signals, that is not the linearized and normalized ones as shown in Figure 4.5, we have to apply the 'non-linear compensation' to the linear model signals. The validation set is given in Example 4.3.

Example 4.3 *Validation of the linear model for control for two sub-models, GF ↔ OUR, and GF ↔ CPR*

The validation data and the model responses are shown in Figure 4.6. As can be seen, the response, from 7.5 [h] to 8.5 [h] is rather good. This period is purely oxidative growth on glucose and the model performs as expected. Before this period there was a slight ethanol production/consumption phase and, quite obviously, our process is outside the model set. The *OUR* (Figure 4.6(a)) is a flat curve (process) and the *OUR*-model is an excited signal. We know that during ethanol production the maximized oxygen uptake is reached such that variations in the glucose flow do not lead to any excitation of the *OUR*. The *CPR* is excited, but differs from the pure oxidative growth on glucose as can be seen in Figure 4.6(b). At the end of the run, around 8.5 [h] and 9 [h], strong non-linearities occur that are not accounted for in the linear model, for an explanation see Example 3.6. The linear model can not follow these non-linearities and will show an offset from the process output for the rest of the run.

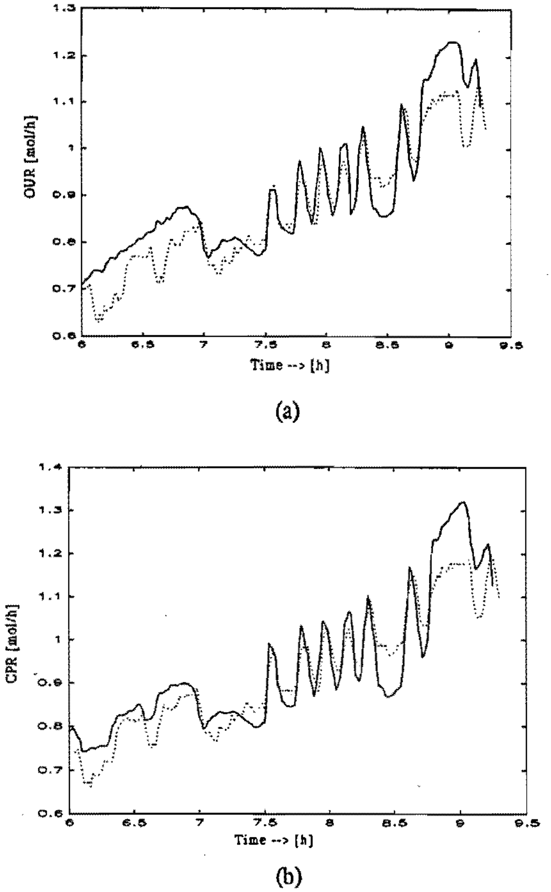


Figure 4.6 *Response of second order models of a fed-batch fermentation run of Saccharomyces cerevisiae with synthetic feed, part (a) is the GF ↔ OUR and part (b) is the GF ↔ CPR; the solid line is the process response, the dotted line the model response* ■

In case of predictive (equation) error models the offset reported in Example 4.3 would be accounted for. It would be worth while to look into non-linear modelling of this non-linear process. Various identification methods are discussed in the next Chapter.

Conclusions

It is found that the dynamics of the fed-batch phase simulation model are almost zero if the culture grows oxidatively on glucose only. The resulting models are either static or low order models. In practice the models are not static due to unexplained dynamics in the culture. The unexplained dynamics are not modelled in the fed-batch simulation model. Note that to obtain a linear model a non-linear compensation is always needed.

Applied to experimental data the non-linear compensation functions rather well, except when changing from one growth situation to another, e.g. from growth on ethanol to growth on glucose only. Using the non-linear compensation three low order models are identified, describing the oxidative growth on glucose well, except when additional dynamic effects occur. Furthermore, as expected, it is found that the resulting models can not describe the process outputs if ethanol production or consumption occurs.

The applicability of the model for process control is not very high as the input has to be corrected by a non-linear compensation, which is hard to implement in a controller if the specific growth rate and biomass concentration are unknown. A possible solution would be an estimation of both process variables, as is shown in Chapter 6. From the results of this model building we learnt that the most important and sensitive part of building a linear model is the non-linear compensation which partly linearizes the process.

References

Andersen, M.Y. and S.B. Jørgensen (1989). Identification of a Simulated Continuous Yeast Fermentation. *Computer Applications in Fermentation Technology: Modelling and Control of Biotechnological Processes*, (Eds. Fish, N.M. *et al.*) London-New York: Elsevier Applied Science, London - New York, pp. 205-209.

Andersen, M.Y., Braband, H. and S.B. Jørgensen (1991). Model Based Control of a Continuous Yeast Fermentation. *Proceedings of the ACC,* Boston, USA, pp.1329-1334.

Axellson, J.P. (1989). Modelling and Control of Fermentation Processes. *Ph.D. Thesis*, Lund Institute of Technology, Sweden.

Backx, A.C.P.M., and A.A.H. Damen (1989). Identification of Industrial MIMO Processes for Fixed Controllers. Part 1 General Theory and Practice. *Journal A*, Vol. 1, pp. 3-12.

Bastin. G. and D. Dochain (1990). *On-Line Estimation and Adaptive Control of Bioreactors*. Elsevier, Amsterdam.

Eykhoff, P. (1974). *System Identification*. John Wiley, London.

Fasol, K.H. and H.P. Jörgl (1980). Principles of Model Building and Identification. *Automatica*, Vol. 16, pp. 505-518.

Grosz, R., Stephanopoulos, G. and K.-Y. San (1984). Studies on On-Line Bioreactor Identification. III. Sensitivity Problems with Respiratory and Heat Evolution Measurements. *Biotechnology and Bioengineering*, Vol. 26, pp. 1197-1208.

Holmberg, A. and J. Ranta (1982). Procedures for Parameter and State Estimation of Microbial Growth Process Models. *Automatica*, Vol. 18, pp. 181-193.

Isermann, R. (1980). Practical Aspects of Process Identification. *Automatica*, Vol. 16, pp. 575-587.

Ljung, L. (1987). *System Identification - Theory for the User.* Prentice-Hall, Englewood Cliffs, New Yersey.

San, K.-Y. and G. Stephanopoulos (1984). Studies on On-Line Bioreactor Identification. II. Numerical and Experimental Results. *Biotechnology and*

Bioengineering, Vol. 26, pp. 1189-1197.

San, K.-Y. and G. Stephanopoulos (1984a). Studies on On-Line Bioreactor Identification. IV. Utilization of pH Measurements for Product Estimation. *Biotechnology and Bioengineering*, Vol. 26, pp. 1209-1218.

Stephanopoulos, G. and K.-Y. San (1984). Studies on On-Line Bioreactor Identification. I. Theory. *Biotechnology and Bioengineering*, Vol. 26, pp. 1176-1188.

Wang, N.S. and G. Stephanopoulos (1984). A New Approach to Bioprocess Identification and Modelling. *Biotechnology and Bioengineering Symposium*, No. 14, pp. 635-656.

5

Non-Linear Identification

Non-linear behaviour is the rule, rather than the exception, in the dynamic behaviour of biotechnical processes. The culture of *Saccharomyces cerevisiae* grown in a reactor is highly non-linear and time-varying due to its changing reactions and underlying non-linearities (Chapter 2). Contrary to chemical processes where the states are defined in terms of well defined physio-chemical properties like temperature, pressure and concentration, the state of biological processes are ill-defined in terms of physiological properties.

There are several options to model the non-linear, time-varying process. One commonly used method is to linearize the process around a working point and identify a linear model of the process as has been done in the previous chapter. The popularity of linear over non-linear models is due to the coherent description of linear processes and the simplicity of calculations.

Bastin and Dochain (1990) proposed a linearizing approach to identify the process. This approach is based on observers and state-reconstruction. In this chapter a different approach is followed by obtaining the model through a non-linear structure identification. Several non-linear structure identification methods are known, see Haber and Unbehauen (1990). Here three methods are highlighted: block-oriented models using Volterra kernels (Haber, 1989), a non-linear model linear-in-the-parameters (Billings *et al.*, 1989), and Radial Basis Functions (Chen *et al.*, 1990). The last technique resembles the use of a neural network. Various neural network applications have been reported (e.g. Linko and Zhu, 1992). Our focus will be on the Radial Basis Functions only as it is based on the same orthogonal estimator as the non-linear model linear-in-the-parameters. Furthermore, numerous researchers (Linko, Rivera) are active in the field of neural networks. All techniques are tested on the simulation model of the fed-batch fermentation process developed in Chapter 3.

The chapter is organised as follows. First the structure identification algorithms are discussed. Next, in Section 5.2, the data used to test the identification models is briefly presented. This data concerns simulation data obtained from the fed-batch model described in Chapter 3, simulation data obtained with the batch model described in Chapter 3, and experimental data from laboratory experiments. The results of using these non-linear identification techniques are presented and discussed in Section 5.3. Finally, in Section 5.4, conclusions are formulated.

5.1 Identification Methods

In this section the structure identification methods are discussed: block-oriented models using Volterra kernels (Haber, 1989), a non-linear model linear-in-the-parameters (Billings *et al.*, 1989), and Radial Basis Functions (Chen *et al.*, 1990). The algorithms are quite different in their approach to modelling the non-linear process. The block oriented method divides the process into a linear (dynamic) part and a non-linear (static) part with corresponding kernels. The non-linear model linear-in-the-parameters has both the static and dynamic part inside the non-linear description. The Radial Basis Function looks like a neural network, generating an output pattern through a non-linear function excited by a few selected input data sets. All of these algorithms will be briefly discussed. For a comprehensive description of these algorithms the reader should consult the main articles cited.

Block-Oriented Models

The basic idea of block-oriented models is that the system can be divided into several blocks: blocks with static non-linear and blocks with dynamic linear terms that are combined into a system through multipliers and summing elements. As the fermentation process is an exponentially increasing process, only quadratic terms will be discussed. The models can be divided into three groups (Haber, 1989):

1) **Hammerstein** models in which the non-linear static term is followed by a linear dynamic term.
2) **Wiener** models in which the linear dynamic term is followed by a static non-linear one.
3) **Wiener-Hammerstein** cascade models which contain dynamic terms both at input and output and a non-linear term in between.

Identification methods for the terms in the block-oriented models are the Volterra method, the frequency method, the correlation analysis method, etc. (see Haber and Unbehauen, 1990). Here only the Volterra kernels will be considered as the other methods are less applicable to the culture of *Saccharomyces cerevisiae* for various reasons, e.g. the frequency method is not applicable due to the low and small range of the process bandwidth.

The non-linear dynamic systems, having polynomial static terms, can generally be described by a continuous Volterra series which can be approximated by a discretization:

$$y(k) = h_0 + \sum_{\kappa_1=0}^{m} h_1(\kappa_1)u(k-\kappa_1) + \sum_{\kappa_1=0}^{m}\sum_{\kappa_1=0}^{m} h_2(\kappa_1,\kappa_2)u(k-\kappa_1)u(k-\kappa_2) \qquad (5.1)$$

where h_0, $h_1(\kappa_1)$ and $h_2(\kappa_1,\kappa_2)$ are the kernels.

Parameter estimation in terms of Volterra series has the disadvantage that many unknown parameters have to be estimated, but on the other hand its advantage is that it makes it possible to use a structure-independent procedure that is valid for quite different models. The kernels can be estimated using an ordinary least squares method. Due to a high degree of non-linearity and a large memory value m the solution can become numerically impossible because of a failing matrix inversion. This problem can perhaps be solved by applying a recursive least squares algorithm. Next the model structure is determined from the characteristics of the kernels. These characteristics are the graphical plots and analytical indices of the kernels (Haber, 1989). Finally the weighting filters can be calculated either from the kernels, or by means of direct parameter estimation of the parametric model of known structure. The method using Volterra kernels is a two layer identification algorithm: first the structure, then the parameters.

The estimation of the Volterra kernels makes the one-step structural identification possible, independent from the model type as defined before. Considering the laboratory process and the non-linear compensation used to linearize the process (Chapter 4), a second degree Volterra kernel description would be sufficient. An exponentially increasing function can readily be described by a linear and quadratic function. This explains the degree of two for the Volterra kernels. Together with the degree a sufficiently high memory m has to be chosen. The memory length m can be compared with the number of FIR parameters chosen. If the memory length is chosen too small, relevant parameters of the process response are not identified. A normal number for the memory length would be 10 to 20. Next we determine the model structure. From analytical measures, see next subsection, an indication about the sort of model structure, Hammerstein, Wiener, or combination, is obtained. If this information is not enough we look at the graphical representation on the first and second order degree Volterra kernels. The plots show patterns belonging to model structures, e.g. a pattern along the main diagonal indicates a Hammerstein model. Having determined the right model structure, the parameters of the parametric model of the known model structure can be estimated. This is done using ordinary least squares estimators.

Analytical Indices.

Next to plotting the kernels and deciding upon the graphical representation there are also the analytical indices (Haber, 1989):

(α) The quadratic kernels of the Hammerstein model differ from zero only at the main diagonal. This feature can be seen from the plot of $h_2(\kappa_1,\kappa_2)$ and from building the index α out of the mean square value of the off-diagonal elements divided by the mean square value of the main diagonal. α only becomes zero for the **Hammerstein** model.

$$\alpha = \frac{\dfrac{2}{m(m+1)} \sum_{\kappa_1=0}^{m-1} \sum_{\kappa_2=\kappa_1+1}^{m} h_2^2(\kappa_1,\kappa_2)}{\dfrac{1}{m+1} \sum_{\kappa_1=0}^{m} h_2^2(\kappa_1,\kappa_1)} \tag{5.2}$$

(β) The linear kernel and the quadratic kernel at the main diagonal are proportional to each other in the **simple Hammerstein** model. This fact can be recognized from the diagrams of $h_1(\kappa_1)$ and $h_2(\kappa_1,\kappa_1)$ and when β becomes zero. $\theta(\kappa_1)$ is the ratio of the second $\{h_2(\kappa_1,\kappa_1)\}$ and first-degree $\{h_1(\kappa_1)\}$ kernels, $\bar{\theta}$ is the average value of the ratio.

$$\beta = \frac{\dfrac{1}{m+1} \sum_{\kappa_1=0}^{m} (\theta(\kappa_1)-\bar{\theta})^2}{\bar{\theta}^2} \tag{5.3}$$

(γ) In the **Wiener** models the sections of the quadratic kernels parallel to the axes are proportional to each other. This is expressed in the fact that γ is zero.

$$\psi(\kappa_1,\kappa_2,\kappa) = \frac{h_2(\kappa_1,\kappa_2+\kappa)}{h_2(\kappa_1,\kappa_2)} \tag{5.4}$$

$$\begin{aligned} \gamma &= \max_{\kappa_2,\kappa} \gamma(\kappa_2,\kappa) \\ &= \max_{\kappa_2,\kappa} \frac{\dfrac{1}{m+1} \sum_{\kappa_1=0}^{m} (\psi(\kappa_1,\kappa_2,\kappa) - \bar{\psi}(\kappa_2,\kappa))^2}{\bar{\psi}^2(\kappa_2,\kappa)} \end{aligned} \tag{5.5}$$

(δ) The quadratic Volterra kernels along the main diagonal are proportional to the squares of the linear kernels in the **simple Wiener** models. Here δ becomes zero.

$$\phi(\kappa_1) = \frac{h_2(\kappa_1,\kappa_1)}{h_1^2(\kappa_1)} \tag{5.6}$$

$$\delta = \frac{\dfrac{1}{m+1} \sum_{\kappa_1=0}^{m} (\phi(\kappa_1)-\bar{\phi})^2}{\bar{\phi}^2} \tag{5.7}$$

One can draw also conclusions from the contours of $h_2(\kappa_1,\kappa_2) = (\epsilon) =$ constant.

(ϵ)

(a) In **Hammerstein** models, the points of the contours of a second- degree kernel can only lie on the main diagonal, because they are zero at any other place.

(b) In the **simple** and **generalized Wiener** models the contours are perpendicular to the main diagonal.

(c) In **Wiener** models, the contours do not close in if the linear terms before the multiplier are overdamped.

(d) In **Wiener** models, the contours are closed - eventually many times, and they may form 'islands' - if one of the linear terms before the multiplier is an underdamped one.

(e) In **Wiener-Hammerstein** cascade models, the contours close around a point of the main diagonal if all linear terms in the model are overdamped.

(f) In **Wiener-Hammerstein** cascade models, the contours may close around more points at the main diagonal - may build 'islands' - if the linear term after the multiplier is of the underdamped type.

(g) In **Wiener-Hammerstein** cascade models, the contours may close in around any points outside the main diagonal if at least one of the linear terms before the multiplier is of the underdamped type.

Next there is a criteria based upon the rank of a matrix $H_2(\kappa_1,\kappa_2)$ formed from quadratic kernels:

(ζ)

(i) The rank of the **simple** and **generalized Wiener** model is *1*.

(ii) The rank of the **extended Wiener** model is *2*.

(iii) The rank of the **Hammerstein** and **Wiener-Hammerstein** models is $m + 1$.

$$H_2(\kappa_1,\kappa_2) = [h_2(\kappa_1,\kappa_2)] \tag{5.8}$$

The Non-Linear Model Linear-in-the-Parameters

The non-linear model is assumed to be a differential equation linear in the parameters. The structure identification is performed by selecting all significant components amongst all possible ones. This can be done by using an orthogonal-forward-regression method, see Desrochers and Mohseni, 1984; Korenberg *et al.*, 1988; Chen *et al.*, 1989. The process itself is divided into a process part (no noise elements) and a noise part. Firstly, the structure of the process part is estimated and secondly, the structure of the noise part is determined. Next the actual estimation of the parameters is carried out over the combined process and noise parts of the model.

The system can be represented as an m outputs and r inputs non-linear model:

$$y(t)=f(y(t-1),...,y(t-n_y),u(t-1),...,u(t-n_u),e(t-1),...,e(t-n_e))+e(t) \tag{5.9}$$

where

$$u(t)=\begin{bmatrix}u_1(t)\\ \vdots\\ u_r(t)\end{bmatrix},\ y(t)=\begin{bmatrix}y_1(t)\\ \vdots\\ y_m(t)\end{bmatrix},\ e(t)=\begin{bmatrix}e_1(t)\\ \vdots\\ e_m(t)\end{bmatrix} \tag{5.10}$$

are the system input, output and noise respectively; n_y, n_u and n_e are the maximum lags in the output, input and noise; $\{e(t)\}$ is assumed to be a white sequence; and $f(\cdot)$ is some vector-valued non-linear function. The non-linear form of $f(\cdot)$ is generally unknown. In order to use the model for identification, a means of parameterization is required and a polynomial expansion of degree L of $f(\cdot)$ is formed:

$$y(t)=\sum_{j=1}^{n}\theta_j x_j(t)+e(t) \tag{5.11}$$

where $x_i(t)$ are the polynomials of degree 0 to L. One can write the model $f(\cdot)$ as a combination of a process part $f^p(\cdot)$ and a noise $f^n(\cdot)$ part:

$$\begin{aligned}y(t)=&f^p(y(t-1),...,y(t-n_y),u(t-1),...,u(t-n_u))\\&+f^n(y(t-1),...,y(t-n_y),u(t-1),...,u(t-n_u),e(t-1),...,e(t-n_e))+e(t)\end{aligned} \tag{5.12}$$

where $f^p(\cdot)$ includes all terms $\theta_j x_j(t)$ which do not contain noise elements. The remainder of the terms are included in $f^n(\cdot)$.

For the general model, delayed noise terms $e(t-j)$ are included in the model and these have to be estimated using the prediction errors or residuals $\varepsilon(t-j)$. The orthogonal-forward regression algorithm of the estimator ensures that the selection of the process and noise model parameters can be decoupled. Which terms will be included in the process model will not be affected by whatever noise model is produced later because of the orthogonal-forward regression algorithm that selects terms on basis of their contribution to the output signal. Initial residuals and a tolerance are then computed based on the process model and the model structure can then be re-selected with noise terms included. A revised residual sequence is calculated and so is a new tolerance, and an improved process and noise model is determined. A few iterations are often enough to determine the final model. Finally by calculating some simple correlations it is determined whether the model is adequate. Tests have shown that less than 10 iterations, typically 5-6, are sufficient for the algorithm to converge with a tolerance of 10^{-3}.

Radial Basis Functions (RBF)

The RBF approximation is a traditional technique for interpolating in multidimensional space (Powell 1985, Michelli 1986) and the theoretical properties of the method have been carefully investigated. Broomhead and Lowe (1988) introduced a generalized form of the RBF expansion with n inputs and a scalar output implementing a mapping $f_r:\mathbb{R}^n \rightarrow \mathbb{R}$ according to

$$f_r(x) = \lambda_0 + \sum_{i=1}^{n_r} \lambda_i \phi(\|x - c_i\|) \tag{5.13}$$

where $x \in \mathbb{R}^n$, $\phi(\cdot)$ is a function from $\mathbb{R}^+$ to $\mathbb{R}$, $\|\cdot\|$ denotes the euclidean norm, λ_i, $0 \leq i \leq n_r$, are the weights or parameters, $c_i \in \mathbb{R}^n$, $1 \leq i \leq n_r$, are the RBF centres and n_r is the number of centres, see Figure 5.1. The functional form $\phi(\cdot)$ is assumed to have been given and the centres c_i are some fixed points in n-dimensional space. The centres c_i must be appropriately distributed over the input domain, and they are usually chosen either to be subsets of the data or distributed uniformly in the input domain (see for another possibility the paper of Hofland *et al.*, 1992). Typical choices for $\phi(\cdot)$ are the thin-plate-spline function $\phi(v) = v^2\log(v)$, the gaussian function $\phi(v) = \exp(-v^2/\beta^2)$, the multiquadric function $\phi(v) = (v^2+\beta^2)^{1/2}$, and the inverse multiquadric function $\phi(v) = \dfrac{1}{(v^2+\beta^2)^{1/2}}$ where β is a real constant. Why such choices of non-linearity form good RBF approximations is shown, for example, by Powell (1987). It is important to emphasize that the performance of RBF networks depends critically upon the given centres.

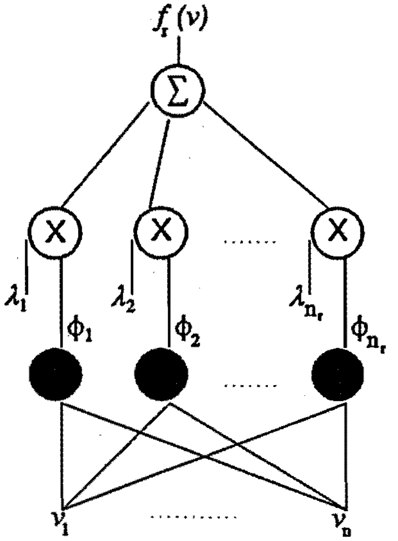


Figure 5.1 *The Radial Basis Functions network model*

A brief description will be given of how the RBF expansion, Eq. (5.13), can be used to realize or to approximate some unknown and complex nonlinear mapping $f_r:\mathbb{R}^n \rightarrow \mathbb{R}$. Assume for the time being that the centres are adjustable parameters and consider the two layer network in Figure 5.1. The surface generated by the gaussian function looks like a 'bump' in the two dimensional case. If the first layer of the network consists of enough gaussian functions, sufficient 'bumps' can be generated in appropriate positions by adjusting the parameters of the first layer. A variety of continuous surfaces from $f_r:\mathbb{R}^n \rightarrow \mathbb{R}$ can be well approximated by adding up a series of 'bumps'. In an RBF network the centres, that is the adjustable parameters, are fixed. If, however, a sufficient number of centres are used and they are suitably distributed over the input domain, it is possible that the RBF network will provide reasonably representations for various functions. The same orthogonal-forward-regression technique used with the non-linear model linear-in-the-parameters can be used to select the centres of the RBF-network.

5.2 Simulation Data

The structure identification algorithms are tested on data obtained with the fed-batch model described as in Section 3.4 with the initial settings as in Table 3.2. In this case the air flow and stirrer speed remain constant as we are only interested in the *OUR* and *CPR* models, or a combination of both measurements, the *RQ*. The input signal is the glucose flow, enriched by a *PRBS*-sequence, which can be arbitrarily fast, as we do not have any constraints on sensors or sampling devices. The exact excitation of the various data sets is found in Table 5.1.

Table 5.1 *Simulation settings for data generated for the non-linear identification algorithms using the fed-batch simulation model*

	oxidative growth experiment 1	fermentative growth experiment 2	combined growth experiment 3
μ [h^{-1}]	0.15	0.25	0.23
AF [l.h^{-1}]	8	8	8
SS [rpm]	1250	1250	1250
time [h]	14.4	7.2	7.5

Both the *GF* and *PRBS* signals will be used as input sequences for the structure identification algorithms described in Section 5.1.

Instead of using the fed-batch phase simulation model to generate data for testing the non-linear identification algorithms one could also use the batch phase simulation model excited with fed-batch input data. The justification for using a batch phase model to simulate fed-batch phase growth is the higher non-linearity of the batch phase model compared to the fed-batch phase simulation model as described in Chapter 3, plus the non-accounted for non-linear terms in the fed-batch simulation model due to the rapid changes of the glucose concentrations in the broth (Chapter 3 and 4). Here we present only 1 data set, comparable to the 'combined growth', experiment 3, data set of Table 5.1 with the same settings, only the used simulation model is the batch model from Section 3.3.

The third data set concerns the experimental data set of Figure 4.4, where a

PRBS sequence was applied to the *GF*. This data set is comparable to the 'oxidative growth', experiment 1 of Table 5.1, the growth rate is 0.15 [h^{-1}], the *AF* and *SS* are exponentially increasing to ensure high enough *DOT* levels.

A fourth data set is concerned with experimental batch data. This batch data set has no excitation, there is no input, all ingredients have been added before (see Chapter 2 and 3). This batch data set will only be used with the non-linear model linear-in-the-parameters and the Radial Basis Functions. The aim of these investigations is to see whether a suitable prediction for the specific growth rate and the biomass concentration can be obtained.

5.3 Results

The data described in the previous section will be used to test the non-linear structure identification algorithms on their applicability to identify the structure and parameters of the fed-batch phase of the laboratory process. The tests will be done in the same order as the algorithms are presented in Section 5.1. As already mentioned in the previous Section not all of the data is used with every algorithm; the Examples will denote what has been used.

Most of the results of the different non-linear structure identification algorithms are comparable. However, one has to remember that all algorithms have different backgrounds, i.e. ideas on which they are developed. One algorithm might work better for a particular set of signals than the other; our main task is to find out which of these algorithms are suitable for identification of the fed-batch process. At the end of this section we will compare all algorithms on a few key items such as: accuracy of results, complexity, convergence, etc.

Block-Oriented Models

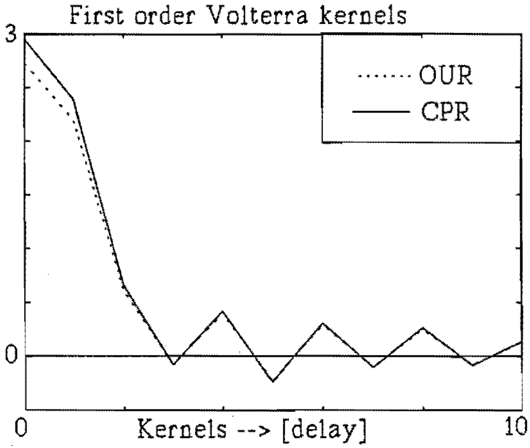
This section describes the use of the structure identification algorithm based on Volterra kernels described in Section 5.1. The algorithm is tested, not shown here, on known structures such as: simple and general Wiener and Hammerstein models and Wiener-Hammerstein models. In accordance with Haber (1989) the following items are estimated or shown: the DC-component, the analytical terms, the second order Volterra kernels, and a contour plot of the second order Volterra kernels. The tests indicate that the routine works and that the analytical and graphical indices do represent the system structure. Furthermore it is noted that the graphical indices give a better indication of the structure than the analytical ones. The analytical indices are derived from the graphical ones.

First the algorithm is tested on data generated by the fed-batch simulation model. We will try to find out what kind of structure is involved and compare that with the knowledge we have about the process (Chapter 2 and 3). Where possible we try to make a model of the process.

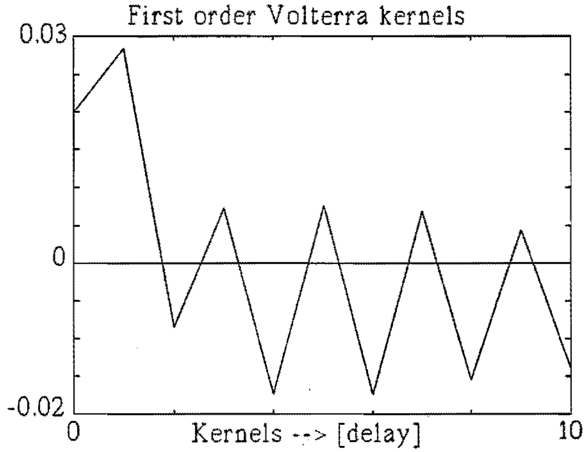
Example 5.1 *Structure identification using Volterra kernels; the data is from the fed-batch simulation model; the input is GF↔ OUR, an exponentially increasing signal with PRBS excitation*

Settings:
Up to and including second degree Volterra kernels
Estimation of weights (h_1 and h_2) by means of least squares
m = 10 (memory of the Volterra kernels)
Data:
U = GF
Y = OUR, CPR, RQ RQ is detrended
μ = 0.15 0.23 0.25 [h^{-1}]
Resulting analytical indices:
The DC-term Volterra kernel = 1.16664
The α-term = 0.90058
The β-term = 25.5852
The γ-term = 29189
The δ-term = 12.1034
The ζ-term = 11

The analytical indices indicate a Hammerstein model. The α-term is almost zero and the ζ-term is equal to 11 indicating a (Wiener-)Hammerstein model. The rest of the analytical indices do not contribute further information. This observation is made for all input (*GF*) - output (*OUR,CPR,RQ*) relations under all specific growth rates of the simulation data. The figures of the analytical indices differ somewhat along the different outputs and specific growth rates but are not very different from Example 5.1 (*GF* ↔ *OUR*).



(a)



(b)

Figure 5.2 *Representation of the first degree Volterra kernels for OUR and CPR outputs part (a) and for the RQ output part (b)*

The contour plot of the second degree Volterra kernels and its graphical 3-dimensional representation do not contain any valuable information. The plot of the first degree Volterra kernels shows that the underlying function, that is for *OUR* and *CPR*, is nothing more than a one (or two) delay linear kernel function. For the *RQ* it is not as clear. The picture has no distinct values for any of the delay kernels except for delay 0 and 1, but the values associated with these delays are not significantly different from the rest. ■

The models obtained, i.e. the *GF* ↔ *OUR* and *GF* ↔ CPR models, from the Volterra kernels based identification of Example 5.1 are the following:

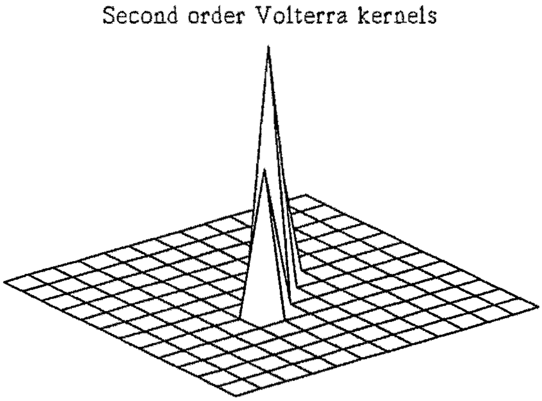
$$CPR(t),\ OUR(t) = h_0 + h_1(0)\cdot GF(t) + h_1(1)\cdot GF(t-1) \tag{5.14}$$

This model is linear and could be estimated by any linear estimation algorithm. There are some drawbacks using this model structure as from the discussion it follows that the model is valid for all specific growth rates (parameters may vary). Firstly, the *RQ* can not be modelled by this function. Secondly the *OUR* is modelled incorrectly as we know that in case of ethanol production the *OUR* will not be excited by the *GF* signal. Using the model (Eq. 5.13) we still find excitation due to $h_1(0)$ and $h_1(1)$.

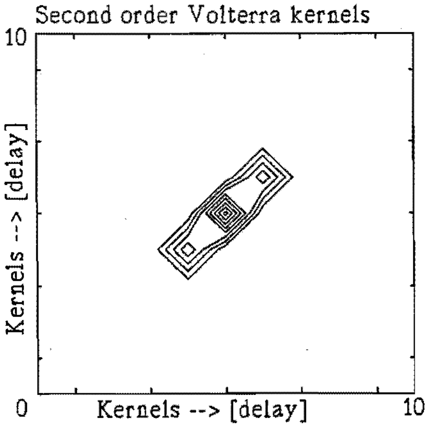
If we try to identify the process as in Chapter 4 using a non-linear compensation, but this time not compensating for the output signals, we could find a "real" non-linear function. In Example 5.2 we use the excitation signal, *PRBS*, as input instead of the *GF* and once again look at the outputs *OUR*, *CPR*, and *RQ*. The non-linear compensation, i.e. the experimental increasing function Eq. (4.11), is used in many identification and control schemes (Axellson, 1989 and O'Conner *et al.*, 1992).

Example 5.2 *Structure identification using Volterra kernels; the data is from the fed-batch simulation model; the input is PRBS only*

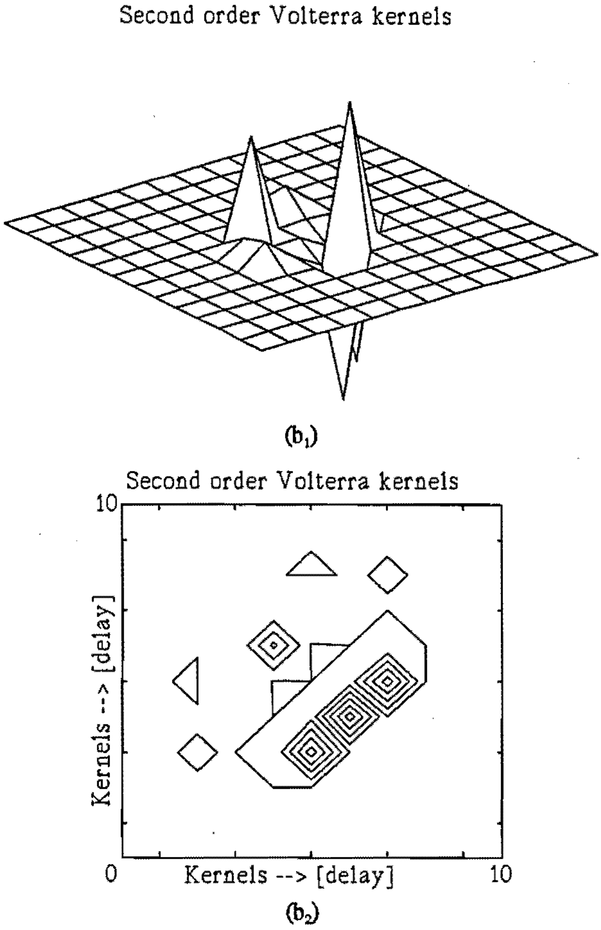
Settings:
See Example 5.1
Data:
See Example 5.1, only U = *PRBS*
Resulting analytical indices:
The DC-term Volterra kernel = 0.00328
The α-term = 0.06651
The β-term = 10
The γ-term = 161.462
The δ-term = 10
The ζ-term = 3



(a_1)



(a_2)



(b_1)

(b_2)

Figure 5.3 *3-dimensional and contour plots of the second degree Volterra kernels for CPR and RQ outputs, the specific growth rate is 0.15 [h^{-1}] or 0.23 [h^{-1}] part (a) and 0.25 [h^{-1}] part (b)*

The analytical indices indicate again a Hammerstein model (α and ζ term), as in Example 5.1. Now the graphical plots of the second degree Volterra kernels confirm these indications as can be seen in Figure 5.3. Specially for the μ = 0.15 [h^{-1}] and 0.23 [h^{-1}] structure (part a) this is noticed. For the μ = 0.15 [h^{-1}] and 0.23 [h^{-1}] the structure of the *PRBS* ↔ *OUR*, *PRBS* ↔ *CPR*, *PRBS* ↔ *RQ* relations remains the same. For the pure ethanol production growth rate (μ = 0.25 [h^{-1}]) there is no *PRBS* ↔ *OUR* relation any more and the *PRBS* ↔ *CPR*, *PRBS* ↔ *RQ* relations change, as can be seen in Figure 5.3, part (b). The Hammerstein model becomes a bit more complex, as Wiener elements come into the picture. ■

The difference with Example 5.1 is that in this case the *PRBS* ↔ *OUR* can be modelled. Note that the input data set has been compensated for with the same non-linear compensation as used in Section 4.2. The resulting model for the $GF \leftrightarrow RQ$ or $GF \leftrightarrow OUR$, $GF \leftrightarrow CPR$ relation for $\mu = 0.15$ [h^{-1}] looks like:

$$\begin{aligned} RQ(t) &= h_0 + h_1(0) \cdot PRBS(t) + \ldots + h_1(8) \cdot PRBS(t-8) \\ &+ h_2(4) \cdot PRBS^2(t-4) + h_2(5) \cdot PRBS^2(t-5) + h_2(6) \cdot PRBS^2(t-7) \end{aligned} \qquad (5.15)$$

This model is quite strange as it has a delay of 3 samples for the quadratic terms although we know from the simulation results that there is no delay incorporated in the model. These quadratic terms, however, make the model non-linear. The model structure, Hammerstein, i.e. a non-linear function followed by a linear one, is quite understandable if we relate the results to the yeast culture. The culture of *Saccharomyces cerevisiae* is fed with glucose in order to grow. This growth is achieved by budding of the yeast cells which can be seen as a multiplication; one parent cell 'generates' one daughter cell, this can be seen as the non-linear part of the model. During the growth the yeast cell excretes carbon-dioxide and takes up oxygen at the same time, which is measured using the effluent-gas of the fermenter and a mass-spectrometer (Chapter 2). The measurement can be seen as the linear part of the model, following the non-linear part.

Another possibility to test the algorithm is to use data generated from the batch phase simulation model. This model is more non-linear due to the influence of the enzymatic terms (see Section 3.3). Using the data as described in Section 5.2 we obtain the results as presented in Example 5.3.

Example 5.3 *Structure identification using Volterra kernels; data generated using the batch simulation model*

Settings:
See Example 5.1, only $m = 20$
Data:
$U = GF$
$Y = RQ$ RQ is detrended
$\mu = 0.23$ [h^{-1}]
Resulting analytical indices:
The DC-term Volterra kernel = $9.8 \cdot 10^{-2}$
The α-term = 0.30376
The β-term = $9.5 \cdot 10^{-2}$
The γ-term = 9573.77
The δ-term = 1.86435
The ζ-term = 21

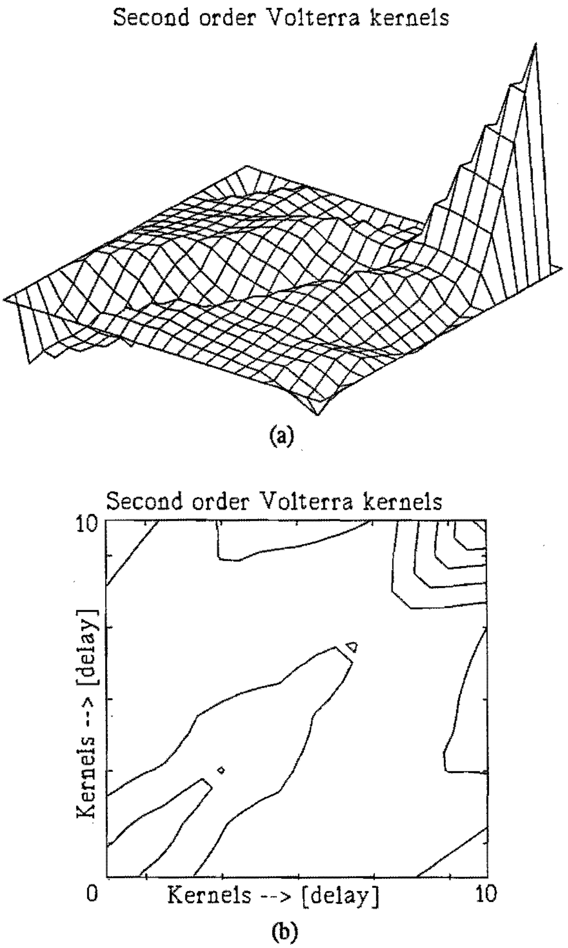


Figure 5.4 *3-dimensional (a) and contour plots (b) of the second degree Volterra kernels for the GF ↔ RQ relation*

This time a combination of Wiener and Hammerstein model is indicated: both α and β terms are low; the ζ-term indicates this as well. The graphical plots give the indication of a Wiener-Hammerstein model. Note that this data set only contained data for $\mu = 0.23$ [h^{-1}], also, no model has been made for *GF ↔ OUR, GF ↔ CPR.* ■

The enzymatic non-linearities that are involved give a slightly different presentation, but still a Hammerstein model. This time a complete model structure, a Wiener Hammerstein model, together with parameters is constructed. The linear dynamic terms of the model are estimated using a recursive least squares algorithm using UDU factorization in order to avoid numerical instability.

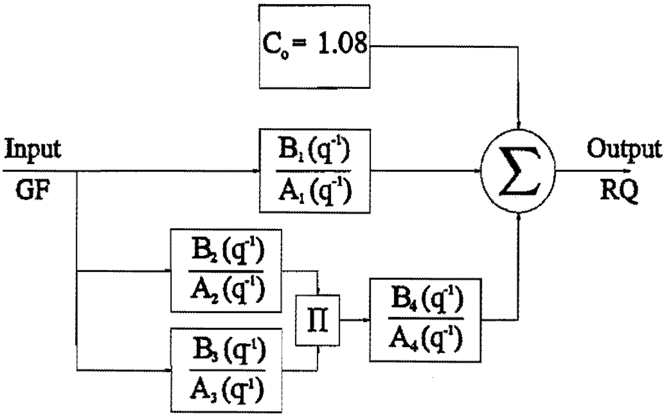

$C_0 = 1.08$
Input
GF
$\frac{B_1(q^{-1})}{A_1(q^{-1})}$
Σ
Output
RQ
$\frac{B_2(q^{-1})}{A_2(q^{-1})}$
Π
$\frac{B_4(q^{-1})}{A_4(q^{-1})}$
$\frac{B_3(q^{-1})}{A_3(q^{-1})}$

Figure 5.5 *Block structure of the Wiener-Hammerstein model*

Table 5.2 *Parameters associated with Wiener-Hammer-stein model*

Block #	$A(z^{-1})$	$B(z^{-1})$
1	$1 - 0.95 \cdot z^{-1}$	$0.08 \cdot z^{-1}$
2	$1 - 0.002 \cdot z^{-1}$	$1.30 \cdot z^{-1}$
3	$1 - 0.007 \cdot z^{-1}$	$1.60 \cdot z^{-1}$
4	$1 - 0.03 \cdot z^{-1}$	$300 \cdot z^{-1}$

The actual model is an extended Wiener-Hammerstein model according to Figure 5.5 and Table 5.2. The choice of the extended Wiener-Hammerstein model was dictated by the presence of 'fast' and 'slow' dynamics in the system (see Section 3.3). The model consists of 4 different dynamic blocks, a summer, a multiplier and a constant. The linear dynamic block (B_1/A_1) represents the 'slow' dynamics of the system, associated with the enzyme dynamics. The other three blocks are related to the 'fast' dynamics, the yeast growth.

Finally we use the experimental data to identify a non-linear model. Example 5.4 will illustrate this.

Example 5.4 *Structure identification using Volterra kernels, experimental fed-batch data*

Settings:
See Example 5.1
Data:
U = *PRBS, GF*
Y = *OUR, CPR, RQ* — *RQ* is detrended
μ = 0.15 [h^{-1}]
Resulting analytical indices:
The DC-term Volterra kernel = 2.587
The α-term = 1.208
The β-term = 121.9
The γ-term = $1.13 \cdot 10^9$
The δ-term = 81537
The ζ-term = 21

The results of the structure identification indicate (ζ-term) a Hammerstein model. The graphical presentations of any of the relations ($GF \leftrightarrow OUR$, $GF \leftrightarrow CPR$, and $GF \leftrightarrow RQ$) are all not informative at all. This indicates that the relation might be very simple, first order model, or too difficult to identify using the Volterra kernels. The presence of the noise, both at input and output signals deteriorates the performance of the algorithm as well. ■

The data generated with the fed-batch simulation model from Section 3.4 was quite helpful in hinting at the underlying structure. This knowledge could then be exploited with the batch simulation model data (Chapter 3.3.). Using experimental data for identification did not result in any model. The method as presented here is thus of limited practical use for identification of the laboratory process.

Non-Linear Model Linear-in-the-Parameters

This Section describes the use of the structure identification algorithm based on the orthogonal forward-regression algorithm of Chen *et al.* (1989). The algorithm is first checked on a known process described by Billings (1989). In accordance with Billings (1989) the following items are tested; the sum of squared output-error, the input-output correlation, and the parameter estimates. These items are tested with various settings of the stop criterium for the process model, and for the convergence, as these are the only selection criteria on the algorithm. As the algorithm is completely automatic, it doesn't need any interactive handling. The setting of the stop criterium should be chosen carefully. If one tries to obtain a process model that describes the input-output data set very accurately, the model could easily be overparameterized. An adequate stop criterium for the process model would be little higher than the standard deviation of the noise.

If, however, both process and noise model have to be estimated a very accurate process model is the right choice. First the algorithm will find a process model satisfying the given accuracy. Next the resulting error will be treated as noise and added to the input domain for estimation of the noise model. A new stop criterium is calculated. The noise model is estimated together with the process model yielding a new equation error and stop criterium. The algorithm repeats these steps and finds an optimal value after iteration. If no noise parameters are involved, the best way to use the algorithm is by an interactive iterative procedure. The outcome of the algorithm can then be checked against the correlation functions. If the functions fall within the 95% confidence interval the model is regarded as adequate. The output generated by the algorithm indicates which input/output data have been selected and how big the contribution of this data is to the process output. Then the user can decide whether to incorporate all selected terms or skip less relevant terms that contribute to the process output but not substantially. This choice should be made afterwards as the parameter values are estimated during the structure identification. Note that the input and output signals should be detrended, if possible, and not be oversampled.

The algorithm is implemented in the MatLab environment and has a menu-driven structure. Part of the menu will be shown in the "Example window" under the name Settings. This menu, see Example 5.5, concerns the algorithm settings for the identification of the process part (1-7) and the settings for the noise part (9-11). With switch 8 one can chose to have the noise part identified or not. Switch 12 is the display parameter to show the obtained results on screen or not.

The algorithm is again first tested on the data generated by the fed-batch simulation model presented in Section 5.2. Due to the fact that the algorithm will automatically generate a model these models will be presented as results. We compare the results with the knowledge we have about the process.

Example 5.5 *Structure identification using the non-linear model linear-in-the-parameters; data generated by the fed-batch simulation model, delayed output terms included*

```
Settings:
     1)   Order of non-linearity (1=linear, 2=square)     : 2
     2)   Number of process inputs                        : 1
     3)   Number of process outputs                       : 3
     4)   Selected process output for identification      : 1-3
     5)   Minimal delay of inputs                         : 0
     6)   Maximum degree of process model                 : 3
     7)   Termination tolerance for process model         : 0.01
     8)   Calculate noise model ( 0=No, 1=Yes )           : 0
     9)   Maximum number of iterations for noise model    : 10
     10)  Maximum degree of noise model                   : 1
     11)  Stop criterium for noise model                  : 0.01
     12)  Display parameter ( 0=Off, 1=On )               : 1
```

Data:

U = GF + delayed terms of OUR, CPR, and RQ

Y = OUR, CPR, RQ RQ is detrended

μ = 0.15 [h^{-1}]

Estimation data = 0 - 6.5 [h],

Validation data = 6.5 - 14.4 [h].

Results:

$OUR(t) = 0.926 \cdot CPR(t)$

$CPR(t) = 1.08 \cdot OUR(t)$

$RQ(t) = 1.08 + 1.3 \cdot RQ(t-1) - 390 \cdot RQ(t-2)^2$

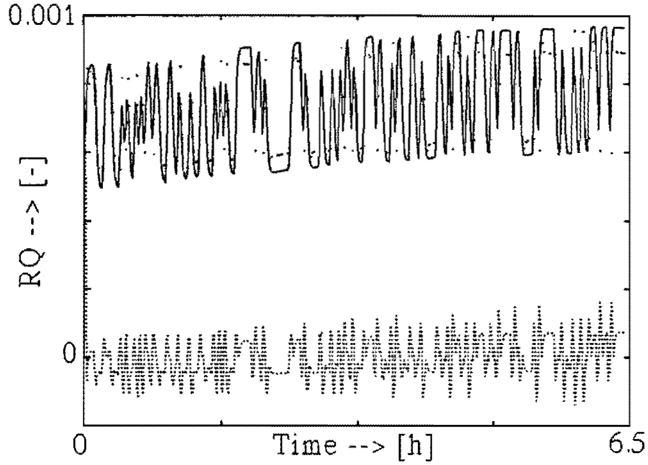


Figure 5.6 *Plot of the calculated RQ and its model response; the error is denoted as a dotted line*

A linear function is found for both *OUR* and *CPR*, independent of the input, depending only on each other. For the *RQ* also an auto-regressive model is found, although non-linear in its second lag term. Figure 5.6 shows the estimation data for the *RQ* model, the validation is not shown because of the auto-regressive model, it is self predictable. ■

The results describe the models found with the given settings. As all data was available the structures chosen are not very strange. From Chapter 3 we know that with pure oxidative growth, as in Example 5.5, there is a static relation between the *OUR* and the *CPR* expressed by the *RQ*. From Section 3.2 we know that the *RQ* should be constant, around 1.08 for oxidative growth on glucose only and for this yeast. The relation for the *RQ* is like most predictive error models, heavily dependent on the latest information available, i.e. *RQ(t-1)*.

In order to eliminate the above mentioned influence of auto-regressive data, a FIR model will be identified, using only the delayed input combinations. The results are as denoted in Example 5.6

Example 5.6 *Structure identification using the non-linear model linear-in-the-parameters; data generated with the fed-batch simulation model, only delayed input terms included*

Settings:
Same as in Example 5.5
Data:
Same as in Example 5.5 only $U = GF$
Results:
$\text{OUR} = 6.143 \cdot \text{GF}(t)$
$\text{CPR} = 6.633 \cdot \text{GF}(t)$

Two linear models for both the *GF* ↔ *OUR* and *GF* ↔ *CPR* relation are found. The *GF* ↔ *RQ* relation could not be identified up to the requested accuracy and the model obtained did not give a reasonable prediction of the *RQ* response to the *GF* data. The identification could have been done with a normal linear algorithm. ■

Now the results are directly linked with the input *GF*, multiplied with a constant for *OUR* and *CPR*, which can be concluded from the equations of the fed-batch simulation model (Eq. 3.16). The *RQ* can not be reconstructed in this way, it needs lagged outputs as in Example 5.5.

Next we look at the combined growth data, $\mu = 0.23$ [h^{-1}], of the fed-batch simulation model. As the process will be more non-linear than the previous one, we expect a more complex model to be identified. We know that taking into account only lagged inputs is not enough for the *RQ* model. So, we take a combination of lagged input and output terms. The results are denoted in Example 5.7.

Example 5.7 *Structure identification using the non-linear model linear-in-the-parameters; data generated with the fed-batch simulation model, combined growth*

Settings:
Same as in Example 5.5 only
7) Termination tolerance for process model : $2 \cdot 10^{-4}$
Data:
Same as in Example 5.6 only $\mu = 0.23$ [h^{-1}]
Estimation data 0 - 5 [h], validation data 5 - 7.5 [h]

Results:

$OUR(t) = 0.0093 + 0.818\cdot OUR(t-1) +$
$0.0073\cdot OUR(t-1)\cdot GF(t-2) + 4.364\cdot GF(t) - 4.366\cdot GF(t-1) -$
$0.272\cdot OUR(t-2)^2$
$CPR(t) = -0.0281 + 0.952\cdot CPR(t-1) +$
$0.128\cdot CPR(t-1)\cdot GF(t-2) - 0.614\cdot CPR(t-2)^2 - 0.006\cdot GF(t-2)^2 +$
$5.291\cdot GF(t) - 5.308\cdot GF(t-1) + 0.204\cdot CPR(t-3)^2$
$RQ(t) = -0.516\cdot RQ(t-1) - 17.32\cdot RQ(t-2) + 6.06\cdot RQ(t-3) + ..$

The *GF* ↔ *OUR* and *GF* ↔ *CPR* relations are modelled by non-linear functions. Some of the model parameters are almost identical but have a different sign, like the terms $4.364\cdot GF(t)$ and $4.366\cdot GF(t-1)$ in the *GF* ↔ *OUR* model. This has the effect that a change in the input is only noticeable at one time instant, which is the case with the steep changes in the *PRBS* signal. A *GF* ↔ *RQ* model is also found, but only the first three terms are given, there are quite a lot more, but the minimum accuracy could not be attained. The three terms represent the process for about 70 %; whereas the others did not contribute more than 2 - 3% to the error reduction. ■

For both the *GF* ↔ *OUR* and *GF* ↔ *CPR* model we find models that are highly non-linear but still are very much dependent upon the auto-regressive terms. This is even more true for the *GF* ↔ *RQ* relation, that is entirely based on the auto-regressive parts. The non-linearities of switching between ethanol production and consumption can not be modelled adequately by the *GF* ↔ *RQ* using the non-linear model linear-in-the-parameters.

If we look at the results for the fermentative growth ($\mu = 0.25$ [h^{-1}]), see Example 5.8, the models are less complex.

Example 5.8 *Structure identification using the non-linear model linear-in-the-parameter; data generated with the fed-batch simulation model, fermentative growth*

Settings:
Same as in Example 5.5
Data:
Same as in Example 5.5 only $\mu = 0.25$ [h^{-1}]
Estimation data 0 - 5 [h], validation data 5 - 7.2 [h]
Results:

$OUR(t) = 0.473\cdot OUR(t) + 5.73\cdot GF(t) - 2.87\cdot GF(t-1) -$
$18.94\cdot GF(t)^2 + 3.544\cdot GF(t-1)\cdot OUR(t-1)$
$CPR(t) = 3.944\cdot GF(t) + 0.364\cdot CPR(t-1)$
$RQ(t) = 1.23\cdot RQ(t-1) - 5.93\cdot RQ(t-2)^2 + 1.97\cdot RQ(t-1)\cdot RQ(t-3)$
$+ 10.72\cdot GF(t)^2 - 12.28\cdot GF(t-1)\cdot GF(t-2) +$
$5.04\cdot GF(t-2)\cdot RQ(t-2) - 0.212\cdot RQ(t-2)$

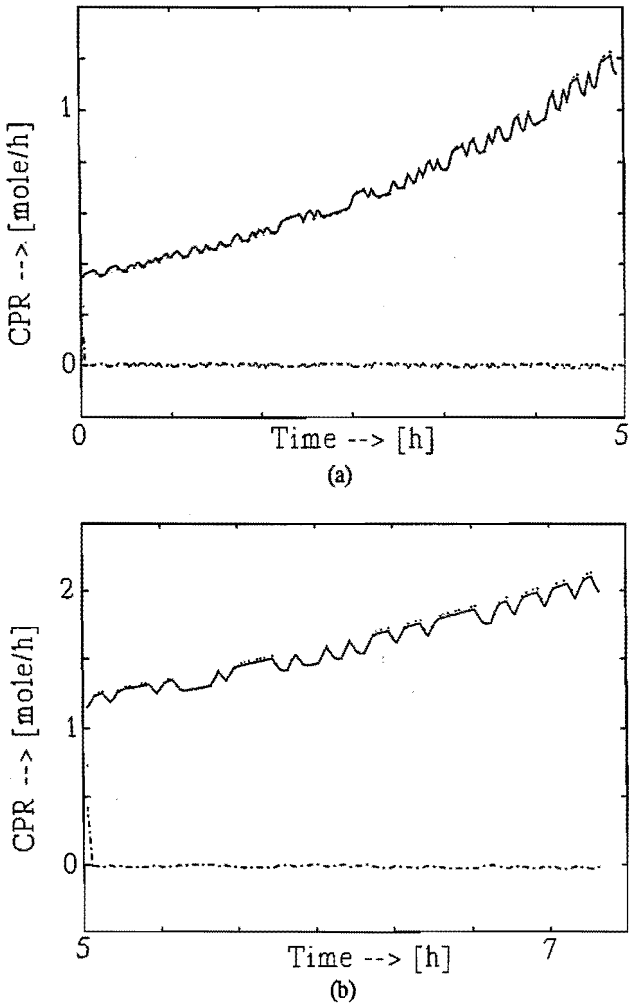


Figure 5.7 *Estimation (a) and validation (b) data for the estimated GF ↔ CPR model for fermentative growth*

The *GF* ↔ *OUR* relation is a complex model of 5 terms, and is modelled by a non-linear function. The *GF* ↔ *CPR* relation is just an ARX model which could again be identified by a linear identification algorithm. The prediction of this model is quite good as can be seen from Figure 5.7 part (a), estimation data, and part (b), validation data. Contrary to the previous models the *GF* ↔ *RQ* relation has not only auto-regressive terms but also lagged input terms. ■

The *GF* ↔ *OUR* relation is simply nothing more than an exponentially increasing function as long as the oxygen supply is high enough. The model relies thus heavily on the auto-regressive parts (50 %) and the input has to be smoothened such that no excitation is seen. The *GF* ↔ *CPR* relation is quite simple, as it is built from an oxidative part, the *CPR(t-1)*, and a fermentative part, the *GF(t)*. The *GF* ↔ *RQ* relation is a more complex one, and non-linear but, contrary to the other two Examples (5.6 and 5.7), this time it is modelled.

Whether the method will still work when more non-linearities are involved can be checked when using the batch simulation data set as presented in Section 5.2 and used with Example 5.3. The results of our findings are presented in Example 5.9

Example 5.9 *Structure identification using the non-linear model linear-in-the-parameters; data generated with the batch simulation model*

Settings:
Same as in Example 5.5
7) Termination tolerance for process model : 0.1
Data:
Same as in Example 5.3
Results:
$RQ(t) = 1.08 + 0.8545 \cdot RQ(t-1) + 3.31 \cdot GF(t) -$
$2.941 \cdot GF(t-1) + 19.21 \cdot RQ(t-1) \cdot GF(t) + 422.4 \cdot GF(t)^2 -$
$309.9 \cdot GF(t-1)^2$

Here we are only interested in the *GF* ↔ *RQ* relation that can reconstruct the output signal for about 90 % (termination tolerance for process model : 0.1). Once again, like in Example 5.7, some of the terms seem to compensate each other. Due to the low termination criterium a model could be made, contrary to Example 5.7. ■

The *GF* ↔ *RQ* model is a non-linear one where it is seen that the contribution of the non-linear terms to the signal is small in comparison with the linear terms. Furthermore, the parameter values of the non-linear terms are big in comparison with the linear terms. The original signal is reconstructed for about 90 %. Note that the system has no noise incorporated such that only the process terms are selected and the user has to define the stop criterium for the process model such that this matches his needs.

Next we look at the experimental data. First the fed-batch experimental data is presented in Example 5.10.

Example 5.10 *Structure identification using the non-linear model linear-in-the-parameters; experimental fed-batch data*

Settings:
Same as in Example 5.5

Data:
Same as in Example 5.5 only $\mu = 0.15$ [h^{-1}]
Results:
OUR $= 0.364 \cdot GF(t)$
CPR $= 0.375 \cdot GF(t)$

As with the simulation model data, only the *GF* ↔ *OUR* and *CPR* can be identified, the *GF* ↔ *RQ* is not identifiable. The models found are, again, just linear ones, thus a linear identification algorithm could do the job as well. If we try to incorporate a noise model we find that the complete model, process and noise model, converges to a simple linear process model. ■

The results are comparable to that found in Example 5.6 with the oxidative growth on glucose. The only difference is that some parameters are different. Once again, for this data set, oxidative growth, a linear estimation technique, together with non-linear compensation (Chapter 4) appears to be the best solution.

In conjunction with the estimator to be presented in the next Chapter, we tried to predict, using non-linear identification algorithms, the biomass and specific growth rate. This was done using the *OUR, CPR*, lagged measurements, and estimates of both the biomass and specific growth rate.

Two batch experiments have been selected for the estimation and validation of the method as described in Section 5.2. The following measurements and estimations are available, biomass concentration (estimated and measured, i.e. optical density and dry weight), the *CPR*, the *OUR* and a prediction of the specific growth rate of the biomass. The data has been pre-processed: first peaks (outliers) have been removed, next filtering has been carried out because of the noisy character of the data and finally normalization has been performed to enhance estimation and validation. The noisy character of the output would normally be of no concern if it was consistent during the total run. However, the normal input and output signals of a fed-batch phase show an exponentially increasing trend. The noise on the output is additive and does not increase during an experiment. Consequently the signal to noise ratio is lower at the beginning and increases during the experiment. This is reflected at the specific growth rate that is not estimated well enough in the beginning, showing a very noisy pattern, thus filtering is necessary. The results of the structure determination is noted in Example 5.11.

Example 5.11 *Structure identification using the non-linear model linear-in-the-parameters; experimental batch data*

Settings:
Same as in Example 5.5 only
7) Termination tolerance for process model : $7 \cdot 10^{-3}$
Data:
Same as in Example 5.9

Results:

$$\mu(t) = 0.3721\cdot\mu(t-1) - 4.5993\cdot10^{-02}\cdot\mu(t-3) - 2.095\cdot OUR(t-1) - 1.8427\cdot10^{-03}\cdot CPR(t)^2 - 3.0365\cdot10^{-02}\cdot CPR(t-2)^2 + 0.3929\cdot\mu(t-1)^2 - 0.4554\cdot\mu(t-3)\cdot CPR(t-2) + 7.108\cdot\mu(t-1)\cdot OUR(t-1) + 0.2993$$

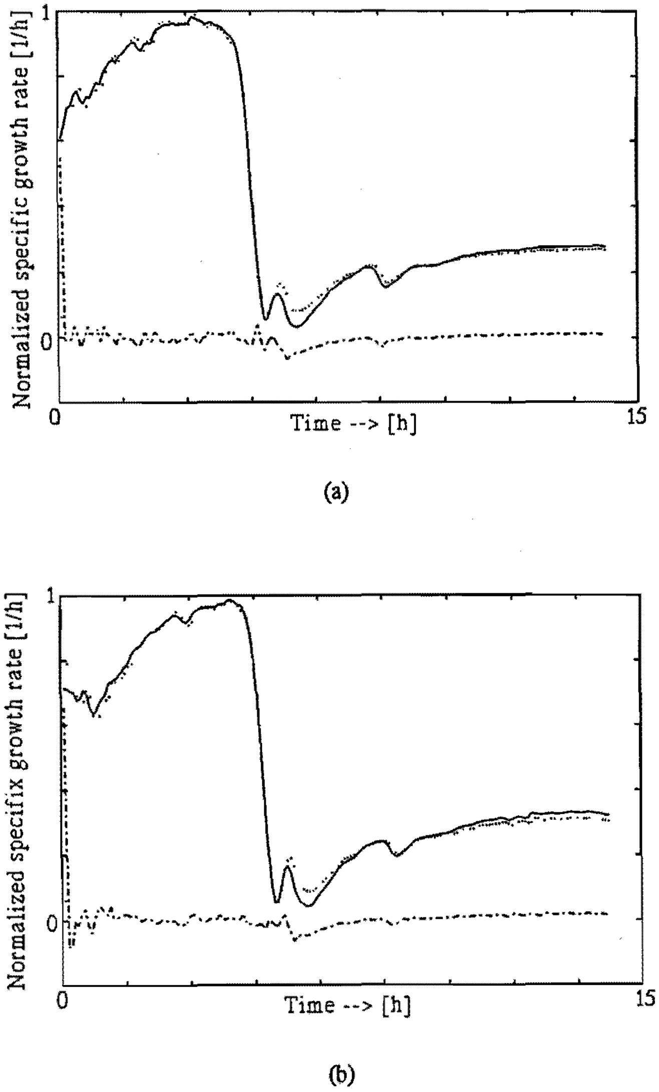


Figure 5.8 *Estimation (a) and validation (b) plots for the modelled specific growth rate based on OUR and CPR data*

Figure 5.8 shows the results of the non-linear model linear in the parameters for the specific growth rate. The model gives a good indication what is the most important variable, i.e. $\mu(t-1)$, as was found in other Examples. The variable $\mu(t-1)$ contributes for more than 95% to the output signal, which is of course evident as the two outputs do not change dramatically during the fermentation. If one tries to estimate a model for the biomass concentration this causes such a low sum of squared errors that one has to discard this model. It represents an integrator ($biomass(t) \approx biomass(t-1)$) and none of the actual input signals is taken into account, see Figure 5.14. The results of the biomass identification are omitted. ■

Once again the models tend to be auto-regressive as can be seen from the importance of the term $\mu(t-1)$ and *biomass(t-1)* for both models. Increasing the accuracy from $7 \cdot 10^{-3}$ to $5 \cdot 10^{-3}$ does not yield a better model, the model seems to be overparameterized. Two more terms are selected but the validation set shows a severe degradation at the end and the more accurate model can thus be regarded as useless

The non-linear model linear-in-the-parameters provides a low complexity model that is quite easy to obtain. It is noticed during all the Examples that the auto-regressive terms play an important role, and it also points out the biggest drawback of these models namely one step ahead prediction. The method however is very flexible to the choice of model structure and estimation of the parameters, whether the process is non-linear or not. Not all non-linear processes can be modelled, which again limits the application of this method.

Radial Basis Functions

This section describes the use of the Radial Basis Functions identification algorithm based on orthogonal basis. The algorithm is first checked on a known process described by Chen *et al.* (1990). In accordance with Chen *et al.* (1990) the following items are tested: the sum of squared output-error, the input-equation error correlation, and the validation data. These items are tested with various settings of the stop criterium for the process model and for the convergence. Beside these items, the algorithm can be tuned with a maximum of selected centres, both for the process and noise model. The same applies for the non-linear model linear-in-the-parameters algorithm as for this algorithm: the stop criterium should be chosen carefully. The algorithm automatically converges to a model if noise elements are involved, like the previous method. The outcome of the algorithm can be checked against the correlation functions or against validation data.

This method will follow the same line as the previous methods: first fed-batch simulation data, next batch simulation data, and finally experimental data. Note that the identification results are given in Tables as they contain too much data to be captured in the Example window. The algorithm is also implemented in the MatLab environment and has a menu driven structure. Part of the menu will be shown in the Example window under the name Setting. Particularly Example 5.12 will contain a lot of menu information that will be omitted in the next Examples if the information does not change.

Example 5.12 *Structure identification using RBF functions estimation and validation of GF ↔ OUR using the RBF-network*

Setting:

1)	Number of process inputs	1
2)	Number of process outputs	3
3)	Selected process output for identification	1,2
4)	Minimal delay of inputs	0
5)	Maximum lags of process inputs	1
6)	Maximum lags of process outputs	1
7)	Terminating tolerance for process model	10^{-4}
8)	Maximum number of centres for process model	10
9)	Calculate noise model (0=No, 1=Yes)	0
10)	Maximum number of iterations for noise model	10
11)	Maximum lags of noise model	1
12)	Stop criterium (loss-function)	0.1
13)	Maximum number of centres for noise model	10
14)	Display parameter (0=Off, 1=On)	1

Used non-linear function:

thin-plate-spline function $p(v) = v^2 \log10(v)$

Data:

μ = 0.15

U = *GF*

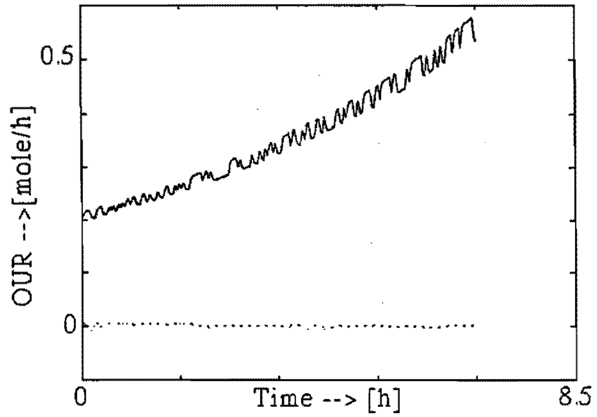
Y = *OUR, CPR, RQ* *RQ* detrended

Estimation set: 1:300

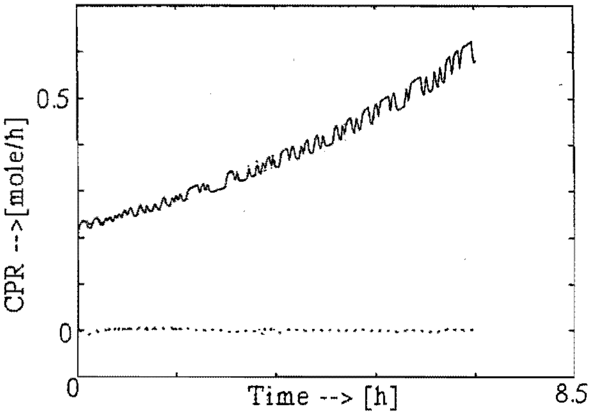
Validation set: 301:865

Results:

See discussion



(a)



(b)

Figure 5.9 *Estimation results for OUR (a) and CPR (b) of the RBF-network using the simulation data of Table 5.1 and* $\mu = 0.15$

Table 5.3 *Results of the RBF function identification. GF ↔ OUR model for oxidative growth*

Selected term	Weight estimate	Error reduction ratio
1	-3.88	0.99
219	4.48	0.0059
206	-102	0.000059
215	99	0.00021
12	-249	0.00059
283	-5.43	0.00023
11	247	0.000028
261	2.7	0.000019
397	3.93	0.000013
376	-4.37	0.000068

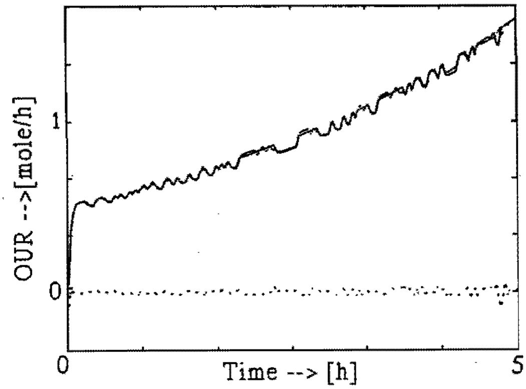
The settings of the number of centres with the estimation of the *GF* ↔ *OUR*, *GF* ↔ *CPR* functions is 10. An example of such a model is given in Table 5.3. The centres consist of 4 points of the input space (*GF, OUR, CPR, RQ*). The term that contributes most to the function is the first data point, which already generates 99% of the function output. Although the accuracy is high, near 10^{-4}, the validation data shows a completely erratic behaviour of the model (not shown here). Nevertheless, this minimum number of centre points gives the best solution; increasing the number of centre points deteriorates the model. It should be pointed out that no model of the *GF* ↔ *RQ* relation could be made within a reasonable number of centres (less than 25) and a normal accuracy (10^{-2}). ■

The outcome of the identification is not good. Use of another non-linear function instead of the thin-plate function may improve the results, but with the other three functions additional calculations for every centre have to be done, causing a vast amount of calculations. The outcome of this exercise is not predictable. The method is found to be very sensitive to the number of centre points (the order of the input space). No rule (of thumb) can be given regarding this number, only minimizing it can keep computations under control.

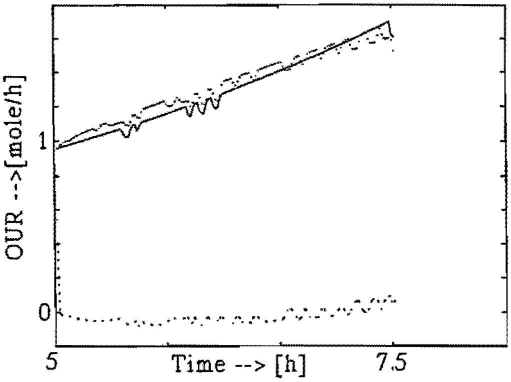
For the rest of the fed-batch simulation identification we keep the number of centres points low. Example 5.13 will show the models found with combined growth data.

Example 5.13 *Structure identification using RBF functions; combined growth data*

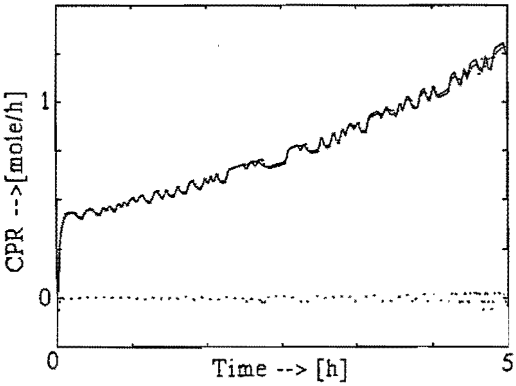
Setting:
As Example 5.12
Data:
As Example 5.12 only:
μ = 0.23
Validation set: 301:451



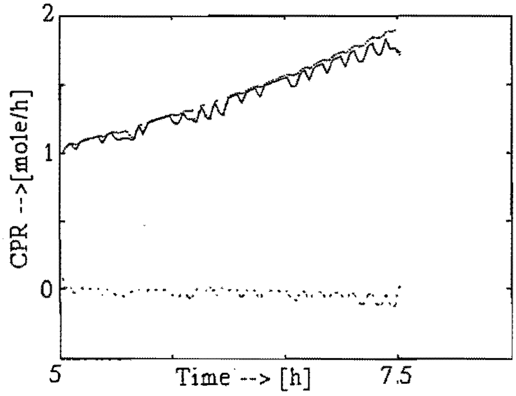
(a_1)


OUR -->[mole/h]
1
0
5
Time --> [h]
7.5

(a₂)


CPR -->[mole/h]
1
0
0
Time --> [h]
5

(b₁)


CPR -->[mole/h]
2
1
0
5
Time --> [h]
7.5

(b₂)

Figure 5.10 *Estimation, part (a_1) and (b_1), and validation, part (a_2) and part (b_2) of the GF ↔ OUR, part (a) CPR Part (b) models*

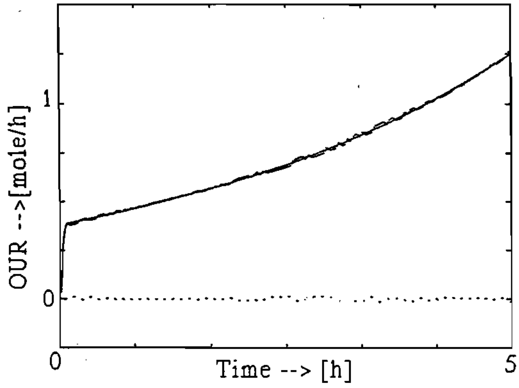
Here we see from Figure 5.10 that both an estimation and a validation set have been used for the *GF* ↔ *OUR,CPR* models. It is seen that the *GF* ↔ *OUR* model is not as good as the *GF* ↔ *CPR* one, due to the non-linearities. The *GF* ↔ *OUR* model is represented by 7 terms with the first one (again) as the most important one (99%), the *GF* ↔ *CPR* model has 10 centres and has as most important factor a centre (#90) that contributes for only 24%. Once again no model could be made of the *GF* ↔ *RQ* relations. ■

The non-linear switching between ethanol production and consumption is noticeable in all outputs. The *OUR* is the most non-linear one as it saturates to a maximum value causing some difficulties for the identification programmes so far. This switching is not so profound in the *CPR*. The models respond rather well to the validation data,but are of no use for other growth rates other than $\mu = 0.23$ [h^{-1}].

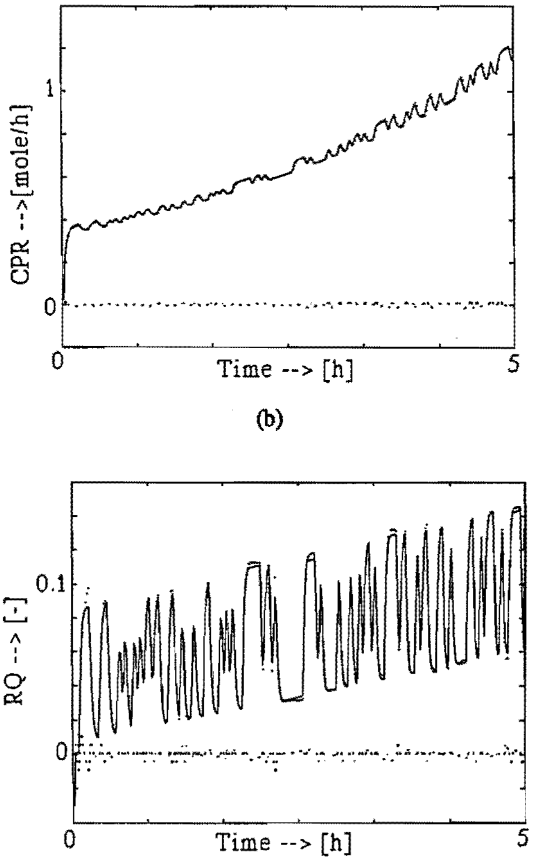
Finally we will consider the fermentative growth rate, $\mu = 0.25$ [h^{-1}]. Here no ethanol is consumed, only produced. The results are denoted in Example 5.14

Example 5.14 *Structure identification using RBF functions; fermentative growth*

Setting:
As Example 5.12
Data:
As Example 5.12 only:
$\mu = 0.25$
Validation set: 301:433



(a)



(c)

Figure 5.11 *Estimation data sets for fermentative growth of OUR (a), CPR (b) and RQ (c) models*

As one can see the *GF* $\leftrightarrow$ *RQ* relation is modelled here; note however that no models could be made using the validation data. For both the *GF* $\leftrightarrow$ *OUR and CPR* models 11 centres are found including the DC-term. Both the first data set and the DC-term are the most important ones (9.5%). For the *GF* $\leftrightarrow$ *RQ* only 9 terms are selected including a DC-term which is again with the first term the most important one (9.3%). ■

The question why the *GF* $\leftrightarrow$ *RQ* relation can be modelled with a $\mu = 0.25$ [h^{-1}] is probably answered by the fact that the system becomes almost linear and thus can be easily modelled except for the exponentially increasing terms. If we compensate for this term a pure linear model will follow, as can be read in Axellson (1989).

So far the overall performance of this identification technique is rather low, but another field where this technique could be used advantageously is in modelling batch fermentations. But before doing so we look at the experimental fed-batch data used with the two other methods as well. The results are presented in Example 5.15.

Example 5.15 *Structure identification using Radial Basis Functions, experimental fed-batch data*

Setting:

7) Terminating tolerance for process model $1 \cdot 10^{-2}$

Data:

Estimation set: 1:151
Validation set: 151:300

Table 5.4 *Results of the experimental fed-batch data*

GF ↔ OUR model		
Selected term	Weight estimate	Error reduction ratio
1	-6.32	0.97
34	-6.28	0.03
GF ↔ CPR model		
Selected term	Weight estimate	Error reduction ratio
1	-6.42	0.96
35	-6.74	0.04
GF ↔ RQ model		
Selected term	Weight estimate	Error reduction ratio
150	7.74	0.047
73	-12.5	0.12
74	49.1	0.021
145	-0.031	0.13
151	-0.22	0.086
130	-10.5	0.097
81	-30.9	0.099
105	3.78	0.078
87	-118	0.02
86	111	0.096

Both the *GF* ↔ *OUR* and *CPR* relations yield a 2-term RBF network. The *GF* ↔ *RQ* relation was limited to only 10-terms in order to keep the complexity of the model low. The results are given in Table 5.4.

Here we see from Figure 5.12 that the estimation of both CPR and RQ is reasonable, but the validation set of CPR is extremely bad. The *GF* ↔ *OUR* model is represented by 2 terms with the first one (again) as the most important one (97%), the *GF* ↔ *CPR* model has also 2 centres. The *GF* ↔ *RQ* model has 10 centres and has an total error reduction factor of only 79.4 %. ■

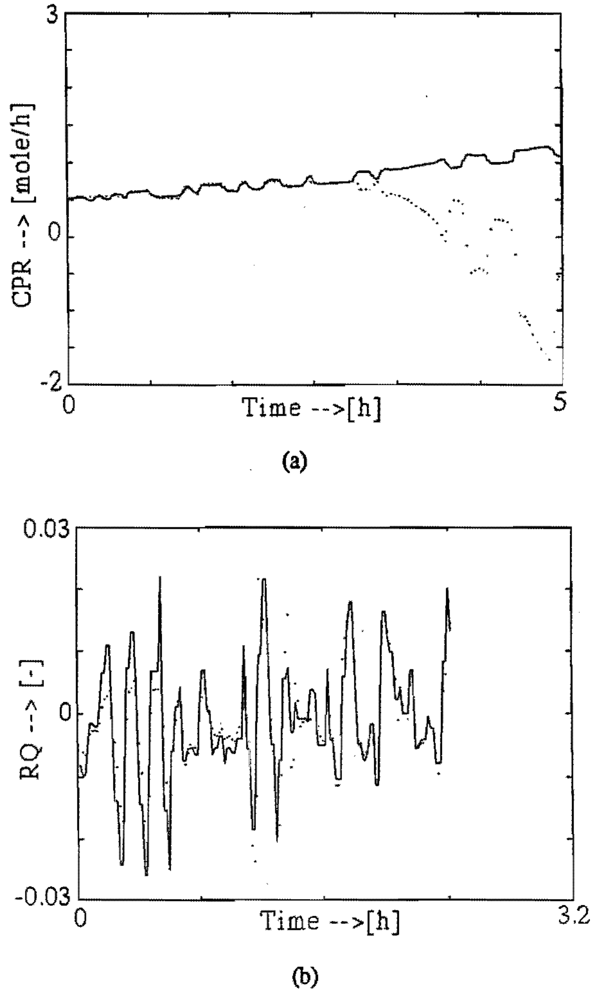


Figure 5.12 *Estimation (0 - 2.5 [h]) and validation (2.5 - 5 [h]) results for CPR (a) and RQ (b) of the RBF-network using experimental fed-batch data*

The RBF-network is capable of modelling all the relations, but the models we obtain are not very suitable. During the validation the models loose track and become very poor. This seems to be caused by the fact that the chosen centres are no longer representing the complete input domain. A possible solution for this problem would be the use of an adaptive strategy. As new data is gathered, it can be compared with the old centres and if the new centres contribute more to the model than the old centres, replace the old centres with the new ones (Chen *et al.*, 1992).

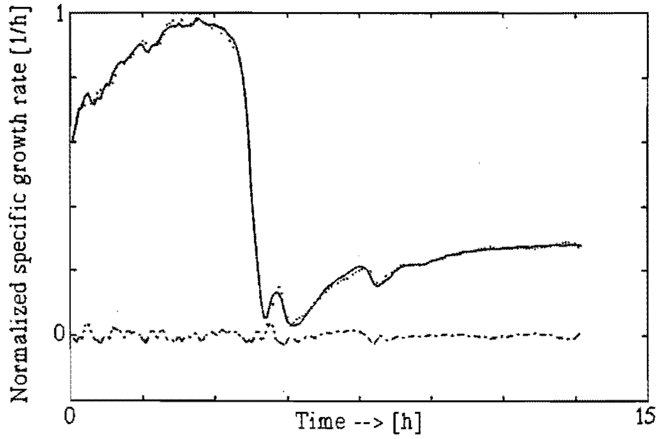
Note that using the option of the algorithm to select a noise model did not lead to a more refined model. To the contrary, after the estimation had converged, the remaining model only consisted of the same process terms as already found in Table 5.4.

The last Example 5.16 is concerned with the reconstruction of the estimated specific growth rate and estimated biomass concentration from the off-gas measurements and lagged estimations.

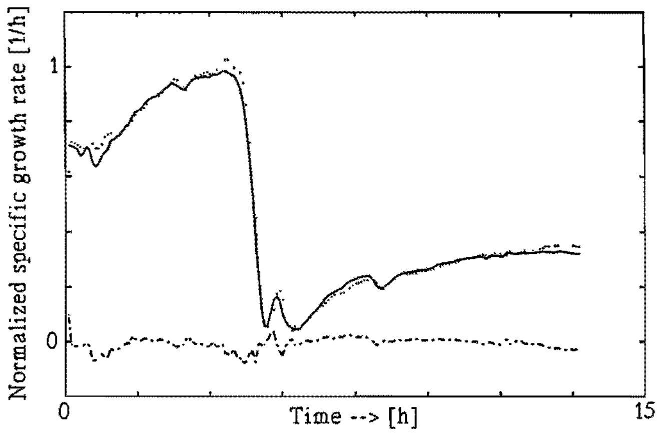
Example 5.16 *Structure identification using Radial Basis Functions; experimental batch data*

Setting:
7) Terminating tolerance for process model *
*) Depending on the model
Data:
Estimation set: batch experiment
Validation set: batch experiment

In Figures 5.13 and 5.14 the results of two RBF networks are presented for the specific growth rate and the biomass concentration.



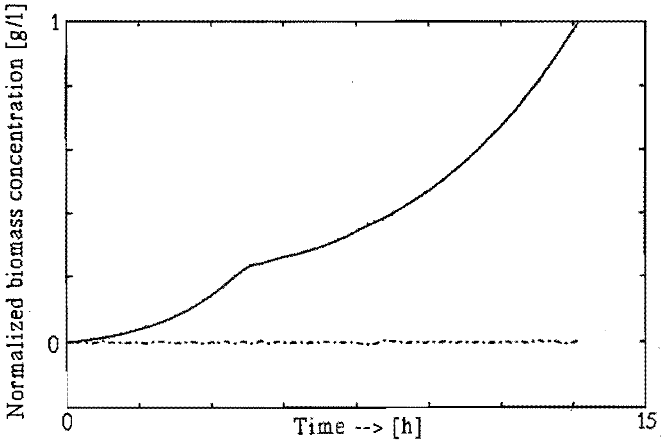
(a)



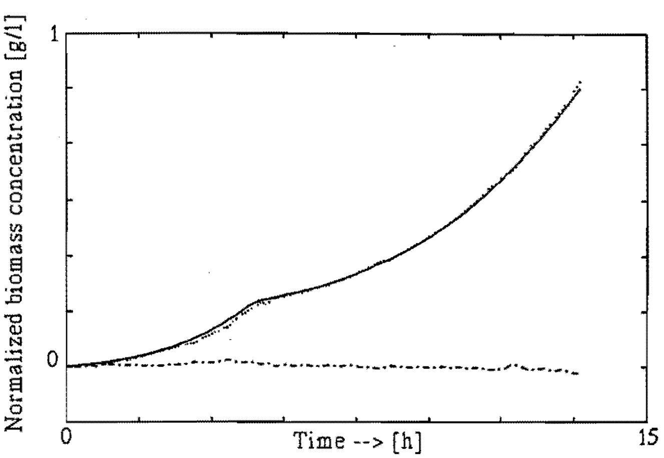
(b)

Figure 5.13 *Estimation results (a) and validation results (b) of the RBF-network identification of the specific growth rate*

The difference between the two models is the accuracy of the process output description for both signals. Increasing the accuracy from 1·10-3 to 5·10-4 for the specific growth rate will result in a network that has one neuron more, 16 instead of 15 neurons.



(a)



(b)

Figure 5.14 *Estimation results (a) and validation results (b) of the RBF-network identification of the biomass concentration*

If we increase the accuracy to $1 \cdot 10^{-4}$ the number of centres increases to 48. We can actually describe the total output with an accuracy of $1 \cdot 10^{-6}$ with not less than 154 centres. Note however that this model is perfect for the estimated output but that using it for any other data set will give poor results because the prediction error is not white. Note that it is known that RBF networks will try to decrease the learning error (estimation error) by fitting the noise rather than the underlying process. The fit does seem to improve but the validation data reveals the misjudgment. ∎

The performance of the model with respect to the validation data is a better way of judging the quality of the model. A good method to take into account the validation data is to stop the increase of centres at the moment the error of the validation data stops decreasing (or starts increasing). The biomass concentration model has an accuracy of $1 \cdot 10^{-4}$ and $1 \cdot 10^{-5}$ based on 10 and 20 centres. Due to the low frequency contents of the biomass concentration output, this output is not modelled as satisfactorily (in terms of whiteness of the prediction error) as the specific growth rate, but if one looks at the total sum of squared errors these models are better.

5.4 Conclusions

The structure identification technique based on Volterra kernels is usable to identify an equation error model for simulation data, generated by the batch phase simulation model. The results are not very promising as the model can not describe the simulation data for more than 70%. The technique can not be applied successfully to the experimental data.

The non-linear model linear-in-the-parameters is capable of modelling transfer functions between the *GF* and *OUR, CPR,* and sometimes the *RQ*. This holds for both simulation data and experimental data. The resulting models are capable of describing the simulation data up to a certain accuracy, but if the specific growth rate is near the critical growth rate, the switching between ethanol production and consumption occurs. This switching is a non-linear function, but can not be described by the models identified by the non-linear model linear-in-the-parameters. Only if the switching is very smooth and rather slow is the identification algorithm capable of identifying an equation error model, with a large influence from the $y(t-1)$ term, which is an integral function thus.

The Radial Basis Function network is also capable of modelling transfer functions between the *GF* and *OUR, CPR,* and sometimes the *RQ*. This holds again for both simulation data and experimental data. The RBF-network creates rather good models for experimental batch phase data, as the data originated from batch phase experiments is rather smooth. Once again the algorithm creates equation error based models of the process.

The non-linear structure identification techniques can help to model part of the laboratory process, but they are not capable of modelling every part of that process and have to be used with specific problems such as sensor failure. Due to their equation error based models they are only capable of one-step ahead prediction, which would be not enough for most control purposes.

We will proceed with some conclusions that are not necessarily based on the data generated by laboratory process or one of its simulation models. These conclusions are general.

The structure identification method based on Volterra kernels gives a nice indication of the structure in terms of blocks. These blocks still have to be identified using normal identification techniques. The algorithm based on Volterra kernels is time consuming, specially if a recursive least squares algorithm has to be used. The technique has no flexibility in dealing with the memory *m*; it has to be chosen beforehand and has to be large enough to get a good impression of the structure and the blocks of the system. The value of the analytical indices is somewhat doubtful as they are derived from the graphical representation and only with rather simple models do they give (useful) information.

The algorithm of Billings *et al.*, (1989) is more flexible, although it has only two selection criteria: the stop-criterium for the process model and the convergence criterium. As has been demonstrated in the test, the algorithm selects its own stop-criterium for process and noise model together; this is sometimes very time consuming. Another advantage of the algorithm of Billings is the automatic estimation of delay terms in the NARMAX model, and the fact that the process can be MIMO.

The Radial Basis Functions network is the most flexible, as almost everything can be selected: complexity of the model, accuracy, convergence, etc. This does not automatically mean that the algorithm will be user-friendly; that depends highly on the detailed knowledge the user has about the applied technique. Furthermore, some signal pre-processing has to be carried out in terms of delays and outliers, which could disturb the applicability of the obtained model. Like the non-linear model linear-in-the-parameters, the RBF-network can provide a MIMO-model.

A comparison of the modelling techniques results in Table 5.5, where a few items are presented. The *complexity of* the non-linear model linear-in-the-parameters is low because the resulting model is built with delayed input/output terms and combinations of these terms; these model terms can be clearly understood. The same holds for the algorithm based on Volterra kernels based algorithm, although the multiplication must still be performed (by the user). As the RBF-network uses centres of data point combinations this is not as simple as the other two.

Due to the fact that for both the non-linear model linear-in-the-parameters and the RBF-network the *accuracy* can be set beforehand, it is normally higher than the algorithm based on Volterra kernels. Furthermore, the fact that for the algorithm based on Volterra kernels no noise model can be incorporated, means that the actually obtained accuracy can be low, specially with non-white noise.

Table 5.5 *Comparison of the non-linear structure identification algorithms*

Criterium	Volterra	Non-linear model linear-in-the-parameters	RBF
complexity	low	low	high
accuracy	low	high	high
insight	high	medium	low
speed	low	high	medium
convergence	?	yes	yes
flexibility	low	medium	high

On the other hand, due to its block-structure the algorithm based on Volterra kernels gives quite a lot *insight* into the process structure, whereas the two others are less informative, especially the RBF-network. The *convergence* of the parameters and structure of the algorithm based on Volterra kernels is not guaranteed because they are selected separately, whereas the non-linear model linear-in-the-parameters and the RBF-network have the most selection criteria, thus the most flexibility. By contrast, the algorithm based on Volterra kernels leaves the user the choice of the structure and the number of parameters with every block involved.

The RBF-network needs to be investigated more profoundly on the most important item: the selection of the centres. If applied in an adaptive control strategy the RBF-network should be able to update its centres with every new sample (Chen et al., 1992). Furthermore, as all of these methods are equation error methods taking into account the process outputs, it is interesting to rewrite these techniques for the output error methodology, taking into account the model outputs. These techniques are candidates for future research.

References

Axellson, J.P. (1989). Modelling and Control of Fermentation Processes. *Ph.D. Thesis*, Lund Institute of Technology, Sweden.

Backx, A.C.P.M., and A.A.H. Damen (1989). Identification of Industrial MIMO Processes for Fixed Controllers. Part 1 General Theory and Practice. *Journal A*, Vol. 1, pp. 3-12.

Bastin, G. and D. Dochain (1990). *On-Line Estimation and Adaptive Control of Bioreactors.* Elsevier, Amsterdam.

Billings, S.A., Chen, S. and M. Korenberg (1989). Identification of MIMO Non-Linear Systems Using a Forward-Regression Orthogonal Estimator. *International Journal of Control,* Vol. 49, No. 6, pp. 2157-2189.

Broomhead, D.S. and D. Lowe (1988). Multivariable Functional Interpolation and Adaptive Networks. *Complex Systems,* Vol. 2, pp. 231-355.

Chen, S., Billings, S.A. and P.M. Grant (1992). Recursive Hybrid Algorithm for Non-Linear System Identification Using Radial Basis Function Networks. *International Journal of Systems Science,* Vol. 21, No. 12, pp. 2513-2539.

Chen, S., Billings, S.A. and W. Luo (1989). Orthogonal Least Squares Methods and their Application to Non-Linear System Identification. *International Journal of Control,* Vol. 50, pp. 1873-1896.

Chen, S., Billings, S.A., Cowan, C.F.N. and P.M. Grant (1990). Practical Identification of NARMAX Models Using Radial Basis Functions. *International Journal of Control,* Vol. 52, No. 6, pp. 1327-1350.

Desrochers, A. and S. Mohseni (1984). On Determining the Structure of a Non-Linear System. *International Journal of Control,* Vol. 40, No.5, pp. 923-938.

Haber, R. (1989). Structural Identification of Quadratic Block-Oriented Models Based on Estimated Volterra Kernels. *International Journal of Systems Science,* Vol. 20, No. 8, pp. 1355-1380.

Haber, R. and H. Unbehauen (1990). Structure Identification of Nonlinear Dynamic Systems - A Survey on Input/Output Approaches. *Automatica,* Vol. 26, No. 4, pp. 651-677.

Hofland, A.G., Morris, A.J. and G.A. Montague (1992). Radial Basis Function Networks Applied to Process Control. *Proceedings of the ACC,* Chicago, USA, pp. 480-484.

Korenberg, M., Billings, S.A., Liu, Y.P. and P.J. McIlroy (1988). Orthogonal Parameter Estimation Algorithm for Non-Linear Stochastic Systems. *International Journal of Control,* Vol. 48, No. 1, pp. 193-210.

Linko, P. and Y.H. Zhu (1992). Neural Network Modelling for Real-Time Variable Estimation and Prediction in the Control of *Glucoamylase* Fermentation. *Process Biochemistry,* Vol. 27, pp. 275-283.

Michelli, C.A. (1986). Interpolation of Scattered Data: Distance Matrices and Conditionally Positive Definite Functions. *Constructive Approximations,* Vol. 2, pp. 11-22.

O'Conner, G.M., Sanchez-Riera, F. and C.L. Cooney (1992). Design and Evaluation of Control Strategies for High Cell Density Fermentations. Submitted to *Biotechnology and Bioengineering.*

Powell, M.J.D. (1985). Radial Basis Functions for Multivariable Interporlation: A Review. *Proceedings of the IMA Conference on Algorithms for the Approximation of Functions and Data,* Shrivenham, UK.

Powell, M.J.D. (1987). Radial Basis Function Approximations to Polynomials. *12th Biennial Numerical Analysis Conference,* Dundee, UK, pp. 223-341.

6

Observers for Control

In the previous two Chapters we presented some methods to make a model of the process. The aim was to develop a suitable control scheme for the process. The models were all based on so called "black box" system descriptions, dealing with almost no a-priori process knowledge, only input-output data. These models could be used for maximizing the yield of the biomass on glucose, so called *RQ*-control. On the other hand, these models are not suitable for controlling the specific growth rate for a certain set-point or pattern. There is no direct relationship between the available inputs (*AF, SS, GF*) and outputs (*OUR, CPR, ethanol*) and the specific growth rate simply because the latter can not be measured on-line. If we want to develop a controller that can control the specific growth rate by manipulating an input variable, estimated or measured data of the specific growth rate must be available. The use of Radial Basis Functions was the best way of building a model for the specific growth rate depending on the glucose feed. Note that estimated specific growth rate data was needed to do so and only a one step ahead prediction could be generated.

Due to the lack of any reliable specific growth rate sensor, the best solution is to observe the specific growth rate based on measurements that are available on-line, e.g. *OUR* and *CPR* (Sonnleitner, 1992). Here we present a simple way of estimating the specific growth rate and its metabolic parts, the oxidative growth on glucose, the reductive growth on glucose and the oxidative growth on ethanol, together with the biomass concentration. A rough estimation of both glucose concentration and ethanol concentration is made as well. All estimations are based on on-line measurements of the oxygen uptake rate and carbon-dioxide production rate plus a-priori knowledge about the stoichiometric relations of the culture *Saccharomyces cerevisiae*. This is the main difference between the previous chapters and this one as here we have to take a-priori knowledge into consideration. This means that where the models of the previous two Chapter were quite generic, the observer presented in this Chapter is not generic at all. If the culture has different stoichiometric constants, the observer has to be altered. Without any doubt this is the main drawback of the presented observer.

This Chapter is organised as follows. Section 6.1. will deal with observers presented in literature. In Section 6.2, the observer will be discussed together with the assumptions made. Then, results of both batch and fed-batch experiments will be given together with two evaluation criteria. The conclusions will be presented in Section 6.4.

6.1 Bio-Process Observers

The feasibility to monitor process parameters of the process is low, see Chapter 2. Most valuable parameters such as glucose and biomass concentration can not be measured on-line. These parameters can be measured off-line, but these off-line measurements are slow, disturb the process, and take a long time to be produced. Due to the fact that the biomass is not measured on-line, related parameters such as the specific growth rate, are not measured as well. There are several sensors available for the specific growth rate and biomass concentration but the best seems to be observer technique (see Sonnleitner, 1992). A nice review on observers for on-line estimation of non-measurable process variables is written by van der Heijden *et al.* (1989). In this Section we will review some of the literature on observers. Furthermore, we will specify the basis of our observer algorithm and what is special about it.

The first estimation techniques reported are those of Cooney *et al.*, (1977) and Zabriskie and Humphrey (1978). The accent of these estimation techniques is on the use of measurements for the indirect estimation of the biomass. They used the total dosed base volume for pH regulation and the used oxygen for estimating the biomass. The approach of Zabriskie and Humphrey (1978) is improved by Mou and Cooney (1983) for a penicillin fermentation. In parallel Aborhey and Williamson (1977) and Holmberg and Ranta (1982) considered the estimation problem from a mathematical viewpoint. Aborhey and Williamson (1977) proposed a non-linear observer based on the Monod equation, where either biomass concentration or glucose concentration is measured, and the other is estimated. In the case that both are measured, one can estimate the parameters. Holmberg and Ranta (1982) considered the case where the biomass is estimated from the off-gas measurements.

The next estimation techniques reported involved some articles on synthesis between mathematical and measurement based observers (Dekkers, 1982; Stephanopoulos and San, 1984; Dochain and Bastin, 1985; Ghoul *et al.*, 1985; Staniškis and Simutis, 1985; Náhlík and Burianec, 1988; Shimizu *et al.*, 1989; Flaus *et al.*, 1991). These techniques use a mathematical procedure and a certain number of indirect or direct measurements together with a Kalman filter for the estimation of variables. Two sorts of algorithms can be distinguished. First there are some algorithms that use the specific growth rate expressed by several other state variables (e.g., Ghoul *et al.*, 1985), secondly there are algorithms that assume that the specific growth rate is a time varying parameter (e.g., Stephanopoulos and San, 1984). In both cases interesting experimental results are obtained. The noise is filtered and some a-priori information is used. There are two inconveniences associated with the methods; the tuning of the Kalman filter, specially the covariance matrices of the noise, and the algorithms are considered not to be robust; no convergence or robustness is proven in these works.

A mathematical observer for non-linear biological systems is presented by Gauthier *et al.*, (1992) that is realised under certain technical assumptions. Furthermore, the inputs should be observable, otherwise no observer can be built. The proposed observer is closely related to high gain observers, has arbitrary exponential decay, and is particularly simple. Chamilothoris (1987), Dochain *et al.* (1989), Dochain and Bastin (1990), and Dochain *et al.* (1992) proposed a method based on a mathematical description of the laboratory process. In this case the specific growth rate is treated as a time varying parameter and is estimated from the filtered, measured biomass concentration. This can also be used if the biomass concentration is measured indirectly. Pomerleau and Perrier (1990) described a method to design a biomass concentration observer and an estimator for multiple specific growth rates. They applied the method to a bakers' yeast fed-batch process where the overall specific growth rate was divided in three components in order to reflect the three main metabolism of the yeast. The method is based on the work done by Bastin and Dochain (1990). The method was evaluated by simulations under different conditions and different sets of measured state variables. Pomerleau and Perrier (1992) presented the experimental validation of their theoretical analyses (Pomerleau and Perrier, 1990) of the estimation of multiple specific growth rates for a fed-batch bakers' yeast process. The accuracy of the estimates of the three specific growth rates involved in this process is verified according to two criteria based on the respiratory quotient and on the evaluation of the ethanol production/consumption rate. The performance of the algorithm is very good and the accuracy is high if the appropriate set of variables is chosen.

The latest step involves neural network, artificial intelligence, fuzzy control. Neural networks to estimate the biomass concentration are discussed by Karim and Rivera (1992; 1992a). Both feed-forward and recurrent network methodologies with their appropriate learning rules are presented. Another feed-forward network approach is presented by Thibault and van Breusegem (1991). Psichogios and Ungar (1992) used a hybrid neural network-first principles modelling scheme to model a fed-batch bioreactor. The hybrid model combines a partial first principles model, the a-priori knowledge, with a neural network which serves as an estimator of unmeasured process parameters that are difficult to model from first principles. This hybrid model has better properties than standard "black-box" neural network models in that it is able to interpolate and extrapolate much more accurately, is easier to analyze and interpret, and requires significantly fewer training sets. The hybrid network model gives better estimates of the parameters and can also make predictions. The results apply when full and partial state measurements are available.

Of all proposed methods only a few try to distinguish between the different growth conditions; oxidative growth on glucose, reductive growth on glucose, and oxidative growth on ethanol. Pomerleau and Perrier (1990,1992) proposed an algorithm to calculate the different growth ways. However the proposed algorithm is rather complex. The aim of the present study is to estimate all specific growth rates and the biomass concentration from as few measurements as possible, using as little a-priori information as possible. The on-line measurements used are the oxygen uptake rate and the carbon-dioxide production rate; the a-priori knowledge concerns the yield coefficients and some ratios between the oxygen uptake rate and carbon-dioxide production rate.

6.2 Model Based Observer

In order to estimate the specific growth rates and the biomass concentration we use a physical model of the process, described in Chapter 3. This physical model is based on four basic pathways Eq. (3.1) - (3.4) for the culture growth. This involves oxidative growth on glucose and ethanol, reductive growth on ethanol, and maintenance.

The on-line measurements that are available are *OUR, CPR,* and *ethanol* concentration. All these measurements are made using the mass-spectrometer as a sensor (see Chapter 2). In order to be able to observe the specific growth rates adequate knowledge about the condition of the culture has to be obtained. The condition of the culture can be defined as producing or consuming ethanol, maintenance, etc. As we have an ethanol concentration measurement available, a simple way to detect the ethanol related conditions would be to take the derivative of the ethanol signal. If positive, the condition is ethanol production; if negative, the condition is ethanol consumption. Due to the high noise level, this solution is not applicable.

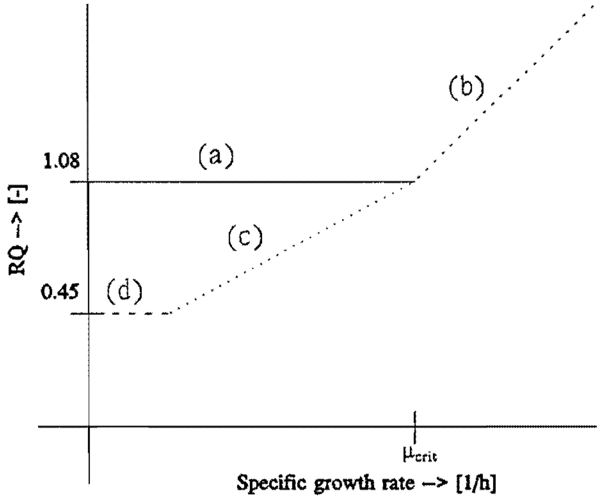


Figure 6.1 *The relation between μ and RQ for all pathways; figures calculated from Eq. (3.1) - Eq. (3.4), (a) oxidative growth on glucose, (b) reductive growth on glucose, (c) oxidative growth on glucose and ethanol, and (d) oxidative growth on ethanol*

Another mechanism to detect the condition of the culture is the respiratory quotient (*RQ*) value (see Figure 6.1). In literature one concludes mostly that ethanol is produced at a $RQ > 1.08$ (b) and consumed at an $RQ < 1.08$ (c,d). For some production strategies the growth on glucose should be without ethanol production. This means that the specific growth rate should always be below the critical specific growth rate, μ_c, i.e. the value at which the culture start producing ethanol. If the specific growth rate is below μ_c and no ethanol is present the line (a) in Figure 6.1 is valid. The *RQ* is equal to 1.08 for any specific growth rate below the critical growth rate. This means that the above described conditional description of the yeast is incomplete. Another, third condition has to be defined.

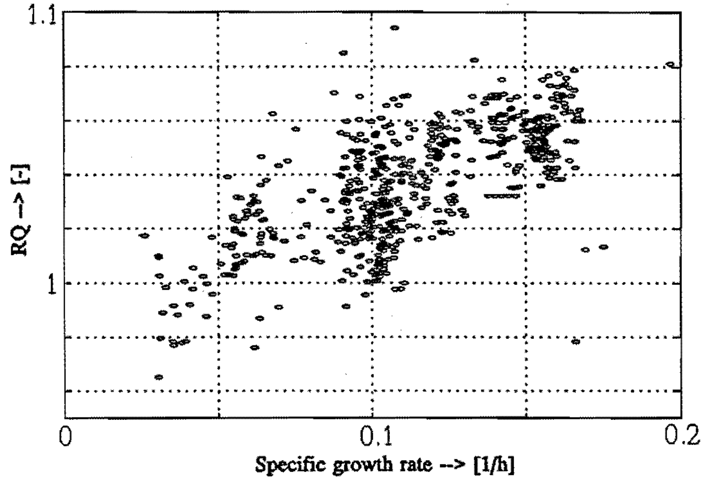


Figure 6.2 *The relation between μ and RQ for oxidative growth; data calculated from fed-batch experiments*

If we stick to the relationship between the specific growth rate and the *RQ* as shown in Figure 6.1 we have a problem. There is one *RQ* value (1.08) for a whole range of specific growth rate values. If we express this in terms of equations we obtain:

$$\begin{aligned} 0 \leq \mu \leq \mu_{crit} &\rightarrow RQ \leq 1.08 \\ \mu > \mu_{crit} &\rightarrow RQ > 1.08 \end{aligned} \tag{6.1}$$

The data from Figure 6.1 is taken from continuous experiments, where the process is in steady state. If we use the data collected from various fed-batch experiments we obtain Figure 6.2. Again there is a (linear) relation between the specific growth rate and the *RQ* but this time it is not a horizontal line. We will assume that the relationship is a linear one as this will simplify calculations. Taking the experimental data into account we can reformulate Eq. (6.1):

$$\begin{array}{lll} 0 \leq \mu \leq \mu_{crit} & 1.00 \leq RQ \leq 1.08 & \textit{only glucose present} \\ 0 \leq \mu \leq \mu_{crit} & 0.45 \leq RQ < 1.00 & \textit{ethanol and glucose present} \\ \mu > \mu_{crit} & RQ > 1.08 & \textit{ethanol production} \end{array} \tag{6.2}$$

Note that the values for the *RQ*-boundaries (0.45, 1.00 and 1.08) are valid for this culture only and should be determined experimentally.

The experimental data is according to the relationship we obtain if we rewrite Eq. (3.1) and (3.4) for the oxidative growth on glucose only:

$$RQ = \frac{c \cdot \alpha \cdot \mu + (b+c) \cdot ms}{a \cdot \alpha \cdot \mu + (a+\alpha \cdot b) \cdot ms} \tag{6.3}$$

This relationship is valid for both fed-batch and continuous cultures, showing that both Figure 6.1 and Eq. (6.1) are incorrect. Furthermore, the relationship found Eq. (6.3) is not linear but a Monod type one. Nevertheless, we will use the linear one.

Now we have a three condition process description, Eq. (6.2), based on measurements of the *RQ* or, rather, based on *OUR* and *CPR*. Here we assume that the resulting *RQ* will be below 1.00 for ethanol consumption. We make a small mistake here, as ethanol will be consumed for 1.00 < RQ < 1.08 as well, but we have a better conditional description for oxidative glucose consumption.

Calculation of μVX

Based on the measurements of the *CPR* and *OUR* the biomass production rate can be calculated using Eq. (3.1) - (3.4). Firstly the *CPR* and *OUR* measurements are corrected for the maintenance effect:

$$\begin{aligned} OUR' &= OUR - ms \cdot V\hat{X} \\ CPR' &= CPR - \frac{ms}{\alpha} \cdot V\hat{X} \end{aligned} \tag{6.4}$$

Here we assume to know the maintenance coefficient and take $V\hat{X}$ as the estimated biomass as the true value is not available. From Eq. (6.2) we know that with a certain ratio of *OUR'* and *CPR'* the yeast is either producing or consuming ethanol or growing oxidatively on glucose only. This is visualised in Figure 6.3. In case (a) there is no ethanol production or consumption, 1.00 < *RQ* < 1.08. If there is ethanol present and it is consumed, case (c), the ratio is smaller than 1.00. On the other hand if ethanol is produced, case (b), the ratio is bigger than 1.08. We can use the *OUR*' and *CPR*' measures, together with the ratio information, to calculate μ*VX*. Furthermore, using the ratios of Eq. (3.1) - (3.4), we can discriminate between oxidative growth on

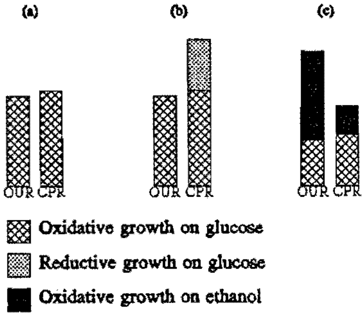


Figure 6.3 *Growth on glucose (a,b,c) and ethanol (c) represented by the OUR and CPR.*

glucose $\mu_o VX$, on ethanol $\mu_e VX$, and reductive growth on glucose $\mu_r VX$. This has been summarised in Table 6.1

The biomass production rate is (including the maintenance term):

$$\mu VX = \mu_o VX + \mu_r VX + \mu_e VX - \frac{ms}{\alpha} VX \tag{6.5}$$

The parameters δ, ζ and γ have to be calculated using experimental results. They denote the boundaries in the *RQ*-μ graph. ζ is the *RQ* value at which ethanol production occurs, γ the lower limit for no ethanol consumption and δ the *RQ* value for pure ethanol consumption. In this study we assume $\zeta = 1.08$, $\gamma = 1.00$ and $\delta = 0.45$

Table 6.1 *The contribution of different growth rates to different pathways*

specific growth rate	ethanol production	ethanol consumption	no ethanol present
$\mu_o VX$ (oxidative growth on glucose)	$\frac{b}{a} \cdot OUR'$	$\frac{b}{a \cdot m - c \cdot k} \cdot (OUR' - \delta \cdot CPR')$	$\frac{b}{c} \cdot CPR'$
$\mu_r VX$ (fermentative growth on glucose)	$\frac{g}{h} \cdot (CPR' - \zeta \cdot OUR')$	0	0
$\mu_e VX$ (oxidative growth on ethanol)	0	$\frac{l}{k-m} \cdot (OUR' - \gamma \cdot CPR')$	0

Estimation of $V\hat{X}$ and $\hat{\mu}$

Having calculated μVX we use the mass balance to estimate $V\hat{X}$ and $\hat{\mu}$. In order to estimate $V\hat{X}$ we use:

$$V\hat{X}(t) = \int_{\tau=0}^{t} \mu VX(\tau) + V\hat{X}(0) \tag{6.6}$$

where $V\hat{X}(0)$ is the initial estimate of the biomass in the fermenter. Eq. (6.6) is a continuous function, yet the *OUR* and *CPR* measurements are sampled data measurements with a sample time of at least 1 minute. The discrete version of Eq. (6.6) could be (Euler approximation):

$$V\hat{X}(k) = V\hat{X}(k-1) + T_s \cdot \mu \cdot VX(k-1) \tag{6.7}$$

where T_s is the sampling time in hours and $V\hat{X}(k)$ denotes the amount of biomass in mole at time kT_s. Note that $\mu VX(k\text{-}1)$ is calculated by using the equations in the previous section. When using Eq. (6.7) one assumes that the *OUR* and *CPR* stay constant during the sample. This is not the case as the yeast grows exponentially. There are several methods to incorporate the exponential increase of the yeast in Eq. (6.7), e.g. through linear interpolation, etc. The time constant of the yeast dynamics, however, shows that the difference between applying a (linear) interpolation or a constant is not of any influence on the accuracy obtained.

Now from the above equations one can easily calculate $\hat{\mu}(k)$

$$\hat{\mu}(k) = \frac{\mu VX(k)}{V\hat{X}(k)} \tag{6.8}$$

Beside the overall specific growth rate μ, the oxidative growth rate on glucose μ_o, the reductive growth rate on glucose μ_r and the oxidative growth rate on ethanol μ_e are estimated together with an a-priori known maintenance term.

The convergence of the estimator for the biomass is guaranteed by its simplicity, see Eq. (6.6) and (6.8). If the estimate of the biomass at $t=0$ is incorrect, the growth rate will be incorrect as well ($V\hat{X}(0)$ too high → $\hat{\mu}(k)$ too low, $V\hat{X}(0)$ too low → $\hat{\mu}(k)$ too high) causing the algorithm to converge to its true value. The convergence time is dependent on the accuracy of the estimation of $V\hat{X}(0)$ and the sampling time.

6.3 Results

The observer is meant to be a basis for control design, especially the case where the control of the specific growth rate is crucial one can not do anything without measuring or observing μ. In a well designed fed-batch phase experiment the oxidative growth on ethanol and the reductive growth on glucose do not normally occur or to a very small extend. In order to test these parts of the specific growth rate a batch phase experiment is included. During a batch fermentation, substantial amounts of ethanol are produced and consumed, see Section 3.3. The batch fermentation results are presented in Example 6.1.

Example 6.1 *Estimation of specific growth rates of a batch experiment of Saccharomyces cerevisiae*

Initial conditions:
  X(0) = 0.25 [g.dm^{-3}]
  S(0) = 30 [g.dm^{-3}]
  E(0) = 0 [g.dm^{-3}]
  V(0) = 4.4 [dm^3]
Measurements:
  OUR [mole.h^{-1}]
  CPR [mole.h^{-1}]
  time = 16 [h]
Observations: (^)
  μ, μ_o, μ_r, μ_e, X, S, and E

The initial conditions are the same for all batch experiments (see Chapter 2). The biomass concentration *X(0)* is measured using optical density measurements. The results obtained with the observer are shown in Figure 6.4 and 6.5. The solid line in Figure 6.4, near the bottom of the plot, represents the assumed maintenance contribution *ms/α* for the yeast. Furthermore one can see that μ_r is greater than zero only in the first part of the batch phase, as long as glucose is present. On the other hand, μ_e is exactly opposite to μ_r, only increasing above zero during the second part of the batch phase when ethanol is consumed. This is according to the theory where glucose uptake inhibits ethanol uptake and ethanol production only occurs if glucose is present in excess of the oxygen bottleneck (Sonnleitner and Käppeli, 1986).

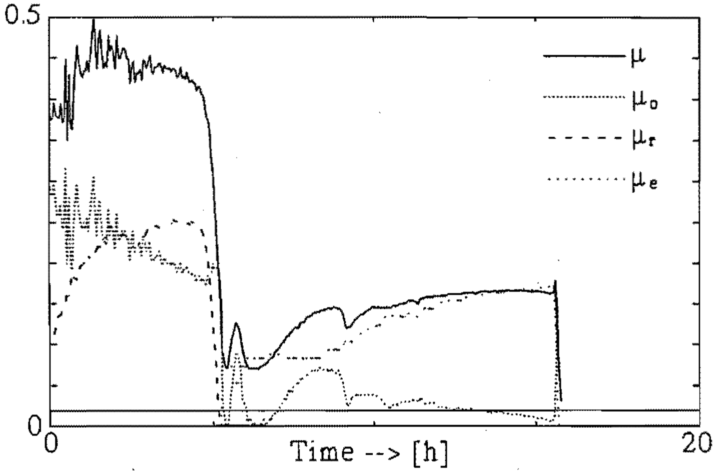


Figure 6.4 *The specific growth rates for batch fermentation. (-) the overall specific growth rate μ, (.) oxidative growth on glucose μ_o, (--) fermentative growth on glucose μ_r and (-.) oxidative growth on ethanol μ_e, the solid line at the bottom is the maintenance*

During the first part of the batch phase (0-5 [h]) the culture grows on glucose, reductively (μ_r) and oxidatively (μ_o). During the second part (5-16 [h]) the yeast should, according to the theory, consume the produced ethanol and grow oxidatively (μ_e). From Figure 6.4 we can see that μ_o does not become zero during the second part, indicating oxidative growth on glucose. The explanation is that extracellular components such as pyruvate, acetate and acetaldehyde are consumed as well (Dantigny *et al.*, 1989) during the second part of the batch phase. Normally these phenomena are imbedded in either the overall growth rate (μ) or the oxidative growth rate on ethanol, μ_e. In our algorithm the uptake of these extracellular components is explained by the oxidative growth rate on glucose μ_o. The production of these extracellular components occurs in the first phase (acetaldehyde) or during the lag phase (pyruvate and acetate). The slow decline of μ_o during the first part of the batch phase indicates that the yeast cells either are inhibited by the ethanol produced or are slowly adapting to the new situation. The growth rate decreases from +/- 0.30 [h^{-1}] to 0.17 [h^{-1}], which is low. Note that effluent oxygen was present during the experiment. The maximum μ_r is 0.25 [h^{-1}] and μ_o is 0.16-0.17 [h^{-1}] where the maximum overall growth rate is 0.45-0.50 [h^{-1}]. All correspond with data found in literature.

The estimation of the amount of biomass is shown in Figure 6.5, as well as the estimated amount of glucose and ethanol. The biomass pattern corresponds with patterns found in literature (e.g. Sweere, 1988). No optical density measurements are available for a validation of the results. The same holds for the glucose and the ethanol pattern. The glucose is not fully consumed when the ethanol consumption starts, whereas it is known that the glucose inhibits ethanol uptake. Also the ethanol consumption is lower than expected, leaving some ethanol in the fermenter. A possible explanation could be incorrect kinetics associated with μ_r and μ_o. Another explanation could be the influence of the extracellular components. ■

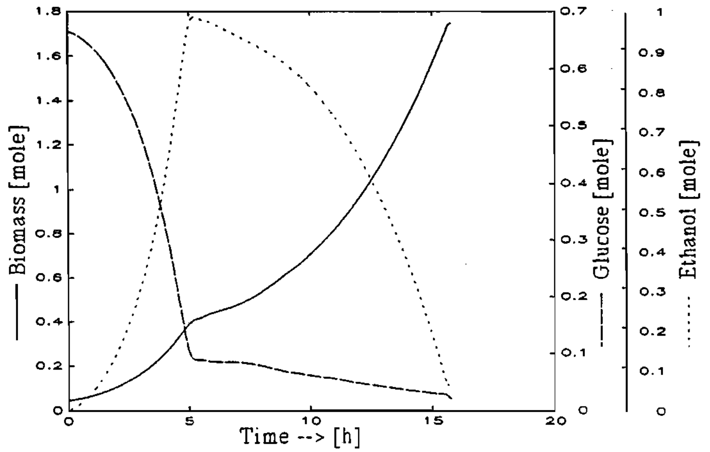


Figure 6.5 *Estimated biomass (-), glucose (--) and ethanol (.) concentrations for a batch fermentation*

The amount of biomass at the end of the batch phase, about 10 [g.dm^{-3}], is close to the results of the dry-weight measurements for several experiments, 10.5 [g.dm^{-3}]. This is an error of about 5 %. This error is caused by disturbances on the *OUR* and *CPR* caused by the mass-spectrometer. Unfortunately these errors can not be corrected on-line due to their random nature.

Now we will present an experimental fed-batch experiment, grown on molasses. The experiment will be used to evaluate the observer presented here. During the experiment a small amount of ethanol was produced and consumed. Furthermore, there were some disturbances that could not be accounted for. The evaluation is done according to two criteria. The first criterion is based on the *RQ*. As is well known, we can calculate the *RQ* from :

$$RQ(t) = \frac{CPR(t)}{OUR(t)} \tag{6.9}$$

but using the estimates of $\hat{\mu}_o$, $\hat{\mu}_r$ and $\hat{\mu}_e$ we can also calculate $R\hat{Q}$:

$$R\hat{Q}(t) = \frac{\frac{c}{b}\cdot\hat{\mu}_o + \frac{h}{g}\cdot\hat{\mu}_r + \frac{m}{l}\cdot\hat{\mu}_e + \frac{ms}{\alpha}}{\frac{a}{b}\cdot\hat{\mu}_o + \frac{k}{l}\cdot\hat{\mu}_e + ms} \tag{6.9}$$

Comparing the estimated $R\hat{Q}$ with the measured RQ gives an indication of the accuracy of the growth rates estimates.

A second criterion is based on the derivative of the low pass filtered ethanol concentration and the estimated rate derived from the two ethanol specific rates, μ_r and μ_e. The low pass filtering of the derivative of ethanol is necessary because of the sensor noise. The filter is a second order Butterworth with a cut off frequency equal to 5 [Hz]. Comparing both measurements also gives an indication of the performance of the estimator.

Example 6.2 *Estimation of a fed-batch culture of Saccharomyces cerevisiae grown on molasses*

Initial conditions:
X(0) = 11.6 [g.dm^{-3}]
S(0) = 2.5 [g.dm^{-3}]
E(0) = 0 [g.dm^{-3}]
V(0) = 5 [dm^{3}]
Measurements:
OUR [mole.h^{-1}]
CPR [mole.h^{-1}]
time = 17 [h]
Observations: (^)
μ, μ_o, μ_r, μ_e, and X

The feed used in this experiment is molasses. The glucose contents of the molasses is 50% and the specific weight of the molasses is 1.36 10^3 [g.dm^{-3}]. The peaks seen in Figure 6.6 part (a) after 10 [h] are caused by disturbances in the off-gas measurements. These disturbances act severely on the estimated specific growth rates because they are not filtered. The effect on the biomass concentration estimate is less profound due to the integral action in estimating the biomass, which acts as a low pass filter. As can be seen from part (a), there are moments where ethanol is produced (μ_r) and others where it is consumed (μ_e), though the contribution of both is small compared to the oxidative growth on glucose (μ_o).

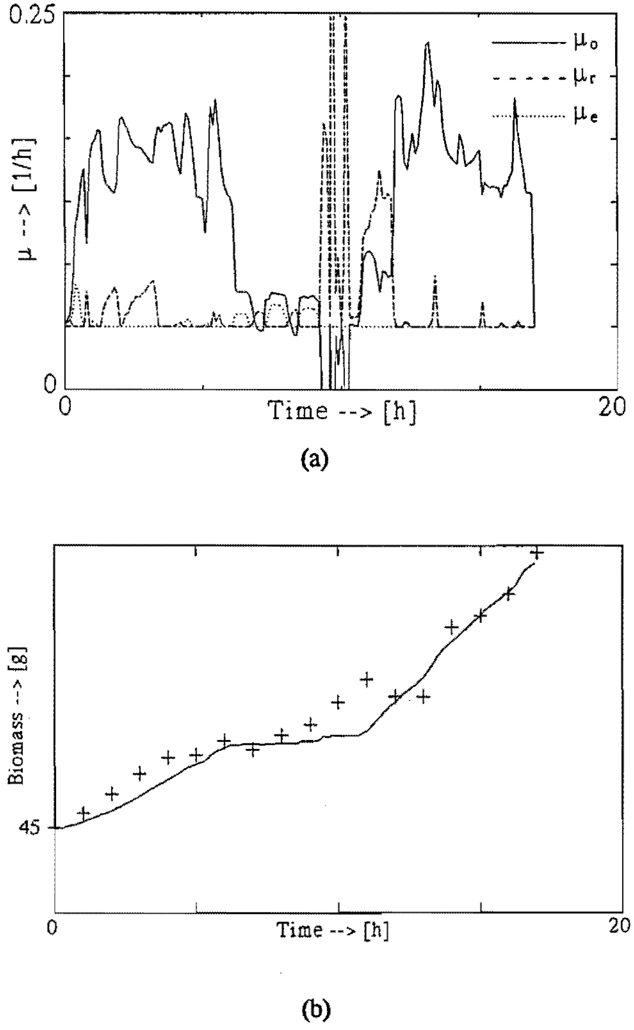
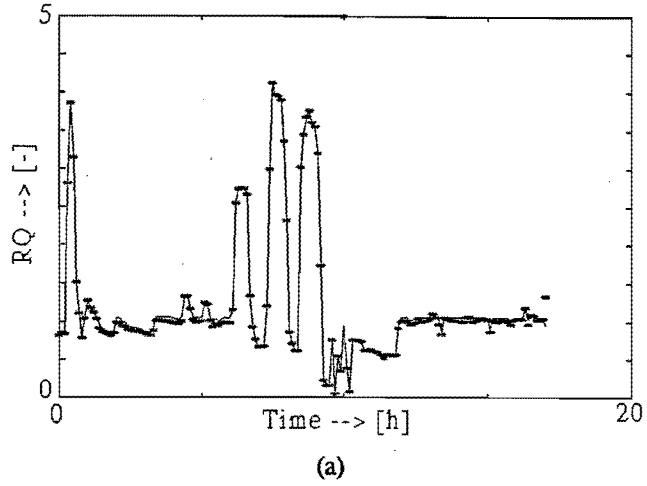


Figure 6.6 *Estimated specific growth rates, part (a), and estimated biomass concentration, part (b); the (+) in part (b) denotes the optical density measurements*

The estimation of the biomass is good if one compares it with the optical density measurements, (+) in part (b) of Figure 6.6. If we test these results against our two criteria we obtain Figure 6.7. The estimated and calculated RQ are almost identical: the small error may be caused by a wrong estimate of the maintenance contribution.



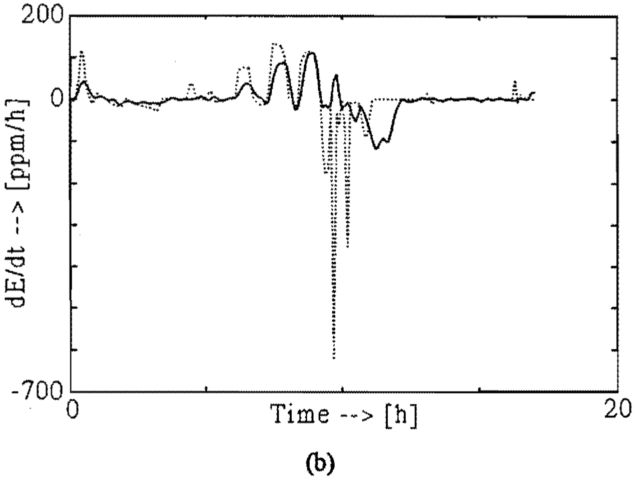


Figure 6.7 *Estimated RQ (-) and calculated RQ (*), part (a), and estimated dE/dt (--) and measured dE/dt (-), part (b)*

The second criterion, the derivative of the ethanol concentration, is less smooth due to the sensor noise. Yet we see that again there is a good agreement between production and consumption phases of the ethanol, except after 10 [h] the estimated dE/dt shows a noisy behaviour due to the wrongly estimated growth rates. ■

6.4 Conclusions

The algorithm presented here is very simple. It is based on a-priori stoichiometric knowledge about the process which makes the algorithm non-generic. Its applicability to other, similar cultures is quite good due to its simplicity. Another advantage of its simplicity is the convergence property.

The algorithm can distinguish between four different pathways: oxidative growth on glucose (μ_o), reductive growth on glucose (μ_r), oxidative growth on ethanol (μ_e), and maintenance. Furthermore, the biomass concentration is estimated together with a rough estimate of the glucose and ethanol concentration, both calculated from the specific growth rates.

Verification against off-line measurements of the biomass yield satisfactory results for both batch and fed-batch experiments, showing the accuracy of the algorithm. From the *RQ* and derivative of the ethanol concentration it is concluded that the specific growth rate and its metabolic parts are estimated correctly. Disturbances on the measurements will lead to wrong specific growth rate estimates. The biomass concentration estimate has low pass properties such that disturbances are attenuated.

The estimated specific growth rates (μ_o, μ_r and μ_e) can facilitate the control design for many of the known control problems, e.g. optimal production of yeast (→ maximizing μ_o and minimizing μ_r), ethanol production (→ tracking of μ_r and minimizing μ_e), etc.

References

Aborhey, S. and D. Williamson (1977). State and Parameter Estimation of Microbial Growth Processes. *Automatica,* Vol. 14, pp. 493-498.

Bastin, G. and D. Dochain (1990). *On-Line Estimation and Adaptive Control of Bioreactors*. Elsevier, Amsterdam.

Chamilathoris, G. (1987). Techniques Adaptatives pour le Suivi et la Conduite des Processus de Fermentation. *Ph.D. Thesis.* Université de Toulouse, France.

Cooney, C.L., Wang, H.Y., and D.I.C. Wang (1977). Computer-Aided Material Balancing for Fermentation Parameters. *Biotechnology and Bioengineering*, Vol. 19, pp 55-67.

Dantigny, P., Ninow, J.L., Marc, I., and J.M. Engasser (1989). Representation of Changes in the Metabolic Pattern of Bakers' Yeast from Measurements of Extracellular Pyruvate, Acetate, Acetaldehyde and Ethanol. *Biotechnology Letters*, Vol. 11, No. 7, pp 515-520.

Dekkers, R.M. (1982). State Estimation of a Fed-Batch Bakers' Yeast Fermentation. *IFAC Symposium Modelling and Control of Biotechnical Processes,* Helsinki, Finland, pp. 201-211.

Dochain, D. and G. Bastin (1985). Stable Adaptive Algorithms for Estimation and Control of Fermentation Processes. *IFAC Symposium, Modelling and Control of Biotechnological Processes*, Noordwijkerhout, The Netherlands, pp. 37-42.

Dochain, D., De Buyl, E. and G. Bastin (1989). Experimental Validation of a Methodology for On-Line State Estimation in Bioreactors. In: *Computer Applications in Fermentation Technology: Modelling and Control of Biotechnological Processes,* (Eds. Fish, N.M. *et al.*) London-New York: Elsevier Applied Science, London - New York, pp. 187-194.

Dochain, D., Perrier, M. and B.E. Ydstie (1992). Asymptotic Observers for Stirred Tank Reactors. *Chemical Engineering Science,* Vol. 47, No. 15/16, pp. 4167-4177.

Flaus, J.M. Cheruy, A., Engasser, J.M., Poch, M. and C. Sola (1991). Estimation of the State and Parameters in Bioprocesses from Indirect Measurements. *Proceedings of the ECC*, Grenoble, France, pp. 1642-1647.

Gauthier, J.P., Hammouri, H. and S. Othman (1992). A Simple Observer For Non-Linear Systems Applications to Bioreactors. *IEEE Transactions on Automatic Control*, Vol. 37, No. 6, pp. 875-880.

Ghoul, A., Pons, M.N., Engasser, J.M. and J. Bordet (1985). Extended Kalman Filtering Technique for the Online Control of Candida Utilis Production. *IFAC Symposium, Modelling and Control of Biotechnological Processes.* Noordwijkerhout, The Netherlands, pp. 165-170.

Holmberg, A. and J. Ranta (1982). Procedure for Parameter and State Estimation of Microbial Growth Process Models. *Automatica,* Vol 18, pp. 181-193.

Karim, M.N. and S.L. Rivera (1992). Application of Neural Networks in Bioprocess State Estimation. *Proceedings of the ACC,* Chicago, USA, pp. 495-499.

Karim, M.N. and S.L. Rivera (1992a). Use of Recurrent Neural Network for Bioprocess Identification in On-Line Optimization by Micro-Genetic Algorithms. *Proceedings of the ACC,* Chicago, USA, pp. 1931-1932.

Keulers M. and G. Reyman (1991). The Application of the GPC Algorithm to a Fed-Batch Fermentation Process. - A Simulation Study -. *Proceedings of the ECB*, Copenhagen, Denmark, pp. 879-884.

Náhlík, J. and Z. Burianec (1988). On-Line Parameter and State Estimation of Continuous Cultivation by Extended Kalman Filter. *Applied Microbiology and Biotechnology*, Vol. 28, pp. 128 - 134.

Mou, D.-G. and C.L. Cooney (1983). Growth Monitoring and Control Through Computer-Aided On-Line Mass Balancing in a Fed-Batch Penicillin Fermentation. *Biotechnology and Bioengineering*, Vol. 25, pp. 225-255.

Pomerleau, Y. and M. Perrier (1990). Estimation of Multiple Specific Growth Rates in Bioprocesses. *AIChE Journal*, Vol. 36, No. 2, pp. 207 - 215.

Pomerleau, Y. and M. Perrier (1992). Estimation of Multiple Specific Growth Rates: Experimental Validation. *AIChE Journal*, Vol. 38, No. 11, pp. 1751-1760.

Psichogios, D.C. and L.H. Ungar (1992). A Hybrid Neural Network-First Principle Approach to Process Modelling. *AIChE Journal,* Vol. 38, No. 10, pp. 1499-1511.

Shimizu, H., Araki, K., Shioya, S., Suga, K. and E. Sada (1989). Optimal Production of *Glutathione* by Controlling the Specific Growth Rate of Yeast in Fed-Batch Culture. *IFAC Symposium, Production Control in the Process Industry*, Osaka, Japan, pp. 67-72.

Shimizu, H., Takamatsu, T., Shioya, S. and K.-I. Suga (1989a). An Algorithmic Approach to Constructing the On-Line Estimation System for the Specific Growth Rate. *Biotechnology and Bioengineering*, Vol. 33, pp. 354-364.

Shimizu, H., Shioya, S., Suga, K.-I., and T. Takamatsu (1989b). Profile Control of the Specific Growth Rate in Fed-Batch Experiments. *Applied Microbiology and Biotechnology*, Vol. 30, pp. 276-282.

Sonnleitner, B. (1992). On Line Measurement of Cell Concentration. *Process Control and Quality*, No. 2, pp. 97-104.

Sonnleitner, B., and O. Käppeli (1986). Growth of *Saccharomyces cerevisiae* is

Controlled by its Limited Respiratory Capacity: Formulation and Verification of a Hypothesis. *Biotechnology and Bioengineering*, Vol. 28, pp. 927-937.

Staniškis, J. and R. Simutis (1986). A Measuring System for Biotechnical Processes Based on Discrete Methods of Estimation. *Biotechnology and Bioengineering*, Vol. 28, pp. 362-371.

Stephanopoulos, G. and K. San (1984). Studies of On-Line Bioreactor Identification, I, II, III and IV. *Biotechnology and Bioengineering*, Vol. 26, pp. 1176-1218.

Swiniarski, R., Lesniewski, A., Dewshi, M.A.M., Ng, M.H. and J.R. Leigh (1982). Progress Towards Esitmation of Biomass in a Batch Fermentation Process. *IFAC Symposium, Modelling and Control of Biotechnical Processes*, Helsinki, Finland, pp. 231-241.

Thibault J. and V. van Breusegem (1991). Modelling, Prediction and Control of Fermentation Processes via Neural Networks. *Poceeding of the ECC*, Grenoble, France, pp. 224-229.

van der Heijden, T.J.M., Hellinga, C., Luyben, K.Ch.A.M. and G. Honderd (1989). State Estimators (Observers) for the On-Line Estimation of Non-Measurable Process Variables. *TIBTECH*, Vol. 7, pp. 205-209.

Wang, H.Y., Cooney, C.L., and D.I.C. Wang (1977). Computer-Aided Bakers' Yeast Fermentations. *Biotechnology and Bioengineering*, Vol. 19, pp. 69-86.

Zabriskie, D.W. and A.E. Humphrey (1978). Real-Time Estimation of Aerobic Batch Fermentation Biomass Concentration by Component Balancing. *AIChE Journal*, Vol. 24, pp.138-146.

7

Control

In this Chapter we will design and test a specific growth rate controller. In the previous Chapters various models, for simulation (Chapter 3) and for identification (Chapter 4 and 5), have been described. As one of our key variables, the specific growth rate, is non-measurable an observer (Chapter 6) has been developed to estimate the specific growth rate. All of these models and the observer are the prerequisites to a specific growth rate controller.

The simulation models of Chapter 3 will be used to test and tune a controller. The test of the controller on simulation models has advantages: quick tuning, no experimental costs, faster than the real experiments, etc. It also has a disadvantage: a model is always a description of reality, so tuning parameters based on simulation models may be entirely wrong in practice.

The models obtained through linear and non-linear identification (Chapter 4 and 5) can serve as models for a (model based) controller. In this work we have chosen to start with a PI-controller. The choice to start with a PI-controller is based on several factors: we tried to keep the control as simple as possible, the PI-controller needs almost no maintenance, its algorithm is easy to understand even for non-control engineers, and we know from the linear models (Chapter 4) that after linearization the resulting process is of low order. The PI-controller needs no (explicit) model to control the process. From Chapter 4 we use the non-linear compensation to linearize the process and apply the PI-controller to control the linearized process. None of the models found in Chapter 5 has been used to control the process. They can be used to serve as one-step ahead predictors to replace wrong sensor readings.

The reason to use the observer is quite obvious: the specific growth rate is not measurable.

The control objective is reformulated from the problem statement in Chapter 1. Essentially there are two control objectives:

- maximizing the yeast production and minimization of the nutrient consumption simultaneously (fed-batch phase)
- tracking of a pre-defined specific growth rate pattern (fed-batch and qualtity phase)

There is a difference between the two objectives although the first could be part of the second. The first objective can be solved using control techniques such as *RQ*-control. The second objective can only be tackled using a specific growth rate controller or by pre-specified feedforward trajectories. In this work we will focus on the second objective.

Together with one of the control objectives the user can set some additional restrictions namely:

- the ethanol concentration does not exceed a predefined user set limit
- the dissolved oxygen tension in the broth does not fall below a pre-described percentage (in order to avoid oxygen limitation).

Many different control approaches have been described, see Section 7.1, to satisfy the control objectives formulated above. Especially a lot of attention has been given to the first objective together with the two restrictions. Few approaches have addressed the problem of tracking some specific growth rate patterns. This can be attributed to the fact that so far no reliable growth rate sensor is available, other than the observer based ones.

The main industrial cost involved in the production of yeast is the price of glucose (nutrient). Less important are the expenditures of the energy required for rotating the stirrer and for blowing air through the broth. Minimising these costs is a secondary goal. At present, a tight μ response is considered most important.

This chapter is organised as follows. A literature review of various control strategies and applications will be treated in the first section. Next, a control strategy will be presented which will be tested on the simulation models (Section 7.2). Some enhancements will be suggested, tested and implemented leading to a final control scheme. This control scheme is then tested on experimental data in Section 7.3. Modifications to the initial controller set-up are made such that the controller can control the specific growth rate. The conclusions, Section 7.4, will close this chapter.

7.1 Literature Review

The first articles on control of biotechnological processes were published in the seventies (Constantinides, 1970; D'ans *et al.*, 1972; Takamatsu *et al.*, 1975; Onho *et al.*, 1976; Cheruy and Durand, 1976; Yamané *et al.*, 1977). These articles concentrated on optimal trajectories for the process such that productivity was maximized. The principle of such an approach is to construct a detailed model, a suitable criterion for the productivity, and then calculate the optimal trajectory by minimizing (or maximizing) the criterion. Some constraints may act on input variables or intermediate ones. The methods used are based on the maximum principle of Pontryagin, the theory of Green, or the dynamic programming principle of Bellman. In the case of continuous operation, for which one can define stationary states, one can restrict oneself to stationary optimisation, which will give the optimal working point. In case of batch or fed-batch phase there are optimal trajectories. This approach is successful if one has a good process model, which is not often the case. The control strategy is of the feedforward type, the procedure is thus very sensitive to disturbances or perturbations.

Somewhat later, late seventies and early eighties, articles were published on the application of computer control to the fed-batch phase of bakers' yeast production (Aiba *et al.*, 1976; Wang *et al.*, 1977; Wang *et al.*, 1979; Veres *et al.*, 1981; Cooney and Swartz, 1982; Pons *et al.*, 1982). The strategies involved maximization of the biomass yield and productivity. This goal requires control of the specific growth rate below the critical value (see Chapter 3) in order to prevent ethanol formation. To meet the second objective, productivity, the specific growth rate should be as close as possible to the critical growth rate. Most of the control schemes are closed loop strategies that obtain the required feedback information from the ethanol concentration or the effluent-gas analysis, *OUR*, *CPR*, and *RQ*. As has been elucidated in Chapter 6, using the *RQ* or the ethanol as feedback information can lead to sub-optimal feeding, having maximum yield but low productivity.

The experimental results reported in the articles published in the seventies (see above) were not very satisfactory. To improve the control performance adaptive schemes were suggested, all for continuous operated fermenters (Biryukov, 1982; Rolf and Lim, 1985; Wu *et al.*, 1985; Harmon *et al.*, 1987; Golden and Ydstie, 1987; Modak and Lim, 1992). Another possibility to improve the performance is to split up the problem. From one part one defines trajectories for principal variables. These profiles are designed through experimental and theoretical models. From the other part a controller is designed to keep the variables on the trajectories (Takamatsu *et al.*, 1985). One of the biggest advantages is that the perturbations are dealt with.

Different controllers are used, the most simple one is the PID-controller (Dairaku *et al.*, 1983; Suzuki *et al.*, 1986; Cardello and San, 1988; Axellson, 1989). These articles show that it is difficult to obtain good regulation during the total duration of the fermentation. Different possibilities to change the parameters of the controller are discussed which improved the controller performance. The results obtained with the PID-controller show that this technique is limited, which is caused by the changing dynamics during the fermentation.

Numerous techniques utilizing an adaptive controller have been reported trying to compensate the dynamics and non-linearities during the fermentation. One can distinguish the linear adaptive controllers (Waterworth and Swanick, 1981; Bastin *et al.*, 1982; Dochain and Bastin, 1984; Williams *et al.*, 1984; Montgommery, 1985; Williams *et al.*, 1986; Montague, 1986; Wu *et al.*, 1987; Chamilothoris *et al.*, 1988; Andersen *et al.*, 1991; Samaan *et al.*, 1991; Zheng and Dahhou 1991) and specific bioprocess control design (Dochain, 1990; Bastin and Dochain, 1990; Dochain and Perrier, 1991; Bastin, 1991). Williams *et al.* (1984) used an adaptive algorithm to control yeast fermentation. The process is linearized around a working point in order to apply a linear technique. Different control strategies are applied and tested: dissolved oxygen tension control, respiratory quotient control, carbon-dioxide production rate control, ethanol production control and combinations of these. The results obtained show that the efficiency is maximized if a combination of above mentioned variables is minimized. Furthermore, it is necessary to estimate variables that can not be measured. Montgommery et al., 1985 considered a linear time variant model. Dochain and Bastin (1990) proposed a specific method for bioprocess control and estimation, based on an external linearization technique.

Mészáros and Báleš (1992) suggested both optimal feedforward and feedback control for a fed-batch cultivation. The solution results in an optimal substrate feed rate profile combined with recursive least squares identification and an adaptive discrete time control strategy. The aim of the control is to maximize the productivity or to improve the quality of the products.

Non-linear control of biotechnical processes is rather popular. Numerous authors have contributed with various algorithms ranging from optimization theory (Axellson, 1989; van Impe, 1993) to partial feedback control (Moubaraki *et al.*, 1991). Manchanda *et al.* (1991) and Proell *et al.* (1991) compared several control techniques, globally linearising control, adaptive control, and non-linear predictive control, to a simulated continuous fermentation process. Moubaraki *et al.* (1991) designed their controller by using partial linearization by feedback and change of coordinates (diffeomorphism). Ramseier *et al.* (1991) used an adaptive nonlinear generic model controller to regulate the cell concentration in a continuous bakers' yeast fermentation. This was first tested on a simulation model, then on an experimental process. Béteau *et al.* (1991) used the L/A control, a nonlinear approach, to an anaerobic digester. Pomerleau and Viel (1992) used an adaptive non-linear control law based on a linearizing approach to control ethanol concentration in an industrial scale bakers' yeast production.

Using an estimator for the specific growth rate and controlling this growth rate at a certain set-point or optimal profile has been extensively reported by Shimizu and coworkers, 1989; 1989a; 1989b; Shioya, 1992. However, they used an Extended Kalman filter to obtain an estimate of the specific growth rate. Zheng and Dahhou (1991) used a model reference adaptive estimator to estimate the biomass concentration and specific growth rate. With these estimates they adjusted a model reference controller. The objective of the model reference controller was to track the substrate and biomass concentrations of the plant. This approach needed substantial input data: both biomass and glucose concentration should be measured on-line, which is not a feasible situation.

In the nineties some application of expert systems, fuzzy control, intelligent control, etc. have been reported. Aynsley *et al.* (1991) described a real-time expert system that comprised a supervisory knowledge base which monitored and controlled individual fermentation units and a scheduling knowledge base which aimed to maximise plant wide productivity. Lochner *et al.* (1992) used a supervisory computer to perform data acquisition, management, evaluation, calculation and analysis as well as control and interaction with subsystems. The supervisory system uses pattern recognition to make automated decision for the control of the fermenter. Siimes, *et al.* (1992) described an object-oriented fuzzy expert system to support on-line control of an automated fermentation plant. Sterback *et al.* (1993) used an expert system to control an industrial scale bioreactor. Zheng *et al.* (1991) presented an intelligent CAD system for fermentation process control. The system is based on three layers: a real-time supervision level, a CAD level, and a learning level.

In this Chapter we propose a controller that is capable of maximizing the yield of the product and is also suitable for specific growth rate control. As the specific growth rate is not measured we use the observer of Chapter 6 as our 'soft-sensor', measuring both ethanol and biomass concentration. Furthermore we try to keep the controller as simple as possible by introducing the non-linear compensation of Chapter 4 as a part of our control strategy, hereby linearizing the control. The remaining control actions can then be tackled using a normal PI-controller. A more detailed description is given in the next Sections.

The proposed control structure is different from the above mentioned strategies. It is comparable with the structure of Axellson (1989), specially the use of non-linear compensation, but is different as the controller proposed by Axellson (1989) only regulated the ethanol concentration in the fermenter whereas the proposed control structure is intended to regulate the specific growth rate of the structure. The use of an observer to tackle the control problem is widely used (e.g. Bastin and Dochain, 1990; Ramseier, 1991). The difference with the published articles so far is the kind of observer, see Chapter 6. The control structure as proposed here has not been presented yet.

7.2 Control Structure

In this subsection we will develop a control structure such that the specific growth rate can be kept at set-point and constraints on the ethanol concentration are met as well. The basic idea of the control structure is shown schematically in Figure 7.1. The user has set a maximum ethanol concentration, E_{max}, which is not to be exceeded. Furthermore, a specific growth rate set-point (μ_{set}) or pattern of set-points is set as well. As long as the ethanol limit is not exceeded, the E-controller simply connects μ_{set} to μ_d, that is the desired specific growth rate. The μ_d is then controlled via the μ-controller. If the ethanol limit is exceeded the E-controller lowers the μ_d such that ethanol production will stop and ethanol consumption can start. Figure shows the basic idea in a schematic overview. An advantage of this structure is, that the priority relation between ethanol limit and specific growth rate set-point is reflected automatically by this set-up. Another potential advantage is that, with this structure, both μ-control and E-control can be realised. μ-control is ensured by setting the ethanol limit to an extremely large value, so that the E-control will never be activated. Setting the μ set-point to a value well above the critical value will result in violation of the ethanol limit. μ_d will be lowered until, hopefully, the ethanol limit is just maintained.

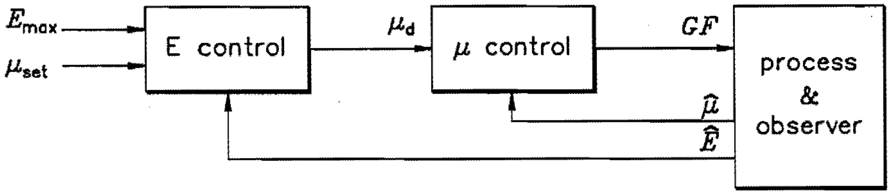


Figure 7.1 *Basic idea for the control structure*

We will now elucidate the schematic overview of Figure 7.1 in a more detailed way as can be seen in Figure 7.2. At the right hand side we see two familiar blocks; firstly the "yeast" block which represents the laboratory process (see Chapter 2) and secondly the "observer" block which is demonstrated in Chapter 6. The "yeast" block can be influenced by the glucose flow (*GF*), to adjust the specific growth rate, and the stirrer speed (*SS*), to adjust the dissolved oxygen tension.

It is a well known fact that controllers for fed-batch phase processes need increasing gain throughout the fed-batch phase process. This simply stems from the fact that the yeast needs more glucose and oxygen towards the end of the fed-batch phase because the total amount of yeast has increased.

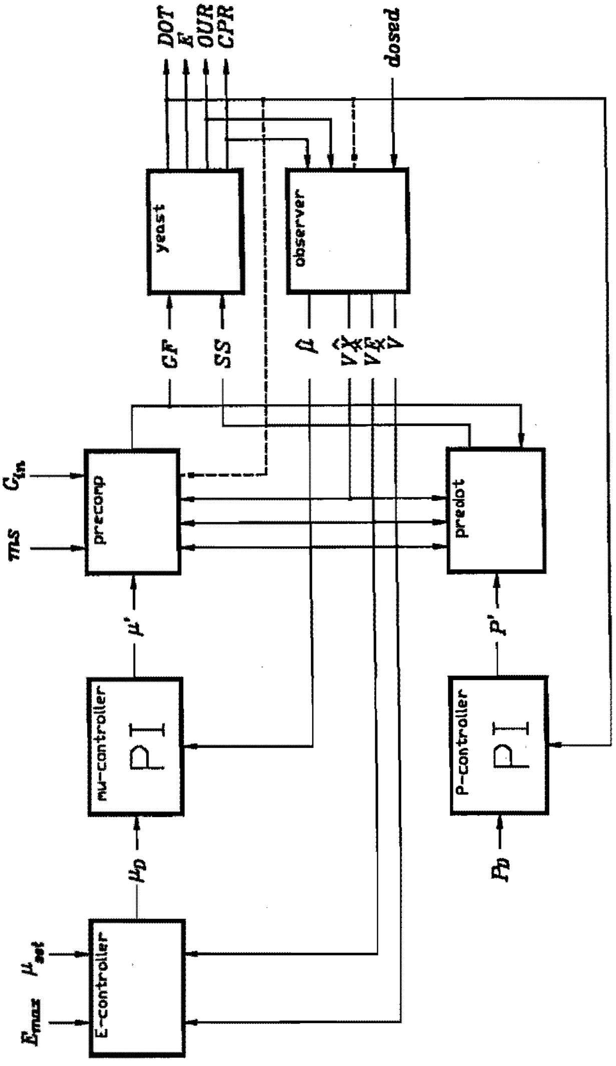

DOT
E
OUR
CPR
dosed
yeast
observer
GF
SS
μ̂
VX̂
VÊ
V
precomp
predot
C_in
ms
μ'
mu-controller
P I
P'
P-controller
P I
μ_D
P_D
E-controller
μ_set
E_max

Figure 7.2 *Control structure based on simulation results*

This non-linearity has already been presented in Chapter 4 as the non-linear compensation. In Figure 7.2 we find this block back as precomp. at the top right hand side just left to the laboratory process, "yeast". A similar block can be made for the oxygen, seen in Figure 7.2 as "predot". From Chapter 4 we know that the non-linear compensation needs the specific growth rate which in this case is provided as the manipulatory value μ'.

If the desired growth rate μ_d is unequal to the estimated growth rate a simple PI-controller will adjust the difference by manipulating the glucose feed. The controller does not need to be sophisticated as we know from Chapter 4 that the linear models, after non-linear compensation, of the transfer function between the glucose feed and the effluent gasses are of low order. The models found in can help in pre-tuning of the PO-controller. The same reasoning holds for the DOT-controller. At the left hand side of Figure 7.2 the E-controller is situated. The E-controller is a supervisory control block as indicated before in Figure 7.1.

Every block of Figure 7.2 will be elucidated, starting at the top right hand side; the *GF* compensation, the μ-controller, the E-controller,down to the bottom left hand side; the stirrer speed compensation, and the DOT-controller. All these blocks will be treated in separate subsections and the choices made will be elucidated with examples. The tests done are based on the batch and fed-batch phase simulation models described in Chapter 3. The batch phase simulation model is also used to test the controller performance because the batch phase simulation model contains more non-linearities and time-variant parameters than the fed-batch phase simulation model. The use of the batch phase simulation model under fed-batch phase conditions can be regarded as an extra, more difficult, test case for the controller performance

Glucose Flow Compensations

Inspection of the fed-batch phase simulation model (Section 3.4) shows that, under practical conditions, the specific growth rate can be considered decoupled from the oxygen concentration, if the dissolved oxygen tension is sufficiently high, say 10%. Furthermore, the yeast's need for glucose is proportional to the total amount of yeast. This is confirmed by analyzing the simulation model. Therefore *GF* is made proportional to $V\hat{X}$ obtained from the observer. The maintenance effect Eq. (3.4) shows that some additional glucose is needed. In steady state, the amount of glucose added per minute should be equal to the amount of glucose consumed per minute. If only oxidative growth on glucose occurs, the ideas mentioned so far result in the compensation:

$$GF = \left[\mu_d + \frac{ms}{\alpha}\right] \cdot \frac{V\hat{X}}{b \cdot S_{in}} \tag{7.1}$$

This compensation is based on an article of O'Connor *et al.* (1992) and on Section 4.2. It will be referred to as compensation 1. The stoichiometric constant *b* can be regarded as the yield of the fermentation process in [mole biomass / mole glucose]. S_{in} is the concentration of glucose in the feed [mole.dm^{-3}] and the glucose feed rate is expressed in [dm^3.h^{-1}].

In Table 7.1 the a-priori knowledge and assumptions for compensation 1 are described. If the assumptions made indeed hold, this scheme is stable: if the specific growth rate is less than the set-point, more glucose is added than is consumed. Glucose is accumulated, its concentration grows and the yeast will grow faster, until the set-point for μ is reached. Simulations point out, that this steady state is reached well within a sample interval. Therefore the control structure in general and the compensation in particular need not compensate for any dynamics in this adaptation process: changing the glucose feed rate has an instantaneous effect on the glucose concentration and the specific growth rate. This simplifies the control design significantly.

Table 7.1 *A-priori knowledge and the assumptions belonging to compensation 1*

- glucose concentration reaches steady state within the sampling interval
- maintenance is known a-priori
- yield on oxidative growth on glucose is known a-priori

If the first assumption of Table 7.1 does not hold, the PI-control part of the μ-control faces additional dynamics. If the maintenance rate is not exact, a constant offset on the realised μ will occur. This will be counteracted by the integral part of the PI-controller. The third assumption finally will give changing offsets on μ if it turns out to be invalid.

Compensation 1 has the drawback that it accounts only for one of three possible pathways: oxidative growth on glucose. Ethanol consumption and production are not considered. Taking into account these pathways as well gives the second compensation. Compensation can be expressed as Eq. (7.2).

$$\lambda_p = \frac{\mu_d + \frac{ms}{\alpha} + (g-b) \cdot z_0}{g}, \quad \lambda_o = \frac{\mu_d + \frac{ms}{\alpha} - \frac{la}{k} \cdot z_0}{b - \frac{la}{k}}, \quad \lambda_e = \frac{\mu_d + \frac{ms}{\alpha} - 0.13 \cdot l \cdot z_1 \cdot z_2}{b} \tag{7.2}$$

$$GF = \max\{\lambda_p,\ \lambda_o,\ \lambda_e\} \cdot \frac{V\hat{X}}{S_{in}}$$

z_0, z_1 and z_2 are defined as

$$z_0 \doteq \frac{0.2}{a} \cdot \frac{DOT}{K_o + DOT}, \quad z_1 \doteq \frac{E}{K_e + E}, \quad z_2 \doteq \frac{K_i}{K_i + S}$$

z_1 should be reconstructed using observed values for E. z_2 should be approximated by a constant close to 1 as normally G is much smaller than K_i. It is also the only possibility, because the value of G can not be measured or reliably observed. Using this simplification, all data required for this compensation are available on-line. In Eq. (7.2), λ_p represents ethanol production. If λ_o is chosen, ethanol is consumed and the oxygen bottleneck is completely filled. This will be called *oxygen limited ethanol consumption.* If λ_e is active, ethanol is consumed and the oxygen bottleneck is not completely filled, the ethanol consumption is limited by the ethanol concentration. This will be called *ethanol limited ethanol consumption.* The situation of no ethanol production or consumption is considered a special case of *ethanol limited ethanol consumption.*

Theoretically the incorporation of all pathways in the compensation may seem a small step, however the amount of required a-priori knowledge has increased substantially. First, the critical specific growth rate, μ_{crit}, i.e. the μ at the switching point between ethanol consumption and ethanol production, needs to be known. Secondly, the ethanol consumption rate as a function of the glucose and ethanol concentrations is used. This rate appears in Eq. (7.2) as $0.13 \cdot l \cdot z_1 \cdot z_2$. Finally, the yield of all pathways is needed. For ethanol consumption and production, these are l [mole biomass / mole ethanol] and g [mole biomass / mole glucose]. Table 7.2 summarizes the assumptions for the second compensation.

Table 7.2 *A-priori knowledge and the assumptions belonging to compensation 2*

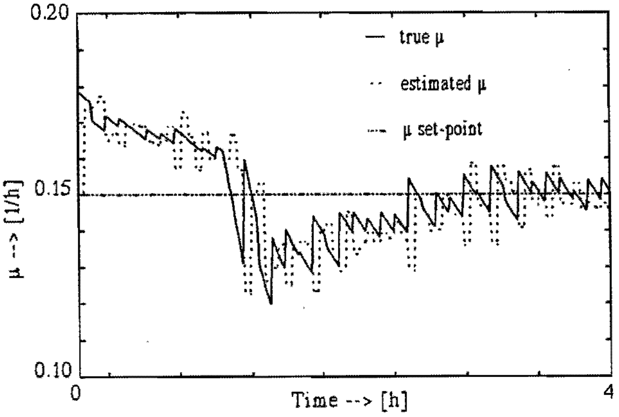
same knowledge as in *compensation 1* plus

- yields of all pathways are known a-priori
- ethanol consumption rate as function of E and S is known a-priori
- critical growth rate μ_{crit} is known a-priori

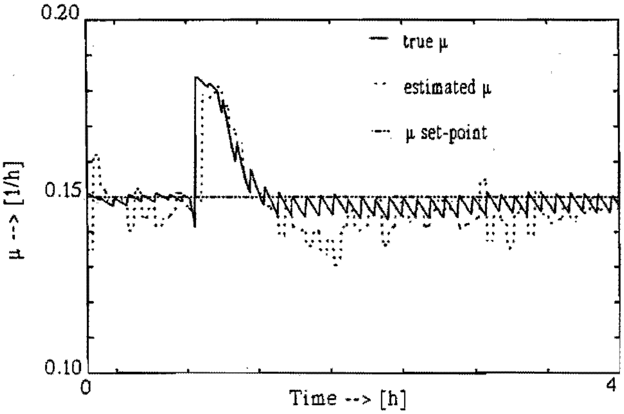
The performance of both compensations was tested by incorporating them in the control structure of Figure 7.1. They are discussed in Example 7.1. The E-controller in Figure 7.1 need not be considered yet, the μ controller is a simple PI-controller. Both controllers will be discussed later on in this chapter.

Example 7.1 *Tests of compensation 1 and 2*

During the first 1.5 hours of this simulation, ethanol was consumed. For the rest of the simulated experiment, ethanol was neither consumed nor produced. The tuning of the PI-controller is the same for compensation 1 and 2. The results are shown in Figure 7.3.



(a)



(b)

Figure 7.3 *Estimated and simulated specific growth rate for constant set-point using compensation 1, part (a), and compensation 2, part (b)*

The response using compensation 1 lies above the set-point for 1.5 [h]. During this span of time, the yeast grows both on glucose and on ethanol. The PI-controller is compensating for the growth on ethanol, pulling the specific growth rate down. The specific growth rate does not manage to reach the set-point within 1.5 [h], though. As soon as all ethanol has been consumed completely, μ falls below its set-point. The compensation achieved so far for growth on ethanol must be undone again. This is realized in the next 1.5 [h]. At three hours from the simulation start, the specific growth rate is tightly on set-point.

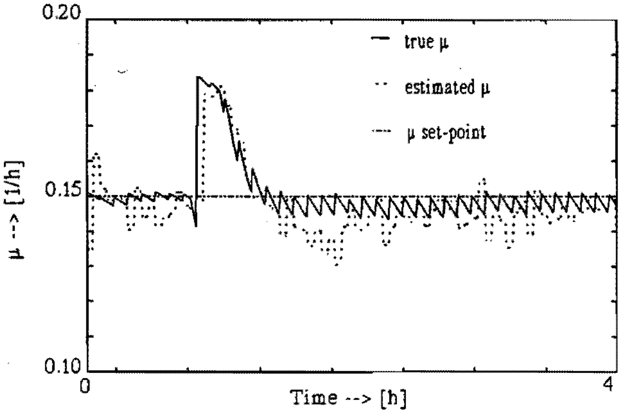
During the first 1.5 [h] of the simulation with compensation 2, the specific growth rate is on set-point. The compensation detects growth on ethanol and compensates correctly for this. The ethanol consumption rate is limited by the oxygen concentration and therefore independent of the exact ethanol concentration. As soon as the ethanol concentration is low enough to have ethanol limited ethanol consumption, control becomes worse. This is caused by the noisy ethanol estimates. These influence the glucose feed rate during the rest of the simulation. The fluctuations in the ethanol estimates cause fluctuations in the glucose feed rate. The specific growth rate is disturbed by these fluctuations. ■

To improve the performance of compensation 2, the sensitivity to errors on the ethanol estimates during ethanol limited ethanol consumption needs to be addressed. The third compensation uses a variant of Eq. (7.2). $\lambda_e(E=\hat{E})$ is used to determine which of λ_p, λ_o or λ_e should be active. In case λ_e is active, $\lambda_e(E=0)$ is used for the calculation of *GF*. This will result in a peak on μ if the compensation switches from oxygen limitation to ethanol limitation. This peak will die out because of the consumption of ethanol, which eventually will result in $E=0$ indeed. The PI-controller contributes to the removing of this peak. The knowledge used by compensation 3 is the same as that for compensation 2. Compensation
]

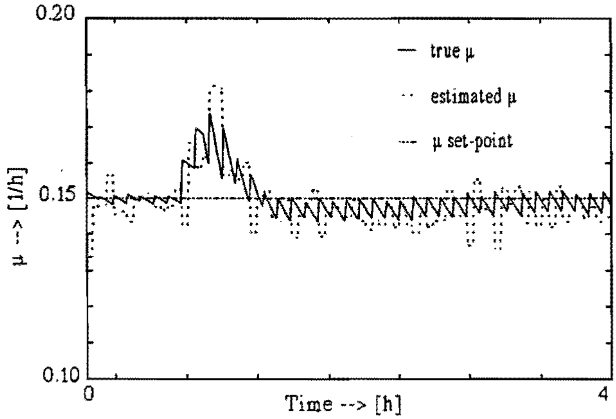
3 still needs the ethanol consumption rate as a function of E and S, but only to choose between ethanol limitation and oxygen limitation, not to dose the glucose feed.

The last compensation, compensation 4, tries to eliminate the need for accurate ethanol estimates by using predictions of the ethanol concentration instead of observer values under ethanol limitation. (Part of) the simulation model is used to generate these predictions. It is acknowledged, that this might not work in practical situations. From Chapter 6 we know that the prediction of the ethanol is not very reliable, see Example 6.1. It seemed nevertheless worthwhile to investigate the possibilities of this approach. The knowledge used by compensation 4 is again the same as that used by compensation 2. This algorithm relies heavily on the ethanol consumption rate as a function of E and S.

Example 7.2 *Tests of compensation 3 and 4*



(a)



(b)

Figure 7.4 *Estimated and simulated specific growth rate for constant set-point*

using compensation 3, part (a), and compensation 4, part (b)

The peak on the specific growth rate at the switch from oxygen limitation to ethanol limitation, which was expected for compensation 3, is indeed observed. As soon as all ethanol has been consumed, the specific growth rate is steady at its set-point. Compensation 4 achieves little gain in performance. The peak at μ is smaller than for compensation 3 but still rather obvious. Its dependence on the validity of the simulation model has however increased significantly. ■

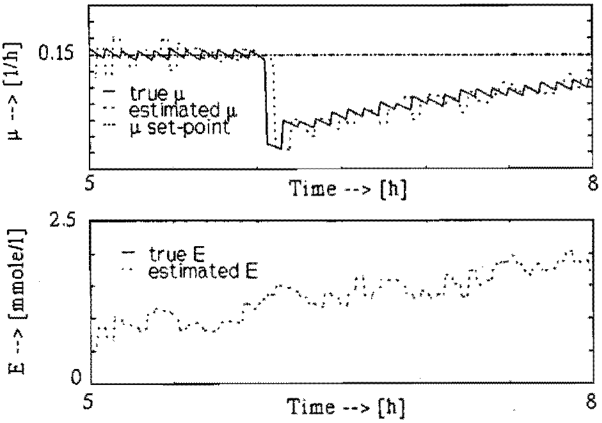
For the discussed simulations the μ set-point was noticeably below its critical value. Consequently the dip in the specific growth rate after all the ethanol has been consumed is considerable. There is still room left in the oxygen bottleneck for growth on ethanol. The presence or absence of ethanol in the broth can make a large difference in the specific growth rate. If the oxygen bottleneck is filled more completely, the switch from oxygen limitation to ethanol limitation produces a smoother response of the specific growth rate.

The control performance appears to be limited by the accuracy of the ethanol estimates. Improving the accuracy of these estimates would improve control. Therefore, some extra knowledge was implemented in the observer. If the μ estimate is well below its critical value, ethanol production can not occur. If μ is significantly above its critical value, ethanol consumption is prohibited. By keeping the margins used around the critical value large enough, the use of this knowledge may stay within justifiable bounds.

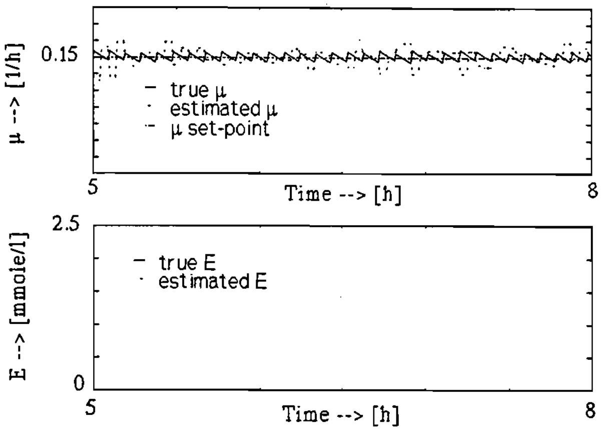
Another way of dealing with a-priori knowledge is to incorporate it in the observer. The next example will discuss this.

Example 7.3 *Observer with exclusion of ethanol production for $\mu < 0.21$ [h^{-1}]*

Excluding ethanol production for μ below 0.21 [h^{-1}] leads to a significant improvement of control performance. In the figures below, performance the observer with the ethanol exclusion is compared with performance obtained by the observer used so far. It is evident, that control has improved.



(a)



(b)

Figure 7.5 *Original observer, part (a), and modified observer, part (b), controller response*

The sudden fall of μ in Figure 7.5, part (a), is caused by an incorrect recognition of the yeast state (ethanol production, oxygen limitation or ethanol limitation). This in turn is caused by the drifting ethanol concentration estimate.

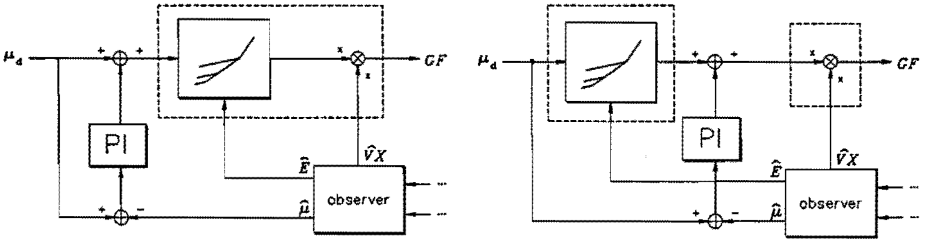
Another possible solution to the problems regarding the accuracy of the ethanol concentration estimates is using an ethanol measurement. Measurements of the ethanol concentration are available in the off-gas. Assuming certain exchange processes are in steady state, the ethanol concentration in the broth can be derived from these data. The noise level is about 100 [ppm] for these measurements. This exceeds by far the error in the observer estimates. Therefore the measurements currently available offer no solution to the sensitivity to errors in the ethanol estimates. ■

μ-controller

The glucose flow compensations are essentially feed-forward controllers of the specific growth rate. They have the usual disadvantages associated with feed-forward control:

- discrepancies between the assumptions implemented in the compensation and the true situation are not corrected or even detected
- external disturbances are not suppressed

For this reason, the compensation should be complemented by some means of feed-back control. A simple PI-controller was chosen for this task. Two conceivable PI-structures are shown in Figure 7.6. The parts within dotted lines are the glucose flow compensation. The essential difference between the first and the second variant is, that the first variant influences the state recognition of the compensation, the second variant does not. It depends on the kind of error that needs to be corrected which of these two would be appropriate. If for example the value of the critical growth rate is in error, the first structure can correct this. If the maintenance rate is wrong, the second scheme would eliminate this error. A combination of both schemes is also possible: the PI-output could be added at both places with different gains. Making either one of the gains equal to zero reduces this structure to one of the variants mentioned above.



(a) (b)

Figure 7.6 *Two variants for a PI-controller structure*

From the simulation environment it can not be predicted, which of both situations will occur in practice. Therefore quite arbitrarily the second structure is chosen. All simulations in this chapter using a μ-PI use the variant of Figure 7.6, part (b). As long as the compensation does not switch between different pathways, both solutions are interchangeable.

Obviously the glucose feed rate is limited to positive values. There is also an upper bound to it, caused by the maximum flow capacity of the glucose pump. If the upper bound is reached, the integrator value in the PI-controller is not allowed to increase any further. Conversely, if the lower bound is reached, the integrator value may not decrease. This measure improves the practical robustness of the design, but does not affect normal performance.

Tuning of the PI poses a theoretical problem. Apart from the delays in the measuring equipment, the nominal transfer from the input of the compensation to the μ estimate can be modelled by a zero order model. Due to numerical errors made in the observer this could result in a higher order transfer function accounting for these numerical errors. To investigate this possibility a PRBS simulation was designed. The result are given in Example 7.4.

Example 7.4 *The influence of the observer based numerical errors on the transfer function of the input of the compensation to the μ estimate*

The influence of observer based numerical errors was investigated by a PRBS excitation of the *GF*. The input signal of the compensation alternated between an ethanol producing specific growth rate and an ethanol consuming growth rate. The alternation between ethanol production and consumption was designed such that during the simulation experiment the ethanol concentration would be at a low, constant level. Keeping the ethanol concentration at a constant level can not be completely controlled, as the ethanol consumption rate is a function of the ethanol concentration, which is not controlled directly in this set-up. Nevertheless the ethanol concentration did not drift away during the simulation. First, second and third order ARMA models of the transfer function of the compensation to the μ estimate were fitted using a least squares equation error criterion. This gave the pole-zero plots of Figure 7.7. The pole-zero cancellation for all model orders is clearly visible, so the nominal transfer of the input of the compensation to the μ estimate is of order zero as was assumed.

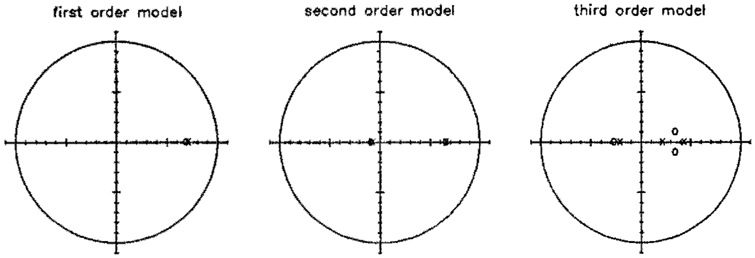


Figure 7.7 *Pole-zero plots for μ_d-$\hat{\mu}$ transfer* ■

An optimal tuning for the delay could be developed, but in practice, the delay in the measuring equipment will not be constant. Presumably it will vary between one and two samples. Tuning the PI based on the noise characteristics offers no solution either. Another possibility would be to model the structure of the modelling errors to be expected. This approach is doubtful, because one can not expect the simulation model to give a reliable indication of the structure of the dynamics it does not model correctly, because the structure of the model is based on sheer data fitting. It is not supported by physical or chemical reasoning. Even if this were allowable, it would be far from straightforward to do. This leaves us little choice than to base the tuning on trial and error. The above mentioned observations imply also, that control design techniques such as pole placement and H_∞ can not be used either.

The tuning parameters found for the simulation model are listed in Table 7.3.

Table 7.3 *PI-controller settings*

μ-controller	K_p	0.2
	K_i	0.1

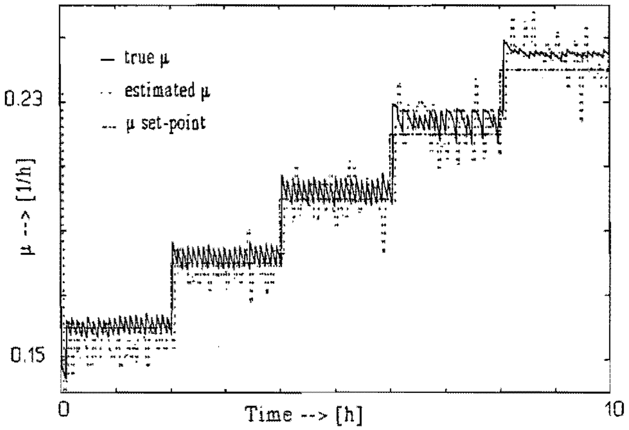
On the actual yeast process, these parameters can be used as a starting point for tuning of the actual process. Then, another balance has to be found between responsiveness on the one hand and robustness and stability on the other hand.

The current structure is able to maintain a constant set-point on the specific growth rate. This was confirmed by simulations (not shown). Further simulations were carried out to assess the performance of the controller for a stair-case profile on the set-point and a ramp signal on the set-point. The results are given in Example 7.5.

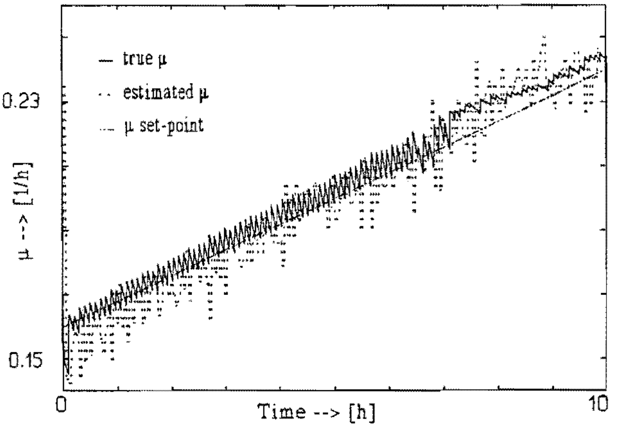
Example 7.5 *Stair-case and ramp profiles for the specific growth rate*

The staircase signal starts at $\mu = 0.16$ [h^{-1}], a growth rate that ensures oxidative growth on glucose. The steps are positive and 0.02 [h^{-1}] big and go up until $\mu = 0.24$ [h^{-1}], a growth rate at which reductive growth on glucose is present. Every stair lasts about 2 [h] and the sampling time is 6 minutes. The ramp signal also starts at μ = 0.16 [h^{-1}] and ends at μ = 0.24 [h^{-1}]. The expected behaviour is that until the critical growth rate, μ = 0.23 [h^{-1}], is reached the controller should be able to control the specific growth rate at the set-point given by the staircase and ramp signals.

The set-point changes of the staircase profile are rapidly followed (see Figure 7.8). Between 6 and 8 hours, the true μ lies somewhat above the set-point. For μ = 0.22 [h^{-1}], very close to the critical value, the observer has difficulties reconstructing the yeast state from the noisy measurements. Because the controller uses the observer values, the bad μ estimates influence the overall control performance as well. This is not due to problems with the controller but due to observer sensitivity.



(a)



(b)

Figure 7.8 *Performance of the μ-controller to a staircase signal, part (a) and a ramp signal, part (b)*

The error that is due to difference between the true μ and its estimated value is stored in the integrator of the PI and passed to the next stair, with μ = 0.24 [h^{-1}], where the offset on the true estimate is clearly seen. The observer has little trouble recognising the yeast state now and the performance of the controller is better.

The response to a ramp signal (Figure 7.8 (b)) on the μ set-point corresponds to the response to the staircase signal. For set-points near the critical growth rate the true growth rate is higher than the observed value. This effect is passed on to the two hours following this situation. The current control structure is capable of following a ramp that passes through a wide range of values: it starts off with a value far below the critical value and ends with a value significantly above it.

It is not advisable to reproduce this simulation in a practical setting, because the ethanol production rate towards the end of the simulation is considerable. The ethanol concentration reaches values of the order 100 [$mmole.dm^{-3}$] (approximately 2000 [ppm]).

The saw-tooth shaped pattern superimposed on the true specific growth rate in each simulation is a result of yeast growth during a sample interval. The glucose feed rate is calculated at the start of each sample interval from the amount of yeast that is present at that moment. During the sample interval the glucose feed rate should actually increase, because the amount of yeast increases but the glucose feed rate stays constant over a sample period. The sawtooth patterns will not be visible in experimental data. This is due to the slightly higher sampling frequency in practice (twice as high). Furthermore, the head space of the fermenter acts as an integrator such that small but sharp fluctuations in the exhaust gas are filtered. ■

We have now determined all the parameters of the specific growth rate part of the controller and will now look at the ethanol controller.

E-controller

The ethanol controller is added to the structure mainly as a safeguard against excessive ethanol concentration. High ethanol concentrations obstruct yeast growth and can eventually destroy the yeast cells. High ethanol concentrations also have a negative influence on the quality of the yeast. If the controller manages to maintain the ethanol concentration on its set-point (E_{max}) without significant overshoot to either side, the current control structure can be used for both ethanol control and μ control, as already mentioned at the beginning of this chapter.

The algorithm used by the controller is

$$\mu_d = \mu_{set} - K_p \cdot (\hat{E} - E_{max})_+ - K_i \cdot i(\hat{E}) \qquad (7.3)$$

with

$$(x)_+ = \begin{cases} x & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \qquad (7.4)$$

and

$$i(\hat{E},k) = \begin{cases} i(k-1) + (\hat{E}-E_{\max}) & \text{if } \hat{E} > E_{\max} \\ p \cdot i(k-1) & \text{if } \hat{E} \leq E_{\max} \end{cases} \qquad 0 \leq p < 1 \tag{7.5}$$

If the ethanol limit is exceeded, Eq. (7.3) acts as an ordinary PI-controller, feeding back both the error and the sum of errors. As soon as the ethanol limit violation has been eliminated, the intervention on μ is gradually turned off, Eq. (7.4). The μ set-point converges back to its original value according to a first order step-response with a pole in p, Eq (7.5).

In order to preserve smooth transitions and account for set-point changes of μ_{set} during ethanol control some precautions are taken. If a μ set-point change occurs when the ethanol limit is exceeded, the output of the ethanol controller is not allowed to increase. Keeping the ethanol on set-point has a higher value than controlling the specific growth rate. The value of *i* is changed so that the set-point change is absorbed completely. Downward changes of the controller output as a result of a μ set-point change are not filtered out. The responsiveness of the controller is further improved by proper initialisation of *i* at the beginning of a limit violation. *i* is set to a value such, that the output of the controller is immediately at a value near the critical growth rate, unless this would increase μ_d. Just like there are physical limitations on the range of values of *GF*, μ_d is limited too. It is not allowed to take on negative values. If μ_d would become negative, the value of *i(k)* is frozen, so that the integrator in the PI does not drift away. This is similar to the freezing of the integrator in the PI of the μ-controller.

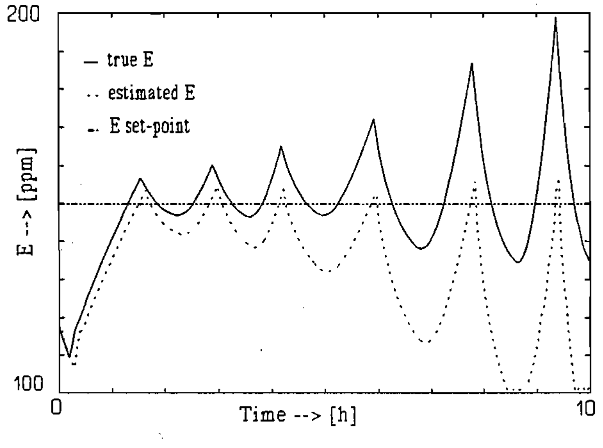
The asymmetric nature of the PI-part of the controller makes conventional PI-tuning criteria impractical. If K_p and K_i are chosen too small, it will take the controller too long to reach μ values for which ethanol is consumed rapidly enough. If they are chosen too big, μ values take on too small values, so that the ethanol intervention stays on too long after the violation has been eliminated. The same holds for the value of p. If it is taken too close to one, the decrease of μ_d does not die out rapidly enough, if it is too small, μ_d overshoots the critical value too far, so that instant production of ethanol occurs and a new intervention can be expected within a few sample instants.

Another problem the E controller faces, is that the ethanol production rate is proportional to the biomass. During an experiment the biomass increases by more than a factor 10, making control more difficult as the experiment progresses. Due to the growth of the biomass the time constants of the process change and the overall system response to changes in the glucose flow change; the response becomes faster. The effect on the E-controller is that the control will not be as tight any more. Nevertheless, control should still be possible.

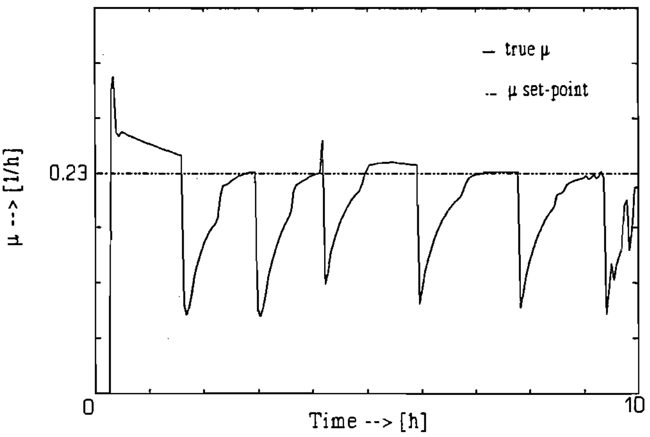
Next we determine the parameters of the E-controller. This is done in Example 7.6.

Example 7.6 *Tuning of the E-controller*

A simulation with a μ set-point of 0.23 [h^{-1}] and an ethanol limit of 100 [ppm] is presented in Figure 7.9. No noise was added to the measurements to make the essential issues stand out more clearly.E control results inherently in μ set-points close to the critical value. It is for these set-points, that the observer has most difficulties in recognising the pathways currently followed. This degrades the ethanol concentration estimates in Figure 7.9, part (a).



(a)



(b)

Figure 7.9 *E-controller performance for the critical growth rate; the ethanol response, part (a), and the specific growth rate response, part (b)*

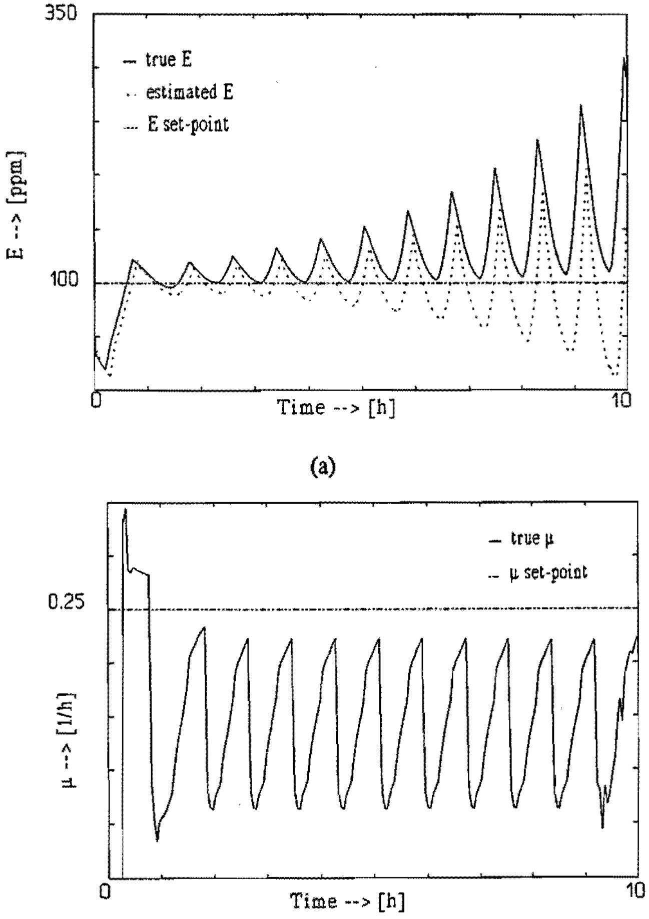
The increasing overshoot in the ethanol concentration is caused by the increasing biomass concentration. The ethanol production rate is proportional to this concentration. In Figure 7.9, part (b) shows, that the pattern on the specific growth rate remains fairly constant. ■

Things get even worse if a larger μ set-point is chosen as can be seen in Example 7.7.

Table 7.4 *PI-controller settings ethanol part*

	K_p	12
Ethanol-controller	K_i	4
	p	0.9

Example 7.7 *Tuning high specific growth rates and ethanol concentrations*



(a)

(b)

Figure 7.10 *E-controller performance for high growth rate; the ethanol response, part (a), and the specific growth rate response, part (b)*

> The larger μ set-point causes the μ-controller input to return more rapid to growth rates causing ethanol production after an ethanol controller intervention. The ethanol limit is now exceeded more quickly. The oscillations in Figure 7.10 clearly have a higher frequency than the one in Figure 7.9. The violation of the ethanol limit is also more severe. ■

From these few simulations it is already clear, that the current structure is not suited for strict ethanol control. It could be used as a mere safeguard if the ethanol estimates could be improved, but realising true ethanol control by setting the μ set-point to large values is certainly beyond the capabilities of this controller. The specific growth rate is not an adequate input to control the ethanol concentration for the following reasons:

- The specific growth rate is not the only factor determining the ethanol consumption rate. The ethanol concentration itself is an important factor too, as is the biomass concentration.
- The performance of the μ-controller is not such, that the true specific growth rate is always instantly equal to the set-point. Running the ethanol control at a lower sample frequency than the μ control would ensure that the specific growth rate would be at set-point. The lower sampling frequency could result in an ethanol concentration violation going undetected for longer than necessary. The slopes in Figure 7.9 indicate, that this is not a useful option.
- The critical growth rate is not known exactly and may even vary somewhat over different experiments due to slight changes in the biomass composition (see Section 3.2). This is certainly the most important shortcoming of this set-up.

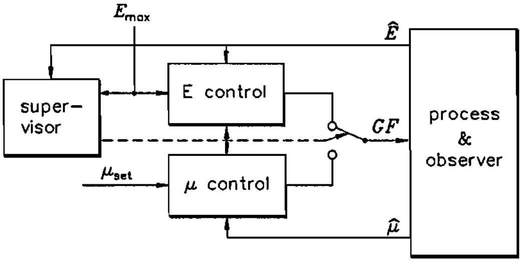


Figure 7.11 *Alternative to Figure 7.2*

An alternative to the structure of Figure 7.2 could be the one of Figure 7.11. In this set-up, the E-controller no longer uses the specific growth rate as a process input, but acts directly upon the glucose feed rate. The approach of Figure 7.11 may look promising at first, but it has some of the disadvantages from which the set-up of Figure 7.2 suffered as well. In the Figure 7.11 set-up, the critical glucose feed rate is not known. This gives the same problems in tuning the E-controller. Also, the problem of higher ethanol production and consumption rates when the yeast grows occurs in the same way in the structure of Figure 7.11 as in the structure of 7.2.

Now no distinct solution can be offered for the ethanol control problem. If the controller is seen as just a safeguard, the controller developed here meets its targets. If ethanol control is desired, one might look into the structure of Figure 7.11. For the block E-control one of the ethanol controllers described in the literature could be used. See for example (Axellson, 1989).

Stirrer Speed Compensation

The principle of the stirrer speed compensation is the same as that of the glucose flow compensation. The yeast is considered to be in one of a limited number of conditions. The offset correction of the compensation is changed according to the condition that is detected. Based on the simulation model, only two conditions are distinguished. These are *maximum oxygen consumption* and *sub-maximum oxygen consumption*. Maximum oxygen consumption occurs during ethanol production and oxygen limited ethanol consumption. Sub-maximum oxygen consumption takes place during ethanol limited ethanol consumption.

Under maximum oxygen consumption the oxygen consumption rate in [mole $O_2.h^{-1}$] is equal to

$$-d\frac{DOT_{1a}}{dt} = \left[0.2 \cdot \frac{DOT}{K_o + DOT} + ms\right] \cdot X \tag{7.6}$$

This is, apart from the biomass-dependence, constant.

Under sub-maximum oxygen consumption, the oxygen consumption rate is equal to

$$-d\frac{DOT_{1b}}{dt} = \left[0.5 \cdot a \cdot \frac{S}{K_s + S} + 0.13 \cdot z_1 \cdot z_2 + ms\right] \cdot X \tag{7.7}$$

The glucose consumption rate is taken to be in steady state. As the glucose concentration in the feed exceeds by far the concentration in the broth, in good approximation it holds, that

$$\frac{GF}{VX} \cdot S_{in} = 0.5 \cdot \frac{S}{K_s + S} \tag{7.8}$$

z_1 and z_2 should be reconstructed like it was done for the glucose flow compensations. Contrary to Eq. (7.6), Eq. (7.7) is far from constant. It depends, amongst others, on the current glucose feed rate and the ethanol consumption rate.

The oxygen that is mixed into the broth per unit time, is given by

$$\frac{dDOT_2}{dt} = C_{DOT} \cdot SS^{C_{ss}} \cdot AF^{C_{af}} \cdot \left(DOT_{max} - DOT\right) \tag{7.9}$$

This should compensate Eq. (7.6) or Eq. (7.7), whichever is appropriate, and also drive the oxygen concentration in the desired direction. If the desired net time derivative of the oxygen concentration is denoted $dDOT_s/dt$, the overall compensation algorithm is

$$SS = \left[\frac{\min\left\{d\frac{DOT_{1a}}{dt}, d\frac{DOT_{1b}}{dt}\right\} + d\frac{DOT_2}{dt}}{C_{DOT} \cdot AF^{C_{af}} \cdot (DOT_{max} - DOT)}\right]^{1/C_{ss}} \tag{7.10}$$

The state detection algorithm simply chooses the smallest one from Eq. (7.6) and Eq. (7.7). The dilution effect has been neglected.

The number of assumptions and the required a priori knowledge is considerable:

- The glucose concentration is in steady state.
- Current glucose feed rate
- Current amount of biomass
- Ethanol consumption rate as a function of E and S
- Size of the oxygen bottleneck

Further, measurements or estimates are required for

- yeast concentration
- ethanol concentration
- oxygen concentration

These are available, either from the observer or from direct measurements.

The oxygen control runs at twice the sampling rate of the μ-control and the observer. At sample instants for which no new observer estimate is available, a simple zero order hold on the estimates was tested. Also linear interpolation of the estimates was tried.

Example 7.8 *Simulation results of the stirrer speed compensation*

Simulations point out that, again, the ethanol concentration is the bottleneck in the control performance. This is needed to reconstruct z_l. Either the ethanol estimate from the observer is used, or the ethanol concentration predicted by compensation 4. Together with the choice between interpolating linearly between two samples or using no interpolation at all, this gives four options for the ethanol concentration estimates / predictions.

This linear interpolation was not significantly better than the non-interpolated case. It was decided not to use the interpolation. As the ethanol predictions were, (on the simulation model!) more accurate than the ethanol estimates, best results were obtained when using these predictions without interpolation. They will be presented in the next section. ■

DOT Controller Algorithm

To compensate for errors in the stirrer speed compensation and to suppress external disturbances, a PI-controller is put in series with the compensation. This PI-controller compares the measured *DOT* with the set-point. As the compensation requires the time-derivative of the *DOT* instead of the *DOT* itself, the PI-controller is followed by a numerical differentiator using Euler backward. If the *DOT* is above its set-point, more energy is spent on stirring the mixture and blowing air through it than is needed. If it is below its set-point, ethanol production is likely. This means glucose is used less efficiently, as more glucose is needed to grow yeast on ethanol formed during earlier fermentative glucose consumption than to grow yeast on glucose directly. It also disturbs the specific growth rate controller, as this assumes, that the *DOT* does not influence the specific growth rate. This is only the case if the *DOT* is not too low. The PI controller takes above mentioned observations into account by an asymmetric weighing of deviations from the set-point: deviations towards no oxygen are weighed more heavily than deviations towards saturated oxygen tension.

The tuning of the PI-controller for dissolved oxygen tension is based on simulation results. As it was the case for the μ- and E-controllers, insufficient data are available to tune the PI theoretically. Further manual tuning of the controller on the real process is unavoidable.

Example 7.9 *Simulation results of the DOT-controller*

A typical controller response was obtained in a simulation were all paths were followed. The achieved oxygen concentration is plotted in Figure 7.12. The peak on the oxygen concentration after 0.4 [h] is a result from the transition from oxygen limited ethanol consumption to ethanol limited ethanol consumption. This is only detected by the controller after it occurs. The time constants associated to the oxygen concentration are so small, that if such a switch occurs during a sample interval, it will be noticeable at the end of the interval. The mentioned peak is over-compensated, resulting in a dip during the next sample instant. This oscillatory behaviour dies out within a few samples. Then the oxygen concentration is restored to its set-point.

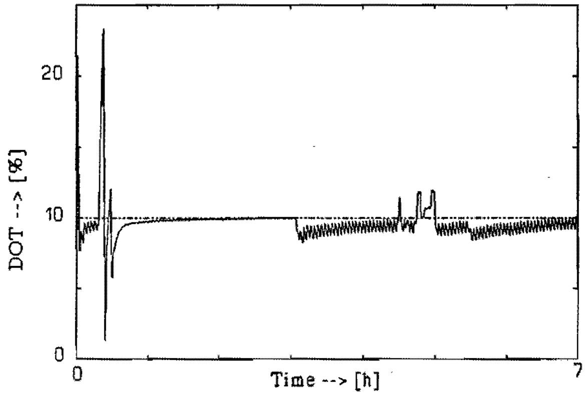


Figure 7.12 *Oxygen controller performance*

After about three hours μ is set to a value for which limited ethanol production occurs. Apparently the need for oxygen is more dependent on the amount of biomass in this case as it was during ethanol consumption, as the saw-tooth pattern observed on the specific growth rate in Example 7.5 now appears on the oxygen concentration as well. Apart from this effect, the oxygen controller has no difficulties with ethanol production. This also applies to the situation in which significant ethanol production occurs. This is demonstrated in the simulation between five and seven hours since the simulation start. ■

Good results have been attained by Lee *et al.* (1991) controlling the *DOT* using an adaptive controller manipulating both the *AF* and the *SS*. However, the proposed algorithm is rather complex and is too extensive to be considered here as a simple way to keep *DOT* on setpoint. Some less successful approaches are given in: Yano *et al.*,1979; Yano *et al.*,1981; Ko *et al.*, 1982; Chen *et al.*, 1985; Clark *et al.*, 1985.

7.3 Results

In Section 7.2 a control structure was designed and tuned based on simulation results. The control goals were to maintain a constant set-point on μ and then to track changing set-points. Furthermore, when these goals were realised, provisions were added to obey a limit on the ethanol concentration and to keep the dissolved oxygen tension on set-point. Now the obtained controller will be tested on experimental data. The fed-batch phase is entered in a known and reproducible way as described in Chapter 2. The initial conditions for the fed-batch phase are summarised in Table 7.5.

Table 7.5 *Initial fed-batch conditions*

Quantity	value	unit
Biomass concentration	10.5	[g.dm^{-3}]
Volume	4.45	[dm^3]
Ethanol concentration	0	[ppm]
Glucose concentration	0	[g.dm^{-3}]
Dissolved oxygen tension	>90	[%]
pH	5.0	[-]
Temperature	30	[^{0}C]

During the experiments the control of the dissolved oxygen concentration in the broth is realised by a primary controller. The proposed DOT-controller of Section 7.2 will not be tested in the following experiments, because the emphasis of the control design is on controlling the specific growth rate.

We will start with a check on some of the a-priori knowledge of the compensations. One of the crucial points is the critical growth rate, the switching point between ethanol production and consumption. The determination is quite simple as this is the maximum oxidative growth rate on glucose minus the (known) maintenance effect. A simple experiment with a specific growth rate set-point higher than the assumed critical growth rate is sufficient, see Example 7.10. Together with the critical specific growth rate the assumptions on the steady state of the glucose concentration and the maintenance are checked.

Example 7.10 *Check on a-priori knowledge*

The critical growth rate during the simulations was approximately 0.23 [h^{-1}]. It seemed safe to choose a μ set-point of 0.25 [h^{-1}] when trying to generate some ethanol production, not too much as this could possibly inhibit the oxidative growth on glucose, see Example 6.1. Experimental results confirmed our settings. A little bit of ethanol was produced during the experiment. The set-point was kept on 0.25 [h^{-1}] for about 4 [h], see Figure 7.13. The critical specific growth rate was identified to be approximately 0.23 [h^{-1}].

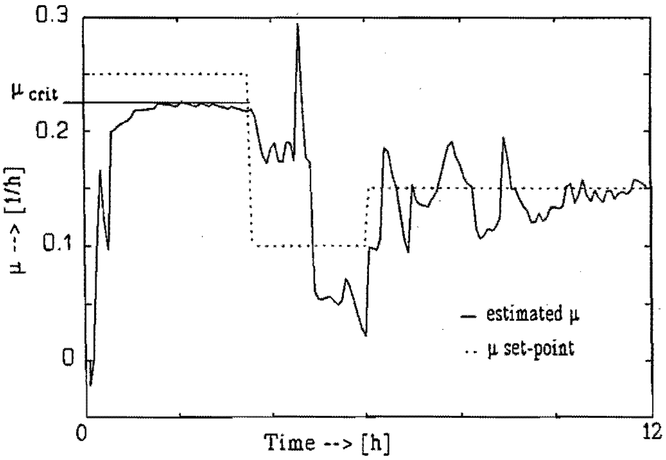


Figure 7.13 *Estimating the critical growth rate μ_{crit}* ■

As mentioned before the proposed controller for dissolved oxygen was not implemented, as a primary controller was already available. The settings of this primary controller should be checked for influence on the process and controller due to changes in the air flow and stirrer speed. This is done in Example 7.11.

Example 7.11 *Air flow and stirrer speed influence*

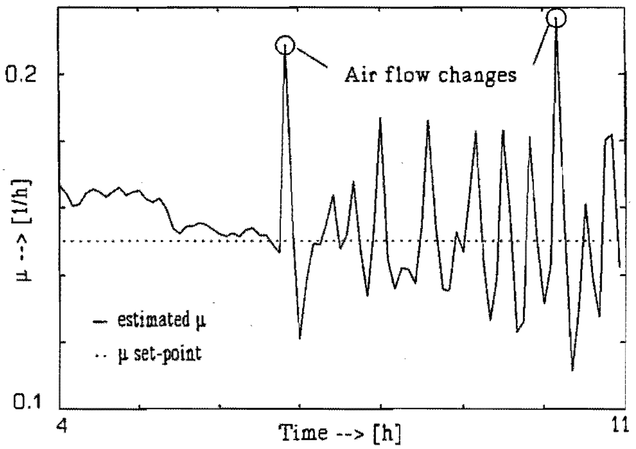


Figure 7.14 *Air flow and stirrer speed influence on μ estimates*

For high air flow rates, the noise level on *OUR* and *CPR* and consequently on μ are higher. The setting of the stirrer speed does not influence the noise on the off-gas measurements. The changes in the air flow rate itself cause a large peak on the μ estimate. Both air flow changes in Figure 7.14 were due to the increase of the *AF*. All other operating parameters were constant. The system starts oscillating after the first air flow raise. The behaviour of the μ estimates coincides with the stirrer speed behaviour.

This means that as the amount of glucose is raised, the amount of dissolved oxygen is decreased causing the controller to raise the stirrer speed in order to keep the dissolved oxygen on set-point. Next, due to changes in the equilibrium between the liquid and gas-phase, extra carbon-dioxide is going from liquid to gas-phase, causing an peak on the *CPR*, thus on the specific growth rate. This then causes the controller to act on the glucose flow, lower it, and the opposite will happen, causing these oscillations.

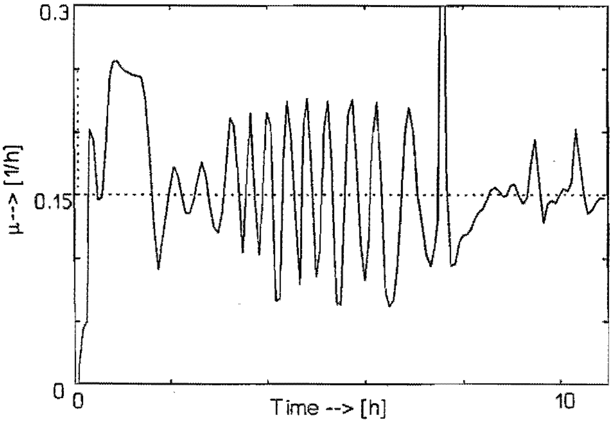
Air flow and stirrer speed changes disturb the *OUR* and *CPR* measurements. The higher air flow apparently amplifies the disturbances introduced by the stirrer speed influence. This amplification can be seen as follows. If there is a change in the stirrer speed the equilibrium between liquid and gas in the head space of the fermenter changes causing a disturbance. This disturbance is amplified if at the dame time (or shortly before or after) an increase in the *AF* occurs, forcing more oxygen and carbon-dioxide to evaporate from the liquid into the gas in the head space of the fermenter. As explained before, these disturbances are not due to physiological changes but due to physical changes (Section 3.2). To verify this hypothesis, the oxygen controller was switched off during the next experiments. Stirrer speed and air flow were set manually to values high enough to keep the dissolved oxygen tension above 20%. Again changes in the air flow and stirrer speed were made, causing the observer to "peak" at the time of the changes. This confirmed the hypothesis. ■

Slowly changing the air flow and stirrer speed reduces the oscillatory effect on the μ estimates to acceptable levels. Some rate limiting should therefore be implemented in the oxygen controller. As the oxygen controller was not our main target a simple solution was found: the oxygen control was in manual mode for the rest of the experiments. This means that the user will provide the fermenter with air flow and stirrer speed settings such that no oxygen limitation occurs.

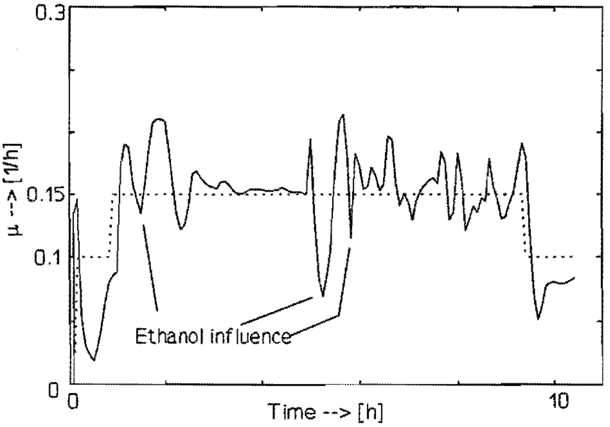
The proportional and integral feed-back gain of the PI controller in the μ-controller were initially tuned based on simulations. Some early experiments were dedicated to tune these parameters based on experimental data. If feed-back gains are too low, responsiveness is needlessly sacrificed. It they are too high, disturbances will cause badly damped oscillations or even undamped oscillations and instabilities.

Example 7.12 *PI tuning*

First the controller was tested with the settings found from the simulation data as described in Example 7.4 and Table 7.3. It was found that the settings were too sluggish specially for the specific growth rate controller. The parameters of the PI-controller were tuned more aggressive and the result of this can be seen in Figure 7.15, part (a). Due to the aggressive tuning and the fact that the actual delay was two samples and the algorithm was corrected for a delay of one sample, oscillations were found with a period of four samples. In Figure 7.15, part (b), the final PI tuning performance is shown. The dips in μ marked 'ethanol influence' are caused by faulty ethanol concentration estimates. This problem will be addressed in the next Example.



(a)



(b)

Figure 7.15 *Too aggressive, part (a) and final PI-tuning, part (b)*

Nevertheless, this figure demonstrates, that the PI is now tuned such, that disturbances on μ cause damped oscillations. Later in the experiment control becomes worse. This can be attributed to the stirrer speed and air flow interference. The oxygen controller was not switched to manual mode in this experiment. ■

The parameters of the PI-controller are given in Table 7.6. Compared to the ones found with the simulation data little has changed, see Table 7.3 and 7.4.

No ethanol production occurs in the current series of experiments. The true ethanol concentration will be zero. If the compensation tries to compensate for ethanol consumption because it sees a non-zero ethanol concentration, extra noise is introduced in the glucose feed rate. For an example of this, see the dips marked 'ethanol influence' in Figure 7.14. It was therefore decided to use compensation 3, which does not compensate for growth on ethanol in the glucose feed rate, unless there is excess amount of ethanol. It was also decided to have the PI act upon the compensation according to Figure 7.6 part (a) instead of part (b). This avoids that the compensation assumes ethanol production if values for μ become too high. The extra a-priori knowledge incorporated in the observer, as presented in Example 7.3, was removed. The original observer as presented in Chapter 6 is used during the rest of the experiments.

Table 7.6 *PI-controller settings*

μ-controller	K_p	0.1
	K_i	0.1
Ethanol-controller	K_p	9
	K_i	3
	p	0.9

During the experiments it was observed, that set-point changes were always accompanied by a considerable overshoot. To find out whether this was caused by the yeast or by the control, the glucose feed rate was studied.

Example 7.13 *Normalized glucose feed rate*

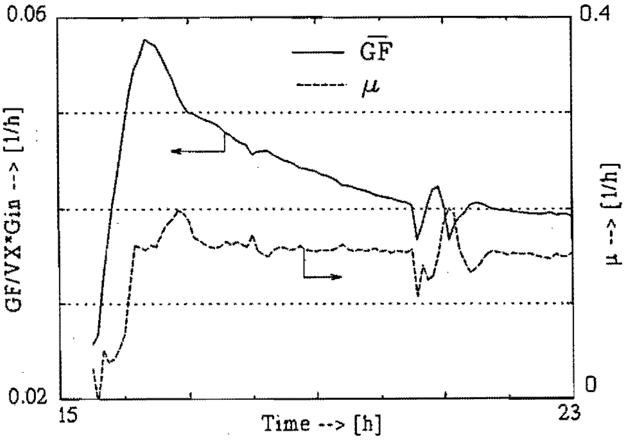


Figure 7.16 *Apparent increase in yield*

To remove the effects related to growth, the glucose feed rate was normalised to the estimated amount of yeast first. This showed, that the amount of glucose needed per mole biomass seemingly decreased during an experiment, see Figure 7.16.

Theoretically several factors can cause the decrease of the yield. If the initial amount of biomass is estimated too low, more glucose is ***seemingly*** needed per mole yeast to achieve a certain growth. As yeast grows, the error on the initial estimate becomes of less relative importance. The normalised glucose feed rate should then converge to a constant value. Attempts to give μ and $\overline{GF}$ a similar pattern show, that only unrealistically large errors on VX_0 could give the results of Figure 7.16. Off-line measurements of VX_0 using dry weight data contradict this possibility.

An error in the specific maintenance rate can also deform the graph of $\overline{GF}$ to the pattern shown in Figure 7.16. The maintenance rate should have a value over three times larger than its current value to match μ and $\overline{GF}$. This results in significantly lower estimates of the amount of yeast, as more yeast is consumed. This is contradicted by the dry weight data as well, which indicates that the observer gives reliable estimates of the amount of biomass.

Another possibility is a non-linearity in the pump. If $\overline{GF}$ remains constant over a period of time, *GF* actually increases because of growth. If the yield of the pump for large values of the glucose flow is higher than expected, this explains the observed anomaly. There should be a severe non-linearity, however, to give the result shown. From t = 17 [h] to t = 21 [h], the amount of yeast has not even doubled. Calibration experiments for the pump operated the pump from 10% up to 90% of its full capacity. Their results do not support the assumed non-linearity.

Apart from measuring errors a further explanation is, that the yeast adapts to the requested growth rate. Finding a physical mechanism for such an adaptation is a challenge for biochemists and biotechnologists. The adaptation is certainly not covered by Eq (3.1) - Eq. (3.4). From a control point of view, the time constant of the adaptation is large enough to be absorbed by the PI-controller. The current control structure can keep μ constant, albeit a little above its set-point. As long as the observed phenomenon is not modelled adequately, it seems best to accept this solution. Before the effect can be incorporated in the control, it should be clear, whether the effect is re-triggered after a new set-point change, whether time-constants are the same for different specific growth rates and so on. ■

Now it is clear, that one of the assumptions for all compensations does not hold. The yields of all pathways are not known constants. It was decided to leave any dynamic effects out of the compensation and tune it based on experimental steady state values. Steady states can only be expected if neither ethanol production nor consumption occurs. Therefore the experimental tuning can only cover oxidative growth on glucose and maintenance. Several experiments were carried out. The μ set-point was set to various values and was kept constant until a steady state seemed to be reached. A straight line was fitted to the experimental data using a least squares criterion. The slope of this line was considered most important, as any errors in the offset will be corrected by the integral part of the PI-controller. This will then influence control performance no further. The tuning found was

According to Eq. (3.1) - Eq. (3.4), the value of 0.2005 in Eq. (7.11) is equal to $1/b$, so

$$\frac{GF}{VX}\cdot S_{in} = 0.2005\cdot\mu_d+0.0093 \tag{7.11}$$

that b should be equal to 4.99. The value used by the observer is 3.65. These values differ 37%. Observer performance seems too accurate to support the value of 4.99. A possible explanation is, that the actual yeast behaviour is not in the model set of Eq. (3.1) - Eq. (3.4). This means no physical interpretation for the value of 0.2005 is available. Before bothering biochemists for providing a realistic interpretation of Eq. (7.11), measuring errors should be ruled out. Several factors can cause such errors.

- the yield of the pump may be higher than assumed.
- the true concentration of glucose in the fed-batch feed may be higher than expected.
- the observed value for VX is possibly too large.

The decrease in weight of the glucose supply is logged. The logged data seem to confirm, that the yield of the pump is too high. The total volume of the broth at the end of an experiment is lower than the integration of the assumed glucose feed rate predicts. This can be attributed to either a lower yield of the pump than expected or evaporation effects. The latter are quite likely, as blowing air through the broth enhances evaporation. The hypothesis of higher pump yield is contradicted only by the calibration experiments. These take half an hour to an hour, depending on the number of data points used for calibration and the time taken per data point. The pump is driven in a feed-forward way only. Errors occurring if the pump is used over a longer period of time are therefore not detected. For this reason, the calibration experiments may not be totally reliable. Nevertheless it is felt, that further backup for this hypothesis is needed before accepting it. This could be done by means of a calibration experiment covering several hours.

It is highly unlikely, that the concentration of glucose in the fed-batch feed is higher than expected. The error in the amount of glucose in the fed-batch feed is less than 1 ‰. The confidence in the observed values is also such, that an error of 37% in VX seems impossible. From other experiments it is known, that the yeast produces ethanol if the glucose concentration increases suddenly. As a rule of thumb, an increase of 0.1 [g.dm^{-3}] in the glucose concentration is sufficient to trigger this effect. This may obstruct set-point changes. In order to avoid this effect, μ set-point changes are filtered by a finite impulse response filter. A description of the FIR-filter design is given in Example 7.14.

Example 7.14 *Filtering set-point changes*

The step response of the FIR-filter should start and finish smoothly. This suggests the triangular impulse response shown in Figure 7.17. A typical value for the length of the impulse response is taken in this Figure. At every set-point change it is verified whether the current length is sufficient to keep the maximum change in glucose feed rate below a critical value. This critical value is 0.1 [g.dm^{-3}] glucose divided by the sampling time, yielding an upper bound on the change in the glucose feed rate. As the transfer from $\Delta\mu_{set}$ to ΔGF is proportional to the amount of yeast, the length of the impulse response will increase towards the end of the experiment. Large set-point changes will also induce an increase in the impulse response length. The DC-gain of the filter is always equal to 1: set-point changes should be smoothened, not changed in magnitude.

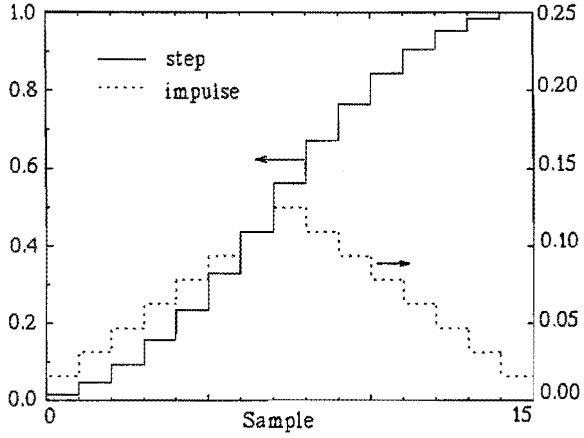


Figure 7.17 *FIR-filter response* ■

Unfortunately, the FIR-filter approach proved unsuccessful in removing the overshoot from set-point changes. The overshoot remained even roughly the same. The only effect of this filter turned out to be a pointless delay on the set-point change.

As overshoot remained on the response to step-changes after re-turning the compensation and rate limiting μ_{set}, the suspicion arose that the overshoot was caused by not modelled yeast dynamics. To investigate this, set-point changes were requested more or less in open loop: the PI-controller was switched off during these trials. The experiments are described in Example 7.15.

Example 7.15 *Dynamic compensation*

As *GF* still depends on *VX* the process is not completely operated in open loop. μ_{set} was changed from 0.048 [h^{-1}] to 0.055 [h^{-1}]. Larger changes could not be investigated, because they caused the specific growth rate to drift away to ethanol producing values. The suspicion was confirmed: the change was accompanied with overshoot (see Figure 7.18).

The dynamic behaviour of the yeast will be accounted for in the control structure by means of a block "dynamic compensation." Together with the FIR filter of the previous paragraph, the final structure will be as depicted in Figure 7.21. The dynamic compensation should compensate the yeast dynamics just discovered. It operates on a feed-forward basis. Therefore a model was fitted to the experimental data using a least squares output error criterion. A second order model proved sufficient to describe the overshoot. The model poles and zeros are given in Table 7.7. Removing the zero in -57.7310 deteriorated the fit, so it was decided to leave it in the model. The transfer function of the dynamic compensation is the inverse of the model. The unstable pole in -57.7310 is relocated to 1/-57.7310.

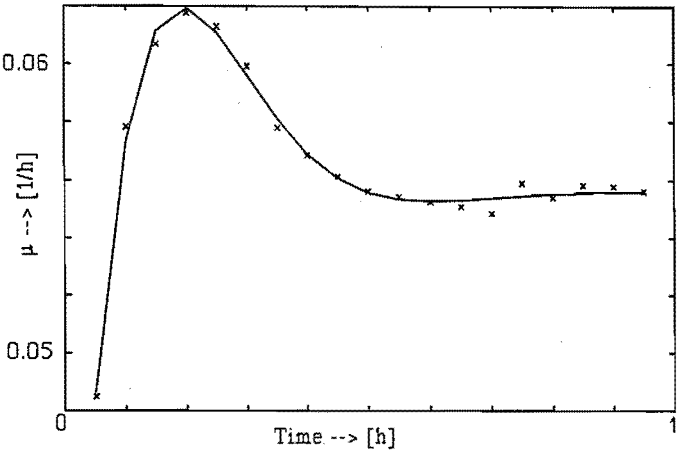
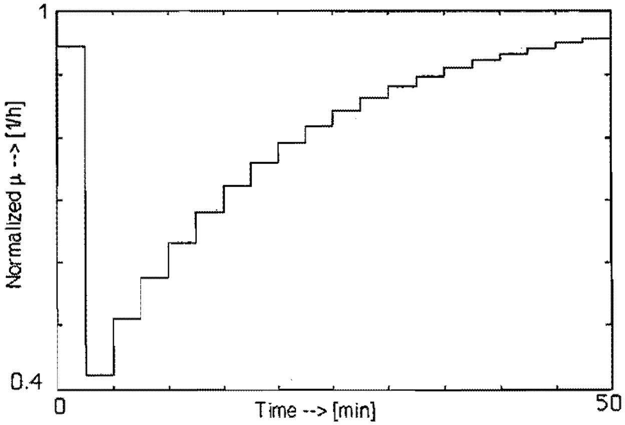


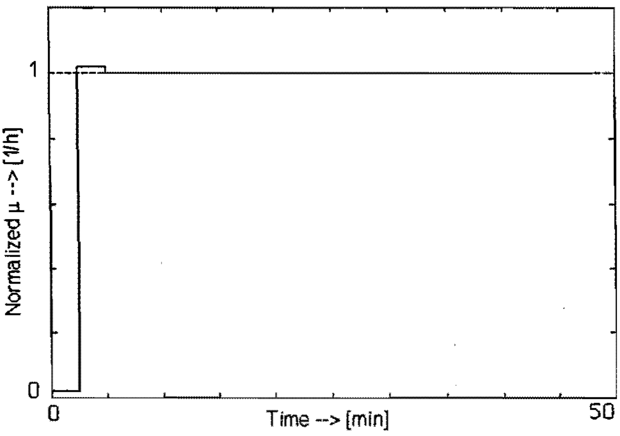
Figure 7.18 *Step response to open loop excitation of the GF*

Table 7.7 *Model poles and zeros and resulting controller*

model		controller	
zeros	poles	zeros	poles
0.6430+0.2581i	0.8422	0.8422	0.6430+0.2581i
0.6430-0.2581i	-57.7310	1/-57.7310	0.6430-0.2581i



(a)



(b)

Figure 7.19 *Step-responses without and with controller*

As the compensation operates on a feed-forward basis, its sample rate need not be limited by the sample rate of the measuring equipment. To improve control performance, the compensation is run at twice the sample rate of the measuring equipment. The transfer function is converted to continuous time and then back to discrete time. Theoretically, even higher sample rates can be used.

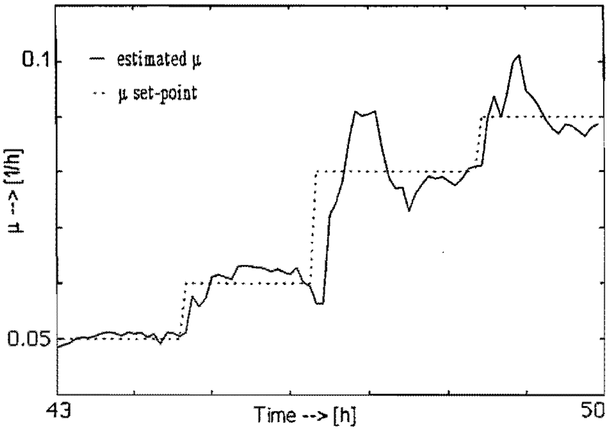


Figure 7.20 *Response to small steps*

The higher the sample rate, however, the more sensitive the compensation will be to errors in the linear model. This is undesirable, because the model is based on a few samples only and because the true process is probably not in the model set used. The compensation response and the compensation and process response to a step change are depicted in Figure 7.19 part (a) and (b). The response of the process is improved by the addition of a dynamic compensation. In Figure 7.20 results are shown for steps of the order 0.01 [h^{-1}]. This is comparable to the size of the step on which the prediction model is based.

Within roughly one or two hours a set-point change is completed. In previous experiments, set-point changes took 3 hours and longer. Overshoot is still present but small compared to experiments without dynamic compensation.

For large steps (not shown in Figure 7.20), large overshoot is still present on the response. Yeast behaviour for large steps is apparently not a linear extrapolation from the behaviour for small steps. This is not surprising. Remarkable about the response to large steps with dynamic compensation is, that no ethanol production occurs, even though the specific growth rate temporarily reaches values for which ethanol production was observed in other experiments. ■

The final control structure is presented in Figure 7.21. Compared to Figure 7.2 some changed are made. The DOT-controller is no longer present. A local controller or manual changes of the *SS* and *AF* regulate the *DOT*. The μ-controller and the compensation block of Figure 7.2 have merged in one μ-controller block of Figure 7.21. All the functions of the two separate blocks in Figure 7.2 are preserved. The compensation is now at the right hand side, the multiplication of $V\hat{X}$ and the output of the summer above the PI-block, and at the left hand side, the static compensation block. The PI is still present, as part (b) of Figure 7.6, but with a 2 sample delay correction. The ethanol controller remains the same, only an additional filter, FIR-filter, is added to smoothen set-point changes. However, the FIR-filter does not give a better controller performance, see Example 7.24.

With this final controller two final experiments are done; one with specific growth rate set-point changes during a few hours and one with a specific growth rate ramp set-point signal. The results of the two experiments are discussed in Example 7.16.

Example 7.16 *Final experiments*

The FIR-filter and the dynamic compensation are turned off during these experiments to get an idea of the performance of the "basic" controller. The specific growth rate is not chosen too high, $\mu = 0.1$ [h^{-1}], in order to be able to perform a long, in time that is, experiment. Using higher specific growth rates could give complications such as oxygen limitation of viscosity effects. The perturbation on the specific growth rate is a PRBS-sequence with a minimum pulse duration of 8 [min] and an amplitude of 0.01 [h^{-1}]. The sampling time of the controller was 3 [min]. The results are depicted in Figure 7.22

During the first 5 [h] almost every positive set-point change results in a peak on the specific growth rate due to overshoot. As can be seen not every step has the overshoot so compensating for it with the dynamic compensation would not always leas to a successful control action. The overshoot at the negative set-point changes is not very big during the complete experiment. After 4-5 [h] the culture seems to have adapted to the PRBS sequence and follows the changes quite good, the overshoot is now more or less proportional to the duration of the PRBS blocks. Why the overshoot becomes more or less proportional to the duration of a PRBS block is not known.

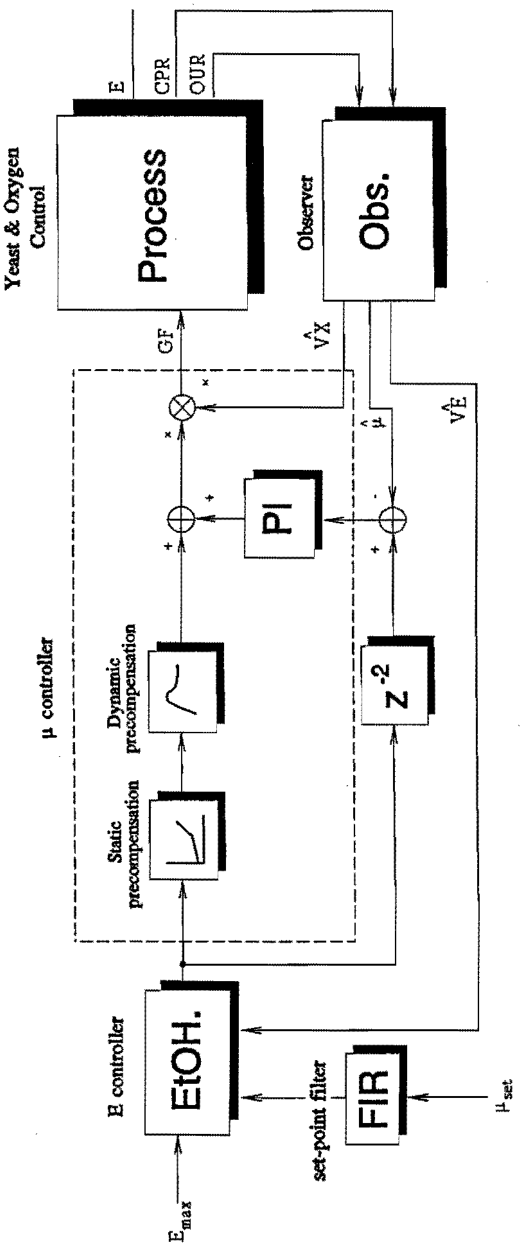

Yeast & Oxygen Control
Process
E
CPR
OUR
Observer
Obs.
GF
$\hat{VX}$
$\hat{\mu}$
$\hat{VE}$
x
x
+
+
+
-
PI
z^{-2}
Dynamic precompensation
Static precompensation
μ controller
E controller
EtOH.
set-point filter
FIR
μ_{set}
E_{max}

Figure 7.21 *Final control structure*

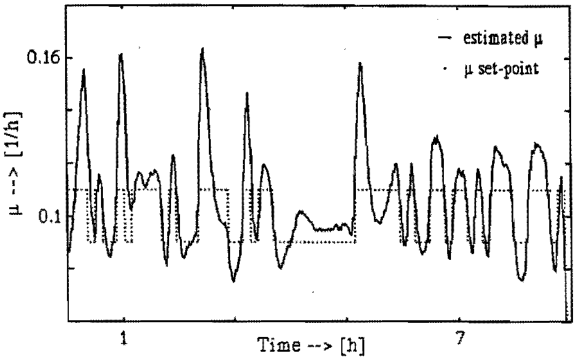


Figure 7.22 *Specific growth rate set-point changes*

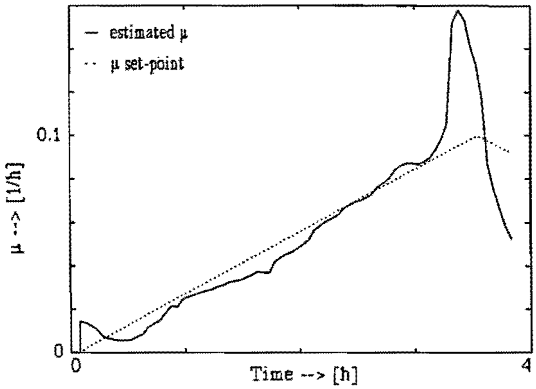


Figure 7.23 *Specific growth rate ramp set-point signal*

Another interesting experiment is to let the specific growth rate set-point follow a ramp signal. A slowly increasing ramp was applied, starting from $\mu = 0.0$ [h^{-1}] increasing over 3.5 [h] to $\mu = 0.1$ [h^{-1}]. The results are given in Figure 7.23. During the first two hours the control is not as tight as it should be due to adaptation of the yeast to the new situation. After these two hours the specific growth rate follows the ramp quite good. At the end, near $\mu = 0.09$ [h^{-1}] a strange phenomenon occurs, the growth rate increases very rapidly. Why this happens is not known. A possible explanation could be the degradation of trehalose in the yeast itself, see Keulers *et al.*, 1993. ■

The above result show that the controller is able to control the specific growth rate on set-point but has trouble if 'special' effects occur.

7.4 Conclusions

The control structure is able to maintain a constant set-point on μ while obeying a limit on the ethanol concentration in simulations. Performance deteriorates if ethanol production or consumption occurs. In the simulations this could be counteracted by incorporating extra a priori knowledge in the control. If the limit on the ethanol concentration is exceeded, the ethanol concentration is successfully brought back within the allowed range. However, the overshoot is too large to consider the current control structure as a candidate for ethanol control.

In experiments, occasional ethanol production or consumption indicates that control is poor under these conditions. Accurate information on the ethanol concentration is badly needed then. New yeast properties were discovered: the yield for oxidative growth on glucose increases throughout an experiment, but converges to a constant value. The time constant associated to this behaviour is large enough to be handled by the control structure.

Changes in air flow and stirrer speed disturb the oxygen uptake rate and carbon-dioxide production rate measurements, and consequently the specific growth rate estimates. To enable μ control the oxygen controller should be switched to manual mode, until a rate limitation has been implemented on AF and SS.

Open-loop experiments uncovered dynamic effects during μ changes. Compensating for these dynamics by a predictive feed-forward controller improves performance. A prediction model is currently only available for small steps. Models for larger steps should be identified in further experiments.

Only part of the yeast behaviour is covered by the experiments carried out. Ethanol production and consumption have not been investigated properly. The control structure has provisions for these situations but these are tuned to the simulation model. Identification experiments should be carried out first to tune the structure to the real process.

The implemented control uses no explicit adaptation scheme. Other control structures use adaptive schemes to cover for the different dynamics for different metabolic pathways. A major drawback of this approach is, that control is deteriorated immediately after a change in the used pathways: the adaptation requires several samples to converge to a new model. As the change can be detected using RQ, it is possible to eliminate this by changing the model used on a feed-forward basis. This is what the compensations in our control structure try to do. Further, the on-line identification suffers from the fact, that the input signal of the process is not very rich. External disturbances are either non-stationary peaks or slowly changing offsets. These disturbances certainly do not excite the process persistently, so that the on-line identification results may be unreliable. For our particular case, there is another argument not to use adaptive schemes.

The two concepts mentioned above could be mixed: a state recognition algorithm changes the models to be used on a feed-forward basis, an adaptation scheme tunes these models to the actual situation. In this case, it is again doubtful whether the models for ethanol production and consumption ever get a chance to converge.

The current control suffers from inaccurate ethanol estimates. A relatively large sensitivity to errors in these estimates can be expected for sub-optimal growth based on physical reasoning: a small amount of ethanol can make a big difference in the overall growth rate. This is probably best demonstrated at the end of the batch phase. When all ethanol has been consumed the oxygen consumption rate falls to a low level within a minute or so (this can be seen from the oxygen concentration in the broth). There is no gradual changeover when the ethanol concentration decreases. Other control techniques will have to cope with this problem as well.

References

Agrawal, P., Koshy, G. and M. Ramseier (1989). An Algorithm for Operating a Fed-Batch Fermentor at Optimum Specific Growth Rate, *Biotechnology and Bioengineering*, Vol. 33, pp. 115-125.

Aiba, S., Nagai, S. and Y. Nishizawa (1976). Fed Batch Culture of *Saccharomyces cerevisiae:* A Perspective of Computer Control to Enhance the Productivity in Bakers' Yeast Cultivation. *Biotechnology and Bioengineering,* Vol. 18, pp. 1001-1016.

Andersen, M.Y., Brabrand, H. and S.B. Jørgensen (1991). Model Based Control of a Continuous Yeast Fermentation. *Proceedings ACC*, Boston, USA, pp. 1329-1334.

Axellson, J.P. (1989). Modelling and Control of Fermentation Processes, *Ph.D. Thesis,* Lund Institute of Technology, Sweden.

Aynsley, M., Peel, D., Hofland, A.G., Montague, G.A. and A.J. Morris (1991). Real-Time Expert System in the Control of a Fermentation Plant. *Proceedings of the ECC*, Grenoble, France, pp. 236-241.

Bastin, G. (1991). Nonlinear and Adaptive Control in Biotechnology: A Tutorial. *Proceedings of the ECC,* Grenoble, France, pp. 2001-2012.

Bastin, G. and D. Dochain (1990). *On-Line Estimation and Adaptive Control of Bioreactors*, Elsevier, Amsterdam.

Bastin, G., Dochain, D., Haest, H., Installé, M. and Ph. Opdenacker (1982). Modelling and Adaptive Control of a Continuous Anaerobic Fermentation Process. *IFAC Symposium, Modelling and Control of Biotechnical Processes*, Helsinki, Finland, pp. 299-306.

Béteau, J.F., Ferret, E.V. Lakrori, M.L. and A.C. Chéruy (1991). Bioprocess Control: an Original Approach Taking into Account some Bioprocess Constraints. *Proceedings of the ACC,* Boston, USA, pp. 1335-1340.

Biryukov, V.V. (1982). Computer Control and Optimization of Microbial Metabolite Production. *IFAC Symposium, Modelling and Control of Biotechnical Processes,* Helsinki, Finland, pp. 135-144.

Cardello, R.J. and K.Y. San (1988). The Design of Controller for Batch Bioreactors. *Biotechnology and Bioengineering*, Vol., 32, pp. 519-526.

Chamilothoris, G., Renaud, P.Y., Sévely, Y. and P. Vigie (1988). Adaptive Predictive Control of a Multistage Fermentation Process. *International Journal of Control,* Vol. 48, No. 3, pp. 1089-1105.

Chen, J., Tannahill, A.L. and M.L. Shuler (1985). Design of a System for the Control of Low Dissolved Oxygen Concentrations: Critical Oxygen Concentrations for *Azotobacter vinelandii* and *Escerichia coli. Biotechnology and Bioengineering,* Vol. 27, pp 151-155.

Cheruy, A. and A. Durand (1979). Optimisation of Eythromycin Biosynthesis by Controlling pH and Temperature. - Theoretical and Practical Aspects Application. *Biotechnology and Bioengineering,* Vol. 19, pp. 303-320.

Clark, T.A., Hesketh, T. and T. Seddon (1985). Automatic Control of Dissolved Oxygen Tension via Fermenter Agitation Speed. *Biotechnology and Bioengineering*, Vol. 27, pp. 1507-1511.

Constantides, A., Spencer, J.L. and E.L. Gaden (1970). Optimum Profiles for Batch Penicillin Fermentation. *Biotechnology and Bioengineering*, Vol. 12, pp. 1081-1090.

Cooney, C.L. and J.R. Schwartz (1982). Application of Computer Control to Yeast Fermentation. *IFAC Symposium, Modelling and Control of Biotechnical Processes*, Helsinki, Finland, pp. 243-251.

D'Ans, G., Kokotovic, P. and D. Gottlieb (1971). A Non-Linear Regulator Problem for a Model of Biological Waste Treatment. *IEEE Transactions on Automatic Control*, pp. 341-347.

Dairaku, K., Izumuto, E., Mourikawa, H., Shioyta, S. and T. Takamatsu (1983). An Advance Micro Computer Control System in a Baker's Yeast Fed-Batch Culture Using a Tubing Method. *Journal of Fermentation Technology*, Vol. 61, No. 2, pp.189-196.

Dochain, D. (1990). Multivariable Adaptive Control of Nonlinear Completely Mixed Bioreactors. *Proceedings of the ACC*, San Diego, USA, pp. 2661-2666.

Dochain D. and G. Bastin (1984). Adaptive Identification and Control Algorithms for Nonlinear Bacterial Growth Systems. *Automatica*, Vol. 20, No. 5, pp. 621-634.

Dochain D. and M. Perrier (1991). Nonlinear Adaptive Controllers for Non Minimum Phase Bioreactors. *Proceedings of the ACC*, Boston, USA, pp. 1311-1316.

Golden, G.C. and B.E. Ydstie (1987). Nonlinear Adaptive Optimization of Continuous Bioreactors. *IFAC worldcongress*, Munich, Germany, pp. 102-113.

Harmon, J., Svoronos, S.A., and G. Lyberatos (1987). Adaptive Steady State Optimization of Biomass Productivity in Continuous Fermenters. *Biotechnology and Bioengineering*, Vol. 25, pp.335-344.

Keulers, M. and G. Reyman (1991). The Application of the GPC Algorithm to a Fed-Batch Fermentation Process. A Simulation Study. *Proceedings of the ESC*, Copenhagen, Denmark, pp. 879-884.

Keulers, M., Ariaans, L. and M.L.F. Giuseppin (1993). The Dynamic Step Response of a Fed-Batch Bakers' Yeast Process. *Preprints of the 6th European Congress on Biotechnology*, Florence, Italy.

Ko, K.Y.-J., McInnis, B.C. and G.C. Goodwin (1982). Adaptive Control and Identification of the Dissolved Oxygen Process. *Automatica*, Vol. 18, No. 6, pp. 727-730.

Landau, I.D., Samaan, M. and M. M'Saad (1990). Robust Performance Oriented Adaptive Control for Biotechnological Processes, a Tutorial. *Proceedings of the ACC*, San Diego, pp. 2684-2687.

Lee, S.C., and Hwang, Y.B., Chang, H.N. and Y.K. Chang (1991). Adaptive Control of Dissolved Oxygen in a Bioreactor. *Biotechnology and Bioengineering*, Vol 37, pp. 597-607.

Lochner, G., Sonnleitner, B. and A. Fiechter (1992). Software and Implementation for Automated Decision Making in Bioprocess Control. *Process Control and Quality*, Vol. 2, pp. 257-274.

Manchanda, S., Willis, M.J., Tham, M.T., Di Massimo, C. and G.A. Montague (1991). An Appraisal of Nonlinear Control Philosophies for Application to a Biochemical Process. *Proceedings of the ACC*, Boston, USA, pp. 1317-1322.

Mészáros, A. and V. Báleš (1992). A Contribution to Optimal Control of Fed-Batch Biochemical Processes. *Bioprocess Engineering*, Vol. 7, pp. 363-367.

Modak, J.M. and H.C. Lim (1992). Optimal Mode of Operation of Bioreactors for Fermentation Processes. *Chemical Engineering Science*, Vol. 47, No. 15/16, pp. 3869-3884.

Montague, G.A., Morris, A.J. Wright, M., Aynsley, M. and A.C. Ward (1986). Online Estimation and Adaptive Control of Bakers' Yeast Fermentation. *IEE Proceedings*, Vol. 133, Pt. D, No. 5, pp. 340-345.

Montgommery, P.A., Williams, D. and B.H. Swanick (1985). Control of Fermentation Process by a On-Line Adaptive Technique. *IFAC Symposium, Modelling and Control of Biotechnological Processes.* Noordwijkerhout, The Netherlands, pp. 111-119.

O'Connor, G.M., Sanchez-Riera, F. and C.L. Cooney (1992). Design and Evaluation of Control Strategies for High Cell Density Fermentations, submitted to *Biotechnology and Bioengineering.*

Ohno, H., Nakanishi, E. and T. Takamatsu (1976). Optimal Control of a Semibatch Fermentation. *Biotechnology and Bioengineering*, Vol. 18, pp. 847-864.

Pomerleau, Y, and G. Viel (1992). Industrial Application of Adaptive Nonlinear Control for Bakers' Yeast Production. *IFAC Symposium, Modelling and Control of Bioetechnical Processes*, Colorado, USA, pp. 315-318.

Pomerleau, Y., Perrier, M. and D. Dochain (1989). Adaptive Nonlinear Control of the Bakers' Yeast Fed-Batch Fermentation, *Proceedings of the ACC*, Pitsburg, pp. 2424-2429.

Pons, M.N., Bordet, J. and J.M. Engasser (1982). A Micro-Mini Computer Hierarchical Control System for a Laboratory Scale Fermentor. *IFAC Symposium, Modelling and Control of Biotechnical Processes*, Helsinki, Finland, pp. 259-265.

Poulisse, H.N.J. and C. van Helden (1985). Adaptive LQ Control of Fermentation Processes, *IFAC Symposium, Modelling and Control of Biotechnological Processes*, Noordwijkerhout, The Netherlands, pp. 43-47.

Proell, T., Hilaly, A., Karim, M.N. and D. Guyre (1991). Comparison of Different Optimization and Control Schemes in an Industrial Scale Microalgae Fermentation. *Proceedings of the ACC*, Boston, USA, pp. 1323-1328.

Queinnec, I., Dahhou, B. and M. M'Saad (1991). On Adaptive Control of Fed-Batch Fermentation Processes. *Proceedings of the ECC*, Grenoble, France, pp. 1636-1641.

Ramseier, M., Agrawal, P. and D.C. Mellichamp (1991). Nonlinear Adaptive Control of Fermentation Processes Utilizing *a priori* Knowledge. *Proceedings of the ACC*, Boston, USA, pp. 1305-1310.

Rolf, M.J. and H.C. Lim (1985). Experimental Adaptive On-Line Optimization of Cellular Productivity of a Continuous Bakers' Yeast Culture. *Biotechnology and Bioengineering*, Vol. 27, pp. 1236-1245.

Samaan, M., Dahhou, T., Queinnec, I. and M. M'Saad (1991). Experimental Results in Adaptive Control of Biotechnological Processes. *Proceedings of the ACC*, Boston, USA, pp. 2679-2683.

Shimizu, H., Araki, K., Shioya, S., Suga, K. and E. Sada (1989). Optimal Production of *Glutathione* by Controlling the Specific Growth Rate of Yeast in Fed-Batch Culture. *IFAC Symposium, Production Control in the Process Industry*, Osaka, Japan, pp. 67-72.

Shimizu, H., Takamatsu, T., Shioya, S. and K.-I. Suga (1989a). An Algorithmic Approach to Constructing the On-Line Estimation System for the Specific Growth Rate. *Biotechnology and Bioengineering*, Vol. 33, pp. 354-364.

Shimizu, H., Shioya, S., Suga, K.-I., and T. Takamatsu (1989b). Profile Control of the Specific Growth Rate in Fed-Batch Experiments. *Applied Microbiology and Biotechnology*, Vol. 30, pp. 276-282.

Shioya, S. (1992). Optimization and Control in Fed-Batch Bioreactors. *Advances in Biochemical Engineering Biotechnology*, Vol 46, pp. 111-142.

Siimes, T., Nakajima, M., Yada, H., Asama, H., Nagamune, T., Linko, P. and I. Endo (1992). Object-Oriented Fuzzy Expert System for On-Line Diagnosing and Control of Bioprocesses. *Applied Microbiology and Biotechnology*, Vol. 37, pp. 756-761.

Sterback, Z. and J. Votruba (1993). An Expert System Applied to the Control of an Industrial Scale Bioreactor. *Chemical Engineering Journal*, Vol. 51, No. 2, pp. B35-B47.

Suzuki, T., Yamane, T. and S. Shimizu (1986). Control of Carbon Source Supply and Dissolved Oxygen by Use of Carbon Dioxide Concentration of Exhaust Gas in Fed-Batch Culture. *Journal of Fermentation Technology*, Vol. 64, No. 4, pp. 317-326

Takamatsu, T., Shioya, S., Okada, Y. and M. Kanda (1985). Profile Control Scheme in a Bakers' Yeast Fed-Batch Culture. *Biotechnology and Bioengineering*, Vol. 27, pp. 1675-1686.

Takamatsu, T., Hashimoto, I., Shioya, S., Mizuhara, K, Koike, K. and H. Ohno (1975). Theory and Practice of Optimal Control in Continuous Fermentation Process. *Automatica*, Vol. 11, pp. 141-148.

van Impe, J. (1993). Modelling and Optimal Adaptive Control of Biotechnological Processes. *Ph.D. Thesis*, Katholieke Universiteit Leuven, Belgium.

Verbruggen, H.D., Eeldering, G.H.B. and P.M. van Broecke (1985). Multiloop Controlled Fed-Batch Fermentation Processes Using a Selftuning Controller. *IFAC Symposium, Modelling and Control of Biotechnological Processes*, Noordwijkerhout, The Netherlands, pp. 121-126.

Veres, A., Nyeste, L., Kurucz, I., Kirchknopf, L., Szigeti, L. and J. Holló (1981). Automated Fermentation Equipement 1. Program-Controlled Fermentor. *Biotechnology and Bioengineering*, Vol. 23, pp. 391-404.

Wang, H.Y., Cooney, C.L. and D.I.C. Wang (1977). Computer-Aided Bakers' Yeast Fementations. *Biotechnology and Bioengineering*, Vol. 19, pp. 69-86.

Wang, H.Y., Cooney, C.L. and D.I.C. Wang (1979). Computer Control of Bakers' Yeast Production. *Biotechnology and Bioengineering*, Vol. 21, pp. 975-995.

Waterworth, G. and B.H. Swanick (1981). Investigation into the Adaptive Control of a Fermentation Process. *Transaction of the Institute of Measurement and Control*, Vol. 3, No. 1, pp. 39-56.

Williams, D. and Montgomery (1986). Multivariable Adaptive Control of Bakers' Yeast Fermentation. *IEE Proceedings*, Vol. 133, Pt. D, No. 5, pp. 247-253

Williams, D., Yousefpour, P. and B.H. Swanick (1984). Online Adaptive Control of a Fermentation Process. *IEE proceedings*, Vol. 131, Pt. D, No. 4, pp. 145-156.

Williams, D., Yousefpour, P. and E.M.H. Wellington (1986). On-Line Adaptive Control of a Fed-Batch Fermentation of *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering*, Vol. 28, pp. 631-645.

Wu, W.-T., Chen, K.-C. and H.-W. Chiou (1985). On-line Optimal Control for Fed-Batch Culture of Bakers' Yeast Production. *Biotechnology and Bioengineering*, Vol. 27, pp. 756-760.

Wu, W.-T., Jang, W.-D. and S.-C. Wu (1987). On-line Control for Cultivation of *Saccharomyces cerevisiae* via Weighted Moving Identification. *Chemical Engineering Journal*, Vol. 36, B1-B6.

Yamané, T. and Kume, T., Sada, E. and T. Takamatsu (1977). A Simple Optimization Technique for Fed-Batch Culture. *Journal of Fermentation Technology*, Vol. 55, No. 6, pp 587-598.

Yano, T., Kobayashi., T. and S. Shimizu (1981). Control System of Dissolved Oxygen Concentration Emmploying a Microcomputer. *Journal of Fermentation Technology*, Vol. 59, No. 4, pp. 295-301.

Yano, T., Mori, H., Kobayashi., T. and S. Shimizu (1979). Control System for Dissolved Oxygen Concentration in Aerobic Cultivation. *Journal of Fermentation Technology*, Vol. 57, No. 2, pp. 91-98.

Zheng, F.Y. and B. Dahhou (1991). Model Reference Adaptive Estimation and Control Applied to a Continuous Flow Fermentation Process. *IFAC Symposium, Advanced Control of Chemical Processes*, Toulouse, France, pp. 89-94.

Zheng, F.Y., Quennec, I., Dahhou, B., Pourciel, J.-B. and G. Goma (1991). FPC-ICAD: an Intelligent CAD for Fermentation Process Control. *Proceedings of the ECC*, Grenoble, France, pp. 86-91.

8

Conclusions

In this thesis a controller has been developed for the control of the specific growth rate of *Saccharomyces cerevisiae* grown in a fed-batch operated bioreactor. This controller has been developed along the lines of building simulation and estimation models, testing structure identification techniques on the data of the simulated process, the design of an observer for non-measurable process variables, applying the models and observer in practice, and finally the developing of a controller.

The first part of our strategy concerned the process description, where insight into the yeast itself, the kinetics and dynamics of the yeast, and the important process variables such as biomass concentration, glucose concentration, and ethanol concentration, was attained. The major conclusion from this part of the work is that there is still a lack of reliable sensors for measuring on-line essential process variables such as biomass concentration and glucose concentration.

Yet, on the basis of experimental data and a physiological/physical model of *Saccharomyces cerevisiae* two simulation models were obtained: one for batch phase growth and one for fed-batch phase growth. These models describe reality fairly well, , they describe most of the transients occurring, but they are sensitive to the geometry of the used fermenter, especially for the dissolved oxygen tension. The dissolution of oxygen in liquid depends on mixing items like baffles, on location of probes and inlet tubes. The geometry dependent parameters have to be identified for every major change of the fermenter system.

One of the other findings is that for the batch phase simulation model not all the effects occurring can be modelled. Effects such as pyruvate excretion can be measured indirectly but the differential and algebraic equations of the simulation model do not describe the effects. As a result of this limited description of the effects occurring during the diauxic phase, which consists of the excretion and uptake of several other components, is not modelled well. It is thought that expanding the batch phase simulation model with additional equations, modelling the excretion and uptake of several components (i.e. pyruvate, acetaldehyde, and acetate) does not yield a better model as the extracellular presence of these components is strongly dependent on the initial glucose concentration.

The fed-batch phase simulation model does not suffer from these limitations and is a more simple model, because the excretion of components is not noticed during the fed-batch phase. The limitations of the fed-batch phase simulation model are found in the dynamic response of the yeast to glucose. Certain internal storage effects are not modelled and the simulation model gives a wrong dynamic response to some glucose flow signals.

Although some "special effects" of the yeast growth, both for batch phase and fed-batch phase growth, can not yet be modelled, the resulting simulation models are considered to be an adequate description of the batch phase and fed-batch phase laboratory process. The simulation models can be used to test (structure) identification and control techniques.

The linearized model of the fed-batch growth of *Saccharomyces cerevisiae* involves only oxidative growth and is a relatively simple second order model but needs the exponential compensation as a linearizing effect. The use of the model is restricted; only oxidative growth is modelled, and it can not be used in other growth regions. The obtained model is an output error model, which means that its prediction horizon is long, as could be seen from the validation data. The overall conclusion is that a non-linear compensation is required to obtain the linear model, which makes implementation of the model for control control difficult. Note that the non-linear compensation depends on the specific growth rate and, thus, is a dynamic compensation.

As a linear identification technique was not suited for identification of a control model, several non-linear structure identification techniques were applied. For the block-oriented structure identification technique using Volterra kernels it was found that the method provided the user with a nice indication of the structure of the process, but not all non-linearities of the laboratory process could be described by this method. Furthermore, the method has two separate layers, first structure identification, next parameter estimation. The non-linear model linear-in-the-parameters is much more flexible and provides the user with a structure and parameters at the same time. Although the physiological insight in the process is less than with the Volterra kernels, the resulting model is not at all complex as it is built-up out of a combination of lagged output and input signals. The Radial Basis Function network algorithm is even more flexible than the non-linear model linear-in-the-parameters algorithm at the price of a highly complex model and a low insight in the process structure. For both the non-linear model linear-in-the-parameters and the Radial Basis Function network algorithm it holds that they converge according to a user set criterium and that the accuracy to be obtained can be set. Note that the accuracy can be in conflict with the convergence if output noise is present. As a total conclusion the Volterra kernels based identification algorithm was not able to provide suitable models, whereas the last two methods were capable of modelling, in a one step ahead sense, the specific growth rate for a batch fermentation. These models are, however, not very suited for control design as they are one step ahead predictors.

Another way of estimation or prediction of the specific growth rate is to make use of an observer. An observer is developed for the specific growth rate and its metabolic parts plus the biomass concentration. Both for the batch and the fed-batch phase the observer gives accurate indication of the metabolic parts of the specific growth rate and the biomass concentration. Non-modelled extracellular components such as pyruvate are taken into account by the oxidative growth on glucose, which gives some errors in the readings for the batch. The main drawback of this observer is that it is dependent on the stoichiometry of the yeast and thus is non-generic.

Based on this observer a controller is designed that is capable of controlling the specific growth rate. The controller is built-up around the non-linear compensated process using the linear model identification, and a simple PI-controller. Fast varying patterns of set-points can not be followed if the amplitude is too large, as non modelled dynamic effects occur (storage/degradation of intracellular trehalose) that are not represented in the observer. Adjusting the controller to these effects can be accomplished, but leads to a rather complex control scheme, as these effects are non-linear and unpredictable. Slowly varying patterns of set-points with a small amplitude as well a ramp signal can be followed. The proposed structure for ethanol control could be enhanced with the metabolic specific growth rates μ_r and μ_e or replaced by a simple ethanol regulator.

The experiments shown throughout the thesis are considered to be reasonably good representations of the industrial scale process. On an industrial scale the glucose (molasses) feed rate has abrupt changes due to the instant addition of large amounts of feed. The feed rate is more a sequence of pulses, which can be seen as small batch phases or a non smooth fed-batch phase. The presented controller should be able to work on industrial scale as well if one readjusts the PI tuning parameters to the new situation. The PI parameters have to be readjusted due to the fact that the time constants change. There is one big difference with the laboratory scale fermenter: the fermenter is no longer mixed homogeneously. However, if one uses effluent-gas analysis the gas in the head space of the fermenter will act as an integrator, averaging all the reactions taking place.

The controller can handle small set-point changes but the laboratory process is sensitive to large changes in the glucose feed rate (or specific growth rate). A more sophisticated controller would not be able to regulate the process better than a PI-controller would. The storage events occurring during specific growth rate control are not yet understood completely, so no model for control can be made to compensate these events. One could argue that based on input-output data a model is readily available, but the storage events only occur at positive increases of the glucose feed rate and at certain specific growth rates. Some biotechnically based research has to be done before this problem can be solved. If, however, the problem is solved and control can be made even more tight, more sophisticated controllers could work as well and give better results than the PI-controller.

Due to the possibility of (tight) specific growth rate control, new possibilities and problems were encountered, such as low specific growth rate control and dynamic responses to 'high' concentration of extracellular glucose. These facts could not be noticed until recently, because they could not be measured or seen at all. Using the proposed control scheme new, prosperous and interesting research fields can be explored. These research fields include flavour and enzyme production at low, controlled specific growth rate. Furthermore, more fundamental research can be performed towards the quality of the culture *Saccharomyces cerevisiae* as most of the states can now be monitored on-line by the observer.

List of Examples

Example 1.1 *Control of intracellular events; a future approach* 4

Example 3.1 *Transition from batch phase to fed-batch phase of the laboratory process* 40
Example 3.2 *Experimental batch fermentation with* $S_0 = 30$ *[g.dm*$^{-3}$*]* 47
Example 3.3 *Experimental batch fermentation with* $S_0 = 5$ *[g.dm*$^{-3}$*]* 49
Example 3.4 *Experimental batch fermentations and the simulation model responses* 54
Example 3.5 *Experimental fed-batch phase data using glucose as feed* 62
Example 3.6 *Experimental fed-batch phase data used as validation set for the fed-batch phase simulation model* 66

Example 4.1 *Experimental data using the non-linear compensation and a PRBS as input signals for the glucose flow* 80
Example 4.2 *Estimation of linear models for control of three input-output data sets, GF ↔ OUR, GF ↔ CPR, and GF ↔ DOT* 83
Example 4.3 *Validation of the linear model for control for two sub-models, GF ↔ OUR, and GF ↔ CPR* 86

Example 5.1 *Structure identification using Volterra kernels. The data is from the fed-batch simulation model; the input is GF, an exponentially increasing signal with PRBS excitation* 100

Example 5.2 *Structure identification using Volterra kernels, the data is from the fed-batch simulation model; the input is PRBS only* 102

Example 5.3 *Structure identification using Volterra kernels; data generated using the batch simulation model* 104

Example 5.4 *Structure identification using Volterra kernels; experimental fed-batch data* 107

Example 5.5 *Structure identification using the non-linear model linear-in-the-parameters; data generated by the fed-batch simulation model, delayed output terms included* 108

Example 5.6 *Structure identification using the non-linear model linear-in-the-parameters; data generated with the fed-batch simulation model, only delayed input terms included* 110

Example 5.7 *Structure identification using the non-linear model linear-in-the-parameters; data generated with the fed-batch simulation model, combined growth* 110

Example 5.8 *Structure identification using the non-linear model linear-in-the-parameters; data generated with the fed-batch simulation model, fermentative growth* 111

Example 5.9 *Structure identification using the non-linear model linear-in-the-parameters; data generated with the batch simulation model* 113

Example 5.10 *Structure identification using the non-linear model linear-in-the-parameters; experimental fed-batch data* 113

Example 5.11 *Structure identification using the non-linear model linear-in-the-parameters; experimental batch data* 114

Example 5.12 *Structure identification using RBF functions estimation and validation of GF ↔ OUR using the RBF-network* 117

Example 5.13 *Structure identification using RBF functions; combined growth data* 119

Example 5.14 *Structure identification using RBF functions; fermentative growth* 121

Example 5.15 *Structure identification using Radial Basis Functions; experimental fed-batch data* 123

Example 5.16 *Structure identification using Radial Basis Functions. Experimental batch data* 125

Example 6.1 *Estimation of specific growth rates of a batch experiment of Saccharomyces cerevisiae* 138

Example 6.2 *Estimation of a fed-batch culture of Saccharomyces cerevisiae grown on molasses* 141

Example 7.1 *Tests of compensation 1 and 2* 155

Example 7.2 *Tests of compensation 3 and 4* 157

Example 7.3 *Observer with exclusion of ethanol production for* $\mu < 0.21$ *[h^{-1}]* 158

Example 7.4	*The influence of the observer based numerical errors on the transfer function of the input of the compensation to the μ estimate*	160
Example 7.5	*Stair-case and ramp profiles for the specific growth rate*	161
Example 7.6	*Tuning of the E-controller*	164
Example 7.7	*Tuning high specific growth rates and ethanol concentrations*	165
Example 7.8	*Simulation results of the stirrer speed compensation*	169
Example 7.9	*Simulation results of the DOT-controller*	170
Example 7.10	*Check on a-priori knowledge*	171
Example 7.11	*Air flow and stirrer speed influence*	172
Example 7.12	*PI tuning*	173
Example 7.13	*Normalized glucose feed rate*	175
Example 7.14	*Filtering set-point changes*	177
Example 7.15	*Dynamic compensation*	178
Example 7.16	*Final experiments*	181

List of Figures

Figure 1.1	*Multidisciplinary nature of biotechnology*	1
Figure 1.2	*A simple representation of a yeast cell and its main inputs and outputs*	3
Figure 1.3	*Different levels for controlling Saccharomyces cerevisiae*	4
Figure 1.4	*The concentration of biomass as a function of time during the three phases of laboratory model process*	6
Figure 1.5	*Outline of the thesis by Chapters*	9
Figure 1.6	*A schematic overview of the models used in this thesis, the solid lines between the boxes represent models based on experimental data, the dashed lines between the boxes represent models based on simulation data*	10
Figure 2.1	*A simplified metabolic map involving the glycolysis of sugar to pyruvate and its connection with the ethanol production of the TriCarboxylic Acid (TCA) cycle and connected respiration (after Enfors et al. 1990)*	14
Figure 2.2	*The fermenter*	16
Figure 2.3	*The mass-spectrometer measurement cycle*	19
Figure 2.4	*Schematic overview of the experimental side with the fermenter, the CDAS unit, the μVAX, and the mass-spectrometer*	20

Figure 2.5 *Ethanol measurement disturbed with hourly inserted reference gas* 22
Figure 2.6 *Correlation between cell dry-weight and optical density. Error bars indicate standard deviations.* 26

Figure 3.1 *Bottle-neck principle* 38
Figure 3.2 *Batch and begin of fed-batch phase; big arrows denote possible detection points; the plot is to illustrate the transition, no values are given* 40
Figure 3.3 *A schematic overview of part of the batch phase simulation model* 44
Figure 3.4 *Experimental batch phase results of Saccharomyces cerevisiae and a synthetic medium with* $S_0 = 30\ [g.dm^{-3}]$ 48
Figure 3.5 *Experimental batch phase results of Saccharomyces cerevisiae and a synthetic medium with* $S_0 = 5\ [g.dm^{-3}]$ 51
Figure 3.6 *Experimental batch fermentation results of Saccharomyces cerevisiae in synthetic medium with* $S_0 = 30\ [g.dm^{-3}]$ 56
Figure 3.7 *Experimental batch fermentation results of Saccharomyces cerevisiae in synthetic medium with* $S_0 = 5\ [g.dm^{-3}]$ 58
Figure 3.8 *A schematic overview of part of the fed-batch phase simulation model* 60
Figure 3.9 *Experimental fed-batch phase data of Saccharomyces cerevisiae with synthetic media as feed* 64
Figure 3.10 *Experimental fed-batch phase data of Saccharomyces cerevisiae with synthetic feed, this is a validation set* 67

Figure 4.1 *The Bode plot of the simulation model taken at* $t = 27\ [h]$ *of the validation data presented in Section 3.4.* 76
Figure 4.2 *A schematic overview of the experimenter* 79
Figure 4.3 *Fed-batch fermentation data of Saccharomyces cerevisiae with synthetic feed* 81
Figure 4.4 *Singular values of the estimated Markov parameters* 82
Figure 4.5 *Linearized and normalized experimental data and second order model response of a fed-batch phase of Saccharomyces cerevisiae with synthetic feed* 85
Figure 4.6 *Response of second order models of a fed-batch fermentation run of Saccharomyces cerevisiae with synthetic feed* 87

Figure 5.1 *The Radial Basis Functions network model* 97
Figure 5.2 *First degree representation of the Volterra kernels for OUR and CPR outputs part (a) and for the RQ output part (b)* 101
Figure 5.3 *Second degree 3-dimensional and contour plots of the Volterra kernels for CPR and RQ outputs* 103
Figure 5.4 *Second degree 3-dimensional (a) and contour plots (b) of the Volterra kernels for the GF ↔ RQ relation* 105
Figure 5.5 *Block structure of the Wiener-Hammerstein model* 106

Figure 5.6 *Plot of the calculated RQ and its model response, the error is denoted as a dotted line* 109
Figure 5.7 *Estimation (a) and validation (b) data for the estimated GF $\leftrightarrow$ CPR model for fermentative growth* 112
Figure 5.8 *Estimation (a) and validation (b) plots for the modelled specific growth rate based on OUR and CPR data* 115
Figure 5.9 *Estimation results for OUR (a) and CPR (b) of the RBF-network using the simulation data of Table 5.1 and* $\mu = 0.15$ 118
Figure 5.10 *Estimation, part* (a_1) *and* (b_1)*, and validation, part* (a_2) *and (part* b_2*) of the GF* $\leftrightarrow$ *OUR, part (a) CPR Part (b) models* 121
Figure 5.11 *Estimation data sets for fermentative growth of OUR (a), CPR (b) and RQ (c) models* 122
Figure 5.12 *Estimation (0 - 2.5 [h]) and validation (2.5 - 5 [h]) results for CPR (a) and RQ (b) of the RBF-network using experimental fed-batch data* 124
Figure 5.13 *Estimation results (a) and validation results (b) of the RBF-network identification of the specific growth rate* 126
Figure 5.14 *Estimation results (a) and validation results (b) of the RBF-network identification of the biomass concentration* 127

Figure 6.1 *The relation between* μ *and RQ for all pathways; figures calculated from Eq. (3.1) - Eq. (3.4)* 134
Figure 6.2 *The relation between* μ *and RQ for oxidative growth; data calculated from fed-batch experiments* 135
Figure 6.3 *Growth on glucose (a,b,c) and ethanol (c) represented by the OUR and CPR.* 136
Figure 6.4 *The specific growth rates for batch fermentation. (-) the overall specific growth rate* μ*, (.) oxidative growth on glucose* μ_o*, (--) fermentative growth on glucose* μ_r *and (-.) oxidative growth on ethanol* μ_e 139
Figure 6.5 *Estimated biomass (-) , glucose (--) and ethanol (.) concentrations for a batch fermentation* 140
Figure 6.6 *Estimated specific growth rates, part (a), and estimated biomass concentration, part (b)* 142
Figure 6.7 *Estimated RQ (-) and calculated RQ (*), part (a), and estimated dE/dt (--) and measured dE/dt (-), part (b)* 143

Figure 7.1 *Basic control structure idea* 151
Figure 7.2 *Control structure based on simulation results* 152
Figure 7.3 *Estimated and simulated specific growth rate for constant set-point using compensation 1, part (a), and compensation 2, part (b)* 156
Figure 7.4 *Estimated and simulated specific growth rate for constant set-point using compensation 3, part (a), and compensation 4, part (b)* 157
Figure 7.5 *Original observer, part (a), and modified observer, part (b), controller response* 159

Figure 7.6 *Two variants for a PI-controller structure* 160
Figure 7.7 *Pole-zero plots for μ_d-$\hat{\mu}$ transfer* 161
Figure 7.8 *Performance of the μ-controller to a staircase signal, part (a) and a ramp signal, part (b)* 162
Figure 7.9 *E-controller performance for the critical growth rate; the ethanol response, part (a), and the specific growth rate response, part (b)* 165
Figure 7.10 *E-controller performance for high growth rate; the ethanol response, part (a), and the specific growth rate response, part (b)* 166
Figure 7.11 *Alternative to Figure 7.2* 167
Figure 7.12 *Oxygen controller performance* 170
Figure 7.13 *Estimating the critical growth rate μ_{crit}* 172
Figure 7.14 *Air flow and stirrer speed influence on μ estimates* 172
Figure 7.15 *Too aggressive, part (a) and final PI-tuning, part (b)* 174
Figure 7.16 *Apparent increase in yield* 175
Figure 7.17 *FIR-filter response* 178
Figure 7.18 *Step response to open loop excitation of the GF* 179
Figure 7.19 *Step-responses without and with controller* 180
Figure 7.20 *Response to small steps* 180
Figure 7.21 *Final control structure* 182
Figure 7.22 *Specific growth rate set-point changes* 183
Figure 7.23 *Specific growth rate ramp set-point signal* 183

List of Tables

Table 2.1 *Medium composition for batch and fed-batch* 17
Table 2.2 *Operating conditions during batch and fed-batch phase of important variables* 18
Table 2.3 *Set-point, maximum deviation from the set-point and operation region of temperature and pH for the laboratory process* 22
Table 2.4 *Temperature, pH and anti-foam dependence of the pH and DOT-measurement* 22
Table 2.5 *Set-points, maximum deviation from the set-points, settling times for the set-points and operation regions of the outputs for the three phases of the laboratory process* 23
Table 2.6 *Physical limitations, operation regions and maximum rate of change for the three phases of the laboratory process* 25
Table 2.7 *Some dry-weight and optical density measurements* 27
Table 2.8 *Available actuators and sensors and their specifications and characteristics* 28

Table 3.1 *Stoichiometric constants based on the cell yields* 37
Table 3.2 *Free run settings of the inputs* 41
Table 3.3 *The estimated gains of the staircase experiments* 41
Table 3.4 *Time constants, bandwidth and delay times of the process; all times are in minutes* 42

Table 3.5 *Estimated K_La parameters* 52
Table 3.6 *The estimated parameters of the batch phase simulation model* 53
Table 3.7 *Relative sum of squared errors for both simulation models* 58
Table 3.8 *Initial conditions for the batch phase simulations models* 58
Table 3.9 *The glucose feed rate PRBS-settings* 65
Table 3.10 *The parameters of the fed-batch phase simulation model* 65

Table 4.1 *Poles, zero and K-factors of the SISO transfer functions* 77
Table 4.2 *Model orders found with ARX-structure determination* 83
Table 4.3 *Relative output error of different model order for all outputs for output error (OE) and Box Jenkins (BJ) models. n_C and n_D only for BJ-models* 86

Table 5.1 *Simulation settings for data generated for the non-linear identification algorithms using the fed-batch simulation model* 98
Table 5.2 *Parameters associated with Wiener-Hammerstein model* 106
Table 5.3 *Results of the RBF function identification. GF ↔ OUR model for oxidative growth* 118
Table 5.4 *Results of the experimental fed-batch data* 123
Table 5.5 *Comparison of the non-linear structure identification algorithms* 129

Table 6.1 *The contribution of different growth rates to different pathways* 137

Table 7.1 *A-priori knowledge and the assumptions belonging to compensation 1* 154
Table 7.2 *A-priori knowledge and the assumptions belonging to compensation 2* 155
Table 7.3 *PI-controller settings* 161
Table 7.4 *PI-controller settings ethanol part* 165
Table 7.5 *Initial fed-batch conditions* 170
Table 7.6 *PI-controller settings* 175
Table 7.7 *Model poles and zeros and resulting controller* 179

Samenvatting

De productie van bakkers gist is niet alleen belangrijk voor de bakker. In de industrie worden meer en meer producten gemaakt met behulp van gist. Dit zijn onder meer enzymen, geur- en smaakstoffen en proteïnen. De meet en regel ingenieur moet er onder meer voor zorgen dat de procescondities gunstig zijn zodat de productie van bovengenoemde producten kan plaatsvinden. Het promotie onderzoek omvat modelleren, identificatie en regeling van de produktie van *Saccharomyces cerevisiae* (bakkers gist).

Voordat we kunnen modelleren, identificeren of regelen dienen we het proces zelf grondig te bestuderen. Het proces is zowel niet-lineair als tijd-variant. Er is onderzoek gedaan naar proces-, opnemer- en regelaar-karakteristieken. Dit is onder meer gedaan om de limitaties van het proces voor identificatie en regeling te weten te komen.

Zowel de "batch" als de "fed-batch" groei van *Saccaromyces cerevisiae* zijn gemodelleerd en de modellen zijn gevalideerd met data uit laboratorium experimenten. Dit zijn modellen, op fysische en fysiologische grondslag. De modellen zijn gebruikt als simulatie modellen om identificatie- en regelstrategieën te testen voor gebruik op laboratorium schaal.

Het doel van dit werk is het regelen van de groei van *Saccharomyces cerevisiae*. Om het proces te kunnen regelen dienen er regel modellen gemaakt te worden. Rond een werkpunt van het proces is een lineair model gemaakt door toepassing van niet-lineaire compensatie van bekende niet lineaire verschijnselen. Dit is gedaan met beproefde identificatietechnieken. De manier van toepassen van deze technieken is nieuw in de biotechnologie.

Verder is een aantal niet lineaire identificatietechnieken (onder meer: "Volterra kernels", "Radial Basis Functions" en niet lineaire modellen lineair- in-de-parameters) toegepast op het proces en hieruit is een aantal modellen afgeleid. Dit zijn voorspellende modellen; als regelaarmodellen voor dit proces zijn ze minder zinvol.

Om regelstrategieën te kunnen toepassen is het wenselijk dat niet meetbare proces-parameters, zoals gistconcentratie en groeisnelheid, gemodelleerd worden. Tot dusver is het niet mogelijk gebleken de specifieke groeisnelheid van de gist te modelleren. De specifieke groeisnelheid beschrijft de groei van de gist en is dus een belangrijke proces variabele. Door middel van een "observer" kan er een schatting van de specifieke groeisnelheid gemaakt worden. Dit gebeurt met een eenvoudige, maar doeltreffende strategie, gebruik makend van gemeten procesdata.

Met behulp van deze schatter is een regelaar voor de oxidatieve en de fermentatieve groei van *Saccharomyces cerevisiae* ontworpen. De regelaar is op laboratorium schaal getest en er zijn goede resultaten voor de oxidatieve groei behaald. Tijdens het gebruik van deze regelaar zijn er nieuwe, onverwachte dynamische effecten opgetreden.

Curriculum Vitae

12 january 1962	Geboren te Voerendaal.
mei 1981	Gymnasium β aan het Bernardinus college te Heerlen.
sept 1981 - april 1989	Studie elektrotechniek (oude stijl) aan de Technische Universiteit Eindhoven. Afstudeerrichting Meet en Regeltechniek. Afstudeerwerk: *Modelling and Selftuning Control Application for Fed-Batch Bakers' Yeast Fermentation,* onder leiding van Prof. dr. ir. P. Eykhoff en dr. G. Reyman (Unilever).
januari 1989 - maart 1989	Buitenlands verblijf aan de Techniche Universiteit van Tallinn, Estland. Werk verricht onder leiding van dr. M. Min.
juni 1989 - september 1993	Promotie werk onder leiding van Prof. dr. ir. P. Eykhoff, dr. ir. A.J.W. van den Boom, dr. ir. M.L.F. Giuseppin (Unilever), dr. G. Reyman (Unilever, tot 1992) en dr. ir. R.S.A. Soeterboek (Unilever, vanaf 1992).

STELLINGEN

behorende bij het proefschrift van

Marc Keulers

Eindhoven, 2 september 1993

1
'White-box' gebaseerde regelaars voor het regelen van biotechnologische processen voldoen beter dan 'black-box' gebaseerde regelaars.

2
Het model van Dantigny *et al.* (1992) pretendeert een 'fed-batch' model te zijn. Daar het model gebaseerd is op data van 'batch' en 'continuous' fermentaties kan het model niet gelden voor 'fed-batch' fermentatie.

> **Dantigny, P., Ziouras, K. and J.A. Howell** (1992). A Structured Model of Bakers' Yeast Fed-Batch Growth. *IFAC-Symposium on Modelling and Control of Biotechnical Processes*, Keystone, Colorado, USA, pp. 223-226.

3
De meet- en regeltechniek is ten onrechte ondergewaardeerd in de biotechnologie; methodieken uit de meet- en regeltechniek zoals het ontwerpen van lineaire modellen en op kennis gebaseerde waarnemers van biotechnologische processen kunnen het inzicht van biotechnologen in deze processen verhogen.

> Hoofdstuk 1,4,6 en figuur 1.1 van dit proefschrift.

4
Bij de maximalisatie van de produktie van gist dient men uit te gaan van een regeling gebaseerd op metingen van de ethanol concentratie in de fermentor en niet een die gebaseerd is op afgeleide metingen zoals de 'Respiratory Quotient'.

6
De produktie van gist kan gezien worden als een 'dual control system' (Fel'dbaum, 1960,1961;1965; en Eykhoff, 1974). Dientengevolge zullen adaptieve regelstrategieën niet voldoen.

Fel'dbaum, A.A. (1960,1961). Dual Control Theory. *Automn. Remote Control*, Vol. 21, pp.874-880, 1033-1039; Vol. 22, 1-12,109-121.
Fel'dbaum, A.A. (1965). *Optimal Control Systems*, Academic Press, New York, 452 pp.
Eykhoff, P. (1974). *System Identification*, John Wiley and Sons, New York, pp.13-14.

6
Het gebruik van een gemodificeerde "hopscotch"-methode voor het simuleren van de beweging van superfluïde wervels is af te raden, aangezien het ongelijk behandelen van de punten op de wervel oscillaties veroorzaakt.

Tsubota, M. and S. Maekawa (1992). Reconnection of Quantized Vortices in Superfluid ^{4}He. *Journal of The Physical Society of Japan*, Vol. 61, No. 6, pp. 2007-2017.
Aarts, R.G.K.M. (1993). A Numerical Study of Quantized Vortices in He II. *Ph.D.-Thesis*, Eindhoven University of Technology, pp. 48-51.

7
Verdere automatisering en verhoging van de toegevoegde waarde bij de produktie van goederen in hoog-ontwikkelde landen bevordert de rassendiscriminatie in het land zelf.

8
In tijden van economische recessie neemt de correlatie tussen inkomen en discriminatie toe.

9
De Nederlandse taal wordt als cultuurgoed door Nederlanders dermate ondergewaardeerd dat buitenlandse werknemers in Nederland die geen Nederlands of Engels spreken eerder Engels leren dan Nederlands. Het verdient dan ook aanbeveling deze buitenlandse werknemers enige tijd te laten samenwerken met Nederlanders die alleen Nederlands willen spreken.

10
Een grote stroom drugstoeristen doet Rotterdam aan, al of niet geholpen door drugskoeriers. Om beter op dit 'toerisme' in te kunnen spelen en om de lokale 'drugsscene' beter te kunnen beheersen zou de stad een 'high'-centrum moeten bouwen. De rest van de stad zou dan vervolgens 'drugsvrij' verklaard moeten worden. Dit zal niet alleen een positieve uitwerking hebben op het 'normale' toerisme maar ook de leefbaarheid in de stad zelf bevorderen.

11
Het meenemen van Nederlands voedsel door Nederlanders naar een buitenlandse vakantiebestemming getuigt niet alleen van een 'kruideniers' mentaliteit, maar is ook discriminerend ten aanzien van de buitenlandse keuken.

Van Kooten en De Bie, Keek op de Week.

12
Het kubisme als artistieke stroming wordt vaak ten onrechte geassocieerd met de schilderkunst (Picasso en Braque); zij vindt haar ontstaan echter in de architectuur van Frank Lloyd Wright van rond 1900.

Kunstschrift, (1993), No. 3, pp. 10-11.

13
Het leven is niet-lineair. Dientengevolge zou ieder boek dat handelt over oplosmethodieken voor 'lineaire, tijd-invariante systemen' de waarschuwing moeten bevatten: "Oplosmethodieken voor lineaire, tijd-invariante systemen kunnen stress (hoofdpijn, maagpijn, etc.) veroorzaken".