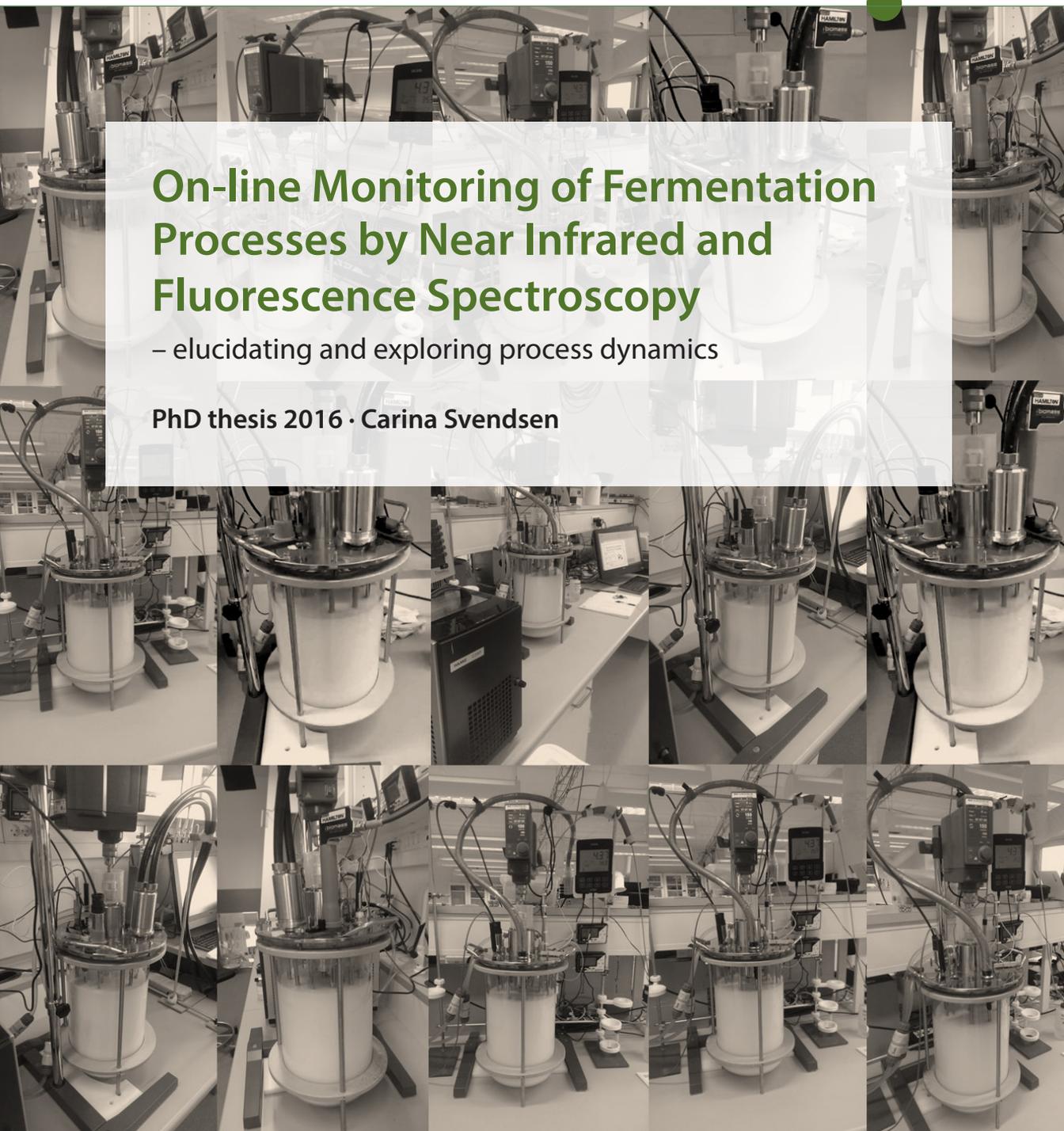




On-line Monitoring of Fermentation Processes by Near Infrared and Fluorescence Spectroscopy

– elucidating and exploring process dynamics

PhD thesis 2016 · Carina Svendsen



On-line Monitoring of Fermentation Processes by Near Infrared and Fluorescence Spectroscopy

-elucidating and exploring process dynamics

PHD THESIS · 2016

Carina Svendsen

**University of Copenhagen · Faculty of Science
Department of Food Science (FOOD)
Chemometrics and Analytical Technology section (CAT)
Rolighedsvej 26 · 1958 Frederiksberg C · Denmark**

Title:

On-line Monitoring of Fermentation Processes by Near Infrared and Fluorescence Spectroscopy
-elucidating and exploring process dynamics

Submission:

31st of May, 2016

The PhD School of The Faculty of Science, University of Copenhagen

Defence:

9th of September, 2016

Frederiksberg Campus, The Faculty of Science, University of Copenhagen

Supervisors:

- Frans van den Berg, Associate Professor
- Thomas Skov, Associate Professor
Chemometrics and Analytical Technology (CAT), Department of Food Science
Faculty of Science, University of Copenhagen, Denmark

Opponents:

- Marina Cocchi, Associate Professor
Analytical Chemistry, Department of Chemical and Geological Science
University of Modena and Reggio emilia, Italy
- Rodrigo Bibiloni, Senior Team Manager
Cultures & Enzymes Team
Arla Foods Amba, Denmark
- Åsmund Rinnan, Associate Professor
Chemometrics and Analytical Technology, Department of Food Science
Faculty of Science, University of Copenhagen, Denmark

Cover illustration:

"A lactic acid fermentation" by Carina Svendsen

PhD thesis 2016 © Carina Svendsen. All rights reserved.

ISBN 978-87-93476-44-8

Printed by SL grafik, Frederiksberg, Denmark (www.slgrafik.dk)

“Without data you’re just another person with an opinion”

W. Edwards Deming

PREFACE

This PhD thesis, titled *On-line Monitoring of Fermentation Processes by Near Infrared and Fluorescence Spectroscopy -elucidating and exploring process dynamics*, has been carried out at the Chemometrics and Analytical Technology section (CAT), Department of Food Science, Faculty of Science, University of Copenhagen, under supervision by Associate Professor Frans van den Berg and Associate Professor Thomas Skov. The project was partly sponsored by BIOPRO (www.biopro.nu), which is founded by Region Zealand, The European Regional Development Fund, The Innovation Foundation and Partners. I am grateful to my supervisors, Frans and Thomas, CAT and BIOPRO for giving me the opportunity to work within the field of fermentation monitoring for the past three years.

A special thanks to Frans, who has always been supporting, motivating and enthusiastic in talks and discussions on experimental set-up, process monitoring and advanced modelling of fermentation processes. He is a truly dedicated expert within PAT and process monitoring. I have learned a lot from Frans and he has given me advice and support, even when I was located in Lisbon or in Brazil, which I appreciate very much.

During my PhD studies I had a four months research stay at the BioEngineering Research Group, Institute for Biotechnology and BioSciences, Instituto Superior Técnico in Lisbon, Portugal hosted by Professor José Cordoso Menezes. Thanks for hosting me and thanks to the 4-Tune team and especially to Francisca and Maria for enriching my stay.

Although Frans has been the main support throughout this project, my colleagues at CAT have always shown their willingness to discuss topics within spectroscopy, chemometrics, MATLAB or project/article writing in general. That is why CAT is a unique research group to work in and that is why I am very thankful to you guys working at CAT.

Last but not least, I would like to thank the ones I care most about. Thanks for listening to my frustrations about molecules vibrating in a certain manner, red MATLAB errors and yoghurt bacteria. You have always supported me and believed I could finalise whatever I decided to start. Tak Jens Erik, mor, far og Ronni.

*Carina Svendsen
Frederiksberg C, May 2016*

ABSTRACT

Monitoring and control of fermentation processes is important to ensure high product yield, product quality and product consistency. More knowledge on on-line analytical techniques such as near infrared and fluorescence spectroscopy is desired in the fermentation industry to increase the efficiency of on-line monitoring systems.

The primary aim of this thesis is to elucidate and explore the dynamics in fermentation processes by spectroscopy. Though a number of successful on-line lab-scale monitoring systems have been reported, it seems that several challenges are still met, which limits the number of full-scale systems implemented in industrial fermentation processes. This thesis seeks to achieve a better understanding of the techniques near infrared and fluorescence spectroscopy and thereby to solve some of the challenges that are encountered.

The thesis shows the advantages of applying real-time monitoring of bioprocesses and it also highlights that the applied techniques with different measurement orders deliver specific but also complementary sources of information. Furthermore, it was shown that valuable process information can be obtained both by near infrared spectroscopy and fluorescence spectroscopy, which provide indirect and direct measurements, respectively.

Based on the measurements obtained by near infrared spectroscopy it was found that variation in scatter and in the absorption can be obtained from the same near infrared spectrum. By kinetic modelling, it was possible to capture both physical and chemical changes appearing in a lactic fermentation process. The physical changes were associated with the textural transformation appearing during the gel formation and chemical changes were associated with the biological conversion reactions, which take place during the fermentation process.

The results presented in this thesis also highlight that pH changes have a major effect on the fluorescence intensities, which can influence the quantifications of the relevant components negatively. When the pH was either increased or decreased, manually, during the measured process, a clear increase or decrease was observed in the fluorescence landscapes. This thesis presents a correction strategy based on a chemometric modelling approach, where weighted non-linear regression and weighted PARAFAC analysis are combined.

Based on the research conducted in this PhD project, it is concluded that near infrared spectroscopy can provide valuable physical and chemical real-time information during yoghurt fermentation. Also, it is concluded that fluorescence data must be evaluated carefully if pH changes appear in the measured system. However, such data can still be applied for on-line monitoring if corrections or preventive measures during the quantification are carried out. The findings presented in this thesis have enabled the possibility of obtaining a better process understanding and to ease monitoring and controlling of fermentation processes.

RESUMÉ

Monitorering og kontrol af fermenteringsprocesser er vigtigt for at sikre højt produktudbytte, produktkvalitet og ensartethed. Mere viden om on-line teknikker, såsom nærinfrarød spektroskopi og fluorescensspektroskopi, er på verdensplan ønsket af fermenteringsindustrien for at øge effektiviteten og forståelsen af on-line monitoreringssystemer.

Det primære formål med denne afhandling er at belyse og udforske dynamikken i fermenteringsprocesser ved brug af spektroskopi. På trods af, at flere publicerede studier har præsenteret succesfulde on-line laboratoriemonitorerings-systemer, viser det begrænsede antal af implementerede monitoreringssystemer i industrielle fuldskalafermenteringsprocesser, at der stadig eksisterer udfordringer. I denne afhandling søges at opnå en bedre forståelse af teknikkerne nærinfrarød- og fluorescensspektroskopi og derved at reducere udvalgte udfordringer, der findes ved on-line monitorering af fermenteringsprocesser.

Afhandlingen omhandler og identificerer fordelene ved at benytte real-tidsmonitorering af bioprocesser. Det fremhæves også, at de benyttede teknikker, med varierende data-dimensioner, leverer forskellige niveauer af information, der både er specifikke såvel som komplementerende. Ydermere, vises det, at værdifulde procesinformationer bliver opnået både via nærinfrarød og fluorescensspektroskopi, som giver henholdsvis indirekte og direkte målinger.

Baseret på nærinfrarød spektroskopi data bliver det vist, at variation i *scatter* og i absorption kan findes i det samme nærinfrarøde spektrum. Via kinetik modellering er det muligt at modellere både de fysiske og kemiske ændringer, som optræder i en mælkesyrefermentering. De fysiske ændringer er associeret med de teksturændringer, som finder sted under geldannelsen, mens de kemiske ændringer er relateret til de biologiske omdannelsesreaktioner, som ligeledes finder sted under fermenteringsprocesser.

De resultater, der er præsenteret i denne afhandling bekræfter også, at pH ændringer har en stor indflydelse på fluorescensintensiteten, hvilket kan påvirke kvantificeringen af relevante komponenter negativt. Når pH enten øges eller sænkes manuelt, under den målte proces, bliver en tydelig intensitet-stigning eller -fald observeret i fluorescenslandskaberne. Denne afhandling præsenterer de væsentlige ændringer observeret i fluorescenslandskaberne og præsenterer en korrektionsstrategi baseret på en kemometrisk modelleringstilgang, hvor metoder-

ne "vægtet ikke-linenær regression" og "vægtet PARAFAC analyse" er kombineret.

Baseret på forskningen udført i denne ph.d.-afhandling, konkluderes det, at nærinfrarød spektroskopi kan give værdifuld fysisk og kemisk real-tidsinformation under en yoghurtfermentering. Det konkluderes også, at fluorescens data bør evalueres med forsigtighed, hvis pH ændringer finder sted i den målte proces. Dog kan sådanne data stadig benyttes til on-line monitorering, hvis korrektioner eller forebyggende databehandling under kvantificeringen bliver udført. De opnåede resultater præsenteret i denne afhandling gør det muligt at opnå en bedre procesforståelse og gøre monitorering og kontrol af fermenteringsprocesser lettere.

LIST OF PUBLICATIONS

➤ **PAPER I**

Carina Svendsen, Thomas Skov and Frans W. van den Berg

Monitoring fermentation processes using in-process measurements of different orders.

Journal of Chemical Technology and Biotechnology, 90 (2015), 244-254.

➤ **PAPER II**

Carina Svendsen, Tomasz Cieplak and Frans W. van den Berg

Exploring process dynamics by near infrared spectroscopy in lactic fermentations.

Journal of Near Infrared Spectroscopy (2016), *article in press*

➤ **PAPER III**

Carina Svendsen, Thomas Skov and Frans W. van den Berg

Weighted PARAFAC and non-linear regression for handling intensity changes in fluorescence spectroscopy caused by pH fluctuations.

Applied Spectroscopy (2016), *article published online*

POSTERS

➤ **POSTER I**

Carina Svendsen, Frans W. van den Berg and Thomas Skov

Implementation of On-line Monitoring of Cells in Fermentation Processes.

Scandinavian Symposium on Chemometrics (SSC13), Sweden, 2013

➤ **POSTER II**

Carina Svendsen, Thomas Skov and Frans W. van den Berg

A chemometric approach for correction of intensity changes in fluorescence spectroscopy caused by pH fluctuations.

Scandinavian Symposium on Chemometrics (SSC14), Italy, 2015

OTHER PUBLICATIONS BY THE AUTHOR

- Lise Søndergaard, Mia Ryssel, Carina Svendsen, Erik Høier, Ulf Andersen, Marianne Hammershøj, Nils Arneborg, Lene Jespersen
Impact of NaCl reduction in Danish semi-hard Samsøe cheeses on development and autolysis of DL-starter cultures
International Journal of Food Microbiology, 213 (2015), 59-70.
- Xiangqian Zhu, Carina Svendsen, Kristina B. Jaepelt, Paul J. Moughan, Shane M. Rutherford
A comparison of selected methods for determining eicosapentaenoic acid and docosahexaenoic acid in cereal-based foods
Food Chemistry, 125 (2011), 1320-1327.
- Martin Thorup Nielsen, Carina Svendsen, Line Thorsen, Mogens Jakobsen
Wet heat treatment of *Cronobacter sakazakii* and detection of viable cells using RT-PCR and Propidium monoazide for distinction between dead and viable cells
International ICFMH Symposium Food Micro 2010

LIST OF ABBREVIATIONS

ATR	Attenuated Total Reflectance
CCP	Critical Control Point
CPP	Critical Process Parameter
CQA	Critical Quality Attributes
CER	Carbon dioxide Evolution Rate
EMSC	Extended Multiplicative Signal Correction
FDA	Food and Drug Administration
FT	Fourier Transformed
GMP	Good Manufacturing Practice
HACCP	Hazard Analysis and Critical Control Point
IR	InfraRed
LAB	Lactic Acid Bacteria
<i>L. Bulgaricus</i>	<i>Lactobacillus delbrueckii</i> subsp. <i>Bulgaricus</i>
MSC	Multiplicative Scatter Correction
NLR	Non-Linear Regression
NIR	Near InfraRed
NIRS	Near InfraRed Spectroscopy
OD	Optical Density
OD ₆₀₀	Optical Density (at wavelength 600 nm)
OUR	Oxygen Uptake Rate
PARAFAC	PARAllel FACtor Analysis
PAT	Process Analytical Technology
PC	Principal Component
PCA	Principal Component Analysis
pO ₂	Partial pressure of dissolved oxygen
pCO ₂	Partial pressure of dissolved carbon dioxide
SNV	Standard Normal Variate
<i>S. thermophilus</i>	<i>Streptococcus thermophilus</i>
SG	Savitzky-Golay derivation

TABLE OF CONTENTS

Preface.....	i
Abstract	iii
Resumé	v
List of Publications.....	vii
List of Abbreviations	ix
Table of Contents	xi
1. Introduction.....	1
1.1 Background	1
1.2 Aims	3
1.3 Outline of the Thesis.....	3
2. The Lactic Fermentation Process.....	5
2.1 Introduction	5
2.2 Characteristics of the applied Starter Bacteria	5
2.3 The Metabolisms and Growth Associations of LAB in Yoghurt Fermentation.....	6
2.4 Industrial Yoghurt Manufacture	9
3. On-line Monitoring of Fermentation Processes	13
3.1 Process Analytical Technology.....	13
3.2 On-line State-of-the-art Methods for Fermentation Monitoring.....	14
3.3 Study I – Exploratory Study on Fermentation Monitoring using Measurements Techniques of different Orders.....	26

4. Near Infrared Spectroscopy for On-line Monitoring of Fermentation Processes.....	31
4.1 Introduction	31
4.2 Fundamental of Near Infrared Spectroscopy	31
4.3 Sampling Modes	34
4.4 Data Processing for NIRS Data	37
4.5 Chemical and Physical Information	44
4.6 Study II – Exploring Process Dynamics by NIR Spectroscopy in Lactic Fermentations	47
5. Fluorescence Spectroscopy for On-line Monitoring of Fermentation Processes	53
5.1 Introduction	53
5.2 Fundamental of Fluorescence Spectroscopy	53
5.3 Factors influencing the fluorescence intensity signal	58
5.4 Data Processing for Fluorescence Data	62
5.5 Study III – Evaluation of on-line Fluorescence Spectroscopy applied in Fermentation Processes	64
6. Conclusions	71
7. Perspectives	73
8. Reference List	77

Paper I

Paper II

Paper III

Poster I

Poster II

CHAPTER 1

INTRODUCTION

1.1 Background

Monitoring and controlling processes is of great importance in all industries. This statement also includes the food- and pharma bioprocess industries, where real-time monitoring techniques in combination with chemometrics have been widely explored during the past decades (Lourenco, et al., 2012). Bioprocesses use living microorganisms to produce a broad range of products such as metabolites, cells, proteins, hormones, enzymes or e.g. fermented food products. To ensure high product yield, product quality and product consistency, the main key is to monitor and control the metabolism(s) of the applied microorganism(s). In order to monitor such a process successfully, measurement techniques describing physical and/or chemical process variables qualitatively and/or quantitatively must be applied. The obtained measurements must be modelled in a manner that makes sense, thus process variation of interest can be gained, process understanding can be obtained, and successful process control can be achieved (Sonnleitner, 2013).

At many production sites nowadays, simple real-time monitoring systems are already implemented. These systems rely on measurements providing a limited number of physical and chemical variables such as pH, temperature and headspace gas composition, typically O₂ and CO₂. These few chosen variables provide important and valuable real-time information about the bioprocess, which in many cases can help us recover a batch from being under abnormal operation conditions to become stable and back under normal conditions. As an example, pH is a direct measure for acidification in a lactic fermentation, which is linked to the metabolism and growth of the cells (De Brabandere and De Baerdemaeker, 1999). Likewise, the O₂ uptake and CO₂ production are closely related to the cell growth and metabolism of e.g. yeast cells under aerobic conditions (Furukawa, et al., 1983). Fermentation is however a complex process and additional chemical and physical information would likely improve the monitoring system and allows for even better process understanding and optimisation possibilities as well as improved process control (Figure 1.1). A better monitoring capability would also increase the product safety in terms of microbiological risks, chemical contamination and possible traceability during the production. The benefits are thus not only there for the industries, but also for mankind, in which an advanced real-time approach allows process optimization and thereby reduction of resource consumption leading to e.g. energy saving and a limited carbon footprint, which are

key parameters for an increased sustainability. There is no doubt that the motivation for additional real-time monitoring is high and therefore more research within real-time monitoring techniques and data acquisition has been devoted over the last decades (Pohlscheidt, et al., 2013).

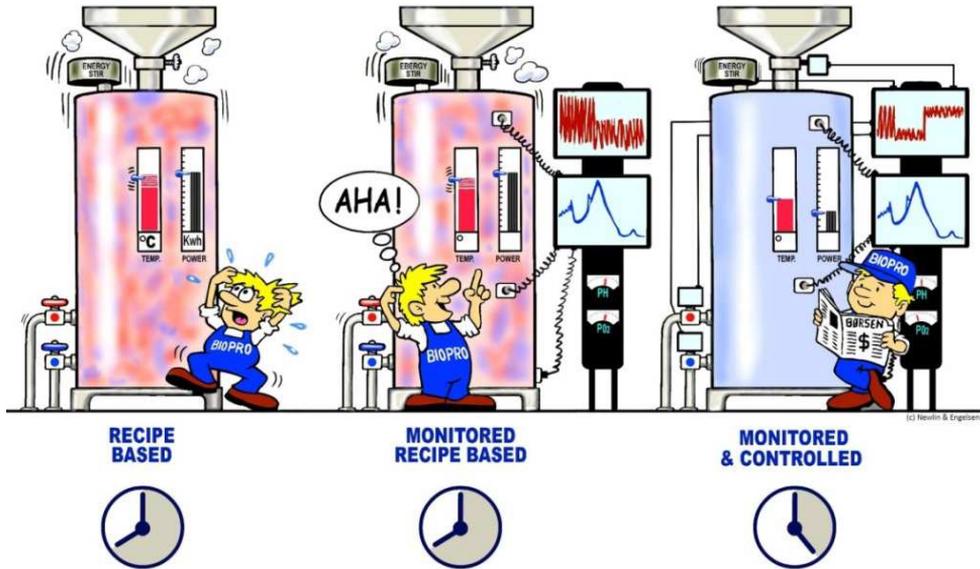


Figure 1.1: Illustration of non-monitored versus monitored and controlled bioreactors by ©Newlin & Engelsen

The Food and Drug Administration (FDA) has encouraged the food- and pharma industry to use Process Analytical Technology (PAT) in order to increase process understanding as well as process optimization and consistent product quality by introducing the PAT guidance in 2004 (FDA, 2004). Since then, an increasing focus has been on various analytical techniques including on-line sensors such as rapid and non-destructive spectroscopic methods like near infrared spectroscopy and fluorescence spectroscopy. Even though several research studies on fermentation processes (Faassen and Hitzmann, 2015; Lopes, et al., 2015) confirm that spectroscopic methods have provided chemical and/or physical information successfully, several challenges are still met when implementing those techniques at a full scale production site. Hence, a need for supplementary understanding is still needed before the step from lab-scale studies to full scale implementation can become reality.

In this thesis fermentation processes are monitored by on-line spectroscopy in combination with simple sensors such as pH and temperature. The thesis seeks to further elucidate the advantages of near infrared and fluorescence spectroscopy for fermentation monitoring. Furthermore, some disadvantages of the methods are clarified and solutions are suggested for better process understanding. Hereby, I believe that the gap between lab-scale monitoring and

full-scale monitoring implementation will decrease to some extent.

1.2 Aims

The primary aim of this thesis is to elucidate and explore the dynamics in fermentation processes by near infrared spectroscopy and fluorescence spectroscopy. The thesis specifically seeks to address the following aims:

- *Investigate and compare a number of suitable on-line techniques providing different data structures and process information for fermentation processes.*
- *Explore the possibility of obtaining chemical and physical information of a lactic fermentation process by near infrared spectroscopy.*
- *Elucidate fluorescence spectroscopy and its stability towards environmental disturbances and explore the possibilities for correcting these interferences.*

New insights on the already well-studied near infrared and fluorescence spectroscopy methodologies, applied in bioprocesses, will allow us to acquire an improved toolset for an on-line monitoring and control systems, which will again provide us with new knowledge on fermentation processes. Hereby, the process understanding can potentially be improved, and in turn allow for better optimization and control. The research conducted in this thesis is based on lab-scale experiments (albeit with relatively large volumes) using two different model systems. The first model system consists of a lactic fermentation process in which various monitoring techniques providing different data orders are applied and explored. An additional study, carried out on the same model system, focussing on near infrared spectroscopy and its capability to provide dynamic information about the physical changes in the lactic fermentation process. A second model system is built to explore the performance of fluorescence spectroscopy when it is experiencing environmental disturbances. It is based on a simple light induced riboflavin degrading process. Different chemometrics modelling methods are applied and considered depending on the applied on-line technique. Furthermore, a new correction strategy is suggested for disturbances observed in the fluorescence landscapes. In all investigations, I aim to obtain a better process understanding and to ease monitoring and controlling of fermentation processes.

1.3 Outline of the Thesis

This thesis is written based on three articles either published in or accepted for publication by peer-reviewed journals. The main focus is on near infrared spectroscopic and fluorescence

spectroscopy and their ability, in combination with the right chemometric methods, to provide process-relevant information about the fermentation process. A brief overview of the chapters included in the thesis follows.

Two of the written articles (Paper I and Paper II) are based on a lactic fermentation process and therefore Chapter 2 provides a description of the employed lactic acid bacteria cultures and their metabolisms. Since the end product of this lactic fermentation is yoghurt, an overview of the industrial yoghurt manufacture is also presented.

Chapter 3 gives an overview of the state-of-the-art monitoring techniques applied in fermentation processes. First, classical methods are presented and then a number of spectroscopic methods are described. The main focus is on near infrared and fluorescence spectroscopy. The chapter ends with an introduction to Paper I and a short evaluation of the applied techniques.

In Chapter 4, near infrared spectroscopy in relation to monitoring of lactic fermentations is in focus. The fundamental of the method is first described and hereafter different sampling modes and data treatments are discussed. Additionally, chemical versus physical information obtained by near infrared spectroscopy is elucidated. Finally, Paper II is shortly introduced and discussed.

The focus in Chapter 5 is on fluorescence spectroscopy, where the fundamentals and factors influencing the method are described. Hereafter, data processing for fluorescence spectroscopy is presented and Paper III is briefly introduced and discussed.

Finally, in Chapter 6, the conclusions of the conducted research are summarized and in Chapter 7 perspectives and ideas for further research are presented.

The thesis is complemented by the manuscripts, appended after the chapters.

CHAPTER 2

THE LACTIC FERMENTATION PROCESS

2.1 Introduction

Lactic Acid Bacteria (LAB) are the prime agents in producing fermented milks and dairy products via the lactic fermentation process (Walstra, et al., 2005b). The group of LAB causes rapid acidification of the raw material through the production of lactic acid, and they also contribute to the preservation as well as the flavours, texture and probiotic properties of dairy products such as sour milk, sour cream and cheese (Leroy and De Vuyst, 2004). In the production of the fermented milk product yoghurt, the lactic fermentation is carried out with the starter bacteria *Streptococcus thermophilus* (*S. thermophilus*) and *Lactobacillus delbrueckii* subsp. *bulgaricus* (*L. bulgaricus*). Since the model system applied for Studies I and II is based on a simplified yoghurt production, the present chapter focuses on the lactic acid fermentation where yoghurt is the end product.

2.2 Characteristics of the applied Starter Bacteria

Characteristics of Streptococcus thermophilus

The microorganism *S. thermophilus* is a Gram-positive, spherical-shaped (Figure 2.1) and thermophile bacteria. Further characteristics of the species *S. thermophilus* include that it is anaerobic, has a growth optimum near 45°C and it is mainly used for the manufacture of hard cheeses and yoghurt. It grows well in milk where it ferments the milk sugar lactose (Harnett, et al., 2011). The organism requires free amino acids, including glutamic acid, histidine, methionine, cysteine, valine, leucine, isoleucine, tryptophan, arginine and tyrosine. The essential amino acids vary a bit from strain to strain, but all strains grow well on media containing hydrolysed proteins (Harnett, et al., 2011).

Characteristics of Lactobacillus delbrueckii subsp. bulgaricus

The microorganism *L. bulgaricus* is also Gram-positive and thermophile, but compared to the spherical-shaped *S. thermophilus*, *L. bulgaricus* is rod-formed as shown in Figure 2.1. Furthermore, this bacteria is anaerobic homo-fermentative and has a growth optimum around 43°C. Apart from being utilized in yoghurt production, it is used in curd preparation for Swiss-type cheeses and hard Italian-type cheeses (Rizzello and De Angelis, 2011). Their growth is in general improved when CO₂ and bicarbonate are available (Gürakan and Altay, 2010).

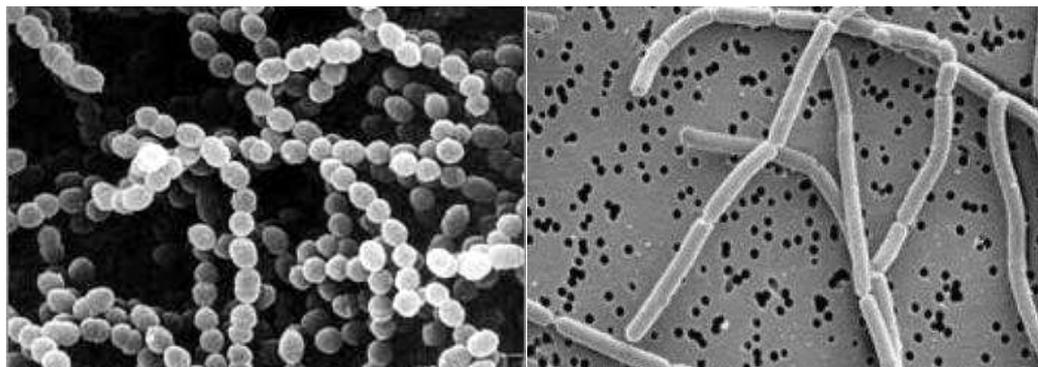


Figure 2.1: The morphology of the spherical-shaped *S. thermophilus* (left) and the rod-formed *L. bulgaricus* (right). Picture left: (Vivatfor, 2016), right: (Kahl, 2016).

2.3 The Metabolisms and Growth Associations of LAB in Yoghurt Fermentation

The fermented milk product yoghurt is an acid gel produced by the fermentation of milk with a bacterial culture containing the lactic starter bacteria *S. thermophilus* and *L. bulgaricus*. Before the fermentation process is started, the milk is pasteurized and other optional ingredients, such as flavours, vitamins, sweeteners, stabilizers and colour additives are added. Energy and nitrogen is required by the yoghurt starter bacteria to maintain their life cycle and the main energy source for growth of *S. thermophilus* and *L. bulgaricus* in milk is the carbon source of the disaccharide sugar lactose (Walstra, et al., 2005a).

Carbohydrate metabolism

The main energy source for the growth of lactic acid bacteria in milk is the carbohydrate lactose, which is metabolised via the homo- or the hetero-fermentative metabolic pathway, depending on the specific starter bacteria (Tamime and Robinson, 2007c). *S. thermophilus* and *L. bulgaricus* both ferment lactose homo-fermentative and they both transport lactose into the cell without any chemical modification via the cytoplasmic protein proteases. In the cell, lactose is hydrolysed to glucose and galactose by the enzyme β -galactosidase. Only the glucose moiety of lactose is utilized, whereas galactose is secreted from the cells and left in the growth medium (Harnett, et al., 2011). However, if the glucose level gets low, galactose can be metabolised by *S. thermophilus* (Hutkins and Morris, 1987). Glucose is catabolised via the glycolysis, also known as the Embden-Meyerhof pathway, where pyruvate is the end product (Nauth, 2004). The pyruvate is then synthesised to lactate by the enzyme lactate dehydrogenase, which lactic acid bacteria possess (Tamime and Robinson, 2007c). The described metabolism is illustrated in Figure 2.2. The homo-fermentative pathway yields two moles of lactic acid and two moles of ATP per mole glucose consumed (Von Wright and Axelsson, 2012).

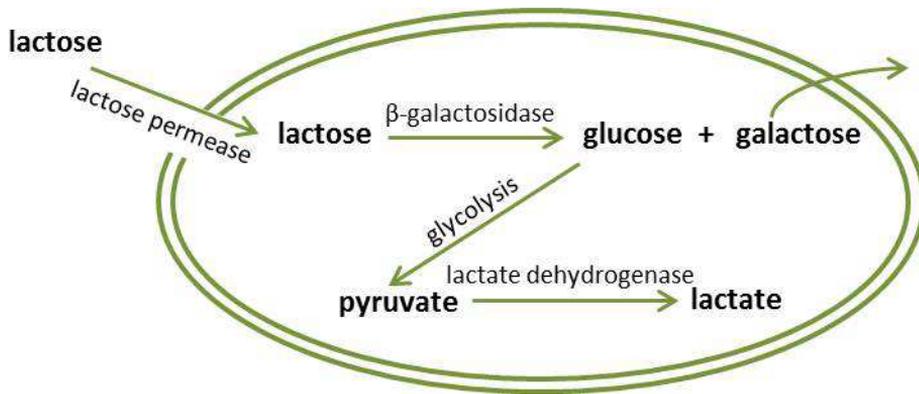


Figure 2.2: Metabolic pathway of lactose (intracellular). Lactose is transferred into the cell and converted to glucose, which is further metabolised into pyruvate via the glycolysis. Pyruvate is finally converted into lactate/lactic acid.

Two different isomers of lactic acid (Figure 2.3) are formed during the fermentation, where *S. thermophilus* produces L(+) lactic acid and *L. bulgaricus* produces D(-) lactic acid (Tamime and Robinson, 2007c).

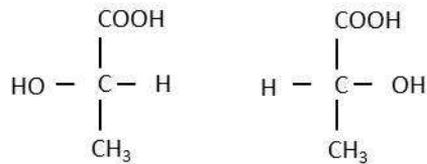


Figure 2.3: The two isomers of lactic acid L(+) lactic acid (left) and D(-) lactic acid (right).

Growth associations between S. thermophilus and L. bulgaricus in yoghurt production

In mixed starter culture with *S. thermophilus* and *L. bulgaricus* a symbiotic growth relationship exists between the two bacteria. Carbon dioxide, formate, peptides and several amino acids released from casein are among the compounds that are involved in this process (Nauth, 2004). Since the proteolytic system of *S. thermophilus* is more limited compared to most other dairy starter bacteria, it is an advantage to pair *S. thermophilus* with another more proteolytic starter bacterium such as *L. bulgaricus* (Harnett, et al., 2011). Hence, *L. bulgaricus* possesses the enzyme protease that releases amino acids from casein and these amino acids stimulate the growth of *S. thermophilus*. In turn, *S. thermophilus* produces carbon dioxide and formate, which stimulates the growth of *L. bulgaricus* (Figure 2.4).

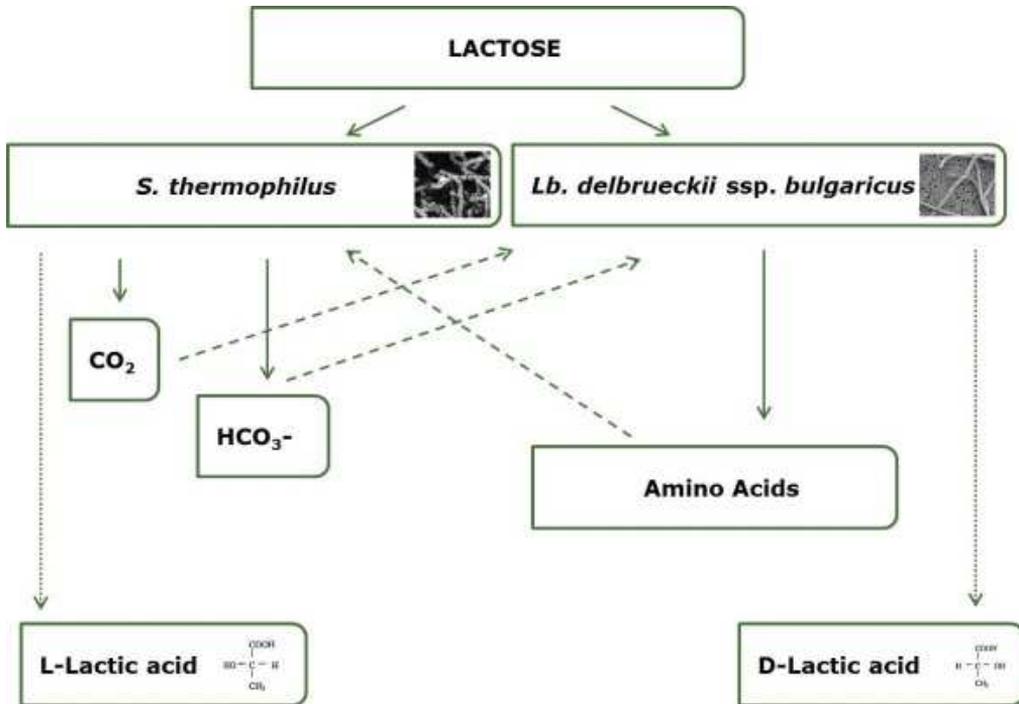


Figure 2.4: Growth associations between the two starter bacteria *S. thermophilus* and *L. bulgaricus*.

This associated growth results in greater acid production and flavour development compared to the growth of a single starter bacteria. In the early stage of the fermentation process, *S. thermophilus* grows faster and removes additional oxygen. Due to the increasing concentration of lactic acid the growth of *S. thermophilus* will decrease, and the more acid tolerant *L. bulgaricus* will then increase in cell number (Nauth, 2004). The growth of the two bacteria in relation to each other is illustrated in Figure 2.5. When mixed cultures of *S. thermophilus* and *L. bulgaricus* are utilized more tyrosine is released than the potential of the sum of the individual starter bacteria. Mixed cultures also produce a larger amount of acids, compared to the sum of the acids being produced by each of the two bacteria (Walstra, et al., 2005a).

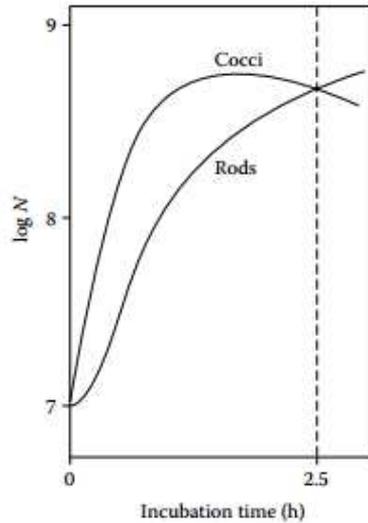


Figure 2.5: The growth of *S. thermophilus* (cocci) and *L. bulgaricus* (rods) over time in a mix culture (Walstra, et al., 2005a).

2.4 Industrial Yoghurt Manufacture

The Production of Yoghurt

In industrial large scale yoghurt production plants several manufacture steps are involved. Each step of the production chain, from raw materials to packaging, has to take place in accordance with Good Manufacturing Practice (GMP) to guarantee the quality and traceability of the product (Béal and Helinck, 2014). In the text below, the industrial production steps in yoghurt manufacture are briefly described.

Milk pre-treatments: initially, a fat standardization is carried out due to seasonal milkfat variation and also to reach the fat content of the final yoghurt product, which can vary from 0.1% to 10%. It is done either by removing part of the fat by mechanical separation or by adding full cream. Other methods can also be applied and most of them are fully automatic systems. Next step is fortification of the solids-not-fat content in the milk, which improves the gel strength of set yoghurt and viscosity of stirred yoghurt. It can be done in several ways, e.g. by addition of milk or whey powder. Hereafter additives, such as sweeteners, stabilizers, thickeners and preservatives are added. Since raw-milk is an oil-in-water emulsion, where fat globules are distributed in a skim-milk phase, the milk is homogenized by a mechanical treatment where high pressure is applied resulting in smaller fat globules, and hence an increase in number and surface area of the fat globules is achieved. From a microbiological point of view, the milk must be heat treated to destroy the pathogenic microflora. Finally, the milk is cooled to the desired fermentation temperature of around 42-43°C (Tamime and Robinson, 2007a; Özer, 2010).

Milk Fermentation: to start the fermentation process, the starter culture containing a mix of *S. thermophilus* and *L. bulgaricus* is inoculated into milk at the desired fermentation temperature. The amount of starter culture can vary a lot from culture to culture depending on the required textural and sensory characteristics of the final yoghurt product. The inoculation is most commonly added as freeze-dried or frozen cultures. If the end product is set yoghurt (Figure 2.6), the inoculated milk will be filled into cups and here the fermentation will proceed. If the end product is stirred yoghurt (Figure 2.6) the inoculated milk will be fermented in a tank. These tanks are typically water jacketed and warm water as heating medium is then circulated during the fermentation. The tanks are equipped with temperature and pH recorders and when a pH around 4.5-4.6 is reached the coagulum is stirred gently and pumped to the filling machine (Tamime and Robinson, 2007a; Özer, 2010).

Set-style yoghurt is inoculated in the package and is packed directly after the inoculation. Thereby, the fermentation is proceeding in the packaging.

Stirred-style yoghurt is defined by having total solid content not higher than 11%. It has a lower viscosity and is more drinkable compared to set-styles. To reduce the viscosity it is further homogenized.



Figure 2.6 Definition of the two yoghurt styles; set and stirred (Gürakan and Altay, 2010).

During the fermentation process the gel formation is taking place. It was previously described how the starter bacteria utilize lactose as energy source and via the sugar metabolism lactose is mainly converted into lactic acid. The casein micelle found in milk has an isoelectric point around pH 5.15 and when this pH is reached during the fermentation process, the micelles become unstable and an aggregation of the micelles will start (Özer, 2010). When the pH further decreases, a contraction of the casein aggregates appears, which results in larger casein particles compared to the native casein micelles (Tamime and Robinson, 2007a). Around pH 4.65 the aggregation is completed and thiol-disulfide bridges will link α -lactalbumin and β -lactoglobulin with the κ -casein, and thereby a gel network will be formed. In more physical terms the gel formation can be described in four steps: (1) an initial lag period with low viscosity, (2) a fast viscosity change, (3) a high viscosity, where the gel formation has completed, and (4) a syneresis stage, which is the death phase of the starter bacteria (Özer, 2010).

Parameters such as fermentation temperature, solids level of the milk, the starter bacteria combination and mechanical handling are all affecting the fermentation time. Likely, the starters' capability of producing acid and their inoculation rate are critical parameters in the gelation kinetics (Walstra, et al., 2005a).

Downstream Treatments and Packaging: when the desired pH has been reached, the fermenting milk is rapidly cooled in order to hinder further metabolic activity of the starter bacteria. In the downstream process cooling, partial dehydration of the coagulum, smoothing of the coagulum, and addition of food additives takes place. The final step is packaging, where selection of packing material is of importance for the protection of the product (Özer, 2010).

An overall graphical illustration of the production plant is seen in Figure 2.7.

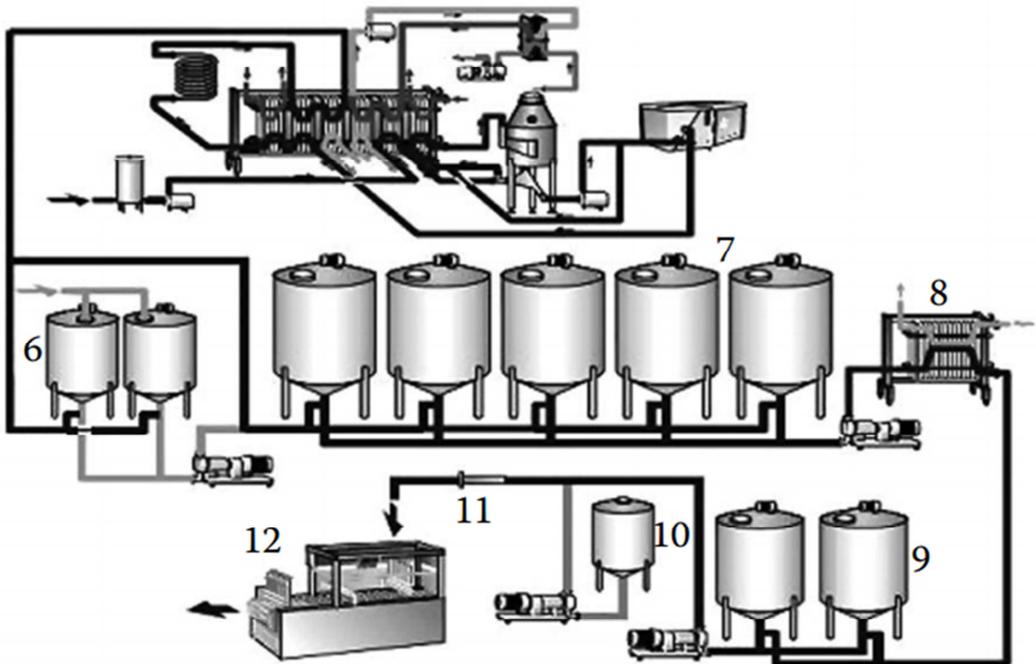


Figure 2.7: Production line for stirred yoghurt. The milk pre-treatment steps are not shown in the figure. 6, bulk starter tanks; 7, fermentation tanks; 8, plate cooler; 9, buffer tanks; 10, fruit flavour tank; 11, inline yoghurt/fruit mixer; 12, filling machine (Tamime and Robinson, 2007b).

Control measurements during the industrial manufacture

In order to ensure reproducibility of the processes, control of all operations is necessary. Initially, the raw materials (milk, milk powder, starters and, if applicable, fruit preparations) are controlled by performing various tests such as temperature, fat- and protein levels of the milk on arrival, total colony count and acidification activity of the starters. During the manufacture, controls have to be carried out to ensure the quality of the products and the repeatability of the production. The Hazard Analysis and Critical Control Point system (HACCP), described by food safety management systems such as ISO 22000, is implemented in all dairy manufacturing companies. In this way identification, evaluation and control of hazards is systematically determined in the operations or steps that are critical to food safety. Several Critical Control Points (CCP) are defined during the different steps in the yoghurt manufacture, however, in this thesis focus is kept on the fermentation step. An important parameter in the fermentation step is pH. In set-style yoghurt fermentation, pH is measured manually, which most likely is done by random sampling of some packages. This way of controlling the pH can be a potential issue due to non-representativeness. In stirred-style yoghurt fermentation, samples are periodically removed from the fermentation tank, or in more modern production sites pH probes are installed in the tank, in order to measure the pH as a function the fermentation time (Özer and Kirmaci, 2010).

Automatic systems allow for controlling the main functions of the industrial process. By applying monitoring devices such as probes and sensors that are connected to a controller, it is possible to compare the measured value to e.g. the desired fermentation temperature. If a significant change or difference to the normal operating conditions is observed, the “controller” increases the heating or the cooling to limit the difference between the measure value and the target value. This form of automatic controlling delivers full traceability of the process operations and thereby ensures full quality control and food safety (Béal and Helinck, 2014).

To my knowledge, most dairies are operated recipe-driven supported by quality assurance/HACCP measurements. And even though the dairy industry seems to promote a more automatic yoghurt production, presently only limit information about the yoghurt itself is gained, and in quality assurance end-product testing rather than the production dynamics are in focus. In the optimal production site the production team would benefit from a real-time measure of product performance in terms of chemical composition, textural properties and e.g. a microbial contamination. Any desired parameter of the yoghurt can be analysed offline in the laboratory, but such measurements cannot replace a rapid instrumental in-line technique allowing the production team or an automatic controlling system to adjust the process when unwanted changes or disturbances are observed (Tamime and Robinson, 2007b).

CHAPTER 3

ON-LINE MONITORING OF FERMENTATION PROCESSES

3.1 Process Analytical Technology

The Food and Drug Administration (FDA) has since 2004 actively promoted the consideration of Process Analytical Technology (PAT) in food and pharmaceutical manufacturing processes. In the Guidance for Industry, PAT is defined as: “a system for designing, analysing, and controlling manufacturing through timely measurements (i.e., during processing) of critical quality and performance attributes of raw and in-process materials and processes, with the goal of ensuring final product quality” (FDA, 2004). Based on the guidance it is expected that the concept and technology behind PAT will help the food and pharmaceutical industry in its development towards more advanced process control.

Fermentation monitoring techniques and process monitoring techniques in general can be classified as off-line, at-line and on-line/in-line, based on the location of the measurement system in relation to the bioreactor or the process line of interest. Off-line measurements include sampling, which can be either manual or automatic, sample transfer to the laboratory, followed by time-delayed laboratory analysis. Such techniques could e.g. be liquid- or gas chromatographic measurements, which are widely used for the analysis of complex mixtures, in which the chemical components are separated before detection. Even though such techniques are accurate and provide valuable process information, they require several steps of sample preparation and are in general time consuming and retrospective, meaning that real-time knowledge about the process cannot be gained by these measurement techniques (Lourenco, et al., 2012). At-line measurements also include sampling, but in this case the withdrawn sample is analysed close to the process, in order to minimize the time delay. At-line also implicitly assumes that the analysis method is less sophisticated/more fail-safe and can be performed by the process operators, and not just by dedicated lab technicians. On-line measurements cover those techniques which are linked directly to the process (Figure 3.1). Normally no sample extraction is involved and measurements are obtained in real-time or at least fast compared to the process dynamics. If the sensor is located directly in the process the monitoring technique can be further classified as an in-line measurement (Callis, et al., 1987). On-line techniques can also be defined as in-situ or ex-situ, where in-situ measurements are collected e.g. in the fermentation broth, whereas ex-situ represent measurements where the on-line probe is not in direct contact with the sample e.g. measurements through a

glass-window in the bioreactor.

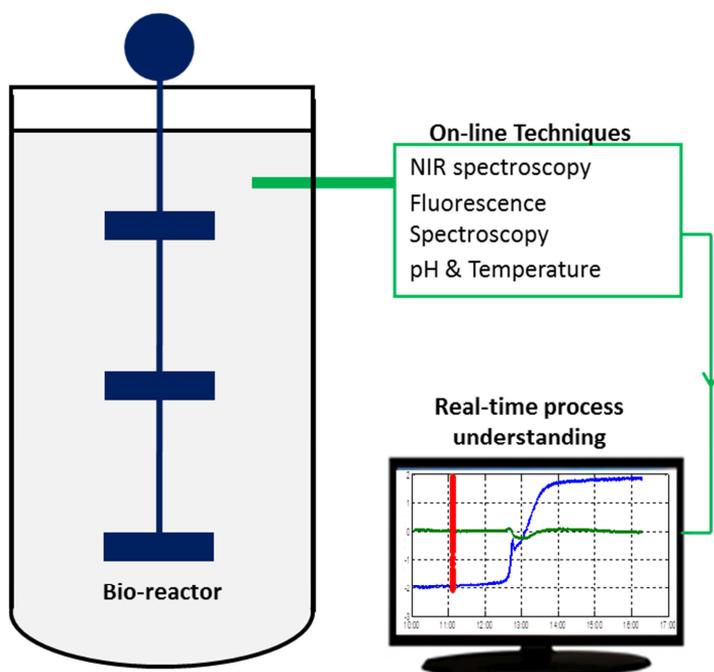


Figure 3.1: Illustration of bioreactor where the real-time information gain from the applied on-line techniques combined with meaningful data handling can provide us with a better process understanding.

Real-time monitoring is considered as an essential approach for effective bioprocess control, which is needed to increase efficiency, productivity and reproducibility. Thereby, quality control can be improved and environmental pollution can be reduced, hence an overall optimization of the cost may appear (Alford, 2006). But, biological processes such as fermentations are complex and due to variation of inoculant the reaction kinetics and biological activity cannot be predicted or calculated at an exact value in advance. Therefore online observations are needed, in which process information can be gathered as the fermentation progresses (King, 2014).

3.2 On-line State-of-the-art Methods for Fermentation Monitoring

In the section on industrial yoghurt manufacture (Paragraph 2.4) it was stated that the only on-line technique applied in the industrial yoghurt fermentation is pH. In this section an overview of the state-of-the-art techniques applied in other industrial food/pharmaceutical fermentation processes are presented. In yoghurt fermentations lactic acid is produced via an anaerobic pathway, meaning that oxygen is not required. But since the aerobic fermentation process is commonly applied in various pharmaceutical productions, this process is also taken

into account in this section. In aerobic fermentation processes the oxidation of glucose can be written as:



where oxygen and the carbon source glucose is consumed and water and carbon dioxide are among other intermedia and metabolites being produced (Ratledge, 2006). In bioprocess monitoring there are in general three types of variables to measure:

- *Physical variables* including e.g. pressure, temperature, viscosity, foam, stirring speed and flow rate.
- *Chemical variables* covering pH, O₂, CO₂, N₂, volatiles gases, nutrients and metabolites.
- *Biological variables* covering biomass concentration, cell metabolism and cell morphology.

All three types of variables are closely related and in some cases biological variables might be derived from chemical variables. In the following text classical techniques commonly used in industrial fermentation processes are presented. Hereafter, a brief overview of spectroscopic techniques applied in bioprocesses, primarily based on lab-scale research, is given.

Monitoring by Classical Methods

For each of the mentioned physical, chemical and biological variables several methods are available. In Figure 3.2 a fermentation plant is illustrated, representative of the systems used in bio-industrial (e.g. enzymes) or biopharmaceutical applications. The figure also includes some of the most important chemical variables, which are monitored with in-line sensors or by a gas analyser. Partial pressure of dissolved oxygen (pO₂), partial pressure of dissolved carbon dioxide (pCO₂), pH- and Optical Density/OD-value are among the chemical variables being measured by in-line sensors, whereas a mass spectrometer measuring the headspace gasses can be applied for control of the oxygen uptake rate (OUR) and carbon dioxide evolution rate (CER) (Stanbury, et al., 1999; Sonnleitner, 2013).

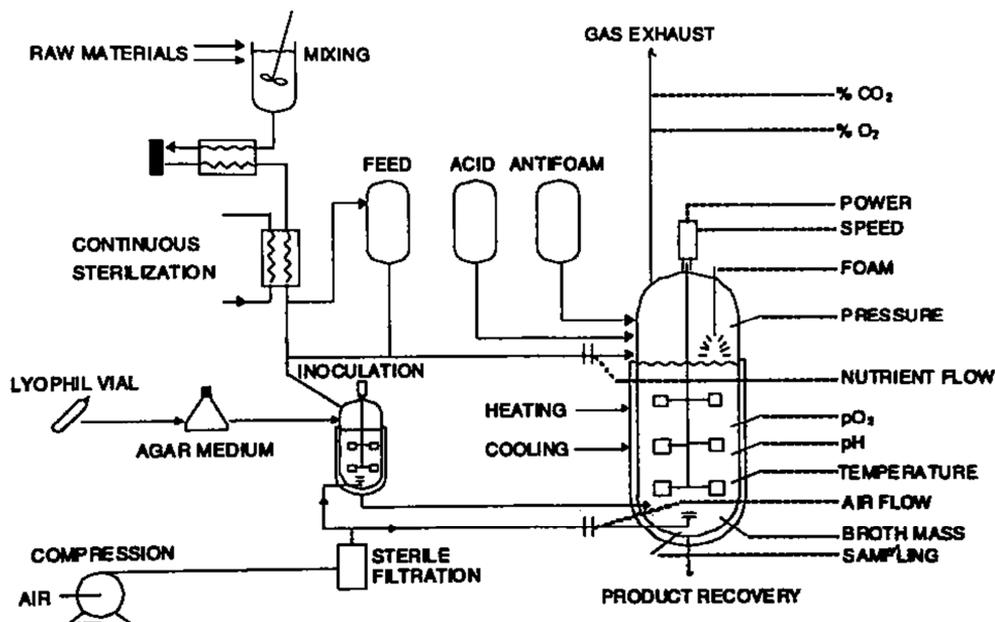


Figure 3.2: An industrial fermentation plant (Heijnen, 2015).

pH: Since the pH of an actively growing culture in a batch fermentation will not remain constant for very long, there is a need for pH control, in order to keep the cells at optimal cultivation conditions. Furthermore, metabolic processes can be highly susceptible to even slight changes in the pH. To maintain the optimal pH for growth, compounds such as ammonia or sodium hydroxide are continuously added (Stanbury, et al., 1999; King, 2014).

Temperature

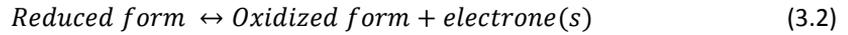
Similar to pH, the temperature control is important to maintain optimal growth conditions (King, 2014).

Optical Density

Optical density is a standard indicator for cell growth. Samples are often being withdrawn and analysed at-line by a spectrophotometer at wavelength 600 nm or 620 nm. However, optical fibre and in-line probes are also available and can be applied in bioprocesses (Sonnleitner, et al., 1992; Lam and Kostov, 2009).

Redox

The Redox potential is a measure of the oxidation-reduction potential of a biological system and is depending on the equilibrium of:



As the fermentation process is a dynamic system it is not at redox equilibrium until the end of the growth cycle. This measurement technique has however been associated with interpretation difficulties (Stanbury, et al., 1999).

Oxygen

As oxygen is consumed and carbon dioxide is formed during aerobic fermentation processes, the dissolved oxygen and carbon dioxide concentrations are indicators for the respiratory activity of the cells. Hence, the measure of oxygen is a measure for the potential of growth. Parameters such as agitation¹, aeration² rate and the composition of the gas phase are also dependent on the oxygen level. The electrodes which can be applied for oxygen monitoring do not directly measure the concentration, but the partial pressure of the dissolved oxygen. However, the concentration can be determined by the following equilibrium:

$$P(O_2) = C(O_2) \cdot P_T \quad (3.3)$$

where $P(O_2)$ is the partial pressure of dissolved oxygen measured by the sensor, $C(O_2)$ represents the concentration and P_T is the total pressure. If the total pressure of the gas equilibrium with the fermentation broth varies, the actual reading of the partial pressure of dissolved oxygen is affected even though no changes appear in the gas composition. The temperature can also influence the dissolved oxygen reading by approximately 2.5 % per °C. This is mainly caused by increases in the permeability in the electrode membrane and therefore many electrodes have built-in temperature sensors that compensated for such changes (Stanbury, et al., 1999; Sonnleitner, 2013; Biechele, et al., 2015).

Carbon dioxide

Carbon dioxide (CO₂) is not only an indicator for the respiratory activity of the microorgan-

¹Agitation is important to ensure suspension of the biocatalyst and attain a relatively homogeneous environment in the bioreactor. Agitation can be applied by various types of impellers (Chisti, 2006).

²Aeration is normally applied by sparing air bubbles into the bottom of the bioreactor (Chisti, 2006).

isms. It can also affect the microbial growth in different ways due to its appearance in the catabolism as well as in the anabolism³. The partial pressure of carbon dioxide ($p\text{CO}_2$) can be measured indirectly as pH by a bicarbonate buffer separated from the growth medium by a gas-permeable membrane (Locher, et al., 1992).

Gas analyser

Oxygen and carbon dioxide can for instance be measured by mass spectrometers, which allow rapid on-line monitoring of gasses in the headspace of the bioreactor. Dissolved gasses such as O_2 , CO_2 and CH_4 , but also volatile compounds such as methanol, ethanol, acetone and other simple organic molecules can be measured. Inlet and exit-gas analysis can provide concentrations of carbon dioxide and oxygen in the entry and exit gasses. Also, if the flow rate is determined, the oxygen uptake of the system, the carbon dioxide evolution rate and the respiration rate of the microbial culture can be determined. The correlation equation for the oxygen uptake rate is:

$$\text{OUR} = q_{\text{O}_2} \cdot C_X = \frac{1}{Y_{\text{XO}}} \cdot \frac{dC_X}{dt} + m_{\text{O}_2} \cdot C_X \quad (3.4)$$

where Y_{XO} is the yield of oxygen consumed for cell growth, m_{O_2} is oxygen consumption and C_X is the biomass concentration. The equation for carbon dioxide evolution follows the same format (Mitchell, et al., 2000; Garcia-Ochoa, et al., 2010).

Monitoring by Spectroscopic Methods

The common aspect for the above mentioned techniques is that they are all producing a univariate output (though, the gas analyser/mass spectrometer could be classified as a multivariate method measuring several analytes or target molecules). Due to the fact that bioreactors are complex multivariable systems, where substrates are consumed and products and intermediate metabolites are formed, several limitations might appear in the process understanding when univariate methods are applied. Also, retrieving additional off-line information about concentrations of reagents and products often requires sample preparation, which delays the analysis and thereby the process control. In this section spectroscopic measurements are presented. Due to the fact that all spectroscopic methods are based on interaction between electromagnetic waves and molecules, spectroscopic techniques are suitable for qualitative and quantitative analyses. The range of applicable wavelengths, where spectroscopic techniques can measure, ranges from 190 nm to 1 m (Beutel and Henkel, 2011), as

³The combined processes of catabolism and anabolism are known as the metabolism, where the catabolism can be defined as the degradation and the anabolism as the biosynthesis in which cell materials are build up (Ratledge, 2006).

depicted in Figure 3.3.

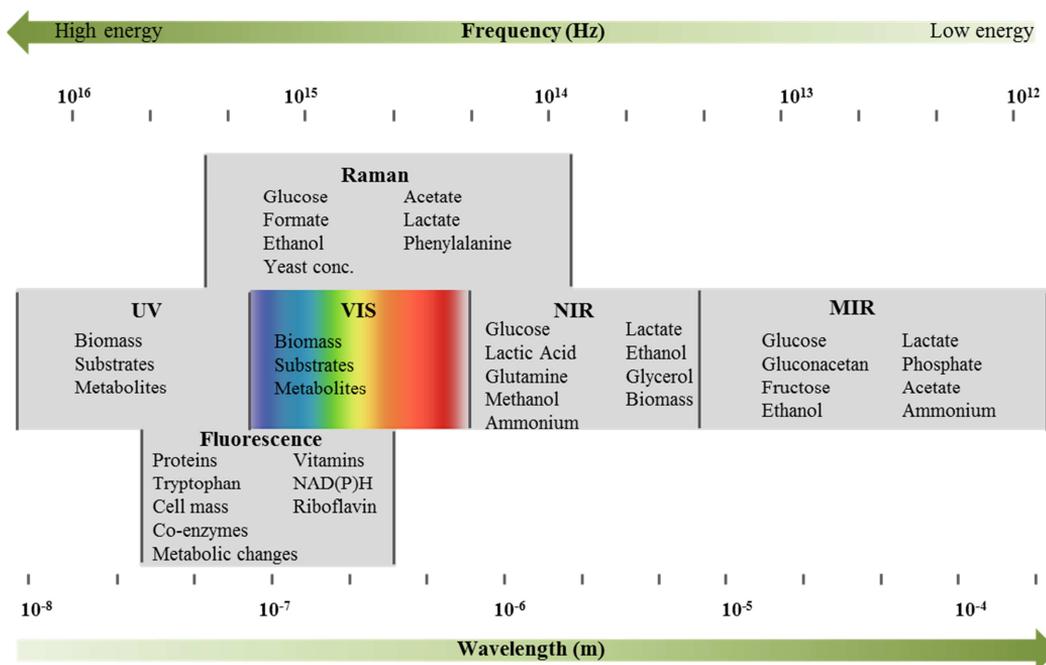


Figure 3.3: The Electromagnetic spectrum, where some of the compounds detected in bioprocesses by the various spectroscopic techniques are stated; UV/VIS, Fluorescence, Raman, NIR and MIR spectroscopy are illustrated.

In the following text, focus is not on the fundamentals of each technique, but rather on the biological, chemical and physical information that can be obtained real-time in fermentation processes.

UV/VIS spectroscopy

UV/Vis uses ultraviolet (UV) and visible (VIS) light (190-740 nm) to excite electrons of molecules, where the spectral region for UV is typically defined from 190 to 400 nm and the spectral region for VIS is from 400 to 750 nm. A selection of analytes, substrates, metabolites and products can be measured by UV/Vis spectroscopy (Herman, 2000; Pons, et al., 2004). In this spectral region transmittance measurements at 600 nm or 620 nm are commonly carried out for biomass monitoring. These measures are, as previously mentioned, known as optical density measurements i.e. OD_{600} values, and they have been applied to various cell cultures. (Abou El-Magd, et al., 2010; Alves-Rausch, et al., 2014; Schmidt-Hager, et al., 2014). A high resolution UV/Vis spectrophotometer for monitoring bioprocesses at-line has also been described in the literature (Noui, et al., 2002).

Fluorescence spectroscopy

In fluorescence spectroscopy fluorophores are excited by light of a specific frequency, and molecules are thereby elevated to a higher energy level. When the molecules returns to their ground state, photons are emitted at other frequencies, meaning that each fluorescence compound has a specific excitation wavelength and a specific emission wavelength. Several molecules, including fluorophores such as tryptophan, vitamins and co-enzymes, appearing in food and bioprocesses have fluorescence properties (Marose, et al., 1998). The very first on-line sensors were based on one excitation and one emission wavelength, which only allows measurements of one fluorophore at a time (Zabriskie and Humphrey, 1978). Such measurements only provide little information for a bioprocess containing several biological compounds because the full excitation-emission landscapes are always broad and there is thus a high chance for overlapping signals in univariate recordings. Real-time measurement of several fluorophores in the fermentation broth is possible, when approaches like 2D fluorescence spectroscopy are applied in combination with an on-line probe (Herman, 2000; Lakowicz, 2006a). Such an on-line approach for industrial applications was developed by DELTA Lights & Optics and named BioView (The BioView sensor, DELTA Light & Optics, Denmark). The BioView sensor is capable of capturing automatic optical measurements in industrial bioprocesses, where the environment can be harsh, due to e.g. high temperatures and moisture (Stärk, et al., 2002). A number of studies have investigated the potential of bioprocess monitoring by on-line 2D fluorescence.

One study applied a *Bacillus polymyxa* batch cultivation to evaluate if the BioView sensor was able to give successful determinations of cell mass and the produced compound polymyxin. In order to achieve different growth patterns, the concentrations of phosphate and nitrogen were varied in the medium. The cultivations were run for 50 hours and were carried out in a bioreactor with a working volume of 41 L. The stirring was kept at 900 rpm, aeration was applied to control the dissolved oxygen level, and pH was maintained at 7 by addition of base. The study obtained successful determinations of polymyxin, however the calibration model for the biomass was not successful, since it was not possible to use a common model for all cultivations under various growth conditions (Lantz, et al., 2006).

A study based on filamentous fed-batch cultivation was conducted to monitor the biomass by on-line BioView fluorescence measurements. The cultivations were carried out with the microorganism *Streptomyces coelicolor* and with a working volume of 5 L, which was increased to around 9 L at the end of the fed-batch phase. The stirring was kept at 600 rpm and the aeration was kept constant. The pH was maintained at either 5.9 or 6.9 and the cultivations continued for 80-110 hours. They found that fluorescence failed to produce reliable estimations of the dry cell weight. The author suggests that the failure might be explained by the differences in pH, which are known to have a large effect on the fluorescent properties of

some fluorophores (Ronnest, et al., 2011). Though filamentous cultivations are considered to be more challenging, it seems that the largest challenge was the pH differences.

Another filamentous study managed to measure the cell mass concentration and the lipase activity (Haack, et al., 2007). This study was based on fungus *Aspergillus oryzae* fermentations. The batch cultivations were carried out in a reactor with a working volume of 4 L. The stirring speed and the aeration rate were kept constant at 800 rpm and 1.0 vvm, respectively. The pH was controlled at 6 by adding acid or base. Fed-batch cultivations were carried out under comparable conditions. In addition to the successful estimations of cell mass and lipase activity, the authors also report, that they observed an increase in the fluorescence intensity, when they increased the stirring rate. They assumed that this is caused by the position of the probe. Other studies have shown that agitation and aeration can lead to unstable spectra when the probe is inserted vertically instead of in the side of the reactor (Li and Humphrey, 1992).

Likewise, monitoring of viable cells and the concentration of the produced recombinant glycoprotein was successfully carried out in mammalian fed-batch fermentations (Teixeira, et al., 2009). The cell culture consisted of baby hamster kidney cells and both batch and fed-batch fermentations with different feeding strategies were carried out in a 2 L bioreactor. The pH was maintained at 7.2 and the agitation rate was kept at 60 rpm.

Apart from bacterial, fungus and mammalian cell fermentations, fluorescence spectroscopy has also been applied to yeast and various food fermentations. As the research presented in this thesis is based on a model system with lactic acid bacteria, it should also be mentioned that previous studies have applied fluorescence spectroscopy in yoghurt. A study by Christensen *et al.* (2005) evaluated yoghurt samples by a BioView spectrometer. The yoghurt samples were measured during a storage experiment and riboflavin, tryptophan and the degradation product from riboflavin, lumichrome were assigned. Furthermore, they managed to build a regression model for riboflavin, and thereby illustrate that fluorescence spectroscopy is a potential method for rapid determination of riboflavin.

Only a limit number of studies have used fluorescence spectroscopy to measure yoghurt samples/productions, whereas several studies applying fluorescence spectroscopy on other cultivations types exist. A recent review study (Faassen and Hitzmann, 2015) has gathered most of the investigations which have monitored cultivations by fluorescence spectroscopy. In Table 3.1 the studies based on the BioView spectrometer are listed.

Table 3.1: Overview of studies applying BioView Spectrometer in cultivation processes.

Type	Organism	Cultivation	Reference
Bacteria	<i>Escherichia coli</i>	Batch	(Clementsitsch, et al., 2005)
	<i>Escherichia coli</i>	Fed-Batch	(Johansson and Liden, 2006)
	<i>Bacillus polymyxa</i>	Batch	(Lantz, et al., 2006)
	<i>Klebsiella pneumonia</i>	Batch	(Rossi, et al., 2012)
	<i>Aspergillus oryzae</i>	Batch & Fed-Batch	(Haack, et al., 2007)
	<i>Bacillus</i>	Fed-Batch	(Mortensen and Bro, 2006)
	<i>Streptomyces coelicolor</i>	Fed-Batch	(Odman, et al., 2010)
	<i>Streptomyces coelicolor</i>	Fed-Batch	(Ronnest, et al., 2011)
Fungi	<i>Pichia pastoris</i>	Batch	(Surribas, et al., 2006b)
			(Hisiger and Jolicoeur, 2005a)
			(Surribas, et al., 2006a)
	<i>Saccharomyces cerevisiae</i>	Fed-Batch	(Hantelmann, et al., 2006)
		Batch	(Odman, et al., 2009)
			(Haack, et al., 2004)
<i>Claviceps purpurea</i>	Batch	(Masiero, et al., 2013)	
<i>Boehl, et al., 2003)</i>			
Mammalian	<i>Chinese Hamster Ovar Cells</i>	Batch	(Bonk, et al., 2011)
Plant	<i>Eschscholzia California</i>	Batch	(Hisiger and Jolicoeur, 2005b)
	<i>Catharantuhus roseus</i>		

The mentioned studies on fluorescence spectroscopy were all carried out with the BioView instrument, which measure the excitation wavelengths ranging from 270 to 550 nm and the emission wavelengths ranging from 310 to 590 nm, both in steps of 20 nm. Acceptable monitoring performance was reached for bacteria, filamentous bacteria and fungi and mammalian cells.

Raman spectroscopy

This spectroscopic technique is another form of vibrational spectroscopy, which is based on shifted wavelength scattering of molecules appearing after excitation by a monochromatic light source (Becker, et al., 2006). Several parameters such as glucose, acetate, formate, lactate and phenylalanine can be monitored in fermentation processes by Raman spectroscopy, which was e.g. illustrated by a study based on an *Escherichia Coli* cultivation (Lee, et al., 2004). Similarly, a study based on yeast cultivation showed that glucose, ethanol and yeast concentration could be determined by Raman spectroscopy (Iversen, et al., 2014). Though Raman spectroscopy seems to be useful for fermentation monitoring, the strong fluorescence activity of the many biological molecules is a problem, since they overshadow the Raman bands when a laser in the visual range (< 830nm) is used. Lasers using longer wavelengths (e.g. 1064nm) would eliminate part of the fluorescence issue, but their low efficiency and e.g. potential risk in an industrial surrounding makes practical use very limited. Since the fluorescence com-

pounds cannot be eliminated from the system, a quite strong signal is needed to determine analytes by Raman (Beutel and Henkel, 2011).

NIR spectroscopy

Near infrared spectroscopy (NIRS) is a measurement technique with the potential of rapid and accurate determination of chemical composition, also in biological systems. It covers the spectral area from about 800 to 2500 nm and the technique is based on different vibrational modes, overtones and combination vibrations. The main functionalities targeted are O-H bonds of alcohols, C-H bonds of aliphatic and aromatic carbon compounds and N-H bonds of proteins (Landgrebe, et al., 2010; Beutel and Henkel, 2011). The relatively low cell density anaerobic fermentations for production of lactic acid and ethanol are considered to be some of the simplest fermentation systems, due to gently agitation and usually non-aeration (Scarff, et al., 2006). In such a system, where the microorganism *Lactobacillus casei* was applied, NIRS has been implemented on-line to control glucose, lactic acid and biomass (Vaccari, et al., 1994). NIRS was applied in batch, fed-batch and continuous fermentations. The fermentations were carried out in a fermentor with a working volume of 3 L. Stirring at 120 rpm was applied and nitrogen gas was passed through the headspace to maintain anaerobic conditions. The authors conclude that the applied on-line system permits control of the process and ease the optimisation of the process.

Another study compared at-line and in-situ NIRS measurements in a fed-batch cultivation, where the bacterium *Escherichia coli* was applied (Arnold, et al., 2002). The cultivations were carried out in a 1.5 L fermentor with an aeration rate of 1 vvm and agitation of 800 rpm. The transmittance spectra were collected either at-line with a pathlength of 5 mm or in-situ with a pathlength of 0.5, 1 and 2 mm. For the in-situ approach a fibre optic probe was applied, which means that the spectral region above 2100 nm is unusable as residual hydrogen bonds in the fibres have a detrimental effect on the spectra. However, successful monitoring for both at-line and in-line settings was achieved. The authors also evaluated various pathlength for the on-line approach. They concluded that an increased pathlength could improve the signal by increasing the absorption, but in practice it also means an increase of the interference from the air bubbles entering the probe-head, which resulted in noisy spectra. For that reason the best reproducibility was obtained for the smallest pathlength of 0.5 mm.

NIRS has also been implemented in aerobic mammalian cell fermentations for on-line monitoring of glucose, lactate, glutamine and ammonia (Arnold, et al., 2003). Chinese hamster ovary cells were applied and the cultivation was carried out in a 2 L reactor. A transmission probe with a pathlength of 1.2 mm was placed in the fermentation broth in order to obtain in-situ measurements. The authors report that NIRS in-situ in this kind of process is challenging, since it is hard to achieve accurate determination of the key analytes (glucose, lactate, gluta-

mine and ammonia), which are found at very low levels. Though the levels are low compared to ex-situ studies, it is concluded that in-situ NIRS potentially provide a highly effective means of improving the monitoring and control of the cultivation.

The final study to be mentioned here was able to provide information about the biomass in fermentation processes with the pathogen bacteria *Bordetella pertussis* (Soons, et al., 2008). The batch cultivation was carried out in a 7 L reactor with a working volume of 4 L. The measurements were collected by transmission NIRS with a pathlength of 5 mm. The authors concluded that the on-line estimations of biomass were less robust towards noise and various disturbances from the environment. They also compared the NIRS data with data obtained from a soft sensor for dissolved oxygen, which seems to be less sensitive towards the surroundings than the NIRS. Since NIRS, in comparison to the soft sensor, has the potential of determining other key parameters in the cultivation the authors recommended using joint modelling of data from the two sensors in order to obtain complementary information and the best monitoring system.

Most studies on NIR spectroscopy applied for bioprocess monitoring are based on lab-scale fermentations, but a recent study applied NIRS in a full scale industrial reactor of 50 m³ (Alves-Rausch, et al., 2014). The culture grew aerobically at an agitation speed of 100 rpm and a relatively high aeration rate between 0.4 and 1 vvm. The authors report that the high aeration, which is needed for this kind of cultivation, leads to air bubbles passing the probe and thereby a decrease in the measured absorption is seen, since less sample amount is exposed to the NIRS detector. However, good performance in determining key analytes was achieved.

A review has gathered most of the studies, which have monitored cultivations by NIRS (Cervera, et al., 2009). In Table 3.2 some of the key studies are listed together with more recently published investigations.

Table 3.2: Overview of studies applying NIRS in cultivation processes

Type	Organism	Cultivation	Sampling/Mode	Instrument	Reference
Bacteria	<i>Bacillus</i>	Full scale, 50 m ³	On-line (in-situ)	Sartorius Stedim Biotech	(Alves-Rausch, et al., 2014)
	<i>Bordetella pertussis</i>	Batch, 7 L, work vol:4 L	On-line (in-situ) Transmission, 5 mm	Bruker Optics Matrix F	(Soons, et al., 2008)
	<i>Escherichia coli</i>	Fed-Batch 1.5 L	At-line (cuvette): Transmission, 0.5 mm On-line (in-situ) Transmission: 0.5, 1, 2 mm	Foss- NIRSystems	(Arnold, et al., 2002)
	<i>Lactobacillus casei</i>	Batch, 3 L Fed-Batch Continuous	On-line (ex-situ)	InfraAlyzer, Bran- Luebbe co.	(Vaccari, et al., 1994)
	<i>Streptomyces coelicolor</i>	Batch, 3 & 4 L	On-line (in-situ) Transflectance, 0.5 mm	ABB Bomem	(Petersen, et al., 2010)
	<i>Streptomyces fradiae</i>	Batch, 10 L	At-line (cuvette) Transmission, 1 mm	Foss- NIRSystems	(Arnold, et al., 2000)
Fungi	<i>Pichia pastoris</i>	Batch, 15 L	At-line (cuvette) Transmission, 0.5 mm Reflectance, 4 mm	Foss- NIRSystems	(Crowley, et al., 2005)
		Fed-Batch, 3 L	On-line (ex-situ) Transmission, 1 mm	ASL Analytic	(Kim, et al., 2015)
		Fed-Batch, 1.6/3 L	On-line (ex-situ) Transmission, 1 mm	ASL Analytic	(Goldfeld, et al., 2014)
Mammalian	Chinese Hamster Ovary	Fed-Batch, 2 L	On-line (in-situ) Transmission, 1.2 mm	Foss- NIRSystems	(Arnold, et al., 2003)
		Batch	Transflection, 1 mm	Antaris II MX, Thermo Fisher	(Clavaud, et al., 2013)

These studies clearly point out the potential of NIRS in monitoring fermentation processes. It should however be noticed that NIRS is an indirect method, so the above mentioned parameters are all found because spectral information correlates well with off-line measurements representing the specific parameters.

MIR spectroscopy

MIR radiation excites fundamental vibrations of functional groups from organic compounds and the technique covers the spectral area from 1300 (the overlap with/separation from the NIRS range is treated differently by different authors) and up to 15000 nm. MIRS can be implemented for in-line measurements by use of optical fibre probes, albeit much less flexible

compared to e.g. NIRS (Lourenco, et al., 2012). Studies applying attenuated total reflectance (ATR) probes have shown that glucose and ethanol can be determined in-line in *Saccharomyces cerevisiae* fermentation (Bogomolov, et al., 2015), glucose and lactate can be determined in-line in mammalian cell cultivation (Rhiel, et al., 2002) and based on a *Gluconacetobacter xylinus* fermentation calibration models providing fructose, acetate, ethanol, gluconacetan, ammonium and phosphate concentrations were made (Kornmann, et al., 2004). Though on-line reflectance measurements by MIR spectroscopy are possible, an important disadvantage of the use of infrared spectroscopy is the large amount of absorbance by water, which can mask important information in the infrared region (Stuart, 2004).

3.3 Study I – Exploratory Study on Fermentation Monitoring using Measurements Techniques of different Orders

This section is based on the peer reviewed paper “Monitoring fermentation processes using in-process measurements of different orders” (Svendsen, et al., 2015) published in Journal of Chemical technology and Biotechnology.

Aim

In order to explore the advantages and variations of different ordered measurement techniques a lactic fermentation process was monitored by both univariate and multivariate techniques providing different data structures. BRIX and pH were applied as univariate techniques, whereas NIR and fluorescence spectroscopy were applied as multivariate techniques. BRIX and pH can be classified as zero-order measurements, which only provide one data point per measurement. When measuring BRIX or pH over time, the single measurements each representing a time point, can be organized in a data set.

As illustrated in Figure 3.4, the gathered data array will be a one-way data array, a vector. NIR spectroscopy provides an absorbance value at each wavelength between 1000 to 1800 nm, meaning that the technique provides a first order data outcome for each measurement. This results in a two-way array when several measurements over time are gathered into one data set. The outcome for one fluorescence measurement is a two dimensional landscape, hence fluorescence is a technique providing data of second order for each recoding, resulting in a three-way array when several measurements representing various times are gathered in one set. Multivariate data analysis (PCA⁴ and PARAFAC⁵) was applied on the first- and second-order data sets, and the different measurement signals or principal/latent variables derived of these were combined by a multiblock strategy.

⁴Principal Component Analysis (PCA) is described in Paragraph 4.4 Data Processing for NIRS Data.

⁵Parallel Factor Analysis (PARAFAC) is described in Paragraph 5.4 Data Processing for Fluorescence Spectroscopy Data.

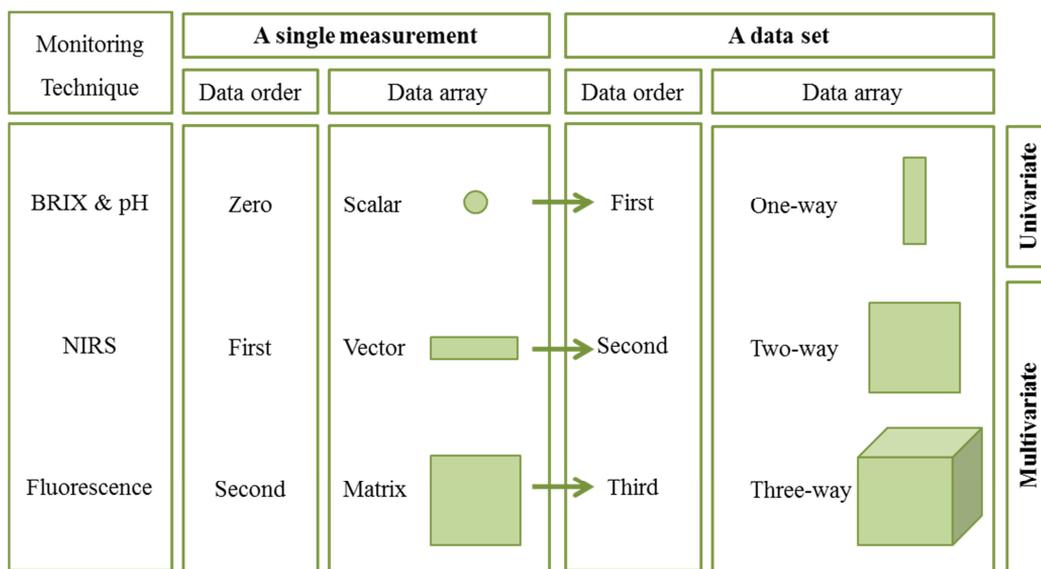


Figure 3.4: Overview of the three different data structures applied in this study.

Illustrations of the Experimental set-up

This study is based on lactic acid fermentations with the starter bacteria *Streptococcus thermophilus* (*S. thermophilus*) and *Lactobacillus delbrueckii* subsp. *bulgaricus* (*L. bulgaricus*). After 5 hours of fermentation the fermented milk product yoghurt was obtained. The frozen starter culture applied is pictured in Figure 3.5 and the experimental set-up is shown in Figure 3.6.



Figure 3.5: Frozen starter culture containing a mixture of the lactic acid bacteria *S. thermophilus* and *L. bulgaricus*.

Available instruments of different measurement orders were used. BRIX data were collected frequently via manual at-line measurements by a DR-103 (Index Instruments Limited, Cambridgeshire, UK) and pH data were obtained with an in-line pH-meter (MadgeTech, Inc, Warner, NH-US) measuring every 30 sec. NIR spectroscopy data were collected by a ABB Bomem spectrometer (ABB Bomem, Quebec, QC, Canada) measuring with a time interval of 60 sec, and the fluorescence measurements were collected on-line and non-invasive through the glass fermenter wall with a BioView spectrofluorometer (DELTA Light and Optics, Hørsholm, Denmark) measuring every 120 sec. Further details on the experimental set-up are given in Paper I.

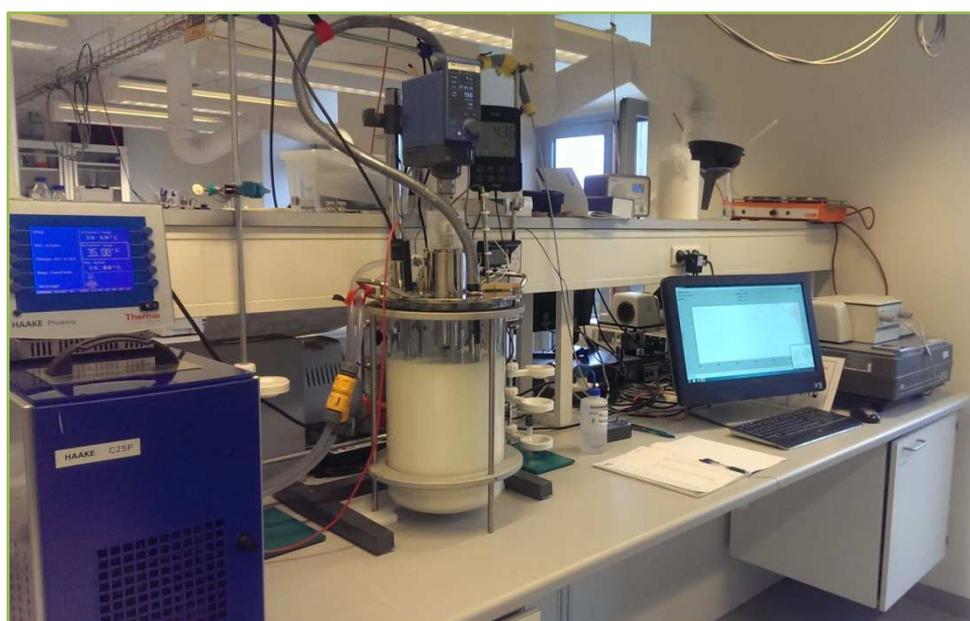


Figure 3.6: The experimental set-up. From the left; the temperature was maintained by the blue HAAKE Phoenix pumping water bath connected to the bioreactor; the 15 L bioreactor containing the growth medium skimmed milk (working volume 11 L). In-line probes (pH, temperature and NIRS) are situated in the top and enter the reactor via ports. Fluorescence probe (not visible) was located on the site and was measuring through the glass fermenter. An IKA EUROSTAR 60 control motor is controlling the stirring speed, also situated at the top; on-line NIRS spectra can be followed on the screen. The ABB Bomem spectrometer connected with optic fibres reflectance probe placed in the fermenter broth.

Results

As formulated in Paper I, the zero-order pH and brix measurements decreased in a smooth and logical pattern from 6.4 to 4.4 and from 10.5% to 6.2%, respectively. These values are Critical Quality Attributes (CQA), confirming that the fermentation process is progressing over time in accordance with biological and engineering intuition. The first-order NIR measure-

ments modelled with PCA showed an increasing trend over time, which corresponds to the growth of the lactic bacteria. Based on the second-order fluorescence measurements modelled by PARAFAC with its mathematical uniqueness properties, three distinctive fluorescence compounds were found to vary over fermentation time. Most probably these three compounds represent riboflavin, tryptophan and lumichrome or NADH. Using multiblock PCA the combined sensor signals identified two distinguished and reproducible time profiles for all batch runs.

The most interpretable chemical information was obtained by fluorescence spectroscopy due to the uniqueness properties of second-order measurements. The first-order technique NIR spectroscopy also provided valuable process information, though the process trends only can be interpreted indirectly and if interfering species had been encountered they could not have been modelled. The multiblock data set provided by zero-, first- and second-order measurements recorded over time highlighted important relationships among the different variables (sensors) that provide chemical information when multivariate data analysis is applied.

Discussion, Concluding Remarks and Perspectives

The study illustrates that more process information can be obtained by the applied first- and second-order measurements compared to the applied zero-order measurements. However, it is important to keep in mind that the right combination of zero-order measurements, such as pH and the previously mentioned standard methods OD, oxygen and carbon dioxide, also can provide valuable process information, especially if several zero-order measurements are combined.

The study confirms that on-line measurements can provide us with real-time process understanding throughout the process. The process measurements in the lactic fermentation process provide real-time process data, from which process trends can be compared to previous/historical runs and thereby reduce uncertainty and potentially product variations. This means that in a larger production line, it is possible to get closer to the production target and in that way potentially increase the production yield and improve the product quality and consistency. Additionally, the process trends allow a high traceability, which enable us to detect process break down, such as leaks, in real-time. These advantages are not possible by off-line measurements carried out away from the process, where all information is delayed. When considering off-line measurement uncertainty, we have to include the sample error, which in comparison to on-line measures might make the overall error or uncertainty larger for off-line measurements. Furthermore, it should also be remembered that on-line measurements often are obtained with a lot shorter sampling interval than off-line measures, which means that a single measurement outlier measured on-line does not have the same influence as an off-line measure, which is sampled with a larger sample interval. Thus, the

overall measurement error theoretically is lower for on-line measurements and often the process dynamics can be more easily reconstructed even from noisy, high-frequency measurements.

Though several of the spectroscopic on-line approaches have shown promising results for fermentation monitoring in lab scale and pilot plant research platforms, it does not seem that these methods have been successfully implemented in the industry. Some challenges still exist, which complicates the achievement of a reliable on-line spectroscopic system. To overcome some of these challenges further research is needed. Therefore more detailed studies on NIR and Fluorescence spectroscopy, respectively, has been carried out (Paper II and Paper III).

Even though certain on-line multivariate techniques, such as NIR spectroscopy, provide indirect information, the data values gathered over time can provide us with information about the process dynamics, which also provide important information about the process trend. In order to find out, what is actually causing the dynamics obtained by the NIR spectra, they can be compared with off-line reference measurements providing quantitative analytical information. Since the NIRS data in this study seems to be consistent with the pH- and BRIX, it is considered that the NIRS data might represent the growth of the bacteria. In order to explore this further a study focussed on monitoring of lactic fermentation by NIR spectroscopy was carried out. The study is presented in Chapter 4.

Though the current and several other studies reveal that on-line process monitoring provide importance process understanding, some requirements must be fulfilled for in-line sensors in order to gain successful data. They must be long-term stable and should not be easily affected or changed by disturbances from the surrounding system, e.g. pH changes. Also, the dynamic sensitivity range of the sensor has to meet the encountered parameter range seen during the process. Optical and spectroscopic techniques do show promising perspectives due to the broad range of information given by the various wavelength areas (Beutel and Henkel, 2011).

From the previously presented research studies we learned that various factors can influence the fluorescence signal intensity. It was e.g. reported that agitation, aeration and change in pH can have an effect. Since only a limit aeration of 150 rpm or no agitation was applied in the lactic fermentations, this is not an issue in my investigations. However, the pH might be a major issue, since the pH is systematically decreasing as lactic acid is being produced during the fermentation. As such changes do not meet the requirements towards stability of sensors and as it seems that the sensors are easily affected by the surrounding system, our applied fluorescence spectroscopy system was further tested. Therefore, the fluorescence spectroscopy technique was further explored and studied during disturbances in form of pH fluctuations. This study is further presented in Chapter 5.

CHAPTER 4

NEAR INFRARED SPECTROSCOPY FOR ON-LINE MONITORING OF FERMENTATION PROCESSES

4.1 Introduction

A large number of studies have shown that NIR spectroscopy is a potential technique for fermentation monitoring. It can be applied on-line and thereby provide real-time measurements which makes process optimization and quality assessment along the process possible (Cervera, et al., 2009).

In this chapter, the fundamental of NIR Spectroscopy will shortly be addressed. Next, sample modes and data processing are elucidated. Chemical and physical information obtained by NIRS in fermentation processes are then clarified and finally Paper II on “Exploring process dynamics by near infrared spectroscopy in lactic fermentations” is introduced and discussed.

4.2 Fundamental of Near Infrared Spectroscopy

Near infrared (NIR) spectroscopy is a method which determines the absorption in the NIR region from about 800 to 2500 nm, situated in between the visible and fundamental infrared region of the electromagnetic spectrum (Figure 4.1).

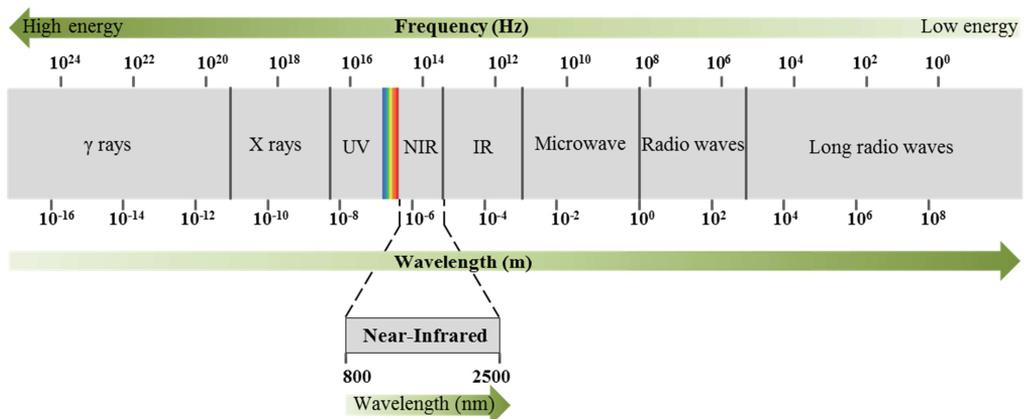


Figure 4.1: Illustration of the electromagnetic spectrum, where the near infrared region covers the wavelengths from 800 to 2500 nm.

The absorption in the NIR region is based on vibrational energy changes corresponding to overtones and combinations of fundamental vibration of molecules observed in the infrared (IR) region (Scarff, et al., 2006; Harris, 2010).

The term vibrational energy refers to oscillations of atoms, which are connected and oscillated through their molecule bonding. The energy (E) of such a system is explained as:

$$E = \frac{h}{2\pi} \sqrt{\frac{k}{\mu}} \quad (4.1)$$

where k is the force constant of the bond being probed, h is Planck's constant and μ is the reduced mass of the atoms involved in this bond, described as:

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \quad (4.2)$$

in which m_1 and m_2 represent the masses of the two atoms. The vibrational energy changes are caused by variations in the so called dipole moment of the molecules. The dipole moment appears when transition between energy states of molecular vibrations corresponds to a change in the molecule's polarity. When molecules are irradiated they can absorb light photons, which have the same frequency as the vibrating bond of the concrete molecule. Hence, molecules are relatively selective about the light that they absorb. Since the vibrational frequency of the vibrating bond is equal to the energy different between two vibrational states, the energy difference (ΔE) can be explained if the vibration is assumed to be harmonic, as seen below:

$$\Delta E = E_{v_2} - E_{v_1} \quad (4.3)$$

where E_{v_1} and E_{v_2} each represent a vibrational state. The relation between the energy (E) and the frequency (ν) is then described as below (Harris, 2010):

$$E = h\nu \quad (4.4)$$

Since light energy is directly linked to the frequency or wavelength of the light, only certain light frequencies or wavelengths can be absorbed by a specific molecule, which provides physical and chemical information about the measured material. Thus, vibrational spectroscopy is

appropriate for qualitative analysis. Quantitative analyses are also possible by NIRS based on the amount of light absorbed at each wavelength. Since the absorbed light cannot be determined directly, it has to be calculated from the transmitted or reflected light. The absorbance is defined as follows:

$$A = \log \frac{I_0}{I_t} \quad (4.5)$$

where I_0 is the intensity of the incident radiation, and I_t is the intensity of the transmitted radiation. The Lambert Beer's law is a fundamental law for the quantitative absorption spectroscopy describing how absorbance is directly proportional to the concentration of light absorbing species (Griffiths, 2002). It is defined as:

$$A = \epsilon bc \quad (4.6)$$

in which the absorbance (A) is equal to the conditions of the concentration (c), the (effective) pathlength (b) and the molar absorptivity (ϵ), which tells how much light is absorbed at a particular wavelength for a given substance (Harris, 2010).

The absorbance measured by NIRS is mainly based on overtones and vibrational combinations, whereas the fundamental vibrations can be measured by Infrared (IR) spectroscopy. The vibrations vary due to the inter-molecular bindings. As an example the CH_2 bindings are shown in Figure 4.2.

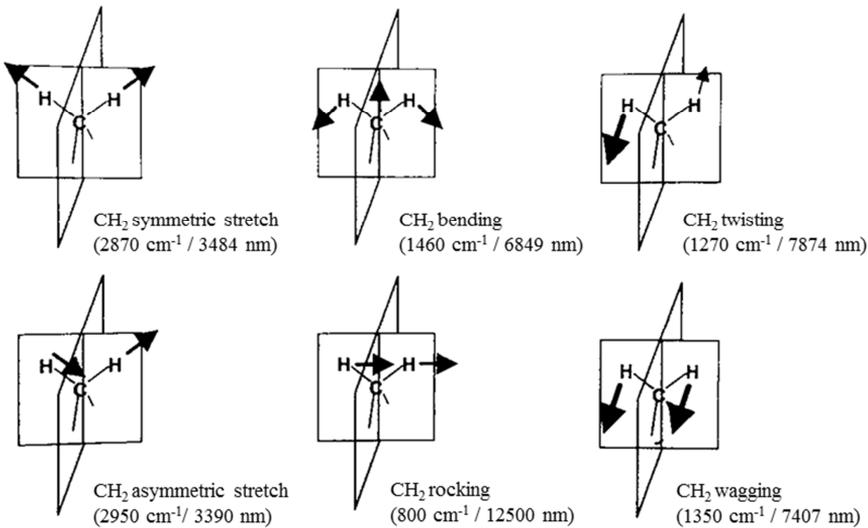


Figure 4.2: Vibrational modes of CH_2 -group, their common names and approximately frequencies are illustrated (Miller, 2001).

The overtones are formed when photons excite to the second or third energy level. But only those photons with twice or three times the amount of energy that was needed to elevate a molecule to the energy level corresponding to a fundamental absorption can excite to the second or third energy level, which cause the first and second overtone. The frequencies of the overtone bands are nearly equal to multiples of frequency of the fundamental vibrations (Table 4.1). Similarly, the vibrational combination bands appear when the absorbed photon excites two or more vibrations at the same time (Table 4.1). The energy of the absorbed photon has to be the same as the energies of the combined vibrations, in order to make the excitation of two or more vibrations happening. Also, the frequency of a combination band is nearly the same as the sum of the frequencies of the corresponding single vibrations (Miller, 2001).

Table 4.1: Vibrational combinations and overtone regions of CH bands

	CH	CH ₂	CH ₃
Combination bands	4400-4150 cm ⁻¹ 2273- 2410nm	4450-4200 cm ⁻¹ 2247-2381 nm	4515- 4220 cm ⁻¹ 2215-2370 nm
1 st overtone	6000-5600 cm ⁻¹ 1667-1786 nm	6200-5700 cm ⁻¹ 1613-1754 nm	6400-5900 cm ⁻¹ 1563-1695 nm
1 st overtone combinations	7100-6900 cm ⁻¹ 1408-1449 nm	7300-7000 cm ⁻¹ 1370-1429 nm	7400-7250 cm ⁻¹ 1351-1379 nm
2 nd overtone	8750-8000 cm ⁻¹ 1143-1250 nm	8800-8200 cm ⁻¹ 1136-1220 nm	8900-8400 cm ⁻¹ 1124-1190 nm
3 rd overtone	11700-11000 cm ⁻¹ 855-909 nm	11200-10700 cm ⁻¹ 893-935 nm	11000-10500 cm ⁻¹ 909-952 nm
4 th overtone	13500-13100 cm ⁻¹ 741-763 nm	13750-13250 cm ⁻¹ 727-755 m	14000-13500 cm ⁻¹ 714-741 nm

4.3 Sampling Modes

The NIRS instrumentation can be applied both in transmittance and reflectance mode, as illustrated in Figure 4.3. It should be noted that the mechanical pathlength is fixed in the transmittance mode (assuming no scattering takes place) while the effective pathlength (hence, the penetration/interaction depth of the light with the material) in the reflectance mode depends on the sample broth. For industrial process monitoring the NIRS instruments is conveniently applied with optical fibres, which are connected to an online probe. The reflectance mode can be applied in the process either by measuring through a window or by entering a probe into the system, whereas transmittance mode is only possible by using a probe in the process or a flow-loop (a so-called fast-loop) connected to the process.

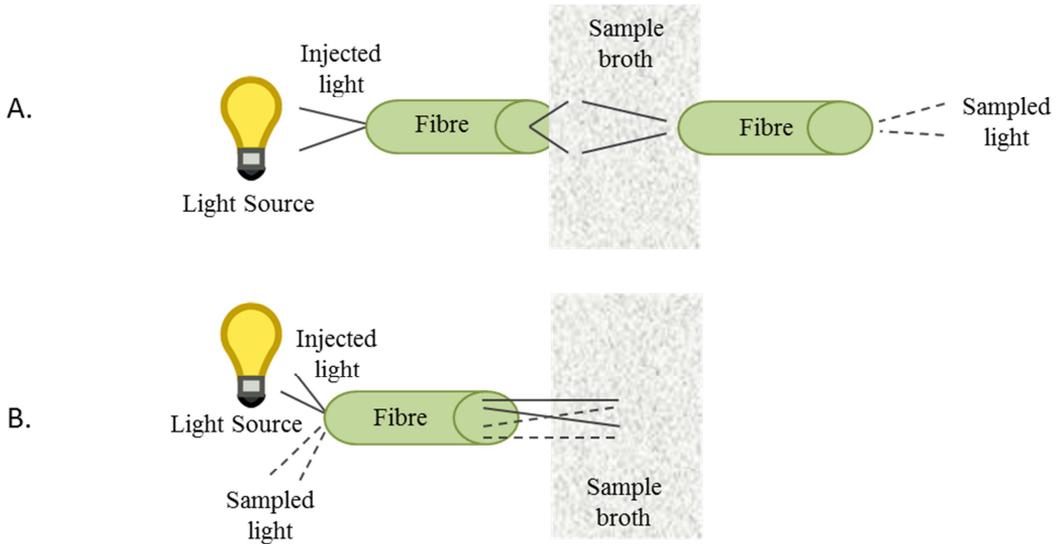


Figure 4.3: Illustration of Transmittance spectroscopy (A) and Reflectance spectroscopy (B).

Prior to Studies I and II, where in-line NIR spectroscopy was applied, transmittance and reflectance spectroscopy were applied simultaneously in a five hours yoghurt fermentation. The raw spectra collected during the fermentation process are shown in Figure 4.4.

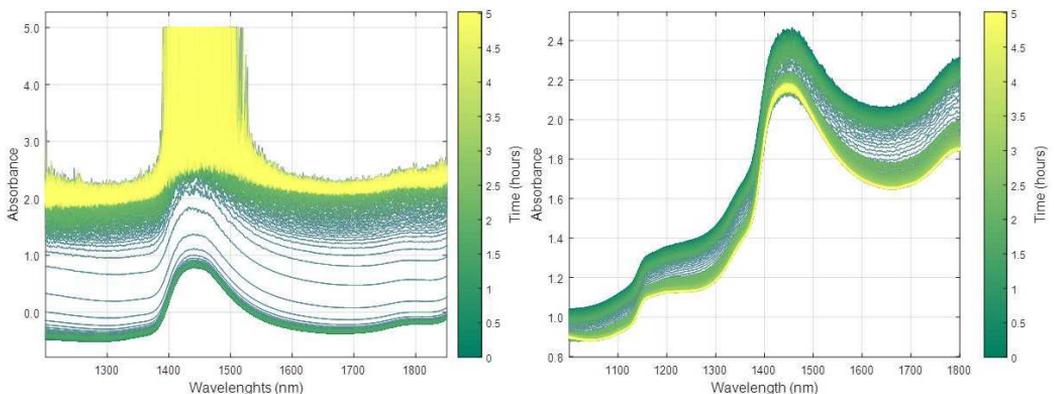


Figure 4.4: Raw NIRS spectra collected over time by transmittance spectroscopy, pathlength 1 mm (left) and reflectance spectroscopy (right) in a lactic fermentation process.

The spectra obtained by the transmittance probe become saturated after approximately one hour and 47 minutes of fermentation, which indicates that the broth turns into a more complex, opaque manner, probably due to a higher biomass concentration or due to the viscosity

change happening during the yoghurt fermentation process. A solution is to change the NIRS mode from transmittance to reflectance. A study by Crowley et al. (2005) investigated the use of transmission versus reflectance spectroscopy in a high cell density fed-batch industrial *Pichia pastoris* bioprocess. They observed that over a certain biomass concentration only a small amount of light reached the transmission detector, because most of it was reflected back due to the high biomass concentration. For that reason they concluded that transmittance was not suitable for high cell density bioprocesses, which correspond well with our observations from Figure 4.4. Additionally, a study by Aernouts et al. (2011) compared the transmittance and reflectance mode for determination of fat, crude protein, lactose and urea in raw milk. It turned out, that the best monitoring of fat and crude protein were obtained by the reflectance measurements, whereas the transmission mode provided the best predictions for lactose. None of the modes provided acceptable measures for urea. Furthermore, the authors concluded that transmittance was only suitable with a very small pathlength, due to high absorbance of the water band as well as strong light scattering from fat globules. This leads us to the assumption that the saturated spectra in Figure 4.4 both can be caused by an increased cell density and due to the high water content in milk in combination with scatter from the fat globules. However, a smaller amount of fat globules are expected in our skimmed milk yoghurt fermentation compared to raw cow's milk and it is expected that the size of the existing fat globules will remain constant throughout the acidification (Belitz, et al., 2004a). However, it must be remembered that during the acidification a gel network is formed, and this particle network might have a great influence on the spectral saturation observed. Thus, several related factors might cause the saturation. The study of Crowley et al. (2005) also suggested switching from transmittance to reflectance mode at the certain cell density threshold, where transmittance becomes inappropriate, which in general led to improved calibration models for the process control parameters. They outlined that it is important not to oversimplify the measured process by selecting one mode of NIRS and use this throughout the process. It can be discussed how convenient it is to switch from one mode to another during process monitoring. However, two probes, one transmittance and one reflectance, should then be applied, which means that more maintenance of the equipment and the models are required. Alternatively, a transreflectance probe (Von Barga, 1996), combining transmittance and reflectance, can be applied. For on-line measurements incident light is either transmitted through the sample and scattered back via a mirror or simply reflected back before it reaches the mirror, corresponding to transmission and reflectance components (Kawano, 2002). The approach is illustrated in Figure 4.5.

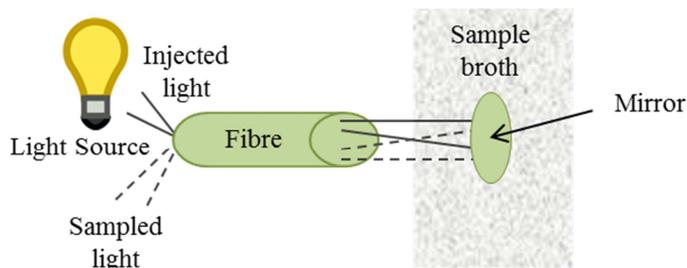


Figure 4.5: Illustration of transreflectance spectroscopy, where transmittance and reflectance are combined.

In-line monitoring of biomass, glucose and ethanol in both anaerobe and aerobic fermentation by transreflectance mode was applied by Princz et al. (2014). The in-line measurements were collected with a pathlength of 0.5 mm. They concluded that it was possible to obtain successful predictions, but they also outlined that high density of cells was problematic, without giving further details. Furthermore, the transreflectance mode for yoghurt fermentation monitoring has been studied by Grassi et al. (2013) using a Fourier Transform (FT)-NIR spectrometer equipped with a fibre-optic probe with 1 mm pathlength. The authors found that NIRS was capable of describing the curd development appearing during the acidification. Though the transreflectance mode seems to be suitable for on-line fermentation monitoring, it should be noticed that both of the studies mentioned applied a relatively small pathlength, which increase the chance to have unwanted bubbles or larger particles being trapped in the light path. Additionally, the transreflectance data might lead to more difficult data interpretation compared to transmission and reflectance spectroscopy. However, it cannot be excluded that the transreflectance approach might increase the versatility of the collected data.

4.4 Data Processing for NIRS Data

In order to explore and interpret the NIRS data, multivariate data analysis is applied. It is thereby possible to elucidate trends and relationships in the obtained spectra, which can be difficult or impossible to see in the raw spectra. When gathering NIRS spectra collected over time, the data structure becomes a two-way array, which is the simplest multivariate data arrangement. Such data can be arranged in a matrix (\mathbf{X}) with K variables and N objects (Smilde, et al., 2004a), where the objects can be the samples measured at certain time points and the variables can represent the wavelengths region at which the samples have been analysed. The most basic analysis among the multivariate analyses is Principal Component Analysis (PCA), which can be described as a projection method extracting the systematic variation found in \mathbf{X} . The matrix is decomposed into a sum of matrix products, where one matrix is called scores (\mathbf{T} of size PCs times N) and the other is called loadings (\mathbf{P} of size PCs times K). The variation not described by the conducted PCA model is found in matrix \mathbf{E} . In matrix notation PCA can be written as:

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad (4.7)$$

PCA is a bilinear method summarising the variation of the data matrix \mathbf{X} by creating new directions in the original data, which are constructed as linear combinations of the original variables (Wold, et al., 1987). The first created direction, also called the first principal component (PC1), describes the maximum variance found in the original data. The second direction, called the second principal component (PC2), is found orthogonal in respect to the first direction and this is continued as long as descriptive variation is found (Esbensen, et al., 2000). A two component model, which describes a two dimensional subspace or plane in the original K and N dimensional spaces, can be graphically illustrated as shown in Figure 4.6.

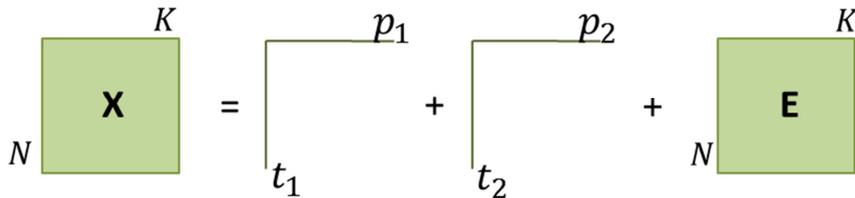


Figure 4.6: A PCA model decomposing \mathbf{X} into a score vector (\mathbf{T}), a loading vector (\mathbf{P}) and the residuals (\mathbf{E}).

where t_1 and t_2 represent the score vectors and p_1 and p_2 represent the loading vectors in PC1 and PC2, respectively. The score value for the first component is defined by the projection of the original position of a sample onto PC1 (specifically, the direction determined by the loading vector for PC1), and likewise the second score value for this sample is found by projecting onto PC2 (Eriksson, et al., 2006). The projection of observations principle in a two component model is illustrated in Figure 4.7.

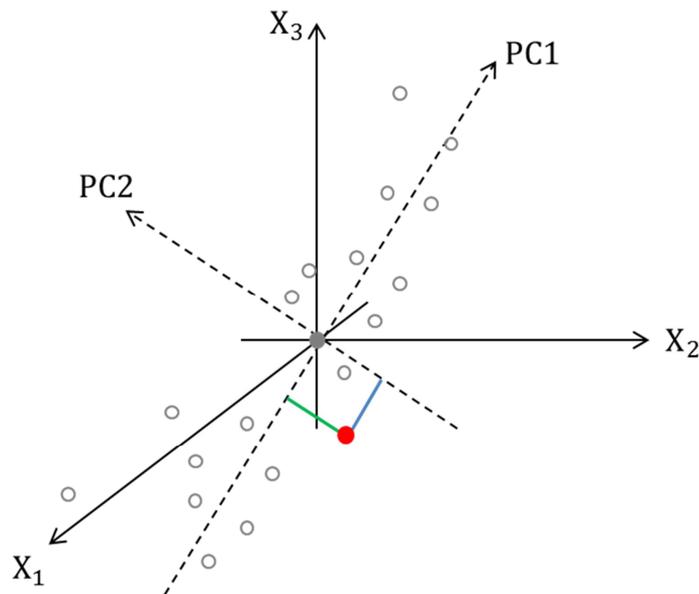


Figure 4.7: A two component model, where the first and second principle component (PC1 and PC2) pass through the average-point of the data cloud. The blue line represents the score value on PC1 and the green line represents the score value on PC2 for the projection of the observation highlighted in red.

The obtained scores and loadings can be visualised in a score plot and loading plot, respectively, where the score plot shows the similarity between the observations and the loading plot displays the relation between the variables and the role they play in finding the PCs (Eriksson et al. 2006a). Before the multivariate data analysis is carried out, it is important to investigate if pre-processing of the data is required. Spectral data can be pre-processed in several ways and available methods can be combined in numbers of strategies depending on the intention (Rinnan, 2014). In order to pre-process the data in a meaningful way, the first step is to inspect the raw data to highlight and visualize potential artefacts that are irrelevant for the analysis and can therefore safely be removed (e.g. a variable baseline). Spectral data are commonly mean centred, in which the average value of each variable is subtracted from the full data, variable-wise (Figure 4.8). When the data are mean centred the cloud of score points in a two dimensional subspace will have its centre in the point (0,0) where the PC's meet.

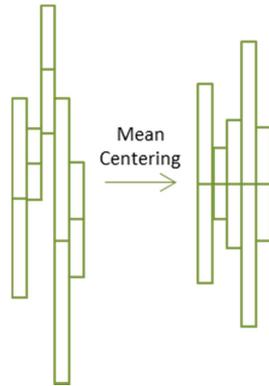


Figure 4.8: After mean centering all variables will have the same mean value at zero.

Further pre-processing can be carried out in order to correct for scatter. Scatter is often unwanted and it is removed in order to extract the chemical information of the spectra. A number of methods have been developed for scatter corrections, but two of the most applied scaling strategies are Standard Normal Variate (SNV) and Multiplicative Scatter Correction (MSC). In the SNV application the correction is applied to each spectrum individually:

$$\mathbf{X}_{\text{corr}} = \frac{\mathbf{X}_{\text{org}} - m_i}{s_i} \quad (4.8)$$

where \mathbf{X}_{org} is the original sample spectrum measured, \mathbf{X}_{corr} is the corrected spectrum, m_i and s_i are the mean and the standard deviation, respectively, of the spectrum to be corrected. Hence, the mean is subtracted from every data point of the sample spectrum and each point is divided by the standard deviation (Barnes, et al., 1989; Næs, et al., 2002). In Figure 4.9 raw NIRS spectra and SNV pre-processed spectra are plotted, respectively. The spectra were collected during a yoghurt fermentation of 5 hours by a NIRS equipped with an in-line reflectance probe (Paper II).

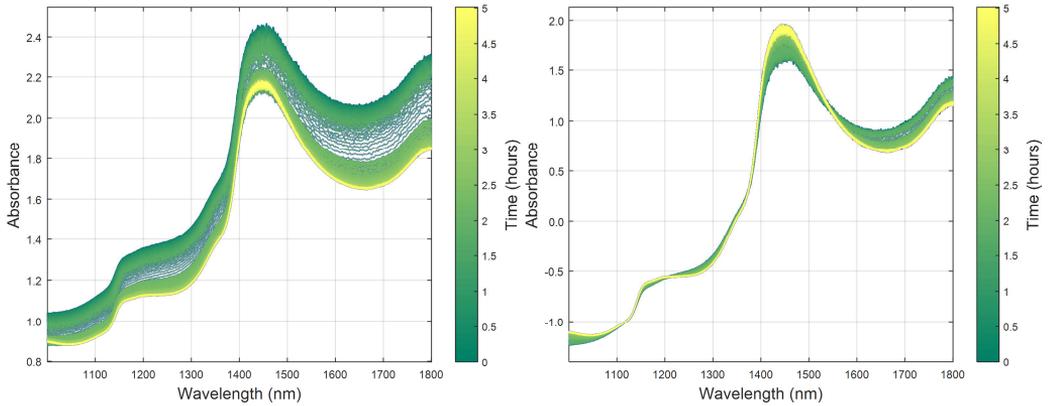


Figure 4.9: NIRS spectra obtained during a yoghurt fermentation: raw spectra (left), SNV pre-processes spectra (right).

When inspecting the raw NIRS spectra a large variation among the recordings is seen more or less for all wavelengths, however it also seems that the main variation is caused by an off-set, which is decreasing over fermentation time. From the SNV pre-processed data no such off-set is observed and the main variation is now observed around 1400 to 1500 nm, 1600 to 1800 nm as well as in the region around 1275 nm, 1175 nm and below 1075 nm.

The MSC pre-processing method, which originally was developed for NIRS data, affects the spectra in a very similar way as SNV. The MSC technique consists of two steps, the first of which is fitting all spectra against a common reference spectrum:

$$\mathbf{X}_{org} = a\mathbf{X}_{ref} + b \quad (4.9)$$

where \mathbf{X}_{ref} is defined as the reference spectrum and is often represented by the mean spectrum of all spectra in the data set. Furthermore, a is a multiplicative correction factor, b is an additive correction factor. From each spectrum a (the slope) and b (the intercept) can be defined and thereby the corrected spectrum \mathbf{X}_{corr} can be determined (Geladi, et al., 1985; Næs, et al., 2002):

$$\mathbf{X}_{corr} = \frac{\mathbf{X}_{org} - b}{a} \quad (4.10)$$

When applying MSC on the NIRS raw spectra shown in Figure 4.9 (left), a very similar output, to the once obtained by SNV, is obtained (not shown). Although the SNV and MSC techniques make different assumptions about the multiplicative variations in the spectral data, the tech-

niques perform similar. However, since the SNV is calculated for individual samples, where the spectrum is corrected due to its own mean value, and MSC is calculated data-set wise, where the spectrum is corrected due to the mean of all spectra in the data set, there might be cases where the two methods perform differently. Such cases could include situations where offsets among replicates exists, in which MSC would correct for such offset, but SNV would probably not. However, since on-line NIRS measurements do not provide us with replicates, it is not an issue to be considered here. The a and b values obtained by MSC from each spectrum, each representing a time point, can be graphically visualised as a function of fermentation time (Figure 4.10).

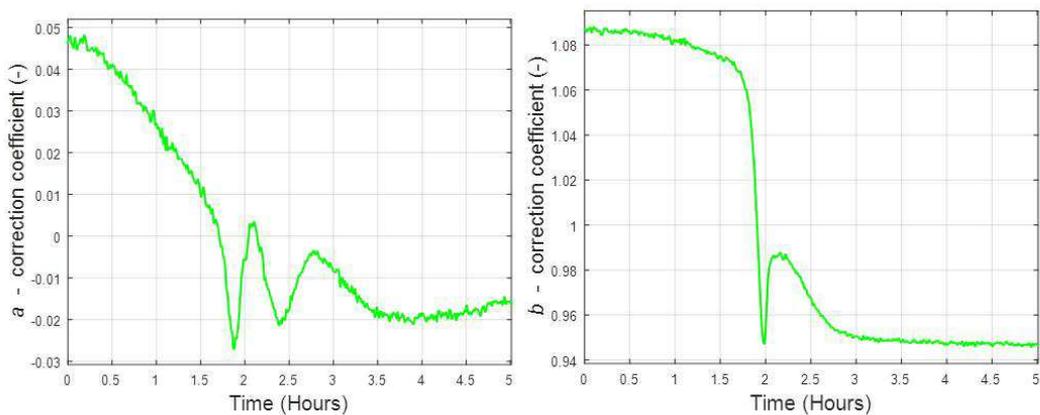


Figure 4.10: The correction values a (left) and b (right) for all spectra, each representing a time point, obtained from the MSC pre-processing method.

The a values represent the multiplicative correction on the spectrum and when plotting them versus fermentation time a dynamic trend, which is overall decreasing, is observed (Figure 4.10, left). The b values, representing the additive correction of the spectra, provide a much more steep decrease over time, illustrating that a fast change after two hours of fermentation time appears (Figure 4.10, right). As the offset most likely is physically related and probably caused by a scatter change in the fermentation, it leads us to the assumption that b seems to be related to the viscosity change, which is also discussed in Paper II. It could be discussed whether a is representing the chemical information, related to the growth or the pH drop. The drop seen for the b values is very similar to one of the kinetic profile described in Paper II, whereas a is less alike the second profile fitted in Paper II. Another option could be that a contains a combination of physical and chemical information, since the trend seems to be a combination of the two kinetic profiles, chemical and physical, fitted in Paper II. It is hard to tell and more investigation should be conducted before anything can be assumed with certainty. Nevertheless, this output for a and b was very representative for all fermentation batches included in Paper II.

In order to improve the separation of physical and chemical information an extended and modified version of MSC, denoted Extended Multiplicative Signal Correction (EMSC), was developed (Martens and Stark, 1991; Martens, et al., 2003).

$$\mathbf{X}_{\text{corr}} = a\mathbf{X}_{\text{ref}} + \mathbf{b}_H\mathbf{H} + \mathbf{c}_S\mathbf{S} + \mathbf{d}_P\mathbf{P} \quad (4.11)$$

where the matrix \mathbf{H} consists of different baseline profiles, such as offsets, linear baseline slope and curved baseline components, which is extracted from the data. The matrix \mathbf{P} also consists of profiles which are extracted and denoted as not of interest for the analytical investigation. On the other hand, \mathbf{S} consists of profiles, which are of interest and are kept in the data for further modelling. The advantages of EMSC compared to MSC are the ability to apply better estimates for baseline correction and removal of spectral artefacts, which are not of interest (Martens, et al., 2003; Miller, 2010). The EMSC method was also applied to our raw NIRS yoghurt spectra and the output is shown in Figure 4.11.

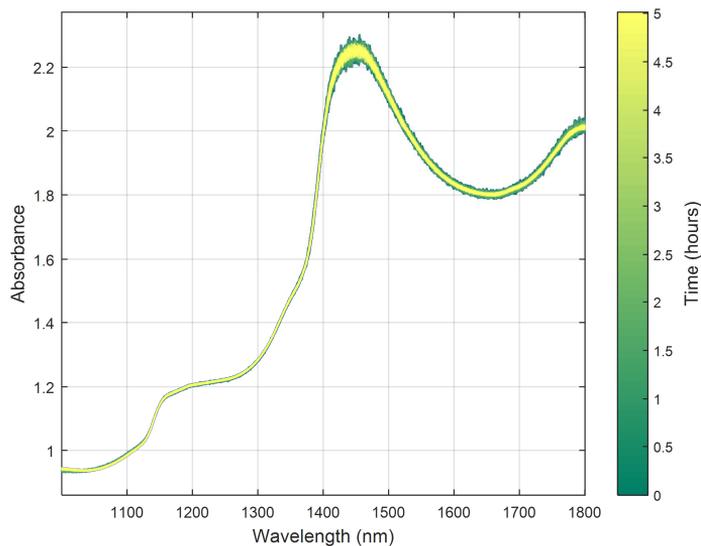


Figure 4.11: The yoghurt NIR spectra collected over time are pre-processed by the EMSC method.

The EMSC pre-processed spectra are quite different from the SNV (Figure 4.9, right) and the MSC pre-processed data, as no variation among the spectra seems to appear for the EMSC pre-processed data. When no variation is obtained, no process dynamics can be determined from further modelling. It basically seems that all information is lost, indicating that the main information existing in the NIRS data are caused by physical changes.

Finally, the Savitzky-Golay derivation (SG) method should shortly be mentioned when considering pre-processing methods for near-infrared spectroscopy. The method is used for smoothing and derivatizing a set of sample spectra. When applying the first derivative, offset variation among the sample spectra will be removed and by the second derivative possible slope effect will be removed. In order to calculate the derivative, a polynomial (of a specific order) is fitted on the raw spectra in a defined window. The method requires the user to define the window size, which is the number of points used to calculate the polynomial (Savitzky and Golay, 1964; Rinnan, et al., 2009).

4.5 Chemical and Physical Information

Chemical Information – based on absorbance

The average spectrum of all raw spectra in Figure 4.9 (left) was found in order to simplify the chemical interpretation. Because of the high signal of the overtone (1450 nm) of water, spectral bands related to other milk components are difficult to see. Since the derivative treatment is considered a pre-processing method for solving overlapping peaks issues (Laporte and Paquin, 1999), the data were treated by the Savitzky-Golay pre-processing method using a window size of 15, 2nd order polynomial fitting and the second derivative of the average fermentation spectrum (Figure 4.12). Thereby, the characteristic absorption peaks are more clearly separated and peak assignment can be carried out.

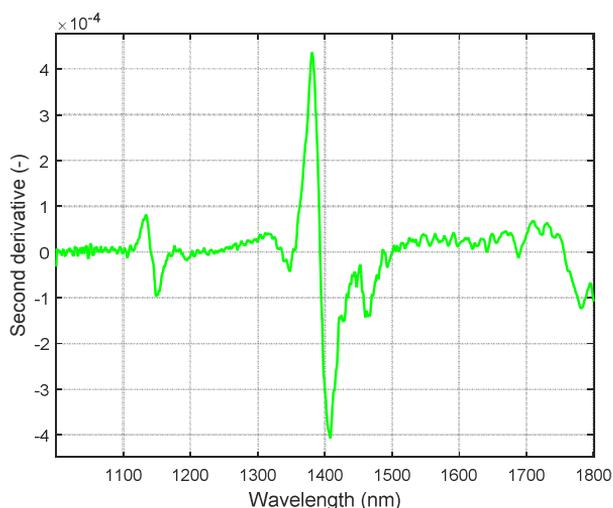


Figure 4.12: The second derivative of the average NIRS reflectance spectrum from the spectra shown in Figure 4.9

The study of Aernouts et al. (2011) has shown that the absorbance around 1960 nm in a reflectance spectrum could be related to the fat content. As our spectra were measured within the

range 1000-1800 nm, the 1960 nm peak cannot be assigned. Furthermore, the milk used for the yoghurt fermentation is based on a skimmed milk with low fat content, so even though a broader measurement range was applied, it would most likely be difficult to detect the fat. Furthermore, Aernout et al. (2011) reported that protein is associated with the region around 1650 nm. In addition it was not possible to predict lactose from the reflectance spectrum, but the study concluded that lactose predictions can be found by transmission between 1000 and 1700 nm, which include the first O-H stretch vibrations for sugars at wavelength 1490 nm. The study by Czarnik-Matusiewicz et al. (1999) assigned the first overtone of amide B at 1638 nm and the first overtone of amide A at 1584 nm in milk samples. By reflectance spectroscopy it is also possible to determine the composition of fatty acid acids in cow milk, which was shown by Coppa et al. (2010). Quite a big wavelength region from 700 to 2500 nm was applied, but the study demonstrated that it is possible to predict various fatty acid groups, including saturated fatty acids, monounsaturated fatty acids, polyunsaturated fatty acids, unsaturated fatty acids and conjugated linoleic acid as well as individual fatty acids in liquid and oven dried milk samples. Though, it has been reported that the C-H of lipids and the N-H of proteins can be detected by NIRS in the wavelength region equal to what has been applied in our study, these bands are hard to assign in our average spectrum. Though the aggregation of micelles followed by contraction of the aggregates during the acidification results in larger casein particles, it is hard to imagine that an actual variation of protein concentration should appear. Likely, it is not expected that the fat concentration should change during the fermentation process, meaning that no variation in these wavelength regions is expected. For that reason these wavelength are as such not of importance for yoghurt fermentation monitoring. On the other hand, lactose is expected to be consumed over time and could therefore be a suitable compound to measure, in order to monitor and control the lactic fermentation process. Both the C-H and the O-H bonds of lactose have been detected by NIRS (Wu, et al., 2008), however the absorbance are most commonly assigned in lower NIRS regions (<1000 nm) than what we have applied. Though, the first O-H stretch vibrations for sugars could be expected at wavelength 1490 nm (Aernouts, et al., 2011), it seems that region is found at a shoulder on the water peak and it is hard to see a clear variation both in the raw and the pre-processed spectra. In the loadings obtained by the PCA model the sugar region does not seem to have an influence here. Our spectra can in general be assigned as a classical water spectrum in which the overtone at 1450 nm (Laporte and Paquin, 1999) is very dominating. However, it seems that the scatter and the off-sets which are also very dominating in the obtained spectra could contain valuable information, which is elucidated in the following section on physical information.

Physical Information – based on scatter

In order to understand the scatter phenomenon it can be imagined that the light source entering the sample is divided into two fluxes, where one flux is radiation travelling through the

sample, which is related to the already described transmission technology and the other flux is radiation scattered backward meaning that the light is being remitted and this phenomenon is related to the reflectance technology. The reflectance principle is also related to the scatter phenomenon, which can be hard to separate from the reflectance technique, since the processes that cause scatter are essentially the same as those that cause the reflectance. Reflection and scatter can however be differentiated, since light is being reflected when the refractive index changes. This occurs when light meet the appearing molecules or rather their bindings. On the other hand scatter of light is affected by interactions among light and particle surfaces, and the surroundings will influence the light's interaction with an analyte (Næs, et al., 2002). In other words, scatter is a phenomenon taking place at an interface and the scatter from a particle is depending on its surface area and its refractive index. This means that smaller particles, having a larger surface area/volume ration than larger particles, provide more scatter per unit mass than larger particles (Dahm and Dahm, 2001).

The scatter is often considered as an unwanted phenomenon that complicates the spectroscopic measurements and interpretation of the results. Therefore the scatter is often reduced by various mathematical processing, as described in Paragraph 4.4. Despite this, a few studies confirm that scatter may deliver quantitative information. A study by Bogomolov et al. (2012) showed that scatter in the short wavelengths of the NIR region (up to 1000 nm) could provide quantitative analyses of fat and protein in bovine milk. The milk samples, which were varying in fat and protein content, were measured in transmission mode with a 5.5 mm pathlength. The collected spectra had a varying offset that seemed to be related to both the fat and the protein content; however fat and protein could still be separated from each other due to slope differences. A higher amount of fat globules leads to a higher amount of particles and thereby a stronger scatter is observed. This fits well with the fact that the offset differs in relation to the fat content. The sample spectrum with the highest fat content revealing most scatter has the lowest offset, whereas the sample spectrum with the lowest fat content and the weakest scatter has the highest offset. This observation is related to our data, where the sample spectra collected in the end of the fermentation, expected to provide more scatter, have the lowest offset (Figure 4.9). Bogomolov et al. (2012) further concluded that the best predictions for fat and protein were obtained without any pre-processing of the spectra. Though the authors partly conclude that scatter in the low wavelengths NIR region can be applied for fat and protein determination, they narrowed down the wavelength region to the VIS region within 600 to 700 nm, in order to obtain better predictions. In addition to this sample set, Bogomolov and Melenteva (2013) introduced the variation of fat globules size by systematically varying the homogenization degree of the milk samples. As add on to the previous study they concluded that it also was possible to monitor fat globule sizes by scatter in the low wavelength region of NIRS.

4.6 Study II – Exploring Process Dynamics by NIR Spectroscopy in Lactic Fermentations

This section is mainly based on Paper II “Exploring process dynamics by near infrared spectroscopy in lactic fermentations”, where further thoughts and discussions are presented. Additional results not included in the manuscript are also presented to elucidate some clarifications.

Aim

The aim of the study is to explore the process dynamics in yoghurt fermentations by the on-line NIR spectroscopy. The industrial yoghurt production is nowadays only monitored by pH, but additional process information might be gained by applying NIR to such systems.

The Experimental Set-up

A total of seven lactic fermentation batches were conducted for this study. They were carried out at different temperatures, in which one batch was conducted at 32 °C, four at 35 °C and two at 37.5 °C. The model system used is the same as the system applied in the previous study (Paragraph 3.3), where the lactic acid bacteria *L. bulgaricus* and *S. thermophilus* were applied. In addition to pH, Brix and NIRS measurements, samples were in this study withdrawn during the fermentation process in order to determine cell growth and lactic acid concentrations off-line. The cell growth was measured by qPCR analyses, whereas the D- and L-lactic acid concentrations were found by using an enzyme kit in combination with UV spectroscopy. Further details are given in Paper II.

Results

Due to the variable fermentation temperatures various trends were observed from the batches. A slower drop in pH was observed for the batches with lower temperature, whereas as a faster drop in pH was observed for the batches with a higher temperature. The same output was seen in the modelled NIRS data, where the process dynamics were delayed for the batches with lower temperature and while the dominant changes for the batches with a higher temperature were taking place earlier (Figure 4.13).

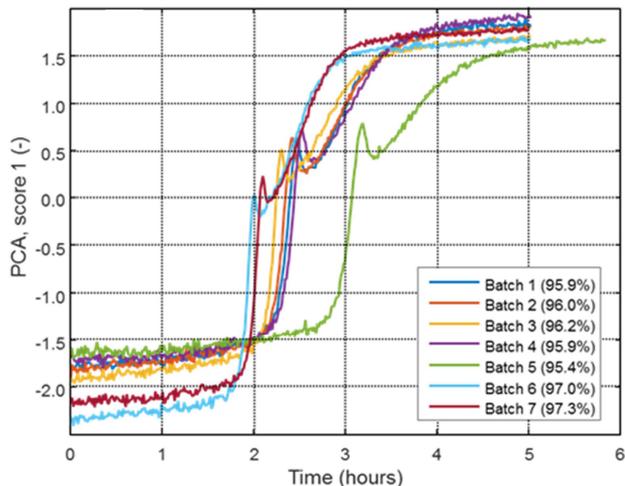


Figure 4.13: The first principal component of the modelled NIRS data are shown for the seven batches. Batch 1-4 had a fermentation temperature of 35°C, Batch 5 was carried out at 32 °C and Batch 6 -7 at 37.5 °C. The explained variance is given in the brackets.

To further explore the dynamics explained by NIRS kinetic profiles were fitted to the PCA profiles. The bump on the first principal component profile (Figure 4.13) is hard to describe by a kinetic model, and this part was thus excluded. This is illustrated in Figure 4.14c, where only the dark blue part was used for the kinetic modelling. A separate kinetic profile was fitted for the first part of the PC1 profile (green line in Figure 4.14c) and for the last part (black line in Figure 4.14c).

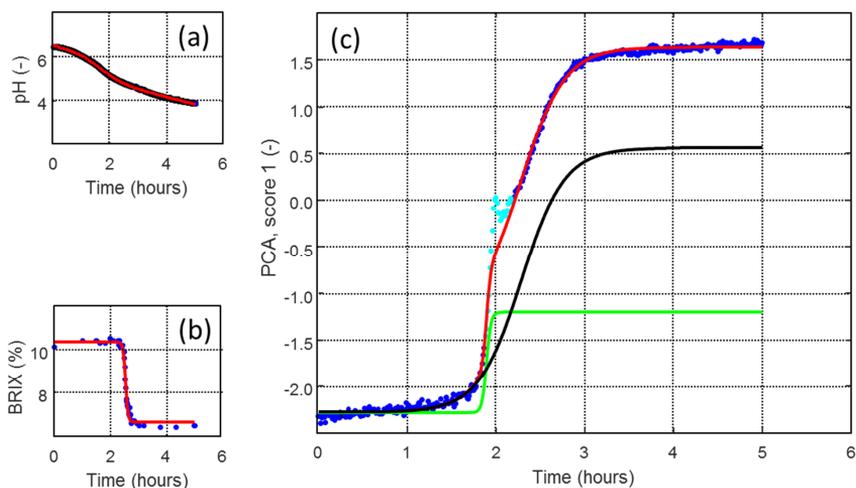


Figure 4.14: The profiles for pH (a), Brix (b) and PCA scores on PC1 of the modelled NIRS data in cyan and dark blue dots (c) are plotted. The red lines (a, b and c) represent the kinetic fit of the data points. The green and black lines (c) represent the kinetic fit on the first and last part, respectively, of the dark blue PC values. The light blue values were excluded from the kinetic fittings.

The first part is very similar to the Brix values (Figure 4.14b), whereas the last part is more similar to the S-curved pH drop (Figure 4.14a).

Additional Results, Discussion, Concluding Remarks and Perspectives

Near Infrared spectroscopy has previously proved to be ideal for analysing the chemical properties of many different sample types, both liquid and solid samples. Its high performance has led to the capability of measuring highly scattering and complex matrices, though it seems that the majority of NIRS studies on milk analyses are based on short wavelength NIR region in combination with the visible region. In this region several milk components have been identified and successful predictions have been made. However, this study confirms that valuable process parameters can be obtained in the wavelength region from 1000 to 1800 nm. The preliminary results showed that near infrared transmission spectroscopy might be applicable for on-line measurements when a very small pathlength is used. However, this may give rise to various issues with bubbles and the fact that yoghurt becomes more viscous, which potentially can get stocked in the small pathlength. These issues can be a huge disadvantage for the quality of the measurements. It is therefore suggested to use reflectance measurements when working with a high cell density or very viscous processes. From the literature various examples are found where the main components fat, protein and lactose are determined (Aernouts, et al., 2011). A few studies also suggest that information can be gained from the scatter (Bogomolov, et al., 2012). However, to our knowledge not many studies manage to achieve both chemical and physical information from the same NIRS data set.

The presented results indicate that both chemical and physical information are obtained by NIRS. As illustrated, it was hardly possible to differentiate the physical and chemical information by any of the applied pre-processing methods. However, the MSC pre-processing method did give us the hint that more information than the off-set scatter could be found, in which the b-value was very similar to the Brix-values, whereas the a-value seemed to contain some chemical information. Scatter is in many situations unwanted and various pre-processing techniques for removing all existing scatter has been developed, but in this study removing scatter was not an optimal solution.

PCA is normally applied to extract the systematic variation and ease the interpretation of these and an ideal multivariate model would be capable of describing one kinetic profile per principal component. It is therefore notable that both kinetic profiles, in this study, are described by the same principal component (PC1) and it can be considered if any model decisions could have improved the model. Before the chemical rank of the PCA model was decided, a few models with different ranks were investigated. Though the number of components was increased it was not possible to separate the two different kinetics profiles from each other at higher PCA levels. This study suggests how to separate two kinetic profiles, described by the same principal

component, from each other, though an ideal model would describe one kinetic profile per principal component, in order to ease the interpretation. It can be speculated why PCA is not capable of separating the two dynamic profiles from each other, but it very much seems that the two profiles behave very similar, hence the variation is changing within the same pattern, which cannot be distinguished by the model. It could be considered whether other multivariate models are capable of separating the dynamics and thereby simplifying the interpretation of the chemical and physical information. Multivariate Curve Resolution (MCR) was designed to separate chemical analytes from each other based on for example unique chemical spectra or distinguishable time-trends (Tauler, et al., 1993). Also, MCR often applies non-negativity, which can be an advantage when interpreting chemical trends. It could be considered whether the physical information could be assigned to a specific wavelength and likely whether the chemical information could be assigned to another specific wavelength region, which thereby would force the separation of the two dynamic trends. This means that we again have to combine a multivariate model with manual constructions. It is however hard to really assign a specific part of the spectrum, as the PCA loadings showed that the whole spectrum had an influence on the explained variance. After all there is a good reason why the two profiles cannot as such be separated, and that it because they are strongly correlated and one effects the other. This supports the facts that the cell growth leads to an increase of the lactic acid production and the acidity is causing the viscosity change, meaning that the chemical and physical variables are closely connected.

By applying reflectance spectroscopy instead of transmittance it was possible to measure the dynamics throughout the fermentation process. And as long as the measurements did not become saturated, it seems that valuable information can be obtained. Apart from the data presented in Paper II, OD₆₀₀ measurements were collected by an on-line UV instrument (Elution 220 UV-visible spectrometer, Thermo Scientific, Denmark) during the fermentation process. Optical Density (OD) is a standard indicator for cell growth and therefore a good gauge for whether the fermentation process is progression or not (Sonnleitner, et al., 1992). The UV-results shown in Figure 4.15, turned out to become saturated in the same way as the transmittance NIRS measurements. However, further interpretation of the UV-data indicates that two different kinetic profiles might also be observed here in which the first part might represent the cell growth. But after 2 hours, where the gel formation starts, it is texture being described. Around 2.5 hours the texture has become too viscous and no signal can be achieved with this measurement system. By OD we are thus not able to monitor the cell growth after 2 hours due to the textural change, unless we fit kinetic models to the data in the same way as the kinetic profiling strategy carried out for the NIRS data.

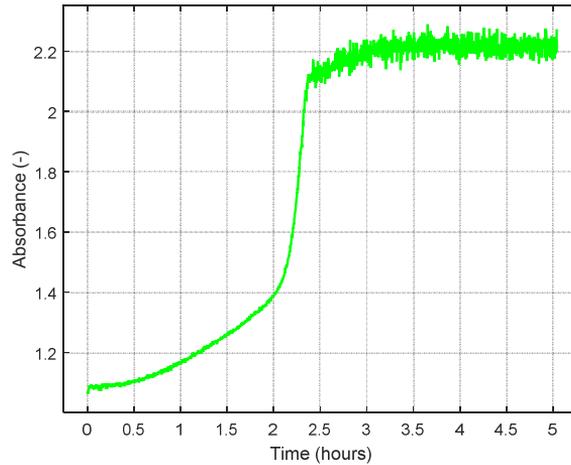


Figure 4.15: Optical density measurements (OD_{600}) versus fermentation time.

The Brix analysis, or more general refractive index, is also an example of a measurement technique which can be interpreted wrongly. The Brix values are commonly applied for sugar determination in the food industry, especially in the wine industry (Nagodawi.Tw, et al., 1974). It is a scale which is expressed in degrees and it measures the percentage by weight of sugar in water at a given temperature (Schaschke, 2014). In Paper I we trusted the Brix values, which are shown in Figure 4.16, and the Brix data was presented as being representative for the sugar consumption. However, the Brix measurements are collected by a refractometer, which operates due to the refractive index. Hence, there is a chance that scatter effect will influence the actually Brix degree measured. In Paper II it is assumed that the Brix values are correlated to the gel formation and due to the scatter effect observed by NIRS, we also assume that the Brix must be affected by the scatter to some extent. Furthermore, we know from the sugar metabolism that lactic acid is being produced as sugar is being consumed, which means that a more comparable decrease pattern of the sugar/Brix relation to pH than the one seen in Figure 4.16 is to be expected.

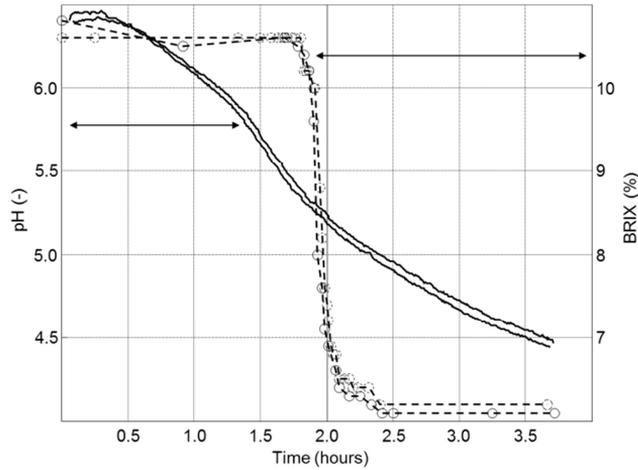


Figure 4.16: The Brix and pH values versus fermentation time, which are also presented in Paper I (Svendsen, et al., 2015).

This leads us to the assumption that the gel-formation also seems to be manifested in the Brix values. From these examples we have learned that measurement techniques based on refracting index must be interpreted carefully and it is of great importance that you are familiar with the measured process in order to gain the right process understanding.

FLUORESCENCE SPECTROSCOPY FOR ON-LINE MONITORING OF FERMENTATION PROCESSES

5.1 Introduction

Fluorescence spectroscopy has over the past 30 years been increasingly applied in biological sciences. It is known for being highly sensitive as well as specific, and is therefore a suitable spectroscopic method for quantitative determination of fluorescence compounds (Lakowicz, 2006a). Additionally, fluorescence spectroscopy has shown promising results within on-line monitoring of various fermentation processes (Lantz, et al., 2006; Teixeira, et al., 2009).

In this chapter the fundamental of fluorescence spectroscopy will shortly be presented. Next, various parameters affecting the fluorescence measurements will be examined and multivariate data analysis for fluorescence data is described. Finally, Paper III on “Weighted PARAFAC and non-linear regression for handling intensity changes in fluorescence spectroscopy caused by pH fluctuations” is introduced and further discussed.

5.2 Fundamental of Fluorescence Spectroscopy

Luminescence, which can be divided into phosphorescence and fluorescence, is the emission of light from any substance, and occurs from electronically excited states. The deviation of the two categories is depending of the nature of the excited states. Phosphorescence represents the emission of light from excited triple states, whereas fluorescence represents the emission of light from excited single states (Lakowicz, 2006a). The processes that happen between the absorption and emission of light are commonly illustrated by a Jablonski diagram as shown in Figure 5.1.

In the Jablonski diagram, S_0 , S_1 and S_2 represent the ground, first and second electronic singlet states of the molecule, respectively. At each of these levels the fluorophores can exist in various vibrational energy levels, illustrated as 0, 1, or 2 (Figure 5.1). The energy of the first excited triplet state (T_1) is normally lower than the energy of the first excited singlet state (S_1). The excitation process of the molecule occurs via absorption ($h\nu_A$) either from S_0 to S_1 or from S_0 to S_2 . No matter if the excitation process results with the molecule being in the first or the second excited single state, the molecule can end up at any of the excited vibrational states. The excited molecule can return to its ground state via various combinations of energy steps. Two of

these steps are fluorescence ($h\nu_F$) and phosphorescence ($h\nu_P$) which both involve the release of a photon of radiation. However, the fluorescence step (and also the phosphorescence) can only occur from the lowest vibration energy level of either S_1 or S_2 . If the molecule is not situated at the lowest level, a vibrational relaxation will occur, where the additional vibrational energy is lost. Thus, the molecule ends up in the lowest energy level. An expanded version of these vibrational relaxations is the internal conversion, where a vibrational energy lost results with the molecule passing to a lower energy electronic state. Finally, it is to be mentioned that an intersystem crossing process can appear, where the spin of an excited electron is reversed, and the molecule is transferred from S_1 to T_1 (Skoog, 1998; Lakowicz, 2006a).

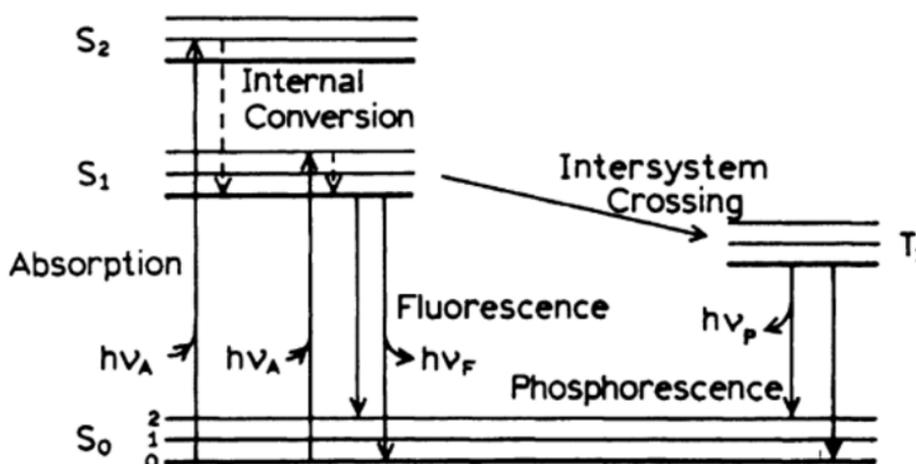


Figure 5.1: The Jablonski diagram (Lakowicz, 1999).

The Jablonski diagram describes the absorption (the excitation energy) and the fluorescence (the emission energy), but the actual relation between the excitation and emission spectral output is presented in the following. A fluorophore is normally excited from the singlet ground state to higher vibrational level of S_1 or S_2 and the return to the ground state typically occurs to a higher excited vibrational ground state level. Hence, more energy is needed for radiation to be absorbed by the molecule than what is emitted by the molecule and therefore molecules absorb radiation at lower wavelengths than the radiation they emit (Lakowicz, 2006a). In some cases the distance between the vibrational levels for the excitation and emission process are roughly the same and if the transition probabilities are similar, the emission spectrum will approximately be a mirror of the excitation spectrum (Harris, 2010).

Fluorescence spectral data can be presented as an excitation spectrum, where the excitation wavelengths are varied and the emitted light is measured at one certain wavelength (λ_{em}). It

can also be obtained as an emission spectrum, where the emitted radiation is measured based on a specific excitation wavelength. Such an emission spectrum is illustrated in Figure 5.2, where one excitation wavelength (310 nm) is applied and an emission wavelength range from 350 to 590 nm is shown for a vitamin solution.

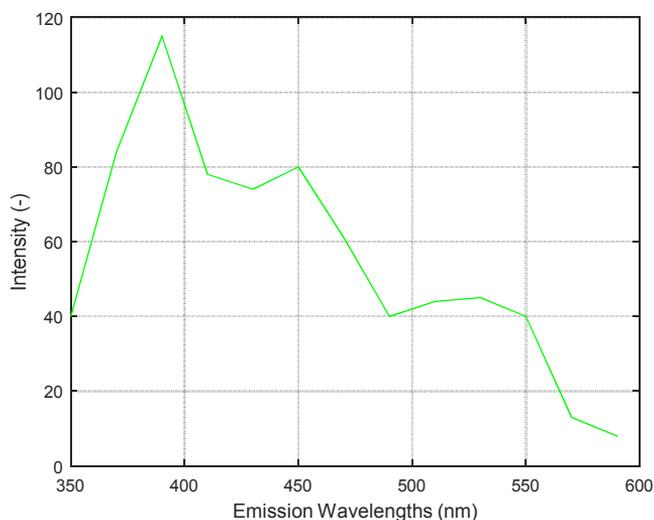


Figure 5.2: An emission spectrum (350-590 nm) of a vitamin solution measured at excitation wavelength 310 nm.

If several combinations of excitation and emission wavelengths are applied this results in a 2D landscape (Figure 5.3), where the emission and excitation wavelengths represent mode one and mode two, respectively (Lakowicz, 2006a; Harris, 2010). Compounds that fluoresce strongly and thus give the most intense signals are those containing aliphatic and alicyclic carbonyl structure or highly conjugated double-bond structures (Skoog, 1998). This means that intensities of different fluorophores within one mixture cannot be directly compared. Nevertheless, if a sample or a process is measured over time the relative intensity measured can still provide us with e.g. dynamics, from where it can e.g. be determined whether a compound is increasing, constant or decreasing during processing.

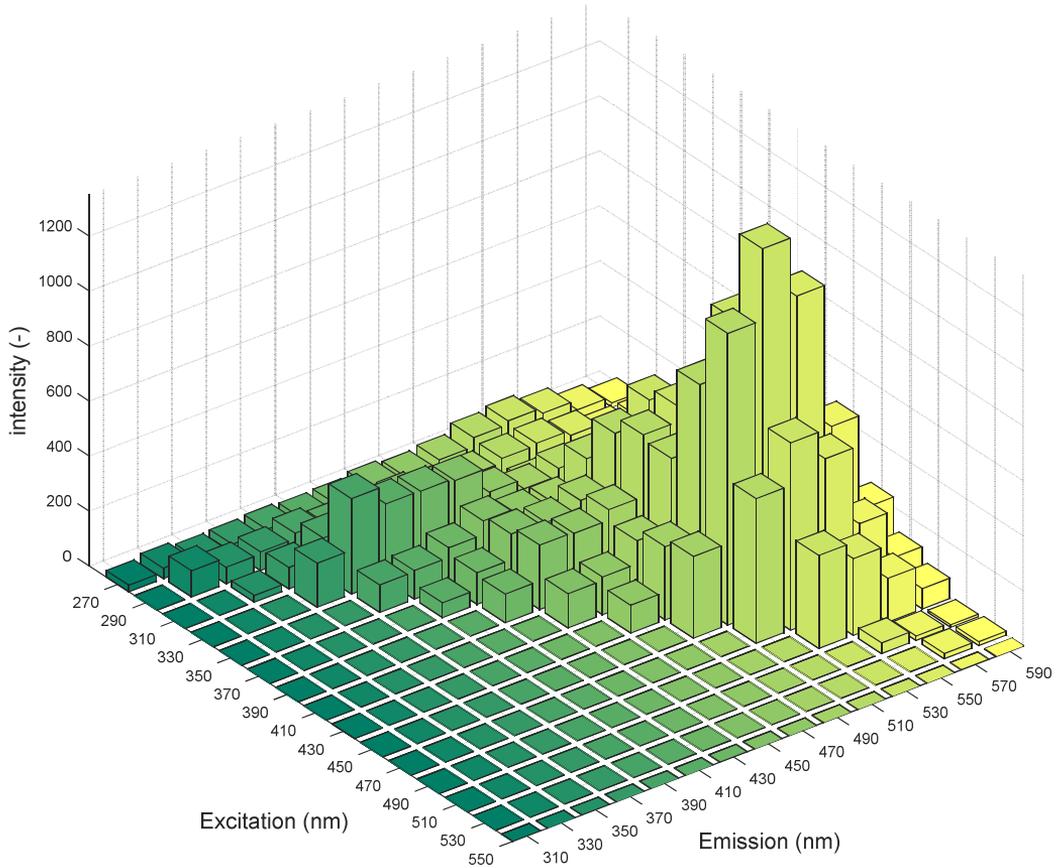


Figure 5.3: A 2D Fluorescence landscape of a vitamin solution.

If a process is measured over time by fluorescence spectroscopy (as in Paper I) a number of 2D fluorescence landscapes will be obtained. In order to follow a specific compound over time, the intensity of a selected peak with a certain excitation- and emission wavelength can be plotted as shown in Figure 5.4. Hence, the fluorescence intensity of that fluorophore can be determined during process- or reaction time. If the fluorescence spectrometer is equipped with an optical fibre, instead of the classical sampling technique by cuvettes, on-line monitoring can be applied, which allows us to follow the fluorescence intensity in real-time (Sablinskas, et al., 2003).

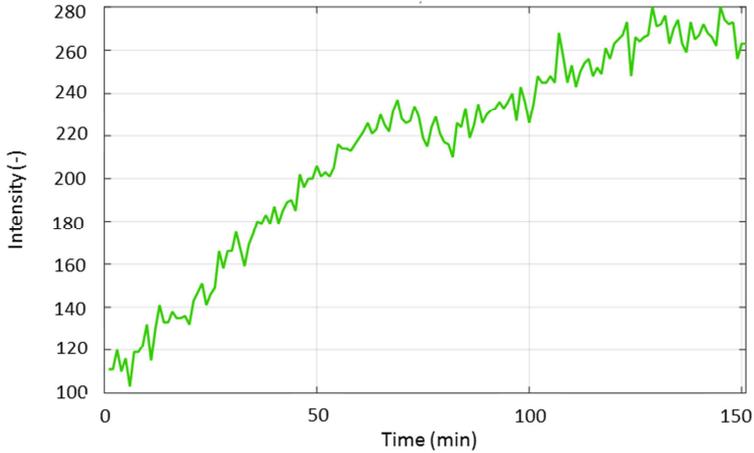


Figure 5.4: The fluorescence intensity of a vitamin solution followed over time at excitation wavelength 400 nm and emission wavelength 450 nm.

The definition of fluorescence intensity (I_f) of a molecule can be described by the intensity (I_a) of absorbed light and the quantum yield of fluorescence (ϕ_f) of the specific molecule:

$$I_f = \phi_f I_a \quad (5.1)$$

where I_a can be derived from Lambert Beer's law for absorption of light ($A = \varepsilon \cdot C \cdot l$):

$$I_a = I_0 - I_t = I_0 (1 - 10^{-\varepsilon Cl}) \quad (5.2)$$

in which I_0 and I_t represent the incident and transmitted light intensities, ε is the molar absorptivity, C is the concentration of the absorbing compound and l is the optical depth of the sample, which in our case is the fermentation broth. Finally, the quantum yield of fluorescence (ϕ_f) from Equation 5.1, which is the fluorescence efficiency, can be defined by following equation:

$$\phi_f = \frac{k_f}{k_f + k_d} \quad (5.3)$$

where k_f is the rate of fluorescence and k_d is the competitive deactivating processes appearing in the measured system (Schulman, 1985).

An important feature of fluorescence is the high sensitivity, which means that reliable detection of fluorescent materials is possible even at small concentrations. The technique can also be highly specific because a limited number of molecules absorb and re-emit light, and even though other compounds in the measured sample broth do absorb and emit light, it is extremely unlikely that compounds fluoresce at the exact same wavelength combinations (Herman, 2000).

5.3 Factors influencing the fluorescence intensity signal

Numerous parameters can affect the fluorescence intensity signal. As the concentration of the fluorescence compound is assumed to be exclusively related to the intensity of the fluorescence signal during quantitative investigations in accordance with Equation 5.2, environmental disturbances affecting the intensity can have a major influence on determination of the concentration. Therefore, it is of great importance to be aware of these influencing factors and the major ones are presented below.

Quenching

The term fluorescence quenching covers any process that decreases the fluorescence intensity of a sample (Lakowicz, 2006c). In a fluorescent system an external molecule can act as a quencher and reduce the fluorescence intensity. As a result the quantum yield, which is the number of photons emitted, is reduced or in some situations eliminated. Various molecular interactions can result in quenching, but the quenching process is normally classified into two main categories: static or dynamic (Albani, 2007b).

The static quenching refers to complexations between an interfering species and a potential fluorophore in the ground state. The formed complex is non-fluorescent and e.g. the quenching of the fluorescence of salicylic acid by complexation with iron (III) is an example. Dynamic quenching, also called diffusional quenching, is characterized by interactions between the quencher and the fluorophore of interest that appears subsequently to the excitation process and during the lifetime of the excited state. Due to the interactions, the excited molecule becomes inactive and returns to the ground state without emission of a photon. The quenching of fluorescence of a potential fluorophore by dissolved oxygen is an example of dynamic quenching (Sharma and Schulman, 1999; Lakowicz, 2006c). Since available oxygen is required in aerobic fermentation processes, there is a risk that oxygen will make fluorescence measurements challenging due to the possible quenching process.

Temperature

Some fluorophores can be temperature dependent, which means that a change in temperature can affect the fluorescence lifetime and intensity. Since the temperature can change the energy

levels of the ground and excited states of the fluorophore the fluorescence intensity can be affected (Li and Humphrey, 1992). In other words, a temperature variation can cause complete or partly de-excitation of the fluorophore. The so-called rate constant (k_i) is related to the de-excitation due to the temperature effect. The k_i -value's relation to temperature can be described by the Arrhenius theory:

$$k_i = Ae^{\left(\frac{-E}{RT}\right)} \quad (5.4)$$

where A is the temperature independent pre-factor (s^{-1}), E is the Arrhenius activation energy ($kcal \cdot mol^{-1}$), R is the molar gas constant ($J \cdot K^{-1} \cdot mol^{-1}$) and T is the temperature (K). This means that a temperature increase leads to an increase of the k_i -value. As the relationship between k_i and the fluorescence lifetime (τ_0) can be described as;

$$\frac{1}{\tau_0} = k_r + k_i \quad (5.5)$$

where k_r is the radiative constant, the relationship between the temperature and the fluorescence lifetime (τ_0) can thus be described as

$$\frac{1}{\tau_0} = k_r + Ae^{\left(\frac{-E}{RT}\right)} \quad (5.6)$$

Since the radiative constant is at least 10 times smaller than $\frac{1}{\tau_0}$, it only has a minor influence on the relation between the temperature and the fluorescence lifetime (Albani, 2007b). From Equation 5.6, it is shown that a temperature increase leads to a decrease of the fluorescence lifetime. Though fluorescence intensity usually decreases when the temperature increases, for some fluorophores the fluorescence intensity actually increases with temperature. This might be explained by an increase in the energy level of the excited state (Li and Humphrey, 1992).

Solvent interaction

The polarity of the solvent has a large influence on the emission of the fluorophore. The interactions between the solvent and the fluorophore happen via electrostatic interactions and hydrogen bonds. It is usually the difference between the energies of the ground state and the excited state, which contribute to the intensities and spectral position of the fluorescence signals. In the case where the solute has a greater polarity in the excited state than in the ground

state, the fluorescence appears at longer wavelengths in a solvent with high polarity. This happens because a solvent of high polarity will stabilize the excited state to a higher degree than in the ground state. However, if the polarity is lower in the excited state than in the ground state, a solvent of high polarity will stabilize the ground state to a higher degree than the excited state. This situation tends to cause a shift to shorter wavelengths of the absorption (Schulman, 1985; Sharma and Schulman, 1999). These shifts are also known as red and blue shifts, where red shifts refer to a spectral shift towards higher wavelengths and blue shifts refer to a spectral shift towards lower wavelengths (Lakowicz, 2006b).

Inner filter effect

Under optimal circumstances the fluorescence intensity is linearly proportional to the concentration of the determined fluorophore (Equation 5.2). But at a certain point this linearity ends due to the inner filter effect. Under non-linear conditions the incident light is absorbed by other species than the fluorophore (the primary inner filter effect) and the emitted light is re-absorbed (the secondary inner filter effect). The Inner filter effect term refers to a high optical density, which causes a decrease of the fluorescence quantitation, in which the emission peak might be shifted and a decrease in the fluorescence intensity can be observed (Kubista, et al., 1994; Albani, 2007c). When applying fluorescence spectroscopy in bioprocesses, the inner filter effect can easily be a problem due to the increased biomass, which means an increase in the optical density. But also larger molecules or particles can be a reason why emitted light can be reflected or scattered (Li and Humphrey, 1992).

pH

The pH value of the measured sample or fermentation broth, which is a major subject in the presented research, can have a strong effect on the fluorescence intensity. Since the electronic distribution of acids and bases vary from the excited state to the ground state, the acidity and basicity for the same molecule might be different in these two states (Sharma and Schulman, 1999).

When the species of interest has been excited to the S_1 state, it may happen that the electronic charge of an acidic or basic functional group will change in distribution. Hence, the acidity of the functional group will change, and the acidity of the same functional group will differ in S_1 compared to S_0 . Normally, such changes in electronic charge distribution will only occur to functional groups bonded directly to an aromatic ring. When the change occur it is sufficient to cause distinction between the pK_a -value in S_0 and in S_1 (Valeur, 2001).

If a protonation of the functional group appears during the transit from S_0 to S_1 , the energy difference from S_0 to S_1 will decrease. This energy change will result in a shift in the spectral data to longer wavelengths. On the other hand, a dissociation process will result in a shift to

shorter wavelengths (Sharma and Schulman, 1999). Change of pH in the analysed sample broth does not only cause shifts in the fluorescence spectrum, it can also cause changes in the fluorescence intensity signal. An examples is the compound phenol (C_6H_5OH), which fluoresces in the wavelength region 285-365 nm with a relative intensity of 18, whereas the ionized phenolate ion ($C_6H_5O^-$) fluorescence in the wavelength region 310-400 nm with a relative intensity of 10 (Skoog, 1998).

So far this intensity change phenomenon is mostly described for proteins. The complex tertiary structure of proteins is altered when the protein is dissolved in a solution with a pH far away from its own physiological pH. The alteration of the protein can either be a partial or full denaturation, in which the tertiary structure is unfolded. Since the characteristics of an unfolded protein differs significantly from an non-unfolded protein, the fluorescence intensity, among other fluorescence emission parameters, changes (Albani, 2007a).

A few studies have modelled the effect of pH on fluorescence, where it is assumed that the intensity of a fluorophore is proportional to its concentration, and the total fluorescence (F_t) is the sum of the fluorescence from the alkaline form (A^{-1}) and the acidic form (HA), as described in the following equation:

$$F_t = K_1[A^{-1}] + K_2[HA] \quad (5.7)$$

where K_1 and K_2 are constants that are proportional to the incident light intensity, quantum efficiency, molar absorption and the light path (Guilbault, 1990; Li and Humphrey, 1992). If the fluorophore of interest only exists in the forms A^{-1} and HA , the total concentration (C) of the fluorophore can be define as:

$$C = [A^{-1}] + [HA] \quad (5.8)$$

In this way the Henderson-Hasselbalch equation (Po and Senozan, 2001) can be applied in order to relate the concentrations of A^{-1} and HA with the pH as illustrated here:

$$pH = pK_a + \log \frac{[A^{-1}]}{[HA]} \quad (5.9)$$

where pK_a is the dissociation constant of the fluorophore. By combing Equations 5.7, 5.8 and 5.9 the total fluorescence of a fluorophore exciting on both its acidic and basic form can be defined.

5.4 Data Processing for Fluorescence Data

Fluorescence data can be relatively complex and contain loads of information. However, only information valuable for the process or measured system is of interest. In order to describe the data in a more condensed manner than the original data array, multivariate data analysis is applied. Fluorescence data can be classified as higher order arrays and for such data PCA is not appropriate. PARAllel FACtor analysis (PARAFAC) is multi-way method applicable for 2D fluorescence data sets (Bro, 1997) but also for data of even higher dimensionalities.

Multi-way data are characterized by having several dimensions, such as fluorescence having a variable set in the excitation wavelength direction and a variable in the emission wavelength direction. As illustrated in Figure 5.5, PARAFAC decomposes the three-way array ($\underline{\mathbf{X}}$) into three matrices (\mathbf{A} , \mathbf{B} and \mathbf{C}), which all are called loadings. The variation not captured by the model is given in the residuals $\underline{\mathbf{E}}$.

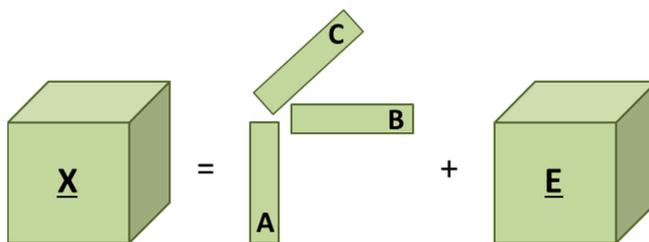


Figure 5.5: A PARAFAC model decomposing $\underline{\mathbf{X}}$ into three vectors (\mathbf{A} , \mathbf{B} , \mathbf{C}) and the residuals ($\underline{\mathbf{E}}$).

In three-way terminology, it is common not to distinguish between scores and loadings because they are treated equal numerically (Bro, 1997). However, the first loading matrix (\mathbf{A}) can for convenience also be recognized as the score or concentration matrix, whereas the two remaining matrices (\mathbf{B} and \mathbf{C}) are known as loadings. When modelling data by PARAFAC, the chemical rank or the number of factors must be defined. This can be done by e.g. judgement of the residuals or applying core consistency diagnostics⁶ (Bro, 1998b). These techniques can very well be applied as an indicator, but it is always important to bring in external knowledge of the modelled data, before deciding on the number of factors. Based on such knowledge following parameters could be considered: *What measurement technique has been applied? What system is modelled? How many chemical components are expected to be detected in the data?* By inspection and considering the data, it often becomes easier to define the chemical rank of the system. A PARAFAC model with F factors can be written as:

⁶The core consistency diagnostic helps determination of the model complexity of low-rank trilinear data, where a high core consistency near 100 % suggests that the model fit is good and a low or negative core consistency alludes that the model is over-fitted (Bro, 1998b).

$$x_{ijk} = \sum_{f=1}^F a_{if} b_{jf} c_{kf} + e_{ijk} \quad (5.10)$$

where a_{if} , b_{jf} and c_{kf} represent the elements of the three matrices (**A**, **B** and **C**), and e_{ijk} is a residual element. In the case of fluorescence data, x_{ijk} represents the intensity of sample number i at variable (emission wavelength) number j and at variable (excitation wavelength) number k (Bro, 1997; Smilde, et al., 2004b). In the case where F is equal to three, the PARAFAC model is a three-component model and can thus be graphically illustrated as seen in Figure 5.6.

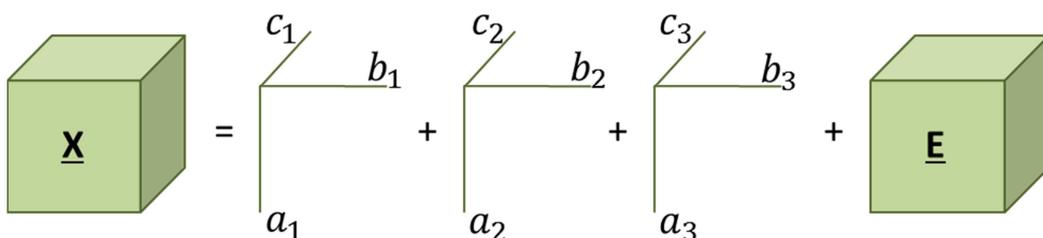


Figure 5.6: An illustration of a three-component PARAFAC model.

The PARAFAC model can also be written in matrix notation as seen in Equation 5.11, in which each slab of the three-way array $\underline{\mathbf{X}}$ ($I \times J \times K$) is given as \mathbf{X}_k :

$$\mathbf{X}_k = \mathbf{A} \mathbf{D}_k \mathbf{B}^T + \mathbf{E}_k \quad (5.11)$$

Each slab has the size $I \times J$ and is modelled by \mathbf{A} ($I \times F$) representing the matrix with the scores (first mode), \mathbf{B}^T ($J \times R$) representing the first set of loadings (second mode) and \mathbf{D}_k representing the second set of loadings (third mode), where \mathbf{D}_k is the diagonal matrix ($R \times R$) containing the weights for the k th slab of $\underline{\mathbf{X}}$ (Kiers, et al., 1999; Smilde, et al., 2004c).

In bilinear methods, such as PCA and MCR⁷, the calculated model can be rotated without changing the solution. This means that even though the decomposed loadings might reflect the pure spectra, it is not possible to find the original spectra, because the solution can be rotated. This rotational freedom is a well-known problem for bilinear methods, which complicates the interpretations. Contrary, if the data are trilinear and the right number of components is used the true underlying spectra will be found by PARAFAC. This is a major advantage in which the

⁷Multivariate Curve Resolution (MCR) is a multivariate data analysis, which can be applied for two-way arrays in the same manner as PCA. However, MCR differ from PCA in its application of additional constraints on the solution (Tauler, 1995).

calculated model is unique, which means that the estimated model cannot be rotated without a lower fit (Bro, 1997; Bro, 1998a).

Several studies have proven that PARAFAC is a comprehensive method for decomposing fluorescence arrays. For example, one study showed that it is possible to distinguish between different dissolved organic matters collected from diverse marine habitats (Murphy, et al., 2008). Another study was able to separate milk samples, treated with instant infusion pasteurization, according to the variations in heat treatment by using fluorescence landscapes and PARAFAC (Hougaard, et al., 2013). Furthermore, PARAFAC was applied to estimate the dioxin content of fish oils based on fluorescence data (Pedersen, et al., 2002). The mentioned studies confirm the advantages of using PARAFAC for analysing fluorescence spectroscopy data and exemplify that PARAFAC is applicable for interpretation of those data.

5.5 Study III – Evaluation of on-line Fluorescence Spectroscopy applied in Fermentation Processes

This section is based on the paper “Weighted PARAFAC and non-linear regression for handling intensity changes in fluorescence spectroscopy caused by pH fluctuations” accepted for publication in the peer-reviewed journal *Applied Spectroscopy*.

Aim

Fluorescence spectroscopy was in previously studies found to be a promising spectroscopic technique for on-line monitoring of fermentation processes. In Paper I, it was also illustrated that fluorescence spectroscopy was capable of monitoring specific chemical compounds important for the control of the fermentation process. However, since various parameters may influence the fluorescence intensity, there is a risk that such interference can cause unreliable determinations and thereby decrease the quality of the quantitative measurements. To examine one of these parameters, pH fluctuations were introduced in a riboflavin degradation process monitored by fluorescence spectroscopy. The influence on the fluorescence intensity was elucidated and a correction strategy for handling the intensity shifts was suggested.

Illustrations of the Experimental Set-up

A fermentation process includes a number of complex sub-processes, such as textual changes, anabolic and catabolic processes where metabolites are consumed and produced. In order to study the manually introduced pH interferences, a simple model system on riboflavin degradation was employed. Vitamin effervescent tablets containing various water-soluble vitamins were dissolved in 1.5 L of water. The multivitamin tablets (Optisana, Kolding, Denmark) are pictured in Figure 5.7.



Figure 5.7: Vitamin tablets (left) and their packages (right) applied for the model system are pictured.

The experimental set-up is illustrated in Figure 5.8. The degradation process was carried out in a fermenter, which was temperature controlled by an external water bath. The fermenter was enclosed with a black cover (not shown in the figure), in order to limit riboflavin breakdown by light from the surroundings. A light source was introduced via a port in the top of the fermenter and was turned on in order to start the controlled light induced degradation process.

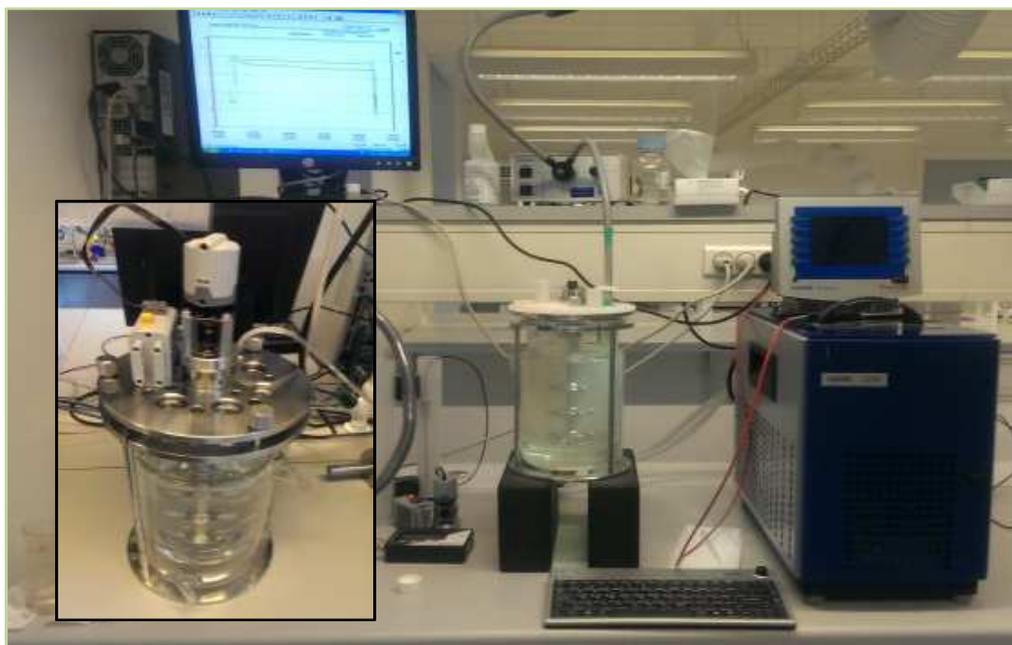


Figure 5.8: The experimental set-up. From the left; the 2 L bioreactor containing the vitamin tablet dissolved in water; the temperature was maintained by the blue HAAKE Phoenix pumping water bath connected to the bioreactor. In-line probes (pH, temperature and fluorescence) are situated in the top and enter the reactor via ports. A small motor was controlling the stirring speed, also situated at the top; a light source was introduced in the top of the fermenter; on-line pH and Temperature profiles can be followed on the screen.

In total 12 experiments were conducted. Seven of these were carried out under Normal Operating Conditions (NOC) and the remaining five batches were carried out under Abnormal Operating Conditions (AOC) where a pH disturbance was introduced. The photo-degradation of riboflavin and its main products lumiflavin and lumichrome were monitored by fluorescence spectroscopy. The structures of the three compounds are illustrated in Figure 5.9.

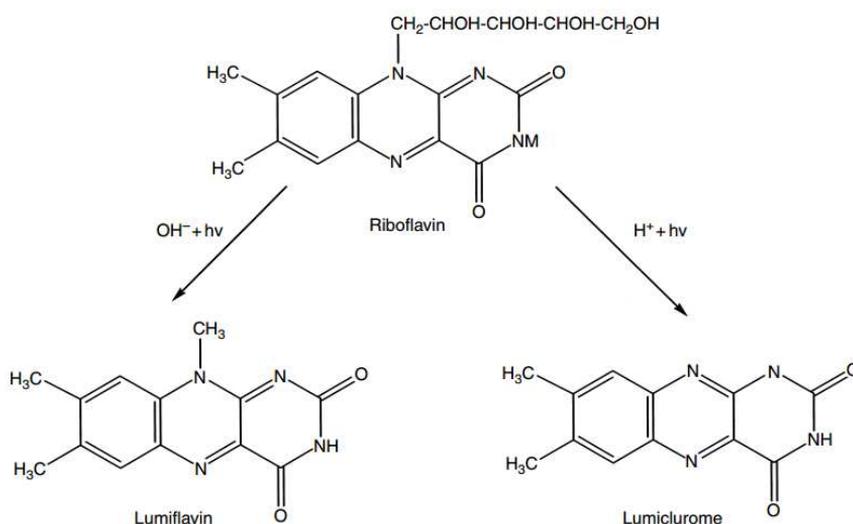


Figure 5.9: The conversion of riboflavin into lumiflavin and lumichrome (Eitenmiller, et al., 2008).

Riboflavin (7,8-dimethyl-10-ribityl-isoalloxazine), better known as vitamin B₂, belongs to the class of water-soluble vitamins and is found in a variety of food products (Belitz, et al., 2004b). The main absorption bands of Riboflavin in the UV and visible wavelength region are around 170, 350 and 440 nm. All three absorptions belong to the singlet transition and emit fluorescence around 530 nm (Pan, et al., 2001).

Results

During the AOC batches acid was added in various amount followed by addition of base in order to reach the original pH again. A fluorescence landscape from one of the AOC batches is illustrated in Figure 5.10. Raw data inspection was carried out by plotting various combinations of excitation and emission wavelengths versus fermentation time. Hereby the intensity of a certain fluorophore can be inspected. As illustrated in Figure 5.10, there is a clear shift in fluorescence intensity for some of the peaks.

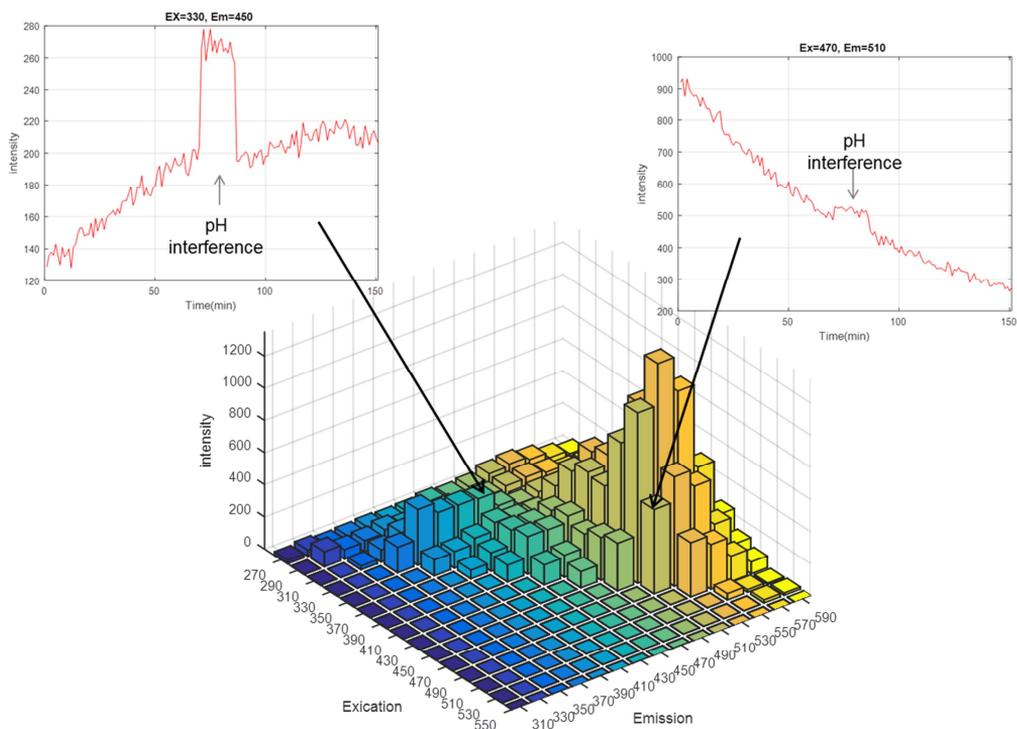


Figure 5.10: A fluorescence landscape from the riboflavin degradation process. Two fluorescence peaks at excitation/emission 330/450 nm and 470/510 nm are plotted versus fermentation time.

It is clear that special care must be taken when interpreting the data, as a false intensity/concentration signal might appear if the pH is changing. Therefore a correcting strategy was suggested in order to filter away these intensity shifts. The strategy is presented in Figure 5.11.

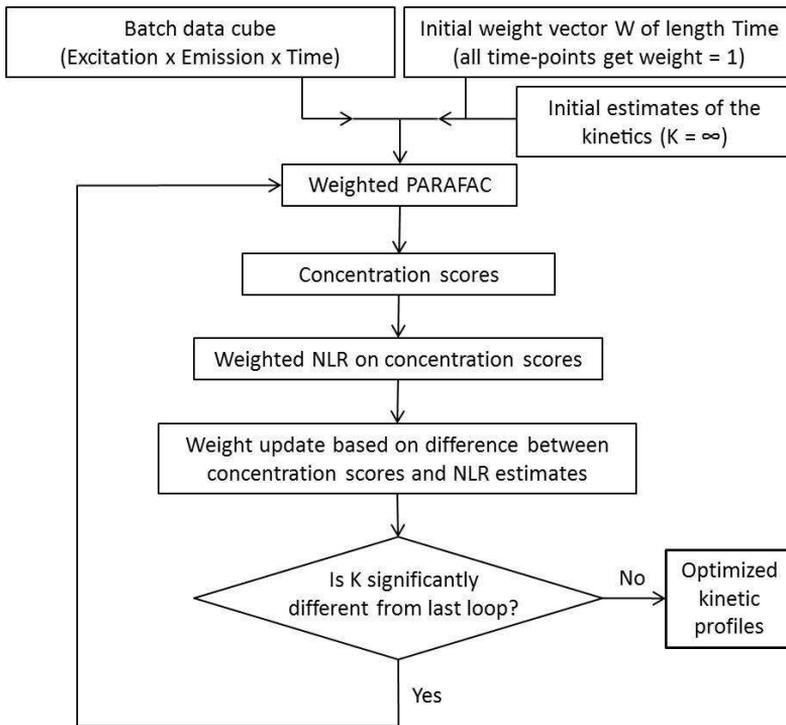


Figure 5.11: Flowchart of the modelling approach.

The suggested correction strategy combines weighted PARAFAC and weighted non-linear regression (NLR). Firstly, the fluorescence landscapes collected over time for one batch should be stacked into a three-way data structure and an initial weight-vector with value 1 for each measurement/time is initialized. To obtain a kinetic profile that is corrected for any shifts or disturbances a new weight vector is determined inside the loop. The initial estimate of the reaction constant K is set to the value infinite. A PARAFAC model is then calculated using the present weight vector, where all excitation-emission-combinations for one time point/landscape get the same weight as determined for that time point. The calculated PARAFAC intensity scores, which provide the relative concentration profile for each of the detected chemical compounds, are obtained from the weighted PARAFAC model. Hereafter, the model parameters/kinetic coefficient of the concentration scores are estimated by weighted NLR. A new weight vector is calculated based on the difference between the concentration scores, obtained from the weighted PARAFAC model, and the kinetic profiles obtained from the weighted non-linear least-squares. Thereby, a weight between the first score and the kinetic profile of compound one is obtained for each time point in the batch run. Likewise, the difference between the second score and the second kinetic profile and the difference between the third score and the third kinetic profile are calculated. The three outcomes are squared and the average is determined. The average was rescaled to values between 0 and 1 and defined as the

new weight vector. If the K value is significantly different from the previous estimate a new loop/iteration will be started. If the K value is not significantly different from the previous estimate the algorithm has converged and the results are presented.

Discussion, Concluding Remarks and Perspectives

The intensity changes caused by pH changes are not very well described by the existing literature, so it is hard to pinpoint what is actually causing them. However a few related attendances will be discussed here. It is known that the fluorescence intensity changes, when a protein is unfolded (Albani, 2007a). Though, vitamins cannot be denaturated as proteins, it could be discussed whether a charge change of riboflavin is appearing, in which a protonation or a deprotonation might appear under the pH changes and thereby causes a characteristic change of the molecule. The rates of protonation and dissociation can be fast enough to compete with the fluorescence, and thereby the excited state of the fluorescence molecule. This means that the fluorescence intensity may be caused by the quantities of the acid or the conjugated base in the excited state instead of the quantities in the ground state (Schulman, 1971).

Several activities can appear during the excitation and emission steps and since riboflavin belongs to the compound group of flavins, it can appear in several forms. Flavins, which are characterized by having an isoalloxazine ring and ribityl side chain, can exist in three different redox states and depending on the pH of the solution each of these redox species can exist in a cationic, neutral and anionic form, which are illustrated in Figure 5.12 (Weimar and Neims, 1975; Müller, 1991).

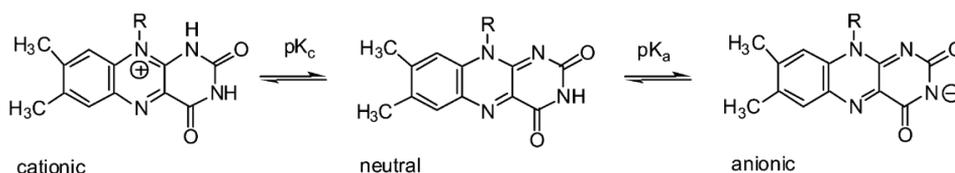


Figure 5.12: An equilibrium scheme between cationic, neutral and anionic flavin species (Islam, et al., 2003).

It has been reported that these forms fluorescence differently in which the cationic form is nearly non-fluorescent and the anionic form is weakly fluorescent, whereas the neutral form is quenched at low pH due to excited-state protonation to the cationic form (Drossler, et al., 2002). Though, it is hard to know which form the riboflavin exists in during the degradation process in the study, these various changing forms could also be a reasonable explanation for the intensity shifts in the applied system. Furthermore, it has even been reported that the pK_a value might differ in the ground state versus the excited state (Schulman, 1971), which only complicates defining the state of riboflavin during a measurement. However, if the pK_a value is

reached within the pH change and the characteristic of the compound thereby change, this could potentially cause an intensity shift. In other words the molecule's transitions between the electronic states might be changed, which could give rise to intensity change.

Another study reports that the fluorescence quantum yield of riboflavin in a aqueous solution differ due to the pH, where the quantum yield was more or less constant in the pH range from 4-9, but below pH 4 the quantum yield was increased (Islam, et al., 2003). As a higher quantum yield leads to larger fluorescence intensity, as seen in Equation 5.1, the intensity of riboflavin should be increased at a lower pH.

The few presented explanations are probably all more or less related to the fluorescence shifts in the riboflavin degradation. No matter what, I believe that the obtained results regarding the issues on intensity changes in fluorescence spectroscopy are of major importance, as the quality of the quantification of the determined chemical compounds can be dramatically affected. I also believe that it important to inform the analytical scientists working with fluorescence spectroscopy that they have to be aware of these measurement issues. And one way to overcome such issues in in-line fluorescence monitoring systems could be to apply the suggested correction strategy.

CHAPTER 6

CONCLUSIONS

The primary aim of this PhD project was to elucidate and explore the dynamics in fermentation processes by NIR and fluorescence spectroscopy. The objective was to decrease the gap between successful research studies conducted at lab scale and industrial full scale implementation of advanced on-line monitoring systems.

The first study (Paper I) outlined the advantages of applying real-time monitoring of bioprocesses and it also highlighted that the applied techniques with various measurement orders deliver different sources of information. Even though zero order measurements, such as pH, provide important information, first order (NIRS) and second order (fluorescence spectroscopy) measurements provided complementary evidence on the microbial state, which increases the process understanding. In a future perspective this allows for better control and enables a much faster optimization effort and error handling. NIRS provided indirect measurements, from which the dynamics of the process could be described, whereas fluorescence spectroscopy allowed quantification of components throughout the fermentation process. It is concluded that both NIR and fluorescence spectroscopy are applicable for real-time monitoring of yoghurt fermentation and thereby provide better fermentation control and process understanding during yoghurt manufacture.

Near infrared spectroscopy has previously demonstrated to be ideal for on-line monitoring of fermentation processes. This PhD thesis provided NIRS results, which show that NIR measurements reveal a quantifiable dynamic trend from which consistency and quality control can be derived. In comparison to pH, which is traditionally applied as the control measure in the industrial production of yoghurt, NIRS does not only provide additional information about the lactic fermentation process, but it also delivers both chemical and physical insight. Pre-processing methods are often required for removal of scatter, in order to model the variation obtained from the absorbance only. In some cases, scatter information obtained from NIRS can however be used for process monitoring. From the results obtained in this thesis, it is concluded that useful variation both in scatter and in the absorption can be present in the NIRS spectra. In combination with kinetic modelling it is possible to model both physical changes, due to the textural changes appearing during the gel formation, and chemical changes, which are related to the biological conversion reactions. The research conducted in this PhD project showed that NIRS can provide valuable physical and chemical information for on-line monitoring of yoghurt

fermentation.

Fluorescence spectroscopy has a high sensitivity and in comparison to near infrared spectroscopy, it is a direct method, which can provide real-time quantitative determinations of molecular compounds throughout the fermentation process. It is however also sensitive to a multiple of external factors which can influence the measurements in a undesirable way. This PhD project elucidates and confirms that pH changes have a major effect on the fluorescence intensities, which can influence the quantifications of relevant key components negatively. When the pH was either increased or decreased, manually, during the light induced degradation process of riboflavin, a clear increase or decrease in the fluorescence landscapes was observed. This thesis presents the major changes seen in the fluorescence landscapes and concludes that fluorescence data must be evaluated carefully if pH changes happen in the measured system. Furthermore, this thesis concludes that such data can still be applied for on-line monitoring if corrections during the modelling stage are carried out. Such a correction strategy, based on a chemometric modelling approach where weighted non-linear regression and weighted PARAFAC are combined, was developed in this PhD thesis and thereby it was possible to compensate for fluorescence intensity shifts.

Based on the above mentioned findings, I believe that this thesis has reached conclusions which may enable improved process control, which potentially allows for a faster detection of abnormal batch conditions, and might also enable better consistency of the product quality in an industrial setting. In addition, I conclude that these findings provide knowledge for the research as well as the industry, in which even better non-invasively cellular physiologic states can be followed both by in-line NIRS, when measuring yoghurt fermentation, and by in-line fluorescence spectroscopy when monitoring processes where pH changes appear during the process.

CHAPTER 7

PERSPECTIVES

The work presented in this thesis and the appended articles and manuscripts illustrates that near infrared spectroscopy and fluorescence spectroscopy have high potential for real-time monitoring of lactic fermentation processes. Furthermore, it presents approaches to extract more information on the dynamics obtained by NIRS, and to gain fluorescence data of higher quality when pH fluctuations are present in the measured process. Nevertheless, more research and effort is needed to overcome the existing challenges and issues concerning implementation of advanced on-line monitoring systems in industrial full-scale fermentation processes.

Knowledge sharing between academia research and industrial manufacturing is a main point for the development of successful monitoring and control approaches. Ideas and practical limitations can hereby be discussed and evaluated before academia research is conducted, which is valuable for the quality of the work. However, when sharing knowledge with e.g. analytical chemists or other scientists, I often meet a certain doubt and scepticism when talking about dynamics and indirect measurements. It seems that analytical scientists find it very uncomfortable to define scientific statements via process dynamics. I believe that more research within dynamics could help convincing scientists that it is a valuable tool for process monitoring. In other words, it seems that one of the main challenges is to achieve acceptance from the scientists, working in the industrial manufacturing, in order to ease the implementation of advanced in-line monitoring strategies.

Off-line state-of-the-art methods for controlling and validating industrial fermentation processes such as GC-MS and HPLC are specific, sensitive and quantitative techniques, which can provide loads of detailed information on the metabolic state of a fermentation. Metabolomics e.g. is a hot research topic these days and GC-MS is a main analytical tool for providing metabolomics profiles of various biological systems. I guess that an approach like GC-MS would meet less scepticism, if it was introduced on-line. It is possible to implement GC-MS and HPLC on-line via a fast-loop and/or a sample port. However, we still have to be aware of complex sample pre-treatment, in which potential errors can be introduced. And it is furthermore not unusual that one measurement takes at least 20 minutes to perform due to the high complexity of fermentation samples which need a proper retention time in order to separate the various compounds optimal. The relatively long measurement duration required increases the measurement inter-

vals. NIRS and fluorescence spectroscopy on the other hand do not meet such challenges. They are not only capable of collecting measurements with an interval of a few minutes, which increase the quality of the process monitoring, but they also do not require further analytical steps, where errors can be introduced, especially in an on-line situation. Therefore, I still believe that on-line probes providing multivariate spectral information have great potential for process monitoring in the food- and pharma industry.

The study concerning intensity shifts in fluorescence landscapes caused by pH fluctuations was based on a model system where the light induced degradation process of riboflavin was followed. This chemistry was applied in order to keep the system as simple as possible. For further work, it would be of great interest to apply the suggested correction strategy on a fermentation system, where pH is decreasing over time due to natural progress. Further research is needed to elucidate how the fluorescence compounds present in the lactic fermentation are affected over time when the pH decreases and if this influences their quantification significantly. If the fluorescence compounds found in the lactic fermentation are actually affected in a similar way as the compounds presented in the riboflavin degradation, the correction strategy might be suitable for the yoghurt fermentation to get a more representative estimation of the dynamics. In addition, pH changes might also appear due to dosage stops, over-adjustment because of slow mixing, and weighting errors in the solvent and buffer preparation step. For that reason, it would be interesting to test the correction strategy in other fermentation systems, especially full scale, in order to be capable of evaluating the approach further.

In GMP manufacturing the criticality of process parameters is often determined and evaluated in order to ensure a uniform production and thereby a uniform product. Based on this evaluation, selection of Critical Quality Attributes (CQA) and rating of severity is done. A CQA can be a physical, chemical, biological or microbiological property or characteristic that should be within an appropriate limit range to ensure the desired product uniformity and quality. Based on the CQAs Critical Process Parameters (CPP) can be defined. The CPPs have an impact on a critical quality attribute and should therefore be monitored and controlled to ensure that the final product has the desired quality. These CPPs could easily be monitored by NIR or/and fluorescence spectroscopy, which would allow for real-time estimation of CPP throughout the process with measurement intervals less than a minute. Based on the work presented in this thesis, I would assume that the gel formation in the yoghurt formation is a potential CPP, as the right viscosity is an important parameter, for the final yoghurt product.

As the concept of GMP is to ensure patient and consumer safety as well as to ensure uniform products, in my opinion, it is just a matter of time before in-line monitoring by e.g. NIR or fluorescence spectroscopy will be implemented in full-scale productions. They are ideal PAT methods, which allow for more sophisticated process controls and thereby for e.g. implementation of a feeding strategy if needed, plus they also give more information on the metabolic state of

the applied microorganism, than the classical monitoring sensors. These mentioned facts all agree with the concept of GMP allowing the process to operate under GMP requirements.

REFERENCE LIST

Abou El-Magd, R.M., Sasaki, C., Kawazoe, T., El-Sayed, S.M., Yorita, K., Shishido, Y., *et al* (2010): Bioprocess development of the production of the mutant P-219-L human D-amino acid oxidase for high soluble fraction expression in recombinant Escherichia coli. *Biochemical Engineering Journal*. Vol. 52:2-3, pp. 236-247.

Aernouts, B., Polshin, E., Lammertyn, J. & Saeys, W. (2011): Visible and near-infrared spectroscopic analysis of raw milk for cow health monitoring: Reflectance or transmittance? *Journal of Dairy Science*. Vol. 94:11, pp. 5315-5329.

Albani, J.R. (2007a): *Binding of TNS on Bovine Serum Albumin at pH 3 and pH 7*. First ed. In: J.R. Albani (ed.): *Principles and Applications of Fluorescence Spectroscopy*. Blackwell Science, Oxford, UK, pp. 210-219.

Albani, J.R. (2007b): *Fluorescence Quenching*. First ed. In: J.R. Albani (ed.): *Principles and Applications of Fluorescence Spectroscopy*. Blackwell Science, Oxford, UK, pp. 139-159.

Albani, J.R. (2007c): *Fluorescence Spectroscopy Principles*. First ed. In: J.R. Albani (ed.): *Principles and Applications of Fluorescence Spectroscopy*. Blackwell Science, Oxford, UK, pp. 88-114.

Alford, J.S. (2006): Bioprocess control: Advances and challenges. *Computers & Chemical Engineering*. Vol. 30:10-12, pp. 1464-1475.

Alves-Rausch, J., Bienert, R., Grimm, C. & Bergmaier, D. (2014): Real time in-line monitoring of large scale Bacillus fermentations with near-infrared spectroscopy. *Journal of Biotechnology*. Vol. 189, pp. 120-128.

Arnold, S.A., Crowley, J., Vaidyanathan, S., Matheson, L., Mohan, P., Hall, J.W., *et al* (2000): At-line monitoring of a submerged filamentous bacterial cultivation using near-infrared spectroscopy. *Enzyme and Microbial Technology*. Vol. 27:9, pp. 691-697.

Arnold, S.A., Crowley, J., Woods, N., Harvey, L.M. & McNeill, B. (2003): In-situ near infrared spectroscopy to monitor key analytes in mammalian cell cultivation. *Biotechnology and Bioengineering*. Vol. 84:1, pp. 13-19.

- Arnold, S.A., Gaensakoo, R., Harvey, L.M. & McNeil, B. (2002): Use of at-line and in-situ near-infrared spectroscopy to monitor biomass in an industrial fed-batch *Escherichia coli* process. *Biotechnology and Bioengineering*. Vol. 80:4, pp. 405-413.
- Barnes, R.J., Dhanoa, M.S. & Lister, S.J. (1989): Standard Normal Variate Transformation and De-Trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy*. Vol. 43:5, pp. 772-777.
- Béal, C. & Helinck, S. (2014): *Chapter 5 - Yoghurt and Other Fermented Milks*. First ed. In: R.C. Ray & D. Montet (eds.): *Microorganisms and Fermentation of Traditional Foods*. CRC Press, Boca Raton, FL., pp. 141-187.
- Becker, T., Hitzmann, B., Muffler, K., Pörtner, R., Reardon, K.F., Stahl, F., *et al* (2006): *Future Aspects of Bioprocess Monitoring*. first ed. In: R. Ulber & D. Sell (eds.) : *White Biotechnology*. Springer-Verlag, Berlin Heidelberg, pp. 249–293.
- Belitz, H.D., Grosch, W. & Schieberle, P. (2004a): *Milk and Dairy Products*. Third ed. In: H.D. Belitz, W. Grosch & P. Schieberle (eds.) : *Food Chemistry*. Springer, Berlin, pp. 505-550.
- Belitz, H.D., Grosch, W. & Schieberle, P. (2004b): *Vitamins*. Third ed. In: H.D. Belitz, W. Grosch & P. Schieberle (eds.) : *Food Chemistry*. Springer, Berlin, pp. 409-426.
- Beutel, S. & Henkel, S. (2011): In situ sensor techniques in modern bioprocess monitoring. *Applied Microbiology and Biotechnology*. Vol. 91:6, pp. 1493-1505.
- Biechele, P., Busse, C., Solle, D., Scheper, T. & Reardon, K. (2015): Sensor systems for bioprocess monitoring. *Engineering in Life Sciences*. Vol. 15:5, pp. 469-488.
- Boehl, D., Solle, D., Hitzmann, B. & Scheper, T. (2003): Chemometric modelling with two-dimensional fluorescence data for *Claviceps purpurea* bioprocess characterization. *Journal of Biotechnology*. Vol. 105:1-2, pp. 179-188.
- Bogomolov, A. & Melenteva, A. (2013): Scatter-based quantitative spectroscopic analysis of milk fat and total protein in the region 400-1100 nm in the presence of fat globule size variability. *Chemometrics and Intelligent Laboratory Systems*. Vol. 126, pp. 129-139.
- Bogomolov, A., Dietrich, S., Boldrini, B. & Kessler, R.W. (2012): Quantitative determination of fat and total protein in milk based on visible light scatter. *Food Chemistry*. Vol. 134:1, pp. 412-418.
- Bogomolov, A., Hessling, M., Wenzel, U., Princz, S., Hellmuth, T., Bernal, M.J.B., *et al* (2015): Development and testing of mid-infrared sensors for in-line process monitoring in biotechnology. *Sensors and Actuators B-Chemical*. Vol. 221, pp. 1601-1610.

Bonk, S., Sandor, M., Rudinger, F., Tscheschke, B., Prediger, A., Babitzky, A., *et al* (2011): In-situ microscopy and 2D fluorescence spectroscopy as online methods for monitoring CHO cells during cultivation. *BMC Proceedings*. Vol. 5 Suppl 8, pp. 76.

Bro, R. (1998a): *Chapter 5 - Background*. First ed. In: R. Bro (ed.) : *Multi-way Analysis in the Food Industri, Models, Algorithms ans Applications, Doctoral Thesis*. University of Amsterdam, Amsterdam, pp. 1-56.

Bro, R. (1998b): *Chapter 5 - Validation*. First ed. In: R. Bro (ed.) : *Multi-way Analysis in the Food Industri, Models, Algorithms ans Applications, Doctoral Thesis*. University of Amsterdam, Amsterdam, pp. 99-134.

Bro, R. (1997): PARAFAC. Tutorial and applications. *Chemometrics and Intelligent Laboratory Systems*. Vol. 38:2, pp. 149-171.

Callis, J.B., Illman, D.L. & Kowalski, B.R. (1987): Process Analytical-Chemistry. *Analytical Chemistry*. Vol. 59:9, pp. 624A–637A.

Cervera, A.E., Petersen, N., Lantz, A.E., Larsen, A. & Gernaey, K.V. (2009): Application of Near-Infrared Spectroscopy for Monitoring and Control of Cell Culture and Fermentation. *Biotechnology Progress*. Vol. 25:6, pp. 1561-1581.

Chisti, Y. (2006): *Bioreactor design*. third ed. In: C. Ratledge & B. Kristiansen (eds.): *Basic Biotechnology*. Cambridge University Press, Cambridge, UK, pp. 181-200.

Christensen, J., Becker, E. & Frederiksen, C. (2005): Fluorescence spectroscopy and PARAFAC in the analysis of yogurt. *Chemometrics and Intelligent Laboratory Systems*. Vol. 75:2, pp. 201-208.

Clavaud, M., Roggo, Y., Von Daeniken, R., Liebler, A. & Schwabe, J. (2013): Chemometrics and in-line near infrared spectroscopic monitoring of a biopharmaceutical Chinese hamster ovary cell culture: Prediction of multiple cultivation variables. *Talanta*. Vol. 111, pp. 28-38.

Clementschtisch, F., Jurgen, K., Florentina, P. & Karl, B. (2005): Sensor combination and chemometric modelling for improved process monitoring in recombinant E. coli fed-batch cultivations. *Journal of Biotechnology*. Vol. 120:2, pp. 183-196.

Coppa, M., Ferlay, A., Leroux, C., Jestin, M., Chilliard, Y., Martin, B., *et al* (2010): Prediction of milk fatty acid composition by near infrared reflectance spectroscopy. *International Dairy Journal*. Vol. 20:3, pp. 182-189.

Crowley, J., Arnold, S.A., Wood, N., Harvey, L.M. & McNeil, B. (2005): Monitoring a high cell density recombinant *Pichia pastoris* fed-batch bioprocess using transmission and reflectance near infrared spectroscopy. *Enzyme and Microbial Technology*. Vol. 36:5-6, pp. 621-628.

Czarnik-Matuszewicz, B., Murayama, K., Tsenkova, R. & Ozaki, Y. (1999): Analysis of near-infrared spectra of complicated biological fluids by two-dimensional correlation spectroscopy: Protein and fat concentration-dependent spectral changes of milk. *Applied Spectroscopy*. Vol. 53:12, pp. 1582-1594.

Dahm, D.J. & Dahm, K.D. (2001): *Chapter 1 - The Physics of Near-Infrared Scattering*. Second ed. In: P. Williams & K. Norris (eds.) : *Near-Infrared Tecnology in the Agricultural and Food Industries*. American Association of Cereal Chemists, Inc., Minnesota, USA, pp. 1-17.

De Brabandere, A. & De Baerdemaeker, J. (1999): Effects of process conditions on the pH development during yogurt fermentation. *Journal of Food Engineering*. Vol. 41:3-4, pp. 221-227.

Drossler, P., Holzer, W., Penzkofer, A. & Hegemann, P. (2002): pH dependence of the absorption and emission behaviour of riboflavin in aqueous solution. *Chemical Physics*. Vol. 282:3, pp. 429-439.

Eitenmiller, R.R., Landen, W.O. & Ye, L. (2008): *Chapter seven: Riboflavin*. Second ed. In: R.R. Eitenmiller, W.O. Landen & L. Ye (eds.): *Vitamin Analysis for the Health and Food Sciences*. CRC Press, Boca Raton, FL, pp. 325-360.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C. & Wold, S. (2006): *Chapter 3 - PCA*. In: L. Eriksson, E. Johansson, N. Kettaneh-Wold, J. Trygg, C. Wikström & S. Wold (eds.): *Multi- and Megavariate Data Analysis, Basic Principles and Applications*. Umetrics Academy, Umeå, Sweden, pp. 39.

Esbensen, K., Guyot, D., Wested, F. & Houmøller, L. (2000): *Principal Component Analysis (PCA) - Introduction*. Fourth ed. In: K. Esbensen, D. Guyot, F. Wested & L. Houmøller (eds.) Camo ASA, Oslo, Norway, pp. 19.

Faassen, S.M. & Hitzmann, B. (2015): Fluorescence Spectroscopy and Chemometric Modeling for Bioprocess Monitoring. *Sensors*. Vol. 15:5, pp. 10271-10291.

FDA (2004): *"Guidance for Industry PAT — A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance"*. Food and Drug Administration, Rockville, USA, Rockville, MD.

Furukawa, K., Heinzle, E. & Dunn, I.J. (1983): Influence of Oxygen on the Growth of *Saccharomyces-Cerevisiae* in Continuous Culture. *Biotechnology and Bioengineering*. Vol. 25:10, pp. 2293-2317.

Garcia-Ochoa, F., Gomez, E., Santos, V.E. & Merchuk, J.C. (2010): Oxygen uptake rate in microbial processes: An overview. *Biochemical Engineering Journal*. Vol. 49:3, pp. 289-307.

Geladi, P., Macdougall, D. & Martens, H. (1985): Linearization and Scatter-Correction for Near-Infrared Reflectance Spectra of Meat. *Applied Spectroscopy*. Vol. 39:3, pp. 491-500.

Goldfeld, M., Christensen, J., Pollard, D., Gibson, E.R., Olesberg, J.T., Koerperick, E.J., *et al* (2014): Advanced near-infrared monitor for stable real-time measurement and control of *Pichia pastoris* bioprocesses. *Biotechnology Progress*. Vol. 30:3, pp. 749-759.

Grassi, S., Alamprese, C., Bono, V., Picozzi, C., Foschino, R. & Casiraghi, E. (2013): Monitoring of lactic acid fermentation process using Fourier transform near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*. Vol. 21:5, pp. 417-425.

Griffiths, P.R. (2002): *Introduction to Vibrational Spectroscopy*. In: J.M. Chalmers & P.R. Griffiths (eds.): *Handbook of Vibrational Spectroscopy, Vol. 1*. John Wiley & Sons Ltd., Chichester, UK, pp. 33-43.

Guilbault, G.G. (1990): *Practical Fluorescence*. Second ed. Marcel Dekker, Inc., New York.

Gürakan, G.C. & Altay, N. (2010): *Chapter 3 - Yoghurt Microbiology and Biochemistry*. First ed. In: F. Yildiz (ed.): *Development and Manufacture of Yoghurt and Other Functional Dairy Products*. CRC Press, Boca Raton, FL., pp. 97-122.

Haack, M.B., Lantz, A.E., Mortensen, P.P. & Olsson, L. (2007): Chemometric analysis of in-line multi-wavelength fluorescence measurements obtained during cultivations with a lipase producing *Aspergillus oryzae* strain. *Biotechnology and Bioengineering*. Vol. 96:5, pp. 904-913.

Haack, M., Eliasson, A. & Olsson, L. (2004): On-line cell mass monitoring of *Saccharomyces cerevisiae* cultivations by multi-wavelength fluorescence. *Journal of Biotechnology*. Vol. 114:1-2, pp. 199-208.

Hantelmann, K., Kollecker, A., Hull, D., Hitzmann, B. & Scheper, T. (2006): Two-dimensional fluorescence spectroscopy: A novel approach for controlling fed-batch cultivations. *Journal of Biotechnology*. Vol. 121:3, pp. 410-417.

Harnett, J., Davey, G., Patrick, A., Caddick, C. & Pearce, L. (2011): *Lactic acid bacteria | Streptococcus thermophilus*. Second ed. In: W.F. Fuquay, P.F. Fox & P.L.H. McSweeney (eds.): *Encyclopedia of Dairy Sciences*. Academic Press, San Diego, CA, pp. 143-148.

Harris, D.C. (2010): *Fundamentals of Spectrophotometry*. Eighth ed. In: D.C. Harris (ed.): *Quantitative Chemical Analysis*. W. H. Freeman and Company, California, pp. 393-418.

Heijnen, J.J. (2015): *Chapter 8 - Fed-batch fermentation and transport phenomena*. 28th ed. In: J. Pronk & J.J. Heijnen (eds.): *Advanced course - Microbial physiology and fermentation technology*. Delft University of Technology, Delft, chap. 8.

Herman, H. (2000): *UV-visible, fluorescence and chemiluminescence spectroscopies in chemical analysis*. 1st ed. In: J.M. Chalmers (ed.): *Spectroscopy in process analysis*. CRC Press, Boca Raton, FL, pp. 16-52.

Hisiger, S. & Jolicoeur, M. (2005a): A multiwavelength fluorescence probe: Is one probe capable for on-line monitoring of recombinant protein production and biomass activity? *Journal of Biotechnology*. Vol. 117:4, pp. 325-336.

Hisiger, S. & Jolicoeur, M. (2005b): Plant cell culture monitoring using an in situ multiwavelength fluorescence probe. *Biotechnology Progress*. Vol. 21:2, pp. 580-589.

Hougaard, A.B., Lawaetz, A.J. & Ipsen, R.H. (2013): Front face fluorescence spectroscopy and multi-way data analysis for characterization of milk pasteurized using instant infusion. *Lwt-Food Science and Technology*. Vol. 53:1, pp. 331-337.

Hutkins, R.W. & Morris, H.A. (1987): Carbohydrate-Metabolism by *Streptococcus-Thermophilus* - a Review. *Journal of Food Protection*. Vol. 50:10, pp. 876-884.

Islam, S.D.M., Susdorf, T., Penzkofer, A. & Hegemann, P. (2003): Fluorescence quenching of flavin adenine dinucleotide in aqueous solution by pH dependent isomerisation and photo-induced electron transfer. *Chemical Physics*. Vol. 295:2, pp. 137-149.

Iversen, J.A., Berg, R.W. & Ahring, B.K. (2014): Quantitative monitoring of yeast fermentation using Raman spectroscopy. *Analytical and Bioanalytical Chemistry*. Vol. 406:20, pp. 4911-4919.

Johansson, L. & Liden, G. (2006): A study of long-term effects on plasmid-containing *Escherichia coli* in carbon-limited chemostat using 2D-fluorescence spectrofluorimetry. *Biotechnology Progress*. Vol. 22:4, pp. 1132-1139.

Kahl, A. (2016): Picture of *Lactobacillus delbruekii subsp. bulgaricus*. At: http://bioweb.uwlax.edu/bio203/s2007/kahl_ambe/.

Kawano, S. (2002): *Sampling and Sample Presentation*. First ed. In: H.W. Siesler, Y. Ozaki, S. Kawata & H.M. Heise (eds.): *Near-Infrared Spectroscopy - Principles, Instruments, Applications*. Wiley-VCH, Weinheim, Germany, pp. 115-124.

Kiers, H., Ten Berge, J. & Bro, R. (1999): PARAFAC2 - Part I. A direct fitting algorithm for the PARAFAC2 model. *Journal of Chemometrics*. Vol. 13:3-4, pp. 275-294.

Kim, S., d'Anjou, M., Lanz, K.J., Evans, C.E., Gibson, E.R., Olesberg, J.T., *et al* (2015): Real-time monitoring of glycerol and methanol to enhance antibody production in industrial *Pichia pastoris* bioprocesses. *Biochemical Engineering Journal*. Vol. 94, pp. 115-124.

King, J.P. (2014): *Instrumentation and Control Systems*. Third ed. In: C.C. Todaro & H.C. Vogel (eds.): *Fermentation and Biochemical Engineering Handbook*. Elsevier Inc., Oxford, pp. 401-414.

Kornmann, H., Valentinotti, S., Duboc, P., Marison, I. & von Stockar, U. (2004): Monitoring and control of *Gluconacetobacter xylinus* fed-batch cultures using in situ mid-IR spectroscopy. *Journal of Biotechnology*. Vol. 113:1-3, pp. 231-245.

Kubista, M., Sjoback, R., Eriksson, S. & Albinsson, B. (1994): Experimental Correction for the Inner-Filter Effect in Fluorescence-Spectra. *Analyst*. Vol. 119:3, pp. 417-419.

Lakowicz, J.R. (2006a): *Chapter 1 - Introduction to Fluorescence*. Third ed. In: J.R. Lakowicz (ed.): *Principles of Fluorescence Spectroscopy*. Springer, New York, USA, pp. 1-26.

Lakowicz, J.R. (2006b): *Chapter 16 - Protein Fluorescence*. Third ed. In: J.R. Lakowicz (ed.): *Principles of Fluorescence Spectroscopy*. Springer, New York, USA, pp. 529-575.

Lakowicz, J.R. (2006c): *Chapter 8 - Quenching of Fluorescence*. Third ed. In: J.R. Lakowicz (ed.): *Principles of Fluorescence Spectroscopy*. Springer, New York, USA, pp. 277-330.

Lakowicz, J.R. (1999): *Chapter 1: Introduction to fluorescence*. Second ed. In: J.R. Lakowicz (ed.): *Principles of Fluorescence Spectroscopy*. Springer, New York, USA, pp. 1-24.

Lam, H. & Kostov, Y. (2009): *Optical Instrumentation for Bioprocess Monitoring*. First ed. In: G. Rao (ed.): *Optical Sensor Systems in Biotechnology*. Springer, Berlin, pp. 125-142.

Landgrebe, D., Haake, C., Hoepfner, T., Beutel, S., Hitzmann, B., Scheper, T., *et al* (2010): On-line infrared spectroscopy for bioprocess monitoring. *Applied Microbiology and Biotechnology*. Vol. 88:1, pp. 11-22.

Lantz, A., Jorgensen, P., Poulsen, E., Lindemann, C. & Olsson, L. (2006): Determination of cell mass and polymyxin using multi-wavelength fluorescence. *Journal of Biotechnology*. Vol. 121:4, pp. 544-554.

Laporte, M.F. & Paquin, P. (1999): Near-infrared analysis of fat, protein, and casein in cow's milk. *Journal of Agricultural and Food Chemistry*. Vol. 47:7, pp. 2600-2605.

Lee, H.L.T., Boccazzi, P., Gorret, N., Ram, R.J. & Sinskey, A.J. (2004): In situ bioprocess monitoring of *Escherichia coli* bioreactions using Raman spectroscopy. *Vibrational Spectroscopy*. Vol. 35:1-2, pp. 131-137.

Leroy, F. & De Vuyst, L. (2004): Lactic acid bacteria as functional starter cultures for the food fermentation industry. *Trends in Food Science & Technology*. Vol. 15:2, pp. 67-78.

Li, J.K. & Humphrey, A.E. (1992): Factors Affecting Culture Fluorescence when Monitoring Bioreactors. *Journal of Fermentation and Bioengineering*. Vol. 74:2, pp. 104-111.

Locher, G., Sonnleitner, B. & Fiechter, A. (1992): Online Measurement in Biotechnology - Techniques. *Journal of Biotechnology*. Vol. 25:1-2, pp. 23-53.

Lopes, M.B., Goncalves, G.A.L., Felicio-Silva, D., Prather, K.L.J., Monteiro, G.A., Prazeres, D.M.F., *et al* (2015): In situ NIR spectroscopy monitoring of plasmid production processes: effect of

producing strain, medium composition and the cultivation strategy. *Journal of Chemical Technology and Biotechnology*. Vol. 90:2, pp. 255-261.

Lourenco, N.D., Lopes, J.A., Almeida, C.F., Sarraguca, M.C. & Pinheiro, H.M. (2012): Bioreactor monitoring with spectroscopy and chemometrics: A review. *Analytical and Bioanalytical Chemistry*. Vol. 404:4, pp. 1211-1237.

Marose, S., Lindemann, C. & Scheper, T. (1998): Two-dimensional fluorescence spectroscopy: A new tool for on-line bioprocess monitoring. *Biotechnology Progress*. Vol. 14:1, pp. 63-74.

Martens, H., Nielsen, J.P. & Engelsen, S.B. (2003): Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures. *Analytical Chemistry*. Vol. 75:3, pp. 394-404.

Martens, H. & Stark, E. (1991): Extended Multiplicative Signal Correction and Spectral Interference Subtraction - New Preprocessing Methods for Near-Infrared Spectroscopy. *Journal of Pharmaceutical and Biomedical Analysis*. Vol. 9:8, pp. 625-635.

Masiero, S.S., Trierweiler, J.O., Farenzena, M., Escobar, M., Trierweiler, L.F. & Ranzan, C. (2013): Evaluation of Wavelength Selection Methods for 2d Fluorescence Spectra Applied to Bioprocesses Characterization. *Brazilian Journal of Chemical Engineering*. Vol. 30:2, pp. 289-298.

Miller, C.E. (2010): *Chemometrics in Process Analytical Chemistry (PAT)*. Second ed. In: K. Bakeev (ed.): *Process Analytical Technology, Spectroscopic Tools and Implementation Strategies for the Chemical and Pharmaceutical Industries*. Wiley, New Jersey, USA, pp. 353-438.

Miller, C.E. (2001): *Chapter 2 - Chemical Principles of Near-Infrared Technology*. Second ed. In: P. Williams & K. Norris (eds.): *Near-Infrared Technology in the Agricultural and Food Industries*. American Association of Cereal Chemists, Inc., Minnesota, USA, pp. 19-37.

Mitchell, D.A., Berovic, M. & Krieger, N. (2000): Biochemical engineering aspects of solid state bioprocessing. *Advances in Biochemical Engineering/Biotechnology*. Vol. 68, pp. 61-138.

Mortensen, P.P. & Bro, R. (2006): Real-time monitoring and chemical profiling of a cultivation process. *Chemometrics and Intelligent Laboratory Systems*. Vol. 84:1-2, pp. 106-113.

Müller, F. (1991): *Chemistry and Biochemistry of Flavoenzymes*. First ed. CRC Press, Boca Raton, FL.

Murphy, K.R., Stedmon, C.A., Waite, T.D. & Ruiz, G.M. (2008): Distinguishing between terrestrial and autochthonous organic matter sources in marine environments using fluorescence spectroscopy. *Marine Chemistry*. Vol. 108:1-2, pp. 40-58.

Næs, T., Isaksson, T., Fearn, T. & Davies, T. (2002): *Scatter correction of spectroscopic data*. First ed. In: T. Næs, T. Isaksson, T. Fearn & T. Davies (eds.): *A user friendly guide to multivariate calibration and classification*. NIR Publications, Chichester, pp. 105-125.

Nagodawithana, T.W., Castellano, C. & Steinkraus, K.H. (1974): Effect of Dissolved-Oxygen, Temperature, Initial Cell Count, and Sugar Concentration on Viability of *Saccharomyces Cerevisiae* in Rapid Fermentations. *Applied Microbiology*. Vol. 28:3, pp. 383-391.

Nauth, K.R. (2004): *Yogurt*. In: Y.H. Hui, L. Meunier-Goddik, Å.S. Hansen, J. Josephsen, W.K. Nip, P.S. Stanfield & F. Toldrá (eds.): *Handbook of Food and Beverage Fermentation Technology*. Marcel Dekker, Inc., New York, pp. 125-145.

Noui, L., Hill, J., Keay, P.J., Wang, R.Y., Smith, T., Yeung, K., *et al* (2002): Development of a high resolution UV spectrophotometer for at-line monitoring of bioprocesses. *Chemical Engineering and Processing*. Vol. 41:2, pp. 107-114.

Odman, P., Johansen, C.L., Olsson, L., Gernaey, K.V. & Lantz, A.E. (2010): Sensor combination and chemometric variable selection for online monitoring of *Streptomyces coelicolor* fed-batch cultivations. *Applied Microbiology and Biotechnology*. Vol. 86:6, pp. 1745-1759.

Odman, P., Johansen, C.L., Olsson, L., Gernaey, K.V. & Lantz, A.E. (2009): On-line estimation of biomass, glucose and ethanol in *Saccharomyces cerevisiae* cultivations using in-situ multi-wavelength fluorescence and software sensors. *Journal of Biotechnology*. Vol. 144:2, pp. 102-112.

Özer, B. (2010): *Chapter 2 - Strategies for Yogurt Manufacturing*. In: F. Yildiz (ed.): *Development and Manufacture of Yogurt and Other Functional Dairy Products*. CRC Press, Taylor & Francis Group, Florida, pp. 47-96.

Özer, B. & Kirmaci, H.A. (2010): *Chapter 8 - Quality Attributes of Yogurt and Functional Dairy Products*. First ed. In: F. Yildiz (ed.): *Development and Manufacture of Yoghurt and Other Functional Dairy Products*. CRC Press, Boca Raton, FL., pp. 229-265.

Pan, Y.L., Pinnick, R.G., Hill, S.C., Niles, S., Holler, S., Bottiger, J.R., *et al* (2001): Dynamics of photon-induced degradation and fluorescence in riboflavin microparticles. *Applied Physics B-Lasers and Optics*. Vol. 72:4, pp. 449-454.

Pedersen, D.K., Munck, L. & Engelsen, S.B. (2002): Screening for dioxin contamination in fish oil by PARAFAC and N-PLSR analysis of fluorescence landscapes. *Journal of Chemometrics*. Vol. 16:8-10, pp. 451-460.

Petersen, N., Oedman, P., Padrell, A.E.C., Stocks, S., Lantz, A.E. & Gernaey, K.V. (2010): In Situ Near Infrared Spectroscopy for Analyte-Specific Monitoring of Glucose and Ammonium in *Streptomyces coelicolor* Fermentations. *Biotechnology Progress*. Vol. 26:1, pp. 263-271.

Po, H.N. & Senozan, N.M. (2001): The Henderson-Hasselbalch equation: Its history and limitations. *Journal of Chemical Education*. Vol. 78:11, pp. 1499-1503.

- Pohlscheidt, M., Charaniya, S., Jenzsch, M., Bork, C., Noetzel, T.L. & Luebbert, A. (2013): *Bioprocess and Fermentation Monitoring*. First edn. In: M.C. Flickinger (ed.): *Upstream Industrial Biotechnology*. John Wiley & Sons, Inc., Hoboken, New Jersey, USA, pp. 1471-1491.
- Pons, M.N., Le Bonte, S. & Potier, O. (2004): Spectral analysis and fingerprinting for biomedica characterisation. *Journal of Biotechnology*. Vol. 113:1-3, pp. 211-230.
- Princz, S., Wenzel, U., Miller, R. & Hessling, M. (2014): Data Pre-Processing Method to Remove Interference of Gas Bubbles and Cell Clusters During Anaerobic and Aerobic Yeast Fermentations in a Stirred Tank Bioreactor. *Journal of Applied Spectroscopy*. Vol. 81:5, pp. 855-861.
- Ratledge, C. (2006): *Chapter 2 - Biochemistry and physiology of growth and metabolism*. Third edn. In: C. Ratledge & B. Kristiansen (eds.): *Basic Biotechnology*. Cambridge University Press, New York, pp. 25-54.
- Rhiel, M., Ducommun, P., Bolzonella, I., Marison, I. & von Stockar, U. (2002): Real-time in situ monitoring of freely suspended and immobilized cell cultures based on mid-infrared spectroscopic measurements. *Biotechnology and Bioengineering*. Vol. 77:2, pp. 174-185.
- Rinnan, A. (2014): Pre-processing in vibrational spectroscopy - when, why and how. *Analytical Methods*. Vol. 6:18, pp. 7124-7129.
- Rinnan, A., van den Berg, F. & Engelsen, S.B. (2009): Review of the most common pre-processing techniques for near-infrared spectra. *Trac-Trends in Analytical Chemistry*. Vol. 28:10, pp. 1201-1222.
- Rizzello, C.G. & De Angelis, M. (2011): *Lactic acid bacteria| Lactobacillus spp.: Lactobacillus delbrueckii Group*. Second edn. In: W.F. Fuquay, P.F. Fox & P.L.H. McSweeney (eds.): *Encyclopedia of Dairy Sciences*. Academic Press, San Diego, CA, pp. 119-124.
- Ronnest, N.P., Stocks, S.M., Lantz, A.E. & Gernaey, K.V. (2011): Introducing process analytical technology (PAT) in filamentous cultivation process development: comparison of advanced online sensors for biomass measurement. *Journal of Industrial Microbiology & Biotechnology*. Vol. 38:10, pp. 1679-1690.
- Rossi, D.M., Solle, D., Hitzmann, B. & Zachia Ayub, M.A. (2012): Chemometric modeling and two-dimensional fluorescence analysis of bioprocess with a new strain of *Klebsiella pneumoniae* to convert residual glycerol into 1,3-propanediol. *Journal of Industrial Microbiology & Biotechnology*. Vol. 39:5, pp. 701-708.
- Sablinskas, V., Steiner, G. & Hof, M. (2003): *Applications*. First edn. In: G. Gauglitz & T. Vo-Dinh (eds.): *Handbook of Spectroscopy, Vol. 1*. Wiley-VCH, Weinheim, pp. 89-468.
- Savitzky, A. & Golay, M.J.E. (1964): Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*. Vol. 36:8, pp. 1627-1639.

Scarff, M., Arnold, S.A., Harvey, L.M. & McNeil, B. (2006): Near Infrared Spectroscopy for bio-process monitoring and control: Current status and future trends. *Critical Reviews in Biotechnology*. Vol. 26:1, pp. 17-39.

Schaschke, C. (2014): *A Dictionary of Chemical Engineering*. First ed. Oxford University Press, Oxford, UK.

Schmidt-Hager, J., Ude, C., Findeis, M., John, G.T., Scheper, T. & Beutel, S. (2014): Noninvasive online biomass detector system for cultivation in shake flasks. *Engineering in Life Sciences*. Vol. 14:5, pp. 467-476.

Schulman, S.G. (1985): *Luminescence Spectroscopy: An Overview*. First ed. In: S.G. Schulman (ed.): *Molecular Luminescence Spectroscopy: Methods and Applications: Part 1*. John Wiley & Sons, Inc., New York, USA, pp. 1-28.

Schulman, S.G. (1971): pH Dependence of Fluorescence of Riboflavin and Related Isoalloxazine Derivatives. *Journal of Pharmaceutical Sciences*. Vol. 60:4, pp. 628-631.

Sharma, A. & Schulman, S.E. (1999): *Environmental Effects on Fluorescence*. First ed. In: A. Sharma & S.E. Schulman (eds.): *Introduction to Fluorescence Spectroscopy*. John Wiley and Sons, Inc., Canada, pp. 27-68.

Skoog, D.A. (1998): *Molecular Fluorescence, Phosphorescence, and Chemiluminescence Spectroscopy*. Third ed. In: D.A. Skoog (ed.): *Principles of Instrumental Analysis*. Saunders College Publishing, USA, pp. 225-249.

Smilde, A., Bro, R. & Geladi, P. (2004a): *Array definitions and properties*. First ed. In: A. Smilde, R. Bro & P. Geladi (eds.): *Multi-way Analysis with Applications in the Chemical Sciences*. John Wiley & Sons, Ltd., England, pp. 13-34.

Smilde, A., Bro, R. & Geladi, P. (2004b): *Introduction*. First ed. In: A. Smilde, R. Bro & P. Geladi (eds.): *Multi-way Analysis with Applications in the Chemical Sciences*. John Wiley & Sons, Ltd., England, pp. 1-12.

Smilde, A., Bro, R. & Geladi, P. (2004c): *Three-Way Component and Regression Models*. First ed. In: A. Smilde, R. Bro & P. Geladi (eds.): *Multi-way Analysis with Applications in the Chemical Sciences*. John Wiley & Sons, Ltd., England, pp. 57-87.

Sonnleitner, B. (2013): *Chapter 10 - Measurement, monitoring, modelling and control*. Third ed. In: C. Ratledge & B. Kristiansen (eds.): *Basic Biotechnology*. Cambridge University Press, New York, pp. 251-270.

Sonnleitner, B., Locher, G. & Fiechter, A. (1992): Biomass Determination. *Journal of Biotechnology*. Vol. 25:1-2, pp. 5-22.

- Soons, Z.I.T.A., Streefland, M., van Straten, G. & van Boxtel, A.J.B. (2008): Assessment of near infrared and "software sensor" for biomass monitoring and control. *Chemometrics and Intelligent Laboratory Systems*. Vol. 94:2, pp. 166-174.
- Stanbury, P.F., Whitaker, A. & Hall, S.J. (1999): *Chapter 8 - Instrumentation and Control*. Second ed. In: P.F. Stanbury, A. Whitaker & S.J. Hall (eds.): *Principles of Fermentation Technology*. Butterworth Heinemann, Oxford, pp. 215-241.
- Stärk, E., Hitzmann, B., Schügerl, K., Scheper, T., Fuchs, C., Köster, D., *et al* (2002): *In-Situ-Fluorescence-Probes: A Useful Tool for Non-invasive Bioprocess Monitoring*. First ed. In: K. Schügerl & A.-P. Zeng (eds.): *Tools and Applications of Biochemical Engineering Science*. Springer-Verlag, Berlin, Germany, pp. 21-38.
- Stuart, B.H. (2004): *Chapter 2 - Experimental Methods*. First ed. In: B.H. Stuart (ed.): *Infrared Spectroscopy: Fundamental and Applications*. John Wiley & Sons, Ltd, West Sussex, England, pp. 15-44.
- Surribas, A., Geissler, D., Gierse, A., Scheper, T., Hitzmann, B., Montesinos, J.L., *et al* (2006a): State variables monitoring by in situ multi-wavelength fluorescence spectroscopy in heterologous protein production by *Pichia pastoris*. *Journal of Biotechnology*. Vol. 124:2, pp. 412-419.
- Surribas, A., Manuel Amigo, J., Coello, J., Luis Montesinos, J., Valero, F. & Maspoch, S. (2006b): Parallel factor analysis combined with PLS regression applied to the on-line monitoring of *Pichia pastoris* cultures. *Analytical and Bioanalytical Chemistry*. Vol. 385:7, pp. 1281-1288.
- Svendsen, C., Skov, T. & van den Berg, F.W.J. (2015): Monitoring fermentation processes using in-process measurements of different orders. *Journal of Chemical Technology and Biotechnology*. Vol. 90:2, pp. 244-254.
- Tamime, A.Y. & Robinson, R.K. (2007a): *Chapter 2 - Background to manufacturing practice*. Third ed. In: A.Y. Tamime & R.K. Robinson (eds.): *Tamine and Robinson's Yoghurt - Science and Technology*. pp. 13-161.
- Tamime, A.Y. & Robinson, R.K. (2007b): *Chapter 3 - Processing plants and equipment*. Third ed. In: A.Y. Tamime & R.K. Robinson (eds.): *Tamine and Robinson's Yoghurt - Science and Technology*. pp. 162-283.
- Tamime, A.Y. & Robinson, R.K. (2007c): *Chapter 7 - Biochemistry of fermentation*. Third ed. In: A.Y. Tamime & R.K. Robinson (eds.): *Tamine and Robinson's Yoghurt - Science and Technology*. pp. 535-607.
- Tauler, R. (1995): Multivariate curve resolution applied to second order data. *Chemometrics and Intelligent Laboratory Systems*. Vol. 30:1, pp. 133-146.

Tauler, R., Kowalski, B. & Fleming, S. (1993): Multivariate Curve Resolution Applied to Spectral Data from Multiple Runs of an Industrial-Process. *Analytical Chemistry*. Vol. 65:15, pp. 2040-2047.

Teixeira, A.P., Portugal, C.A.M., Carinhas, N., Dias, J.M.L., Crespo, J.P., Alves, P.M., *et al* (2009): In Situ 2D Fluorometry and Chemometric Monitoring of Mammalian Cell Cultures. *Biotechnology and Bioengineering*. Vol. 102:4, pp. 1098-1106.

Vaccari, G., Dosi, E., Campi, A.L., Gonzalezvara, A., Matteuzzi, D. & Mantovani, G. (1994): A Near-Infrared Spectroscopy Technique for the Control of Fermentation Processes - an Application to Lactic-Acid Fermentation. *Biotechnology and Bioengineering*. Vol. 43:10, pp. 913-917.

Valeur, B. (2001): *Effects of intermolecular photophysical processes on fluorescence emission*. First ed. In: B. Valeur (ed.): *Molecular Fluorescence: Principles and Applications*. Wiley, Weinheim, Germany, pp. 72-124.

Vivatfor (2016): Picture of *Streptococcus Thermophilus*. At: <http://vivatfor.com/bulk-probiotics/61-streptococcus-thermophilus-n008.html>

Von Bargen, K. (1996): *Transflectance Probe having Adjustable Window Gap Adapted to Measure Viscous Substances for Spectrometric Analysis and Method of use*. United States Patent 5708273. Appl. No. 08/64724. Foss NIRSystems, Inc. ed.

Von Wright, A. & Axelsson, L. (2012): *Lactic Acid Bacteria: An Introduction*. Fourth ed. In: S. Lahtinen, A.C. Ouwehand, S. Salminen & A. Von Wright (eds.): *Lactic Acid Bacteria: Microbiological and Functional Aspects*. CRC Press, Boca Raton, FL., pp. 1-16.

Walstra, P., Wouters, J.T.M. & Geurts, T.J. (2005a): *Chapter 22 - Fermented Milks*. Second ed. In: P. Walstra, J.T.M. Wouters & T.J. Geurts (eds.): *Dairy Science and Technology*. CRC Press, Boca Raton, pp. 551-573.

Walstra, P., Wouters, J. & Geurts, T. (2005b): *Chapter 13 - Lactic Fermentations*. Second ed. In: P. Walstra, J. Wouters & T. Geurts (eds.): *Dairy Science and Technology*. CRC Press, Boca Raton, USA, pp. 357-398.

Weimar, W.R. & Neims, A.H. (1975): *Physical and Chemical Properties of Flavins; Binding of Flavins to Protein and Conformational Effects; Biosynthesis of Riboflavin*. First ed. In: R.S. Rivlin (ed.): *Riboflavin*. Plenum Press, London, pp. 1-47.

Wold, S., Esbensen, K. & Geladi, P. (1987): Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems*. Vol. 2:1-3, pp. 37-52.

Wu, D., He, Y. & Feng, S. (2008): Short-wave near-infrared spectroscopy analysis of major compounds in milk powder and wavelength assignment. *Analytica Chimica Acta*. Vol. 610:2, pp. 232-242.

Zabriskie, D.W. & Humphrey, A.E. (1978): Estimation of Fermentation Biomass Concentration by Measuring Culture Fluorescence. *Applied and Environmental Microbiology*. Vol. 35:2, pp. 337-343.

PAPER I

Monitoring fermentation processes using in-process measurements of different orders

Carina Svendsen, Thomas Skov and Frans W. van den Berg

Journal of Chemical Technology and Biotechnology, 90 (2015), 244-254.

Monitoring fermentation processes using in-process measurements of different orders

Carina Svendsen,* Thomas Skov and Frans W. J. van den Berg

Abstract

BACKGROUND: In-process monitoring of fermentation processes (at-line, on-line or in-line measurements) is essential to control productivity and ensure high product quality. A number of different monitoring techniques are available for this purpose and one possible categorization among this variety of techniques is based on the different structures generated by the measurements and the potential of these with respect to holding and extracting process information. In this study lactic fermentation processes is monitored by different techniques (brix, pH, NIR- and fluorescence-spectroscopy) providing different data structures (zero-, first- and second-order). Multivariate data analysis (PCA and PARAFAC) was applied on the first- and second-order data sets, and the different measurement signals or derivatives of these were combined by a multiblock strategy. The aim of this work is to present and clarify the advantages and variations of the different data structures.

RESULTS: The zero-order pH and brix measurements (a commonly used measure for total sugar content in wine fermentations) decreased in a smooth and logical pattern from 6.4 to 4.4 and from 10.5% to 6.2%, respectively – provided valuable critical quality attributes, communicating the fermentation process is progressing over time in accordance with biological and engineering intuition. The first-order NIR measurements modelled with PCA showed an increasing trend over time on PC1. This increasing trend corresponds to the lactic bacterial growth. This trend could be distinguished by statistical modelling from a second trend (PC2), reproducible for all production batches. Based on the second-order fluorescence measurements modelled by PARAFAC and its statistical uniqueness properties, three distinctive fluorescence compounds were found to vary over process time. Most probably these three compounds represent riboflavin, tryptophan and lumichrome or NADH. Using multiblock PCA the combined sensor signals identified two distinguished, reproducible time profiles for all batch runs.

CONCLUSIONS: The most interpretable chemical information was obtained by fluorescence spectroscopy due to the uniqueness properties of second-order measurements. The first-order technique NIR spectroscopy also provided valuable process information, though the process trends can only be interpreted indirectly and if interfering species had been encountered they could not have been modelled. The multiblock data set provided by zero-, first- and second-order measurements recorded over time highlighted important relationships among the different variables that provide chemical information when multivariate data analysis is applied. Although, first- and second-order measurements seem to obtain more information than the zero-order measurements, it is important to keep in mind that zero-order measurements can provide valuable information about the process, especially in combination with different sensors.

© 2014 Society of Chemical Industry

Keywords: data order; zero-, first- and second-order techniques; process monitoring; lactic fermentation; NIR spectroscopy; fluorescence spectroscopy

INTRODUCTION

An increasing focus on the development of more efficient and less time-consuming methods to monitor and control fermentation processes and bio-processes in general is seen in industry. Fermentation processes are commonly used for production of pharmaceuticals, enzymes and foods such as fermented milk products. By applying a suitable control strategy and measuring the critical process parameters in *real time*, such as physical, chemical and biological process conditions, it is possible to control productivity and ensure high product quality. Furthermore, a decrease in energy and raw material use plus an increase in yield might be achieved by better control. Since the Food and Drug Administration (FDA) published the process analytical technology (PAT) guideline in 2004 (<http://www.fda.gov/cder/OPS/PAT.htm>), the number of studies concerning (in-process) monitoring of fermentation processes using various advanced techniques has increased rapidly.

Simple sensors for on-line monitoring of pH, temperature, O₂ and CO₂ have been used for decades as biotechnology measurement systems,^{1,2} and are so common that they could be called classical fermentation monitoring methods providing invaluable process information. The near-infrared (NIR) spectroscopy technique is increasingly used to monitor fermentation processes.^{3,4} NIR spectroscopy can be used to determine the concentration of chemical compounds since the functional groups C–H (aliphatic,

* Correspondence to: C. Svendsen, Spectroscopy & Chemometrics group, Department of Food Science, Faculty of Science, University of Copenhagen, Rolighedsvej 26, DK-1958 Frederiksberg C, Denmark. E-mail: carinasv@food.ku.dk

Spectroscopy and Chemometrics group, Department of Food Science, Faculty of Science, University of Copenhagen, Rolighedsvej 26, DK-1958, Frederiksberg C, Denmark

aromatic or alkene), N–H (amine) and O–H are absorbing in the NIR range.⁵ By the sparsely employed fluorescence spectroscopy measurements a number of biological relevant fluorescent compounds, such as proteins and cofactors, can be selectively determined in a fermentation broth,^{6,7} thereby providing critical biological process parameters.

One way of categorizing the different monitoring techniques is via the different outcomes of data arrangements, meaning that not only is the biological, physical or chemical information of interest, but also the data structures obtained of different complexities or orders.⁸ As will be shown in this paper such complexities can be useful for biotechnological production. The simplest data is obtained by so-called zero-order measurements, where one data point or scalar is obtained per measured sample, e.g. pH measurements. In the case where one data vector is obtained per sample, such as NIR spectroscopy where absorbance at different wavelengths is registered, the signals are called first-order measurements (first order tensor). Data obtained from second-order measurements are more complex from a data structure point of view, as a matrix is obtained per measured sample (second order tensor). Fluorescence spectroscopy is an example, where signals at various emission and excitation wavelength combinations are recorded per sample. Applying the different techniques in on-line monitoring of a fermentation process, measurements are carried out continuously over time to follow the process. Thus, the zero-order measurements will become a first-order tensor, and likewise first- and second-order measurements turn into second- and third-order tensors, respectively, when an extra dimension (time in this case) is added.

The main advantages of higher-order measurements in (bio)process monitoring⁸ are symbolized in Fig. 1. In this depiction three measurements are performed denoting time-progress in fermentation. The blue and cyan dots represent zero-order measurements (pH recordings) while the blue and cyan line-segments indicate first-order measurements (NIR spectra). The full contour landscapes (in this case simulated bi-normal distributions with increasing intensities) symbolize second-order measurements such as multi-wavelength excitation-emission measurements. The red dot in recording number two marks the top of an unknown interfering component in the system.

From the (simulated) example in Fig. 1, comparing the blue and cyan results, it should become clear that a zero-order measurement cannot handle or, equally important, warn us about interfering species in the measurement. For example, a univariate OD600 measurement is not able to discriminate which chemical species increase optical density of the system; the measurement is just a summation of all components or processes involved. This observation should not be confused with the powerful concept mentioned before of recording zero-order measurements as a function of time, where the total actual data output would turn into a first-order tensor, i.e. a vector. For example, a pH–time profile in fermentation can tell us a lot about the performance of a batch, i.e. if the run was in accordance with expectation. But it is important to note that this information is first available post-run or when at least sufficient process time has passed. Every instantaneous pH measurement gives only limited information by itself.

For first-order measurements like NIR spectroscopy the situation is different; the cyan vectors in Fig. 1 can identify similarity between profiles and quantify the component of interest. The first-order advantage⁸ tells us that in the case of interfering chemical species quantification might be obscured, but diagnostics is readily available to judge if a new measurement fits with the

expected profiles or not (as can be understood from the blue measurement series). This is the main idea behind modern Multivariate Statistical Process Control strategies⁹ where the first-order advantage is used not only to predict the specific chemistry of interest but also as diagnostic for process performance.¹⁰ A practical disadvantage of first-order methods is that often larger training-sets of several runs are required to estimate the latent structures in the time-series of a batch-profile.

Quantification in the presence of interfering chemical species is possible in second-order measurements. As long as the component of interest and the interfering component do not overlap too much in the landscape, as depicted in Fig. 1, it is possible to estimate unique profiles in the two directions that describe each measurement landscape by stacking the measurements on top of each other (thus forming a three-way array). And from this mathematical decomposition directly follows the quantification of the component (plus, if present, any interfering species). A second major advantage of second-order data is that reliable estimates can be made with much smaller data sets, which can be of great benefit for new monitoring strategies (and at the initial stage of a batch-run).¹¹

Having access to several different data structures/arrangements from different analytical platforms as defined above is common when monitoring a fermentation process. This emphasizes the concept of bringing all data together by merging data and looking for consensus between the different signals. This discipline called multiblock data analysis is a non-trivial task not only due to the many different information sources but also due to the different data structures; some being discrete (e.g. brix measured on grab samples) and others being near-continuous (e.g. in-process NIR and fluorescence). Many multiblock approaches can be found in the literature, but here only the Consensus Principal Component Analysis (CPCA) version is considered.²⁸

The aim of this study is to compare various data structures and orders obtained from different in-process techniques and illustrate the advantages and variation in complexity of the different data formations. For this we use a model fermentation from food science based on the lactic acid bacteria *Lactobacillus bulgaricus* (LB) and *Streptococcus thermophiles* (ST).

EXPERIMENTAL

Model system

Fermentation batches (five) were carried out using the starter culture YF-3331 (a mixture of the lactic acid bacteria *Lactobacillus bulgaricus* (LB) and *Streptococcus thermophiles* (ST)). The culture was provided by Chr. Hansen A/S (Hørsholm, Denmark) and stored at –45°C. The concentration of the starter culture was 0.2%. Skimmed milk powder (Arla Food Ingredients, Sweden) was used as media by dissolving 1200 g in 11 L of water.

The fermentations were carried out in a 15 L in-house modified glass fermenter vessel (Applikon, Delft, The Netherlands) with a working volume of 11 L. The lid of the vessel contained six instrument ports, a thermowell for temperature monitoring and two ports for the hollow baffle-ring for water flow to control the temperature in the fermenter. A stirrer with two flat-bladed impellers and three stationary baffles were situated in the fermenter. The mixing speed was kept constant at 150 rpm, controlled by a IKA EUROSTAR 60 control motor (IKA-Werke GmbH & Co. KG, Staufen, Germany). The temperature was maintained at 40°C and was controlled by a Pt100 probe inserted in the thermowell (filled with water for conduction) connected to a HAAKE Phoenix pumping

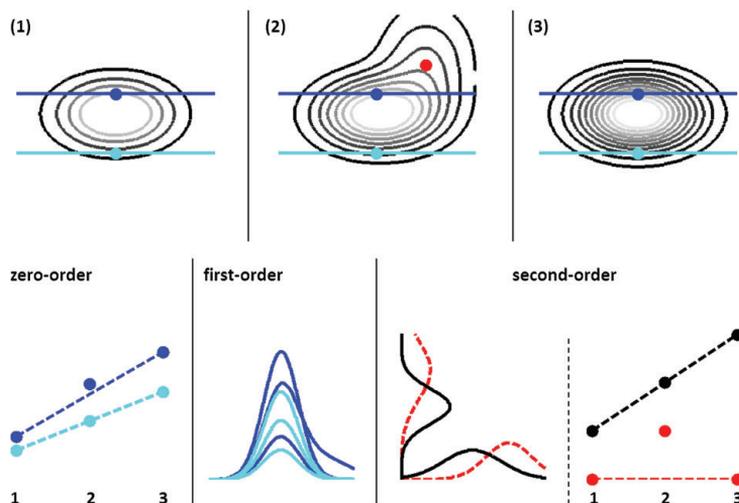


Figure 1. Symbolizing the characteristic difference amid zero-, first- and second-order measurement principles.

water bath (Thermo Scientific, Karlsruhe, Germany). Once the target temperature was reached the instruments were switched on and 22 g of starter culture was inoculated into the media.

Measurements

°Brix (%) measurements

Samples were frequently (but not equidistant or equal over the batch runs) withdrawn manually and °brix was measured at-line by a DR-103 (Index Instruments Limited, Cambridgeshire, UK).

pH measurements

pH measurements were collected in-line with an interval of 30 s, using a pH-meter (MadgeTech Inc., Warner, NH-US) placed in the fermentation broth. The pH probe was calibrated at pH 4.01 and pH 7.00 before use.

NIR spectroscopy

NIR measurements were performed in-line with an ABB Bomem spectrometer (ABB Bomem, Quebec, QC, Canada) equipped with a fibre optic reflectance probe placed in the fermentation broth. The spectral data for each measurement was collected as the average of 64 single beam spectra for each measurement. The spectra were referenced against a white background spectrum (average of 64 scans) collected before the process measurements were started. The fermentation broth was scanned over a range from 1000–1800 nm (10000–5556 cm^{-1} , resolution 8 cm^{-1}) with a time interval of 60 s between measurements.

Fluorescence spectroscopy

The fluorescence measurements were performed on-line/non-invasive with a BioView spectrofluorometer (DELTA Light and Optics, Hørsholm, Denmark). The spectrofluorometer head was situated outside the fermenter, measuring through the glass wall. Fluorescence landscapes were obtained with excitation wavelengths from 270–550 nm and emission wavelengths from 310–590 nm, with an interval of 20 nm, providing a total of 15 excitation and 15 emission wavelengths, every second minute.

Chemometric modelling

All data analyses were performed using Matlab (Matlab R2014a, The Mathworks, Inc. USA). PCA, PARAFAC and CPCA were performed using in-house routines and the PLS-Toolbox (PLS-Toolbox 7.5, Eigenvector Research Inc. USA).

RESULTS

pH and Brix

Figure 2 shows the development of pH in two normal operating condition (NOC) batches, randomly selected from the set of five. The recording and dynamics are in agreement with the well-known profiles of yoghurt cultivation with a pH drop from approximately 6.4 to 4.4 during a processing time of around 3.7 h. The drop in pH with a point of inflection between 1.5 and 2 h into the process indicates that an expected acidification caused by the conversion of lactose into lactate occurs. The end-point of a yoghurt fermentation process is often defined by the pH value and a previous study suggests the continuous control by monitoring the pH.¹² The two runs in the figure also show that the process is very well controlled and reproducible, an observation representative for the larger set.

Figure 2 also presents the zero-order technique brix, measured at-line over time, which shows a sharp decrease from around 10.5% to 6.2% in a narrow window around 2 h run-time. Brix is a summative and unselective measurement concept where all sugars will have a contribution to brix, not only lactose. Moreover, acids generated a measurement response as well. The same observation holds for pH, another zero-order measurement principle, where all acids combined generate the characteristic drop over time. Since a drop in both pH and the sugar content is observed and expected for our system, the pH and brix results can be seen as critical quality attributes as they ensure/indicate that the fermentation is progressing over time according to NOC. However, it should be noted that a brix outcome of 8% can (theoretically) be achieved by an infinite number of concentration-combinations of, for example, lactose and lactic acid and as such the monitoring or

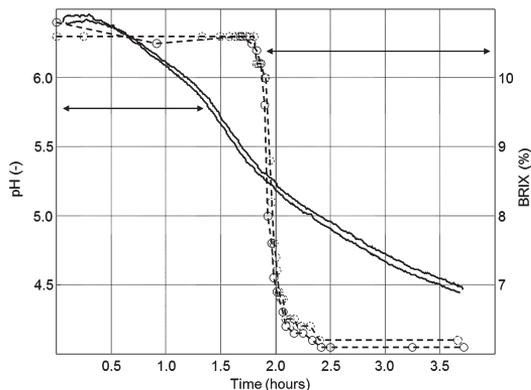


Figure 2. pH (—) and brix (○/—) as a function of time for two representative yoghurt cultivation batches.

control strategy using this signal would have to take the dynamic behaviour (implicitly or explicitly) into consideration. This is also a possible explanation for the unrealistically narrow window of the brix drop; likely there are many different processes taking place at this stage in the batch and zero-order brix measurements are not selective enough to distinguish between them.

NIR-spectroscopy

Figure 3 shows the NIR spectra of one fermentation batch collected over 3.7 h, first-order measurements where the absorbance of the spectra in general decreases over time. The figure shows strong NIR absorption bands in the range between 1125 and 1250 nm and 1350 and 1550 nm. The region between 1150 and 1250 nm can be associated with the second overtone vibration of C–H stretching and the regions between 1350 and 1650 nm might belong to the combination of first overtone of O–H stretching of sugars (1450 nm), first overtone of O–H stretching in water (1440–1470 nm) and third overtone of carbonyl groups (>C=O) (1333–1936 nm).¹³ Further studies report that the sugar content in yoghurt can be predicted by NIR spectroscopy.¹⁴ Thus, it seems feasible to determine sugar contents based on the NIR spectra obtained. However, highly absorbing components such as water and light-scattering particle material causes a major variation in the spectra, while chemical compounds appearing in lactic fermentations such as lactose, galactose and lactate probably only cause minor variances. For that reason multivariate data analysis is applied to identify systematic and principal variation found in the spectra.¹⁵ The raw spectra were pre-processed by standard normal variate (SNV) scaling to remove the sample dissimilarities caused by baseline and effective path-length differences in the NIR reflectance measurements. The region between 1400 and 1500 nm was also removed before modelling because spectra from the beginning of a time-series showed too high absorbance values in this region as observed in Fig. 3, possibly violating Lambert–Beer's law of concentration-linearity. The pre-processed data were modelled by principal components analysis (PCA) in order to find and explore major patterns and trends in the batch data sets. It is worth recalling that pre-processing methods to remove (assumed) artefacts are possible for first-order measurements; at least two responses per single measurement are required to estimate a baseline and remove any undesired offset in the signal. This implies

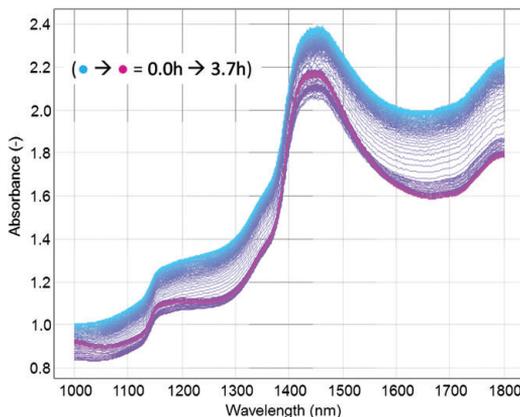


Figure 3. NIR spectra from Batch A (duration 3.7 h, N = 223) coloured/shaded according to time.

that zero-order data cannot be used to detect interfering chemical components, which is a problem when looking at such signals only.

The first principle component vectors shown in Fig. 4(a), computed independently for the two NIR data sets/batch runs, describe an increasing trend over time. This should be interpreted in combination with the variable loadings in Fig. 4(b) due to the sign indeterminacy in PCA.¹⁶ A small bump on the s-shaped curve of the first score is apparent shortly before 2 h for both batches. The increasing tendency is mainly caused by a variation appearing in the baseline slope, around 1350–1400 nm, 1500–1550 nm and to a lesser extent around 1700–1800 nm, which could be interpreted as sugar and the third overtone of carboxyl groups (1333–1936 nm) using the rising flanks of the (eliminated) band centred at 1450 nm as indicator.¹³ The reader should recall that vibrational spectroscopies like NIR are pH dependent and it is notoriously difficult to separate potential change-correlations from true chemistry or biology in reacting systems. Comparing Fig. 4(a) with Fig. 2 we see that the trends show similarity but are sufficiently different to have confidence in the method. The NIR based time-profiles for principle component two (Fig. 4(a)) describes an almost flat baseline for 3.7 h except for a small negative bump followed by a larger positive one at around the (main) inflection point of the first score profile. For NIR the same highly reproducible trends are seen for the two batches, even when modelled separately. The corresponding loading profiles in Fig. 4(b) show primarily a baseline curvature which is, despite the noisy appearance and low percentage variance explained, again very reproducible by a stable measurement principle like NIR. Although we would not describe this as a disturbance or interfering species as classified in the theory section, it does nicely illustrate the first-order principle: compared with the remainder of the batch an 'abnormal event' can be located around 1.9 h into the processes, and we can make this distinction/segmentation based on the spectral pattern in the NIR measurements as summarized in the distinct loading profiles.

Fluorescence spectroscopy

A representative fluorescence excitation-emission landscape – and intrinsic second-order measurement – selected from the measurement series of Batch A is presented as a

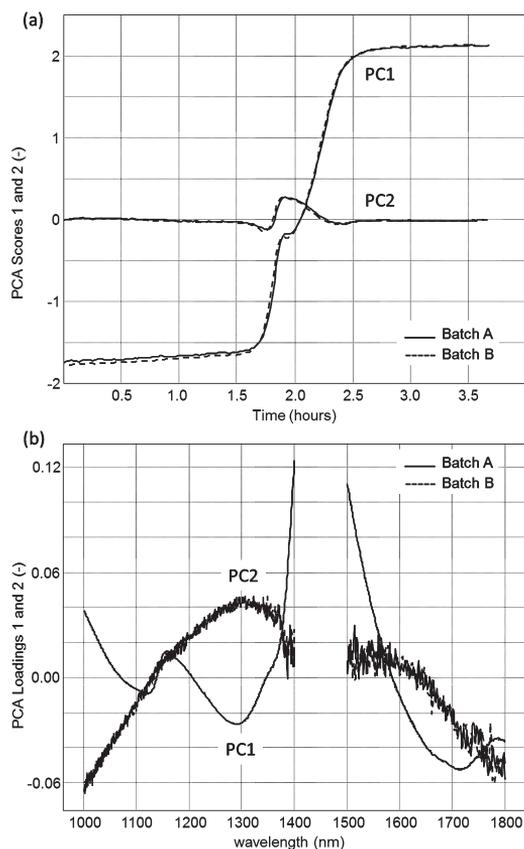


Figure 4. (a) NIR principle component one (PC1) and principle component two (PC2) score values as a function of time for two batches; (b) PC1 (99% variance explained for both batches) and PC2 (<1% variance explained) variable-loading vectors.

pseudo-colour map in Fig. 5(a). By stacking different measurements (samples taken over time) on top of each other we thus get a three-way data structure of excitation-by-emission-by-batch time. The most dominating fluorescence peaks found in the landscape were investigated as a function of time using selected emission-excitation wavelength couples (Fig. 5(b)–(d), effectively ‘taking a tubular vector’ first-order signal from the data cube), to explore if any changes of chemical compounds can be detected in the raw spectra. Despite the erratic pattern, some profiles can be distinguished with transitions close to the expected point around 2 h into the process. To get the most valuable information out of this data structure a tensor algorithm like PARAFAC is required.¹⁷

When the fluorescence data cubes were modelled by PARAFAC making separate models for each of the two selected batches, three components were identified as being significant. This means that three different fluorescence components were found to vary over batch-time. The results from the PARAFAC models are illustrated in Figs 6 and 7, where the loadings and the scores are shown, respectively. In Fig. 6, the three fluorophores are shown as landscapes (reconstructed as outer- or matrix-products of the emission

and excitation loading vectors,¹⁸ see Fig. 1) and the three signals are in agreement with the three signals outlined in Fig. 5. The bias between profiles of the two batches in Fig. 7 can be explained by repositioning of the instrument (outside the reactor vessel, hence looking non-invasively through the glass reactor wall) in-between batch-runs. Note that the fluorescence in this investigation was used as an exploratory/investigative technique rather than for quantification purposes.

The fluorescence component found at the excitation/emission maximum around 470/510 nm (Figs 6(a) and 7(a)) can be identified as riboflavin, a well-known fluorophore in dairy products, whereas the fluorescence component with an excitation/emission top around 310/370 nm (Figs 6(c) and 7(c)) could represent tryptophan, also a well-known fluorescence active compound. Both riboflavin and tryptophan are expected to be found in yoghurt¹⁹ and it has been listed before that riboflavin can be found at wavelengths 445/520 nm and tryptophan at 285/364 nm,²⁰ which agrees well with the results obtained considering these measurements were collected using a process instrument with very limited wavelength resolution. It appears that riboflavin increases over time, sharply in the first hour and right after the sharp drop in brix around 2 h. Some studies^{21,22} found that riboflavin is produced by *Lactobacillus fermentum* and *Lactococcus lactis*, respectively. However, the starter culture used in this study contains the lactic acid bacteria *Lactobacillus bulgaricus* and *Streptococcus thermophilus* and no studies were found describing whether riboflavin is produced or not by these strains. On the other hand, earlier studies also revealed that ATP is found at excitation/emission maximum 292/392 nm (Food Fluorescence Library, www.models.life.ku.dk), which is close to the excitation/emission area where tryptophan has been reported previously (285/364 nm). Therefore, the fluorescence signal found at the excitation/emission maximum around 310/370 nm could represent either ATP or tryptophan. However, ATP is a product in the glycolysis, meaning that an increasing and not a decreasing trend, as seen in the PARAFAC scores in Fig. 7(c), over time are expected. On the other hand, it is well known from the literature that the amino acid tryptophan is a nutrient consumed by lactic acid bacteria.²³ Hence, the trend seen in Fig. 7(c), described by the third PARAFAC score, is most probably tryptophan.

The second PARAFAC score – excitation/emission maximum around 350/450 nm (Fig. 6(b)) – is systematically increasing over time (Fig. 7(b)). The fluorescence profile could be representing lumichrome, which is a degradation product from riboflavin and has previously been defined at an excitation/emission wavelength around 360/450 nm.²⁴ Based on previous studies the profile could also represent NADH, which was reported around 340/465 nm. NADH is formed in the glycolysis, when glucose is converted into pyruvate.²⁵ The glycolysis takes place when the milk sugar lactose, which is a disaccharide containing glucose and its isomer galactose, has been first phosphorylated into lactose-6-phosphate and then further hydrolysed into glucose and galactose-6-phosphate.²⁶

It can be interpreted from the PARAFAC models that the fluorescence signal obtained from yoghurt fermentation arises from riboflavin, tryptophan and lumichrome or NADH. Despite a lower signal-to-noise ratio compared with NIR and an offset due to interfacing issues – there is a very clear similarity between the two batch runs. Furthermore, based on the uniqueness properties and second-order advantage of the PARAFAC model, it is possible to use the score-profiles as illustrated in Fig. 7 directly as pseudo-concentration-profiles over time based on the

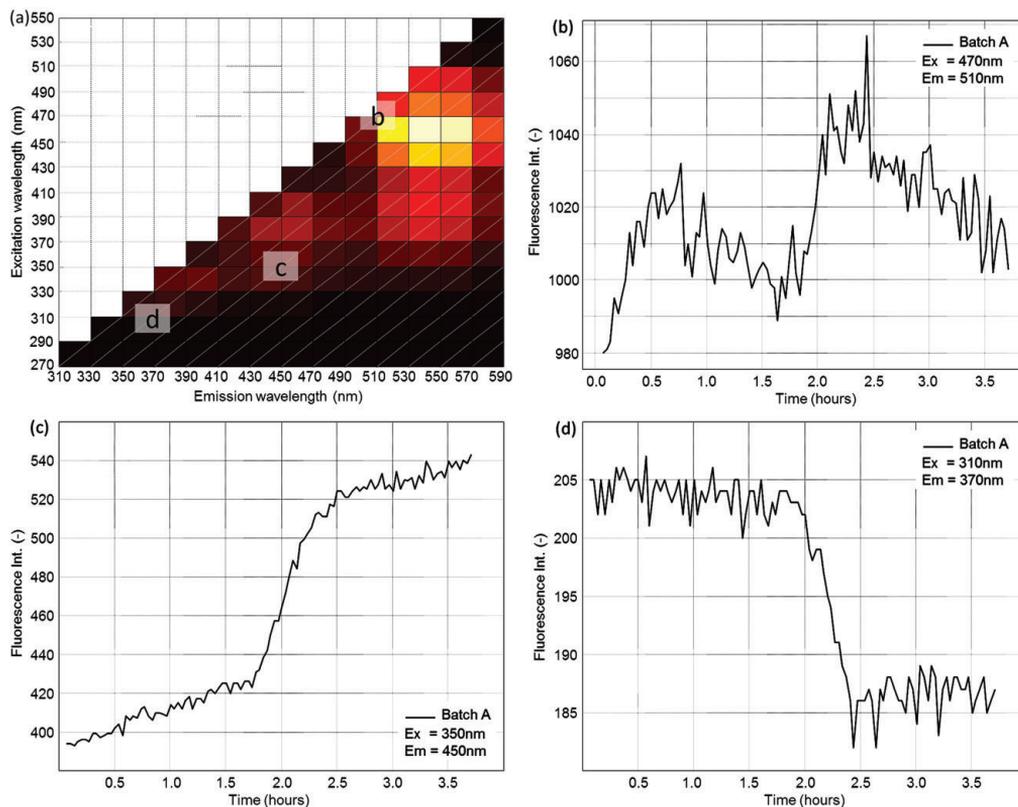


Figure 5. (a) Fluorescence landscapes obtained from Batch A (recorded at 0 h); time profiles for three expected fluorophores found in the raw data at Ex/Em 470/510 (b), 350/450 (c) and 310/370 nm (d).

fluorophores, without limitations like sign ambiguity or interfering species.

Multiblock modelling

In the previous sections different signals (brix, pH, NIR and fluorescence) for different batch runs have been modelled and interpreted separately, where a (visual) comparison could be made as the end-stage of data analysis. However, multiblock analysis can be applied to employ several data blocks simultaneously, thereby simplifying the overall interpretation of the various batch parameters.²⁷ In this section the zero-order pH measurements, the two latent variables obtained from the PCA analysis on the NIR measurements and the three latent variables obtained from the PARAFAC model on the fluorescence measurements are gathered in a single data matrix (six variables in total) for each of five batches (Batch A and B presented before, supplemented with three more runs). The five batches are modelled individually (at the so-called block level) and combined (at the so-called super level) using Consensus PCA (CPCA²⁸), as illustrated in Fig. 8. The block level makes it possible to compare trends and patterns among the batches; the super level gives as outcome the overall trend. The interpretation thus becomes more holistic and simplified. It was chosen not to include the brix measurements, as a minimum of data points were collected compared with the other techniques. Matching

data points from the NIR PCA scores and pH measurements were selected based on the fluorescence time-scale, in order to create vectors with similar length (113 time points for each batch, visually aligned between batches). The blocks were auto-scaled and two principal components were determined.

The block and super scores and block loadings for PC1 and PC2, are plotted in Fig. 9(a) and Fig. 9(b), respectively.

PC1 describes an increase over time as observed before for some individual measurement signals, while the main feature in PC2 is the steep, characteristic pattern between 2 and 2.5 h into the process (Fig. 9(a)). The scores on PC1 seem to be very similar for all batches, while the PC2 score-profile between the batches varies more. It should be noted that the difference in magnitude between block-level and super scores is a consequence of the CPCA algorithm only (which could easily be removed by a post-analysis normalization). Interpretation of the block-loading plots for PC1 shows that Fluor₁, Fluor₂ and NIR₁ are positively correlated with the trend observed for PC1 while Fluor₃ and pH are negatively correlated. This corresponds well with the earlier observations in Figs 4(a) (NIR), 5(b) and 5(c) (Fluorescence). Furthermore, the negative correlations observed are consistent with the decreasing trend seen for the fluorescence data (Fig. 5(d)) and for the pH measurements (Fig. 2).

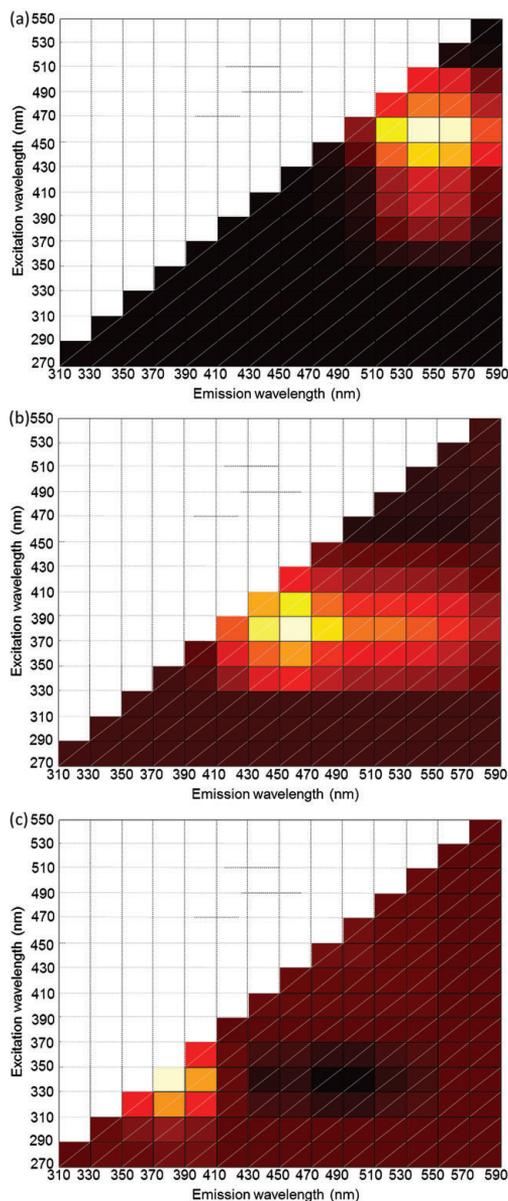


Figure 6. The PARAFAC-loading (Batch A) for (a) factor 1; (b) factor 2; (c) factor 3.

The block loadings for PC2 shows that NIR₂ is negatively correlated with the PC2 trend observed in the score plot, which it is consistent with the bump observed for PC2 for the NIR data (Fig. 4(a)). Furthermore, it is observed that Fluor₁ and Fluor₃ are slightly correlated with the trend observed for PC2.

The block loading plot of PC1 and PC2 (Fig. 10) illustrates that the parameters (pH, NIR₁, NIR₂, Fluor₁, Fluor₂, Fluor₃) are very alike

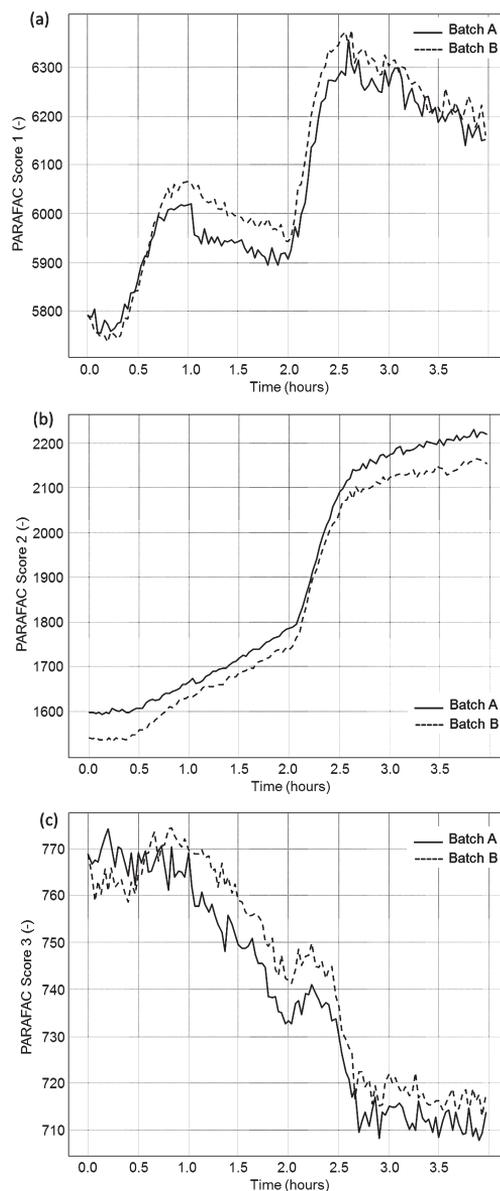


Figure 7. PARAFAC time-score values for two batches of (a) factor 1; (b) factor 2; (c) factor 3.

over the five batches as they are clearly grouped together. Furthermore, NIR₁ and Fluor₂ are highly correlated again indicating that PC1 obtained from the PCA model on the NIR data and factor 2 obtained from the PARAFAC model on the fluorescence data are very alike, and cannot be differentiated from each other in the two component CPCA model. NIR₁ and Fluor₂ are also correlated with Fluor₁ and pH is correlated with Fluor₃ both describing a

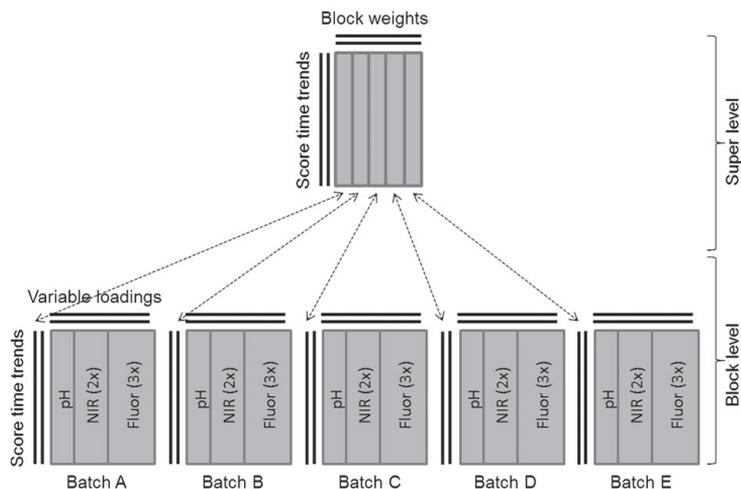


Figure 8. Symbolizing the consensus principal component analysis model.

decreasing trend over time. The last cluster of NIR₂ is separated from the other variables, which corresponds well with the characteristic bump described by PC2 obtained from the PCA model on the NIR data.

DISCUSSION

A single pH or brix-measurement by itself is a zero-order signal, and when these zero-order-measurements are collected as a function of time as was done in the current study, they turn into a vector or first-order tensor.²⁹ Interpretations of a single pH or brix data point does reveal process information for a specific time but this information typically only has value for process monitoring when interpreted in relation to the neighbouring values measured over time. From the pH- and brix-curves in Fig. 2 it is possible to compare the process status of several runs (e.g. under normal operating conditions) but it is not possible to achieve quantification of individual chemical compounds. Besides, these zero order measurements do not allow for checking if the measurements are actually representing the expected chemistry or partly due to interfering compounds. Again, measurements over time can help to elucidate this but there is a risk that process disturbances might not be observed quickly enough.

More information is gained from the NIR spectra. In general, first-order instruments, such as NIR-spectroscopy, are more powerful analytical tools since distinctive signals in the spectrum such as molecule specific vibrations can be used to quantify chemical compounds.⁸ With first-order NIR measurements carried out over time the data-outcome is a second-order tensor or a data matrix. This allows us to follow principal profiles over time which gives the opportunity to monitor the fermentation process with a more comprehensive understanding, at a more abstract or principal level. A related advantage is the ability to detect interfering chemical species based on differing spectral outcome without the need for a direct chemical or biological interpretation.

The first PCA scores obtained from the modelled NIR data (Fig. 4(a)) clearly shows an increasing trend over time. A small break in the rise is observed at time 1.9 h. The curve could very well

describe the bacterial growth of *Lactobacillus bulgaricus* (LB) and *Streptococcus thermophilus* (ST). It can be speculated what is causing the small bump. A possible reason could be that the growth of ST stops at this point and the growth of LB takes over. Earlier studies reveal that ST has a lower pH growth optimum compared with LB, which is in agreement with the PCA-curve (Fig. 4(a)). This could also mean that the curve is not directly linked to the growth as such, but the formation of lactate, where the increase before the bump is caused by L-lactate produced by ST and the increase after the bump is caused by D-lactate produced by LB. On the other hand, if the observed trend describe the bacterial growth, the observed bump on the curve could also be related to the redox potential. Since the NIR measurements are obtained in the yoghurt broth and only limited stirring has been applied a restricted amount of oxygen will be available. Thus, the bump could be caused by the fact that all the available oxygen has been used and the growth, therefore, stops until another pathway is started for further growth. This dual interpretation illustrates a limitation in NIR spectroscopy – the *de facto* workhorse in process analytical chemistry and technology – it is a very powerful but indirect measurement technique that ‘sees everything’ in the biological/food system. Therefore, selectivity in the multivariate/first-order measurement is necessary to ensure that the signal does not contain contributions from other sources than the one of interest.³⁰ As stated previously, in the situation where inferring species appear during a batch run it is not possible to detect them from the univariate signal like brix while we can often spot the issue by multivariate measurements like NIR spectroscopy. But it is not possible to quantify by NIR signals under these circumstances unless the interference problem was present/represented during the calibration stage – this is the reason why (often prohibitively) large calibration sets are obligatory.

The fluorescence measurements are second-order measurements consisting of an excitation wavelength and emission wavelength, thus each measurement has a data matrix structure. Since the fluorescence measurements are obtained continuously over time, the final obtained data for each batch is a three-way data array or a data cube. With second-order measurements follows

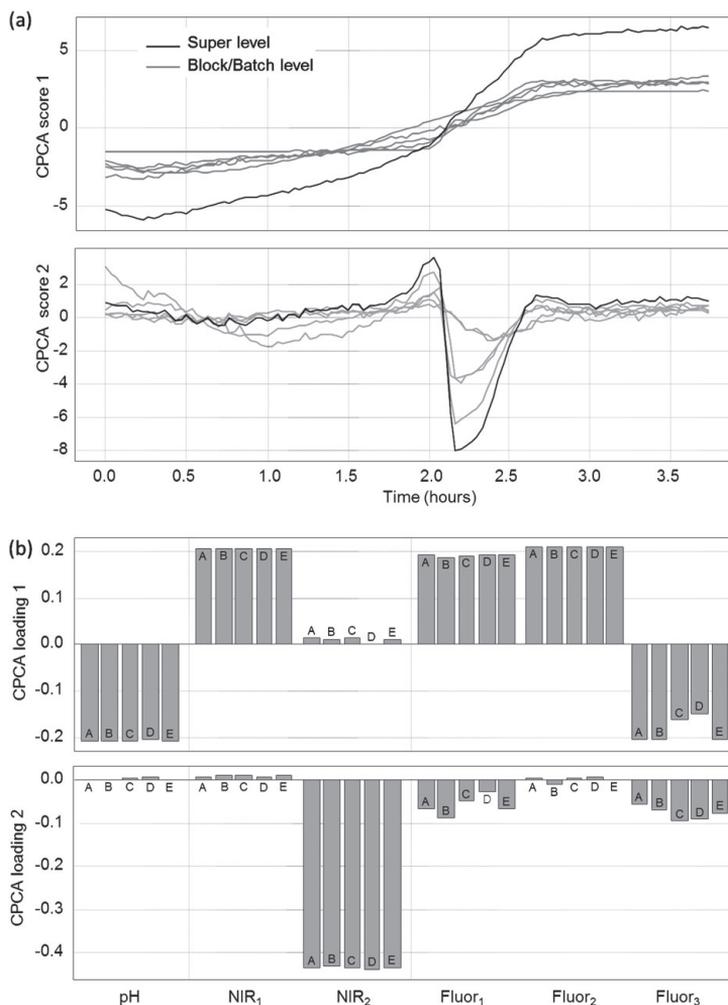


Figure 9. (a) PC1 and PC2 scores for five batches/blocks (grey) and super level (black); (b) block-level loadings for PC1 and PC2.

the second-order advantages,³¹ which means that it is possible to predict concentrations of analytes even when unknown interference appears. Thereby, it also avoids the need to specify each analyte (as long as different components can be sufficiently resolved by the PARAFAC algorithm). Another advantage of second-order measurements, which can also be achieved by first-order measurements, is that more than one analyte can be determined in one measurement.³² Furthermore, the three-way data structures are different from zero- and first-order data because the decomposition of a data cube is often unique, whereas a data matrix decomposition is never unique.⁸

Separate modelling and interpretation of the individual measurements gives valuable information about the process, repeatability of runs, etc. However, when handling a lot of different techniques, each producing various process parameters, a general overview of the different trends and patterns may be

valuable as well. By applying multiblock analysis the combined or consensus interpretation of the trends happening over time and the patterns among the batches as well as among the different process variables becomes a lot more convenient. This more holistic way of analysing data sets consisting of several batches can be an advantage, but it also carries a risk because it might not be the largest variation in data that are the most important for the process. A very useful side effect of multiblock analysis is that the correlation among the process parameters can be easily interpreted. This advantage might even make it possible to perform sensor and variable selection to some extent.

CONCLUSIONS

Based on the in-process techniques utilized and the results obtained it seems that most chemically relevant information

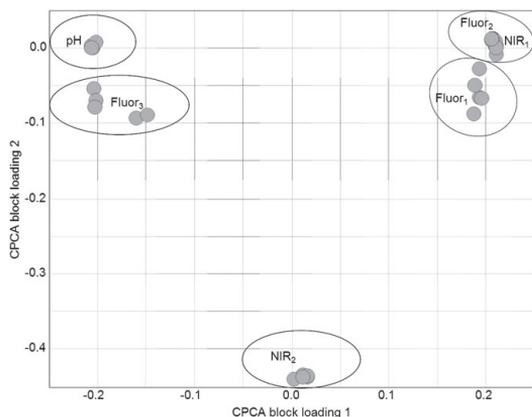


Figure 10. Block loading plot of PC1 vs PC2.

on our system was obtained by the second-order technique fluorescence spectroscopy, where pseudo-concentrations over time could be obtained for three compounds. The increasing trend found from the NIR measurement also seems to be very consistent. However, it can only be interpreted indirectly since NIR results by themselves do not provide sufficient chemical information. If additional runs are conducted to elucidate the observed trend, the NIR results could be a strong technique for monitoring this bio-process, albeit using abstract or principal phenomena. Both first- and second-order measurements collected over time provide a multivariate data set. These data sets might contain important relationships among the different variables that can provide chemical and process-related information when multivariate data analysis is applied, information that cannot be readily observed from the raw data. Though first- and second-order measurements seem to obtain more information than the zero-order measurements, it is important to keep in mind that zero-order techniques such as pH, OD600 and brix can provide valuable information about the process, especially by a combination of different sensors.³³ When handling large amounts of batches and measurement principles, multiblock analysis might be an advantage for easier model interpretation, where comparison of batches and correlations among process parameters are clearly outlined. Hence, a holistic pattern of the collected data is presented in a suitable way.

ACKNOWLEDGEMENTS

The research is sponsored by BIOPRO (www.biopro.nu), which is funded by Region Zealand (Denmark), the European Regional Development Fund (ERDF) and the Partners.

REFERENCES

- 1 Van Steenkiste F, Baert K, Debruyker D, *et al.*, A microsensor array for biochemical sensing. *Sensors Actuators B - Chem* **44**:409–412 (1997).
- 2 Saucedo-Castaneda G, Trejo-Hernandez MR, Lonsane BK, *et al.*, Online automated monitoring and control-systems for Co₂ and O₂ in aerobic and anaerobic solid-state fermentations. *Process Biochem* **9**:13–24 (1994).
- 3 Sivakesava S, Irudayaraj J and Ali D, Simultaneous determination of multiple components in lactic acid fermentation using FT-MIR, NIR, and FT-raman spectroscopic techniques. *Process Biochem* **37**:371–378 (2001).
- 4 Vaccari G, Dosi E, Campi AL, Gonzalezvara A, Matteuzzi D and Mantovani G, A near-infrared spectroscopy technique for the control of fermentation processes - an application to lactic-acid fermentation. *Biotechnol Bioeng* **43**:913–917 (1994).
- 5 Marose S, Lindemann C, Ulber R and Scheper T, Optical sensor systems for bioprocess monitoring. *Trends Biotechnol* **17**:30–34 (1999).
- 6 Boehl D, Solle D, Hitzmann B and Scheper T, Chemometric modelling with two-dimensional fluorescence data for claviceps purpurea bioprocess characterization. *J Biotechnol* **105**:179–188 (2003).
- 7 Ronnest NP, Stocks SM, Lantz AE and Gernaey KV, Introducing process analytical technology (PAT) in filamentous cultivation process development: comparison of advanced online sensors for biomass measurement. *J Ind Microbiol Biotechnol* **38**:1679–1690(2011).
- 8 Booksh KS and Kowalski BR, Theory of analytical-chemistry. *Anal Chem* **66**:A782–A791 (1994).
- 9 Babamoradi H, van den Berg F and Rinnan A, Comparison of bootstrap and asymptotic confidence limits for control charts in batch MSPC strategies. *Chemometrics Intellig Lab Syst* **127**:102–111 (2013).
- 10 Klimkiewicz A, Mortensen PP, Zachariassen CB and van den Berg FWJ, Monitoring an enzyme purification process using on-line and in-line NIR measurements. *Chemometrics Intellig Lab Syst* **132**:30–38 (2014).
- 11 Thygesen JH and van den Berg FWJ, Subspace methods for dynamic model estimation in PAT applications. *J Chemometrics* **26**:435–441 (2012).
- 12 Smilde A, Westerhuis J and de Jong S, A framework for sequential multiblock component methods. *J Chemometrics* **17**:323–337 (2003).
- 13 De Brabandere A and De Baerdemaeker J, Effects of process conditions on the pH development during yogurt fermentation. *J Food Eng* **41**:221–227 (1999).
- 14 Workman J, Interpretive spectroscopy for near infrared. *Appl Spectrosc Rev* **31**:251–320 (1996).
- 15 He Y, Wu D, Feng S and Li X, Fast measurement of sugar content of yogurt using vis/NIR-spectroscopy. *Int J Food Prop* **10**:1–7 (2007).
- 16 Massart DL, Vandeginste BGM, Buydens LMC, De Jong D, Lewi PJ and Smeyers-Verbeke J, *Handbook of Chemometrics and Qualimetrics: Part A (Chapter 1. introduction)*. Elsevier, Netherlands, 1 (1997).
- 17 Wold S, Esbensen K and Geladi P, Principal component analysis. *Chemometrics Intell Lab Syst* **2**:37–52 (1987).
- 18 Murphy KR, Stedmon CA, Graeber D and Bro R, Fluorescence spectroscopy and multi-way techniques. PARAFAC. *Anal Methods* **5**:6557–6566 (2013).
- 19 Bro R, PARAFAC. Tutorial and applications. *Chemometrics Intell Lab Syst* **38**:149–171 (1997).
- 20 Christensen J, Becker E and Frederiksen C, Fluorescence spectroscopy and PARAFAC in the analysis of yogurt. *Chemometrics Intell Lab Syst* **75**:201–208(2005).
- 21 Duggan D, Bowman R, Brodie B and Udenfriend S, A spectrophotofluorometric study of compounds of biological interest. *Arch Biochem Biophys* **68**:1–14 (1957).
- 22 Jayashree S, Rajendhran J, Jayaraman K, Kalaichelvan G and Gunasekaran P, Improvement of riboflavin production by lactobacillus fermentum isolated from yogurt. *Food Biotechnol* **25**:240–251 (2011).
- 23 Hugenholtz J, Sybesma W, Groot M, *et al.*, Metabolic engineering of lactic acid bacteria for the production of nutraceuticals. *Antonie Van Leeuwenhoek Int J Gen Mol Microbiol* **82**:217–235 (2002).
- 24 Chervaux C, Ehrlich S and Maguin E, Physiological study of lactobacillus delbrueckii subsp bulgaricus strains in a novel chemically defined medium. *Appl Environ Microbiol* **66**:5306–5311 (2000).
- 25 Fox J and Thayer D, Radical oxidation of riboflavin. *Int J Vitamin Nutr Res* **68**:174–180 (1998).
- 26 Pratt CW and Cornely K, *Essential Biochemistry (chapter 10 - glucose metabolism)*, 1st edn. John Wiley & Sons, Inc; USA, 302–341 (2004).
- 27 Walstra P, Wouters J and Geurts T, Chapter 13 - lactic fermentations, in *Dairy Science and Technology*, 2nd edn, ed by Walstra P, Wouters J and Geurts T. CRC Press; Boca Raton, USA, 357–398 (2005).
- 28 Wold S, Kettaneh N and Tjessem K, Hierarchical multiblock PLS and PC models for easier model interpretation and as an alternative to variable selection. *J Chemometrics* **10**:463–482 (1996).

- 29 Olivieri AC, Analytical advantages of multivariate data processing. One, two, three, infinity? *Anal Chem* **80**:5713–5720 (2008).
- 30 Bro R, Review on multiway analysis in chemistry - 2000–2005. *Crit Rev Anal Chem* **36**:279–293 (2006).
- 31 Sanchez E and Kowalski BR, Tensorial calibration. II. second-order calibration. *J Chemometrics* **2**:265–280 (1988).
- 32 Gomez V and Callao MP, Analytical applications of second-order calibration methods. *Anal Chim Acta* **627**:169–183 (2008).
- 33 Luttmann R, Bracewell DG, Cornelissen G, *et al.* Soft sensors in bio-processing: a status report and recommendations. *Biotechnol J* **7**:1040–1048 (2012).

PAPER II

**Exploring process dynamics by near infrared spectroscopy in
lactic fermentations**

Carina Svendsen, Tomasz Cieplak and Frans W. van den Berg

Journal of Near Infrared Spectroscopy (2016), *article in press*

Exploring process dynamics by near infrared spectroscopy in lactic fermentations

Carina Svendsen^{*a}, Tomasz Cieplak^b and Frans W. J. van den Berg^a

^a) Spectroscopy and Chemometrics section (SPECC), Department of Food Science, Faculty of Science, University of Copenhagen, Rolighedsvej 26, 1958 Frederiksberg C, Denmark. E-mail: carinasv@food.ku.dk

^b) Food Microbiology section, The Department of Food Science, Faculty of Science, University of Copenhagen, Rolighedsvej 26, 1958 Frederiksberg C, Denmark

ABSTRACT: In the industrial production of yoghurt, pH is normally the only on-line technique applied as a real-time monitoring signal to follow the dynamics during the fermentation process. However every dairy company would benefit from an online technique giving information about the chemical composition, physical/textual properties and/or microbial contamination. In this study lactic fermentation batches with the starter bacteria *Streptococcus thermophilus* and *Lactobacillus delbrueckii* subsp. *bulgaricus* are explored by online NIR spectroscopy. The dynamics obtained from near infrared (NIR) spectroscopy is elucidated in order to explain what causes the variation of the indirect measurements. The results show that the viscosity change has a large impact on the scatter, which affects the NIR data. It seems that the variation found by NIR spectroscopy is caused both by scatter and absorbance, where the scatter gives us information about the textural change happening, and the absorbance gives us information about the biomass formation plus the conversion of sugar into lactic acid.

Keywords: • Near Infrared Spectroscopy • On-line Measurements • Yoghurt • Fermentation • Monitoring • *Streptococcus thermophilus* • *Lactobacillus delbrueckii* subsp. *bulgaricus* • Dynamics • Growth associations

INTRODUCTION

In the production of fermented milk products, such as yoghurts and cheeses, mixed-strain starter cultures with selected lactic acid bacteria are applied as standard. Composing the stain-mixture is far from trivial because it determines the flavour and physical perception of the end product. In yoghurt production the bacterial strains *Streptococcus thermophilus* (*S. thermophilus*) and *Lactobacillus delbrueckii* subsp. *bulgaricus* (*L. bulgaricus*) are, as a minimum, inoculated into heat treated milk to start the fermentation. In this process the natural milk sugars are consumed by the bacteria and thereby converted into lactic acid via the bacterial metabolisms. The lactose is transported into the cell without any chemical modification via cytoplasmic protein proteases. Inside the cell lactose is hydrolysed to glucose and galactose. Glucose is further metabolised via the glycolysis and converted into pyruvate, which is then synthesized to lactate via the enzyme lactate dehydrogenase. The lactate is finally converted into lactic acid, which can exist in two different isomer forms, the L- and the D-isomer, where *S. thermophilus* mainly produces L(+) lactic acid and *L. bulgaricus* mainly produces D(-) lactic acid¹. In the beginning of the fermentation process *S. thermophilus* is more active than *L. bulgaricus*. Later in the process the vitality of *S. thermophilus* will decrease because of the increasing concentration of lactic acid and the more acid tolerant *L. bulgaricus* will become the dominant strain. Hence, the first part of the fermentation is driven by *S. thermophilus*, whereas the last part of the fermentation is mostly driven by *L. bulgaricus*². During the metabolic processes several interactions between the two bacterial strains appear, meaning that a so-called growth association between *S. thermophilus* and *L. bulgaricus*

in the dynamics of yoghurt fermentation exist, where each organism provides nutritional compounds which benefit the other^{3,4}. This growth symbiosis or protocoeperation relationship has been well studied⁵. It has e.g. been reported that *S. thermophilus* provides formic acids and carbon dioxide stimulating the growth of *L. bulgaricus*^{6,7}, whereas *L. bulgaricus* provides various amino acids stimulating the growth of *S. thermophilus*^{4,8}. The yoghurt fermentation process does not only results in production of lactic acid and various aroma and flavour compounds characteristic for the yoghurt product, but it also leads to a textural change. When the lactic acid is released from the cells, the acidity increases and causes milk proteins to denature and thereby coagulating into a solid mass, also called curd⁹. The coagulation starts around pH 5.15, where the milk proteins have an isoelectric point. At pH 4.65 the coagulation is completed and thiol-disulfide bridges will link α -lactalbumin and β -lactoglobulin with the κ -casein, which will result in gel network formation¹⁰. Many investigations have identified the growth symbiosis and physical changes in the system as very characteristics for the yoghurt fermentation and as a crucial factor for the end product, both from a taste and textural perspective.

In industrial yoghurt fermentation it is of interest to achieve as high a consistency and reproducibility as possible to ensure that the consumers will receive a maximum quality product, which does not change from day to day. And this is again far from a trivial task because starter cultures of mixed strains are living systems. At present the only real-time measurements applied to follow the fermentation dynamics in yoghurt production is in-line pH. For that reason various parameters that may affect the metabolism of the starter bacteria (hence, starting conditions of the batch) or the end product characteristics (hence, post-run quality assurance) are also determined. Some of the most commonly

used parameters for quality control of the raw material, during the fermentation process and of the end product, are inferential measurements such as end-pH, water activity and chemical measurements (protein and fat) of the milk¹¹. Determination of the key compounds involved in lactic acid fermentation - residual sugars, biomass of the cultures, etc. - can be time consuming and often requires sample preparation and the use of off-line techniques such as gas- or liquid-chromatography. More or less any chemical desired parameter can be measure off-line in the laboratory, but such measurements cannot replace a fast in-line technique allowing the production team or an automatic control system to regulate the process when unwanted changes or disturbances are observed¹². The production site would benefit from a real-time measure giving chemical and/or physical information to follow the dynamics of the fermentation and indicators about microbial contamination.

Near infrared (NIR) spectroscopy is a measurement technique with the potential of allowing rapid and accurate determination of chemical composition. Previously studies have shown that NIR spectroscopy can be used for quality control of milk products¹³. The study of Grassi *et al.* (2013)¹⁴ showed that at-line NIR spectroscopy can be applied for the determination of curd development during fermentation. The study of Vaccari *et al.* (1994)¹⁵ claims that lactic acid, glucose and biomass concentration can be measured from NIR spectra. Additionally, Navrátil *et al.* (2004)¹⁶ suggests that NIR spectroscopy has a potential for on-line monitoring of and valuation of process quality of yoghurt fermentations, and the investigation by Lyngaard *et al.* (2012)¹⁷ illustrates how NIR spectra can be used to follow (cheese) curd formation real-time. In the present study the fermentation dynamics of yoghurt as well as the interactions between the starter bacteria *S. thermophilus* and *L. bulgaricus* and the physico-chemical change were investigated using NIR spectroscopy.

Experimental system

Fermentation batches were carried out using the starter culture YF-3331 (a mixture of the lactic acid bacteria *L. bulgaricus* and *S. thermophiles* with a concentration of 0.2%(w/w)). The culture was provided by Chr. Hansen A/S (Hørsholm, Denmark) and stored at -45 °C. Skimmed milk powder was used as media by dissolving 1200 gram (Arla Food Ingredients, Denmark) in 11 L of water. The fermentations were carried out in a 15 L in-house modified glass single-wall fermenter vessel (Applikon, Delft, The Netherlands) with a working volume of approximately 12 L. The lid of the vessel contained 6 instrument ports, a thermowell for temperature monitoring and a hollow flow-through baffle-ring for temperature regulation inside the fermenter. A stirrer with two flat-bladed Rushton impellers and three baffles for optimal mixing were situated in the fermenter. The mixing speed was kept constant at 150 rpm and controlled by an IKA EUROSTAR 60 control motor (IKA-Werke GmbH & Co. KG, Staufen, Germany). It should be mentioned that in industrial yoghurt production stirring is normally not applied. In our set-up it was however chosen to apply gentle stirring to make sure that the broth was homogenous and thereby gain more representative samples. The temperature was maintained at the set point using a Pt100 probe inserted in the thermowell (filled with water for conduction) and controlled by a HAAKE Phoenix pumping water bath (Thermo Scientific, Karlsruhe, Germany). Once the set temperature was reached the instruments were switch on and 22 g of starter culture was inoculated into the media (*S. thermophilus* 6.8×10^{10} cfu·g⁻¹ and *L.bulgaricus*: 5.5×10^9 cfu·g⁻¹; cfu = colony forming units). Grab samples were frequently, but not equidistant, withdrawn during the fermentation. They were transferred to Eppendorf tubes and stored at -20 °C until further analyses. In total seven batches (Batch 1-7) were performed at varying fermentation temperatures. Batches 1-4 were carried out with the nominal fermentation temperature of 35 °C, Batch 5 was carried out at 32 °C and

Batches 6-7 were carried out at approximately 37.5 °C. The nominal duration of a batch run was 5 hours.

In-line, at-line and off-line measurements

pH measurements

pH measurements were performed in-line with an interval of 30 seconds, using a pH-meter plus temperature logger (MadgeTech Inc., Warner, NH-US) placed in the fermentation broth. The pH probe was calibrated in pH 4.01 and pH 7.00 before use.

NIR spectroscopy

NIR spectroscopy measurements were performed in-line with an ABB Bomem spectrometer (ABB Bomem, Quebec, Canada) equipped with a fibre optic reflectance probe placed in the fermentation broth. The spectral data for each measurement was collected as the average of 64 single beam spectra for each measurement time point. The spectra were referenced against a white Spectralon ([LabSphere, North Sutton, NH-US](#)) background (average of 64 scans) collected before the process measurements were started. The fermentation broth was scanned over the range 10000-5556 cm⁻¹ (1000-1800 nm) with a resolution of 8 cm⁻¹ and a time interval of 60 seconds between the measurements.

Real-Time Quantitative Polymerase Chain Reaction (qPCR) analysis

Real-time qPCR was used for quantification of *S. thermophilus* and *L. bulgaricus* during the fermentation. To remove exopolysaccharides prior to DNA extraction 1 ml of yoghurt sample was mixed with 7.5 ml of MillQ water, 1 ml of 18% sodium citrate and 0.5 ml of 1 M NaOH. After 5 min the mixture was centrifuged for 10 min at 10,000 g at room temperature¹⁸. The pellet was washed in water and used for DNA extraction using the PowerSoil DNA isolation Kit (MO BIO Laboratories, Inc. Caelsbad, CA-US) following the instructions of the manufacturer. The isolated DNA was subsequently

cleaned by PowerClean® Pro DNA Clean-UP Kit (MO BIO Laboratories, Inc. Caelsbad, CA-US). Absolute abundance of *Streptococcus thermophilus* and *L. bulgaricus* were determined using a 7500 Fast Real-time PCR System (Applied Biosystems, Foster City, CA-US). All samples were analysed in two separate runs in duplicate. Two primer sets, specific for *L. bulgaricus*; Ld1F, Ld2R; and *S. thermophilus*; St1F, St2R¹⁹, were used in this assay. The reaction mixture (20 µL) consisted of 1× SYBR green PCR Master Mix (Applied Biosystems), 1 µL of either the primers specific for *L. bulgaricus* or *S. thermophilus* at a final concentration of 0.25 µM, and 5 µL of template DNA. The qPCR temperature profile was as follows: 95 °C for 5 min, followed by 40 cycles of 95 °C for 15 sec, 60 °C for 1 min. The fluorescence acquiring was set in the annealing/extension step. Serial tenfold dilutions of *L. bulgaricus* and *S. thermophilus* pure culture genomic DNA were used to generate standard curves. Prior to DNA extraction of pure cultures, quantification of cell numbers was carried out by cell counting in a Thoma chamber.

Brix measurements

At-line Brix measurements were carried out on grab samples frequently during the fermentation process using a DR-103 meter (Index Instruments Limited, Cambridgeshire, UK).

D- and L-Lactic acid analyses

Determination of D- and L-lactate was carried out with a D- and L- Lactate enzyme kit (K-DLATE 07/14, Megazyme International Ireland, Bray, Ireland) for a selected number of batch runs. From the samples, withdrawn during the fermentation, 1 g was transferred to a 20 ml glass beaker. For sample clarification 10 ml of water, 400 µL of Carrez I (3.6 g of potassium hexacyanoferrate(II)trihydrate ($K_4[Fe(CN)_6] \cdot 7H_2O$) diluted in 100 mL of distilled water), 400 µL of Carrez II (7.2 g of zinc sulphate heptahydrate ($ZnSO_4 \cdot 7H_2O$) dissolved in 100 ml distilled water) and 800 µL of 100 mM NaOH were

added one by one and mixed after each addition. Finally, 7.4 ml of distilled water was added and the solution was mixed and filtered. Hereafter, the absorbance at 340 nm was measured (Elution 220 UV-visible spectrometer, Thermo Scientific, Denmark) to obtain a quantification for D- and L-lactic acid.

Software

Data were analysed using Matlab (Matlab, R2015a, The Mathworks, Inc., Natick, MA-US). The NIR data were modelled by Principal Component Analysis (PCA) and non-linear regression using in-house routines and the PLS-Toolbox (PLS-Toolbox 7.5, Eigenvector Research Inc., Manson, WA-US).

RESULTS AND DISCUSSION

Temperature and pH

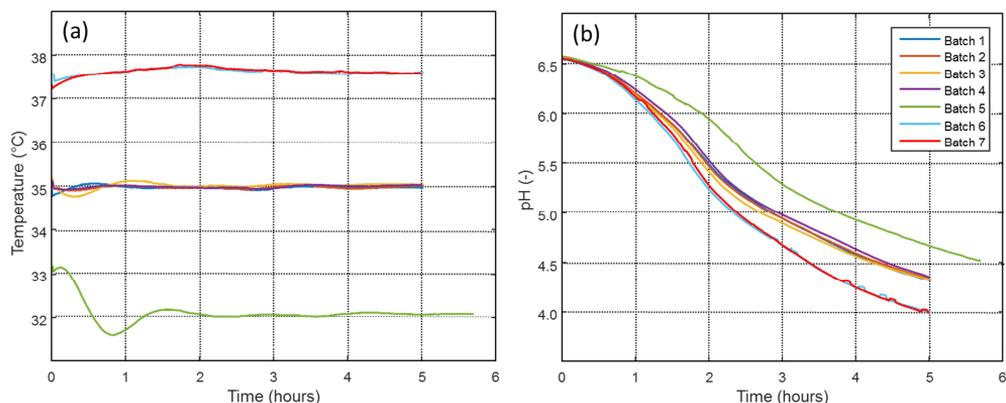


Figure 1: (a) Temperature and (b) pH profiles for the seven batches.

The temperature and pH profiles for the seven batches are plotted over fermentation time in Figure 1. The temperature profile of Batch 5 was unstable in the beginning of the fermentation because the stirring was turned on a little later than under normal conditions. Therefore the broth was not

homogeneous and temperature gradients were probably present until the stirring was turned on. Batch 5 was also run at a lower temperature and it is thus expected to be overall delayed; for this reason it was followed for a longer period (5h50m). All batches have a starting pH of 6.5 and during the fermentation time the pH drops in the characteristic S-shape, confirming that the sugar is converted into acid and the fermentation process is proceeding. This is the conventional way of monitoring fermentation dynamics in yoghurt production. It is e.g. seen that the higher the fermentation temperatures the faster the decrease of pH and vice versa.

NIR Spectroscopy

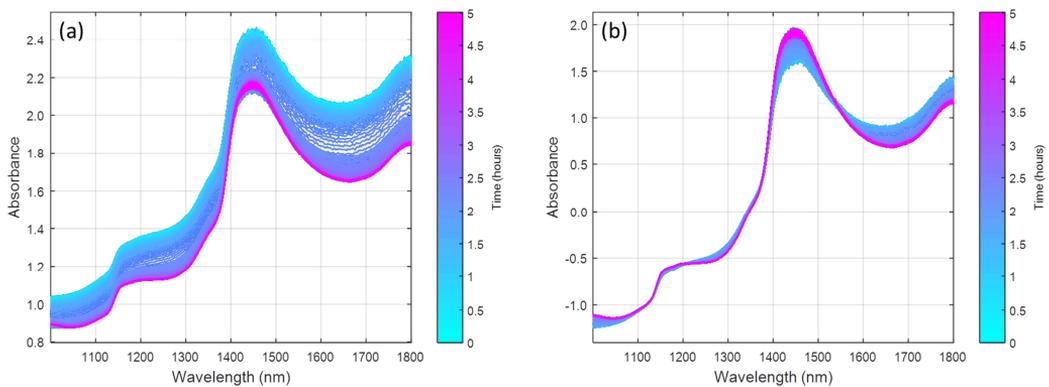


Figure 2: (a) The raw NIR spectra and (b) SNV pre-processed spectra from Batch 6 colored according to time.

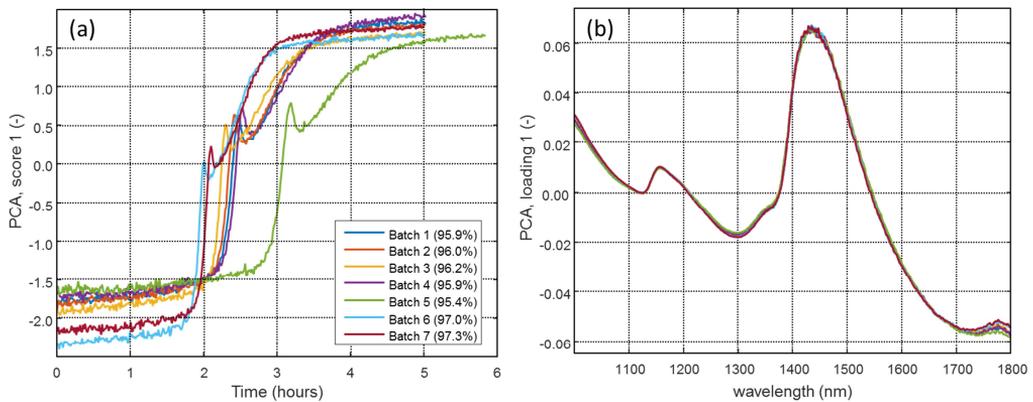


Figure 3: (a) PCA score values for the first principal component (variance explained for each batch in brackets) as a function of time for all seven batches; (b) PCA loading profiles.

In Figure 2a, the raw NIRS spectra for Batch 6 colored by time are shown; these results are representative for all seven batches. In general, the absorbance of the spectra seems to decrease over batch time, indicating that an aggregation of the micelles is happening as a function of the decreasing pH. A contraction of the aggregated casein will occur, which leads to an increase of the casein particle size²⁰, which might be corresponding to the decreasing offset. The strongest absorbance is seen in the region between 1350-1550 nm, which may belong to the third overtone of carbonyl groups ($>C=O$; 1333-1936 nm) and/or the first overtone of O-H stretching from sugars (1450 nm) and/or water (1440-1470 nm)¹³. A peak is also observed in the region between 1150-1250 nm, which can be related with the second overtone of C-H stretching¹³. It seems that the raw spectra are primarily associated with water and the more physical interaction between NIR radiation and the milk matrix in the form of scatter, and only minor information about chemical compounds found in yoghurt fermentations seems present. To enhance the more chemical information and identify the variation in the spectra multivariate data analysis was applied. The spectra were pre-processed by standard normal variate (SNV) scaling to remove artifacts and sample differences caused by baseline changes in the reflectance measurements (Figure 2b).

The pre-processed spectra were then modelled by Principal Component Analysis (PCA), individually per batch run. The first principal component scores, describing on average 96% of the variance, versus fermentation time are pictured for all seven batches in Figure 3a. An increasing trend over time was observed for all batches following the familiar S-curve. However, the batches did differentiate from each other as a clear time shift was seen; the lower the fermentation temperature applied, the more delayed was the profile. Thus, Batch 5 was the slowest and Batch 6-7 the fastest. Another highly reproducible feature was a small bump on the S-shaped curve observed around 2 hours for Batch 6-7, around 2.5 hours for Batch 1-4 and around 3 hours for Batch 5. The pattern can be further interpreted by inspecting the loading plots (Figure 3b), which indicated that the observed spectral variation was mainly caused by absorbance changes in the regions around 1350-1400 nm, 1500-1550 nm and to a smaller extend around 1700-1800 nm. This highly reproducible loading profile, among the different batch runs and PCA models, could be associated with sugar and the third overtone of carboxyl groups, respectively.

As stated, the change described by the PCA modelled NIR spectral measurements does to some extent have the similar characteristic S-shape as seen for the pH measurements. However, the NIR data seemed to combine information which was not available from the univariate method pH, which is normally applied in industrial yoghurt production, and the sharp bump observed on the first principal component for the modelled NIR data is not recognizable in the pH profiles. This might indicate that the first-order NIR spectroscopy method delivers process information, which cannot readily be detected by the zero-order methods²¹.

The dynamics pictured from NIR spectroscopy were very reproducible, which makes it a potential on-line measure for this experimental system and for yoghurt production on an industrial scale. Since NIR

spectroscopy is an indirect method it can be hard to conclude what we actually are measuring and what exactly was causing the dynamics. The bump on the S-curve is of special interest and it can be speculated what caused it. We know from the literature that a growth association between the two starter bacteria appears which might cause a metabolic change around that time where the bump on the first principal component is observed. Additionally, a textual change from liquid to gel is also happening around this time, resulting in a dramatic increase in viscosity. This means that the NIR spectral dynamics can be due to a chemical change or a physical change; in other words the variation can be caused by a change in the absorption or a change in the scattering signal or even a mix of changes is absorption and scatter. The distinction is straightforward if it concerns a non-scattering absorbance spectrum, but when it concerns scattering samples in reflectance mode the principals of scattering and absorption affect each other and the interpretation becomes more complex²². In order to further elucidate what causes the observed dynamics we will look into the results of qPCR- and Brix measurements.

qPCR Measurements

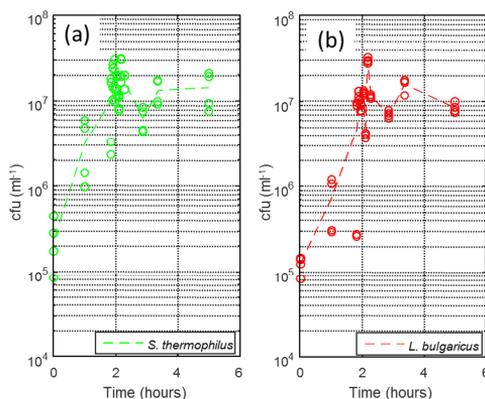


Figure 4: The cfu·ml⁻¹ versus time for Batch 6 for (a) *S. thermophilus* and (b) *L. bulgaricus*. The (o) represent a replicate from one of the two qPCR-analyses. Based on four replicates a mean is calculated, which are connected by the broken lines.

Part of the withdrawn samples from the yoghurt broth was used for qPCR measurements. The cell counts (expressed in colony forming units) are shown on a logarithmic scale as a function of time for *L. bulgaricus* and *S. thermophilus* in Figure 4. The results show that the exponential growth phase is taking place within the first two hours of the fermentation and hereafter the growth of both the bacteria seems to be stationary. As expected from the literature, *S. thermophilus* has a slightly faster growth than *L. bulgaricus*²³. It is notable that the growths reach the stationary phase at the same time, as where the bump on the first principal component for the NIR data is observed, although the uncertainty around the transition phase makes it hard to pinpoint this exactly. This might indicate that the dynamics before the bump on the S-shaped NIR spectroscopy curve could be caused by biomass formation.

Brix Measurement

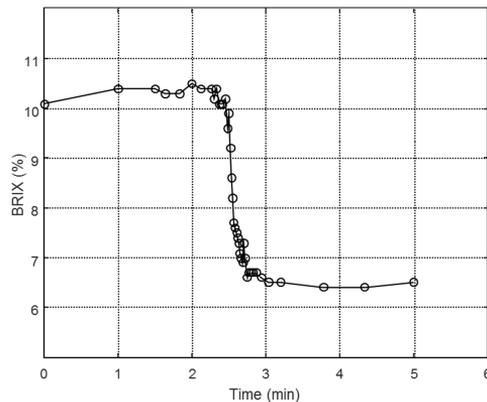


Figure 5: The Brix values for Batch 1 plotted over time.

The Brix results for Batch 1 are shown in Figure 5. This time profile is again representative for all seven batch runs. As seen from the figure, the at-line Brix values were recorded with varying sampling intervals; a higher sample extraction frequency was used on and right after the log phase (around 2h30m into the process for Batch 1, the batch most intensely sampled). The Brix readings decreased from around 10% to 6.5% in this time interval of approx. one hour. Normally Brix, when applied in the

food industry, is associated with e.g. amount of available sugars in wine or the amount of solids in tomato pulp. However, index of refraction, which is the principle behind Brix measurements, is also dependent on the viscosity of the solution among other parameters²⁴. Looking at Figure 5, it is also very unlikely that sugars are consumed at this pace in a biological reaction scheme, and the dynamics of the Brix signal are also incompatible with the dynamics observed for pH in Figure 1b. Therefore, we interpret the results in Figure 5 as an indirect measurement of the viscosity change in the milk-yoghurt curd system, and this agrees well with the observations made during manual sample extraction from the reactor and the known rapid viscosity increase in yoghurts²³.

Lactic Acid Measurements

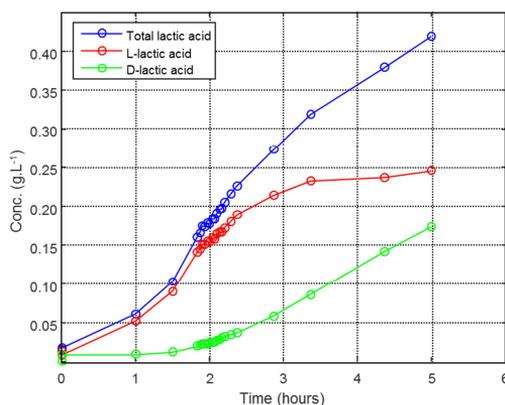


Figure 6: The concentration ($\text{g}\cdot\text{L}^{-1}$) of D-, L- Lactic acid and the sum of lactic acid for Batch 6, plotted as a function of time.

Grab samples from the fermentation broth were used for lactic acid quantification. The concentration of D-, L- and total lactic acid over time are pictured in Figure 6. The concentration of L-lactic acid increased from the very beginning and seemed to stabilize around 3h30m (in Batch 6), whereas the concentration of D-lactic acid started increasing after two hours of fermentation and for the batch run length 5h50m it did not stabilize. The concentration profiles of the D- and L-lactic acid correspond

with the fact that *S. thermophilus*, producing L-lactic acid, was initiating the fermentation, whereas *L. bulgaricus*, producing D-lactic acid, had a delayed metabolic profile compared to *S. thermophilus*. This confirms the assumption that a metabolic shift and an association between the two starter bacteria appeared during the fermentation. This existing growth association between the two starter bacteria is a well-known phenomenon². It should be noticed that the increasing lactic acid concentrations were not identical with bacterial growth seen from qPCR results (Figure 4), which could be expected. The reason for this is that the dynamics of the metabolic activity cannot be paralleled as such with the dynamics of bacterial growth. Initially, the biomass was formed and even though it reached a stationary phase, the metabolic activity was still on going.

Kinetic fitting

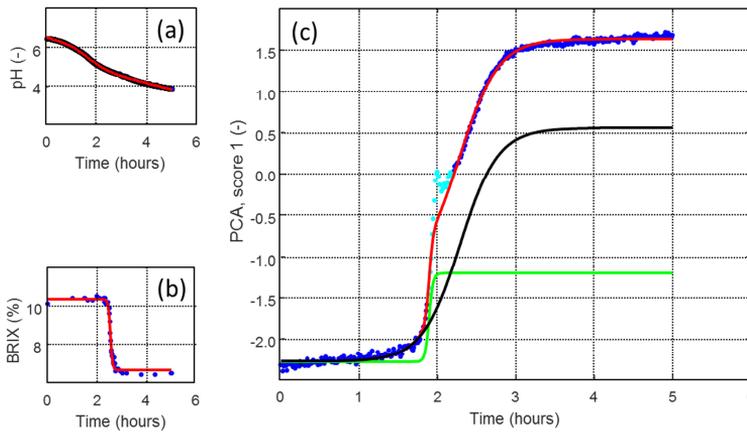


Figure 7: (a) pH for Batch 6 and (b) Brix measurements for Batch 1 with superimposed kinetic fittings; (c) PCA score values for the first principal component (dots) for Batch 6; red line: overall kinetic profile, Equation 1; black and green lines: individual contributions of the kinetic profile. The kinetic fits are made on the dark (blue) markers, whereas the light markers (cyan) were excluded.

In order to further elucidate the dynamics explained by using NIR spectroscopy and its capability of describing physical and/or chemical changes in lactic fermentations kinetic profiles were fitted on some of the measurements. For this purpose a generic mathematical model was fitted by non-linear regression²⁵:

$$Response(t) = \left(\sum_{i=1}^I \frac{b_{1(i)}}{1 + e^{k(i) \times (t - t_{infl.}(i))}} \right) + b_0 \quad (1)$$

where a response was modelled as a function of time t by a sum of I S-curves. Each curve was determined by the kinetic parameter $k(i)$ and the so-called inflection time $t_{infl.}(i)$, where the first determined the speed of change (increase or decrease) in the response signal, while the second term represented when this change happened on the time axis. The auxiliary b -parameters were there to adjust the equation to a particular response where b_0 was an overall offset and $b_{1(i)}$ was the magnitude or amplitude of each response.

Figures 7a and b show the fitting of Equation 1 - with $I = 1$ - on the pH and Brix signals for Batch 6 and 1, respectively. This resulted in the following parameters estimates for pH : $k(1) = 0.411 \text{ h}^{-1}$, $t_{infl.}(1) = 1\text{h}41\text{m}$, $R^2 = 0.9997$ and Brix : $k(1) = 0.044 \text{ h}^{-1}$, $t_{infl.}(1) = 2\text{h}33\text{m}$, $R^2 = 0.9932$. Interpreting these values showed us that the relatively simple model in Equation 1 gave a very small fitting error using only one S-curve, and that the two parameters k and $t_{infl.}$ were intuitive numerical indicators of the observed dynamics.

Looking at the PCA score profile in Figure 7c, two issues are noticeable: (1) the abrupt bump around 2 hours of batch time cannot be captured by a simple and heuristic mathematical equation, and (2) a single S-curve model does not fit well with the observed profiles (see Figure 3a). To solve issue (1) we manually excluded some of the points in the direct vicinity of the bump. Challenge (2) was easily solved by fitting the sum of two S-curves according to Equation 1, resulting in the following parameter list: $k(1) = -0.025 \text{ h}^{-1}$, $t_{infl.}(1) = 1\text{h}54\text{m}$, $k(2) = -0.246 \text{ h}^{-1}$, $t_{infl.}(2) = 2\text{h}59\text{m}$, $R^2 = 0.9998$. As can be seen in Figure 7c, this model described the observed dynamics of the fermentation, as ascertained from the

NIR spectroscopy measurement, extremely well. The overall model splits the response in a fast kinetics – which can be connected with the rapid viscosity change as described previously – and a slow kinetics - which we associate with metabolic conversion. The k-value for the fast kinetics was fairly sensitive towards the manual selection of data points excluded in parameter estimation. The other parameters (time points and inflection and slow kinetics) were however insensitive to this choice. This result gave us an opportunity to quantitatively compare different batches; the combination of in-line NIR spectra and kinetic modelling was a way to evaluate our seven yoghurt batches quantitatively (Table 1).

Table 1: Inflection time ($t_{infl.}$), kinetic parameter (k^{-1}) and R^2 are listed for pH and NIR for all seven batches.

		Batch 1	Batch 2	Batch 3	Batch 4	Batch 5	Batch 6	Batch 7
pH	$t_{infl.}$	1h48m	1h44m	1h41m	1h49m	2h23m	1h41m	1h46m
	k^{-1}	0.418	0.397	0.403	0.328	0.399	0.411	0.411
	R^2	0.9998	0.9998	0.9998	0.9998	0.9998	0.9994	0.9994
NIR	$t_{infl.} (1)$	2h23m	2h22m	2h14m	2h25m	2h59m	1h53m	2h02m
	$k^{-1}(1)$	-0.0186	-0.0233	-0.0221	-0.0250	-0.0346	-0.0190	-0.0167
	$t_{infl.} (2)$	2h44m	2h43m	2h36m	2h51m	3h29m	2h17m	2h23m
	$k^{-1}(2)$	-0.3306	-0.3338	-0.3273	-0.3508	-0.4442	-0.2457	-0.2500
	R^2	0.99954	0.99954	0.99956	0.99954	0.99948	0.99961	0.99962

CONCLUSIONS

Based on the results obtained from the NIR spectral data, it can be concluded that the NIR measurements revealed information that was not only reproducible, but also delivered a quantifiable dynamic trend from which consistency and quality control can be derived. Furthermore, NIR spectra provided additional information about the lactic fermentation compared to pH, which is traditionally applied as control measure in the industrial production of yoghurt. The results based on our yoghurt production model show that both variations in scatter and in the absorption are present in the NIR spectra. From the kinetic modelling it seems that we are able to model both physical changes in terms of textural changes due to the gel formation, and chemical information, which might be related to the biological conversion reactions.

ACKNOGLEDMENT

Thanks to Associate Professor Dennis Sandris Nielsen, PhD-fellow Josue Leonardo Castro Mejia and Professor Finn Vogensen from the Food Microbiology section (Department of Food Science, University of Copenhagen, Denmark) for discussions regarding microbiology and yoghurt fermentation. Thanks to Lisbeth Tina Dahl from the Spectroscopy and Chemometrics section (Department of Food Science, University of Copenhagen, Denmark) for assistants with the D- and L-lactic acid measurements.

REFERENCES

1. A. Y. Tamime, R. K. Robinson, "Chapter 7 - Biochemistry of fermentation", in Tamime and Robinson's Yoghurt - Science and Technology, Ed by A. Y. Tamime, R. K. Robinson, p. 535 (2007).

2. K. R. Nauth, "Yogurt", in Handbook of Food and Beverage Fermentation Technology, Ed by Y. H. Hui, L. Meunier-Goddik, Å. S. Hansen, J. Josephsen, W. K. Nip, P. S. Stanfield and F. Toldrá. Marcel Dekker, Inc., New York, p. 125 (2004).
3. A. Y. Tamime, R. K. Robinson, "Chapter 6 - Microbiology of yoghurt and related starter cultures", in Tamime and Robinson's Yoghurt - Science and Technology, Ed by ; A. Y. Tamime, R. K. Robinson, p. 468 (2007).
4. A. Tamime, H. Deeth, "Yogurt - Technology and Biochemistry", J. Food Prot.,43 (1980).
5. E. S. Bautista, R. S. Dahiya, M. L. Speck, "Identification of Compounds Causing Symbiotic Growth of Streptococcus Thermophilus and Lactobacillus Bulgaricus in Milk", J. Dairy Res, 33 (1966).
6. H. A. Veringa, T. E. Galesloot, H. Davelaar, "Symbiosis in Yoghurt (II). Isolation and Identification of a Growth Factor for Lactobacillus Bulgaricus Produced by Streptococcus Thermophilus", Netherlands Milk and Dairy Journal-Nederlands-Nederlands Melk En Zuiveltijdschrift, 22 (1968).
7. F. M. Driessen, F. Kingma, J. Stadhouders, "Evidence that Lactobacillus-Bulgaricus in Yogurt is Stimulated by Carbon-Dioxide Produced by Streptococcus-Thermophilus", Neth. Milk Dairy J, 36 (1982).
8. A. Zourari, J. P. Accolas, M. J. Desmazeaud, "Metabolism and Biochemical Characteristics of Yogurt Bacteria - a Review". Lait, 72 (1992).
9. F. Yidiz, "Chapter 1 - Overview of Yogurt and Other Fermented Dairy Products", in Development and Manufacture of Yogurt and Other Functional Dairy Products, Ed by F. Yildiz. CRC Press, Taylor & Francis Group, Florida, p.1 (2010).

10. B. Özer, "Chapter 2 - Strategies for Yogurt Manufacturing", in *Development and Manufacture of Yogurt and Other Functional Dairy Products*, Ed by F. Yildiz, CRC Press, Taylor & Francis Group, Florida, p. 47 (2010).
11. C. Béal, S. Helinck, "Chapter 5 - Yogurt and Other Fermented Milks", in *Microorganisms and Fermentation of Traditional Foods*, Ed by R. C. Ray, D. Montet, CRC Press, New York, p. 141 (2014).
12. A. Y. Tamime, R. K. Robinson, "Chapter 3 - Processing plants and equipment", in *Tamime and Robinson's Yoghurt - Science and Technology*, Ed by A. Y. Tamime, R. K. Robinson, p. 162 (2007).
13. S. E. Holroyd, "The use of near infrared spectroscopy on milk and milk products". *J. Near Infrared Spectrosc.*, 21 (2013).
14. S. Grassi, C. Alamprese, V. Bono, C. Picozzi, R. Foschino, E. Casiraghi, "Monitoring of lactic acid fermentation process using Fourier transform near infrared spectroscopy". *Journal of Near Infrared Spectroscopy*. 21 (2013).
15. G. Vaccari, E. Dosi, A. L. Campi, A. Gonzalezvara, D. Matteuzzi, G. A. Mantovani, "Near-Infrared Spectroscopy Technique for the Control of Fermentation Processes - an Application to Lactic-Acid Fermentation". *Biotechnol. Bioeng.*, 43 (1994).
16. M. Navratil, C. Cimander, C. Mandenius, "On-line multisensor monitoring of yogurt and Filmjolk fermentations on production scale". *J. Agric. Food Chem.*, 52 (2004).
17. C. B. Lyndgaard, S. B. Engelsen, F. W. J. van den Berg, "Real-time modeling of milk coagulation using in-line near infrared spectroscopy". *J. Food Eng.* 108 (2012).

18. J. L. W. Rademaker, J. D. Hoolwerf, A. A. Wagendorp, M.C. Te Giffel, "Assessment of microbial population dynamics during yoghurt and hard cheese fermentation and ripening by DNA population fingerprinting", *Int. Dairy J.*, 16 (2006).
19. J. P. Furet, P. Quenee, P. Tailliez, "Molecular quantification of lactic acid bacteria in fermented milk products using real-time quantitative PCR", *Int. J. Food Microbiol.*, 97 (2004).
20. A. Y. Tamime, R. K. Robinson, " Chapter 2 - Background to manufacturing practice ", in *Tamine and Robinson's Yoghurt - Science and Technology*, Ed by A. Y. Tamime, R. K. Robinson, p. 13 (2007).
21. C. Svendsen, T. Skov, F. W. J. van den Berg, "Monitoring fermentation processes using in-process measurements of different orders", *Journal of Chemical Technology and Biotechnology*, 90 (2015).
22. K. D. Dahm, D. J. Dahm, "Separating the effects of scatter and absorption using the representative layer", *Journal of Near Infrared Spectroscopy*, 21 (2013).
23. P. Walstra, J. T. M. Wouters, T. J. Geurts, "Chapter 22 - Fermented Milks", in *Dairy Science and Tachnology*, Ed by P. Walstra, J. T. M. Wouters, T. J. Geurts, CRC Press, Boca Raton, p. 551 (2005).
24. J. E. Brown, B. G. Lipták,"Chapter 8.52 – Refractometers", in *Instrument Engineers' Handbook - Process measurements and Analysis*, Ed by B. G. Lipták, CRC Press, Florida, p. 1620 (2003).
25. D. L. Massart, B. G. M. Vandeginste, L. M. C. Buydens, S. De Jong, P. J. Lewi, J. Smeyers-Verbeke, "Chapter 11 - Non-linear Regression", In *Handbook of Chemometrics and Qualimetrics, 20A; Data Handling in Science and Technology, Part A*, Ed by D. L. Massart, B. G. M. Vandeginste, L. M. C. Buydens, S. De Jong, P. J. Lewi, J. Smeyers-Verbeke, Elsevier Science B.V., The Netherlands, p. 305 (1997).

PAPER III

Weighted PARAFAC and non-linear regression for handling intensity changes in fluorescence spectroscopy caused by pH fluctuation

Carina Svendsen, Thomas Skov and Frans W. van den Berg

Applied Spectroscopy (2016), article published online

Weighted PARAFAC and Non-linear Regression for Handling Intensity Changes in Fluorescence Spectroscopy caused by pH Fluctuations

Applied Spectroscopy
0(0) 1–12
© The Author(s) 2016
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0003702816644610
asp.sagepub.com


Carina Svendsen, Thomas Skov, and Frans W.J. van den Berg

Abstract

Fluorescence spectroscopy is a sensitive and selective technique, which can be of great value in bioprocesses to provide online, real-time measures of chemical compounds. Although fluorescence spectroscopy is a widely studied method, not much attention has been given to issues concerning intensity variations in the fluorescence landscapes due to pH fluctuations. This study elucidates how pH fluctuations cause intensity changes in fluorescence measurements and thereby decreases the quality of the subsequent quantification. A photo-degradation process of riboflavin was investigated by fluorescence spectroscopy and used as a model system. A two-step modeling approach, combining weighted PARALLEL FACtor analysis (PARAFAC) with weighted non-linear regression of the known reaction kinetics, is suggested as a way of handling the fluorescence intensity shifts caused by the pH changes. The suggested strategy makes it possible to compensate for uncertainties in the shifted data and thereby obtain more reliable concentration profiles for the chemical compounds and kinetic parameters of the reaction.

Keywords

Chemometrics, correction strategy for intensity variations, fluorescence spectroscopy, intensity variations, online monitoring, pH fluctuations

Date received: 5 January 2016; accepted: 1 March 2016

Introduction

Fluorescence spectroscopy is an attractive option for the food and biotech industry due to its high sensitivity to organic compounds that contain conjugated bonds. It also has a high selectivity supported by both chemical selectivity and advanced chemometric modeling methods, and is furthermore attractive for its capability to monitor process parameters online and in real time. These aspects are of great importance in the bio-industries for process optimization and ensuring high product quality, productivity and yield. Fluorescence spectroscopy can detect concentrations down to one thousandth of what can be quantified by competitive spectroscopy methods like near infrared vibrational spectroscopy, which makes it a powerful quantitative method. In its most basal definition: fluorescence spectroscopy measure fluorophores in which a molecule first absorbs light (energy) causing an electron to excite from its ground state to an excited singlet state. In the excited state changes and interactions with the molecular environment, such as vibrational relaxation, quenching and energy

transfer, occur. When the electron returns from the excited state to the ground state the molecule emits light, which is then detected by the spectrometer. The intensity of the returned light is determined by the quantum yield, which is the number of emitted photons relative to the number of the absorbed photons. This is dependent on the molecular bond(s) involved in the process. The wavelength of the fluorescence molecular excitation is defined by the energy difference between the ground state and excited singlet state, whereas the energy difference between the excited singlet state and the ground state defines the emission wavelengths, in which the emission by definition has a lower energy than the excitation radiation. Furthermore, each electronic state has several associated vibrational levels, which means that excitation occurs

Department of Food Science, University of Copenhagen, Denmark

Corresponding author:

Carina Svendsen, Rolighedsvej 26, 1958 Frederiksberg C, Denmark.
Email: carinasv@food.ku.dk

over a number of wavelengths, which corresponds to several vibrational transitions. On the other side, the deactivation of the excited state only occurs at the vibrational ground level, whereas the emission, just like the excitation, occurs at several wavelengths as it might reach a number of vibrational levels in the ground state.^{1,2} Molecules which have fluorescence properties from a food and bioprocessing perspective include several biogenic fluorophores such as tryptophan, vitamins (pyridoxine, riboflavin) and co-enzymes (nicotinamide adenine dinucleotide phosphate, NADP+), among others.³

Some of the first fluorescence online sensors were based on one excitation and one emission wavelength allowing measurements of only one fluorophore at a time. In bioprocesses nicotinamide adenine dinucleotide phosphate (reduced form) was commonly measured as the only fluorophore.⁴ These measurements give very limited information for a process typically containing several of the biological constituents mentioned. Moreover, with the broad-feature emission-excitation profiles there could be a severe interference from other chemical constituents making quantification uncertain or even impossible. The more recent development of multi-wavelength fluorescence spectroscopy in combination with sophisticated data analytical methods has allowed simultaneous monitoring of several fluorophores. The technique has among others been applied for monitoring cultivations of bacteria,^{5–7} filamentous fungi,^{8,9} yeast^{10,11} and mammalian cells.^{12,13} In such cultivation processes pH is normally controlled and kept constant at a set-point value. In contrast, in food fermentations such as yoghurt^{14,15} and sourdough production,¹⁶ pH is changing during the fermentation time as a function of the biological process. And even under pH control small pH deviations can be experienced, e.g., during start-up of a batch process or transitions between different phases of a fed-batch operation. A previous study showed that pH changes can affect the fluorescence intensity,¹⁷ meaning that upwards and downwards shifts in the fluorescence intensity can appear. Thus, a dramatic uncertainty in the quantification of the chemical compounds will appear due to the principle behind Lambert–Beers law of an intensity–concentration relationship (even though Lambert–Beers law does not strictly apply in fluorescence spectroscopy). This violation is caused by an alteration of the electron cloud surrounding the conjugated bonds when the pH changes.¹⁷ It thereby seems that fluorescence spectroscopy could be an unreliable technique for quantification of chemical compounds or simply for measuring the dynamics in a process when pH is not under close loop control.

In this paper we first address the effect of pH fluctuations on fluorescence spectroscopy. The degradation process of riboflavin was used as a model system to elucidate how pH disturbances cause intensity shifts in the fluorescence signal. In addition, the study suggests a possible chemometric solution combining weighted PARAllel FACTor

analysis (PARAFAC) and weighted non-linear regression towards an assumed reaction kinetics to compensate for the intensity shifts and thereby providing a more reliable concentration profile and reaction rate constants when monitoring the system.

Experimental

Experimental Setup

The breakdown process of riboflavin and the formation of two degradation products ($A \rightarrow B + C$; a reaction driven by light exposure conducted in an alkaline environment) were investigated by fluorescence spectroscopy. The process will be used as a model batch system to elucidate the effect of intensity shifts in fluorescence landscapes as a function of pH disturbances in a dynamically changing environment. The breakdown process of riboflavin was conducted in a 2 L glass fermenter vessel (Culture Vessel M2, B. Braun, Melsungen, Germany). The vessel was fully enclosed by a black cover to limit any uncontrolled light exposure from the surroundings. The lid of the vessel contained instrument ports, four baffles reaching into the reactor contents close to the wall and a stirrer in the center of the vessel with three Rushton impellers. The mixing speed (100 r/min) was kept constant, controlled by a motor. A water jacket surrounding the fermenter maintained the temperature at 25°C and was controlled by a Pt100 probe inserted directly in the reactor using a HAAKE Phoenix pumping water bath (Thermo Scientific, Karlsruhe, Germany). A multivitamin effervescent tablet (Multivitamin, Optisana, Kolding, Denmark), containing several different vitamins including a fixed amount of riboflavin, was added and quickly dissolved in the glass vessel containing 2 L of water. Hereafter, small quantities of 5 M sodium hydroxide (NaOH) were added to reach a pH around 9. To stabilize the system it was left undisturbed for approximately one hour. The breakdown process of riboflavin was then started (time = 0 minutes) by turning on the light source (Intralux 6000, Volpi, Schlieren, Switzerland). The light was delivered just above the reactor contents by an optical guide (normally used in light microscopy) entering one of the instrument ports of the reactor lid. The kinetics of the photochemical breakdown of riboflavin could be manipulated by the intensity of light exposure and is the main question of interest in our model system. The process was stopped after 150 minutes; each process run will be addressed as a batch.

Normal and Abnormal Operating Conditions Batches

To investigate the effect of pH disturbances on the fluorescence intensity profiles both normal operating conditions (NOC) batches with constant pH around 9, and abnormal operating conditions (AOC) batches with pH interference were carried out. The kinetics or reaction speed was

different for each batch driven by manipulation of the light intensity, hence the efficiency of riboflavin breakdown varied among batches. The pH change in the AOC batches was introduced during the degradation process by either increasing or decreasing pH for a limited time interval by adding small volumes of hydrochloric acid (HCl) (5 M) and/or NaOH (5 M). This was done to mimic operational or equipment failure leading to deviations from the set-point and sluggish or delayed pH control around operational phase transitions. In total seven NOC and five AOC batches were carried out. To show the reproducibility of the riboflavin degradation experiment, the NOC batches were first investigated and compared.

Fluorescence Measurements

The fluorescence measurements were collected in-line with an interval of one minute, using a BioView spectrofluorometer (DELTA Light and Optics, Hørsholm, Denmark). The spectrometer was equipped with a fibre optic quartz-in-stainless steel probe, a so-called 180° geometry catching emitted light after excitation, entering one of the reactor ports. It was placed directly in the reactor solution, looking downwards and away from the light source used for the photochemical breakdown. Fluorescence landscapes (excitation-emission matrices; EEMs) were obtained with excitation wavelengths from 270 nm to 550 nm and emission wavelengths from 310 nm to 590 nm, with an interval of 20 nm, providing a total of 15 excitation and 15 emission wavelengths. This resulted in a data cube $\underline{\mathbf{X}}$ of the size $151 \times 15 \times 15$ for each batch (time \times excitation wavelength \times emission wavelength). No calibration towards concentration values is performed. Instead, score values of a PARAFAC model are treated as relative concentrations, and the reaction kinetics are estimated directly, per individual batch run, from these pseudo-concentrations. This procedure eliminates the need for calibration and calibration maintenance and is insensitive to long-term/batch-to-batch instrumental drift.

pH Measurements

pH Measurements were collected in-line with an interval of five seconds, using a pH logger (MadgeTech Inc., Warner, NH, USA) with the probe placed in the liquid. The pH probe was calibrated with pH 4.01 and pH 7.00 buffers before use and checked before and after each batch with the same buffers to correct for any potential pH offset obtained during the runs.

Data Analysis

All data analyses were performed using Matlab (version R2015a, The Mathworks, Inc., Natick, MA USA). Weighted non-linear regression and weighted PARAFAC

were performed using in-house routines and the N-way toolbox (version 3.20).¹⁸

The PARAFAC analysis decomposes the three-way fluorescence data cube $\underline{\mathbf{X}}$ into a number of trilinear factors and a residual array

$$x(t, i, j) = \sum_{f=1}^F I_f(t) \times EX_f(i) \times EM_f(j) + e(t, i, j) \quad (1)$$

$(t = 0 \dots 150; i = 1 \dots 15; j = 1 \dots 15)$

where $x(t, i, j)$ represents the fluorescence intensity for time point t at excitation wavelength i and emission wavelength j . The EEM data are then decomposed into concentration scores $I_f(t)$, excitation wavelength loadings $EX_f(i)$ and emission wavelength loadings $EM_f(j)$, for each factor or PARAFAC component f . The residual array $e(t, i, j)$ represents the variation not described by the model.¹⁹ A non-negativity constrained was applied in all three modes.¹⁹

Eq. 2(a–c) was used for non-linear estimation/fitting²⁰ of the reaction scheme $A \rightarrow B + C$:

$$I_A(t) = b_{1,A} \cdot e^{-Kt} + b_{0,A} \quad (2a)$$

$$I_B(t) = b_{1,B} \cdot (1 - e^{-Kt}) + b_{0,B} \quad (2b)$$

$$I_C(t) = b_{1,C} \cdot (1 - e^{-Kt}) + b_{0,C} \quad (2c)$$

where I parameters stand for the score intensities (as a function of time) found via the PARAFAC model, the b parameters are included to compensate for magnitude and offset differences in the scores intensities observed for the different chemical species, and K is the reaction rate (common for all species). In total seven model parameters are estimated simultaneously.

To evaluate the relationship between the optimized kinetic profile and the concentration PARAFAC score the concordance correlation coefficient ρ_c was calculated²¹

$$\rho_c(\mathbf{y}_1, \mathbf{y}_2) = \frac{2s_{12}}{s_1^2 + s_2^2 + (\bar{y}_1 - \bar{y}_2)^2} \times 100\% \quad (3)$$

where vectors \mathbf{y}_1 and \mathbf{y}_2 are the series to be compared (e.g., a reference and predicted kinetic profile) based on the means, variances and co-variances of the two series. Unlike the conventional Pearson correlation coefficient the concordance correlation takes into account deviations from the ideal prediction line (angle of 45°, crossing the zero intersect). This is more appropriate when evaluating functional fitting where, e.g. a mismatch of one kinetic parameter can lead to a biased or systematic under or over-fitting, while still achieving a high ordinary Pearson correlation.²¹

Results and Discussion

Normal Operating Conditions Batches

The pH profiles of two of the NOC batches, randomly selected from seven NOC experiments, are shown in

Figure 1a and Figure 2a. Normal operation conditions were applied, hence the pH was kept constant at pH 8.8 (± 0.1) for NOC batch 1 and at pH 9.0 (± 0.1) for NOC batch 2, except for a small gradual decrease, which can be attributed to the formation of carbonic acid from an

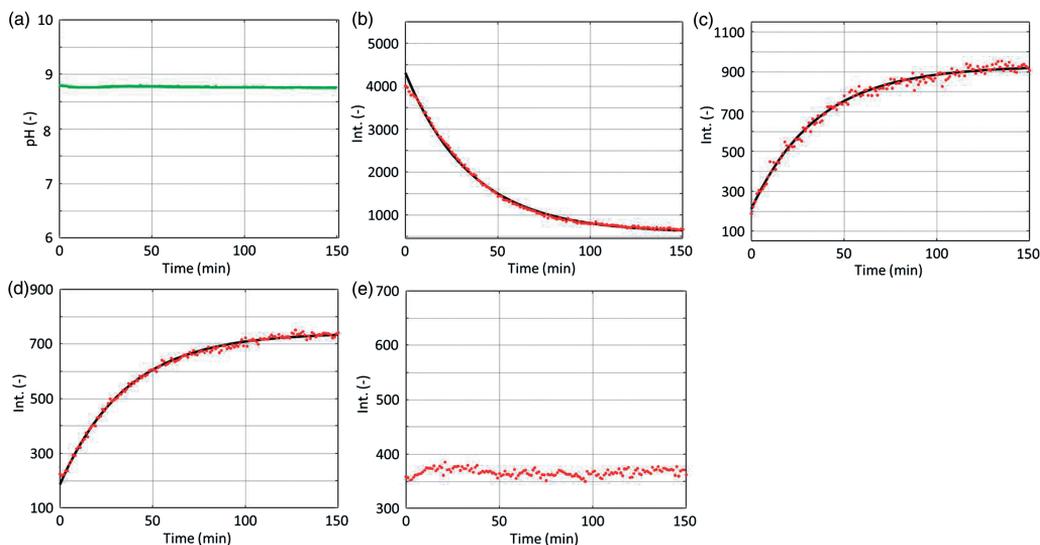


Figure 1. Normal operating conditions batch 1: (a) pH profile, without any disturbance; (b–d) the concentration scores from a PARAFAC analysis (red dots) and the fitted kinetic profiles (solid lines) for compounds 1, 2 and 3, respectively; (e) the concentration scores from a PARAFAC analysis for compound 4. No kinetic profile is given for the non-reactive compound 4.

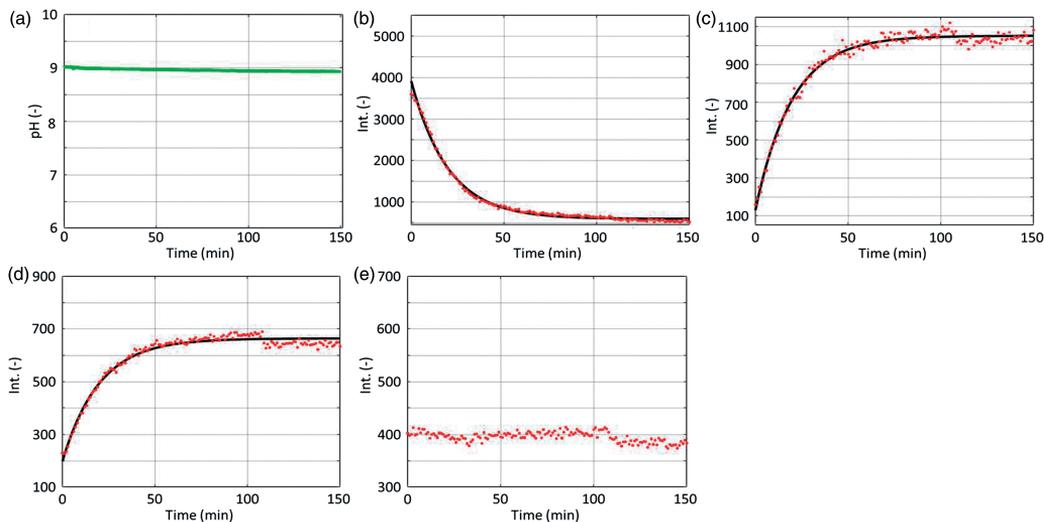


Figure 2. Normal operating condition batch 2: (a) pH profile; (b–d) the concentration scores from a PARAFAC analysis (red dots) and the fitted kinetic profiles (solid lines) for compounds 1, 2 and 3, respectively; (e) the concentration scores from a PARAFAC analysis for compound 4. No kinetic profile is given for the non-reactive compound 4.

increasing concentration of dissolved carbon dioxide due to stirring. The three-way data cubes for the batches were evaluated individually by a PARAFAC model where four factors were found to be the right model complexity, meaning that four different chemical compounds with a fluorescence signal were detected. The intensity scores obtained from the PARAFAC models (Eq. 1) are plotted as a function of batch time in Figure 1b–e and in Figure 2b–e for NOC batch 1 and NOC batch 2, respectively. For both batches, the first score (compound 1) describes a clear decreasing trend over time, the second and the third scores (compounds 2 and 3) describe increases, which are highly correlated (as expected because the formation of the products happens at the same rate), whereas the fourth score (compound 4) seems to represent a non-reacting species. Based on the dynamics of the first three scores, which represent the concentration profiles of the three compounds taking part in the reaction, the kinetic profiles were determined by a non-linear regression model using Eq. 2. The estimated kinetics profiles are included in Figure 1b–d and in Figure 2b–d. As seen, the concentration scores and the estimated kinetic profiles are very similar, indicating that the calculated K values $2.72 \cdot 10^{-2}$ /minute (batch NOC 1; Table 1) and $4.90 \cdot 10^{-2}$ /minute (batch NOC 2; Table 1), for the first order reactions correspond well to the measured kinetics. This also confirms that the detected chemical compounds behave in accordance with a first order system $A \rightarrow B + C$, where the trend is either asymptotically increasing or decreasing over time. It is also worth remembering that a unique determination for the two reaction products $B + C$ can still be found because of the second-order nature of excitation-emission-fluorescence landscapes and PARAFAC data analysis.²²

Further inspection of the scores for NOC batch 2 reveals that a small shift is seen around $t = 110$ minutes. As the shift appears after the reaction has progressed almost to completion, it will not affect the calculated K

value significantly. As the pH was kept constant, the shift cannot be attributed to pH changes as explored in this study. Instead this shift is caused by an automatic intensity correction that the instrument performs by aligning the filter wheel positions after a specific (large) number of measurements (personal communication, DELTA Light and Optics). As a result, all four chemical compounds show a small artificial drop in intensity due to the instrumental adjustment.

The loadings obtained from the PARAFAC model for NOC batch 1 are visualized in Figure 3 by reconstructing the landscapes using the excitation and emission loading vector from Eq. 1 to compute the outer matrix products. The fluorescence compound described by factor one found at excitation/emission wavelength around 460/540 nm (Figure 3a) can be identified as riboflavin,^{23,24} which is described, by the score values, to have a decreasing concentration over time (Figure 1b). This trend corresponds with the fact that riboflavin is being degraded over time, caused by the alkaline environment and the light exposure, as described in the experimental section. Loading two and three found around excitation/emission wavelength 450/540 nm and 400/460 nm can be tentatively identified as expected reaction products lumiflavin and lumichrome;²⁴ (Figure 2c and d). Both compounds show an increasing trend over time in the score plot (Figure 1c and d), corresponding with the fact that these two chemical compounds are products from the breakdown of riboflavin. A fourth compound, representing an unidentified chemical species, which neither shows an increase nor a decrease over time in the score plot (Figure 1e), is found around excitation/emission wavelength 340/400 nm in the loadings (Figure 3d). Notice again that only second-order data like EEMs in combination with a method like PARAFAC can describe a constant component like this.²²

The loadings as seen in Figure 3 are also representative of NOC batch 2 and all other batches. Similarly for all NOC

Table 1. The pH, estimated K values and the concordance correlation coefficient between PARAFAC scores and the kinetic fitted data – calculated per component – for NOC batches.

Batch	pH (\pm)	K value (10^{-2} min^{-1}) for the unweighted/weighted		Concordance correlation coefficient (%) (compound: 1/2/3) unweighted			Concordance correlation coefficient (%) (compound: 1/2/3) weighted		
NOC 1	8.8 (± 0.1)	2.72	2.81	99.9	99.3	99.8	99.9	99.5	99.8
NOC 2	9.0 (± 0.1)	4.90	5.03	99.8	99.0	98.7	99.8	99.1	98.7
NOC 3	8.4 (± 0.3)	3.16	3.24	99.9	98.9	99.8	99.9	99.0	99.8
NOC 4	8.7 (± 0.4)	5.47	5.64	99.8	98.9	99.3	99.8	98.9	99.4
NOC 5	9.1 (± 0.2)	4.25	4.37	99.9	98.4	98.1	99.8	98.6	98.2
NOC 6	8.8 (± 0.1)	3.45	3.54	99.9	99.2	99.6	99.9	99.2	99.6
NOC 7	9.0 (± 0.1)	4.89	5.04	99.8	98.8	99.3	99.9	99.0	99.4

NOC: normal operating conditions.

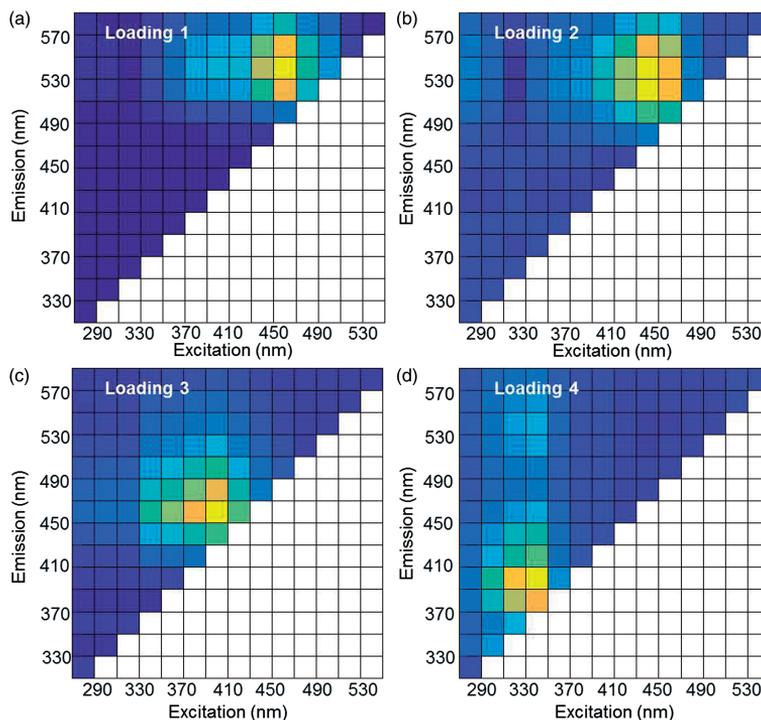


Figure 3. The PARAFAC loadings: (a) factor 1, the reactant riboflavin; (b) factor 2, product lumiflavin; (c) factor 3, product lumi-chrome; and (d) factor 4, a constant, unknown compound.

batches, the fit between PARAFAC scores and estimated kinetic models is good, as seen from the concordance correlation coefficient in Table I. The only thing that varies between the NOC runs of our model system is the kinetic parameter.

Abnormal Operating Conditions Batches

In the AOC batches abnormal conditions in terms of process disturbances – deviations from the set-point – were introduced by manipulating the pH during the riboflavin degradation process. As can be understood from the NOC experiments shown in Table I, pH has no strong or systematic influence on the reaction kinetics, at least within the alkaline pH range where our photochemical breakdown process of riboflavin takes place.

To elucidate how pH fluctuations affect the fluorescence landscapes, raw data inspection was carried out before modeling. Intensity profiles at various excitation/emission wavelength combinations were investigated as a function of time. As seen in Figure 4, which pictures an EEM landscape ($t=74$ minutes, AOC batch A) and two selected excitation/emission combinations (390/450 nm and 450/530 nm), both increasing and decreasing trends over time

are easily identified. Superimposed on this, clear intensity shifts, in opposite directions, appear in both profiles when the rather severe pH interference was introduced for AOC batch A.

In AOC batch A, the pH was first decreased from pH 9 to pH 7 ($t=69$ minutes) and thereafter the pH was raised again to approximately pH 9 ($t=85$ minutes), as depicted in Figure 5a. The rank or optimal model complexity of the PARAFAC model based on the data from AOC batch A remained four, and the loading landscapes were completely similar to those presented in Figure 3. For none of the AOC batches could a spectral shift on the emission or excitation axis be observed by our process instrument, which has a limited spectral resolution. This highlights the difficulty when using a naive approach in process monitoring in which an intensity shift due to pH changes can only be interpreted (incorrectly) as a rise or fall in concentrations. The scores of factors one to four obtained from the PARAFAC model are also illustrated in Figure 5b–e. A very clear intensity shift is observed, which agrees with the time point where acid was added, superimposed on the reaction trends as found in the NOC batches. Moreover, it can be seen that the intensity shifts are different for different chemical compounds. Similar to what was

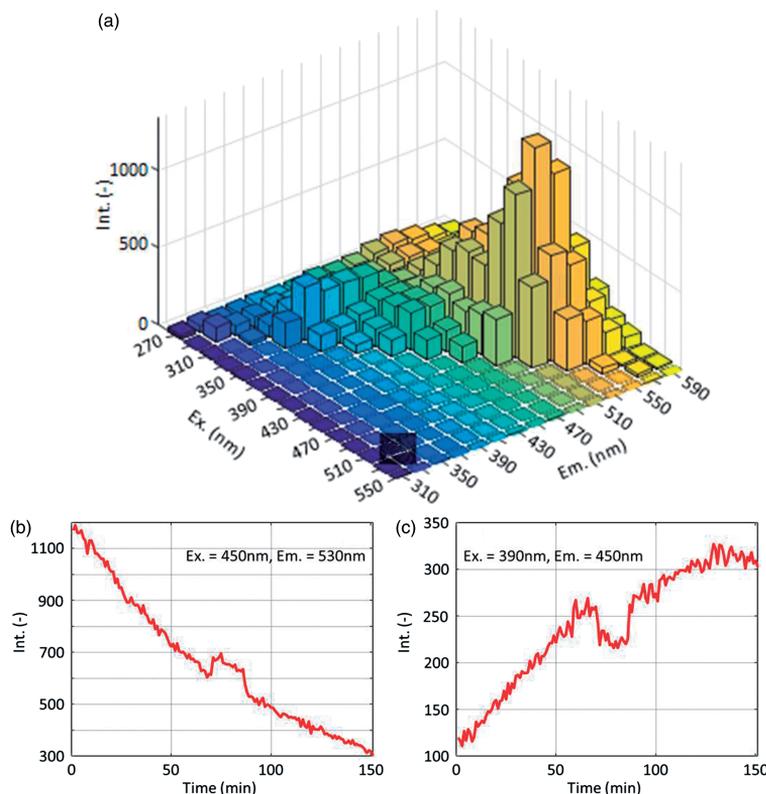


Figure 4. (a) Fluorescence landscape ($t = 74$ minutes) from abnormal operating conditions batch A; (b) a time-intensity profile of excitation/emission wavelength 450/530 nm; and (c) excitation/emission wavelength 390/450 nm.

observed for the NOC runs, the scores of the fourth factor neither show an increasing or a decreasing trend over time. However, this compound is also influenced by the pH change, as a clear intensity shift is observed in the scores.

Calculating the non-linear regression kinetic model parameters from the concentration profiles obtained from the PARAFAC model using data from AOC batch A provided a K value of $0.97 \cdot 10^{-2}$ /minute. As the intensity shift observed in the concentration profile is not presented by the kinetic model (Eq. 2), it might be concluded that the kinetic profile represents the true chemical concentration better than the PARAFAC scores in those time points where the landscapes were intensity shifted (Figure 5b–d), and a form of interpolation over this period would be more realistic to represent the photochemical breakdown process dynamics. Moreover, the estimation of the kinetic parameter is also affected by the intensity shift due to the total sum of squared errors minimization used in the non-linear regression.²⁰ Hence, the PARAFAC scores were more representative for quantification of the riboflavin breakdown in this particular batch run in the areas where no shifts were observed, whereas in the areas where the shift was

observed, the measurements should be down-weighted during model parameter estimation. Therefore, a combined use of the kinetic profile and the PARAFAC scores would be a more effective solution. Based on the difference between the obtained PARAFAC profiles and the expected kinetic profiles a weighted iterative modeling approach is suggested.

The proposed modeling approach, which combines weighted PARAFAC¹⁹ and weighted non-linear regression²⁰ is presented in the flowchart in Figure 6, and detailed as follows: *Data cube* – the fluorescence landscapes collected over time for one batch should be stacked into a three-way data structure. *Initial weight* – an initial weight-vector with value 1 for each measurement/time point is given as input for the algorithm loop. To obtain a kinetic profile that is corrected for any shifts or disturbances a new weight vector is determined inside the loop. *Initial estimates of the kinetics* – the initial estimate of the reaction constant K is set to the (unrealistic) value infinite, the other model b parameters in Eq. 1 are given approximately realistic values based on previous estimates. *Weighted PARAFAC* – a PARAFAC model (similar to Eq. 1) is calculated using the

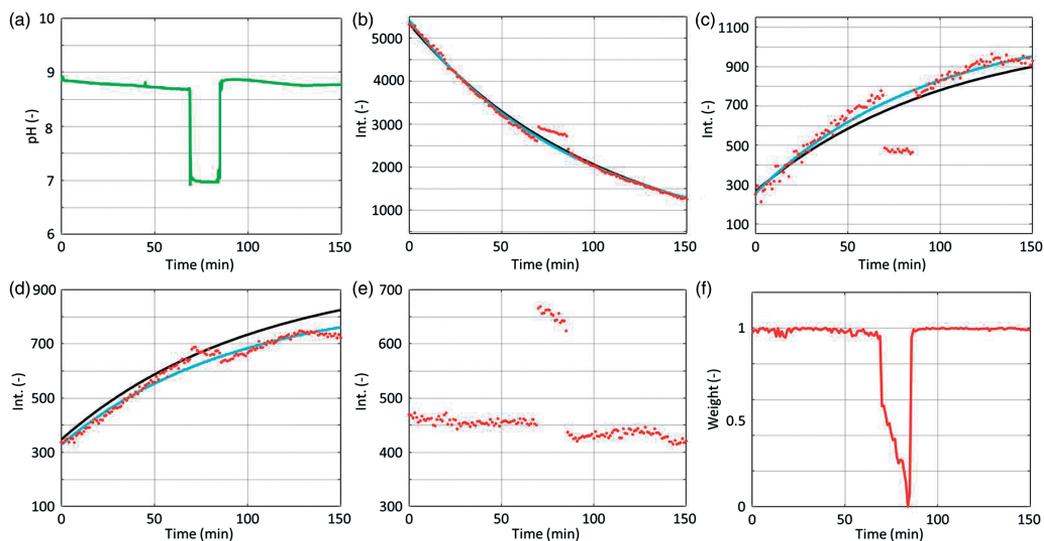


Figure 5. Abnormal operating conditions batch A: (a) pH profile, with pH disturbance at $t = 69\text{--}85$ minutes; (b–d) the concentration scores from a PARAFAC analysis (red dots), the unweighted (black lines) and weighted (cyan lines) kinetic profile for compounds 1, 2 and 3, respectively; (e) the concentration scores for compound 4; and (f) the weight vector.

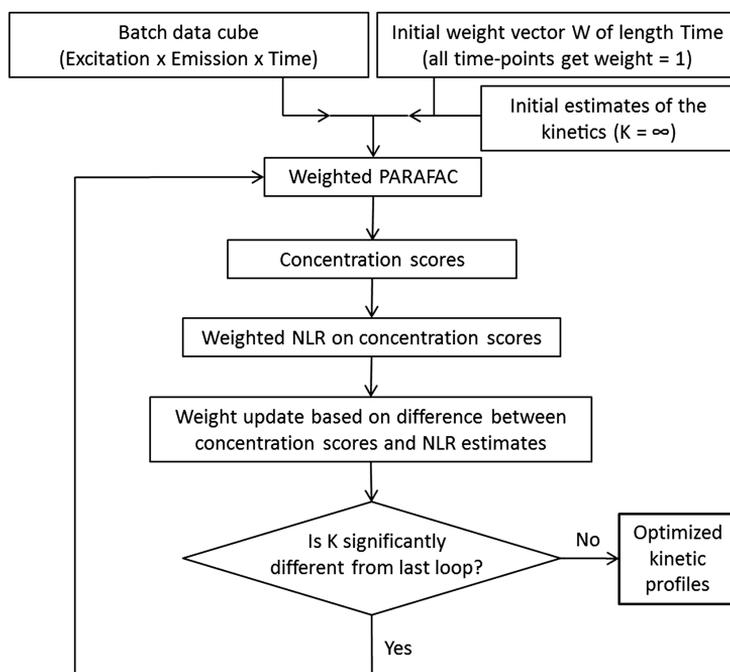


Figure 6. Flowchart of the modeling approach in down-weighting of intensity changes in the fluorescence signal caused by pH fluctuations.

Table 2. The pH, estimated K values and the concordance correlation coefficient between PARAFAC scores and the kinetic fitted data – calculated per component – for AOC batches.

Batch	pH (\pm) ^a	K value (10^{-2} min^{-1}) for the unweighted/ weighted		Concordance correlation coefficient (%) (compound: 1/2/3) unweighted			Concordance correlation coefficient (%) (compound: 1/2/3) weighted		
AOC A	8.8 (± 0.1)	0.97	1.08	99.8	95.4	99.1	99.8	96.1	98.9
AOC B	9.0 (± 0.1)	4.97	5.07	99.8	99.2	99.5	99.8	99.1	99.5
AOC C	8.8 (± 0.2)	1.02	1.06	99.1	98.5	96.1	99.2	98.6	96.3
AOC D	8.8/9.3 ^b (± 0.05)	4.90	5.21	99.2	99.0	99.1	99.2	98.8	99.1
AOC E	8.2 (± 0.2)	2.08	2.12	99.8	96.9	97.2	99.8	96.9	97.2

^apH disturbance not included in the uncertainty/range determination.

^bIn batch AOC D pH was changed from one level to another and kept at the second level throughout the rest of the run.

AOC: abnormal operating conditions.

present weight vector, where all excitation-emission-combinations for one time point/landscape get the same weight as determined for that time point. *Concentration scores* – the PARAFAC intensity scores, which provide the relative concentration profile for each of the detected chemical compounds, are obtained from the weighted PARAFAC model. *Weighted non-linear least-squares regression (NLR) on concentration scores* – the model parameters/kinetic coefficient of the concentration scores, are estimated by weighted NLR using Eq. 1. *Weight update based on difference between concentration scores and NLR estimates* – a new weight vector is calculated based on the difference between the concentration scores, obtained from the weighted PARAFAC model, and the kinetic profiles obtained from the weighted non-linear least-squares. Thereby, a weight between the first score and the kinetic profile of compound one is obtained for each time point in the batch run. Likewise, the difference between the second score and the second kinetic profile and the difference between the third score and the third kinetic profile are calculated. The three outcomes are squared and the average is determined. The average was rescaled to values between 0 and 1 and defined as the new weight vector. *Is the K significantly different from the last loop?* – if the K value is significantly different from the previous estimate a new loop/iteration will be started. If the K value is not significantly different from the previous estimate the algorithm has converged and the results are presented. By applying this approach an optimized and final kinetic profile (cyan lines, Figure 5b–d) delivers a K value of $1.08 \cdot 10^{-2}$ /minute for AOC batch A. When inspecting the weights in Figure 5f it is clear that the time points where the pH interference was introduced is down-weighted. This corresponds perfectly with our aim, as we are interested in knowing the kinetic constant and the true concentration profiles without any intensity contributions from the pH interference. In particular, the fitted profiles for the two reaction products, with their lower intensities, are improved notably, despite the relatively small difference

in unweighted and weighted K estimates for AOC batch A (Table 2).

A second AOC batch (B) with a less dramatic pH change (from pH 9 to 8.7 and back to pH 9) is shown in Figure 7a. The introduced pH disruption was much smaller but minor disturbances in the concentration scores obtained from the PARAFAC model are still observed around $t = 20$ minutes (Figure 7b–d).

The K value for AOC batch B, obtained by calculating the non-linear regression model parameters from the PARAFAC scores (concentration profiles), was $4.97 \cdot 10^{-2}$ /minute. By combining the weighted PARAFAC score and non-linear regression an optimized K value of $5.07 \cdot 10^{-2}$ /minute was obtained. As only a small pH change was introduced, the shift in the PARAFAC scores was less dramatic and the K value obtained before and after applying the weighing strategy is not too different from each other. This corresponds with the weight vector obtained, where only small differences were observed (Figure 7f) and where individual noisy data points and the measurements in the very first part of the batch are down-weighted almost to the same extent as the pH shift.

For completeness the weight strategy was also applied on NOC batches 1 and 2 and when inspecting the weight vector (Figure 8), only the first part of the batches are down-weighted, similar to the weight vector for AOC batch 2. As the weight is rescaled between 0 and 1, it should be remembered that the weight vector is a relative measure within each batch and cannot be compared among batches.

It shows that the suggested strategy of weighting can function as a more general strategy for down-weighting noisy observations, e.g., in the initial part of an exponential profile where small errors can have a considerable influence on parameter estimation for first-order reactions. To explore this further, Table 1 shows the weighted estimates for our NOC runs, where small adjustments of the K values are observed, accompanied by modestly improved

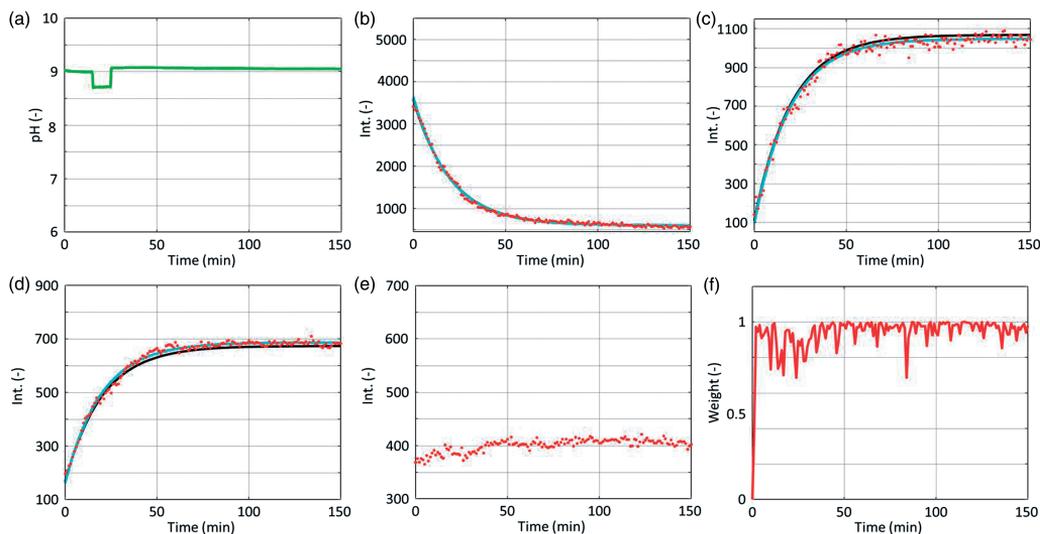


Figure 7. Abnormal operating conditions batch B: (a) pH profile, with pH disturbance at $t = 15\text{--}25$ minutes; (b–d) the concentration scores from a PARAFAC analysis (red dots), the unweighted (black lines) and weighted (cyan lines) kinetic profile for compounds 1, 2 and 3, respectively; (e) the concentration scores for compound 4; and (f) the weight vector.

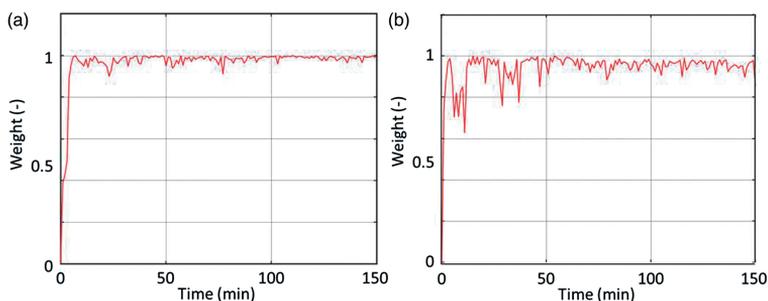


Figure 8. Weight vectors: (a) normal operating conditions (NOC) batch 1, relate to Figure 1; (b) NOC batch 2, relate to Figure 2.

weighted concordance correlation coefficients. It is difficult to compare equally weighted and weighted solutions quantitatively, but in many solutions an improvement similar to the observation made in Figure 7c and d is seen, where an underestimation of the unweighted K values is the trend (not shown). The K values for the unweighted and final weighted fit for the other AOC batches are listed in Table 2. The weighted concordance correlation between the final fit and the concentration scores are also reported. Comparing the pH and the K values again does not reveal any clear trend, confirming that the pH does not seem to affect the speed of the reaction as such.

The score profiles for AOC batch A in Figure 5, obtained from the PARAFAC model, have the potential of providing a relative estimate of the concentration profiles

and reaction kinetics in the data explored in this study. But the pH interference is also manifested as an apparent concentration change because the fluorescence intensity is directly influenced.

This means that the PARAFAC model is perfectly describing the true process changes appearing, but Lambert–Beers law, normally relied on in quantitative spectroscopy, is violated. However, we are not interested in describing the changes caused by the pH fluctuation, but for real-time monitoring of batch processes only in the changes appearing due to true reaction kinetics. It can be discussed why this fluorescence intensity shift is taking place. It might be that the chemical structure changes due to the environmental pH changes,¹ but we would then most probably also expect to observe a shift in the excitation/

emission wavelengths. The latter is hard to discern from the rather crude resolution in our process fluorescence spectrometers. This is, however, not an uncommon situation for filter-based process instruments and will most likely not change in the future. The weighted PARAFAC step in Figure 6 might seem redundant when we state that only the intensity changes (the rank, and thus spectral profiles, stay the same). However, due to the filter/sequential nature of the fluorescence instruments there might be spike disturbances (bubbles, particles), which make landscape down-weighting attractive in practice. Although our study mainly focuses on disturbances caused by pH fluctuations, disturbances can in practice be caused by several process or instrument-related reasons, which can lead to intensity shift in the fluorescence landscapes. Therefore, the suggested weighting strategy can be applied in several cases, where fluorescence outcomes do not follow Lambert–Beers law.

Conclusions

Based on the model system applied in this study, it is concluded that pH fluctuations can have a major impact on the intensity of the fluorescence measurement landscape and thereby affect the concentration profiles obtained from a PARAFAC analysis. When applying fluorescence as a monitoring tool in systems without a fixed pH set-point the user should thus be aware that Lambert–Beers law does not automatically hold and kinetic parameter estimation might be invalid. Our study suggests a chemometric modeling approach, where weighted non-linear regression and weighted PARAFAC are combined, as a solution to compensate for fluorescence intensity shifts. When determining kinetic rates in our model system on the unweighted PARAFAC scores, the intensity shifts can have a negative influence, while calculating the solutions using weights determined based on expected reaction kinetics more representative K values are obtained.

Conflicts of Interest

The authors report there are no conflicts of interest.

Funding

This study is partly funded by BIOPRO (www.biopro.nu), which is sponsored by Region Zealand, the European Regional Development Fund (ERDF) and the industrial partners.

References

- J.R. Lakowicz. "Introduction to Fluorescence". In: J.R. Lakowicz (ed.). *Principles of Fluorescence Spectroscopy*, 3rd ed. Chapter 1, pp. 1-26. New York: Springer, 2006.
- J. Christensen, L. Norgaard, R. Bro, S.B. Engelsen. "Multivariate autofluorescence of intact food systems". *Chem. Rev.* 2006. 106(6): 1979–1994.
- S. Marose, C. Lindemann, T. Scheper. "Two-dimensional fluorescence spectroscopy: A new tool for on-line bioprocess monitoring". *Biotechnol. Prog.* 1998. 14(1): 63–74.
- D.W. Zabriskie, A.E. Humphrey. "Estimation of Fermentation Biomass Concentration by Measuring Culture Fluorescence". *Appl. Environ. Microbiol.* 1978. 35(2): 337–343.
- D.M. Rossi, D. Solle, B. Hitzmann, M.A. Zachia Ayub. "Chemometric modeling and two-dimensional fluorescence analysis of bioprocess with a new strain of *Klebsiella pneumoniae* to convert residual glycerol into 1,3-propanediol". *J. Ind. Microbiol. Biotechnol.* 2012. 39(5): 701–708.
- G. Jain, G. Jayaraman, O. Koekbair, U. Rinas, B. Hitzmann. "On-line monitoring of recombinant bacterial cultures using multi-wavelength fluorescence spectroscopy". *Biochem. Eng. J.* 2011. 58–59: 133–139.
- N.P. Ronnest, S.M. Stocks, A.E. Lantz, K.V. Germaey. "Introducing process analytical technology (PAT) in filamentous cultivation process development: comparison of advanced online sensors for biomass measurement". *J. Ind. Microbiol. Biotechnol.* 2011. 38(10): 1679–1690.
- M. Ganzlin, S. Marose, X. Lu, B. Hitzmann, T. Scheper, U. Rinas. "In situ multi-wavelength fluorescence spectroscopy as effective tool to simultaneously monitor spore germination, metabolic activity and quantitative protein production in recombinant *Aspergillus niger* fed-batch cultures". *J. Biotechnol.* 2007. 132(4): 461–468.
- M.B. Haack, A.E. Lantz, P.P. Mortensen, L. Olsson. "Chemometric analysis of in-line multi-wavelength fluorescence measurements obtained during cultivations with a lipase producing *Aspergillus oryzae* strain". *Biotechnol. Bioeng.* 2007. 96(5): 904–913.
- P. Odman, C.L. Johansen, L. Olsson, K.V. Germaey, A.E. Lantz. "On-line estimation of biomass, glucose and ethanol in *Saccharomyces cerevisiae* cultivations using in-situ multi-wavelength fluorescence and software sensors". *J. Biotechnol.* 2009. 144(2): 102–112.
- S.S. Masiero, J.O. Trierweiler, M. Farenzena, M. Escobar, L.F. Trierweiler, C. Ranzan. "Evaluation of Wavelength Selection Methods for 2D Fluorescence Spectra Applied to Bioprocesses Characterization". *Brazil. J. Chem. Eng.* 2013. 30(2): 289–298.
- S. Bonk, M. Sandor, F. Rudinger, B. Tscheschke, A. Prediger, A. Babitzky, D. Solle, S. Beutel, T. Scheper. "In-situ microscopy and 2D fluorescence spectroscopy as online methods for monitoring CHO cells during cultivation". *BMC Proceedings.* 2011. 5(Suppl 8): 76.
- A.P. Teixeira, C.A.M. Portugal, N. Carinhas, J.M.L. Dias, J.P. Crespo, P.M. Alves, M.J.T. Carrondo, R. Oliveira. "In Situ 2D Fluorometry and Chemometric Monitoring of Mammalian Cell Cultures". *Biotechnol. Bioeng.* 2009. 102(4): 1098–1106.
- E. Becker, J. Christensen, C. Frederiksen, V. Haugaard. "Front-face fluorescence spectroscopy and chemometrics in analysis of yogurt: Rapid analysis of riboflavin". *J. Dairy Sci.* 2003. 86(8): 2508–2515.
- J. Christensen, E. Becker, C. Frederiksen. "Fluorescence spectroscopy and PARAFAC in the analysis of yogurt". *Chemometrics Intellig. Lab. Syst.* 2005. 75(2): 201–208.
- B. Grote, T. Zense, B. Hitzmann. "2D-fluorescence and multivariate data analysis for monitoring of sourdough fermentation process". *Food Control.* 2014. 8–18.
- R. Laane. "Influence of Ph on the Fluorescence of Dissolved Organic-Matter". *Mar. Chem.* 1982. 11(4): 395–401.
- C.A. Andersson, R. Bro. "The N-way Toolbox for MATLAB". *Chemometrics Intellig. Lab. Syst.* 2000. 52(1): 1–4.
- R. Bro. "PARAFAC. Tutorial and applications". *Chemometrics Intellig. Lab. Syst.* 1997. 38(2): 149–171.
- D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke. "Non-linear Regression". In: D.L. Massart, B.G.M. Vandeginste, L.M.C. Buydens, S. De Jong, P.J. Lewi, J. Smeyers-Verbeke (eds). *Data Handling in Science and Technology, Volume 20, Part A, Handbook of Chemometrics and Qualimetrics, Chapter 11*, pp. 305–338. Elsevier B.V.: Amsterdam, 1998, pp.305–338.

21. L.I. Lin. "A Concordance Correlation-Coefficient to Evaluate Reproducibility". *Biometrics*. 1989. 45(1): 255–268.
22. C. Svendsen, T. Skov, F.V.J. van den Berg. "Monitoring fermentation processes using in-process measurements of different orders". *J. Chem. Technol. Biotechnol.* 2015. 90(2): 244–254.
23. A. Surribas, D. Geissler, A. Gierse, T. Scheper, B. Hitzmann, J.L. Montesinos, F. Valero. "State variables monitoring by in situ multi-wavelength fluorescence spectroscopy in heterologous protein production by *Pichia pastoris*". *J. Biotechnol.* 2006. 124(2): 412–419.
24. J. Fox, D. Thayer. "Radical oxidation of riboflavin". *Intl J. Vitamin Nutrition Res.* 1998. 68(3): 174–180.

POSTER I

Implementation of On-line Monitoring of Cells in Fermentation Processes

Carina Svendsen, Frans W. van den Berg and Thomas Skov

Scandinavian Symposium on Chemometrics (SSC13), Sweden, 2013



Implementation of On-line Monitoring of Cells in Fermentation Processes

Carina Svendsen, Frans W.J. van den Berg, Thomas Skov
 Department of Food Science, University of Copenhagen, Denmark
 carinasv@life.ku.dk

The **BIOPRO** project is a new Danish research initiative focusing on the development of tomorrow's biotechnological production systems. The project will conduct research based on the industrial needs to increase yield and sustainability, aimed specifically in full scale production processes. BIOPRO is a close cooperation between leading biotechnological manufacturers Novo Nordisk, Novozymes, DONG Energy, CP Kelco and CAT Science, and the Danish universities University of Copenhagen and the Technical University of Denmark.

A wide gap to bridge in on-line monitoring is *academic proofs-of-principal* and *workable solutions*. The study will explore the potentials of determining the microbial status in biotechnological productions by characterizing and quantifying cells in reaction systems on-line and in real-time. The project objectives will be pursued starting with existing prototype sampling systems and sensors in order to reach our end results, which is an operational on-line automated system that can potentially be used in any biotechnological related production environment.

Techniques

Multi-wavelength fluorescence spectroscopy produces data

Full-Scale Process Reactor

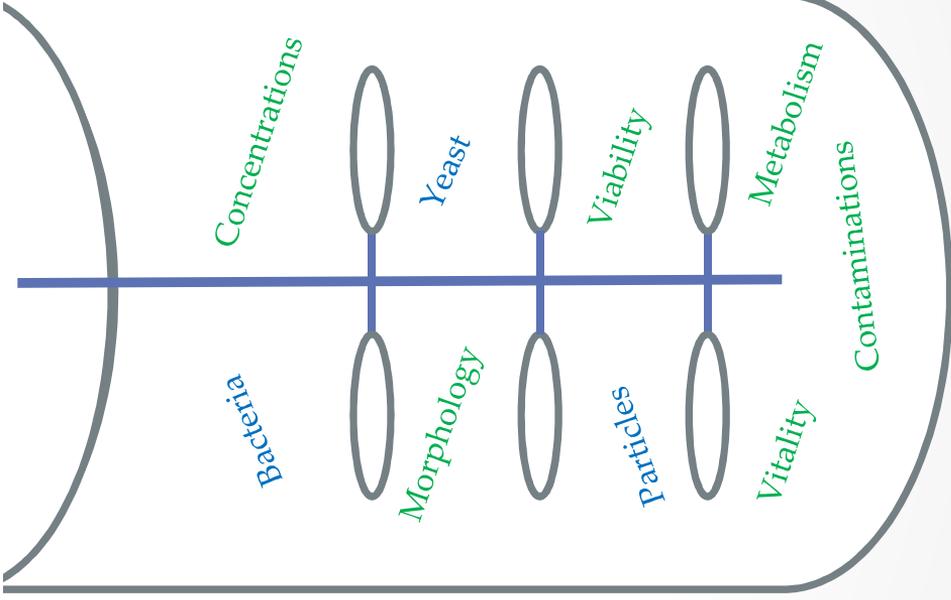
Goals

Implement on-line measurement systems and data

with a three dimensional structure with emission wavelength, excitation wavelength and intensity. The technique can in combination with chemometrics be used for determination of biomass.

Di-electric spectroscopy is able to measure physiological properties of microbial cell suspensions based on the cells ability to accumulate charges when exposed to an electrical field. Thereby, an on-line fingerprint of cell- and culture state can be obtained.

In-situ microscopy is a technique that delivers images of cells in a defined volume inside a bioreactor without sample extraction. The image-based technique is suitable for cell-count, cell-density, chemical solution deposition and morphological changes.



analytical tools for monitoring, control and optimization of cultivation processes in biotech-based production.

Develop a tool to quantify the degree of contamination by living cells in intermediate- and end-products in biotechnology and food production; use measurement information to characterize the (desired and undesired) dynamic behavior of production stages.

Real-time characterization of upstream processes (i.e. fermentations) and downstream process flows (i.e. from separation processes); description of the process with respect to populations of bacteria and yeast in the biomass, viability (ratio between dead and living cells), vitality (state of the cells) and possible infections or contaminations.



NOVOZYMES
Rethink Tomorrow



CPKelco
A HUBER COMPANY

DONG energy **CAT**

POSTER II

A chemometric approach for correction of intensity changes in fluorescence spectroscopy caused by pH fluctuations

Carina Svendsen, Thomas Skov and Frans W. van den Berg

Scandinavian Symposium on Chemometrics (SSC14), Italy, 2015

A chemometric approach for correction of intensity changes in fluorescence spectroscopy caused by pH fluctuations

Carina Svendsen, Thomas Skov, Frans van den Berg

University of Copenhagen, Faculty of Science, Department of Food Science, Spectroscopy & Chemometrics section
Rolighedsvej 26, DK-1958 Frederiksberg C, Denmark
E-mail: carinasv@food.ku.dk



Background

Online monitoring of bioprocesses by spectroscopic methods holds great potential for measuring critical process parameters in real-time. Lately, an increasing focus on fluorescence spectroscopy and its capability of providing valuable information about bioprocesses has been observed. However, our study has shown that uncontrolled pH changes can cause strong intensity shifts in the fluorescence landscapes, which has a major impact on the quantification of the chemical compounds. In the current study the fluorescence shifts were elucidated and a correction strategy for intensity changes caused by pH fluctuations was developed by combining weighted PARAFAC with weighted non-linear regression based on kinetic modeling.

System

- The degradation of Riboflavin, driven by light exposure, and the formation of the two products Lumiflavin and Lumichrome (A → B + C) were measured by fluorescence spectroscopy. Furthermore, a fourth unreactive chemical component was also present in this system.

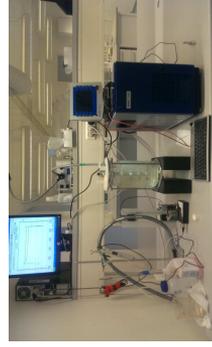


Figure 1: The experimental setup - the Riboflavin solution was exposed to light to induce the degradation process. Notice that the efficiency of the light as "catalyst" determines the reaction kinetics, independent of pH interference.

- Batches having abnormal operation conditions (AOC), with pH changes as interference, were carried out. The pH disturbance was introduced during the degradation process by addition of acid or base as "disturbance".

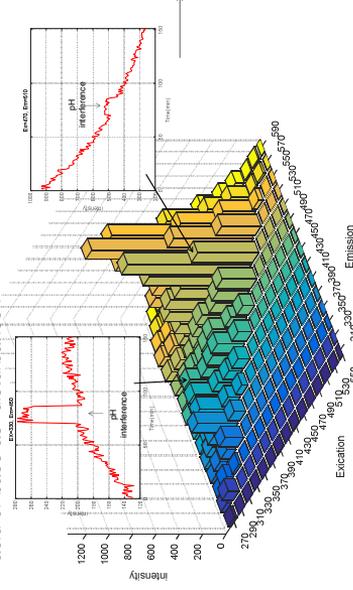


Figure 2: A fluorescence landscape - the spectra at the specific excitation/emission wavelengths, which are affected by the pH disturbance are illustrated and the representing fluorescence areas are outlined with arrows.

- The fluorescence data were modelled by PARAFAC as illustrated in Figure 3. A clear intensity shift, due to the pH interference, is seen in the PARAFAC scores.

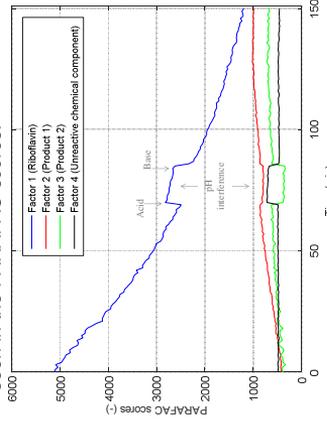


Figure 3: PARAFAC scores of the modelled fluorescence data - the Reactant Riboflavin (Factor 1), the two products (Factor 2 and 3) and unreactive chemical component (Factor 4) are present.

Results

- The pH-profile for batch AOC A (Fig. 4c) and batch AOC E (Fig. 5c) outline the introduced pH interference.
- The pH interference was seen in the PARAFAC scores and did not agree with the kinetic trend (Fig. 4a & 5a).
- Based on this deviation (between scores and kinetic trend), a weight vector was determined and used by both PARAFAC and the non-linear kinetic fitting step (Fig 6).

- The kinetic profiles and K-values obtained after the optimization step (Fig. 6) are shown in Fig. 4b (AOC A) and Fig 5b (AOC E). Also, the final weight vectors are shown (Fig. 4d and Fig. 5d).

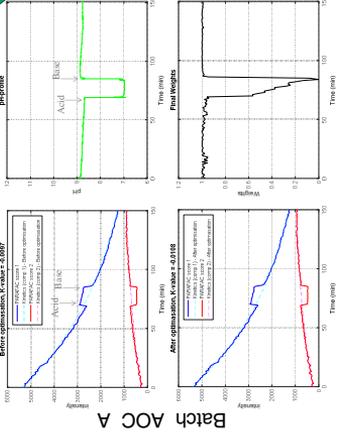


Figure 4: Batch AOC A. The concentration scores and the kinetic profile before (Fig. 4a) and after (Fig. 4b) the optimisation step (Fig. 6), pH profile (Fig. 4c) and the final weight vector (Fig. 4d) are illustrated.

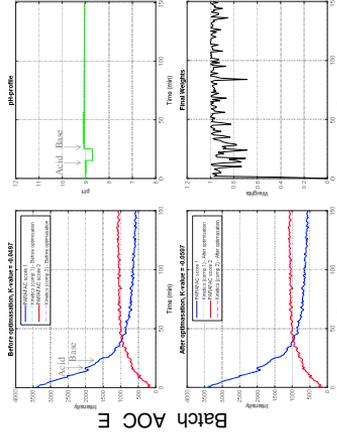


Figure 5: Batch AOC E. The concentration scores and the kinetic profile before (Fig. 5a) and after (Fig. 5b) the optimisation step (Fig. 6), pH profile (Fig. 5c) and the final weight vector (Fig. 5d) are illustrated.

Flowchart of the chemometric approach - for optimisation of the kinetic fit

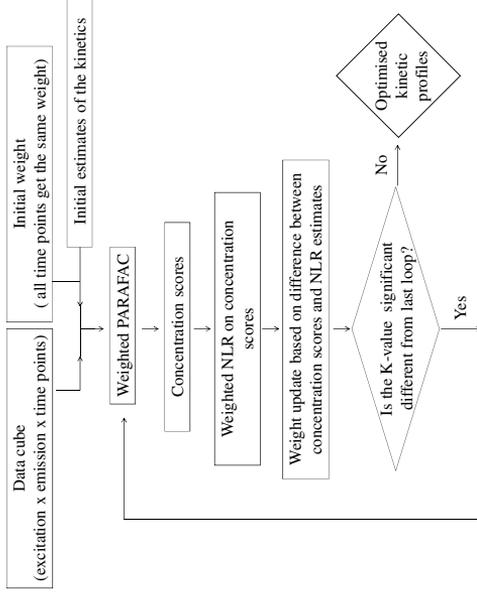


Figure 6: Flowchart of the chemometric approach for correction of intensity changes in the fluorescence signal caused by pH fluctuations.

Conclusions

The chemometric approach suggested in this study combines the kinetic fit based on the weighted PARAFAC profiles with an algorithm modeling the unknown variation. Thereby, it is possible to correct for undesirable intensity changes in Fluorescence caused by pH fluctuations. The approach can be highly relevant in e.g. bioprocesses, where unbiased metabolic profiles are desired for process monitoring.

CARINA SVENDSEN

On-line Monitoring of Fermentation Processes by Near Infrared and Fluorescence Spectroscopy
– elucidating and exploring process dynamics

