# Monitoring Bacterial Growth in Liquid Cultures through the Bulk Optical Parameters in the Near-Infrared Region extracted using the Radiative Transfer Theory

by

**Elitsa Dzhongova**

**A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy**

**School of Chemical Engineering and Advanced Materials**

**Newcastle University**

**May 2010**

*Dedicated to my family – Thank you for always having faith in me*

# Abstract

Near infrared (NIR) spectroscopy offers promise as a monitoring tool for fermentation reactions by its potential to provide information about the physical and chemical state of given system. However, issues associated with confounding effects due to light scattering variations that occur during the fermentation process due to changes in the state of the microorganisms makes it difficult to obtain robust estimation of the analytes of interest.

An approach for removing multiple scattering effects and separating absorption from scattering using the radiative transfer equation (RTE) is proposed in this work. In order to investigate its feasibility and to aid in the development of the technique, the method was applied to a simple system consisting of *Bacillus subtilis* growing in an aqueous solution. In this study optical properties (absorption coefficient $\mu_a$, scattering coefficient $\mu_s$, and anisotropy factor $g$) were estimated and their changes during growth, stationary, and decline phases of *Bacillus subtilis* culture were examined. It was found that the greatest changes were seen during the growth phase, predominantly manifested in the scattering spectra. Effect of sample thickness on the estimation of the optical properties throughout the cultivation was investigated. The extracted absorption and scattering spectra were found to be fairly consistent even though they were obtained from measurements from different sample thicknesses. The extracted optical properties were used to develop Partial Least Squares (PLS) models for prediction of biomass and glucose concentrations. The performance of these models was compared with those obtained using the raw measurements. As a result of the comparison it was revealed that scattering coefficient based models demonstrate good performance while predicting biomass. The accuracy of these models was equal or in some cases greater then the accuracy of the models built on transmittance and reflectance measurements. Glucose, the second analyte of interest in this work, was modeled with limited success and models were able to distinguish only between low and high level of glucose concentration. The effect of sample thickness on PLS models performance was also studied. Results have shown that sample thickness need to be chosen based on the specific analyte of interest in order to achieve good performance of the models.

# Acknowledgments

# Publications from the thesis

## Conference presentations

1. Elitsa Dzhongova, Colin R. Harwood, Suresh N. Thennadil, 'Monitoring bacterial growth through changes in the optical properties in the Near-Infrared wavelength region'. APACT 2-4 May 2007 (Advances in Process Analytics and Control Technology Conference 2007), Edinburgh, Scotland.

2. Elitsa Dzhongova, Colin R. Harwood, Suresh N. Thennadil, 'Application of multiple light scattering theory for monitoring fermentation using NIR spectroscopy'. EuPAT2 13-14 November 2007 (Second pan-European PAT Science Conference), Copenhagen, Denmark.

3. Maria Alejandra Velazco-Roa, Elitsa Dzhongova, Colin R. Harwood, Suresh N. Thennadil, 'Complex refractive index of non-spherical particles in the Vis-NIR region-Application to Bacillus subtilis spores'. OPC 2007 (8th International Congress on Optical Particle Characterization), 9-13July 2007, Graz, Austria.

4. Elitsa Dzhongova, Colin R. Harwood, Suresh N. Thennadil, 'Fermentation reaction characterization using multiple-scattered light in the NIR region'. CAC 2008 (11th Conference on Chemometrics in Analytical Chemistry), 30 June-4 July 2008, Montpellier, France.

## Journal papers

1. Maria A. Velazco-Roa, Elitsa Dzhongova, Suresh N. Thennadil, 'Complex refractive index of nonspherical particles in the visible near infrared region-applications to *Bacillus subtilis* spores'. Applied Optics, v.47, iss.33, pp. 6183-6189 (2008).

2. Elitsa Dzhongova, Colin R. Harwood, Suresh N. Thennadil, 'Changes in the absorption and scattering properties in the Near-Infrared region during the growth of *Bacillus subtilis* in liquid culture. Applied Spectroscopy, v.63, iss.1, pp. 25-32 (2009).

# Contents

## Nomenclature

$a$       additive effect

$b$       multiplicative effect

$g$       anisotropy factor

$i$       sample number in a vector

$I_o$       specific intensity

$k$       wavelength number

$\ell$       pathlength, mm

N       total number of samples

p       phase function

$R^2$       total coefficient of determination

Rd       total diffuse reflectance

$s$       standard deviation

Td       total diffuse transmittance

Tc       collimated transmittance

$\bar{x}_k$       mean value

$x_k$       vector of uncorrected spectral data

$x_k^*$       vector of corrected spectral data

$x_m$       mean spectrum

$y$       vector of dependent variables

$\hat{y}$       vector of fitted values

**Greek Letters**

$\alpha$      bulk absorption albedo

$\sigma_s$      scattering cross section, $mm^2$

$\sigma_a$      absorption cross section, $mm^2$

$\lambda$      wavelength, nm

$\mu_a$      bulk absorption coefficient, $mm^{-1}$

$\mu_s$      bulk scattering coefficient, $mm^{-1}$

$\mu_s'$      reduced scattering coefficient, $mm^{-1}$

$\mu_t$      total extinction coefficient, $mm^{-1}$

$\theta$      scattering angle, rad

$\omega$      solid angle

$\rho$      number density

$\xi$      convergence tolerance

$\tau$      optical depth

$\varepsilon$      source term

# List of Figures

# List of Tables

# List of Pictures

# Introduction

For more than 20 years, considerable research efforts have been dedicated to investigate near-infrared spectroscopy (NIR) as a method for monitoring microbial fermentation processes, since this technique has the potential to provide information about both the chemical and physical state of a system. The fact that NIR measurements could be made with minimum or no-sample preparation, coupled with the availability of robust, easy to operate instruments with high signal-to-noise characteristics, has made this technique a promising tool for at-line and on-line monitoring of microbial growth [1, 2]. Several research groups have investigated the utility of NIR spectroscopy on various fermentation processes such as wine [3, 4] and beer [5] fermentation, and batch, fed-batch and continuous cultivations [6-9] in combination with chemometric techniques.

Many of the studies that have been published identified light scattering as a challenge that is often dealt by applying scatter correction techniques [10]. The main idea behind treating the measured spectra with these techniques is that for calibration purposes, spectra should contain as little irrelevant information (reducing effects such as light scatter) as possible. In this context, multiple scattering of light poses a significant challenge in the development of NIR-based methodologies to extract reliable chemical and physical information contained in the spectra of growing suspensions containing microbial cells. The extent of information that can be extracted from NIR spectra could, in principle, be vastly improved if the scattering and absorption effects can be effectively separated.

In this work an approach to separate the absorption and scattering effects using the Radiative Transfer Equation (RTE) to account for multiple scattering of light [11, 12] is investigated. The requirement for the investigation is needed due to a problem occurring when samples with high concentration of particles are studied through spectroscopic

methods. In such cases, obtaining good calibration models from traditional measurements (transmittance and reflectance) is a challenge, due to difficulties in applying Beer-Lambert's law, with complications arising from nonlinear scattering effects. Multiple scattering events lead to variations in the effective pathlength of the photons and therefore contribute to variations in absorbance measurements. In addition, it could cause potentially large errors in PLS calibration models that are based on the assumption of linear relationship between absorbance and concentration, and could therefore result in lack of robustness.

In this context, absorption and scattering spectra can be viewed as new variables, potentially providing comprehensive description of the state of systems such as growing microbial cultures. The absorption and scattering coefficients are expressed as a measure per unit length which is independent of  the pathlength travelled by photons, a fact considered to be one of the main sources of variations during traditional spectroscopic measurements[13].

 Once extracted, changes in the absorption and scattering properties can be separately monitored during microbial growth cycle. The scattering coefficient will link to changes in properties such as cell size, and biomass concentration, whereas information regarding non-scattering constituents such as glucose, and product concentrations will be contained in the absorption coefficient.

Further, from the point of view of building calibration models for estimating the concentrations of product and nutrients, the removal of the confounding effects due to multiple light scattering from the absorption effects could lead to simpler and more robust models [13]. As a first step towards achieving reliable models for predicting the concentrations of components in a culture medium by separating the absorption and scattering effects, this study focuses on the methodology for extracting the bulk optical

properties and investigates the nature and extent of changes in the optical properties over the course of a bacterial growth cycle.

## 1.1 Aim and contributions.

The main goal of this work is to develop a method for the estimation of optical properties in the Near-Infrared region, for *Bacillus subtilis* growing in a liquid culture, where multiple scattering is taken into account. *Bacillus subtilis* has been extensively studied which led to a very good understanding of its molecular biology, and as such it became a model experimental Gram-positive microbe. In addition to their role as experimental organisms, strains of the genus *Bacillus* are also widely used as sources of products of industrial importance. The main products associated with the genus *Bacillus* are hydrolytic enzymes, antibiotics, insecticides and fine biochemicals. Due to its applications and its well understood molecular biology, this genus became widely used by many specialists from disciplines other than that of microbiology. For the same reasons *Bacillus subtilis* has been chosen in this study as a model for feasibility study of optical properties of suspensions growing in a liquid culture.

Bulk absorption coefficient $\mu_a$, bulk scattering coefficient $\mu_s$ and anisotropy factor $g$ are investigated in order to be used for the extraction of important physical and chemical information from the given system. The optical parameters are obtained by using the inverse adding-doubling method to solve the radiative transfer equation in combination with measurements for total diffuse reflectance, total diffuse transmittance and collimated transmittance using a single integrating sphere setup. The changes in the optical properties of this system during growth, stationary and decline phase were studied using measurements in the wavelength region of 950nm – 1850nm. This range was chosen because it includes the first and second overtone bands of proteins, sugars, and other organics. While it would have been desirable to include the combination band

region (2000-2500nm), the low quality of signal in this region due to high water absorption indicated that with the current setup it would not be possible to obtain measurements with sufficient quality.

The contributions of this research work can be summarised as follows:

1) The development of methodology for the extraction of optical properties in a consistent manner from *Bacillus subtilis* growing culture. The changes of the optical properties were also studied during growth (work was carried out over the exponential part of the growth only), stationary and decline phase using measurements in the Near-Infrared region with sample thickness of 4mm. The results of this work have been published [14].

2) Additional experiments and analysis were conducted in order to examine the effect of sample thickness on the estimation of the optical properties. Further analyses were also carried out to examine how the relationship between the biomass and the optical properties as well as the reflectance and transmittance are affected by the changes in the pathlengths. Spectroscopic measurements and corresponding assays on *Bacillus subtilis* suspensions were made using 2mm, 4mm, and 10mm sample thicknesses. Reflectance and transmittance spectra were measured while the culture was going through growth, stationary and decline phases.

3) Calibration models were built using partial least square regression (PLS) for predicting glucose concentration and biomass during bacterial growth of *Bacillus subtilis* in liquid culture, following a two - step approach. The first step consisted of using the raw measurements (total diffuse transmittance, total diffuse reflectance, and collimated transmittance) to extract bulk absorption coefficient, bulk scattering coefficient and anisotropy factor.

This was followed by building PLS models for predicting glucose and biomass concentration during bacterial growth cycle. Models were built both using the extracted optical properties and directly using the raw measurements (i.e. using the traditionally chemometric approach). The performances of the calibration models developed by these two approaches were compared. Also the effect of sample thicknesses on the performances of the models built using both the approaches was investigated.

## 1.2 Organisation of the thesis.

The thesis is organised as follows:

Chapter 2 presents key components of bioprocesses, classical techniques for analysis, and NIR spectroscopic applications for bioprocess monitoring. A number of issues associated with NIR application to bioprocesses have been briefly discussed including pathlength variations, light scattering effect, and noise in the raw spectral data.

Chapter 3 describes a method for finding the optical properties (scattering coefficient, absorption coefficient and anisotropy factor) of a turbid sample by using diffuse reflectance, diffuse transmittance and collimated transmittance measurements. The general idea for this method is to be applied for monitoring changes for optical properties of *Bacillus subtilis* growing in a liquid culture in order to extract important physical and chemical information from the system. The extraction of optical properties through the inversion of the radiative transfer equation using an iterative adding-doubling method is discussed. Validation of the measurement setup and the Inverse Adding Doubling Method (IAD) is also described. A description of the integrating sphere setup and the experimental methodology used to measure total diffuse reflectance, total diffuse transmittance, and collimated transmittance is also given here.

Chapter 4 introduces a method for the estimation of optical properties in the Near-Infrared region for *Bacillus subtilis* growing in a liquid media. The bulk absorption coefficient $\mu_a$, scattering coefficient $\mu_s$ and anisotropy factor $g$ are obtained by using the inverse adding-doubling method to solve the radiative transfer equation in combination with measurements for total diffuse reflectance, total diffuse transmittance and collimated transmittance using a single integrating sphere setup. The changes in the optical properties of the system were studied during growth, stationary and decline phases.

Chapter 5 describes the experiments and analysis conducted to examine the effect of sample thickness on the estimation of the optical properties. Further analysis is also carried out to examine how the relationships between the biomass and the optical properties as well as the reflectance and transmittance measurements are affected by the changes in the sample thickness. These studies were carried out using spectroscopic measurements that were made using 2mm, 4mm, and 10mm sample thicknesses while the culture was going through growth, stationary and decline phases.

Chapter 6 investigates the effectiveness of predictive models for biomass and glucose concentration based on raw measurements (total diffuse transmittance, total diffuse reflectance, and collimated transmittance) and extracted optical properties (bulk absorption coefficient and bulk scattering coefficient). Comparison of the performance of the constructed models is made and the effect of sample thickness on the performance of PLS models is evaluated.

Finally, a summary of the key findings of this work, the main conclusions and recommendations for future work are presented in Chapter 7.

### References

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

2.  Vaidyanathan, S., G. Macaloney, J. Vaughn, B. McNeil, and L. Harvey, *Monitoring of Submerged Bioprocesses.* Critical Reviews in Biotechnology, 1999. **19**(4): p. 277-316.

3.  Cozzolino, D., R. Dambergs, L. Janik, W. Cynkar, and M. Gishen, *Analysis of Grapes and Wine by Near Infrared Spectroscopy* Journal of Near Infrared Spectroscopy, 2006. **14**(5): p. 279-289.

4.  Cozzolino, D., M. Parker, R. Dambergs, M. Herderich, and M. Gishen, *Chemometrics and visible-near infrared spectroscopic monitoring of red wine fermentation in a pilot scale.* Biotechnology and Bioengineering, 2006. **95**(6): p. 1101-1107.

5.  Engelhard, S., M.U. Kumke, and H.G. Lohmannsroben, *Examples of the application of optical process and quality sensing (OPQS) to beer brewing and polyurethane foaming processes.* Analytical and Bioanalytical Chemistry, 2006. **384**(5): p. 1107-1112.

6.  Arnold, S.A., L.M. Harvey, B. McNeil, and J.W. Hall, *Employing near-infrared spectroscopic methods and analysis for fermentation monitoring and control - Part 1, method development.* Biopharm International-the Applied Technologies of Biopharmaceutical Development, 2002. **15**(11): p. 26-34.

7.  Arnold, S.A., L.M. Harvey, B. McNeil, and J.W. Hall, *Employing near-infrared spectroscopic methods of analysis for fermentation monitoring and control - Part 2, implementation strategies.* Biopharm International-the Applied Technologies of Biopharmaceutical Development, 2003. **16**(1): p. 47-70.

8.  Rathore, S., M. Paulsen, V. Sharma, and V. Singh. *Near-Infrared Spectra Interpretation of Fermentation in a Dry Grind Corn Process.* in *ASAE Annual Internationa Meeting.* 2005. Tampa Convention Center, Tampa, Florida.

9.  Scarf, M., A. Arnold, L. Harvey, and B. McNeil, *Near Infrared Spectroscopy for Bioprocess Monitoring and Control: Current Status and Future Trends.* Critical Reviews in Biotechnology, 2006. **26**(1): p. 17-39.

10. Naes, T., T. Isaksson, T. Fearn, and T. Davies, *A user-friendly guide to multivariate calibration and classification.* 2004: NIR Publications.

11. Ishimaru, A., *Wave Propagation and Scattering in Random Media.* The IEEE/OUP Series on Electromagnetic Wave Theory. 1997: IEEE Press, Oxford University Press. 574.

12. Hulst, C.V.d., *Light Scattering by Small Particles.* 1964, USA: John Wiley&Sons. 470.

13. Steponavicius, R. and S.N. Thennadil, *Extraction of chemical information of suspensions using radiative transfer theory to remove multiple scattering effects: application to a model two-component system.* Analytical Chemistry, 2009. **81**(18): p. 7713-7723.

14. Dzhongova, E., C.R. Harwood, and S.N. Thennadil, *Changes in the absorption and scattering properties in the near-infrared region during the growth of Bacillus subtilis in liquid culture.* Applied Spectroscopy, 2009. **63**(1): p. 25-32.

# Monitoring of bioprocesses

## 2.1. Introduction

Manufacture of wine, beer and yogurt, as well as many pharmaceutical products are based on bioprocess. The successful development of these processes depends on well defined and strictly controlled key factors and conditions and on the quality of the end product. For this, off-line and on-line measurements are employed in order to monitor and provide information regarding changes that may occur in a given process. Off-line analysis are widely used for quality assurance of the final product, however it is the on-line measurements that can provide information about the process in real time, can allow corrective actions to be taken in a timely manner, and can help decrease production costs.

This research concerns near-infrared measurements and their application for bioprocess monitoring specifically fermentation reactions.

Accordingly, the chapter begins with a summary introducing key parameters monitored in bioprocess systems, then followed by a brief discussion on established methods and techniques for analysis and quality control. The use of Near-Infrared (NIR) spectroscopy technique is the focus as already stated because of its potential to provide rapid and real-time information with minimal or no sample preparation. As such, the technique has been investigated for bioprocess monitoring and control[1] by many research groups and these various applications are summarized in a separate section (section 2.3). During these investigations, it has been found that one of the main challenges to deal with during the process arises from multiple scattering of light[2]. As mentioned in Chapter 1, multiple scattering by the cells introduces pathlength variations, which degrade chemical information, and if absorption and scattering can be effectively separated then the extraction of information can be improved significantly. From this perspective, the

origin of multiple scattering, and practical problems associated with the multiple scattering effects are introduced at the end of the chapter (section 2.4).

## 2.2. Bioprocess systems and key parameters

Before discussing the important parameters and monitoring techniques, a brief introduction to fermentation reactions is presented in the next section.

### 2.2.1. Fermentation reactions

In a general sense, fermentation is the conversion of carbohydrate such as sugar into an acid or an alcohol but more specifically, fermentation can refer to the use of yeast to change sugar into alcohol or the use of bacteria for formation of lactic acid. The most simple reaction describing a fermentation process (equation 2.1) is expressed as the anaerobic fermentation of a molecule of glucose to yield two molecules of ethanol and two molecules of carbon dioxide,

$$C_6H_{12}O_6 \rightarrow 2C_2H_5OH + 2CO_2 \qquad (2.1)$$
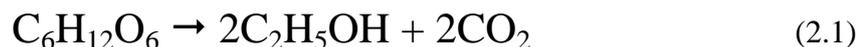
Nearly all fermentation processes are based on microbial growth which is usually initiated by inoculating a microbial organism into a fixed volume of nutrient medium. Figure 2.1 shows a typical growth cycle in batch culture where four main phases can be easily distinguished: lag, exponential, stationary and death phase. The lag phase is the period after inoculation during which there is no observable increase in the biomass. Individual bacteria are in a vegetative state, not yet able to divide, only adapting to the new environmental conditions (pH, temperature, nutrients, etc.). During the exponential growth (exponential or log phase), the living bacterial population increases rapidly with time until it becomes nutrients limited. This is the phase with optimal conditions for

growth. The stationary phase is characterized by slower growth rate as a result of nutrient depletion and accumulation of toxic products. The last of the four phases of the process is the death phase during which living bacterial population decreases with time exponentially, due to a lack of nutrients and formation of toxic metabolic by-products.
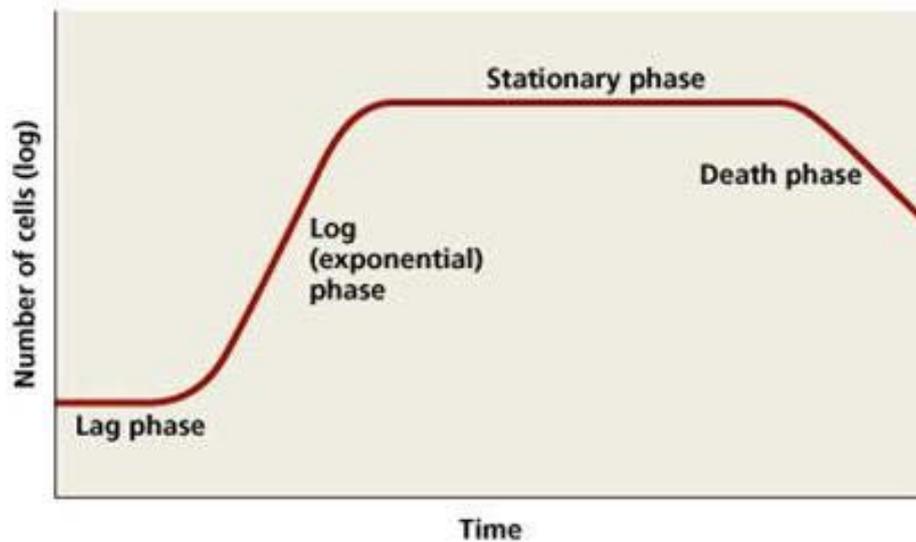


Fig. 2.1 Bacterial growth in batch culture (source: IRSC website, http://faculty.ircc.edu/faculty/tfischer/images/).

The four phases presented above are typical for batch fermentation processes when the media and its components are defined and added at the beginning of the fermentation. The volume of the liquid medium is fixed, and the growth of the microorganism takes place in a closed vessel with limited amount of oxygen. During the period of the incubation no addition of nutrients or removal of effluents is made. Apart from batch cultures, fermentation can be carried out on a fed-batch and continuous culture. In the case of fed-batch process, nutrients are continuously or semi-continuously added to the system while the effluent is removed discontinuously. The specific growth rate of the culture will be slower then in batch cultures and therefore the cultivation time is significantly longer. In the fed-batch process is possible to reach cell densities up to

$OD_{600}$ of 400 in bioreactors which is significantly higher then in shaking batch cultures, in which enough oxygen can be delivered only until $OD_{600}$ from 1 to 2. Continuous fermentation is a process in which the microorganism in the culture is maintained in the exponential growth phase by the continuous addition of fresh nutrient medium and that is exactly balanced by the removal of an equal quantity of fermented material.

The main focus in this research work is on bacterial growth and changes in the absorption and scattering profiles through all four growth phases in a batch culture. The following section will briefly introduce main key parameters typically monitored over the course of a fermentation process.

### 2.2.2. Key parameters in bioprocess monitoring

The success of the bioprocesses and their high quality end product depend on well defined parameters and conditions which requires monitoring during the production process. Three main groups of parameters can be distinguished - physical, chemical and biological – which are described below.

*Physical parameters*

One of the most important parameters for monitoring and control in any bioprocess is *temperature*. It is essential to maintain temperature at an optimal level for the constant activity of the cell during the growth cycle, and this optimum could be different depending on the different bacterial species. After completion of batch fermentation processes, fermentation vessels have to be emptied, cleaned, and sterilized following standard operating procedures where temperature guarantees the effectiveness of heat sterilization process. Sterilization is often required as a final step while preparing nutrient medium, where constant high temperature ($\sim 121^{0}$C) over a period of time (in most

cases ~15 min) ensures sterile conditions for the fermentation medium before inoculation of the bacterial strain.

To achieve efficacious sterilization process, *pressure* is an additional physical factor that has to be monitored. Pressure monitoring is also needed on industrial and pilot plant scale for bioprocesses in order to operate safely and efficient since the majority of the industrial and laboratory equipments are designed to withstand a specified working pressure. In most cases the equipments are often connected with devices for indication, records and control of the pressure.

*Agitation speed* and *power* depend on the type and size of the bio system and are controlled by the rate of rotation (rpm) of the stirrer shaft.

Other essential physical parameters for bioprocess monitoring and control are *viscosity* and *density*, *liquid* and *foam level* as well as *gas* and *liquid flow*.

Presently conventional sensors are widely used for reliable on-line measurements of most physical factors.

*Chemical parameters and components*

The most commonly measured parameter indicating the state of a bioprocess is *pH* due to its importance for optimum cell growth. In actively growing cultures the *pH* will not remain constant for very long and if maximum yield of a product is to be expected, then the *pH* has to be measured and controlled depending on the process. Gunsalus and Niven, in 1942 investigated the effect of three different *pH* on a lactic acid fermentation process, and particularly on the yield of the final products[3]. They found that the yield of lactic acid may fall to 60 percent or less when the reaction is held at *pH* 6.5 or above.

Lactic acid fermentation process is classified in the list of the anaerobic processes where

the factor *redox potential* measures the oxidation-reduction potential of the biological system and is the key point when assuring strict anaerobic conditions. When considering aerobic bioprocess monitoring, *dissolved oxygen* concentration values in the fermentation medium are carefully measured. The main reason is that the maximum biomass formation heavily depends on the dissolved oxygen level which must be kept above the critical concentration ($c_{crit}$ usually is assumed to be less than 0.05 mmol/dm$^3$ [4]).

In bioprocesses, physiology of cells has been found to significantly depend on the concentration of *dissolved carbon dioxide* ($dCO_2$). This is because $dCO_2$ level in the fermentor affects the pH inside the cell which can inhibit cell growth and decrease product formation.

The concentration of various chemical components (both nutrients and products) in a fermentation medium is important for bioprocesses from the point of view of modelling and optimization of a given reaction. Among the most studied nutrient constituents are glucose, lactate, ammonia, and phosphate while products that have been studied include tetracycline, penicillin, amino acids, lactic acid, and ethanol.

*Biological variables*

The key biological variable that is usually monitored in a bioprocess system is the biomass concentration. It is a simple measure of the available quantity of mass, activity, physiological state, morphology or other parameters in a bio-reaction system. The biomass provides information on the rate of growth and product formation in any bacterial culture. Different methods exist when measuring biomass concentration but each one possesses its own limitations. However dry cell weight is accepted to be the *golden standard[5]* for the determination of cell mass concentration and is widely used in the laboratory and industry practices.

The main purpose of this section was to introduce key factors playing a significant role on the bioprocess performance due to their direct connection with the state of the bio system, its physiological behaviour and morphological changes. These variables need to be monitored and controlled in order to better understand modern bioprocesses and to continuously improve their performance. Numerous methods and techniques are applied nowadays while monitoring bioprocesses and the next section aims to present their practical application in both laboratory and industrial scales.

### 2.2.3. Monitoring techniques used in bioprocesses

The rapid development of biotechnology enhanced the need for the development and application of different sensors, electrodes, and sophisticated systems for bioprocess monitoring and assessing their progress. Various methods exist for the monitoring of a fermentation process and for the analysis of the cultivation media and its constituents. While this thesis focuses on the application of NIR spectroscopy for bioprocess monitoring, a brief survey of other methods such as wet chemical analysis, chromatographic methods, and biosensors is given to indicate the range of techniques that are used to monitor such processes.

*Wet chemical analysis method*

Wet chemical analysis forms the basis of greatly improved reference methods and contributes for the decrease of laboratory time usually invested in conventional methods. From the bioprocess perspective, wet chemical analysis demand good sample preparation as part of the total analysis procedure where cells need to be removed from the fermentation media. Continuous Flow Analysis (CFA) represents the early laboratory systems which have been restricted to chemical analysis only. The analysis is performed

in a non-sterile condition where high risk of contamination exists when considering microbiological samples. To improve the analysis and to extend its application to reflect the needs of biotechnology, CFA were replaced by Flow Injection Analysis Systems (FIA). The main advantages of FIA are short response times and reduced costs in comparison with conventional methods. Two types of systems are accessible for laboratory and for industrial applications[6]. Industrial systems are highly reliable however lack of flexibility is the main disadvantage, whereas laboratory systems allow multiple experimental studies due to their flexible configurations. Glucose is among the most measured components[7] in the fermentation processes due to its nutrition importance for bacterial growth. However these systems are still not fully implemented as an on-line industrial fermentation monitoring technique due to high risk of contamination of the bioprocess and large reagent consumption during analysis.

*Chromatographic methods*

Chromatography is a collection of laboratory techniques used for separation and identification of components or solutes of a complex mixture, on the basis of relative amounts of each solute distributed between a moving fluid stream, called the mobile phase, and a contiguous stationary phase. The mobile phase may be either a liquid or a gas, while the stationary phase is either a solid or a liquid (source: Encyclopaedia Britannica).

The type of chromatography in which the mobile phase is a gas (usually helium, hydrogen or nitrogen) is known as gas chromatography (GC). GC is applied in analytical chemistry to monitor gaseous and volatile components, such as aromatics (benzene, toluene etc), flavours and fragrances, and also permanent gases ($H_2$, $N_2$, $O_2$, Ar, $CO_2$, $CH_4$). Gas Chromatography requires very small quantity of samples (less then 1 μl) with

little preparation, however it is difficult to measure and inject such small samples accurately without evaporation for example. The main limitations for GC are the samples, they have to be thermally stable and once separated can not be used for further analysis.

High Performance Liquid Chromatography (HPLC) is a separation technique for analysis of complex mixtures in which the mobile phase is liquid. HPLC operates under high pressure (up to 400atm) and is one of the most powerful tools in analytical chemistry. It has the ability to separate, identify and quantify compounds that are present in any sample that can be dissolved in a liquid. The main advantage is the ability to identify compounds in trace concentration as low as parts per trillion (ppt). It has been applied for the study of various samples: pharmaceuticals, food, cosmetics, forensic samples and industrial chemicals. HPLC allows accurate and reproducible results with a high degree of precision but at a very high price for instrumentation, software and maintenance. It is mainly used off-line while studying bioprocesses, mainly because of its long analysis time.

Other chromatography techniques applied for bioprocess monitoring and control are fast protein liquid chromatography (FPLC), membrane chromatography, polyacrylamide gel electrophoresis and capillary electrophoresis. Despite its attractions, chromatography is more suited for research and development applications, rather than routine industrial process monitoring.

Generally, chromatography technique has many advantages and is widespread in modern analytical chemistry. The technique is very sensitive and reliable (provided that the method and analysis is carried out carefully without any contamination). As already mentioned, complex mixtures can be separated accurately with a very small amount of sample. However one of main disadvantages is that since the method is very sensitive,

improper setup or contamination will provide different results, and analytical reproducibility can vary or even provide very inaccurate results.

*Spectroscopic methods*

Generally spectroscopy is the study of the interaction between radiation and matter as a function of the wavelength. Most common methods assembled in this group are ultra violet-visible (UV-Vis) and infrared spectroscopy, mass spectrometry and nuclear magnetic resonant spectroscopy. However priority is given in this chapter only to UV-Vis and infrared techniques due to their direct relation to the presented research.

UV-Vis spectroscopy is applied in the biochemical laboratory when for example measuring the optical density of cell suspensions. Optical density is an indirect measurement of dry weight of cells and is defined by assessing the turbidity at wavelength around 600nm[8]. Estimation of the biomass concentration with this technique has to be validated for every individual bio system before the start of a real experiment. Glucose concentration can be also monitored at 540nm after the measured sample undergoes reaction process with special set of chemicals (e.g. Sigma Aldrich test kit for glucose measurements). The UV-Vis region of the electromagnetic spectrum starts at 180nm and ends around 770nm, where the infrared region begins.

Infrared region is positioned from 770nm up to 1000μm and is subdivided into Near-Infrared (NIR), Mid or fundamental Infrared, and Far Infrared, named with respect to their relation to the visible spectrum. IR spectroscopy is widely used in both research and industry as a simple and reliable technique for measurement and quality control, and perhaps is the most used method of applied spectroscopy. However, it is rarely possible to identify an unknown compound by using IR spectroscopy alone. Its principal strengths are (1) quick and relatively cheap technique, (2) useful for identifying certain functional

groups in molecules and (3) an IR spectrum of a given compound is unique and can serve as a fingerprint of this compound.

NIR spectroscopy utilizes the near-infrared region. It has wide-ranging applications including pharmaceutical monitoring and medical diagnostics, as well as food and agrochemical quality control. Enormous research efforts have been and are at present concentrated in applying the NIR technique for on-line monitoring of fermentation processes for several reasons: (1) analysis based on NIR can be done within seconds or minutes, and (2) minimal or no sample preparation are required in order to perform the analysis. Since the NIR technique is the focus of this thesis it will be described in greater details in a separate section.

*Biosensors*

Sensors can be divided into three types: (1) physical sensors for measuring distance, mass, temperature, pressure etc; (2) chemical sensors for measuring chemical substances by chemical or physical responses; and (3) biosensors that utilize biological sensing elements. In this section special attention is given to biosensors, while chemical and physical sensors are not considered.

A biosensor is defined as "*a compact analytical device incorporating a biological or biologically derived sensing element (such as an enzyme, antibody, microbe or DNA) either integrated within or intimately associated with a physicochemical transducer*" [9]. This type of device combines two parts: bio element and a sensor element. A specific bio element, for example enzyme, recognizes a specific analyte and the sensor element transduces the change in the biomolecule into an electric signal. Depending on the transducing mechanism biosensors can be of many types, such as electrochemical, optical, calorimetric, acoustic etc. The most commercially successful biosensors are

glucose biosensors. These biosensors have great impact on the quality of life of people with diabetes. They are widely used by people suffering from diabetes for self-monitoring of capillary blood glucose.

Another type of biosensor can be used to monitor cell morphology (known as electric cell-substrate impedance sensing, ECIS), or for DNA detection (known as force amplified biological sensor-FABS, bead array counter-BARC, and force differentiation assay-FDA). At present, biosensors form considerable part of the development of new approaches for rapid detection of microorganisms.

Several advantages are usually associated with biosensors - usually they are less time consuming when compared with conventional methods, more efficient, their design allows analysis of small volume samples, and they are usually very economic.

With the end of this section, various approaches that have found considerable application in the area of bioprocess monitoring were summarised. The purpose of these techniques is to substitute the conventional methods for monitoring due to their relatively long analysis time, need for highly skilled scientist to perform the analysis, and very expensive consumables. Due to this, a significant number of instruments have been developed in the recent years which are used at present for monitoring and control of key variables by sensors, developing mathematical models and are relatively simple devices or equipment.

While advances in relatively simple instrumentation techniques for online monitoring have been made, the reliability and sensitivity of some systems are still questionable. In addition, some analyses are still time consuming, very complex and require the use of environmentally unacceptable solvents. From this perspective, Near-Infrared spectroscopy has the potential to be applied and incorporated into monitoring and control of bioprocesses. NIR spectroscopy has proven to be non-invasive and non-destructive

technique with minimal or no sample preparation where the analysis and results can be obtained within few minutes.

Literature review in the next section introduces various NIR applications utilized for bioprocess monitoring in both, laboratory and industrial scale, along with associated chemometric methods often used to reduce scatter effects from spectral data.

## 2.3 Near-Infrared spectroscopy techniques for bioprocess monitoring

The Near-Infrared (NIR) region of the electromagnetic spectrum is generally considered to lie in the wavelength range from 700nm to 2500nm, and enables the analysis of complex samples in a rapid, non destructive way, without complex sample pre-treatment. During the last two decades, there has been considerable interest in utilizing NIR spectroscopy for monitoring bioprocesses, since this technique has the potential to provide information about both-the chemical and physical state of a system. The fact that NIR measurements could be made with minimum or no-sample preparation, coupled with the availability of robust, easy to operate instruments with high signal-to-noise characteristics, has made this technique a promising tool for at-line and on-line monitoring of microbial growth[10, 11].

In the NIR region, a typical analyte absorbs at more than one wavelength and conversely, the absorbance at a given wavelength may be due to more than one analyte[12]. Multivariate calibration techniques[13], such as partial least squares, are employed to extract relevant chemical information from the sample's spectra based on correlation between absorption of NIR radiation and the analytical data. The correlation can be improved through the use of preprocessing techniques and variable selection methods. Spectra are usually preprocessed to reduce variations not directly related to the compound concentration, such as random noise, baseline drift and light scattering.

In order to understand the impact of the scattering effect on the raw spectra, and to present widely applied solutions for removing this effect, a literature review will introduce different bioprocesses that have been studied, and particular chemometric techniques that have been employed for spectral data treatment.

The utility of NIR spectroscopy has been investigated on various fermentation processes, for example Cozzolino et *al.* reported the use of this technique for analysis of grape and wine [14] and Engelhard et *al.* presented some examples of NIR application to beer brewing[15]. Apart from NIR application to these traditional fermentation processes, the pharmaceutical industry is undergoing essential changes initiated by the US Food and Drug Administration (FDA) with the announcement of their guidance document on Process Analytical Technology (PAT) toolbox in 2004[16]. NIR spectroscopy appeared to be one of the most dominant process analytical chemistry tools in the context of the PAT toolbox, and is widely applicable within the pharmaceutical industry.

In their work, Brimmer and Hall demonstrated the applicability of NIR spectroscopy for bioprocess monitoring[2], where the determination of nutrient levels has been performed by using NIR in the 1100-2500nm region. Spectral samples were collected in reflectance mode because of the highly scattering nature of the fermentation broth and in transmission mode were collected samples from the liquid raw ingredients. In both cases no samples pretreatment or temperature control were applied. The bioprocess itself consisted of fermentation broth, aqueous starch suspension, ammonium nitrate and nutrient oil. As the reaction proceeded, scattering differences occurred due to biomass formation in the fermentation broth, and a multiple linear least-square regression equation was derived to compensate for these effects after it has been shown that results obtained with non-scatter-corrected data were not accurate, especially early in the fermentation when the biomass is still relatively low. Results that have been obtained with scatter-correction equation corresponded to the results obtained from reference

methods. It has been shown that if left uncorrected the scattering variation can seriously impede the NIR determination of oil in this particular process.

While monitoring bioprocesses, Pasteur et *al.*[17], focused their work on a problem associated with light absorption and diffusion of bacterial concentration and its influence on the NIR calibration method. The main reason to initiate the research was the fact that biomass produces multiplicative and additive effects on the NIR spectral measurements. These effects influence the prediction accuracy of concentrations of the analytes in the process under investigation - acetate, butyrate, sucrose, and biomass. In order to improve the prediction accuracy, the authors applied pre - treatment techniques (second derivative and Standard Normal Variate (SNV)). As a result, the best PLS regression method was obtained with SNV correction applied on the transmittance values.

A work for monitoring of acetate, ammonium, biomass, and glycerol was reported by Hall et *al.*[18], where NIR spectroscopy in an industrial Escherichia coli fed-batch fermentation bioprocess was considered. Good models for acetate and glycerol were obtained using multiple linear least-squares regression, whereas ammonium and biomass models needed the use of the more sophisticated partial least-square regression technique.

Tosi et *al.*[19], monitored biomass, glucose, lactic acid and acetic acid in fermentation of *Staphylococcus xylosus*, by a fiber optic probe immersed into the culture and connected to a Near-Infrared instrument. The range of wavelength considered was from 700nm to 1800nm and batch, repeated batch, and continuous fermentations were studied. Spectral data were collected from three different fermentation runs of *Staphylococcus xylosus*, *Lactobacillus fermentum* and *Streptococcus thermophylus* growing in the same medium and under the same conditions. While building the PLS model, raw near-infrared spectra, showing broad absorption bands were processed with second order derivative in order to

reduce baseline offsets introduced by changes in the scattering properties of the sample, due to progressive increase of the cell mass. A unique calibration model was developed for *Staphylococcus xylosus* with successful application to the other two strains.

Cavinato et *al.* have demonstrated monitoring of ethanol concentration[20] and cell density[21] in the short wavelength NIR range (700nm-1100nm) in a fermentation process. In both studies a fiber optic probe was placed on the outside wall of glass fermentation vessel to avoid problems with probe sterilization and in order to attain a more representative sampling volume. Spectral data were collected in diffuse reflectance mode and a PLS model for prediction of ethanol concentration was developed. It was discussed that light scattering effects were affecting the baseline of the spectra. In order to correct for baseline offsets due to increasing biomass, the authors applied second order derivative transformation. A one wavelength (905nm) model for ethanol prediction has been successfully built with standard error of prediction from cross validation 0.27%. While considering cell density, these authors found a reasonably reproducible relation between diffuse reflectance measurements and cell mass concentration. However, prediction ability at higher concentrations of biomass (>40g/L) has been reported to decline. Examination of the residual estimates shows that much of the error appeared at high biomass, in the form of systematic deviation rather than a random one.

Rathore et *al.*[22] studied fermentation broth for a corn dry grind process aiming to determine suitable Near-Infrared wavelength regions for identification of ethanol, total soluble sugars, glycerol, and organic acids. Effects of temperature on spectral absorption were identified. Partial least square (PLS) and Principal Component Analysis (PCA) methods were applied to separate different fermentation constituents and to find suitable wavelengths for their prediction. To overcome shifting in the spectral baselines due to suspended particles and for better accuracy of the prediction models, the fermentation broth was filtered through 0.2μm filter and cells were removed.

Hagman and Sivertsson[23] used NIR spectroscopy to extract quantification information for glucose, lactate, ammonia, lactic acid, and biomass in the region 900nm-1700nm, while monitoring mammalian cell culture. Prediction using calibration models built from NIR measurements components were found to be in good agreement with values measured with conventional methods.

A novel science based approach that compensates for scatter effects in reflectance spectroscopy has been introduced by Kessler et *al.*[24], while studying active pharmaceutical ingredients (API) in tablets. First, absorption and scattering fingerprints were separated for a single tablet. Next, the two spectra were introduced in an alternating least square algorithm (ALS) in a multivariate curve resolution (MCR). As a result from the new approach, when compared with PLS with EMSC, it was found that the hard model constrained MCR-ALS algorithm can yield better predictions for concentration of API.

NIR spectroscopy measurements were collected in absorbance mode and later applied for the determination of very small concentrations (0.2-0.45%w/w) of API in an aqueous media[25]. The selected interval was between 4000-11000cm$^{-1}$ (909-2500nm). Rodrigues *et al.* used second derivative Savitzky-Golay filtering method for the preprocessing of the raw spectral data. The constructed PLS predictive model was robustly developed ($R^2$=0.99) and as a result was successfully transferred and used in different processes.

Rosa *et al.*[26] achieved determination of ranitidine in pharmaceutical products (granulates, cores, and coated tablets) using NIR reflectance measurements The spectra were collected in diffuse reflectance mode with a fiber optic probe in the region of 4000-12000cm$^{-1}$ (833-2500nm). The method was developed in the pharmaceutical industry, with the aim of replacing the conventional reference method.

Hammond[27] reported for the application of spectroscopy techniques in an antibiotic fermentation in industrial plant of Pfizer at Sandwich. Robust multi-component analysis system for at-line and on-line monitoring and control were developed, however details related with analytical data and application of techniques were not included in the paper due to the confidentiality of the work.

The use of NIR diffuse reflectance spectroscopy was described by Rosa *et al.* [28]as a tool for the qualification of *Ginkgo biloba* extract. This application deals with analysis of raw material rather then the pharmaceutical process itself. However it will help in the quality control of the entire process indirectly. The authors reported that variability in some *Ginkgo biloba* propeties have great impact on the granulation process and cause compression issues during the tablet manufacturing. The raw spectral data were collected in diffuse reflectance mode with a fiber optic probe. The region under consideration was in the range 4000-12000cm$^{-1}$(833-2500nm) wavenumbers. It was found that within the investigated parameters, dispersion in water and moisture content can be used successfully for the qualification of the *Ginkgo biloba* extract.

Galante et *al.*[29] proposed NIR light scattering as a noninvasive and nondestructive method for detection of small quantities of microorganisms in liquid pharmaceuticals. They collected spectral data from samples with contamination and samples without contamination. From analysis of this data, it was concluded that changes in the NIR spectra correlate to organism growth and that the spectra also appeared to distinguish between samples contaminated with different types of microorganisms. Further, it was also suggested that optimization is necessary in the sampling procedure for better results, such as configuration of the optical probe, selected samples, and wavelength range scanned. The method that was presented allows prevention of possible contamination that could follow as a result during the application of other analytical methods. It could be

used as an alternative to the conventional microbiological tests where large quantity of cells has to be identified and quantified in a relatively short period of time.

Gonzales-Saiz et *al.*[30], proposed a combination of NIR measurements and multivariate calibration for monitoring of key parameters in an alcohol fermentation process from onion juice. Although several preprocessing techniques were applied to the raw spectral data-including mean-centering, autoscaling, first and second derivatives, orthogonal signal correction, multiplicative signal correction, and standard normal variate - in this case study, the best possible results for all modeled physicochemical variables were obtained without the application of complex pretreatment techniques. Biomass results were reported as very successful (3.04%RMSEC and 4.62%RMSEP) taking into account the existing difficulties when measuring biological variables in a complex system. Also a calibration model for ethanol concentration was successfully built. This research demonstrated the applicability of NIR for real-time monitoring of crucial parameters in alcohol fermentation of onion juice, in spite of the high complexity of the system studied.

The monitoring of grape and wine by NIR spectroscopy is increasingly being studied nowadays. A number of scientific groups have shown NIR as a successful tool for monitoring of such processes. Cozzolino et *al.*[14], published a review article highlighting the most recent applications of the technique for analysis of the composition of grape and wine. The same research group reported recently the application of NIR spectroscopy for monitoring red wine fermentation on a pilot plant scale[31]. They found that while utilizing NIR it was possible to detect changes during fermentation as well as classification of different stages of the fermentation process.

Ferreira et *al.*[32] reported the application of NIR spectroscopy to an industrial bioprocess for the production of clavulanic acid in a complex media where diffuse reflectance spectral data were collected. Two variables selection methods were applied

and results showed that calibration models developed after variable selection have superior predictive ability compared to models developed with no variable selection.

A significant amount of research work in the application of NIR spectroscopy to fermentation has been done by the group of McNeil and Harvey at the Strathclyde fermentation center in Glasgow. Focus has been on addressing the challenges in the application of NIR technique in biotechnology, reviewing the existing approaches, the instrumentation options, and the possibilities for data interpretation and analysis[33]. Comparisons between microbial and mammalian cultivations were presented in terms of the application of NIR spectroscopy [34], with special attention given to animal cell culture and future possibilities for NIR development for controlling and monitoring of those pharmaceutical processes. Practical aspects for spectral method development for fermentation monitoring and control were also discussed in great detail[35], followed shortly by an article presenting the implementation strategies of robust NIR analysis. The work of this group has been particularly focused on various bioprocesses which is summarized in the following paragraphs.

Near-Infrared spectral data of microbial biomass were extracted from different microorganisms: three of them were filamentous fungi, one a filamentous Gram positive bacterium, and Gram negative unicellular bacterium. A detailed study of *Penicillium chrysogenum* biomass was also carried out[36]. While identifying spectral regions useful for the prediction of biomass, influence of scattering and absorption has been found, narrowing the choice of wavelengths to a few regions: 1576nm-1762nm, and 2054nm-2348nm.

Experimental monitoring of culture fluid from a submerged fungal bioprocess was performed, investigating the possibility of implementing Near-Infrared spectroscopy for prediction of concentration of mycelial biomass and other key analytes[37]. The mycelial

culture consisted of filamentous fungus *P. chrysogenum* which produces the antibiotic penicillin G. *P. chrysogenum*. Due to the filamentous nature of the culture, spectral variations were expected originating from the multiple morphological forms of the biomass, such as dispersed hyphae, clumps and aggregates. To minimize the scattering effects on the data analysis, the authors used second derivative spectra for model building.

Another antibiotic production process has been assessed by the same group[38]. The mycelial microorganism *Streptomyces fradiae* was cultivated for the production of tylosin, and NIR measurements were made at-line in transmittance and reflectance mode over the entire Vis-NIR range of 400-2500nm. Considering the same filamentous bacterial cultivation, models for prediction of four key analytes (methyl oleate, glucose, glutamate and ammonium) were built successfully[10] as well as temporally segmented modeling has been applied in order to monitor the product concentration of the process[39]. The influence of morphology of *Streptomyces fradiae* on the Near-Infrared spectra has been also investigated[40].

Cell mass monitoring through at-line and in-situ Near-Infrared spectroscopy has been applied in an industrial fed-batch *Escherichia coli* process[41] and close examination of the relationship between at-line and in-situ biomass models were presented. The use of Near-Infrared spectroscopy was also investigated in a high cell density fed-batch industrial *Pichia pastoris* bioprocess for the production of a therapeutic protein[42], and for the monitoring of biomass, glucose, ethanol and protein content in a high cell density baker's yeast fed-batch process[43].

And finally, NIR spectroscopy has been used for the prediction of mycelial biomass, total sugars, and ammonium in a *Penicillium chrysogenum* bioprocess[44]. The robustness of the predictive ability and the structure of the models were tested and assessed by

artificially introducing analyte and background matrix variations.

The aim of the literature review presented, was to demonstrate the attractions of NIR spectroscopy as a non destructive and non invasive analytical technique that has been applied for the prediction of the concentrations of key components and products from different bioprocesses. Implementation of NIR for monitoring of various bioprocesses mentioned earlier, and studies for the last 30 years have been introduced in this chapter.

Overall, observations indicate that NIR spectroscopy can improve fermentation process operation by providing rapid, non-destructive, multi-constituent analyses of the fermentation broth and media. Experimental data from original spectra need to be pre-processed, due to non-chemical spectral variations and baseline shifts that affect the performance of the NIR analysis. Fermentation broths are usually turbid suspensions with changing concentrations of the main components. With the increase of biomass, for example, the effective pathlength of the photons change dramatically, due to the increasing amount of cells that scatter light. Bubbles are also present in many industrial processes and they act as an additional source for scattering. Scattering of light, temperature, pathlength variations, particles size, various kinds of noise, are all potential features affecting the performance of the NIR measurement. Unfortunately, the majority of the methods used, only reduce or remove scattering form the sample spectra, but actually scattering itself contains information regarding the size and shape of the particles and can be used to improve existing methods[24].

From this perspective, the main focus of the research is to study multiple scattering of light from biological suspensions and to investigate associated with this phenomena changes influencing the raw spectra. Multiple scattering of light by the cells poses a significant challenge in the development of NIR-based methodologies to extract reliable chemical and physical information contained in the spectra of growing suspension of

microbial cells.

## 2.4. Multiple scattering effects

When light hits one or a group of particles certain phenomena can be observed.
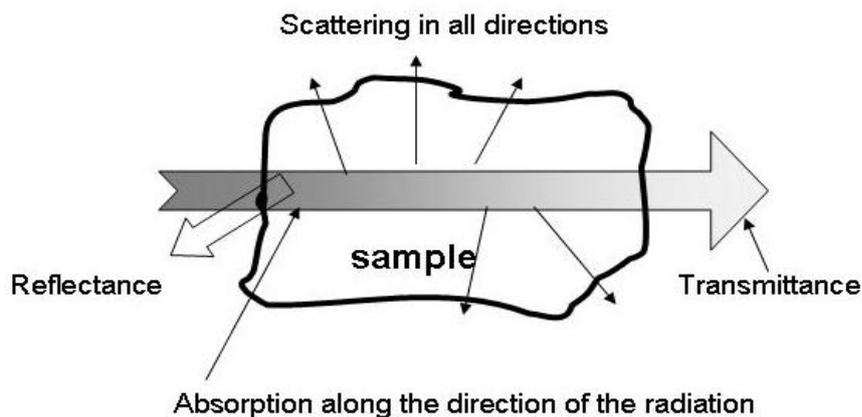


Fig.2.2 Optical phenomena from collection of particles.

Figure 2.2 is a schematic presentation of those phenomena that are considered important for the present work. Reflectance, transmittance, absorption and scattering, all are results of the radiation - matter interaction. Reflectance exists in two forms: diffuse and specular. Reflectance that undergoes scattering is often called diffuse reflectance and unscattered reflectance is called specular (mirror-like) reflectance.

The optical phenomena depend on the chemical composition of the particles, size, shape, orientation, the surrounding medium, the number of particles, and the frequency of the incident beam[45].

Radiation undergoes single scattering when it is scattered by one particle or group of particles with certain distance between them. Multiple scattering is associated with radiation that is scattered many times between groups of particles as shown in figure 2.3.
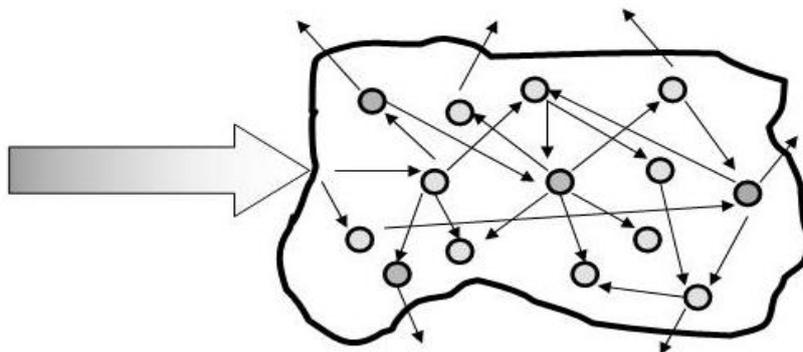
Fig.2.3 Multiple scattering of light from collection of particles.

Multiple scattering of light mathematically has been described by the Radiative Transfer Theory (RTT) in the work of Chandrasekhar[46] and Ishimaru[47]. Light propagation in turbid medium is expressed by the Radiative Transfer Equation (RTE), which is derived from the RTT. It is valid for material with small optical inhomogeneities, uniformly distributed through the medium. According to the assumptions for the equation, every sample is a homogeneous matrix represented by three parameters: bulk scattering coefficient $\mu_s$, bulk absorption coefficient $\mu_a$ and the scattering anisotropy factor $g$. There are no analytical solutions for the transfer equation, and in this case approximations and/or numerical techniques must be developed and utilised in order to obtain results. More detailed information on the RTE and methodology of this work are presented in the next chapter.

## 2. 5. Discussion

Key components, classical techniques for analysis, and NIR spectroscopic application for bioprocess monitoring have been discussed in this chapter. A number of issues associated with NIR application to bioprocesses have been briefly discussed including pathlength variations, light scattering effect, and noise in the raw spectral data. The application of

chemometric techniques is required for NIR analysis because the spectra contain multivariate information due to overlap of chemical absorption bands in the NIR region, and is influenced by the physical characteristics too. First and second derivatives of spectra are used to improve the resolution of overlapping bands and to reduce baseline shifts, in combination with smoothing techniques, such as Taylor and Savitzky-Golay filtering. Corrections for light scattering effects, for pathlength differences and baseline effects are usually made through Multiplicative Scatter Correction (MSC) and Standard Normal Variate transformation (SNV). Multiple scattering is usually related with the increase of biomass. This has been briefly introduced, followed by a comment on the Radiative Transfer Equation as an approach to calculate the scattering and absorption effects from the spectra, and also the possibility to collect both physical and chemical spectral information.

In Chapter 3, the Radiative Transfer Equation and the methodology based on this equation to extract information regarding a growing bacterial culture will be discussed.

## References

1.	Scarf, M., A. Arnold, L. Harvey, and B. McNeil, *Near Infrared Spectroscopy for Bioprocess Monitoring and Control: Current Status and Future Trends.* Critical Reviews in Biotechnology, 2006. **26**(1): p. 17-39.

2.	Brimmer, P.J. and J.W. Hall, *Determination of nutrient levels in a bioprocess using near-infrared spectroscopy.* Canadian Journal of Applied Spectroscopy, 1993. **38**(6): p. 155-162.

3.	Gunsalus, I. and C. Niven, *The effect of pH on the lactic acid fermentation.* The Journal of Biological Chemistry, 1942. **145**(1): p. 131-136.

4.	Ward, O., *Fermentation Biotechnology: principles,processes, products.* 1989, Milton Keynes: Open University Press. 227.

5.	Ratledge, C. and B. Kristiansen, *Basic Biotechnology*. Third ed, ed. C.U. Press. 2006, New York.

6.	Omstead, D., *Computer Control of Fermentation Processes*. 1989: CRC Press. 343.

7.	Umoh, E.F., A.B. vonPutten, and K. Schugerl, *Simultaneous on-line monitoring of glucose and total malto sugar in fermentation processes using an FIA system.* Journal of Chemical Technology and Biotechnology, 1996. **67**(3): p. 276-280.

8.	Burton, Z. and J. Kaguni, *Experiments in Molecular Biology: Biochemical Applications*. 1997, San Diego, California: Academic Press. 223.

9.	Turner, A.P.F., I. Karube, and G.S. Wilson, *Biosensors: Fundamentals and Applications*. 1987, Oxford: Oxford University Press.

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

11.	Vaidyanathan, S., G. Macaloney, J. Vaughn, B. McNeil, and L. Harvey, *Monitoring of Submerged Bioprocesses.* Critical Reviews in Biotechnology, 1999. **19**(4): p. 277-316.

12.	Vaidyanathan, S., S.A. Arnold, L. Matheson, P. Mohan, B. McNeil, and L.M. Harvey, *Assessment of near-infrared spectral information for rapid monitoring of bioprocess quality.* Biotechnology and Bioengineering, 2001. **74**(5): p. 376-388.

13.	Naes, T., T. Isaksson, T. Fearn, and T. Davies, *A user-friendly guide to multivariate calibration and classification*. 2004: NIR Publications.

14.	Cozzolino, D., R. Dambergs, L. Janik, W. Cynkar, and M. Gishen, *Analysis of Grapes and Wine by Near Infrared Spectroscopy* Journal of Near Infrared Spectroscopy, 2006. **14**(5): p. 279-289.

15.	Engelhard, S., M.U. Kumke, and H.G. Lohmannsroben, *Examples of the application of optical process and quality sensing (OPQS) to beer brewing and polyurethane foaming processes.* Analytical and Bioanalytical Chemistry, 2006. **384**(5): p. 1107-1112.

16.	*Process Analytical Technology (PAT) Initiative*. 2004, US Food and Drug Administration: http://www.fda.gov/Cder/OPS/PAT.htm.

17.	Lesteur, M., E. Latrille, J. Roger, V. Bellon-Maurel, C. Gonzales, G. Junqua, and J. Steyer, *A chemometric approach to minimise the diffusive effect of the biomass when using near infrared spectroscopic measurement for the monitoring of bioprocesses*, in *11th International Conference on Chemometrics for Analytical Chemistry*. 2008: Montpellier, France.

18.	Hall, J., B. McNeil, M. Rollins, I. Draper, B. Thompson, and G. Macaloney, *Near-Infrared spectroscopic determination of acetate, ammonium, biomass, and glycerol in an industrial Escherichia coli fermentation.* Applied Spectroscopy, 1996. **50**(1): p. 102-108.

19. Tosi, S., M. Rossi, E. Tamburini, G. Vaccari, A. Amaretti, and D. Matteuzzi, *Assessment of in-line near-infrared spectroscopy for continuous monitoring of fermentation processes.* Biotechnology Progress, 2003. **19**(6): p. 1816-1821.

20. Cavinato, A., D. Mayes, Z. Ge, and J. Callis, *Noninvasive method for monitoring ethanol in fermentation processes using fiber-optic Near-Infrared spectroscopy.* Analytical Chemistry, 1990. **62**: p. 1977-1982.

21. Ge, Z., A. Cavinato, and J. Callis, *Noninvasive spectroscopy for monitoring cell density in a fermentation process.* Analytical Chemistry, 1994. **66**: p. 1354-1362.

22. Rathore, S., M. Paulsen, V. Sharma, and V. Singh. *Near-Infrared Spectra Interpretation of Fermentation in a Dry Grind Corn Process.* in *ASAE Annual Internationa Meeting.* 2005. Tampa Convention Center, Tampa, Florida.

23. Hagman, A. and P. Sivertsson, *The use of NIR spectroscopy in monitoring and controlling bioprocesses.* Process Control and Quality, 1998. **11**(2): p. 125-128.

24. Kessler, W., D. Oelkrug, and R. Kessler, *Using scattering and absorption spectra as MCR-hard model constratints for diffuse reflectance measurements of tablets.* Analytica Chimica Acta, 2009. **642**: p. 127-134.

25. Rodrigues, L.O., J.P. Cardoso, and J.C. Menezes, *Applying Near-Infrared Spectrsocopy in downstream processing: one calibration for multiple clarification processes of fermentation media.* Biotechnology Progress, 2008. **24**: p. 432-435.

26. Rosa, S.S., P.A. Barata, J.M. Martins, and J.C. Menezes, *Development and validation of a method for active drug identification and content determination of ranitidine in pharmaceutical products using near-infrared reflectance spectroscopy: a pharmaceutical release approach.* Talanta, 2008. **75**: p. 725-733.

27. Hammond, S., *NIR analysis of antibiotic fermentations in Making light work: advances in Near Infrared spectroscopy.* 1992, New York: VCH Publishers. 584-589.

28. Rosa, S.S., P.A. Barata, J.M. Martins, and J.C. Menezes, *Near-infrared reflectance spectroscopy as a process analytical technology tool in Ginkgo biloba extract qualification.* Journal of pharmaceutical and biomedical analysis, 2008. **47**: p. 320-327.

29. Galante, L.J., M.A. Brinkley, J.K. Drennen, and R.A. Lodder, *Near-Infrared Spectrometry of Microorganisms in Liquid Pharmaceuticals.* Analytical Chemistry, 1990. **62**(23): p. 2514-2521.

30. Gonzales-Saiz, J.M., C. Pizarro, I. Esteban-Diez, O. Ramirez, C.J. Gonzales-Navarro, M.J. Saiz-Abajo, and R. Itoiz, *Monitoring of alcoholic fermentation of onion juice by NIR spectroscopy: valirization of worthless onions.* Journal of Agricultural and Food Chemistry, 2007. **55**: p. 2930-2936.

31. Cozzolino, D., M. Parker, R. Dambergs, M. Herderich, and M. Gishen, *Chemometrics and visible-near infrared spectroscopic monitoring of red wine fermentation in a pilot scale.* Biotechnology and Bioengineering, 2006. **95**(6): p. 1101-1107.

32. Ferreira, A., T. Alves, and J. Menezes, *Monitoring complex media fermentations with near-infrared spectroscopy: comparison of different variable selection methods.* Biotechnology and Bioengineering, 2005. **91**(4): p. 474-479.

33. Vaidyanathan, S. and B. McNeil, *Near Infrared Spectroscopy - a panacea in pharmaceutical bioprocessing?* European Pharmaceutical Review, 1998. **3**(2): p. 43-48.

34. Crowley, J., A. Arnold, L. Harvey, and B. McNeil, *Near infrared spectroscopy-the key to improved biopharmaceutical manufacturing from animal cell culture?* European Pharmaceutical Review, 2000. **5**(3): p. 39-42.

35. Arnold, S.A., L.M. Harvey, B. McNeil, and J.W. Hall, *Employing near-infrared spectroscopic methods and analysis for fermentation monitoring and control -*

*Part 1, method development.* Biopharm International-the Applied Technologies of Biopharmaceutical Development, 2002. **15**(11): p. 26-34.

36. Vaidyanathan, S., G. Macaloney, and B. McNeil, *Fundamental investigations on the near-infrared spectra of microbial biomass as applicable to bioprocess monitoring.* The Analyst, 1999. **124**: p. 157-162.

37. Vaidyanathan, S., L. Harvey, and B. McNeil, *Deconvolution of Near-Infrared Spectral Information for Monitoring Micelial Biomass and other Key Analytes in a Submerged Fungal Bioprocess.* Analytica Chimica Acta, 2001. **428**: p. 41-59.

38. Vaidyanathan, S., A. Arnold, L. Matheson, P. Mohan, G. Macaloney, B. McNeil, and L.M. Harvey, *Critical evaluation of models developed for monitoring an industrial submerged bioprocess for antibiotic production using near-infrared spectroscopy.* Biotechnology Progress, 2000. **16**(6): p. 1098-1105.

39. Arnold, S.A., L. Matheson, L.M. Harvey, and B. McNeil, *Temporally segmented modelling: a route to improved bioprocess monitoring using near infrared spectroscopy?* Biotechnology Letters, 2001. **23**(2): p. 143-147.

40. Vaidyanathan, S., S. White, L. Harvey, and B. McNeil, *Influence of morphology on the Near-Infrared spectra of mycelial biomass and its implications in bioprocess monitoring.* Biotechnology and Bioengineering, 2003. **82**(6): p. 715-724.

41. Arnold, S.A., R. Gaensakoo, L.M. Harvey, and B. McNeil, *Use of at-line and in-situ near-infrared spectroscopy to monitor biomass in an industrial fed-batch Escherichia coli process.* Biotechnology and Bioengineering, 2002. **80**(4): p. 405-413.

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

43. Finn, B., L. Harvey, and B. McNeil, *Near-Infrared spectroscopic monitoring of biomass, glucose, ethanol and protein content in a high cell density baker's yeast fed-batch bioprocess.* Yeast, 2006. **23**: p. 507-517.

44. Vaidyanathan, S., G. Macaloney, L. Harvey, and B. McNeil, *Assessment of the structure and predictive ability of models developed for monitoring key analytes in a submerged fungal bioprocess using Near-Infrared spectroscopy.* Applied Spectroscopy, 2001. **55**(4): p. 444-453.

45. Bohren, C.F., D.R.Huffman, *Absorption and Scattering of Light by Small Particles.* 1983, New York: John Wiley&Sons.

46. Chandrasekhar, S., *Radiative Transfer.* 1960: London and New York and Dover, Oxford University Press.

47. Ishimaru, A., *Wave propagation and scattering in random media.* 1978, New York: Academic Press.

# Determination of optical properties of biological suspensions using the Radiative Transfer Equation

## 3.1. Introduction

This research proposes a measurement based approach for separating absorption from scattering effects through the rigorous multiple scattering theory. For modeling of multiple scattering in turbid media, Radiative Transfer Theory (RTT) and in particular Radiative Transfer Equation (RTE) are applied. This chapter provides a description of the RTE, followed by the methodology for extracting absorption and scattering properties based on the RTE.

## 3.2. Radiative Transfer Equation

Two distinct theories describe the light propagation in turbid media: (1) the electromagnetic wave theory associated with the solution of Maxwell equations[1-4], and (2) the Radiative Transfer Theory where the light is associated with a propagating energy through a particulate media (turbid media)[1, 4, 5]. In this research, the main focus will be on the use of transfer theory, in particular the Radiative Transfer Equation (RTE). The advantage of the transfer theory lies in its relative computational simplicity and is widely used to model multiple scattering in turbid media. In this section of the chapter fundamental optical properties associated with the Radiative Transfer Equation will be introduced. Accordingly, these main quantities characterize absorption and scattering properties of particulate media, and the extraction of these properties directly addresses issues relating to the multiple scattering of light associated with NIR spectroscopy techniques. Figure 3.1 represents interpretation of scattering of specific intensity and represents schematically the terms of the RTE ([4] - p.156, section 7.3).
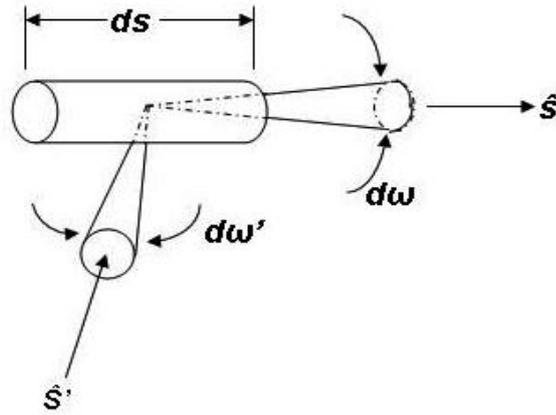
Fig. 3.1 Scattering of specific intensity incident upon the volume *ds* from the direction *ŝ'* into the direction *ŝ*.

The general equation of transfer which governs the radiation field in a medium that absorbs, emits, and scatters radiation is given by [4],

$$\frac{dI(r,\hat{s})}{ds} = -(\mu_a + \mu_s)I(r,\hat{s}) + \frac{\mu_s}{4\pi}\int_{4\pi} p(\hat{s},\hat{s}')I(r,\hat{s}')d\omega' + \varepsilon(r,\hat{s}) \qquad (3.1)$$

where, *I(r, ŝ)* is the specific intensity at a point *r* with radiation incident along direction *ŝ*, $(\mu_s + \mu_a) = \mu_t$ is the bulk extinction coefficient, $\mu_s$ is the scattering coefficient and $\mu_a$ is the absorption coefficient; *p(ŝ, ŝ')* is the phase function, which is a measure of the angular distribution of the scattered light, and *ε(r,ŝ)* is the source term. In the present work, the system under consideration does not emit energy and therefore the source term *(ε)* is zero and equation (3.1) can be rewritten as follows,

$$\frac{dI(r,\hat{s})}{ds} = -(\mu_a + \mu_s)I(r,\hat{s}) + \frac{\mu_s}{4\pi}\int_{4\pi} p(\hat{s},\hat{s}')I(r,\hat{s}')d\omega' \qquad (3.2)$$

For describing the light propagation through a turbid media using equation (3.2), the following parameters also known as bulk optical properties [2] are required: the bulk

scattering coefficient ($\mu_s$), the bulk absorption coefficient ($\mu_a$), and the anisotropy factor ($g$). The latter is the average cosine of the scattering phase function *(p(ŝ,ŝ'))*.

The phase function *(p(ŝ,ŝ'))* is a dimensionless parameter and its integral over all direction is 1 (eq.3.3).

$$\int_{4\pi} p(\hat{s}, \hat{s}')d\omega' = 1$$

(3.3)

The phase function depends on the size and shape of the particles, usually its form is not known, and is often expressed in terms of the cosine of the angle $\theta$ ($\theta$ is the angle between given direction defined by unit vector $\hat{s}$ and the normal to the surface $\hat{s}_o$, Appendix A, figure A.1), characterized by a single parameter $g$ *(‹ cosθ ›)*, eq. 3.4. This parameter is often referred to as the average cosine of the phase function or as anisotropy factor,

$$g = \langle \cos\theta \rangle = \int_{4\pi} p(\hat{s}, \hat{s}')(\hat{s} \cdot \hat{s}_O)d\omega'$$

(3.4)

$$\hat{s} \cdot \hat{s}_O = \cos\theta$$

(3.5)

The anisotropy factor *(g)* is a measure of the asymmetry of the single scattering pattern. For a particle that scatters light, anisotropy factor *(g)* could vary between 1 (extremely forward scattering, $\theta=0^o$), 0 (isotropic scattering-the same in all directions), -1 (highly backward scattering, $\theta=180^o$), and its value depends on the particle size and size distribution.

The next two parameters required for description of light propagation in turbid media, the bulk absorption and bulk scattering coefficients contain all the information for the individual species present in a sample and can be written as,

$$\mu_a(\lambda) = \Sigma \rho_i \sigma_{a,i}(\lambda)$$

$$\mu_s(\lambda) = \Sigma \rho_i \sigma_{s,i}(\lambda)$$

(3.6)

where $\rho$ is the number density or concentration, $\sigma_s$ and $\sigma_a$ are the scattering and absorption cross-section respectively. The bulk absorption/scattering coefficients are functions of the concentration and the absorption/scattering cross-sections respectively of the species present in the sample, at wavelength $\lambda$ of the incident beam. The absorption and scattering cross-sections have dimension of area and they are functions of the orientation of the particle and the state of polarization of the incident light.

In the context of the RTT, the bulk absorption coefficient $\mu_a$ is a quantity that characterizes the amount of the absorbed energy passing through a sample containing absorbing material, and is measured in inverse length (e.g. $mm^{-1}$). Absorption is strongly wavelength dependent, owing to the fact that different atoms and molecules absorb light at different wavelengths. The measured absorption coefficient can be used as an identifier, or fingerprint of a substance, and varies with the concentration of the absorbing component.

The bulk scattering coefficient $\mu_s$ is the amount of scattered energy in all directions when light passes trough a sample, and is expressed in inverse length (e.g. $mm^{-1}$). It is wavelength dependent and is a function of the shape, size and concentration of the particles.

The three parameters, bulk absorption coefficient, bulk scattering coefficient and anisotropy factor can be extracted by inverting the RTE given a set of appropriate measurements at each wavelength. Through the years different methodologies have been developed to invert and solve the RTE and the next section will provide a description of this.

## 3.3. Solving the RTE

There is no analytical solution for the RTE, and it has to be solved numerically. Among the various applied numerical methods, the discrete ordinate, the adding-doubling method, and Monte Carlo method are the most usable when it is necessary to obtain an exact solution for the RTE.

The discrete ordinate was introduced by Chandrasekhar[5] where the discretization of the RTE results in a system of linear equations which is usually solved iteratively, and this method has been mainly used in astrophysical applications. The method is accurate and fast, however with very limited success in media with large optical depth and/or a high degree of scattering. Optical depth (optical thickness) is a measure of transparency and is defined as the negative logarithm of the fraction of radiation (or light) that is not scattered or absorbed on a path.

### 3.3.1. Monte Carlo Method

Monte Carlo method is a technique that can be used to solve the RTE by considering the propagation of individual photons (or photon packets) through the media in a probabilistic way. In this method, a single photon (or photon package) is launched into the media, where the photon trajectory is simulated until the photon is absorbed in the media or leaves it completely. The distribution of light absorption can be then calculated with great accuracy by simulating a large number of photons, however if higher accuracy is needed this will lead to increase in the computation time. The techniques has been applied by several authors[6-9]. Inverse Monte Carlo (IMC) methods can be used to determine the bulk absorption coefficient, bulk scattering coefficient and anisotropy factor over any spectral range, without simplification that might limit their accuracy. Due to its ability to compensate for different measurements and sample geometries, and to account for sources of experimental error such as light lost out the edges of the sample,

IMC is regarded as robust and accurate method for optical properties determination, and it has been widely used in biomedical applications to obtain optical properties of tissue [10-12]. Once again, large computation times are needed in order to acquire adequate signal-to-noise ratio from the Monte Carlo simulations and to ensure proper convergence at each wavelength within the spectral range of interest.

### 3.3.2. Adding-Doubling Method

In this work, the adding-doubling method was selected to solve the RTE because this method has proven to be much less computationally intensive than the MC method and much more stable than the discrete ordinates method.

Van de Hulst[1] introduced the adding-doubling (ADD) method, applicable to plane-parallel geometries and provided a numerical solution of the RTE. The advantage of ADD consists in the accuracy of the solution that can be determined for any combination of optical properties ($\mu_a$, $\mu_s$, and $g$). The method also takes into account anisotropic scattering, arbitrary thick samples and Fresnel boundary conditions for internal reflection, and it has relatively fast computation time.

Layered geometry (plane-parallel) and uniform irradiation are amongst the main limitations, in addition to the condition for homogeneous optical properties at each layer and the fact that the method does not account for light lost at the edges of the sample.

The ADD has been used mainly to solve the RTE in the atmospheric sciences [13], and later extended by Prahl [14-16] in the area of biomedical applications where optical properties from tissue samples were successfully determined under the condition of multiple scattering. In the biomedical applications, the focus was the inversion of the RTE and the method of inversion which used the ADD for solving the RTE is referred to as the inverse adding-doubling method (IAD) [16].

The IAD technique is an iterative method in which initially the values of optical properties are assumed and the reflectance and transmittance for a thin slab are calculated; the slab thickness is then doubled by adding two slabs of the same thickness for which the reflectance and transmittance have identical values. By juxtaposing these two slabs next to each other it is possible to find the total values for reflectance and transmittance from the entire slab with the new doubled thickness. The process continues in this manner until the desired thickness of sample is reached and thus the reflectance and transmittance properties of the entire slab can be calculated. If measurements for reflectance and transmittance are available then the next step is to compare the calculated and measured values. If the total difference between them is within certain interval, values of the optical properties are no longer iterated. The iteration stops when the calculated and measured values for reflectance and transmittance match to within a preset error.

In order to extract the optical properties ($\mu_a$, $\mu_s$ and $g$) from a known sample through the IAD method, it is evident that measurements for reflectance and transmittance need to be available. In practice different measurement techniques are applied and the different options available are introduced below.

## 3.4. Inversion of the RTE using the IAD method

### 3.4.1. Measurements required

There are a number of sophisticated measurement techniques allowing one, two or three parameters to be determined. Parameters refer to the three optical properties, namely, bulk absorption coefficient, bulk scattering coefficient, and the anisotropy factor or a combination of them. Different measurements are usually required in order to obtain these optical properties of interest, and amongst them three main cases can be distinguished: (1) measurement techniques where one of the optical properties can be

directly derived, (2) measurement techniques for derivation of two parameters (two of the optical properties), and measurement techniques for determination of all three optical properties from one sample.

One parameter techniques are related to a group of measurement procedures where one of the optical tissue parameters can be directly derived without any extensive computation or simulations, provided that the measurements are performed on samples that are sufficiently thin so that only single scattering occurs. By measuring only collimated (un-scattered) transmittance (Tc), the total extinction coefficient ($\mu_{t\ =\ \mu_a\ +\ \mu_s}$) can be calculated knowing the thickness of the sample [17, 18]. The scattering and absorption coefficients can also be independently from each other determined, by placing thin samples in the center of an integrating sphere. This case has been experimentally demonstrated by Marchesini [18] and Nelson [19]. The integrating sphere setup will be discussed in greater detail in section 3.5.

The anisotropy factor ($g$) (the third optical property) can be obtained by making angular scattering measurements with a goniometer. Steinke and Shepherd have applied this technique while calculating the anisotropy factor of dilute blood suspensions [20].

Techniques for determining two parameters involve the determination of the following parameters: the reduced scattering coefficient ($\mu_s' = \mu_s(1 - g)$), and the bulk absorption coefficient ($\mu_a$). In this case knowledge of diffuse reflectance and diffuse transmittance is sufficient to determine $\mu_s'$ and $\mu_a$. Spatially-resolved[21] and time-resolved techniques[22] are mostly applied for the determination of the two parameters.

Spatially-resolved measurements technique is based on a reflectance measurement where the sample is illuminated by a continuous light source in the form of a focused spot and the diffuse (scattered) reflectance is recorded at different radial distances.

In time-resolved measurements the source of light is a pulsed laser. After the pulse has propagated through one part of the sample a measurement of its dispersion is conducted.

However, both the spatially and the time-resolved methods could provide information for only two parameters as already mentioned earlier, namely, the reduced scattering coefficient and the bulk absorption coefficient. In fact, theoretically, these methods can be applied for the determination of the anisotropy factor *g* though in practice it is difficult to obtain reliable estimates of *g*. To determine the anisotropy factor and to separate $\mu_s$' into $\mu_s$ and *g*, a third measurement is necessary. Usually this is achieved by taking collimated transmittance measurements. The technique described above is applicable when integrating sphere setup is used for the measurements.

All three optical properties can be obtained through an elegant technique, named the integrating sphere method[8, 16, 23, 24] where in addition to the diffuse reflectance and diffuse transmittance measurements, collimated (unscattered) transmittance is also incorporated as a measurement. An integrating sphere is designed to collect all light flux that enters into it by using a detector mounted on the sphere wall. The measurement setup is restricted however, to samples in the form of thin films or it requires special holder and glass cuvette for liquid or powder samples. The integrating sphere setup is used in this research for measuring total diffuse reflectance and total diffuse transmittance and will be discussed in more detail in section 3.5.

### 3.4.2. Inversion algorithm

This section includes an explanation of the algorithm that was used for inverting the RTE in order to obtain all three optical properties from the system studied.

The Radiative Transfer Equation (RTE) which describes the transport of light through a medium containing particles has been given earlier (3.1). Its main constituents were

underlined to be the bulk scattering coefficient ($\mu_s$ (mm$^{-1}$)), the bulk absorption coefficient ($\mu_a$ (mm$^{-1}$)), and the phase function ($p(\hat{s}, \hat{s}')$) which is a measure of the angular distribution of scattered light and is usually approximated as a function of the anisotropy factor $g$. In this study, the optical properties were extracted using an inversion technique based on the adding-doubling method for numerically solving the RTE [16, 25-28]. To use this approach at least 3 measurements have to be taken at each wavelength. The 3 measurements needed are as already mentioned, total diffuse transmittance (Td), total diffuse reflectance (Rd) and collimated transmittance (Tc). The first two are obtained using an integrating sphere setup while collimated transmittance is measured with standard transmittance accessory.

This research work applies the IAD method for extracting and monitoring optical properties and their variations during growth of *Bacillus subtilis* in a liquid culture. The IAD model was initially introduced by Prahl [16] for calculating optical properties of tissue, and its accuracy has been validated by several research groups[24, 25, 27] in the case of human skin study, as well as apple skin and flesh.

To invert the RTE for the system glass-suspension-glass using the adding-doubling method, Td, Rd and Tc measurements are required as inputs in addition to the refractive index of the glass cuvette and the refractive index of the suspension. The refractive index of the glass cuvette was taken from the manufacturer catalogue [29] while the refractive index of the sample was taken as the average value calculated after measuring the refractive index of several samples with refractometer (these values will be given in chapter 4).

Calculations were carried out using MATLAB 6.5 scientific software with an ADD algorithm already developed in MATLAB by Velazco-Roa and Thennadil[28]. A schematic illustration of the method used in this thesis can be seen in figure 3.2. The program predicts Rd, Td and Tc for a sample which has the guesstimate optical

properties $\mu_a$, $\mu_s$ and anisotropy factor $g$ using the adding-doubling method. In the next step the predicted values are compared with the measured values for the given sample. If the total difference between these values is less/equal than a specified tolerance (in this case study the tolerance is set to be $10^{-6}$), then the guessed optical parameters are accepted as the optical property values corresponding to the sample on which the measurements were made. If the total difference is greater than the specified tolerance, new guess values for $\mu_a$, $\mu_s$ and $g$ are chosen and the process continued until predicted and measured values match within the tolerance, giving the optical properties of the sample. This iterative process and the updating of the guesses are done through the use of the optimization routine "fmincon" available in the Matlab® Optimization Toolbox.
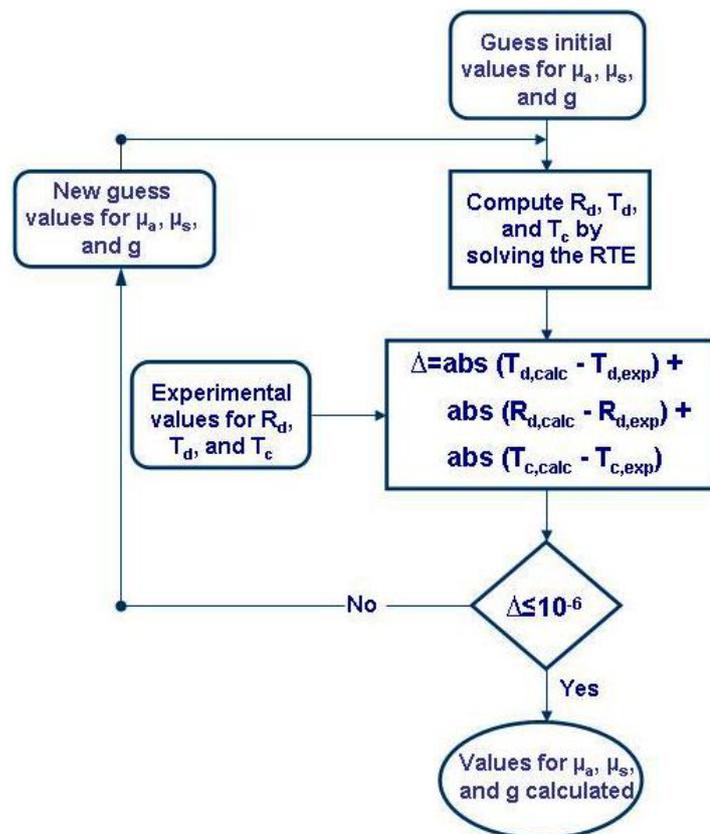


Fig.3.2. Schematic illustration of the algorithm to solve the RTE using the inverse adding-doubling method.

### 3.5. Experimental setup

This section describes the integrating sphere setup with emphasis on the measurement setup for the two main types of configurations: single and double integrating spheres. Description for obtaining measurements introduced in section 3.4.1 for determining the different optical parameters is provided, as well as understanding of instrumental artefacts and sources of error associated with the integrating sphere measurements. The experimental setup for total diffuse reflectance (Rd), total diffuse transmittance (Td) and collimated transmittance (Tc) measurements used in this study to obtain the optical properties of *Bacillus subtilis* growing suspension is also discussed.

#### 3.5.1. Measurements using integrating sphere

The integrating sphere is a simple device that has the ability to collect and measure optical radiation and this fact makes it applicable for a wide range of studies [30, 31]. The most used application, however, is for measurement of diffuse reflectance and diffuse transmittance of different turbid and scattering materials.

An integrating sphere is a closed sphere with few port openings, allowing (1) incoming light to enter, as well as (2) sample, (3) reference standard and (4) detector to be positioned on the sphere wall (figure 3.3). The inside part is coated with a material that should reflect 100% of the incident light in the ideal case.
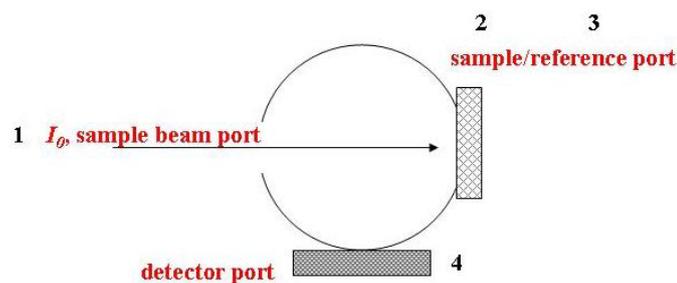


Fig.3.3 Single integrating sphere

Different sphere geometries are available and they differ from each other by the position of the studied sample and the reference standard. In a substitution sphere, one port is available which implies subsequent measurement of the sample and the reference standard. In this case a problem arises when the sample needs to be replaced by the standard. In this case the average reflectance is changed and high accuracy of sample reflectance is hard to be achieved. In order to solve the problem, spheres based on a comparison method have been modeled with ports available at the same time for sample and reference standard. In this way the average reflectance is maintained constant and the accuracy of the measurement is improved.

Comparison spheres also can be divided in two types: (1) sphere with single beam and (2) sphere with double beam. The first one involves having both the sample and the reference on the sphere on the same time. Position of the sample and reference are switched between two measurements to allow each one of them to be in the incident beam, but both samples never leave the sphere and substitution concerns are minimized. A more accurate method for overcoming the substitution effects appears to be a sphere with two identical incident beams - one for the sample and one for the reference. In this type of sphere physical switching between the sample and the reference positions is no longer needed. Illumination of the sample with light beam enables measurements of Td and Rd to be performed (details can be seen from figure 3.7). Simultaneously when transmittance or reflectance spectra are collected from a sample, the second exit port is covered with a calibrated standard with known R and reference measurement is performed.

Several sources of errors have to be considered when operating integrating sphere measurements. The first one is related to the sphere geometry itself, and is due to the fact that the ratio of the area of the various ports of the sphere to the total area of the sphere is

not zero, as it needs to be according to the integrating sphere theory. Fundamental integrating sphere theory and sources of errors in the measurements have been extensively treated by several authors [23, 32, 33]. The most influential sphere error appears to be in the Td measurement due to the fact that the sphere geometry is affected slightly by introducing the sample at the entrance port, and the sample itself can reflect some of the light inside the sphere, which will influence the measurement.

The second important potential source of error arises from lateral losses of light due to the finite size of the ports, where some part of the beam that reflects at wide angles could be intercepted by the sphere wall. This leads to overestimated absorption properties for samples with low absorption [34, 35] where the absolute error in the evaluation of $\mu_a$ is directly determined by the error in the measurement of Rd and Td.

As already has been discussed in section 3.4.1 of the current chapter, collimated transmittance is the third measurement that has to be completed in order to obtain all three optical properties from the IAD. However, collimated beam measurement has its own source of error because the main idea of this measurement is to detect only the light that has penetrated the sample without being scattered. In an ideal case this requires zero collection angles [36] which does not coincide with real instrumental setup, where scattered light might be included in the measurement.

In order to facilitate and increase the accuracy of spectral information collected with integrating sphere, a double integrating sphere set up has been designed with incorporated collimated beam measurement. With the making of this setup (figure 3.4) simultaneous measurements of all three parameters have been achieved. The integrating sphere method, combined with a collimated beam measurement, in practice provide almost the only way to measure absorption coefficient $\mu_a$, scattering coefficient $\mu_s$ and anisotropy factor $g$ of bulk material accurately. The technique was chosen by Troy and Thennadil while studying human skin properties[25], and by Pickering et *al.* for

measuring optical properties of tissue[24]. It could provide simultaneous monitoring for all three optical properties changes, and to enable collection of unique information during analysis of dynamic systems, for example growing biological suspensions.
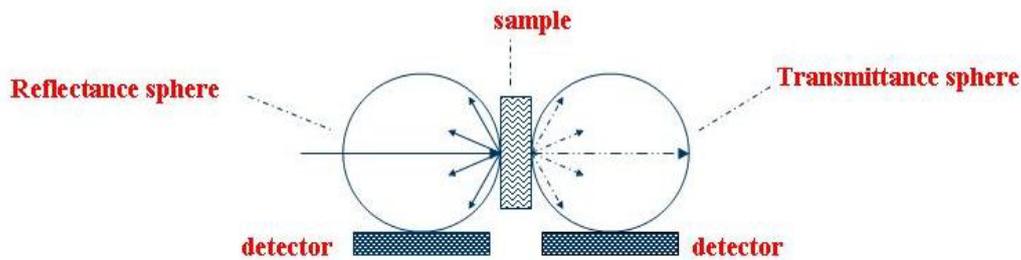


Fig.3.4. Double integrating sphere setup for parallel measurement of Rd, Td and Tc.

However, some of the scattered light from the sample in the double integrating sphere setup, will reflect off the wall of the second sphere, reflect again off the sample, and possibly will interfere with the measurement. It will introduce an additional source of error, and the method should be therefore used with caution. Several authors investigated to what extent the error sources will affect the optical properties estimation. Van Gemert et *al.*[37] developed an inverse Monte Carlo algorithm in order to investigate the light losses in double integrating sphere setup, and concluded that light loss through the exit port is mainly dependent on the anisotropy factor. Quantitative analysis with IAD [38], for light losses have shown below 2% relative error on the scattering coefficient and maximum relative error of absorption coefficient to reach 28%. The authors analyzed influences of (1) absorption coefficient ($\mu_a$) overestimation, (2) sample port diameter and (3) thickness of the sample on the light losses. Recommendation for sample port changes have been made, and advice for high caution while sample thickness is chosen. Finally IAD has been found suitable for estimating optical properties of high scattering and high absorbing samples.

In this section an overview of the integrating sphere method has been given with respect to the general case. Based on published work, light losses and possible error sources were outlined. From the two configurations (single and double integrating spheres), preferences have been given to double integrating sphere, as this technique could provide near-simultaneous results for the three measurements of interest (Td, Rd, and Tc), and helps to avoid change over between the different configurations. This fact reduces the time for sample scanning and increases the accuracy of data taken from dynamic systems, such as growing biological suspensions.

While the double integrating sphere setup offers operational simplicity and avoids the problem of potential errors that can occur when using the single sphere setup due to the requirement of adjustments between different measurements, the latter setup avoids the errors that occur due to "cross-talk" that occurs in a double sphere setup. The major drawback in the single integrating sphere setup is the amount of time it takes to make all the three measurements. This has implications for samples which are prone to change during the measurement time. However, the double integrating sphere was not commercially available at the time this research work was initiated, and although all its advantages comparing to single integrating sphere, the latter was used for this study. The drawbacks of using the single integrating sphere setup were minimized by optimizing the experimental protocol for collecting the spectra. Details of this measurement setup are explained with the next section.

### 3.5.2. Measurements using the CARY 5000

Total Diffuse Reflectance (Rd) and Total Diffuse Transmittance (Td) measurements, were performed using a UV-Vis-NIR spectrophotometer (Cary 5000, product of Varian Scientific Instruments) equipped with external Diffuse Reflectance Accessory (DRA) – 2500, which houses a single integrating sphere. The instrument operates as a double

beam system with a R928 photomultiplier tube (PMT) detector for the UV-Vis region, and the Varian technology - PbSmart$^{TM}$ (cooled PbS photocell), for obtaining low noise performance into the NIR region. The wavelength range of Cary 5000 is between 175nm and 3300nm and between 250nm to 2500nm when the external DRA is attached. The main advantage of the DRA is that this accessory provides means of measuring reflectance and transmittance of opaque and turbid samples, and as such applicable for measuring turbid biological materials. The integrating sphere is fitted with the same detector configuration as the host spectrophotometer. The diameter of the integrating sphere is 150mm. The accessory is equipped with movable mirror system allowing beam adjustment for transmittance or reflectance measurements.

There are three main DRA components: integrating sphere, optics chamber and detector chamber (figure 3.5).
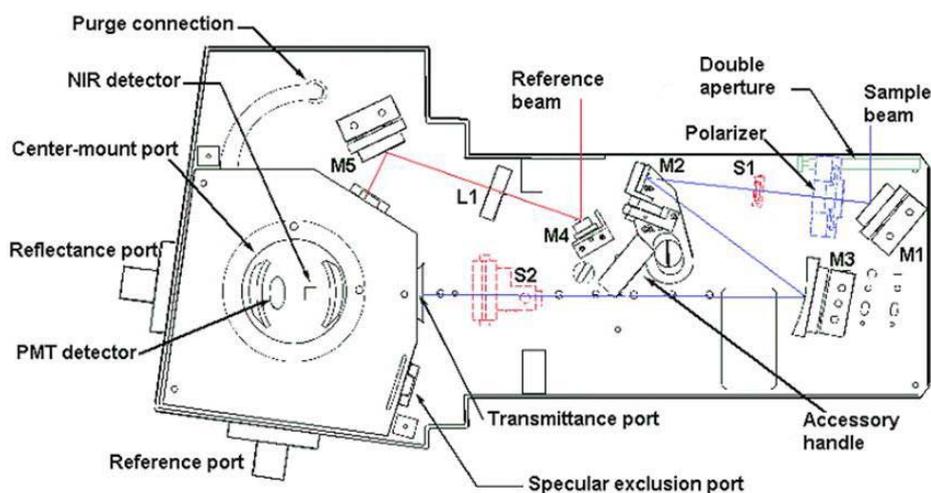


Fig. 3.5. Design of external DRA 2500. Source: Technical manual of Varian.

The optics chamber contains a transfer mirror system that directs the sample beam and the reference beam to their final points. The integrating sphere collects and measures the transmitted/reflected radiation from the sample surface. The sphere is coated on the inside with durable PTFE material with greater than 99% reflectance between 350nm and

1800nm. A specially designed cover fits over the DRA in order to protect the operating environment from any stray light. For the same reason, both sample reflectance and reference ports are fitted with separate magnetic port covers.

The path that light travels with this system within the accessory is illustrated in figure 3.6. M1, M2, and M3 represent three different mirrors the main purpose of which is to guide the sample beam through the sample transmission port of the integrating sphere and onto the sample reflectance port at an $8^0$ angle of incidence. Mirrors M4 and M5 direct the reference beam through the reference beam entrance port onto the reference port.
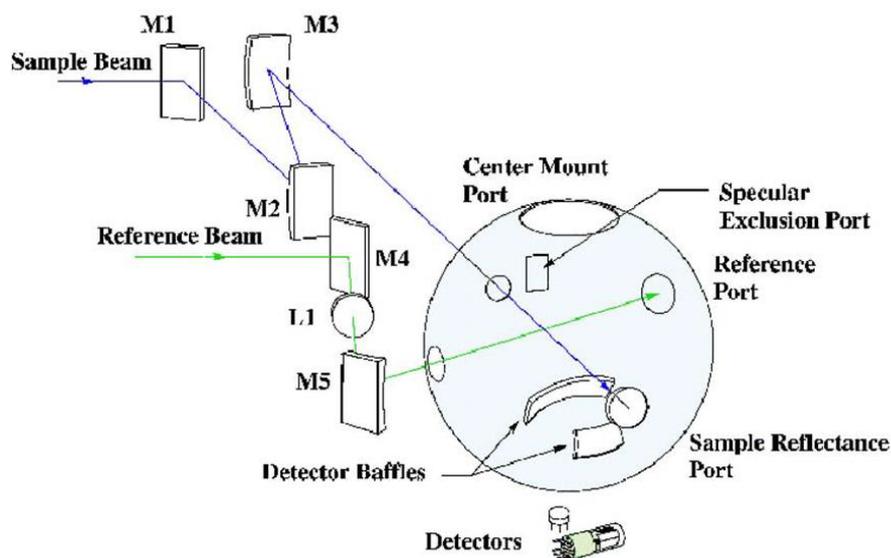
Fig. 3.6. Optical design of external DRA 2500. Source: Technical manual of Varian.

Light entering the accessory is directed to the two entrance ports of the sphere, (1) the entrance port for the reference beam and (2) the entrance port for the sample beam. The readings from the detector are recorded from each separate beam with no physical switching between the sample and the reference placements.

When performing reflectance measurements, two types of baseline corrections are recorded. The first one, zero baseline correction is performed by blocking off the sample beam from the integrating sphere and in this case the detector should register 0% light transmission. For the second baseline correction, measurement from calibrated reflectance standard (provided by Labsphere) is taken in order to verify the accuracy of the instrument and to prevent erroneous results. During this scan, the second calibrated reference standard is placed at the sample reflectance port (figure 3.6) and plays the role of approximately 100% reflectance sample. The exact values for this calibrated standard are provided by Labsphere. The next step is sample measurement which involves replacement of this calibrated standard with the sample itself. Meanwhile the first reference standard does not leave its place from the reference port. During the measurement, light beam first enters the sphere, illuminates the sample, and then the reflection off the sample surface is collected by the sphere as a diffuse reflectance (figure 3.7b).
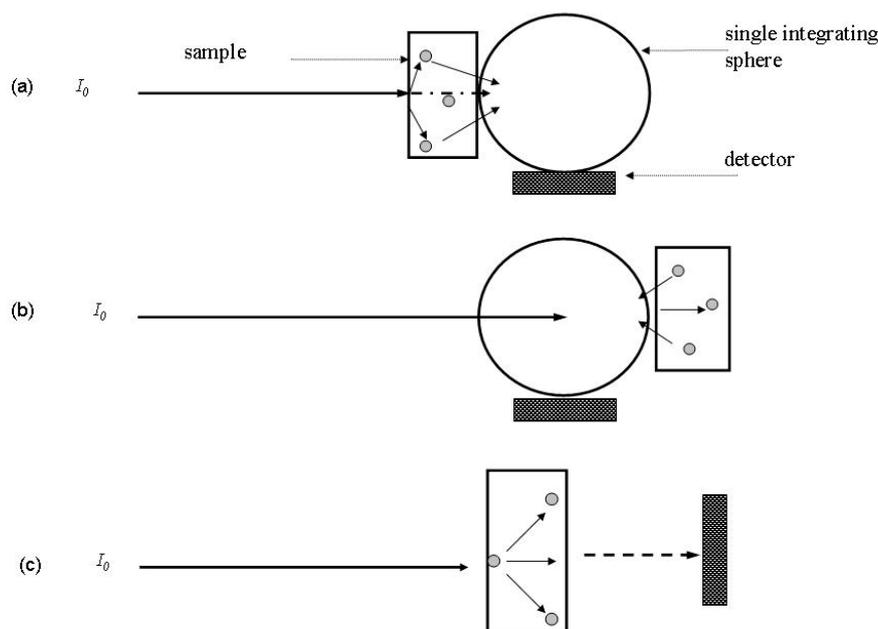


Fig. 3.7. Measurement setup for (a) diffuse transmittance, (b) diffuse reflectance and (c) collimated transmittance.

For the diffuse transmittance measurements, the sample is placed at the entrance port of the sphere. In this case the beam first reaches the sample and the transferred amount of light that enters the sphere is measured as Diffuse Transmittance (figure 3.7a). Baseline correction has been also applied before each transmittance measurements.

For the third measurement, collimated transmittance from a sample, a use of integrating sphere setup is no longer needed. The spectrophotometer has a standard setup that is used to measure in collimated mode (figure 3.7c).

While performing measurements with integrating sphere, the sample is usually placed between two glass slides. In this study, sample's liquid nature made necessary the use of special glass cuvettes. Manufactured by Hellma, cuvettes are produced from synthetic Quartz material free from OH absorption, with high purity and homogeneity. Samples from *Bacillus subtilis* growing suspension were transferred in 2mm, 4mm and 10mm pathlength cuvettes, for scanning in the NIR wavelength region 950-1850nm. All three cells were covered on the top, either with stoppers (2mm and 10mm) or plastic lid (4mm) in order to prevent and control any risk and any exposure to hazards or to avoid possible environmental contamination. These covers also kept the samples protected from external influence (e.g. air, evaporation, external contamination etc.) during the measurements.

For all measurements, the same optimal parameters have been applied for consistency reasons. Scan rate time for the NIR region was kept 600nm/min with data interval of 4nm and average time of 0.4s. Under the condition of these parameters, the instrument required approximately 45 min for a complete scan for all three measurements including time required to change from one measurement configuration to another.

As already mentioned before, a single integrating sphere setup has been used for this experimental work. This configuration necessitates consecutive measurements for Rd, Td and Tc to be performed. Every change over from one measurement to another involves

switching off the instrument, removal of the DRA from the spectrophotometer, and change of the main lenses as they are different for different measurements. Due to these changes in the measurement mode from reflectance to transmittance, then to collimated transmittance, and vice versa, a variation in the measured values is usually introduced. Its origin is mainly related to the adjustment of the sample beam which slightly changes positions after each instrument switch on and switch off. It was also found that the sample beam sometime changes positions accidentally due to internal vibrations of the spectrophotometer. In order to overcome these factors affecting the measurements, incident light spot size and placement were checked before each Td and Rd measurement by following a procedure from Varian operational manual, named 'alignment check'. Translucent white paper is slid into the reflectance/transmission port, in order to check that the sample beam is centered and falls every time completely within the port and the cuvette surface. The same step was repeated for the reference beam at the reference port.

The signal-to-noise ratio and the precision of the measurements can be affected in some cases if the reflected light escapes through the sphere ports. Magnetic covers for the different ports, provided by the manufacturer were used to close the system and care was taken to ensure that they were properly placed to prevent stray light. The inside part of the sphere was observed for ageing or contamination and if not in use, the integrating sphere setup was kept in a special black box to prevent contamination. Prior to each measurement, the sample placement on the outside of the sphere wall was regularly monitored. Any spacing between the sample and the sphere wall can lead to large errors due to loss of reflected light that might escape through the sample port. To avoid incorrect placement, the cuvette holder was kept closely attached to the sphere wall.

To ensure repeatability, light spot size and location were regularly checked. The repeatability of measurements was estimated by performing multiple measurements with a 10% calibrated reflectance standard. Results from these scans are shown in Appendix A

(figure A.2). 10% calibrated reflectance standard  was chosen due to its close reflectance values with the reflectance values of the biological system studied in this work, which took values between 4% and 8%.

In order to validate the applicability and the accuracy of the measurement setup used, and the inverse adding-doubling code, a validation step has been implemented as a predecessor of the dynamic biological system. The validation step is described in the next section.

## 3.6. Validation of the method and the experimental setup for extraction of optical properties

As a first step the experimental setup, measurement protocol, and code for the method were validated using a solution of a low concentration polystyrene microspheres exhibiting similar spectral behavior to the system studied. The model system consisted of polystyrene microspheres with known diameter (0.45 micron) and concentration (0.15% by weight), suspended in water. A similar procedure was used by Troy and Thennadil[25] for the case of a double integrating sphere set up.  The measurements Td, Rd, and Tc were obtained for this system using 1mm quartz glass cuvette purchased from Hellma. The validation procedure is divided into two steps, step (a) and step (b) which are described below:

(a) Theoretical values for scattering coefficient $\mu_s$ (figure 3.8) were calculated from the microsphere radius $r$ and the refractive indices of polystyrene and water using Mie theory [2] in the wavelength region from 350nm up to 2200nm. Experimental values for $\mu_s$ were also obtained by calculating them using the inverse adding-doubling method. Figure 3.8 shows the agreement between theoretically calculated and experimentally obtained values for $\mu_s$. From the two curves obtained for the scattering coefficients it can be seen that both have good agreement over most of the wavelength range considered. Some

differences occur below 900nm and a small peak appears around 1450nm. In this research, the biological system under consideration is studied only in the region 950-1850nm due to the low quality of the spectroscopic data beyond 1850nm. The truncation of the NIR region is due to the fact that high amount of water is present in the sample and the high water absorption leads to the very low signals beyond 1850nm.
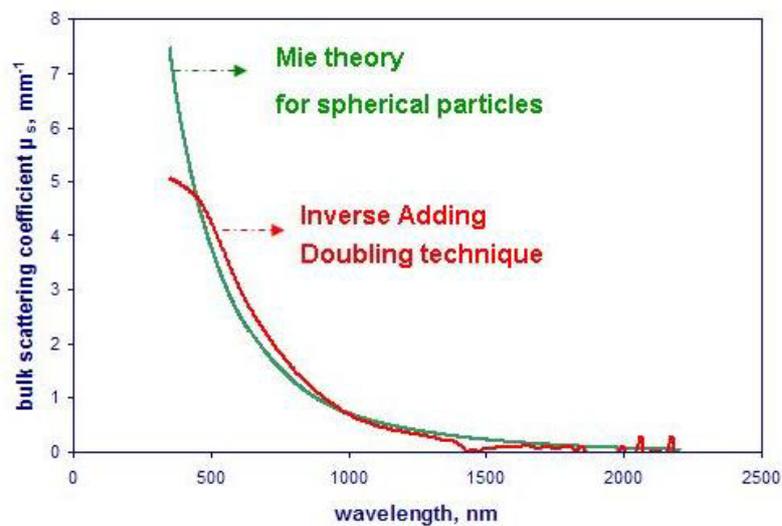


Fig. 3.8 Scattering coefficient $\mu_s$, calculated using inverse adding-doubling techniques and Mie theory for spherical particles.

The agreement between both curves was found satisfactory between 950-1850nm which is the region that was chosen for this study. This range of wavelengths in the NIR region was chosen because it includes the first overtone and second overtone bands of proteins, sugars and other organics. While it would have been desirable to include the combination band region (2000-2500nm), the low quality of signal in this region due to high water absorption indicated that with the current setup it would not be possible to obtain measurements with sufficient quality for the inversion step to be successfully applied.

(b) While the validation approach used above uses a check based on individual particle properties i.e. particle size and concentration and using Mie theory to predict the

scattering coefficient $\mu_s$, the second check was at the level of the bulk optical properties. In other words, the objective was to find out if the inverse adding doubling method (along with the experimental setup) provided optical property estimates that would reproduce the measured spectra to within experimental error. Comparison for calculated and experimental values for all three measurements Td, Rd, and Tc is shown on figure 3.9. By observing all three plots, it can be concluded that the experimental and calculated values are in a very good agreement over almost the entire wavelength region except for a small region around the water absorption peak. Despite this, for the study it was decided not to measure beyond 1850nm because of the high absorption of water from 1850nm up to 2200nm.

Based on the two validation steps, (a) and (b), and the results shown with figures 3.8 and 3.9, it can be concluded that the experimental setup and the inverse adding-doubling method can be successfully applied for the extraction of optical properties from *Bacillus subtilis* growing in a liquid culture which has similar optical properties as the model system of polystyrene used for this validation step.

(a)



(b)



(c)
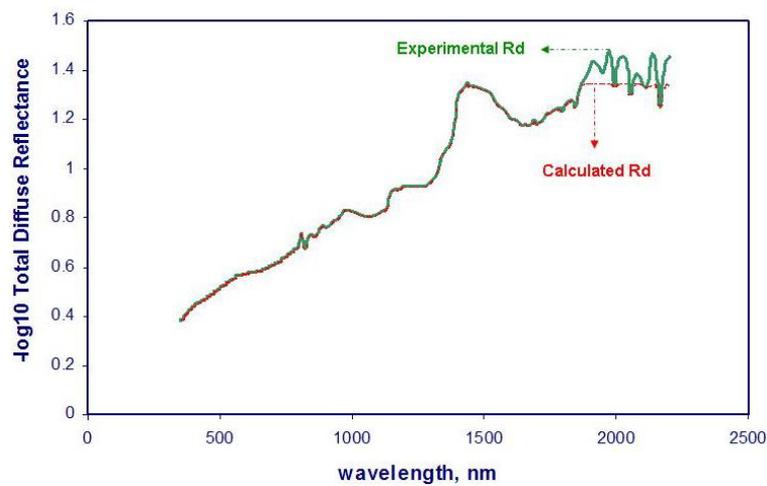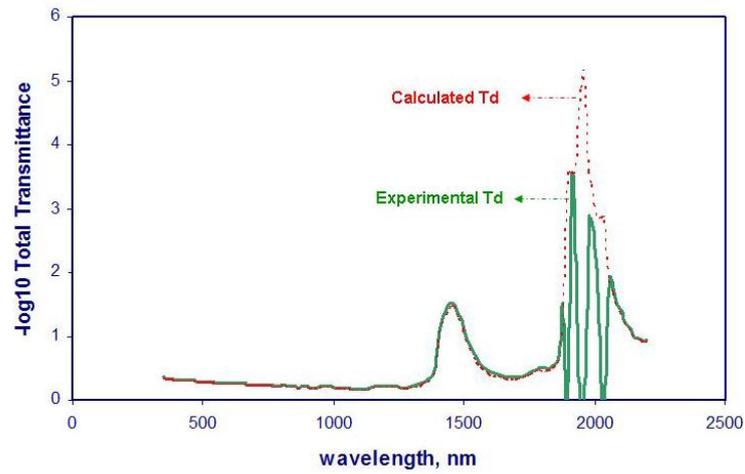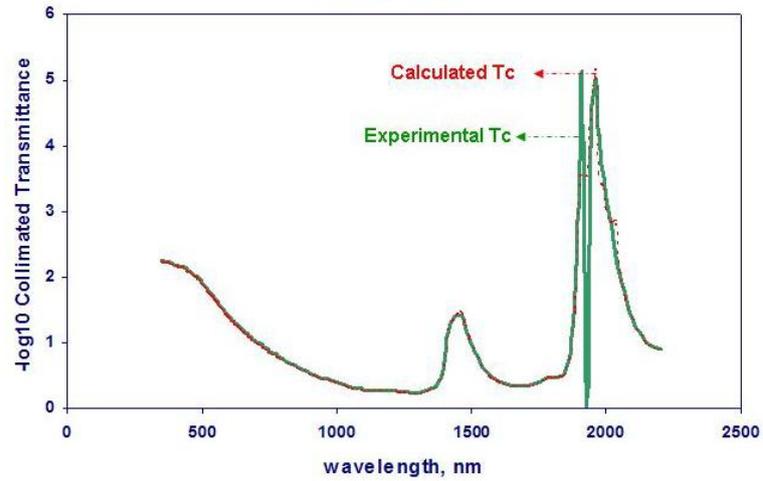
Fig. 3.9. Comparison between experimental and calculated data: shown are (a)
collimated transmittance, (b) diffuse transmittance and (c) diffuse reflectance.

### 3.7. Discussion

In this chapter a method has been described for finding the optical properties (scattering, absorption and anisotropy factor) of a turbid sample by using diffuse reflectance, diffuse transmittance and collimated transmittance measurements. The general idea is to apply this method for monitoring changes in the optical properties of *Bacillus subtilis* growing in a liquid culture in order to extract important physical and chemical information from the system. These optical properties can be obtained by inverting the radiative transfer equation and iterating the adding-doubling method until calculated values for reflectance and transmittance match the measured ones. Validation of the measurement setup and the IAD method was done in order to test the stability and the accuracy of the proposed methodology. Results from the validation have shown good predictive capability in the 950-1850nm wavelength region.

In chapter 4, changes in the optical properties during growth of *Bacillus subtilis* in liquid culture will be investigated.

### References

1.      Hulst, H.C.v.d., *Multiple Light Scattering*. Vol. 1. 1980, New York: Academic Press.

2.      Bohren, C.F., D.R.Huffman, *Absorption and Scattering of Light by Small Particles*. 1983, New York: John Wiley&Sons.

3.      Kerker, M., *The scattering of light and other electromagnetic radiation*. 1969, New York.

4.      Ishimaru, A., *Wave Propagation and Scattering in Random Media*. The IEEE/OUP Series on Electromagnetic Wave Theory. 1997: IEEE Press, Oxford University Press. 574.

5.      Chandrasekhar, S., *Radiative Transfer*. 1960: London  and New York and Dover, Oxford University Press.

6.      Peters, V.G., D.R. Wyman, M.S. Patterson, and G.L. Frank, *Optical properties of normal and diseased human breast tissues in the visible and near infrared.* Physics in Medicine and Biology, 1990. **35**(9): p. 1317-1334.

7.      Graaff, R., M.H. Koelink, F.F.M.d. Mul, W.G. Zijlstra, A.C.M. Dassel, and J.G. Aarnoudse, *Condensed Monte Carlo simulations for the description of light transport.* APPLIED OPTICS, 1993. **32**(4): p. 426-434.

8.      Yaroslavsky, I.V., A.N. Yaroslavsky, T. Goldbach, and H.J. Schwarzmaier, *Inverse hybrid technique for determining the optical properties of turbid media from integrating sphere measurements.* APPlIED OPTICS, 1996. **35**(34): p. 6797-6809.

9.      Jagajothi, G. and S. Raghavan, *Estimation of optical properties in biological tissues using Monte Carlo simulation.* Journal of Mechanics in Medicine and Biology, 2007. **7**(4): p. 449-462.

10.     Prahl, S.A., M. Keijzer, S.L. Jacques, and A.J. Welch, *'A Monte Carlo model of light propagation in tissue', SPIE Institute Series: Dosimetry of Laser Radiation in Medicine and Biology.* 1989: p. 102-111.

11.     Bevilacqua, F., D. Piguet, P. Marquet, J.D. Gross, B.J. Tromberg, and C. Depeursinge, *In vivo local determination of tissue optical properties: applications to human brain.* APPLIED OPTICS, 1999. **38**: p. 4939-4950.

12.     Yaroslavsky, A.N., P.C. Schulze, I.V. Yaroslavsky, R. Schober, F. Ulrich, and H.J. Schwarzmaier, *Optical properties of selected native and coagulated human brain tissues in vitro in the visible and near infrared spectral range.* Physics in Medicine and Biology, 2002. **47**: p. 2059-2073.

13.     Liou, K.N., *An introduction to atmospheric radiation*, ed. 2. 2002: London, Academic Press.

14.     Prahl, S.A., *The Adding-Doubling Method*, in *Optical-Thermal Response of Laser Irradiated Tissue*, A.J. Welch and M.J.C.v. Gemert, Editors. 1995, Plenum Press. p. 101-129.

15.     Prahl, S.A., *Light Transport in Tissue*. 1988, PhD Thesis. University of Texas at Austin: Austin, Texas.

16.     Prahl, S.A., M.J.C. Vangemert, and A.J. Welch, *Determining the Optical-Properties of Turbid Media by Using the Adding-Doubling Method.* Applied Optics, 1993. **32**(4): p. 559-568.

17.     Flock, S.T., B.C. Wilson, and M.S. Patterson, *Total attenuation coefficients and scattering phase functions of tissues and phatom materials at 633nm.* Medical Physics, 1987. **14**(5): p. 835-841.

18.     Marchesini, R., A. Bertoni, S. Andreola, E. Melloni, and A.E. Sichirollo, *Extinction and absorption coefficients and scattering phase functions of human tissues in vitro.* APPLIED OPTICS, 1989. **28**(12): p. 2318-2324.

19.  Nelson, N.B. and B.B. Prezelin, *Calibration of an integrating sphere for determining the absorption coefficient of scattering suspensions.* APPLIED OPTICS, 1993. **32**(33): p. 6710-6717.

20.  Steinke, J.M. and A.P. Shepherd, *Comparison of Mie theory and the light scattering of red blood cells.* APPLIED OPTICS, 1988. **27**: p. 4027-4033.

21.  Dam, J.S., C.B. Pedersen, T. Dalgaard, P.E. Fabricius, P. Aruna, and S. Andersson-Engels, *Fiber-optic probe for noninvasive real-time determination of tissue optical properties at multiple wavelengths.* APPLIED OPTICS, 2001. **40**(7): p. 1155-1164.

22.  Patterson, M.S., B. Chance, and B.C. Wilson, *Time resolved reflectance and transmittance for the non-invasive measurement of tissue optical properties.* APPLIED OPTICS, 1989. **28**(12): p. 2331-2336.

23.  Pickering, J.W., C.J.M. Moes, H. Sterenborg, S.A. Prahl, and M.J.C. Vangemert, *Two Integrating Spheres with an Intervening Scattering Sample.* Journal of the Optical Society of America a-Optics Image Science and Vision, 1992. **9**(4): p. 621-631.

24.  Pickering, J.W., S.A. Prahl, N. Vanwieringen, J.F. Beek, H. Sterenborg, and M.J.C. Vangemert, *Double-Integrating-Sphere System for Measuring the Optical-Properties of Tissue.* Applied Optics, 1993. **32**(4): p. 399-410.

25.  Troy, T. and S. Thennadil, *Optical properties of human skin in the near infrared wavelength range of 1000 to 2200nm.* Journal of Biomedical Optics, 2001. **6**: p. 167-176.

26.  Gebhart, S., W. Lin, and A. Mahadevan-Jansen, *In vitro determination of normal and neoplastic human brain tissue optical properties using inverse adding-doubling* Physics in Medicine and Biology, 2006. **51**: p. 2011-2027.

27.  Saeys, W., M. Velazco-Roa, S. Thennadil, H. Ramon, and B. Nikolai, *Optical Properties of Apple Skin and Flesh in the Wavelength Range from 350 to 2200nm.* APPLIED OPTICS, 2008. **47**(7): p. 908-919.

28.  Velazco-Roa, M.A. and S.N. Thennadil, *Estimation of complex refractive index of polydisperse particulate systems for multiple-scattered ultraviolet-visible-near-infrared measurements.* APPLIED OPTICS, 2007. **46**(18): p. 3730-3735.

29.  Hellma, *Hellma Genaral Catalogue.* 2002: Hellma GmbH & Co. KG.

30.  Carr, K.F., *Integrating sphere theory and applications. Part I: Integrating sphere theory and design.* Surface Coatings International Part B: Coatings Transactions, 1997. **80**(8): p. 380-385.

31.  Carr, K.F., *Integrating sphere theory and applications. Part II: integrating sphere applications.* Surface Coatings International Part B: Coatings Transactions, 1997. **80**(10): p. 485-491.

32.  Jacquez, J.A. and H.F. Kuppenheim, *Theory of the integrating sphere.* Journal of the Optical Society of America 1955. **45**(6): p. 460-470.

33.  Goebel, D.G., *Generalized Integrating Sphere Theory.* APPLIED OPTICS, 1967. **6**(125-128).

34.  Torres, J.H., A.J. Welch, I. Cilesiz, and M. Motamedi, *Tissue Optical Property Measurements - Overestimation of Absorption-Coefficient with Spectrophotometric Techniques.* Lasers in Surgery and Medicine, 1994. **14**(3): p. 249-257.

35.  de Vries, G., J.F. Beek, G.W. Lucassen, and M.J.C. van Gemert, *The effect of light losses in double integrating spheres on optical properties estimation.* Ieee Journal of Selected Topics in Quantum Electronics, 1999. **5**(4): p. 944-947.

36.  Wang, L.H. and S.L. Jacques, *Error Estimation of Measuring Total Interaction Coefficients of Turbid Media Using Collimated Light Transmission.* Physics in Medicine and Biology, 1994. **39**(12): p. 2349-2354.

37.    Vries, G.d., J.F. Beek, G.W. Lucassen, and M.J.C.v. Gemert, *The effect of light losses in double integrating spheres on optical properties estimation.* IEEE Journal of Selected Topics in Quantum Electronics, 1999. **5**(4): p. 944-947.

38.    Zhu, D., W. Lu, S. Zeng, and Q. Luo, *Effect of light losses of sample between two integrating spheres on optical properties estimation.* Journal of Biomedical Optics, 2007. **12**(6): p. 0640041- 0640048.

# Changes in optical properties during the growth of *Bacillus subtilis* in liquid culture

## 4.1. Introduction

In this chapter, a method for the estimation of optical properties in the Near-Infrared region for *Bacillus subtilis* growing in a liquid medium is introduced. Absorption coefficient $\mu_a$, scattering coefficient $\mu_s$ and anisotropy factor $g$ are obtained by using inverse adding-doubling method to solve the radiative transfer equation in combination with measurements for total diffuse reflectance, total diffuse transmittance and collimated transmittance using a single integrating sphere setup.

This approach leads to the separation of the absorption and scattering effects using the RTE to account for multiple scattering of light [1, 2]. Thus changes in the absorption and scattering properties can be separately monitored during the microbial growth cycle. The scattering coefficient is expected to be predominantly related to changes in properties such as cell size, and biomass concentration whereas information regarding non-scattering constituents such as glucose, and product concentrations will be contained in the absorption coefficient. By extracting $\mu_a$ and $\mu_s$, it may be possible to obtain more information for monitoring microbial growth than could be obtained using only the transmittance or reflectance data since by separating the scattering and absorption effects, the issues arising due to the interaction of the two effects can be eliminated to a large extent.

This study focuses on the methodology for extracting optical properties in a consistent manner and investigates their nature and extent of changes over the course of a bacterial growth cycle. For this purpose a simple culture system was established consisting of

*Bacillus subtilis* growing in an aqueous solution (minimum medium). The changes in the optical properties of this system during growth, stationary and decline phase were studied using measurements in the wavelength region of 950nm-1850nm. The reason for using this particular region was explained in chapter 3, as a part of the validation process.

## 4.2. Experimental protocol

The purpose of this section is to describe in details single growth cycle, the type of the used medium, the type of measurements, and assays that have been performed on samples drawn at the same time as those used for spectroscopic measurements.

*Bacillus subtilis* is well known for having very simple growth requirements. It is able to grow on a simple salts medium containing glucose, or other sugars, as a carbon and an energy source and inorganic nitrogen, along with an adequate supply of oxygen. This simple medium has permitted detailed study of the investigated system from spectroscopic point of view by knowing that bacterial cells, water, and glucose are the main components.

In this work, *Bacillus subtilis* culture, growing in liquid culture (minimum medium and trace element solution) was considered. The growth studies were conducted using the Gram-positive bacterium *Bacillus subtilis* strain 168 which was obtained from the Institut Pasteur, Paris. The strain was cultivated in 100ml Spizizen's minimal medium[3] and trace element solution[3] in a 250 ml Erlenmeyer flask. Spizizen's minimal medium contains, per litre: $(NH_4)_2SO_4$-2gr, $K_2HPO_4$-14gr, $KH_2PO_4$-6gr, $Na_3citrate.2H_2O$-1gr, $MgSO_4.7H_2O$-0.2gr. The prepared solution is autoclaved at 15lb/in$^2$ for 30min, cooled down, and add 10ml sterile solution of 50% (w/v) D-glucose, and 10ml of L-tryptophan. The experiment for the growth cycle began with inoculation of pre-warmed culture medium from overnight liquid starter dense culture. The growth cycle was initiated by

inoculating 100ml minimum medium with 1ml of the dense culture in a way that the initial optical density is approximately 0.05. The growth was performed at carefully controlled conditions: temperature $37^0C\pm0.5$, initial pH $7 \pm 0.5$ of the medium, and agitation rate 220rpm maintained in an orbital shaker. In this study, data from the growth, stationary and decline phases were collected from separate cultivations. A total of 9 growth cycles were performed in order to ensure repetition was accurate, and for all these cultures, during the data collection phase, five samples were taken at approximately 2 hour intervals.

On each sample spectroscopic measurements for total diffuse reflectance (Rd), total diffuse transmittance (Td) and collimated transmittance (Tc) were made using a UV-Vis-NIR spectrophotometer (Cary 5000, Varian Scientific Instruments) equipped with an integrating sphere (Diffuse Reflectance Accessory - DRA 2500, figure 3.6, chapter 3). The spectroscopic measurements were made by placing the samples in a special optical glass cuvette. Since the samples were not treated prior to spectroscopic measurements, it can be expected that the changes due to the continuation of bacterial growth cycle will create some variation in the spectra during the measurement period. However, while this impact on the spectral data is not known, it was assumed it was not high enough to have a large effect at least on the bulk optical property estimations provided the measurements are made within a reasonable duration.

Simultaneously with spectroscopic studies biomass and glucose concentration measurements were also performed. There are two reasons for making these measurements. One was to aid in the qualitative investigation as to how the changes in the biomass and glucose concentrations affect the absorption and scattering properties of the growing medium. The second reason was to aid in the development of methodologies for building calibration models on the extracted optical properties for predicting biomass and glucose concentrations using chemometric techniques.

### 4.2.1. Reference assays

The details of how the reference assays for biomass and glucose concentrations were made are given in the next two subsections. In addition, an analysis of the measurement errors and relationship between bacterial growth cycle, biomass formation and glucose concentration depletion are given.

*Biomass (dry weight)*

Biomass was measured gravimetrically and off-line. Aliquots of 5ml growing suspension were filtered without washing through preweighed Millipore filter (pore size 0.45µm). After the culture suspension was filtered, the filters were then dried at $50^0C$ to a constant weight, cooled down in a vacuum desiccator, and then weighed again. The difference between dried filter and pre-weighed filter was expressed as weight of the dry cells per sample volume. The resultant filtrate was used for the glucose concentration assay. With the increase of the biomass, for stationary and decline phases, 5 ml of growing suspension was not possible to be filtered through one filter due to blockage of the filter pores. In this case the 5ml was filtered by filtering 1ml at a time using five filters that were weighed together. The filters were then dried following the same procedure described earlier, and finally weighed together again. The difference between the five dried filters and five pre-weighed filters was used to measure the dry cell weight per 5ml of sample volume. Then, the final results were calculated in terms of weight per 1ml of sample volume.

Figure 4.1 presents the changes in the biomass at growth, stationary and decline phases. The three different symbols (star, triangle and circle) were used to present results obtained for the three different runs from each phase. During the growth process the initial biomass was measured to be about 0.4mg/ml. In the stationary phase it varied

between different runs from 2 mg/ml to 2.4 mg/ml, and a slight decrease was registered in the decline phase with the biomass varying between 1.85 - 2.1 mg/ml.
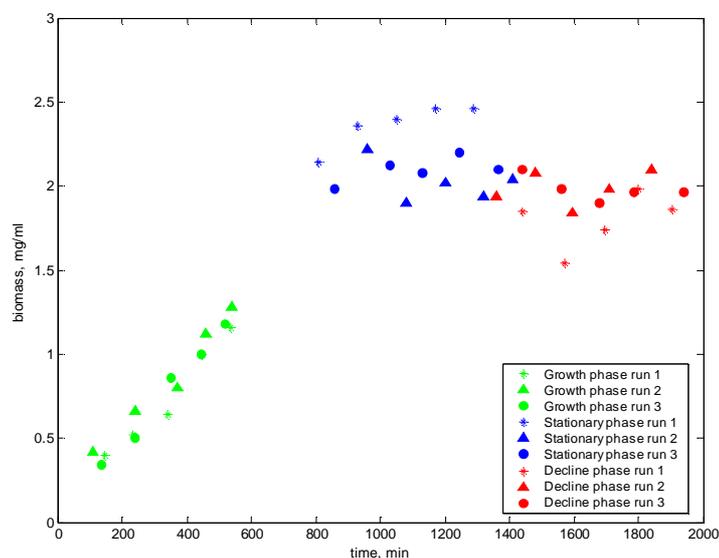


Fig. 4.1. Dry weight of cells profiles for three cultivation runs (green colour-growth phase, blue colour-stationary phase and red colour-decline phase)

The error estimation for biomass measurement has been determined by measuring four replicates from the culture at different growth time. The samples were filtered and measured following the above described procedure and the average measurement error has been estimated to be 8.2% (w/v). There are two possible major sources that contribute to this error. The first one is embedded in the standard deviation of the balance itself. Since the weight of the cells in this experiment is very small (growth phase biomass within 0.4-1.5mg/ml), a possibility for weight values to be affected needs to be borne in mind. The standard deviation of the classical analytical balance used in this experimental work (Mettler Toledo with 320gr maximum capacity) given by the manufacturer is 0.1mg, which compared with the amount of biomass produced at an early stage of the growth cycle suggest and error between 6-25% to be expected. Due to difficulties in obtaining balance with different weighting capacity and respectively higher accuracy, this source of error was dealt by (1) ensuring that the balance is leveled

correctly, and (2) weekly calibration with certified weight standard has been performed in order to prove measurement traceability. The second source of error is related to the possibility for cellular mass and released DNA to clump together, and as a result samples from the same sampling point will have different biomass concentration then others. To reduce the clump effect, the growth was carried out in an orbital incubator at 220 rpm, and the culture was continuously shaken during the growth, however this error is difficult to estimate.

*Glucose concentration*

Glucose solution (50% w/v D-Glucose solution) was prepared with deionized water and sterilized by autoclaving at 10lb/in$^2$. 1ml from this sterile solution was added to 100ml of media in order to obtain Spizizen' minimal medium[3]. Theoretically calculated concentration was determined to be 5mg/ml; however measurements determined the starting concentration to be in the range of 4mg/ml. The origin of this difference was investigated and it was found that the glucose loss is most probably due to issues associated with the autoclaving. Sigma-Aldrich D-Glucose kit (GAGO) for glucose concentration determination was utilized following the instructions provided by manufacturer. Levels of glucose in the current experiment were exceeding the range of the test and this fact implied 50 folds dilution on each sample prior the analysis. Absorption measurements at 540nm were conducted with UV-Vis NIR spectrophotometer (Cary 5000) in 1cm cell for each sample drawn from the culture. Glucose concentration in the growing system considered in this study is the main limitation factor for the growth, owing to the fact that minimal medium typically contains the minimum nutrients possible for colony growth, and a small number of salts. The system studied is known as batch culture where simple conditions exist and it refers to

growth in a fixed volume with a fixed amount of nutrients. For this reason glucose concentration was measured (figure 4.2) only in the growth phase.
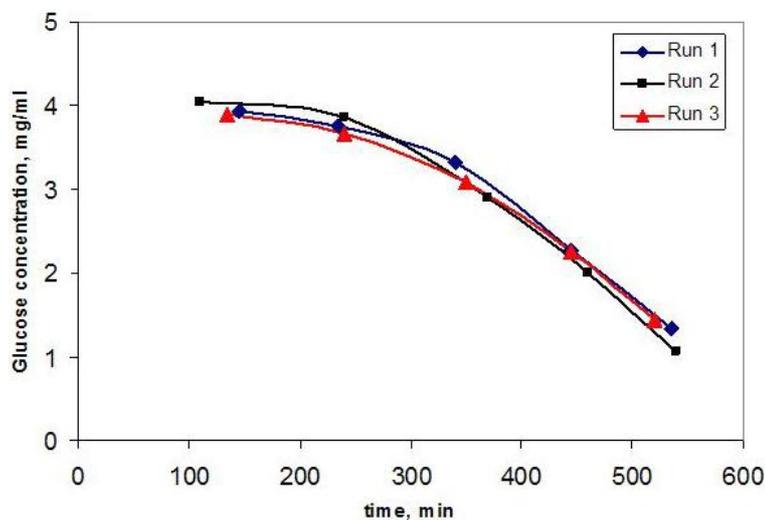


Fig. 4.2. Glucose concentration profiles for three cultivation runs.

From an initial concentration of 4mg/ml, level of glucose during the process has fallen to approximately 1mg/ml after 500 min from the beginning of the growth cycle. This decrease in the glucose level is matched with the increase of the microbial concentration. The residual amount has been rapidly metabolized by the bacteria and indeed by 570 min the concentration was found equal to zero. Figure 4.3 shows the time when complete glucose depletion was found attained, and it concerns glucose concentration during *Bacillus subtilis* single growth cycle from a separate run which was not spectroscopically analyzed. Measurements were performed in replicates of five in order to determine the standard deviation during the entire process and thus the error in the glucose measurements.
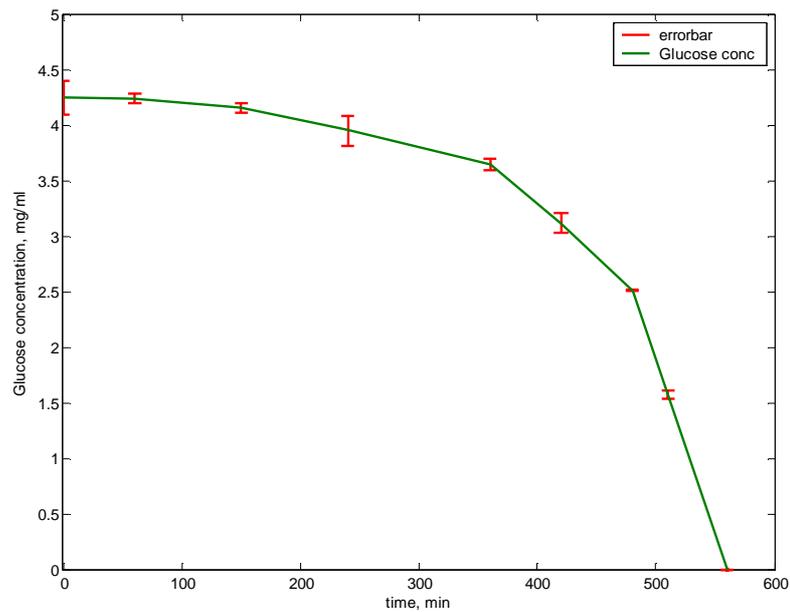
Fig.4.3. Glucose concentration measurements and standard deviation values.

Detailed information for glucose measurements taken through this single run and the percent of error respectively is provided with Table 4.1.

Table 4.1. Error estimation for glucose measurements

| Average glucose concentration, mg/ml | 4.25 | 4.24 | 4.15 | 3.95 | 3.64 | 3.12 | 2.51 | 1.58 |
|---|---|---|---|---|---|---|---|---|
| Error, % | 3.57 | 1.09 | 1.08 | 3.29 | 1.51 | 2.85 | 0.23 | 2.37 |

From Table 4.1, it is seen that the error in the glucose measurements does not exhibit any dependence on the glucose concentration over the course of the growth period. The highest error is experienced at the early stage of the growth cycle (3.57%), and the lowest error (0.23%) appeared to be close before the full exhaust of the glucose concentration. Samples in between have shown errors from 1% to 3% approximately. An average error of 1.87% on the glucose measurements was calculated for the entire process of the glucose depletion.

In summary for this section, following the above presented results, it can be concluded that the average error is relatively low within the order of 1-2% and its main source was accepted to originate from the dilution of the samples and the pipeting techniques used to carry out the dilution.

### 4.2.2. Spectroscopic measurements

Simultaneously with biomass and glucose measurements, each sample drawn from the growing culture was subjected to the following spectroscopic measurements: total diffuse reflectance (Rd), total diffuse transmittance (Td), and collimated transmittance (Tc). The order of the three measurements was arranged in a random manner over the different sample's series aiming to avoid repeated effects being transferred from the measurement onto the later extracted optical properties. As previously mentioned, the change over from one configuration to another results in an increase in measurement time and also introduces variations in the measurements. This is due to the fact that the reflectance accessory needs to be taken out from the spectrophotometer after each measurement when switching from diffuse transmittance or diffuse reflectance measurements to collimated transmittance. The switchover also includes replacement of lenses and the accessory has to be re-installed for making the next diffuse reflectance and transmittance. Further even for the changer over from diffuse reflectance to diffuse transmittance and vice-versa, lenses have to be changed. These changes could potentially lead to large variations from measurement-to-measurement. For the reason of reducing changing over time and measurement errors, as well as improving repeatability on the instrument conditions before each analysis, especially in the case of Rd which turned to be the most sensitive measurement, a measurement scheme has been created which was followed through all the experimental data collection. The sequence of performing spectroscopic measurements is shown with Table 4.2.

Table 4.2 Spectroscopic measurements arrangements

| Sample № / Numerical order | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Sample 5 |
|---|---|---|---|---|---|
| 1 | Tc | Rd | Td | Rd | Td |
| 2 | Td | Tc | Tc | Tc | Tc |
| 3 | Rd | Td | Rd | Td | Rd |

Following this scheme change over between two consecutive Rd and Td measurements have been avoided which helped to improve significantly the repeatability of the results, and helped to minimize the drift influence and the disturbance on the incident light spot. The three measurements were made using UV-Vis-NIR spectrophotometer (Cary 5000. Varian Scientific Instruments) equipped with an integrating sphere (diffuse reflectance accessory, DRA 2500) by placing the samples in a special optical glass cuvette with a pathlength of 4mm. The measurements were made over the wavelength range 950-1850nm with an average integration time of 0.4s and wavelength interval 4nm. Each scan series took approximately 45 min including change over and adjustment of the beam between the different measurements.

## 4.3. Estimation of optical properties

### 4.3.1. Data set

A total of nine growth cycles were performed using the protocol described in section 4.2.2. For the first three cultures, data was collected during the growth phase. For cultures 4-6, data was collected only during the stationary phase, and for cultures 7-9, data was collected during the decline phase. Since all the growth cycle runs were performed under the same conditions, each of the sets of three runs for each phase (growth, stationary, and decline) are essentially replicate runs. For all of these cultures, during the data collection phase, a total of five samples were taken at approximately 2

hours intervals. The progress of the cultivations was also followed by making optical density measurements, which provided information regarding the stage of the growth cycle, which, in turn, ensured that the samples were collected at the appropriate growth phase. Optical Density has been defined as the logarithm to the base 10 of the ratio of the intensity of the incident light to that of the transmitted light. All measurements for optical cell density were measured with UV-Vis NIR spectrophotometer (Cary 5000) in absorption mode at a single wavelength 600nm with a 1cm pathlength cell. Figure 4.4 presents the profile of *Bacillus subtilis* cell replication in a minimal medium for this experimental study.
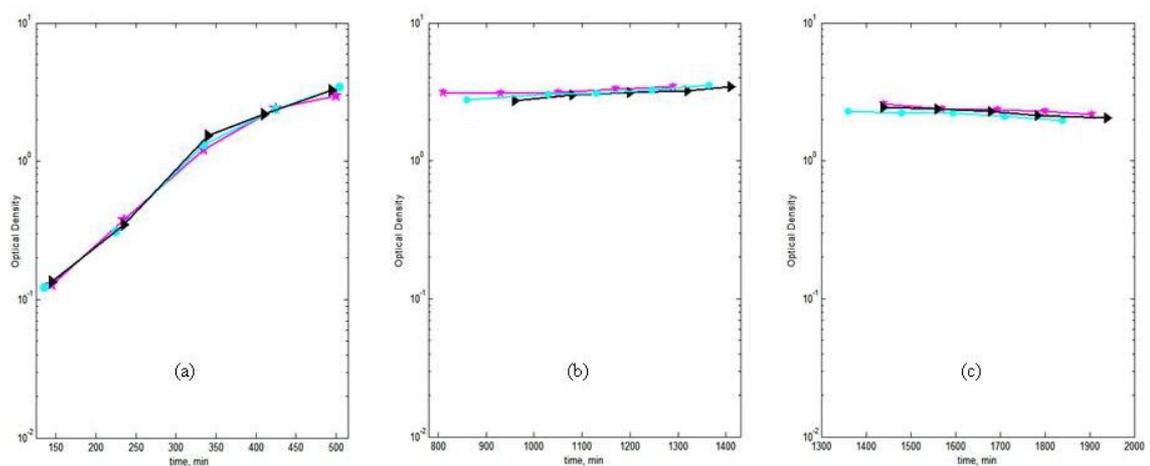


Fig.4.4 *Bacillus subtilis* growth profiles for three consecutive cultivations (Magenta&star-day1; Cyan&circle-day2; Black&triangle-day3). Shown are (a) the growth phase, (b) stationary phase and (c) the decline phase.

Figure 4.4 shows the optical density curves during the cultivation for the different phases and the points on the curves indicate when the samples for measurements were drawn from the flask. While the same recipe was used for all the cultivations, it is seen that for the three runs in each of the growth phases, there are small but distinctive variations in the optical density curves and the points where the samples were drawn are not exactly at the same time point on the curves.

Figures 4.5a-c show the extracted raw spectra plotted in absorbance units, of the samples collected during all the cultivations using the different measurement configurations viz. total diffuse reflectance ($R_d$), total diffuse transmittance ($T_d$) and collimated transmittance ($T_c$). Both Tc and Td are noisy around 1450nm water absorption peak (figure 4.5b and 4.5c). In the case of Rd the spectra appears to flatten out beyond 1400nm. The reason is that in this region, the amount of light reflected from the sample is small compared to the specular reflectance contribution from the glass cuvette. However, the information regarding changes is embedded in the spectrum and can be extracted by accounting for the specular reflectance as is done when extracting the optical properties. It is seen that the total diffuse reflectance spectra are noisier than the other two measurements. Diffuse reflectance spectra from growth, stationary, and decline phase remains similar in shape but do not fall in a particular order with respect to the growth of the suspension i.e. a clear trend in the spectra with respect to the time when the sample was taken was not evident. Figure 4.5 shows the raw near-infrared transmittance spectra of growing suspension over all three days for growth, stationary and decline phases. It is seen that the spectra shows a small change over the growth run and follows a particular order of decrease with the growth time. Despite the fact of biomass formation in the system studied, measurements show that high absorption band of water dominates the spectra at 1450nm and determines the transmission measurements.
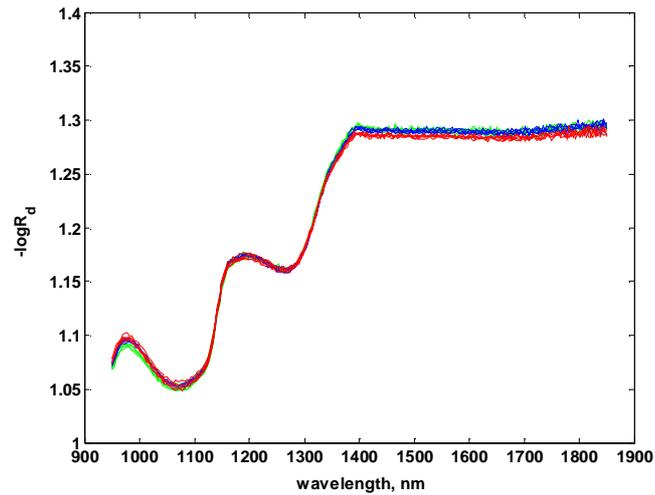
(a)
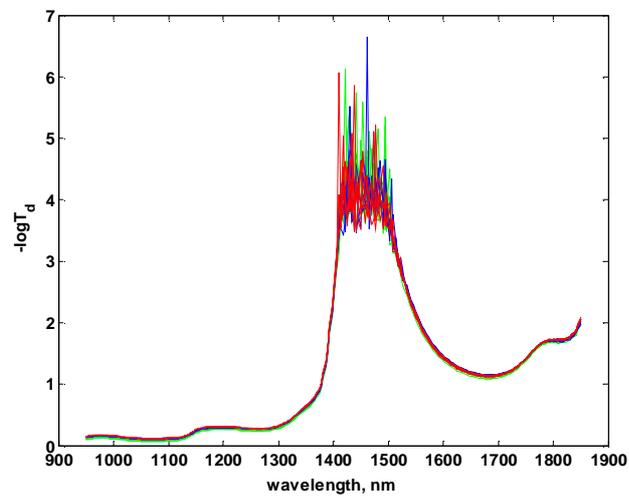


(b)
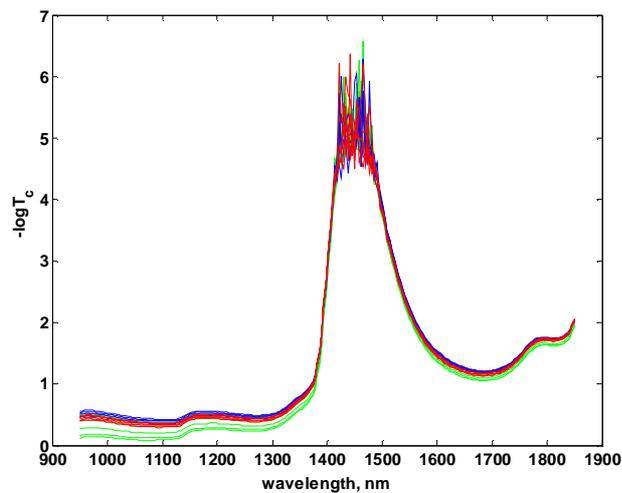


(c)

Fig.4.5. Spectra of samples taken at different stages of the growth cycle (growth phase-green colour, stationary phase-blue colour, decline phase-red colour). (a) Total diffuse reflectance, (b) total diffuse transmittance and (c) collimated transmittance.

Figure 4.5c shows the raw collimated transmittance spectra, which shows greater variation in absorption compared to diffuse and transmittance measurements.

### 4.3.2. Inversion of measurements

The measurements of diffuse reflectance, diffuse transmittance, and collimated transmittance for each sample was inverted to extract the bulk optical properties using the algorithm described in chapter 3. In addition to these measurements, to invert the RTE for the system (glass-suspension-glass) using the adding-doubling method, the refractive index of the glass cuvette and the refractive index of the suspension are also required as inputs. The refractive index of the glass cuvette was taken from the manufacturer catalogue to be 1.523[4]. The refractive index of the sample was taken as an average refractive index, 1.3362, estimated by measuring the refractive index of several samples with a refractometer (Refracto 30Px, Mettler Toledo). The samples were drawn from the culture at various stages of the growth.

The inversion algorithm begins by taking as inputs the refractive index of sample and glass and the thickness of the sample, i.e., the cuvette path length ($\ell$). In addition, initial guess values of the bulk absorption albedo ($\alpha$), optical depth ($\tau$), and the anisotropy factor ($g_o$), of the sample also have to be provided. The albedo and the optical depth are given by the following two equations,

$$\alpha = \frac{\mu_s}{(\mu_s + \mu_a)} \tag{4.1}$$

$$\tau = (\mu_s + \mu_a)\ell \tag{4.2}$$

where $\alpha$ is the bulk absorption albedo, $\tau$ is the optical depth, $\ell$ is the cuvette pathlength, $\mu_s$ is the bulk scattering coefficient, $\mu_a$ and is the bulk absorption coefficient.

The reason for using the albedo and optical depth as the parameters for the inversion instead of directly using $\mu_a$ and $\mu_s$ was because the algorithm is much more stable when the former two parameters were used. Further the adding-doubling equations are naturally cast in terms of albedo and optical depth and once these are extracted the scattering and absorption parameters can be obtained using equtions (4.1) and (4.2). The calculated total diffuse reflectance ($Rd_{calc}$) and total diffuse transmittance ($Td_{calc}$) obtained from the adding-doubling routine and the collimated transmittance obtained from using Beer's law ($Tc_{calc}$) for the input guess values of albedo and optical depth are then compared with the corresponding measured values,

$$\xi = abs(Rd_{calc} - Rd) + abs(Td_{calc} - Td) + abs(Tc_{calc} - Tc) \qquad (4.3)$$

The guess values were then updated and the iterations carried out until convergence ($\xi \leq 1.0 \times 10^{-6}$) has been achieved. This iteration was carried out using the function 'fmincon' of the MATLAB® Optimization toolbox to minimize equation (4.3). Approximate time for the inversion of measurements from one sample has been estimated to be around 20 min. Each sample covered the region 950nm-1850nm.

## 4.4. Results and discussion

Optical properties have been calculated based on the three measurements, namely Rd, Td and Tc, presented in the previous section. As already discussed earlier in this chapter, the values were obtained using an inverse adding-doubling routine.

The extracted optical properties viz. bulk scattering coefficient $\mu_s$, bulk absorption coefficient $\mu_a$ and anisotropy factor $g$ for the samples collected during the growth phase are shown in figures 4.6a-c. In these plots, the gaps in the wavelength region 1360-1550nm are because the inversion using the adding-doubling method does not converge due to the high absorbance in this region. To obtain the optical properties in this region,

sample thickness would have to be reduced by using cuvettes with path lengths smaller than 4mm. It can be seen that the relative change in $\mu_s$ is high compared to the relative changes in $\mu_a$ over the duration of the growth phase. This is expected since, in the culture considered here, the largest variation is due to the increase in biomass. While the biomass also affects $\mu_a$, the absorption due to biomass in the NIR region is small compared to the absorption due to water, therefore the relative changes due to an increase in biomass is small.

The extracted anisotropy factor also exhibits changes during the growth phase. In the early stage of the growth phase, across the wavelength region 950-1350nm, $g$ drops sharply towards zero beyond 1200nm or flattens out in this region. This is physically inconsistent. It is noticed that the wavelength beyond which it falls in this manner increases as the growth cycle progresses. It could be concluded that this effect is due to the fact that the low biomass in the initial stages of the cultivation and the resulting low scattering which characteristically falls off at higher wavelengths could be a factor in the inversion not being effective in extracting the anisotropy factor under such conditions. This effect has previously been evidenced in polystyrene-water system[5] where a sample with 0.15% of polystyrene particles by weight was considered. In that study it was shown that while the estimated values of anisotropy factor was unreliable in the region where it falls off sharply, it did not affect the estimation of $\mu_a$ and $\mu_s$. As the growth cycle progressed, the estimates of $g$ become stable over a larger span of wavelength range. Beyond 1500nm, $g$ stays around 0.8 which is the initial guess value input to the inverse adding-doubling program and thus cannot be expected to be a reliable estimate. Changing the initial guess values (0.7 and 0.9 were tried) resulted in $g$ over this wavelength range to converge to a value close to those guess values, thus reinforcing the conclusion that the estimate of the anisotropy factor in this region is not reliable though this does not affect the estimation of the bulk absorption and scattering coefficients.
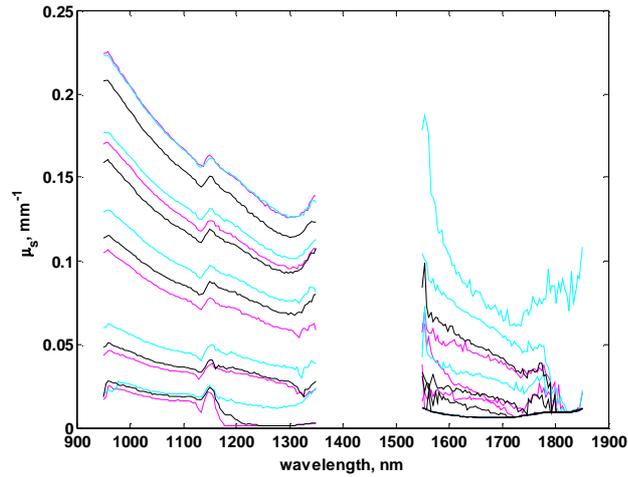
(a)
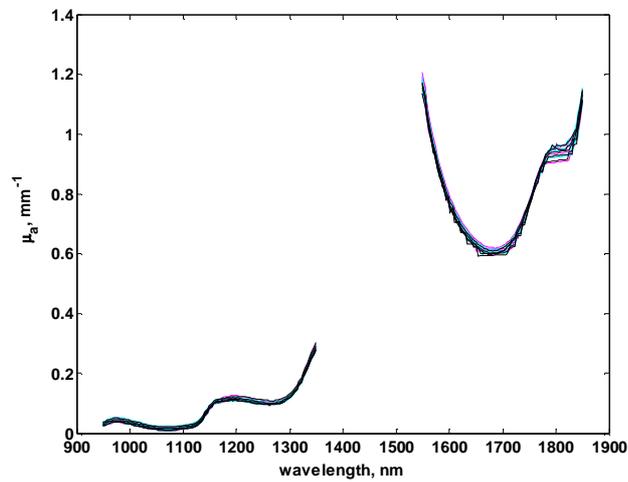


(b)
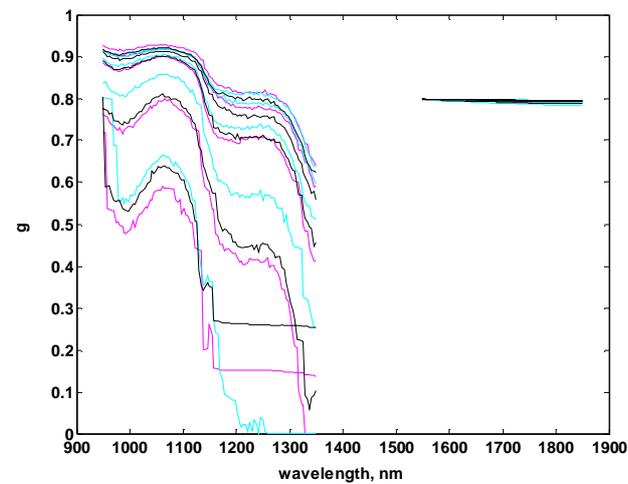


(c)

Fig. 4.6 Optical properties during growth phase for 3 cultivations (Run 1-magenta colour, Run 2-cyan colour, Run 3-black colour). (a) scattering coefficient $\mu_s$, (b) absorption coefficient $\mu_a$ and (c) anisotropy factor $g$.

Examining figure 4.6, it can be seen that the extracted optical properties are consistent over the three cultivations for the growth stage. Since the samples were taken at approximately the same point in the growth cycle during each of the three cultivations which all used the same culture medium and analytical protocol, we would expect the extracted values, if the inversion method is stable, to obtain similar optical properties at these time points. The optical properties corresponding to similar time points are indeed similar. The small differences in the magnitude can be attributed to variations due to small differences in the synchrony of cultures.

Figures 4.7a-c show the evolution of the optical properties during the stationary phase. As would be expected, due to the fact there is very little change in the biomass concentration, the optical properties show a much smaller variation compared to the growth phase. As in the case of the growth stage, the extracted optical properties are consistent over the three cultivations indicating the reliability of the inversion method.

Figures 4.8a-c show the changes in the optical properties during the decline phase. The variations in the optical properties are smaller compared to the growth phase.
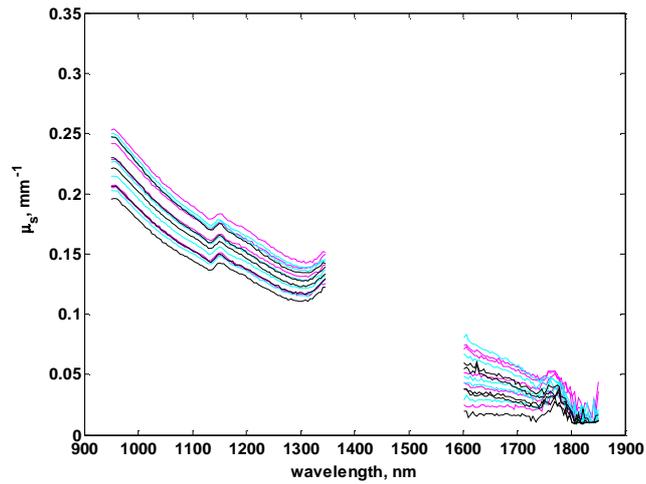
(a)
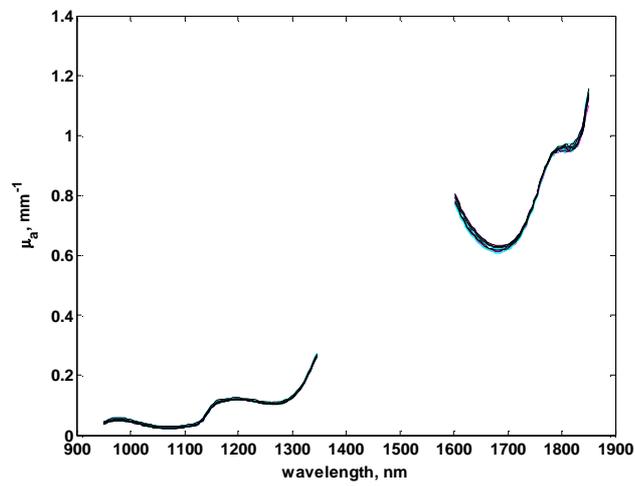


(b)
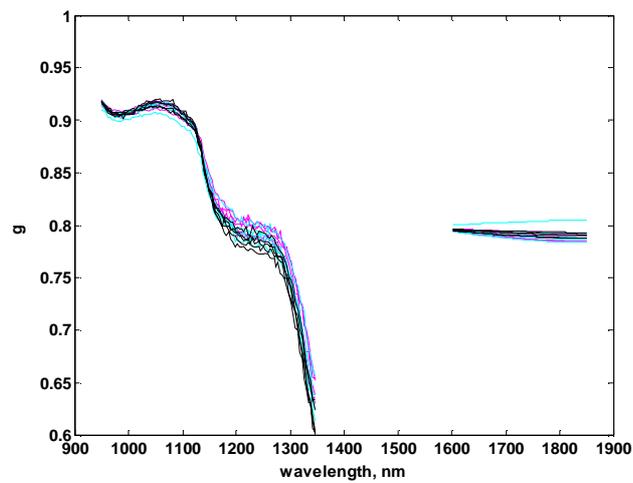


(c)

Fig. 4.7 Optical properties during stationary phase for 3 cultivations (Run 1-magenta colour, Run 2-cyan colour, Run 3-black colour). (a) scattering coefficient $\mu_s$, (b) absorption coefficient $\mu_a$ and (c) anisotropy factor $g$.
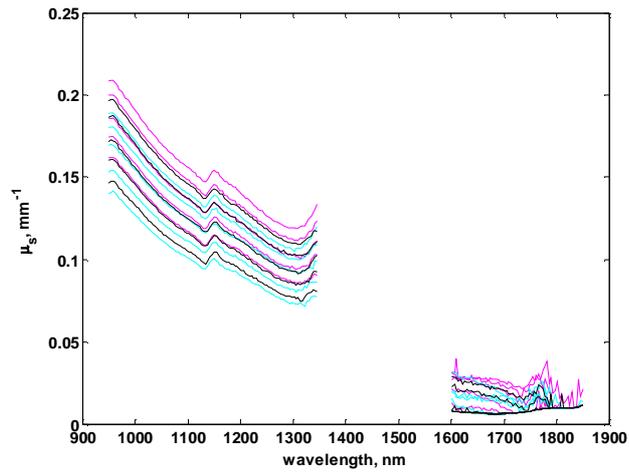
(a)
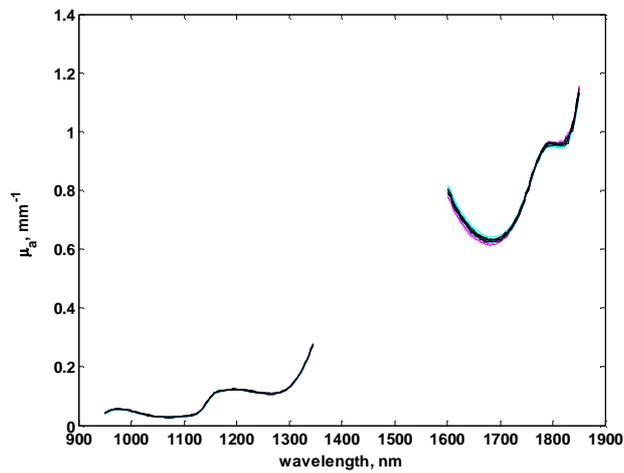


(b)
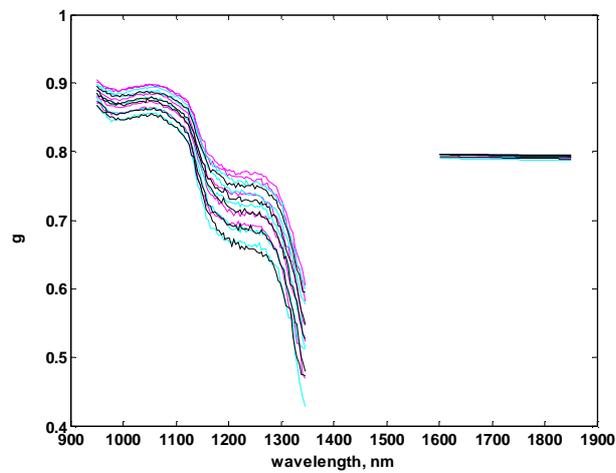


(c)

Fig. 4.8 Optical properties during decline phase for 3 cultivations (Run 1-magenta colour, Run 2-cyan colour, Run 3-black colour). (a) scattering coefficient $\mu_s$, (b) absorption coefficient $\mu_a$ and (c) anisotropy factor $g$.

In order to have a closer look at the "direction" of variations in the optical properties over the course of the cultivation, the extracted values are plotted for two wavelengths 1050nm and 1602nm. These two wavelengths were chosen to examine the behaviour of the optical properties in a scattering dominated region and an absorption dominated region respectively. In figure 4.9, the optical properties at 1050nm are plotted against the biomass concentration. It is seen from figure 4.9a that during the growth phase $\mu_s$ varies approximately linearly with the biomass except for one point with very low biomass concentration. This could be due to the low biomass concentration resulting in a combination of measurement errors and the low scattering levels exacerbating errors in the inversion since the numerical method employed is based on multiple scattering effects being dominant. In the stationary phase, the bulk scattering coefficient does not show strong trends. This is due to the fact that the changes in biomass during this period are not large, trends if any will be swamped by the errors in the measurement of biomass. This characteristic is also seen in the decline phase. Overall, for the same levels of biomass, the values of $\mu_s$ in the decline phase are less than those observed in the stationary phase, possibly indicating differences in the biomass characteristics such as the morphology of viable and non-viable cells, clumping etc. Morphological examination through the growth cycle including calculations for average sizes and cell size distribution are presented later in this section.

In the case of $\mu_a$ (figure 4.9b), the relationship with biomass indicates a slight non-linearity with a curvature observable at the lower end (around the region of 0.5µg/ml) of the biomass concentration. Again samples at the lower cell densities are more likely to be outliers for reasons mentioned when discussing the bulk scattering coefficient in the last paragraph. The values of $\mu_a$ in the stationary and decline phase do not show any clear-cut clustering that was similar to that seen when examining $\mu_s$. This could be due to the lack of significant changes in the concentration of the cellular components over these periods.
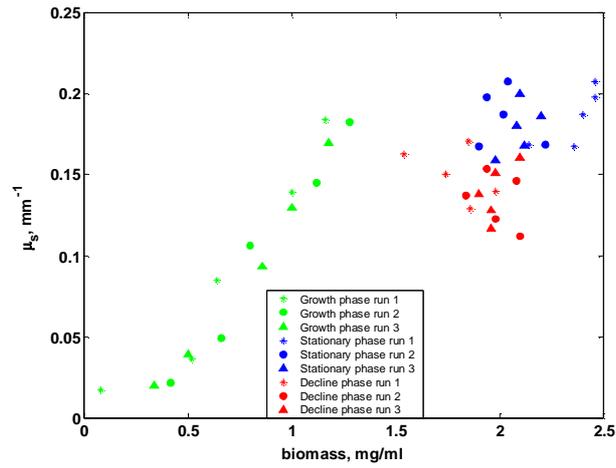
(a)
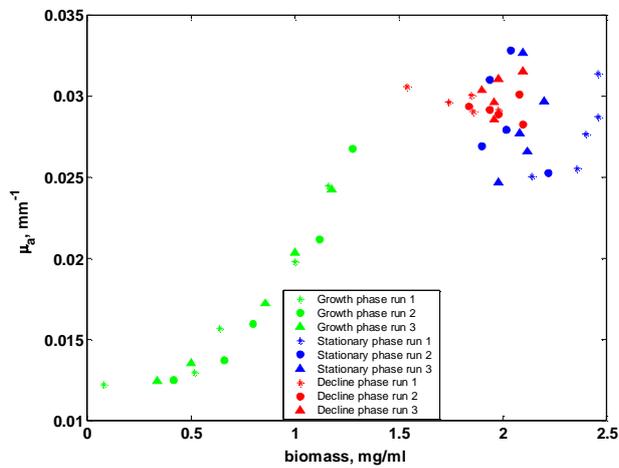


(b)
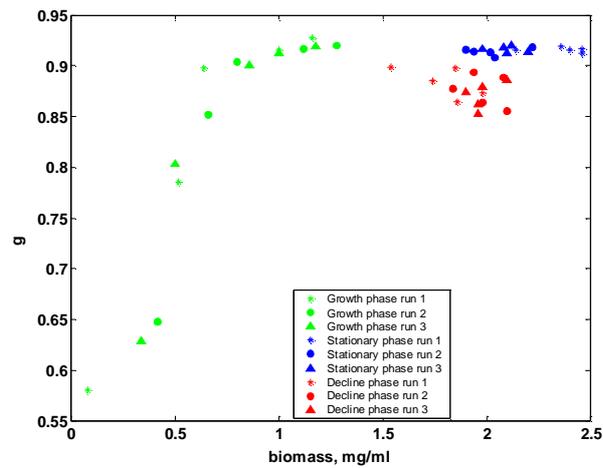


(c)

Fig. 4.9 Optical properties at 1050nm during growth, stationary and decline phase for 3 cultivations (growth phase-green colour, stationary phase-blue colour, decline phase-red colour) versus biomass. (a) scattering coefficient $\mu_s$, (b) absorption coefficient $\mu_a$ and (c) anisotropy factor *g*.

The variation of the anisotropy factor *g* with biomass is shown in figure 4.9c. It is seen that during the growth phase, *g* rises sharply and stabilizes around 0.9 where it remains during the stationary phase. In the decline phase, *g* takes on slightly lower values than observed for the stationary phase. The sharp increase in *g* initially could be due to the rapid increase in rate of cell divisions creating a larger population of freshly divided cells (of smaller sizes) compared to the cells yet to divide into smaller cells. However, more investigation is needed to verify this explanation.
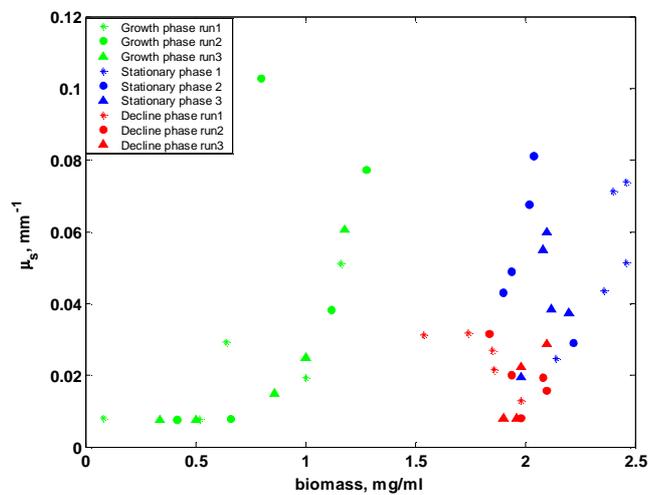


(a)
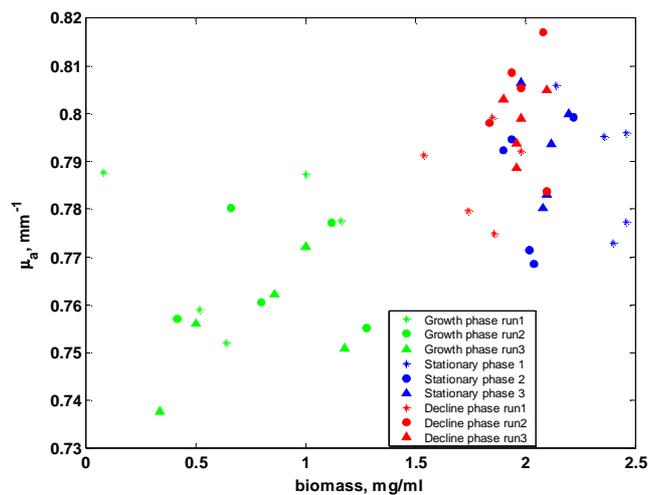


(b)

Fig. 4.10. Optical properties at 1602nm during growth, stationary and decline phase for 3 cultivations (growth phase-green colour, stationary phase-blue color, decline phase-red colour) versus biomass. (a) scattering coefficient $\mu_s$, (b) absorption coefficient $\mu_a$.

In figure 4.10, the optical properties at 1602nm are plotted against biomass concentration. The values for $g$ are not shown since, as discussed earlier, they were not reliable in this region. It is seen in figure 4.10a, $\mu_s$ shows a non-linear relationship in contrast to the trend seen at 1050nm, remaining flat initially, and then rising rapidly with biomass concentration. There are two outliers observed in this region. These are more likely to be due to convergence problems in the inversion rather than to any physical characteristics in the system. As was the case for 1050nm, there are no discernible trends in the bulk scattering coefficient with respect to biomass in the stationary and the decline phase. Compared to the decline phase, $\mu_s$ appears to be slightly higher in the stationary phase. For the bulk absorption coefficient, it is seen that the relationship between $\mu_a$ and the biomass in the growth phase is weaker compared to that at 1050nm. Thus while the absorption and scattering properties change with the biomass concentration, the extent of this variation depends on wavelength.

In order to study the morphology development of the bacterial system during the three phases and to account for possible influence on the optical properties estimation, scanning electron micrograph images defining bacterial size and shape were taken at Electron Microscopy Research Services, Medical School at Newcastle University. All obtained images were then processed with program ImageJ (image processing software; ImageJ 1.37v Wayne Rasband National Institute of Health, USA). The initial microscopic examination of *Bacillus subtilis* grown colonies which is demonstrated with picture 1, revealed that the cells grew as normal rods with regular cell division which has been found in a good agreement with literature publications elsewhere [3, 6, 7].

Images in picture 1, picture 2, and picture 3 were taken from living bacterial cells drawn from the growing culture at the same time with the samples for spectroscopic measurements and reference assays. The biological specimens were fixed chemically for at least overnight in 2% gluteraldehyde in sorenson phosphate buffer in order to prevent

growth continuing while the sample was away from the bulk culture. This step was part of standard operating procedure used in the electron microscopy research laboratory at Newcastle University for preparing bacterial population to be placed under microscope examination.
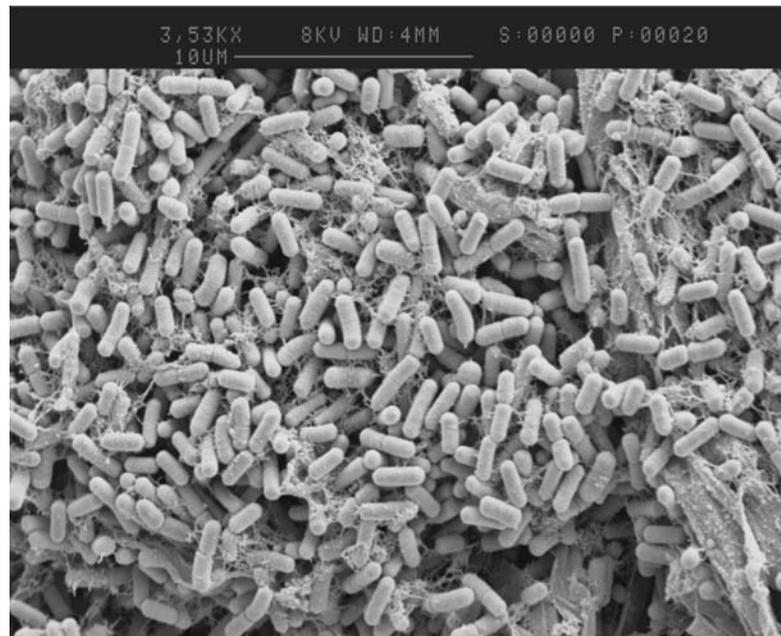


Picture 1. Scanning electron microscopy image from *Bacillus subtilis* suspension during exponential growth phase

From visual inspection of picture 1 we can distinguish two types of cells, newly formed cells and dividing 'parents' cells, all of them with the typical for *Bacillus subtilis* rod shaped profile. The cells undergoing division were actively growing at the moment were fixation has been applied and as a result some of them look ready to separate and other appear to be just beginning to form their cross-walls. Also some small amount of cell debris and particles released into the medium are noticeable. The magnification factor of the image on picture 1 is 3530X, and the grey line in the black background above the picture is equal to 10μm distance.
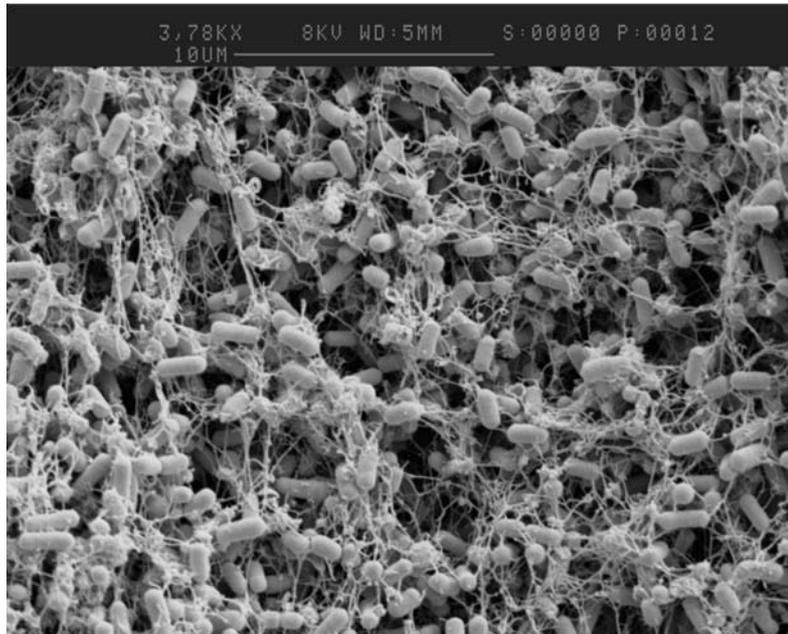
Picture 2. Scanning electron microscopy image from *Bacillus subtilis* suspension during stationary phase

Picture 2 has been taken from a sample drawn from the stationary phase. Dividing cells almost can not be seen, however slight increase on the amount on the cell debris is registered. Cell debris and additional formation different from the cells morphology will be called further in this work 'network'. The magnification factor of the image on picture 2 is 3780X, and the grey line in the black background above the picture is equal to 10μm distance.

When observing picture 3, it can clearly be seen that the amount of network formation has been increased significantly in the decline phase. This slight change in the texture (cells and additional formations) of the bacterial culture can play a role on how the absorption and scattering are distributed in the sample and could be a possible explanation for the decrease of the optical properties values in the decline phase. In addition, cells, both separate and dividing, are difficult to be distinguished due to the coverage by the 'network'. The magnification factor of the image on picture 3 is 3700X, and the grey line in the black background above the picture is equal to 10μm distance.
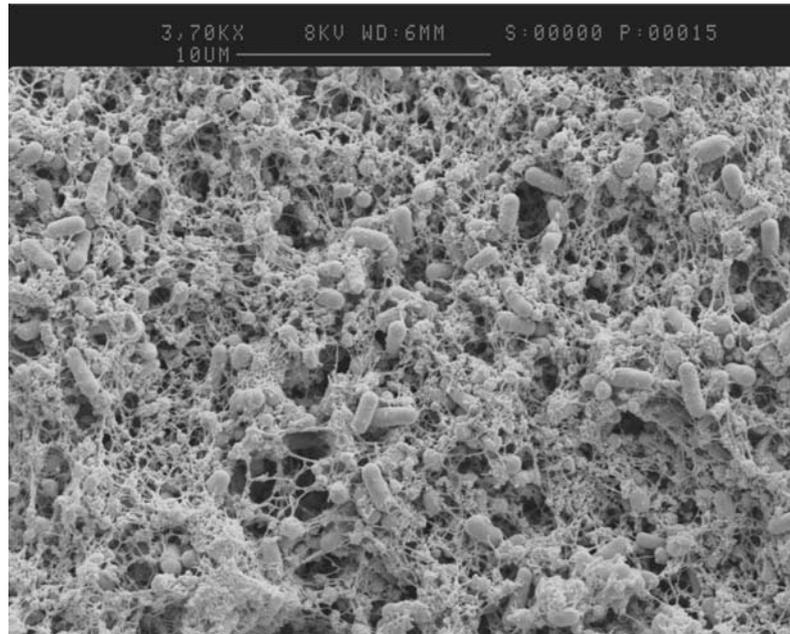
Picture 3. Scanning electron microscopy image from *Bacillus subtilis* suspension during decline phase

From all experiments (three runs for growth phase, three runs for stationary phase and three run for decline phase), approximately 1300 particles have been sorted through image software ImageJ in term of obtaining information about the size characteristics, the average particle size, and the distribution of sizes about that average. All particles studied in the samples exhibit a cylindrical shape which can be concluded from visual inspection however consist a range of sizes, especially heterogeneous in length due to the typical cell elongation before division and the small new born daughter cells. It was found that the average width is $0.7998 \pm 0.0768 \mu m$, and the average length is $2.486 \pm 0.5599 \mu m$ (figure 4.11a and 4.11b). For the length distribution, two types of cells have been considered for measurements, (1) newborn cells and (2) dividing cells. Previous studies have evidenced a significant difference in sizes between them[8] and impact on the cell size distribution between different samples could be expected [9].

(a)                                                                      (b)
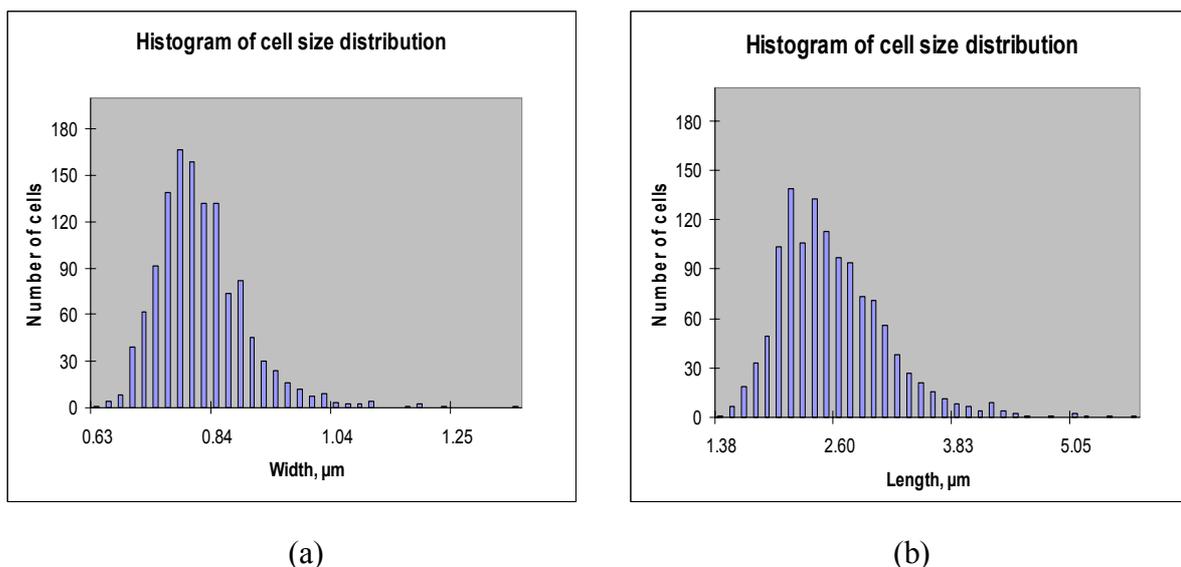
Fig. 4.11. Histogram of *Bacillus subtilis* cells size distribution. Shown are (a) the width distribution and (b) the length distribution.

In conclusion, this study presented extracted optical properties from growing bacterial suspension using the inverse adding-doubling method. The results presented here show that the absorption and scattering properties change during the cultivation of a bacteria and can be consistently extracted from measurements under multiple scattering conditions. The cultivation system considered here was a simple one in which the major change during the cultivation process was the increase in biomass. This is reflected by the fact that the greatest relative change was seen in the scattering properties during the growth phase. Since the wavelength range used in this study (950-1850nm) spans two different regimes, that is, the absorption dominated and scattering dominated regions, the pathlength of the cuvette used may not be optimal for extracting the optical properties with sufficient accuracy over the entire wavelength range. The extreme example occurs around the water absorption peak at the 1450-1550nm region. Since the main aim is to investigate whether this approach can be used to separate absorption and scattering effects so that effective models can be build for monitoring concentrations of nutrients, biomass, etc., further study were initiated to establish the effect of sample thickness on the extracted optical properties. These studies will be described in the next chapter.

**Reference**s

1. Hulst, C.V.d., *Light Scattering by Small Particles*. 1964, USA: John Wiley&Sons. 470.
2. Ishimaru, A., *Wave Propagation and Scattering in Random Media*. The IEEE/OUP Series on Electromagnetic Wave Theory. 1997: IEEE Press, Oxford University Press. 574.
3. Harwood, C.R. and S.M. Cutting, *Molecular Biology Methods for Bacillus*. 1990: John Wiley & Sons Ltd.
4. Hellma, *Hellma Genaral Catalogue*. 2002: Hellma GmbH & Co. KG.
5. Saeys, W., M. Velazco-Roa, S. Thennadil, H. Ramon, and B. Nikolai, *Optical Properties of Apple Skin and Flesh in the Wavelength Range from 350 to 2200nm*. APPlIED OPTICS, 2008. **47**(7): p. 908-919.
6. Parry, J.M., P.C.B. Turnbull, and J.R. Gibson, *A colour atlas of Bacillus species*. 1983: Wolfe Medical Publications Ltd.
7. Gordon, R.E., W.C. Haynes, and C.H.-N. Pang, *The genus Bacillus*. 1973, Washington: Agricultural Research Service, U.S. Dept. of Agriculture.
8. Donachie, W.D., K.J. Begg, and M. Vicente, *Cell length, cell growth and cell division*. Nature, 1976. **264**(25): p. 328-332.
9. Erlebach, C.E., P. Illmer, and F. Schinner, *Changes of cell size distribution during the batch culture of Arthrobacter strain PI/1-95*. Antonie van Leeuwenhoek, International Journal of General and Molecular MIcrobiology, 2000. **77**: p. 329-335.

# Effect of sample thickness on the estimated optical properties

## 5.1. Introduction

In the previous chapter, the optical properties were extracted using a 4mm sample thickness. It was seen that there were regions of the spectrum (1350-1550nm) where due to high noise and/or low signal there were problems in convergence as a result of which reliable estimates of the optical parameters were found difficult to obtain. One approach to improve the estimates is to identify the optimal pathlength (sample thickness) which would provide accurate estimates of the optical parameters. Further, since the ultimate goal of extracting the optical properties is to build robust calibration models for estimating the concentration of various constituents in the sample, it is important to know how the sample thickness will affect the estimation of the optical properties and thus the performance of the corresponding calibration models. This chapter describes the experiments and analysis conducted to examine the effect of sample thickness on the estimation of the optical properties. Further analysis is also carried out to examine how the relationship between the biomass and the optical properties as well as the reflectance and transmittance are affected by the changes in the sample thickness.

In order to study this effect and its influence on the raw measurements and on the extracted optical properties, further spectroscopic measurements and corresponding assays on *Bacillus subtilis* suspensions were carried out by varying the thickness of the sample. Additional measurements with cuvettes of 2mm and 10mm were used for reflectance and transmittance data collection. The spectra was collected while the culture was going through growth, stationary and decline phases following the same protocol as was done for the 4mm sample thickness studies. The raw spectral data was then inverted

to obtain the absorption and scattering coefficients using the inverse adding-doubling technique. For 2mm and 10mm sample thickness studies, measurements were performed in the region 950nm-1850nm, which is the same wavelength range used for the studies with 4mm sample in order to allow comparison.

## 5.2. Data set

The spectral data for the three sample thicknesses are presented in three subsections. The first section 5.2.1 presents diffuse reflectance data obtained with 2mm, 4mm and 10mm samples, section 5.2.2 deals with diffuse transmittance data extracted with 2mm, 4mm, and 10mm samples, and finally 5.2.3.presents collimated transmittance data. Spectral information was extracted from growth, stationary and decline phases, and the experiments were repeated three times. The entire set of data could be presented, however spectral shapes were found similar, and for this reason, data set from only one cultivation run are shown in figures. The aim was to illustrate, what the spectral data from a single cultivation run looks like, and to allow better visualization of the results. Measurements with 4mm sample thickness were reported in chapter 4 and are repeated again in this chapter along with results for 2mm and 10mm, to facilitate comparison and to avoid numerous referrals to the previous chapter. All measurements were conducted with Cary 5000 (see chapter 3 for detailed description).

### 5.2.1. Total diffuse reflectance (Rd) measurements

Figure 5.1 displays raw diffuse reflectance spectra (in absorbance units), extracted for different sample thicknesses (2mm (green), 4mm (red), and 10mm (blue)), where samples were collected at various points of the cultivations. General similarities in spectral curves are observed, however intensities are different. Measurements with 4mm sample thickness have the highest diffuse reflectance, 10mm the lowest, and

measurements with 2mm between the two. The fact that changes in reflectance with pathlength are not ordered in relation to the sample thickness is surprising. This could be possibly due to the different type of glass used to build the cuvettes: the 4mm cuvette was made with glass different from the one used for 2mm and 10mm. As a consequence, the two types of glass having differences in the refractive indexes, may affect the way light is reflected. However, further studies are needed for this to be fully understood. From figure 5.1 it can be seen that variations from one sampling point to another sampling point are very small and while some noise in the spectra is visible, the magnitude of this noise is smaller than the variation over the cultivation run. The high frequency nature of this noise suggests that a significant portion of it can be removed by filtering if necessary. The diffuse reflectance spectra, taken with 2mm cuvette, contains three pronounced valleys (at 1070nm, 1280nm, and 1700nm), whereas in the case of 4mm and 10mm the third valley (at 1700nm) has flattened out. This difference is most likely due to the fact that when the sample thickness is higher, the photons which have penetrated farther into the sample have lesser probability of being reflected back especially in the high absorption region (>1400nm). As a result, the dominant part of the reflected intensity recorded by the detector is due to the portion of incident photons reflected by the front surface of the glass due to Fresnel reflection. The intensity of light that has undergone Fresnel reflection varies weakly with wavelength and thus a flat spectrum is obtained in this spectral region. For the sample thickness of 2mm, the photons which would have otherwise been absorbed at greater depths undergo Fresnel reflection at the back end of the cuvette and thus have greater probability of making it back to the detector.

While observing reflectance spectra obtained for 10mm sample thickness (blue), differences in the intensity levels can be easily distinguished in the region 950nm - 1150nm. The curves correspond to different sampling points from the growth phase.

For samples of 2 mm and 4mm thickness, such significant difference was not registered.
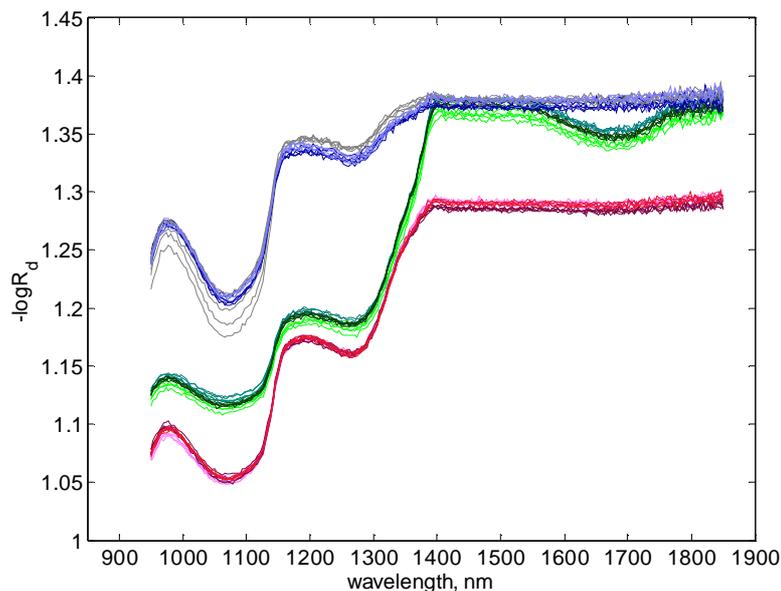


Figure 5.1 Diffuse reflectance spectra taken at different stages of the cultivation and at different samples thickness. Green-**2mm** (light green-growth phase, grey green-stationary phase, and dark green-decline phase), red-**4mm** (rose-growth phase, red-stationary phase, and dark red-decline phase), and blue-**10mm** (grey-growth phase, blue-stationary phase, and dark blue-decline phase).

Changes in $-logRd$ are also associated with the increase of cell density and increase in the scattering level, where variations due to scattering change the average pathlength that photons travel through the sample.

In order to have a closer look at this variation over the course of a cultivation, values for diffuse reflectance (-logRd) were plotted versus biomass at a single wavelength. Two wavelengths were chosen: 1050nm (figure 5.2) and 1602nm (figure 5.3). The choice for the two wavelengths was made on the same basis as with the analysis for 4mm cuvette (presented earlier in chapter 4). The goal was, to examine trends in measurements for the scattering dominated region and absorption dominated region. Figure 5.2 and figure 5.3 display data from three separate cultivations with data collected from growth, stationary and decline phases. The different symbols on the plots distinguish between different

cultivation runs, and colours differentiate between growth (green), stationary (blue), and decline (red) phases.

In figure 5.2, raw reflectance measurements (-logRd) obtained for different sample thicknesses are plotted versus biomass.



(a)



(b)



(c)
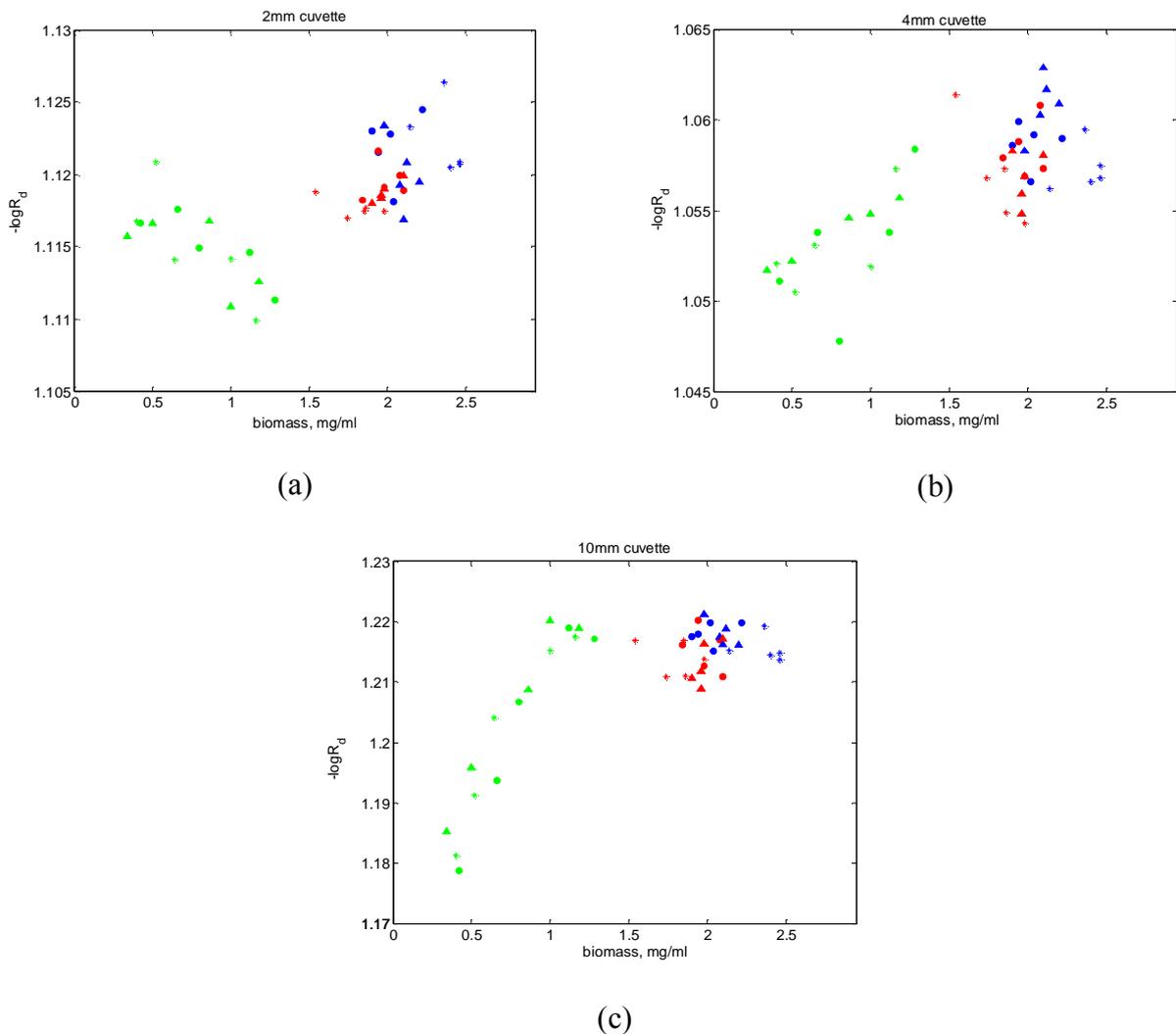
Figure 5.2 Total diffuse reflectance measurements (-logRd) at 1050nm during **growth** (green), **stationary** (blue), and **decline** phase (red) for three cultivations [(star symbols) run 1, (circle symbols) run2, (triangle symbols) run 3] versus biomass. **(a)** Total diffuse reflectance-2mm, **(b)** Total diffuse reflectance-4mm, **(c)** Total diffuse reflectance-10mm.

Measurements with 2mm sample thickness (figure 5.2a), presented in logarithmic form, show a clear trend of decrease over the course of the growth phase. The same fact has been evidenced by Ge *et al*.[1], while monitoring cell density from fed-batch

fermentation of *Saccharomyces cerevisiae*. Decrease in $-\log Rd$ versus cell mass concentration was illustrated by the authors at a single wavelength (810nm). It needs to be clarified for this case that the experimental setup and the microorganism used for the particular study was different from the one used in the presented work. In their study, the diffuse reflectance was collected with a fiber optic bundle on the outside of a glass vessel fermenter, whereas in this work the spectra were collected with an integrating sphere setup. Although two different techniques were used, the same trend for diffuse reflectance has been seen in the scattering dominated region, while considering samples with 2mm thickness from this work with that reported in the above mentioned publication.

For measurements with 4mm (figure 5.2b) and 10mm samples (figure 5.2c) collected during the growth phase, $-\log Rd$ values increase with the increasing biomass which is opposite of the trend exhibited by data from 2mm sample thickness (figure 5.2a). To summarize the information obtained for the growth phase, it was found that sample thickness not only affects the levels of reflected intensity but also affects the nature of relationship of the reflected intensity with respect to biomass concentration.

Moving on to the stationary phase, the diffuse reflectance measurements for 2mm, 4mm, and 10mm sample thicknesses do not show strong trends with respect to biomass, and biomass itself does not change very much during this phase. Similar characteristics were exhibited when results from decline phase were displayed.

Values ($-\log Rd$) from the absorption dominated region (1602nm) are shown as a function of biomass from figure 5.3a through 5.3c. Data points are very scattered on all figures. The relation spectra-biomass, evidenced in the scattering dominated region, does not hold when observing the absorption region. For 2mm samples (figure 5.3a) $-\log Rd$ decreases with the increase in biomass but the relationship is very weak. Values collected from

stationary and decline phases are clustered together as it was in the case of 1050nm and a clear relationship between biomass concentration and -logRd is not present. When observing measurements from 4mm (figure 5.3b) and 10mm (figure 5.3c) sample thicknesses in the growth phase, –logRd exhibits slightly higher values compared to values obtained for stationary and decline phases. For the same figures 5.3b and 5.3c, a weak tendency of decreasing –logRd is seen over the course of the entire cultivation run. This effect is opposite of the trend that was seen for measurements at 1050nm where logarithmic reflectance was found to increase with the increase of biomass.



(a)



(b)



(c)
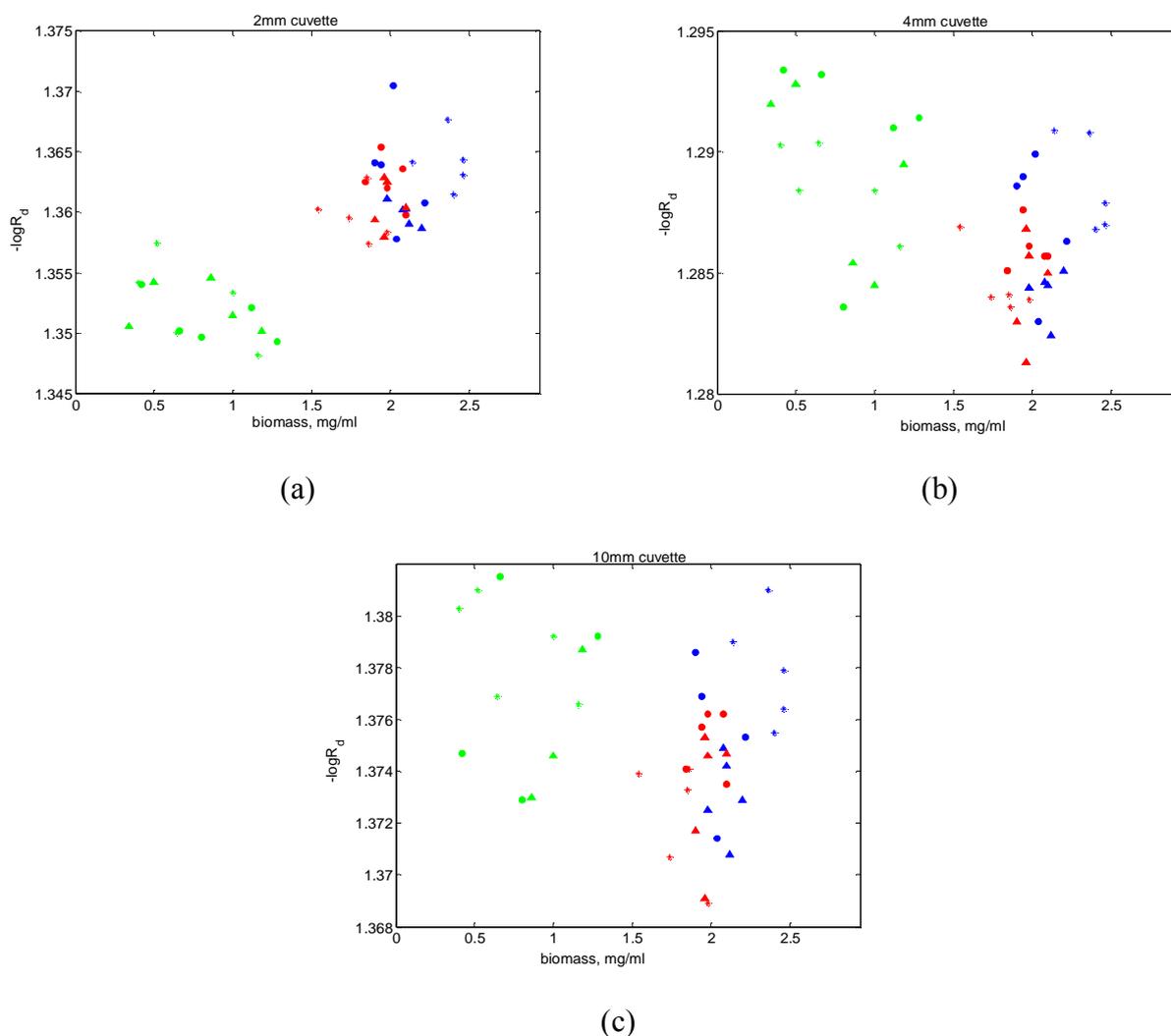
Figure 5.3 Total diffuse reflectance measurements (-logRd) at 1602nm during **growth** (green), **stationary** (blue), and **decline** phase (red) for three cultivations [(star symbols) run 1, (circle symbols) run 2, (triangle symbols) run 3] versus biomass. **(a)** Total diffuse reflectance-2mm, **(b)** Total diffuse reflectance-4mm, **(c)** Total diffuse reflectance-10mm.

It is evident that the relationship between reflectance with respect to biomass is affected both, by the sample thickness and by the wavelength region (whether absorption or scattering dominates the measurement). Based on results (figure 5.1) and analysis (figure 5.2 and figure 5.3) shown in this section, it can be concluded that changes in the sample texture are well captured with measurements obtained from the short wavelength region. Despite the fact that the actual change of −logRd for the range of the measurements is, in this bacterial system, relatively small, it has been demonstrated that general differences exist in the spectra, due to sample thickness. Spectra obtained using 2mm sample thickness show some sensitivity to variations during the bacterial growth, but it is the 10mm sample thickness that is most sensitive amongst the three sample thicknesses studied. In the higher wavelengths, the 2mm sample thickness captures detailed information (a third pronounced valley has been registered) and shows some degree of sensitivity to the variations of the bacterial growth.

### 5.2.2. Total diffuse transmittance (Td) measurements

Figure 5.4 displays raw diffuse transmittance spectra (in absorbance units), extracted for 2mm, 4mm, and 10mm sample thicknesses, and collected during the cultivations. It is seen that measurements with 2mm thickness have the highest values, followed by measurements with 4mm, and the lowest values are for 10mm sample. Diffuse transmittance measurements are very noisy around the 1450nm water peak for both 4mm and 10mm sample thicknesses, whereas measurements with 2mm sample thickness show very little noise, slightly visible only on the top of the peak. For samples of 10mm thickness, the spectral region beyond 1350nm is highly dominated by noise indicating that it would not be possible to obtain useful information from this region at this sample thickness. From visual inspection, it can be seen that the spectra show a small but significant change throughout the cultivation phases, especially for measurements

with 2mm and 4mm samples. In the case of 10mm sample thickness, in the region 950nm-1350nm (figure 5.4b), the transmittance (-logTd) is seen to have a much higher variation over the course of the cultivation run.



(a)



(b)
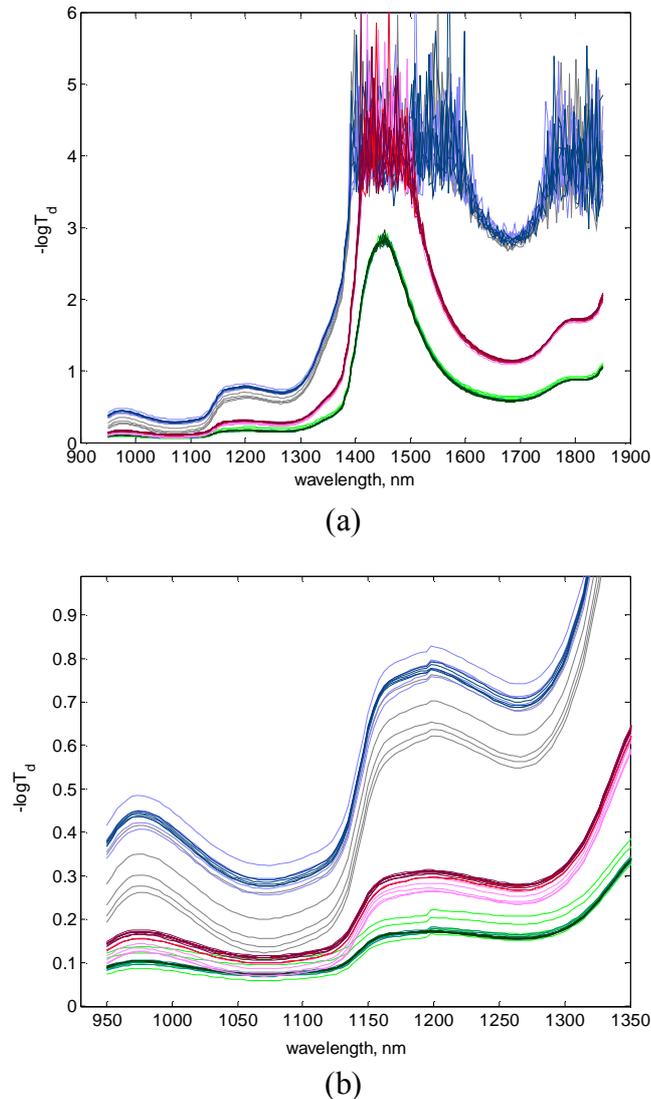
Figure 5.4 Diffuse transmittance spectra taken at different stages of the cultivation and at different samples thickness. Green-**2mm** (light green-growt**h** phase, grey green-stationary phase, and dark green-decline phase), red-**4mm** (rose-growth phase, red-stationary phase, and dark red-decline phase), and blue-**10mm** (grey-growth phase, blue-stationary phase, and dark blue-decline phase). **(a)** Measurements from 950nm to 1850nm, **(b)** magnified view of the region between 950nm and 1350nm.

In order to study these variations in greater detail, values for diffuse transmittance were plotted versus biomass at wavelengths: 1050nm (figure 5.5) and 1602nm (figure 5.6). .

(a)

(b)



(c)
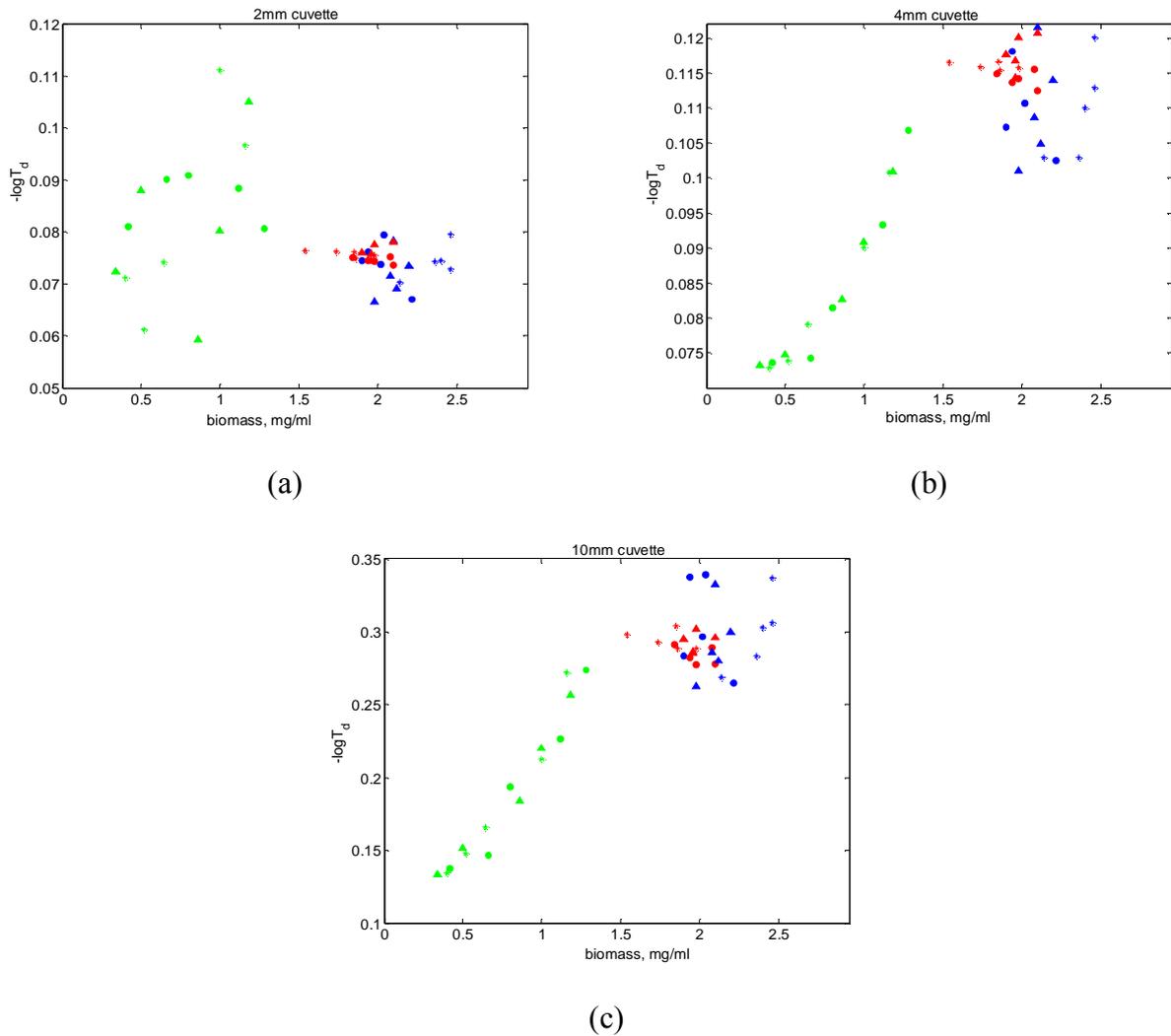
Figure 5.5 Total diffuse transmittance measurements (-logTd) at 1050nm during **growth** (green), **stationary** (blue), and **decline** phase (red) for three cultivations [(star symbols) run 1, (circle symbols) run 2, (triangle symbols) run 3] versus biomass. **(a)** Total diffuse transmittance-2mm, **(b)** Total diffuse transmittance-4mm, **(c)** Total diffuse transmittance-10mm.

At 1050nm wavelength, spectral values (-logTd), obtained with 2mm sample thickness (figure 5.5a) during the growth phase do not indicate a clear trend with the increase in biomass. While observing -logTd values for 4mm (figure 5.5b) and 10mm samples (figure 5.5c) in the growth phase, it can be seen that transmittance increases with biomass. Logarithmic diffuse transmittance indicates a slight nonlinearity with a curvature observable at the lower end (in the region of 0.5mg/ml) as the biomass concentration increases. The values (–logTd) in the stationary phase are very fuzzy for

the region 2-2.5 mg/ml biomass and clear relation could not be observed. Throughout the decline phase, points are grouped in a small cluster in the region of 2mg/ml biomass.

Figure 5.6 in the absorption dominated region (1602nm), shows dependence of $-\log Td$ with biomass.



(a)

(b)

(c)
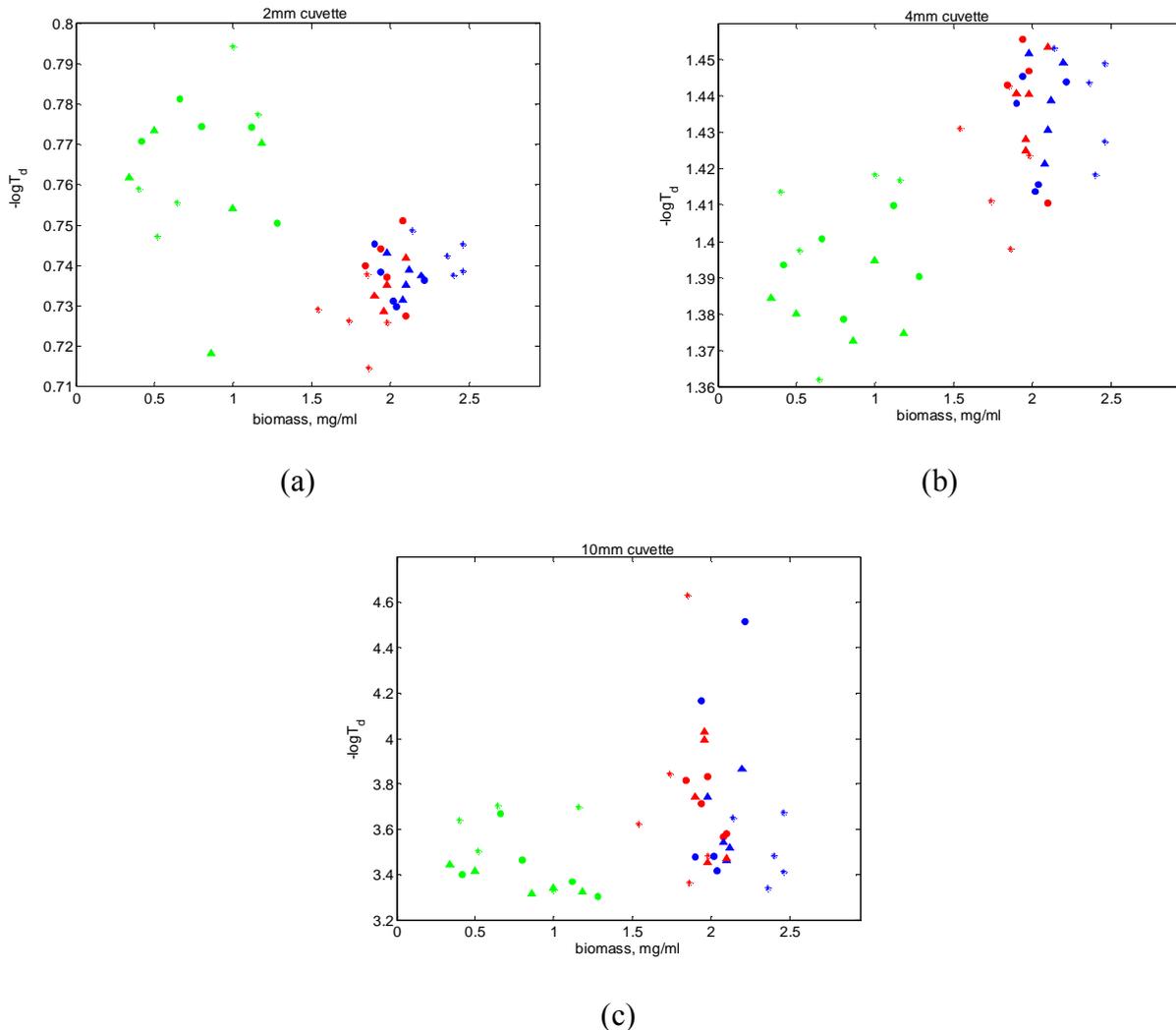
Figure 5.6 Total diffuse transmittance measurements (-logTd) at 1602nm during **growth** (green), **stationary** (blue), and **decline** phase (red) for three cultivations [(star symbols) run 1, (circle symbols) run 2, (triangle symbols) run 3] versus biomass. **(a)** Total diffuse transmittance-2mm, **(b)** Total diffuse transmittance-4mm, **(c)** Total diffuse transmittance-10mm.

Values obtained for 2mm samples thickness (figure 5.6a) show no clear relation between transmittance and biomass. However, when all the three phases are considered together, one could argue that a general trend of decreasing $-\log Td$ with increasing biomass can be

seen. The highest values correspond to the growth phase (green), followed by decrease of –logTd in the stationary phase (blue), and the lowest values were measured throughout the decline phase. For measurements taken with 4mm sample thickness (figure 5.6b), data are fuzzy, repeating the picture for 2mm sample, however a general trend of increase in -logTd with increasing biomass could be seen from the growth phase toward the stationary phase. In the case of 10mm sample thickness (figure 5.6c) no discernible trend can be seen.

Overall, diffuse transmittance spectra show small changes over the course of cultivation that can be most clearly seen for sample of 10mm pathlength in the scattering dominated region. However the longer pathlength leads to very high absorption (–logTd>3), making it unsuitable for analysis beyond 1350nm. In terms of biomass concentration, the –logTd values in the scattering dominated region shows strong correlation in the growth phase, particularly for the 4mm and 10mm sample thicknesses. Thus from the point of view of the diffuse transmittance measurements, the 4mm pathlength provides the best overall measurements when the entire wavelength region is considered in terms signal-to-noise and in terms of correlations with biomass.

### 5.2.3. Collimated transmittance (Tc) measurements

Collimated transmittance spectra (in absorbance units), extracted for 2mm, 4mm, and 10mm sample thicknesses, collected during the cultivations, is displayed in figure 5.7. It is seen that for the region around the water absorption peak (1450nm), the level of noise increases with the increase of the sample thickness (2mm sample thickness at bottom, 10mm sample thickness at top). For 4mm and 10mm samples, the noise in this region is too high to provide any useful information, but in the case of 2mm sample thickness, the spectra are very well pronounced with slightly visible noise.
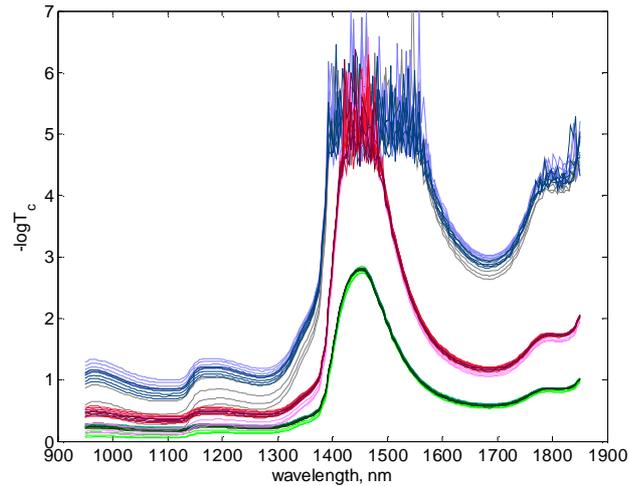
Figure 5.7 Collimated transmittance spectra taken at different stages of the cultivation and at different samples thickness. Green-**2mm** (light green-growth phase, grey green-stationary phase, and dark green-decline phase), red-**4mm** (rose-growth phase, red-stationary phase, and dark red-decline phase), and blue-**10mm** (grey-growth phase, blue-stationary phase, and dark blue-decline phase).

In order to have a closer look at the direction of the variations, values for –logTc were plotted versus biomass at 1050nm (figure 5.8) and 1602nm (figure 5.9). In the scattering dominated region (1050nm, figure 5.8), it can be seen that during the growth phase, -logTc holds a clear linear relation with biomass and rises sharply in the region 0.5-1.5mg/ml biomass concentration. This fact is evidenced for measurements with the three sample thicknesses. In the stationary phase, -logTc values are slightly higher when compared to the ones obtained during the decline phase. Sampling points in both, stationary and decline phases are grouped in two clusters with almost identical form for the three different sample thicknesses.

(a)

(b)



(c)
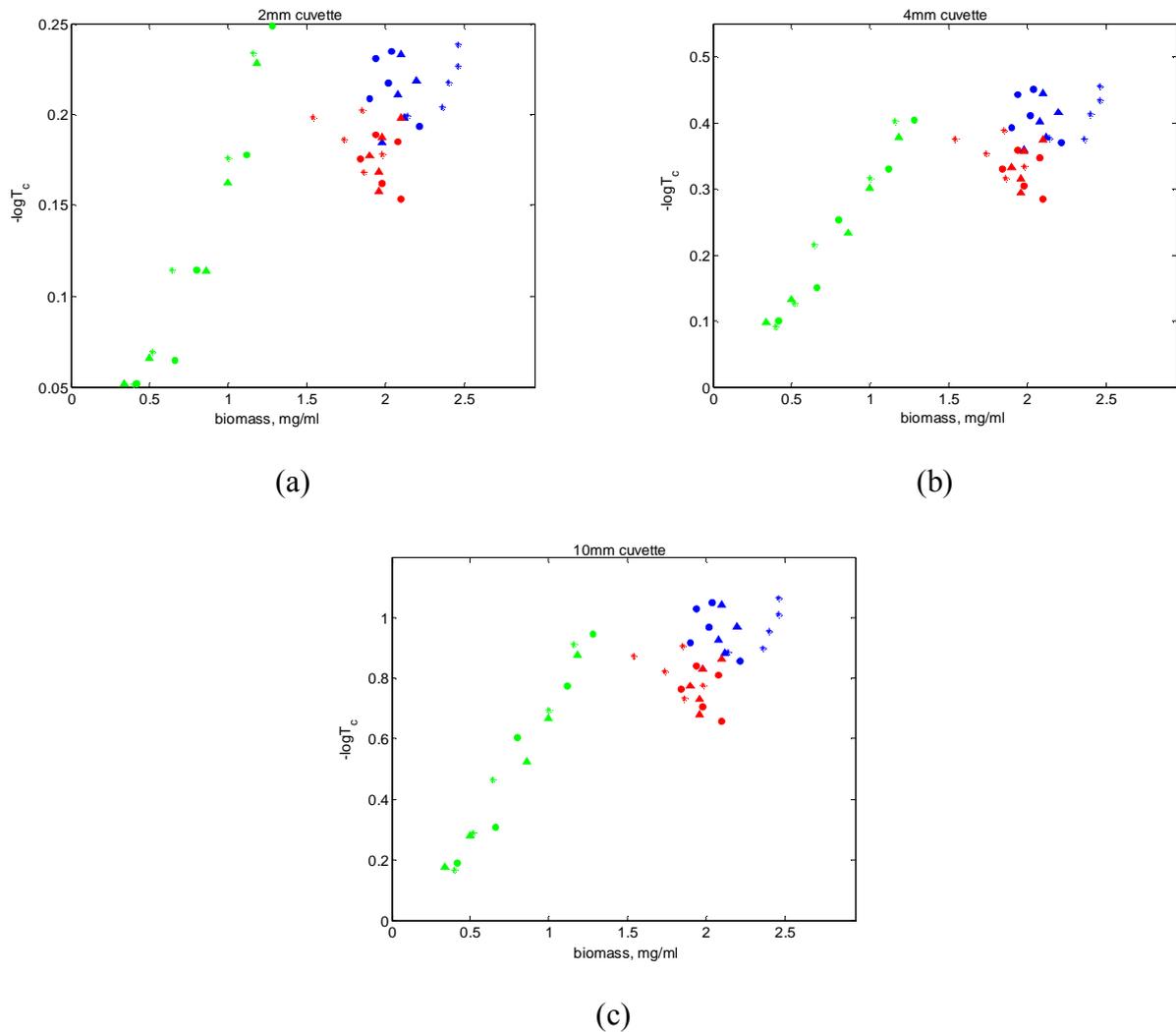
Figure 5.8 Collimated transmittance measurements (-logTc) at 1050nm during **growth** (green), **stationary** (blue), and **decline** phase (red) for three cultivations [(star symbols) run 1, (circle symbols) run 2, (triangle symbols) run 3] versus biomass. **(a)** Collimated transmittance-2mm, **(b)** Collimated transmittance-4mm, **(c)** Collimated transmittance-10mm.

Similar characteristics have been found in the absorption dominated region (1602nm), shown in figure 5.9. In the growth phase, -logTc is seen to vary linearly with biomass for 4mm and 10mm sample thicknesses. A little difference is noticed on the 2mm sample thickness (figure 5.9a): departure from linearity can be seen, with a slight curvature in the middle of the curve, traced with experimental points from the growth phase.

(a)

(b)



(c)
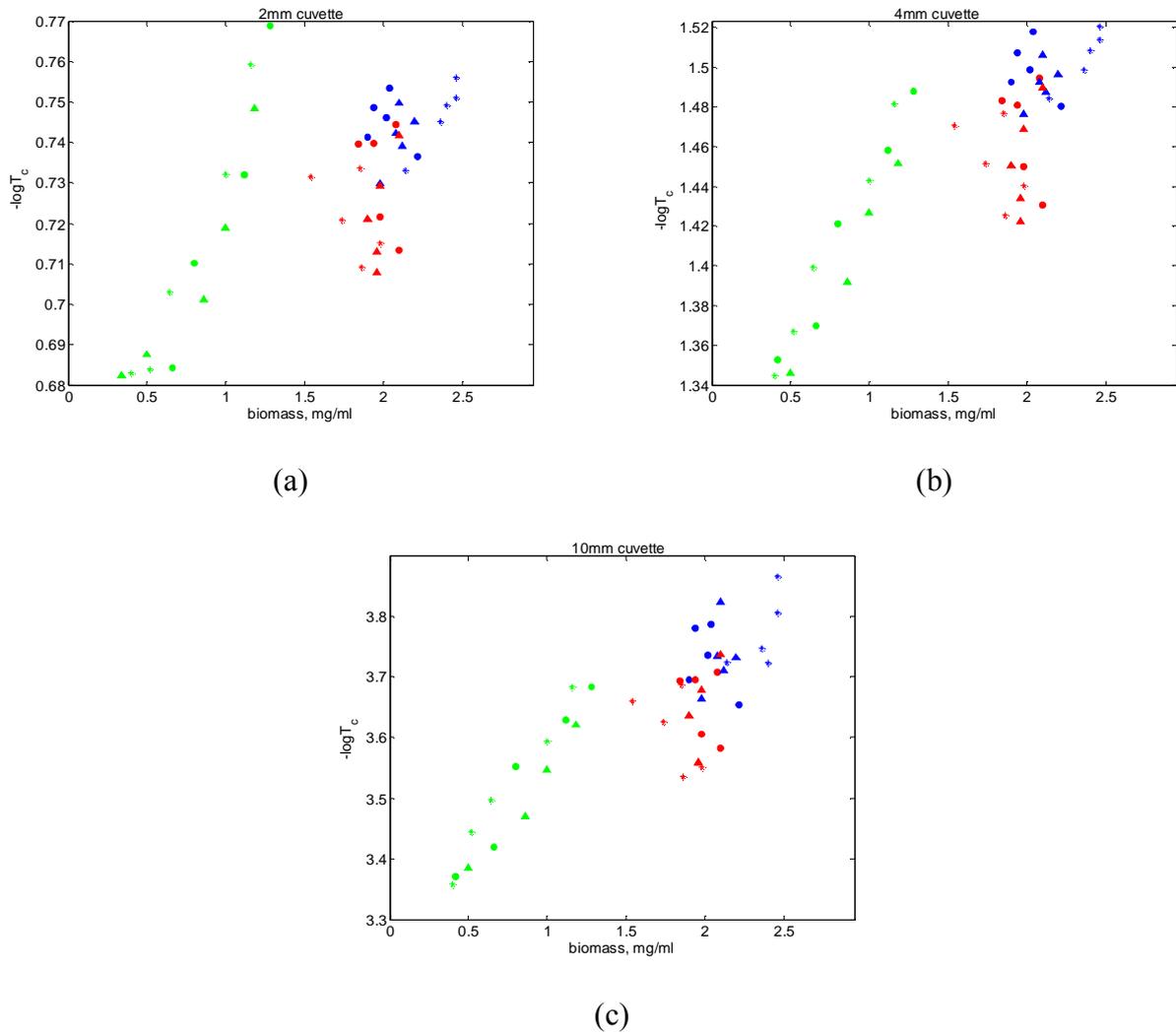
Figure 5.9 Collimated transmittance measurements (-logTc) at 1602nm during **growth** (green), **stationary** (blue), and **decline** phase (red) for three cultivations [(star symbols) run 1, (circle symbols) run 2, (triangle symbols) run 3] versus biomass. **(a)** Collimated transmittance-2mm, **(b)** Collimated transmittance-4mm, **(c)** Collimated transmittance-10mm.

In the stationary phase, -logTc does not show strong trend, most probably due to the fact that changes in biomass are almost negligible in this phase, and if any trend exists, it would be swamped by the error in the biomass determination. The same characteristic is also seen in the decline phase. Overall, values of –logTc in the decline phase are lower than those observed in the stationary phase for the investigated sample thicknesses, most probably indicating changes in the morphology of the culture.

In summary it can be concluded that collimated transmittance spectra obtained with three different sample thicknesses, shows significant changes over the course of cultivation that can be clearly seen in both scattering dominated and absorption dominated regions. In the region of the water absorption peak, spectra obtained with longer pathlength (4mm and 10mm) depicts high absorption values (-logTc>3), which leads to increase in the noise level and makes experimental data from this region unusable for analysis. When the entire wavelength region is considered, all three samples can be used successfully for data collection, however if preference needs to be given, then 4mm and 10mm samples can be chosen. This decision is based on the analysis showing that experimental data from these two sample thicknesses can provide good correlation with biomass. Data obtained with 2mm cuvette also show a high potential in terms of correlation and this shorter pathlength can also be used for measurements of collimated transmittance.

## 5.3. Inversion of measurements and optical properties estimation

Measurements obtained with sample thicknesses of 2mm and 10mm were inverted in order to calculate optical properties using the inverse adding-doubling method in the same manner as was previously done for the case of 4mm sample thickness. For the inversion technique, the refractive index of the glass cuvette was provided by the manufacturer for both 2mm and 10mm cuvettes. Its value is equal to 1.458[2], which is different from the case of 4mm where the refractive index is equal to 1.523[2]. The refractive index (1.3362) of the suspension that was used in chapter 4 remains the same for these calculations. The tolerance for convergence was set to $\xi \leq$ 1.0 x 10$^{-6}$, and function 'fmincon' of the MATLAB® Optimization toolbox was used for the computations. The time taken for the inversion of one spectrum for 2mm and 4mm samples was approximately 20 min, whereas in the case of 10mm pathlength the time was estimated at approximately 10min. The difference in the calculation time is due to

the reduced number of wavelengths used for the inversion in the case of 10 mm sample. The complete data set comprised wavelengths from 950nm up to 1350nm only. The reason for the wavelengths to be truncated is due to the low spectral signal reaching the detector (for the region 1350nm-1850nm), causing problems in the inversion technique. Consequently, the coefficients were only calculated for the lower wavelengths. The subsequent three sections cover results for extracted optical properties together with studies at single wavelengths (1050nm and 1602nm). These investigations include results for samples of 2mm, 4mm, and 10mm sample thicknesses.

### 5.3.1. Bulk scattering coefficient

Figures 5.10a through 5.10c show changes in the scattering coefficient during different phases of a single cultivation, where samples were examined at three sample thicknesses (2mm, 4mm, and 10mm thickness). It can be seen that the relative change in $\mu_s$ over the duration of the growth phase (figure 5.10a) is significant. Further, the bulk scattering coefficient increases with the time at which the sample was taken during the cultivation run. This is expected since, in the cultivation system considered here, the largest variation is due to the increase of biomass which in turn contributes to the increase in scattering coefficient. The scattering coefficient generally decreases with the increase in wavelength, except in the region of the water peak (1450nm) where calculations were highly influenced by the water absorption and reliable estimates could not be obtained. In the region 1550nm-1850nm, the extracted bulk scattering coefficients exhibit a high degree of uncertainty as indicated by the high noise in the scattering spectra in this region, and by the fact that the spectra extracted at different times during the growth phase appear to have inconsistent trends at different wavelengths in this region. This high uncertainty in the extracted bulk scattering coefficients is probably due to the fact that in the absorption dominated region spectra are much less sensitive to changes in

scattering coefficients than to changes in the absorption properties. In the region between 1550nm and 1850nm, only scattering spectra for 2mm and 4mm sample thicknesses are presented, since the high noise in the raw measurements for 10mm sample thickness led to convergence problems when attempting to extract the optical properties in this region. It can be seen that relative levels of scattering are higher for 2mm sample and lower for 4mm sample.

Bulk scattering coefficients calculated with measurements obtained with different sample thicknesses possess similar values in the region 950nm-1350nm despite the fact that the raw spectra differ from each other in intensity for the different sample thicknesses. This is significant since the total diffuse reflectance spectra, shown in section 5.2.1, figure 5.1, exhibited not only different levels in intensity for the different pathlengths but also exhibited different behavior with respect to changes in biomass over the course of the growth phase.

Figure 5.10b shows the evolution of the scattering spectra during the stationary phase. As would be expected, there is very little change in the biomass concentration, and as a result the scattering properties show a much smaller variation compared to the growth phase. Scattering coefficients for different sample thicknesses possess close values in the region 950nm to 1350nm. However, it can be seen that values obtained with 4mm samples are slightly higher when compared with results for 2mm and 10mm. In the region 1550nm-1850nm, scattering coefficients for 2mm sample are higher than scattering coefficients for 4mm sample. Although sets of curves are expected to be similar, refractive index of different glass could be the reason for these differences and further investigations are needed. Results from measurements with 10mm sample could not be obtained in this region.

Figure 5.10c shows changes in scattering properties during the decline phase. The variations here are smaller comparing to the growth phase, and values are clearly lower than the values in the stationary phase.



(a)



(b)



(c)
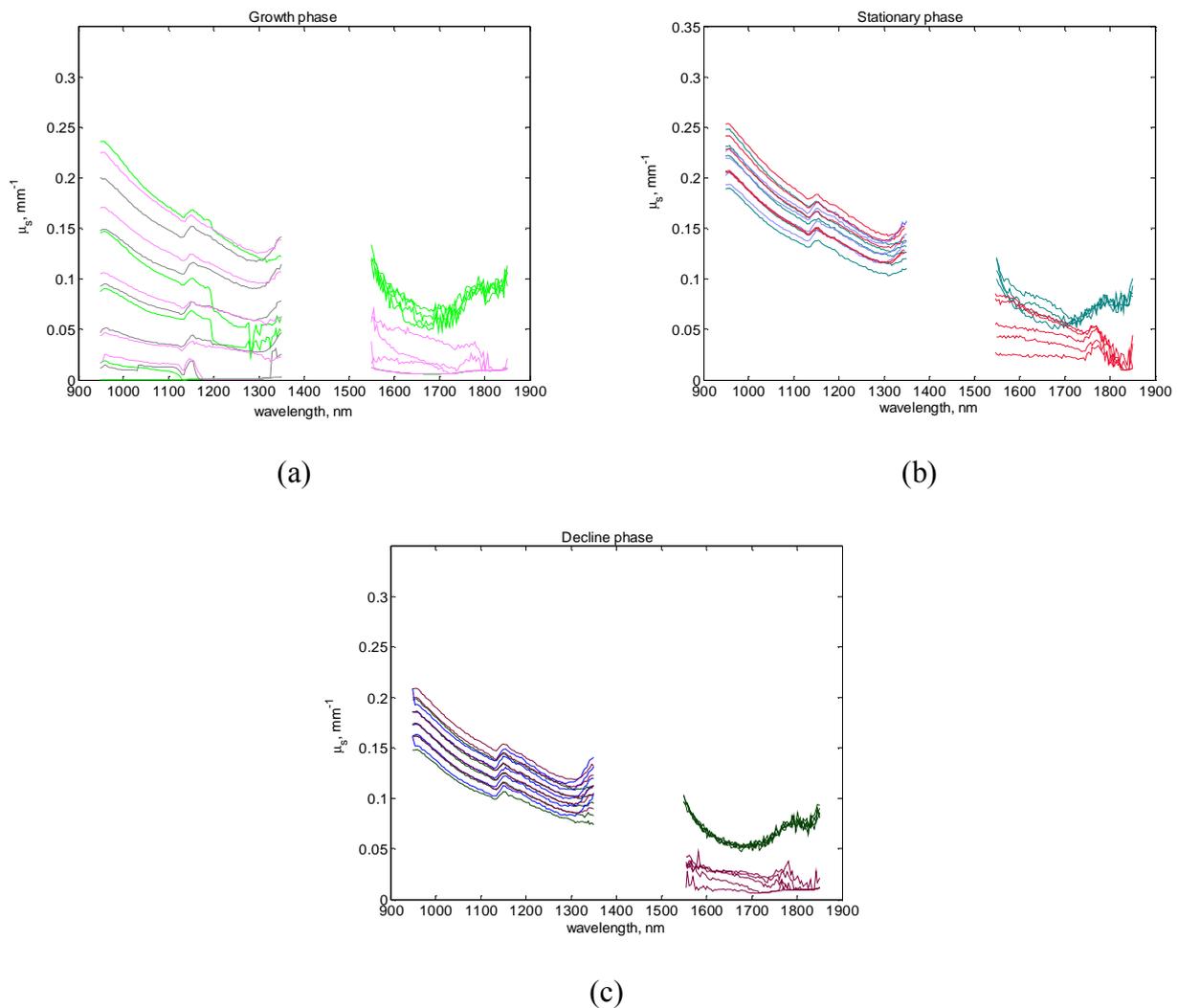
Figure 5.10 Estimated scattering coefficients at different stages of the cultivation and at different samples thickness. **(a)** growth phase (green-2mm, rose-4mm, and grey-10mm), **(b)** stationary phase (grey green-2mm, red-4mm, blue-10mm), and **(c)** decline phase (dark green-2mm, dark red colour-4mm, and dark blue-10mm).

For the region 950-1350nm, it is seen that $\mu_s$ values are slightly higher for 4mm sample thickness, compared to $\mu_s$ values with 2mm and 10mm sample thicknesses, with the last two taking very similar curves. In the region 1550-1850nm, scattering coefficients values for 4mm sample were found lower than values obtained for 2mm. Here, similarly to the

growth and stationary phases, measurements with 10mm sample did not provide reliable

results for $\mu_s$ beyond 1350nm.



(a)



(b)



(c)
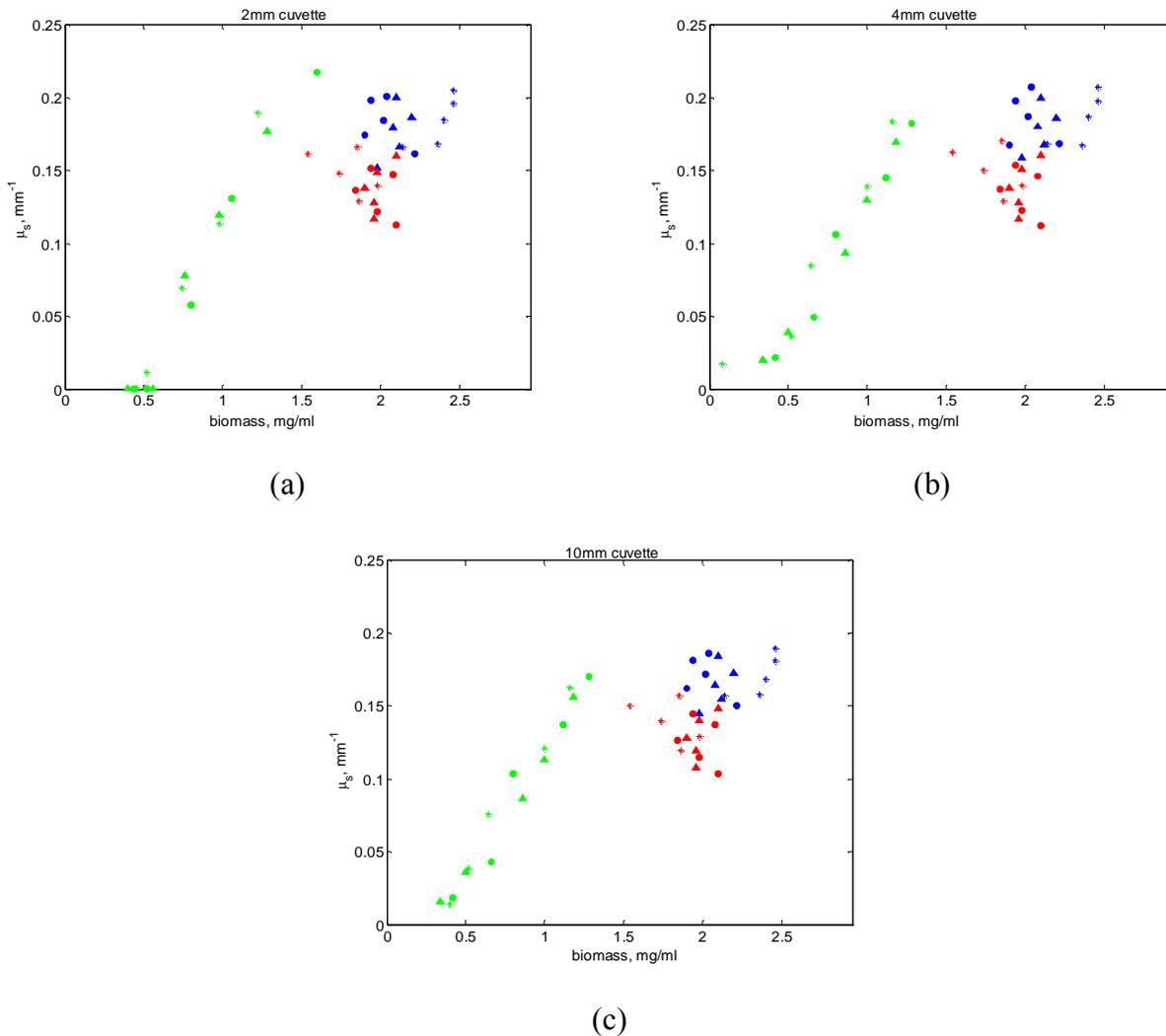
Figure 5.11 Estimated scattering coefficient at 1050nm during **growth** (green), **stationary** (blue), and **decline** phase (red) for three cultivations [(star symbols) run 1, (circle symbols) run 2, (triangle symbols) run 3] versus biomass. **(a)** Bulk scattering coefficient-2mm, **(b)** Bulk scattering coefficient-4mm, **(c)** Bulk scattering coefficient-10mm.

By analogy with the raw measurements, values for $\mu_s$ were displayed versus biomass for

1050nm (figure 5.11) and 1602nm (figure 5.12), in order to examine more closely the

trends in the scattering spectra. From figure 5.11 it can be seen that $\mu_s$ for 2mm, 4mm,

and 10mm show clear linear relationship with respect to the growth phase, while in the

stationary and decline phases this is not evident due to the reasons put forward in the discussion of results with the raw measurements.



(a)                                                    (b)
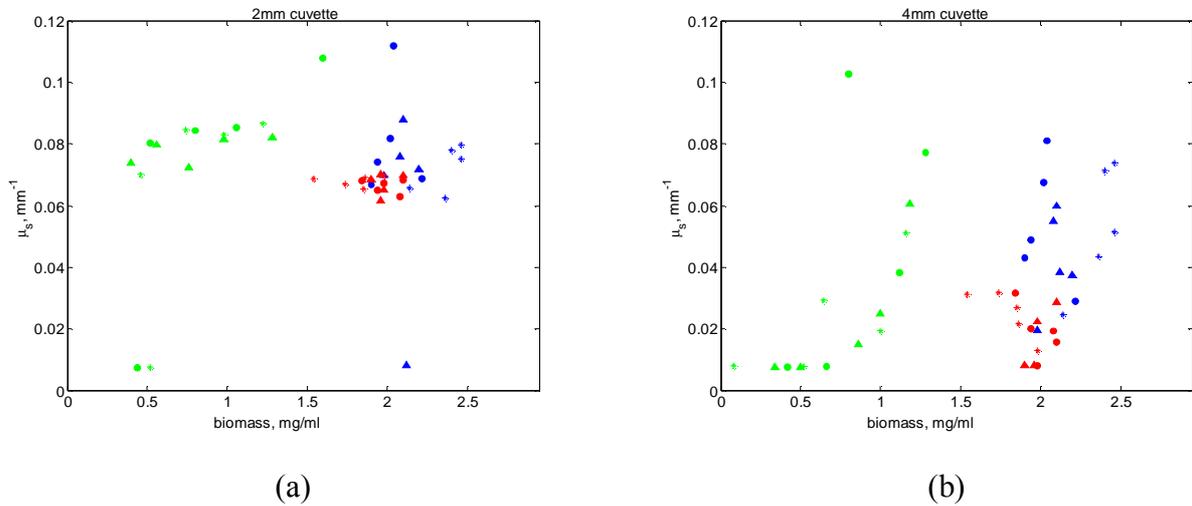
Figure 5.12 Estimated scattering coefficient at 1602nm during **growth** (green), **stationary** (blue), and **decline** phase (red) for three cultivations [(star symbols) run 1, (circle symbols) run 2, (triangle symbols) run 3] versus biomass. **(a)** Bulk scattering coefficient-2mm, **(b)** Bulk scattering coefficient-4mm.

In figure 5.12, $\mu_s$ for 2mm and 4mm sample thicknesses are plotted at 1602nm against biomass. Values for 10mm are not shown since, as described earlier, they were found not reliable in this region. It can be seen that the bulk scattering coefficients for 2mm sample thickness (figure 5.12a) show almost constant behaviour with the increase of the biomass indicating lack of sensitivity to changes in biomass in this region. During the stationary phase, the range of $\mu_s$ values expands a little and a cluster is formed, but overall the changes are very small. In the decline phase, the scattering coefficients were lower than those in the growth and stationary phase. Figure 5.12b shows $\mu_s$ for the 4mm samples. It can be seen that there is nonlinear relationship with respect to biomass in the growth phase. Examining the figure it could be argued that until about 0.5mg/ml of biomass there is no change in the scattering coefficient. It appears that unlike the 2mm samples, in the 4mm samples, since the photons can travel deeper into the suspension, it encounters more scattering events before exiting out of the sample. This raises the contribution

of scattering to an extent that the inverse adding-doubling routine is able to converge to a stable value of $\mu_s$. Two outliers are observed in this region. These are more likely to be due to convergence problem in the inversion rather than to any physical characteristics in the system. As it was in the case of 1050nm, there is no special trend in the bulk scattering coefficient with respect to biomass in the stationary and the decline phase. Compared to the decline phase, $\mu_s$ appears to be slightly higher in the stationary phase.

### 5.3.2. Bulk absorption coefficient

In figures 5.13a through 5.13c, absorption coefficient changes are shown. They are calculated for different sampling points from single cultivation for three sample thicknesses (2mm, 4mm, and 10mm). An important aspect is the comparison between $\mu_a$ for different sample thicknesses, and for this reason, results are displayed together. The absorption band of water dominates the spectra, with relatively low absorption in the region 950-1350nm for all samples studied and for all three different thicknesses, but it is seen to increase beyond 1350nm, which can be attributed to the water absorption in this area. From visual inspection it can be noticed that absorption coefficients determined for each of the sample thicknesses during the growth phase (figure 5.13a) have very small changes compared to the scattering coefficient, despite the differences observed in the raw measurements. The biomass is expected to affect $\mu_a$ and $\mu_s$, but absorption due to biomass in the NIR region is small compared to the absorption due to water, therefore, the relative changes due to an increase in biomass are small.

Absorption coefficients for different sample thicknesses are almost identical in the full region studied (950nm-1850nm). From 1350nm up to 1550nm, absorption coefficient was successfully calculated for 2 mm sample comprising all the stages of the cultivation run. This fact supports the statement made earlier in chapter 4 that decrease in the sample

thickness could bring improvement into the convergence of the program, therefore absorption coefficients can be calculated.



(a)



(b)



(c)
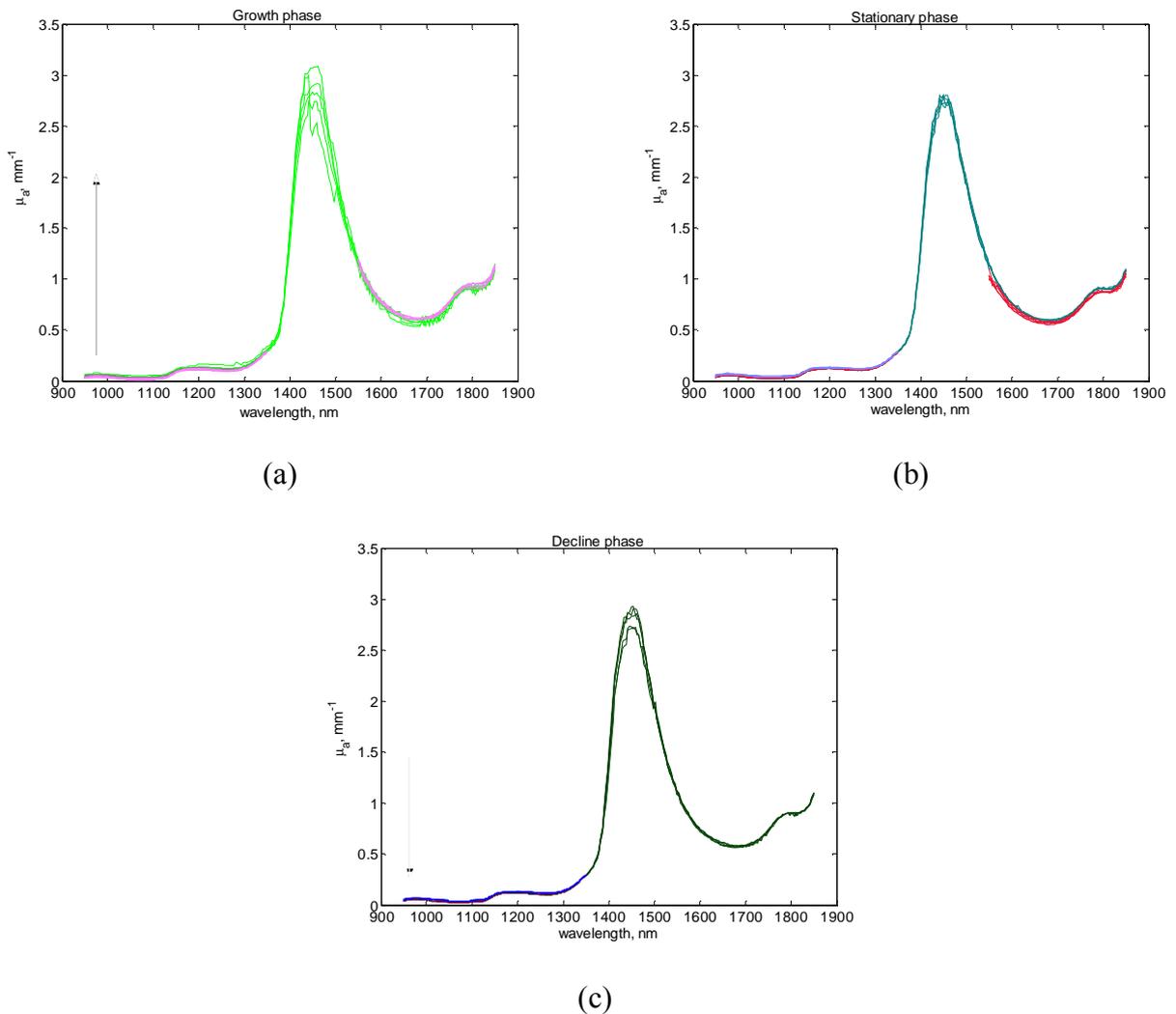
Figure 5.13 Estimated absorption coefficients at different stages of the growth cycle and at different samples thickness. **(a)** growth phase (green-2mm, rose-4mm, and grey-10mm), **(b)** stationary phase (grey green-2mm, red-4mm, blue-10mm), and **(c)** decline phase (dark green-2mm, dark red colour-4mm, and dark blue-10mm).

For 4mm and 10mm sample, meaningful information regarding absorption coefficient could not be extracted in the region of the water peak. Figure 5.13b shows the absorption coefficient during the stationary phase, and results for different sample thicknesses are compared. Again as it was in the case of the growth phase, very small variations can be noticed. Consistent values were obtained for all three samples thicknesses (2mm, 4mm, and 10mm), which also indicates the reliability of the inversion method.

Figure 5.13c shows the results for absorption coefficient during decline phase. The variations are again very small repeating the picture for growth and stationary phases. Absorption coefficient extracted from measurements obtained with 10mm sample could not be calculated for all phases in the region 1350nm-1850nm. In the next figure (5.14a), $\mu_a$ for 2mm sample thickness is shown as function of biomass.



(a)



(b)



(c)
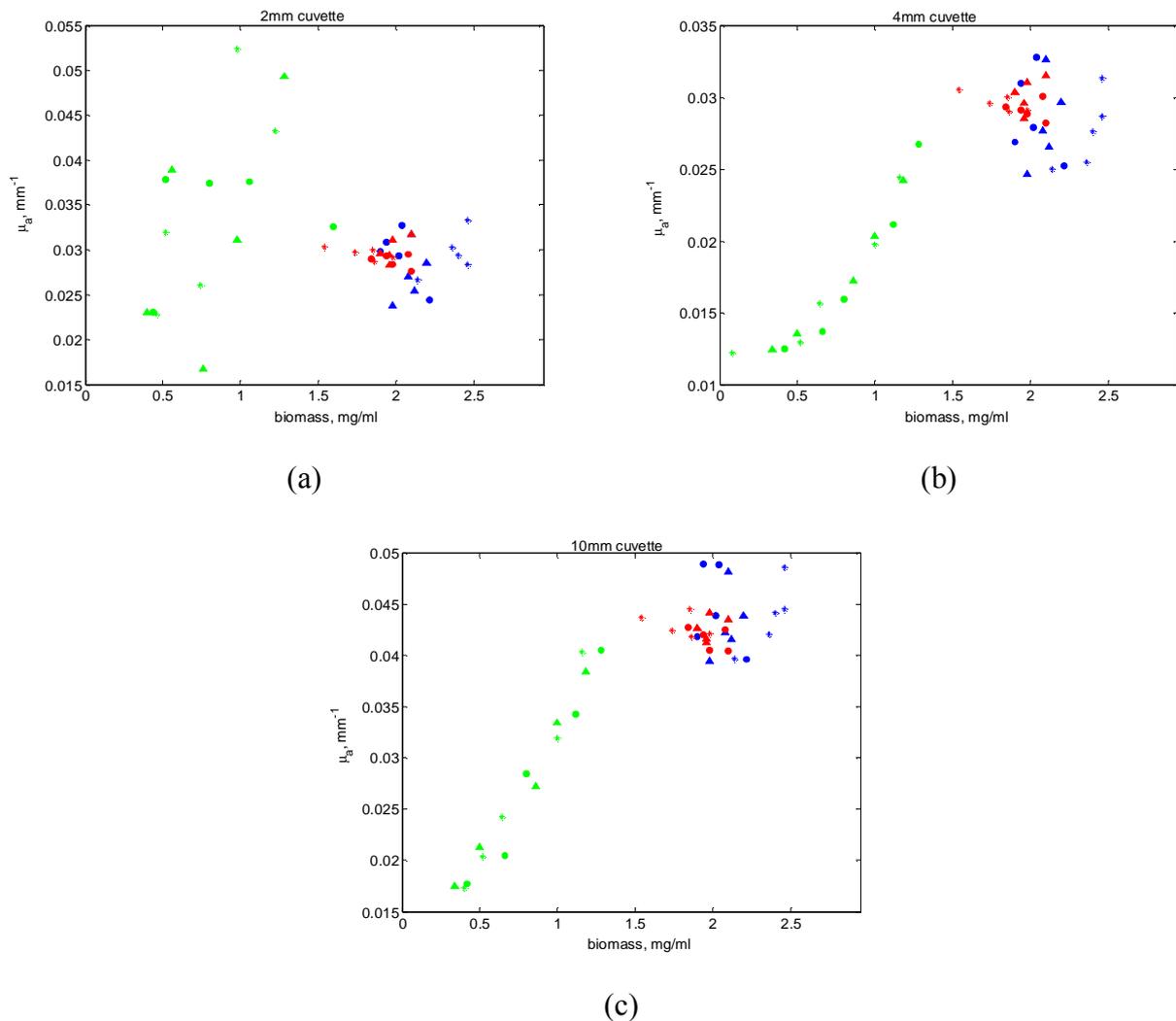
Figure 5.14 Estimated absorption coefficient at 1050nm during **growth** (green), **stationary** (blue), and **decline** phase (red) for three cultivations [(star symbols) run 1, (circle symbols) run 2, (triangle symbols) run 3] versus biomass. **(a)** Bulk absorption coefficient-2mm, **(b)** Bulk absorption coefficient-4mm, **(c)** Bulk absorption coefficient-10mm.

It is seen, in this case, that the relationship in the growth phase is weaker compared to $\mu_a$ shown for samples of thickness 4mm and 10mm. The bulk absorption coefficient is also

seen to increase with the increase of biomass for both 4mm (figure 5.14b) and 10mm (figure 5.14c) sample thicknesses. Such an increase in the bulk absorption coefficient has been also reported by Beauvoit et *al.*[3], while estimating optical properties of commercial *Baker's yeast*. Results from absorption coefficient led to the conclusion that $\mu_a$ appears as a linear function of the biomass when thicker sample is used, although the magnitude of these changes was relatively small with the increase of the cell number.

In figure 5.15, the absorption coefficient at 1602nm is plotted versus biomass. The values for 10mm sample were not extracted in this region, a fact associated with the high level of noise from the raw measurements. It is seen that, $\mu_a$ for 2mm sample thickness (figure 5.15a) does not hold any relation with the biomass whereas the data obtained using 4mm sample thickness shows an increase in the value of the absorption coefficient with the increasing biomass, but the relationship is weaker compared to that at 1050nm.

In summary, it can be concluded that the absorption coefficient is sensitive to changes in the biomass concentration provided the sample thickness is sufficiently large (>2mm) with the sensitivity to biomass concentration decreasing at wavelengths were the absorption of water dominates the spectrum.



(a)                                              (b)
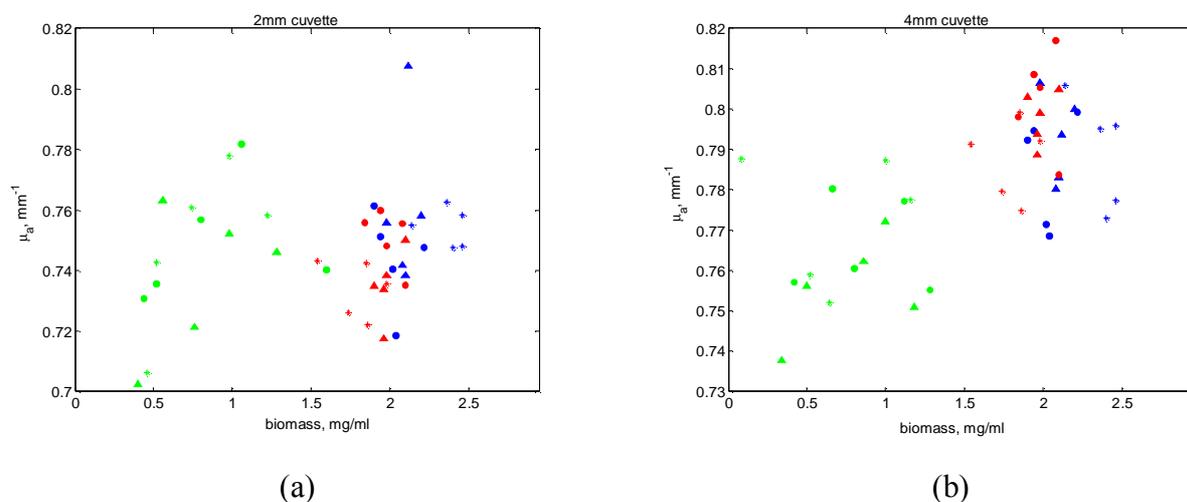
Figure 5.15 Estimated absorption coefficient at 1602nm during **growth** (green), **stationary** (blue), and **decline** phase (red) for three cultivations [(star symbols) run 1, (circle symbols) run 2, (triangle symbols) run 3] versus biomass. **(a)** Bulk absorption coefficient-2mm, **(b)** Bulk absorption coefficient-4mm.

### 5.3.3. Anisotropy factor

Figures 5.16a through 5.16c show the changes in the anisotropy factor *g* for different samples takien  over the course of a cultivation run. Examinations on *g* were performed for three sample thicknesses (2mm, 4mm, and 10mm).



(a)

(b)



(c)
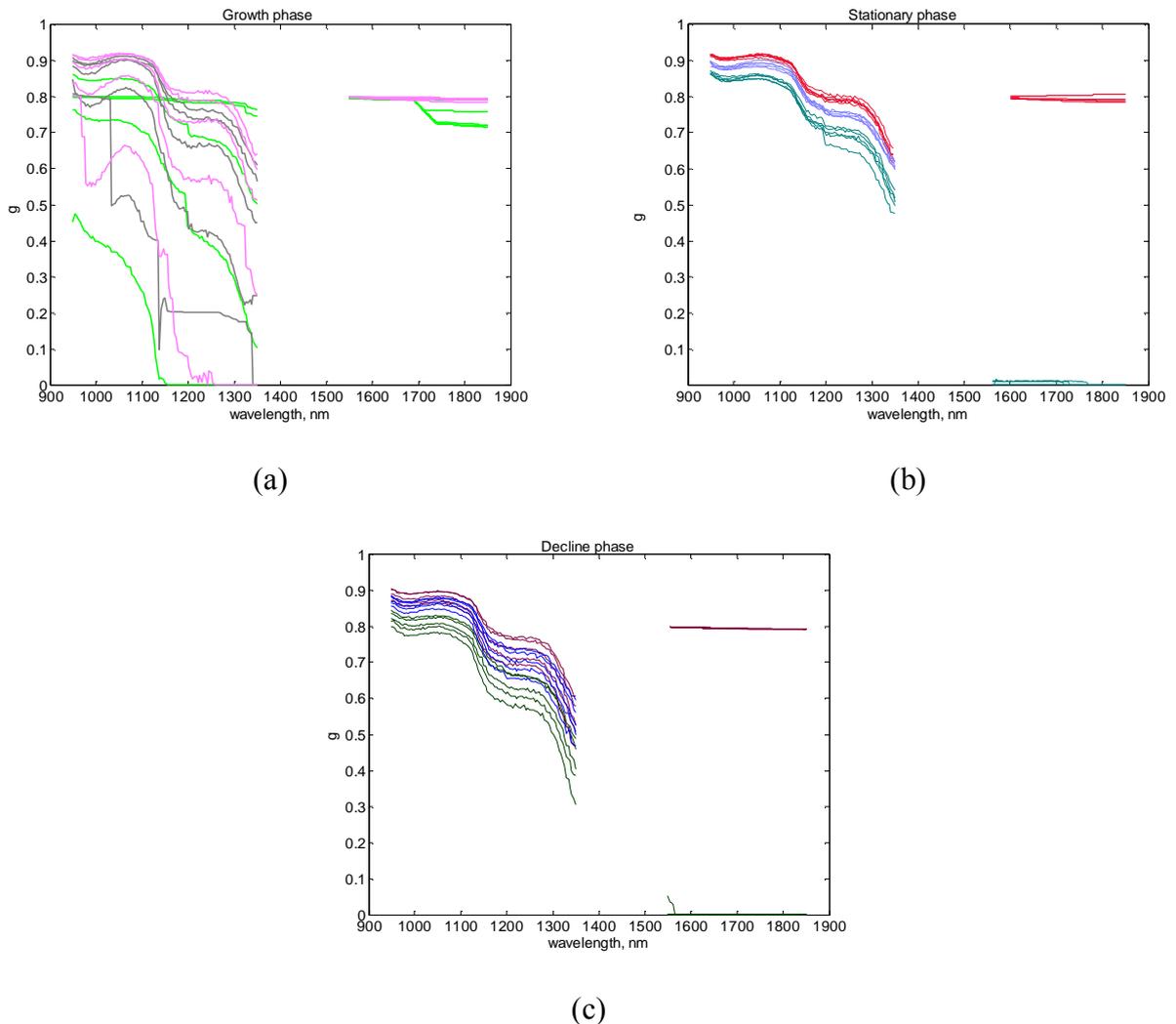
 Figure 5.16 Estimated anisotropy factors at different stages of the growth cycle and at different samples thickness. **(a)** growth phase (green-2mm, rose-4mm, and grey-10mm), **(b)** stationary phase (grey green-2mm, red-4mm, blue-10mm), and **(c)** decline phase (dark green-2mm, dark red colour-4mm, and dark blue-10mm).

For the growth phase (figure 5.16a), the anisotropy factor exhibits the greatest change whereas there is very little change in the stationary and decline phases. In the early stage

of the growth phase, in the region 950nm-1350nm, *g* drops sharply towards zero beyond 1200nm or flattens out in this region. Physically this fact is inconsistent. It has been noticed that the wavelength beyond which *g* falls in this manner increases as the growth cycle progresses. From this, it could be concluded that this effect is due to the fact that the low biomass in the initial stages of the cultivation and the resulting low scattering, which characteristically falls off at higher wavelengths, could be a factor in the inversion not being effective in extracting the anisotropy factor under such conditions. With the progress of the growth, the estimates of the anisotropy factor become stable over a large span of wavelengths. Beyond 1550nm, *g* stays around 0.8 which is the initial guess value input to the inverse adding-doubling program and thus cannot be expected to be a reliable estimate. Once the initial value has been changed with different values (0.7 and 0.9), the *g* in this region was found to accept values close to the new guesstimates, which led to the conclusion that anisotropy factor extraction in this region is not reliable. The change in the initial guess values however did not change the estimates for scattering and absorption coefficients. While examining the anisotropy factor extracted from measurement using different thickness of samples, it is seen that changes are similar for 2mm, 4mm, and 10mm. For all the samples, the values for *g* stay stable between 0.7 and 0.9 in the final stage of the growth. Figure 5.16b shows the anisotropy factor estimates in the stationary phase. It is seen that changes in *g* during this phase is very small. The samples with 2mm thickness exhibited the lowest *g*, the samples with 4mm exhibited the highest values for *g*, and samples with 10mm were found between the first two. Anisotropy factor estimates for the region 1550nm-1850nm were found close to zero or with values close to 0.8, which was the initial guess value used in the inversion technique. Again as it was shown with the growth phase, the anisotropy factor in the region 1550nm-1850nm is difficult to be determined at this stage, and the obtained

values are considered irrelevant. For all sample thicknesses, $g$ possesses values between 0.5 and 0.9.

Figure 5.16c shows the anisotropy factor estimates in the decline phase. Relative changes in $g$ are smaller compared to $g$ values in growth phase, and the results remain similar to the one obtained during the stationary phase. In this phase, the anisotropy factor has values between 0.3 and 0.9 when considering all samples.

The variations of $g$ with biomass at 1050nm are shown in figure 5.17.



(a)

(b)

(c)
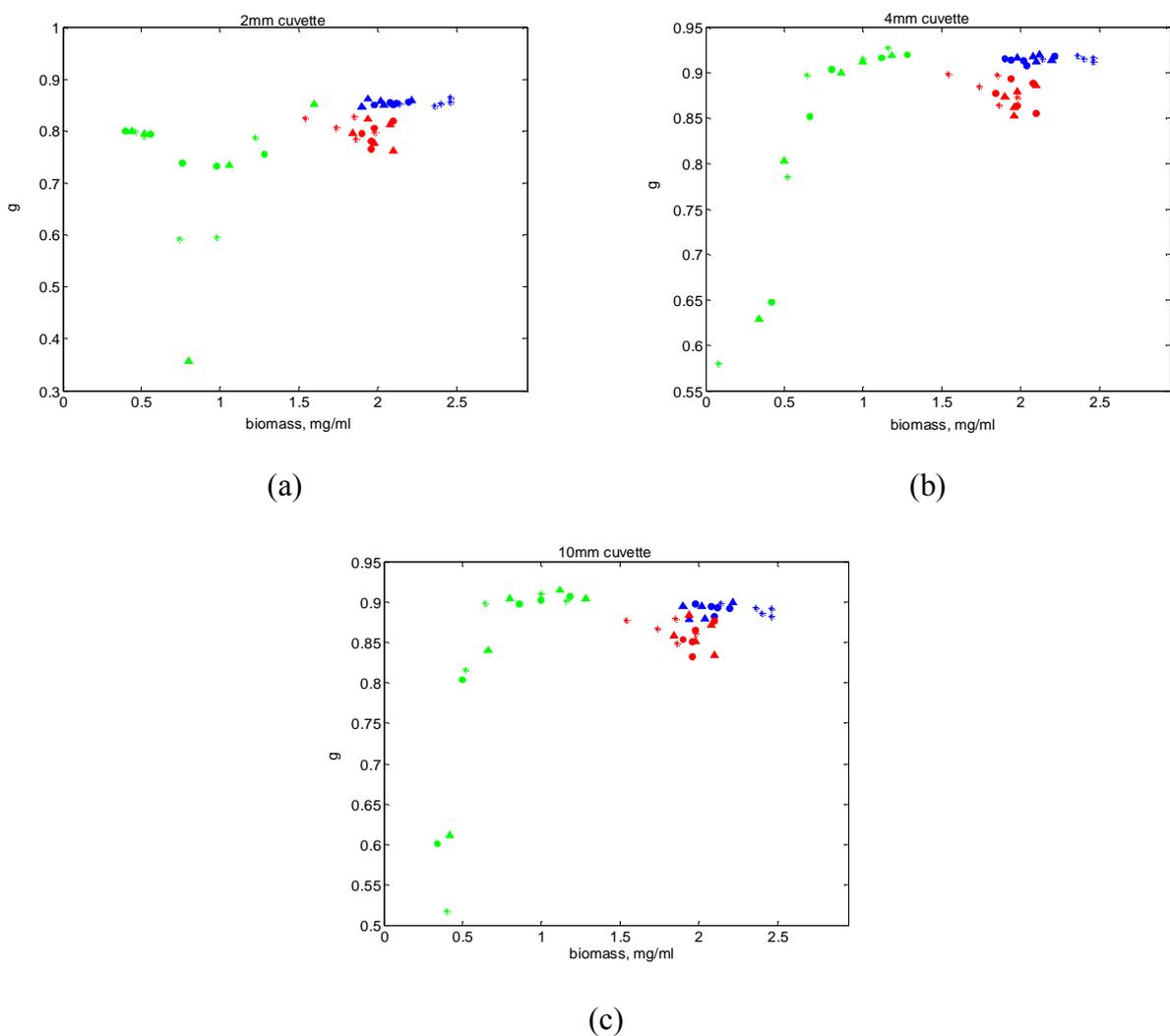
Figure 5.17 Estimated anisotropy factor at 1050nm during **growth** (green), **stationary** (blue), and **decline** phase (red) for three cultivations [(star symbols) run 1, (circle symbols) run 2, (triangle symbols) run 3] versus biomass. **(a)** Anisotropy factor-2mm, **(b)** Anisotropy factor-4mm, **(c)** Anisotropy factor-10mm.

It is seen that during the growth phase with sample of 2mm (figure 5.17a), $g$ takes values close to 0.8 at initial biomass concentrations (around the region of 0.5mg/ml). With the increase of the biomass, the anisotropy factor slightly decreases to 0.75 and then again rises to 0.9. Three point have shown lower values (one point $g$=0.35 and two others $g$=0.6) than the one mentioned, however these three points did not repeat for the three sample preparations from the cultivation runs. They are most likely to be outliers due to convergence issues in the inversion rather then to any physical characteristics in the system. At biomass 1.5mg/ml $g$ is a little less then 0.9, where it remains during the stationary phase. In the decline phase, anisotropy factor takes only slightly lower values ($g$=0.8) than that observed for the stationary phase. The values for $g$, obtained from 4mm (figure 5.17b) and 10mm (figure 5.17c) samples have very similar pattern. During the growth phase for both cases it is seen that $g$ rises sharply and stabilizes around 0.9. It stays in this region during the stationary phase, and decreases slightly in the decline phase. The sharp initial increase in $g$ as was already suggested in chapter 4 could be due to the rapid increase in rate of cells division creating a larger population of freshly divided cells (of smaller size) compared to the cells yet to divide into smaller cells but more investigation are needed in this direction.

**5.4 Discussion and conclusions**

This study is the first to report full optical properties spectra of a microorganism over a wide wavelength range and different samples thicknesses. Previous reports on optical properties of bacterial suspensions found in the literature reported the optical properties at isolated wavelengths only and with limited information. More details regarding these studies are included in table 5.1.

Table 5.1 Optical properties for various liquid biological samples

| Micro-organism | Wavelength, nm | Reduced scattering coefficient, $\mu_s'$ | Absorption coefficient $\mu_a$ | Scattering coefficient, $\mu_s'$ | Mean cell size | Biomass |
|---|---|---|---|---|---|---|
| Commercial Baker's yeast[3] | 780 | 2-5cm$^{-1}$ | 0.02-0.07cm$^{-1}$ | --- | --- | 20-50mg/ml |
| Baker's yeast[4] | 780 | 3.54cm$^{-1}$ | 0.025cm$^{-1}$ | --- | --- | --- |
| US patent[5] | 543 | 3-8.5cm$^{-1}$ | 0.045-0.065cm$^{-1}$ | --- | Diameter 4.7-5.5μm | --- |
| Present study- *Bacillus subtilis* | 950-1850 | --- | 0.01-3mm$^{-1}$ | 0.01-0.25mm$^{-1}$ | Length/Width 2.486/0.7998 | 0.34-2.46mg/ml |

The effect of pathlength on the raw measurements and the extracted optical properties in the region 950-1850nm has been studied in this chapter. A number of comparisons and analysis were made including (a) collection of reflectance and transmittance spectra for 2mm, 4mm, and 10mm samples while the culture was going through growth, stationary, and decline phases, (b) and extraction of optical properties based on the raw measurements. In order to investigate the light distribution as a function of biomass, two main regions were chosen: scattering dominated region (1050nm) and absorption dominated region (1602nm). In terms of raw measurements, for the whole NIR region studied it is difficult to determine only one optimal pathlength that needs to be considered when spectral data from *Bacillus subtilis* growing culture are collected. The shorter wavelengths allow longer sample thicknesses (4mm and 10mm) to be used for higher discrimination between the different stages of the bacterial growth, a fact that is often associated with the weaker nature of the absorption bands in this region. Variations in the diffuse reflectance (figure 5.1) and diffuse transmittance (figure 5.4) measurements are however small, whereas in the case of collimated transmittance (figure 5.7) differences in the spectra due to biomass concentration are well pronounced. The sample

thickness has been found to affect the relationship between diffuse reflectance and biomass concentration particularly when the measurement is performed in the scattering dominated region (short wavelengths; see 10mm sample thickness measurements in figure 5.1). In terms of biomass concentration, diffuse reflectance (figure 5.2), diffuse transmittance (figure 5.5) and collimated transmittance (figure 5.8) show strong correlation in the growth phase especially when samples with higher thicknesses (4mm and 10mm) are analysed. The water absorption peak dominates the region between 1350nm and 1550nm and affects both collimated transmittance and diffuse transmittance resulting in high level of noise that makes this region unusable for analysis when it comes to 4mm and 10mm sample thicknesses. In the longer wavelength region the strong relationship between biomass concentration and spectral information is valid only for the measurement of collimated transmittance (figure 5.9). For diffuse transmittance (figure 5.6) and diffuse reflectance (figure 5.3) this relationship is weaker.

It can be concluded that although the quality of signal at different wavelengths may differ for different thicknesses of the sample, the most appropriate sample thickness for this particular analysis is 4mm. With this sample thickness, the component of interest (biomass) can be effectively studied in both scattering and absorbing regions. The 10mm sample thickness is recommended to be used for analysis in the scattering region only, whereas the 2mm sample thickness is appropriate for absorbing region.

The next step is to examine the effect of sample thickness on the estimation of the optical properties and to determine whether a particular thickness would be better than others from the point of view extracting $\mu_a$ and $\mu_s$. In the region 950nm-1350nm, the extracted scattering coefficients possess similar values for all three sample thicknesses (see 2mm, 4mm, and 10mm in figure 5.10). These results follow the successfully performed validation step in chapter 3 on the measurement setup and the IAD method. The

repeatability of the curves tracing the scattering spectra from the three sample thicknesses, serves as an indicator for the stability and the accuracy of the proposed methodology. For the 2mm, 4mm, and 10mm sample thicknesses, in the growth phase, the scattering coefficients are in a clear linear relationship with respect to biomass, whereas no significant changes were registered in the stationary and decline phases. The absorption dominated region, $\mu_s$ demonstrates a high degree of uncertainty and can not be used as an effective indicator for biomass changes (figure 5.12). Overall, the bulk scattering coefficient has been consistently extracted and all three sample thicknesses can be used for spectral analysis when it comes to the short wavelength region.

The bulk absorption coefficient (figure 5.13) on the other hand, when compared to the scattering coefficient, demonstrated very small changes during the course of the cultivation despite differences observed in the raw measurements. For the 2mm, 4mm, and 10mm sample thicknesses, $\mu_a$ is approximately the same. The consistency of the extracted values indicates a high degree of reliability of the inversion method. The smaller sample thickness (2mm) improved the convergence of the program (values for $\mu_a$ were obtained from the water absorption peak region). When thicker samples were used (4mm and 10mm) in the scattering dominated region, the absorption coefficient was able to exhibit linear relationship (figure 5.14) with the biomass, a fact that could not be seen with 2mm sample. In the absorption dominated region, $\mu_a$ did not demonstrate a strong relationship (figure 5.15) with the increase in the level of biomass.

The variation of the anisotropy factor $g$ was found to exhibit changes during the growth phase that were best captured with 4mm and 10mm sample thicknesses (see figure 5.17) in the scattering region. During the stationary phase, $g$ remained constant with the biomass, and a slight decrease has been registered during the decline phase. Since anisotropy factor depends on the particle size and cell geometry, this decrease in the $g$

values could be related to the changes observed in the morphology of the culture (SEM pictures in chapter 4).

In chapter 5 it was shown that the extracted optical properties were found fairly consistent even though they were obtained from measurements with different sample thicknesses. Samples drawn from bacterial suspension and measured with the three different cuvettes demonstrated differences in the raw spectra in terms of intensity levels, especially in the case of diffuse reflectance measurements (figure 5.1). The results for optical properties, extracted following the proposed approach, have shown that the differences from the raw measurements were not exhibited in the optical properties, and the latest were found almost identical for each of the sample thicknesses: 2mm, 4mm, and 10mm. To put this work into context means that raw measurement (Rd and Td) taken from *Bacillus subtilis* fermentation process are affected by variations arising from multiple light scattering. After separating scattering effects ($\mu_s$) from absorption ($\mu_a$), the obtained identical results led to the conclusion that optical parameters extracted for this biological suspension ($\mu_a$ and $\mu_s$) are independent from variations in the pathlength of the photons that occur due to multiple scattering. It is necessary to investigate whether the newly introduced variables are linearly proportional to the biomass and glucose concentration of the fermentation process studied in this work. For this, chapter 6 will deal with development of calibration models for prediction of constituents of interest. The proposed models will be built on absorption coefficient, scattering coefficient, and on raw measurements. Assessment of the model performance will be provided.

It is difficult to determine one optimal sample thickness at which optical properties can be extracted with respect to the entire wavelength region studied. The analysis from scattering spectra has shown that all three sample thicknesses can be used for collection of raw measurements in the short wavelength region, and the resulting optical properties

will have good reproducibility. For the same scattering region, results for bulk absorption coefficient (figure 5.14) and anisotropy factor (figure 5.17) obtained with 4mm and 10mm samples were found in a good agreement.

**References**

1.      Ge, Z., A. Cavinato, and J. Callis, *Noninvasive spectroscopy for monitoring cell density in a fermentation process.* Analytical Chemistry, 1994. **66**: p. 1354-1362.
2.      Hellma, *Hellma Genaral Catalogue.* 2002: Hellma GmbH & Co. KG.
3.      Beauvoit, B., H. Liu, K. Kang, P.D. Kaplan, M. Miwa, and B. Chance, *Characterization of absorption and scattering properties for various yeast strains by time-resolved spectroscopy.* Cell Biophys, 1993. **23**(1-3): p. 91-109.
4.      Liu, H., M. Miwa, B. Beauvoit, N.G. Wang, and B. Chance, *Characterization of Absorption and Scattering Properties of Small-Volume Biological Samples Using Time-Resolved Spectroscopy.* Analytical Biochemistry, 1993. **213**(2): p. 378-385.
5.      Hielscher, A., J. Mourant, and I. Bigio, *Characterization of highly scattering media by measurement of diffusely backscattered polarized light*, U.S. Pattent, Editor. 2000, Patent number 6011626: United States.

# Models for prediction of biomass and glucose concentration in liquid cultures of *Bacillus subtilis*

## 6.1. Introduction

Near-infrared spectral data are related to the chemical and morphological composition of a given particulate system (in this work a biological system is under consideration). In many cases this relationship is complicated by light scattering effects which due to their nonlinear nature become the main limitation for obtaining accurate and universally applicable multivariate calibration models for prediction of analytes such as glucose and biomass. Accordingly, in the search for an adequate solution to the issues arising due to light scattering effects, tremendous amount of work has been accomplished in the last three decades. Kubelka-Munk transformation[1] is one of the methods widely applied while reflectance measurements are under consideration, as well as the work of Geladi et *al.*[2] when was first introduced the multiplicative scatter correction concept for separating scattering from absorption effects from reflectance spectra of meat.

The approach in this work differs from previously mentioned methods and in fact it is for the first time to be tested on bacterial suspensions. A method called ADD method has been used in order to extract absorption ($\mu_a$) and scattering coefficients ($\mu_s$) from reflectance and transmittance measurements. The methodology of this approach was described in chapter 3 and later (in chapter 4 and chapter 5) applied for the estimation of absorption and scattering spectra of *Bacillus subtilis* growing culture. The accuracy of the method was verified for three different sample thicknesses. The final results obtained for absorption and scattering coefficients from the three sample thicknesses were found in good agreement, with almost identical values. Given the fact that the optical properties can be consistently extracted, the next question was whether they are of sufficient

accuracy such that models for predicting glucose and biomass concentration using the optical properties will perform better than when models are built using the reflectance or transmittance spectra directly.

In this work, calibration models were built using partial least square regression (PLS) for predicting glucose concentration and biomass during bacterial growth of *Bacillus subtilis* in liquid culture. Models were built using the traditional chemometric approach where single measurements are used in conjunction with empirical pre-processing techniques. For the construction of the models, data from three types of measurements collected from each sample were available, namely total diffuse reflectance, total diffuse transmittance and collimated transmittance. Then models were built using a two-step approach. The first step in this approach consisted of using the three measurements to extract the bulk absorption and scattering properties and the anisotropy factor as described in chapter 3. This was followed by building PLS models using the extracted bulk absorption coefficient and bulk scattering coefficient spectra for predicting the glucose and biomass concentration during the bacterial growth cycle. The performances of the models obtained by these various methods were compared.

The main objectives of this chapter are to (1) investigate the effectiveness of the two-step approach using measurements from an integrating sphere setup for predicting biomass and glucose concentration, (2) to compare the performance of models built using the two step approach with models constructed on single measurements (Td, Rd, and Tc), and (3) to evaluate the sample thickness effect on the performance of PLS models.

## 6.2. Materials and methods

### 6.2.1. Cultivation runs

In this study, data from *Bacillus subtilis* fermentations were collected from separate cultivation runs. A total of 5 growth cycles were performed, and for all these cultures, during the data collection phase, samples were taken at approximately 2 hour intervals. Precise details for the conditions for the cultivations can be found in chapter 4 (section 4.2). The main idea that makes cultivation runs from this chapter different from the cultivation runs described in chapter 4, is that a different strategy for initial glucose concentration has been applied by using different initial glucose concentrations for each of the runs: 8mg/ml for cultivation run 1, 16mg/ml for cultivation run 2, 25mg/ml for cultivation run 3, 34mg/ml for cultivation run 4, and 43mg/ml for cultivation run 5. This is a much wider concentration range than in the previous experiments. While estimating optical properties in chapter 4 and chapter 5, all the cultivation runs started with a fixed initial glucose concentration of 4mg/ml.

The strategy of progressively increasing the initial glucose concentration in the nutrient medium was followed in order to ensure that the concentration range is well spanned and to test whether a model developed using one cultivation run could be applied to succeeding runs run under different initial glucose rates and therefore, different growth rates. Optical density curves during the cultivation of 5 different runs, together with the points indicating the time when samples were drawn from the flask are shown in Appendix B.

### 6.2.2. Assays

At 2 hour intervals through the cultivation, simultaneously with the spectra collection, samples were withdrawn from the culture and subjected to off-line analysis to determine

the concentration of glucose and biomass. Both reference assays for biomass and glucose concentration were discussed in chapter 4 (section 4.2.1). Appendix C shows the biomass (Appendix C.1) and glucose concentration (Appendix C.2) ranges for the data set used for this study.

### 6.2.3. Near-Infrared data collection

During the cultivation runs, for each sample and at the same time as the withdrawal of samples for offline assays, samples for NIR analysis of the culture were collected and processed immediately on the Cary 5000 spectrophotometer equipped with a single integrating sphere. In this way match between reference data and spectral data was achieved. The samples were directly introduced for measurements with no pre-treatment or manipulation, and each one was scanned in the wavelength range from 950-1850nm at 4nm intervals with three different sample thicknesses: 2mm, 4mm and 10mm (details of the experimental setup have been covered in chapter 3). The samples were acquired at approximately every 2 hour intervals over 10 hours growth period for cultivation run (1)-(2), and over 30 hours growth period for cultivation runs (3)-(5). The data set for cultivation runs (1)-(2) also consist of five samples each, while the data set for cultivation runs (3)-(5) consist of ten samples each. This difference in the sample number is due to the fact that cultivation runs (1)-(2) were started with lower glucose concentration and its depletion appeared faster compared to runs (3)-(5). Finally, the total data set consists of 40 samples from five cultivation runs of *Bacillus subtilis* growing in a liquid culture. For each sample, biomass concentration, glucose concentration, raw near-infrared measurements data, and estimated absorption and scattering coefficients were included in the data set.

### 6.2.4. Chemometric methods and data analysis

Calibration models were built using PLS regression method in combination with cross-validation (CV). Cross-validation is a technique based only on calibration data and is done by deleting samples from the calibration set. It starts with deleting sample number one, and then the calibration is performed using the remainder of the samples in the dataset. The calibration model is tested by predicting the analyte concentration using the spectrum from the first sample and comparing the value with the value obtained from the reference assay of that sample. After this step is completed, the first sample is put back into the calibration set, the second sample is left out and the first step is repeated. This procedure is repeated until all the samples have been left out once.

In this work segmented CV was used where instead of deleting only one sample, the CV step was performed based on 'leave one cultivation run out'. Segmented CV is considered useful if there are certain structures in the dataset[3], as it is for this case where batch culture of bacterial growth is under consideration. It is expected that if replicates samples exist in the data matrix, RMSECV computed by full cross-validation might be overoptimistic due to similar samples represented in both calibration and test set. It is recommended to combine those replicates in a segment that can be deleted from the calibration data. This approach was adapted such that the cross-validation scheme was applied by leaving out samples from an entire cultivation run. In other words, samples from one cultivation run were taken out from the calibration set, and calibration was performed on the remaining samples and the process repeated till each of the 5 cultivation runs were left out once.

Prior to building PLS models from the spectra, several pre-processing techniques were tried, such as first derivative, second derivative and a number of signal correction methods: (1) ordinary multiplicative scatter correction (MSC), (2) piecewise MSC,

(3) standard normal variate (SNV), (4) extended multiplicative scatter correction (EMSC), and (5) inverse signal correction (ISC). For the EMSC, two forms were considered; the first extended multiplicative scatter correction technique is based on a second degree wavelength dependent polynomial, denoted as EMSCW, and the second has been denoted as EMSCL for which a log term for the wavelength dependence was used[4]. A brief description of these techniques is included in Appendix E. Further, combinations of the various techniques with first and second derivative processing were also considered.

PLS models combined with different pre-processing techniques were constructed for predicting glucose and biomass concentrations. The models were built using the spectra from different measurement configurations (Td, Rd and Tc) and spectra from extracted optical properties ($\mu_a$ and $\mu_s$). The spectral data in combination with data from reference assays were processed using Matlab code in order to build the PLS models. The process of developing the PLS models included three main steps: (1) choice of analyte to predict (biomass or glucose), followed by choice of cuvette pathlength i.e. sample thickness, (2mm, 4mm, or 10mm) and spectral data to be used (Rd, Td, Tc, $\mu_a$, and $\mu_s$), (2) choice of pre-processing methods that have been described earlier in this section. The last step (3) of the process was the application of partial least squares (PLS) regression. Figure 6.1 shows a flowchart illustrating various combinations that have been applied.
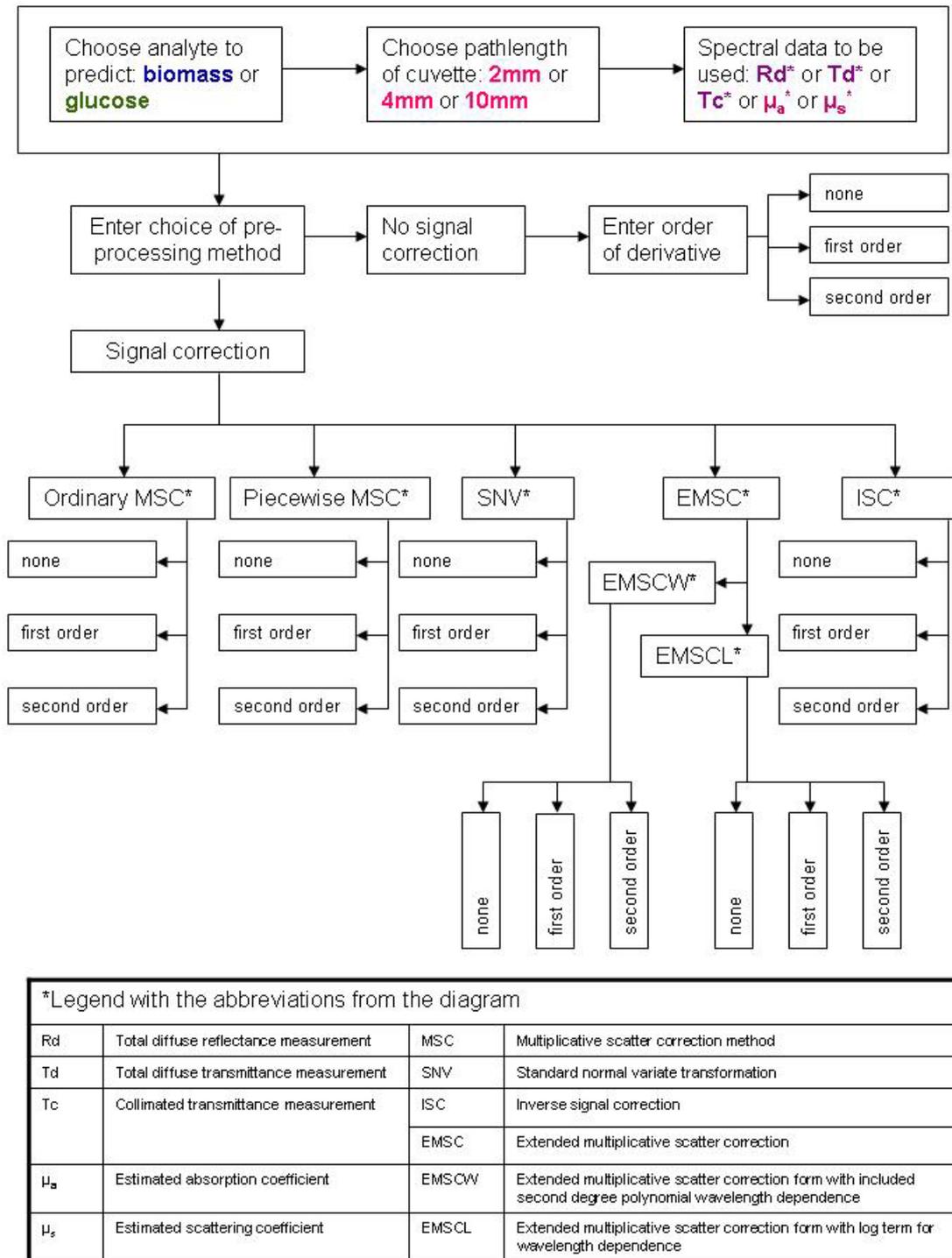
Fig. 6.1. Flow diagram showing the steps in the calibration model building and analysis using different pre-processing techniques carried out in this study.

Individual PLS models were created for prediction of glucose and biomass concentrations and the relative performance of the models was assessed by (1) the

multiple coefficient of determination R$^2$, which is a measure of the goodness of the fit of the model to the calibration data, and (2) the root mean square error of cross validation (RMSECV) which indicates the predictive ability of the model within the calibration set and the optimal number of latent variables (LV's) to be used. The estimate of root mean square error based on cross validation technique is called root mean square error of cross-validation and is defined by the following equation

$$RMSECV = \sqrt{\sum_{i=1}^{N}(\hat{y}_{cv,i} - y_i)^2 / N} \qquad (6.1)$$

where, $\hat{y}_{cv,i}$ is the estimate for $y_i$ based on the calibration equation with sample $i$ deleted, and $N$ is the total number of samples. The number of the LV's to be included in the PLS models was assessed by cross-validating the calibration set. After the RMSECV of the calibration model was calculated, they were plotted versus the number of LV's for further assessment. The optimum number of the LV's was found at the minimum of the plot, however, in some cases a smaller number of LV's resulted in approximately the same error, and therefore this smaller number was selected. Matlab 6.5 scientific software was used for the calculations. Results from the analysis are presented in the next section.

## 6.3. Results and discussion

The next two sections include PLS models for biomass and glucose prediction. Models were constructed on raw measurements obtained using three different sample thicknesses, followed by models based on extracted optical properties for the same three sample thicknesses. A complete set of spectra recorded during the cultivations is

shown in Appendix D, figure D1, including information from Td, Tc, Rd for 2mm, 4mm, and 10mm samples. From the raw spectral data for each sample, the optical properties were extracted following the methodology already described in chapter 4, and results are shown in Appendix D, figure D2. As was already shown in chapter 4 and chapter 5, significant noise was registered for measurements based on different sample thicknesses in the region of the water absorption peak (1450nm). The noise is expected to degrade the signal-to-noise ratio and to decrease the accuracy and the precision of a given calibration model which could introduce biases into the model and will lead to increased number of LV's[5]. In order to avoid such issues, regions dominated by noise were excluded aiming to prevent their influence on the constructed PLS models. With these facts in mind, two regions were considered in the present study: for data collected with 2mm and 4mm sample thicknesses, the regions were from 950nm to 1350nm, and from 1550nm to 1850nm, whereas for data collected with 10mm sample thickness, only the region from 950nm to 1350nm could be used due to the very noisy signal in the region of the water absorption peak (1450nm) and beyond (Appendix D). These regions were chosen for models constructed on raw measurements as well as for models constructed on absorption and scattering coefficients aiming to obtain comparable results and to achieve consistency in the process of analysing the different data.

### 6.3.1. Models for the prediction of biomass

The cultivation runs were designed so that as broad a range of biomass concentrations as possible could be obtained for this particular type of culture and growth media. The aim was to make a collection of samples that can be used for the construction of reliable predictive models. The results presented in this section were obtained after testing each one of the pre-processing techniques described in section 6.2.4 ('Chemometric methods

and data analysis') using the data set obtained from the five cultivation runs of *Bacillus subtilis*. Models for prediction of biomass were built based on six different parameters, total diffuse transmittance (Td), total diffuse reflectance (Rd), collimated transmittance (Tc), scattering coefficient ($\mu_s$), absorption coefficient ($\mu_a$), and bulk extinction coefficient ($\mu_t = \mu_a + \mu_s$), and the criteria used for assessing the utility of the constructed models included LV's, RMSECV and the multiple coefficient of determination. Initially models were not constructed on $\mu_t$, however following the analysis of the influence of sample thickness on the optical properties, it has been found that both absorption and scattering coefficients contain information related to the biomass concentration. Since $\mu_t$ is function of both $\mu_a$ and $\mu_s$ it was felt that bulk extinction coefficient which is the sum of the bulk scattering and absorption coefficient would provide a parameter with enhanced information since it would then include information contained in both the aforesaid parameters. Since $\mu_a$ and $\mu_s$ are pathlength normalised, $\mu_t$ is also pathlength normalised meaning that the pathlength variation due to multiple light scattering have been removed. Table 6.1 reports the performance of PLS models for biomass prediction based on different input variables. While all pre-processing techniques described previously were tried out, only the techniques providing the best results is included in the table and later presented in the figures. The table itself is divided in three sub-sections, each one including data for different sample thicknesses: 2mm, 4mm, and 10mm. In all cases the raw spectral data (Tc,Td, and Rd) were converted to absorbance units (-$\log_{10}$Tc, -$\log_{10}$Td, and $-\log_{10}$Rd) prior to building the PLS models. RMSECV curves for models constructed on absorption coefficient, scattering coefficient, bulk extinction coefficient, collimated transmittance, diffuse reflectance, and diffuse transmittance for measurements taken with 2mm sample thickness are presented in figure 6.2. The results for 4mm sample thickness are presented in figure 6.4 and results for 10mm sample thickness are presented in figure 6.6.

Table 6.1 Results for PLS calibration models performance for prediction of biomass.

| Input spectral data | Pre-processing technique | Derivative choice | Number of LV's | RMSECV (mg/ml) | $R^2$ |
|---|---|---|---|---|---|
| **Cuvette pathlength – 2mm; wavelength region 950-1350nm&1550-1850nm** | | | | | |
| Td | none | none | 5 | 0.595 | 0.934 |
| Rd | EMSCW | none | 3 | 0.799 | 0.880 |
| Tc | Ordinary MSC | none | 5 | 0.314 | 0.982 |
| $\mu_s$ | none | none | 3 | 0.357 | 0.976 |
| $\mu_a$ | none | none | 3 | 0.770 | 0.889 |
| $\mu_t$ | none | none | 4 | 0.511 | 0.951 |
| **Cuvette pathlength – 4mm; wavelength region 950-1350nm&1550-1850nm** | | | | | |
| Td | ISC | none | 3 | 0.636 | 0.924 |
| Rd | none | none | 2 | 0.672 | 0.915 |
| Tc | Ordinary MSC | none | 7 | 0.219 | 0.991 |
| $\mu_s$ | none | none | 2 | 0.557 | 0.942 |
| $\mu_a$ | EMSCL | none | 5 | 0.898 | 0.852 |
| $\mu_t$ | ISC | none | 4 | 0.492 | 0.955 |
| **Cuvette pathlength – 10mm; wavelength region 950-1350nm** | | | | | |
| Td | none | none | 3 | 0.510 | 0.951 |
| Rd | ISC | none | 2 | 0.592 | 0.934 |
| Tc | Ordinary MSC | none | 3 | 0.480 | 0.957 |
| $\mu_s$ | ISC | none | 5 | 0.860 | 0.866 |
| $\mu_a$ | Ordinary MSC | none | 5 | 0.729 | 0.901 |
| $\mu_t$ | Ordinary MSC | none | 7 | 0.216 | 0.991 |

Among the models built with measurements obtained with 2mm sample thickness, there are two models which give slightly better estimation, namely models constructed on $\mu_s$ (figure 6.3c) and Tc (figure 6.3d). As can be seen from the figures, biomass estimates show relatively good agreement with experimental values. 5 LV's (figure 6.2) were chosen for the PLS model developed on collimated transmittance, and the model

based on the scattering coefficient was constructed with only 3 LV's. The RMSECV for the model built using $\mu_s$ is slightly higher compared to that obtained from the model built on Tc. It can be seen from the two PLS models (figure 6.3c and figure 6.3d) that for low biomass concentration the actual and the predicted values are in good agreement, however when observing higher biomass values, it is clear that the error between the results from the reference assay and the results from the predicted values is expanding. A plausible explanation is that this is due to the error in the biomass measurement itself where an average error of 8.2% has been estimated (chapter 4, section 4.2.1). An additional measurement challenge in the higher biomass values region was the fact that the total of 5 ml sample could not be filtered through one filter only due to blockage of the filter pores. For this reason the sample was separated in five samples of 1 ml and each sample was filtered in a consecutive order on 5 different filters, which could have led to an increase in the measurement error.
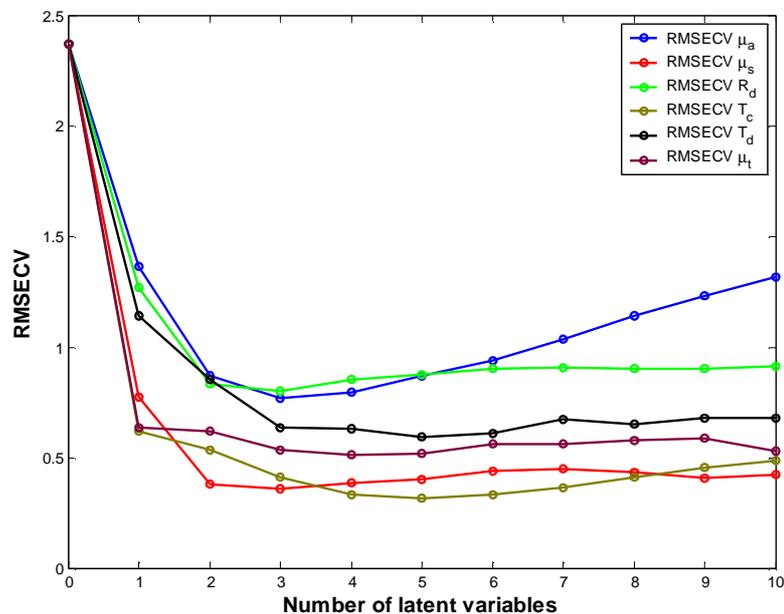


Fig. 6.2. RMSECV curves for $\mu_t$, $\mu_a$, $\mu_s$, Tc, Rd, and Td variables for prediction of biomass while raw measurements were taken with 2mm sample thickness.
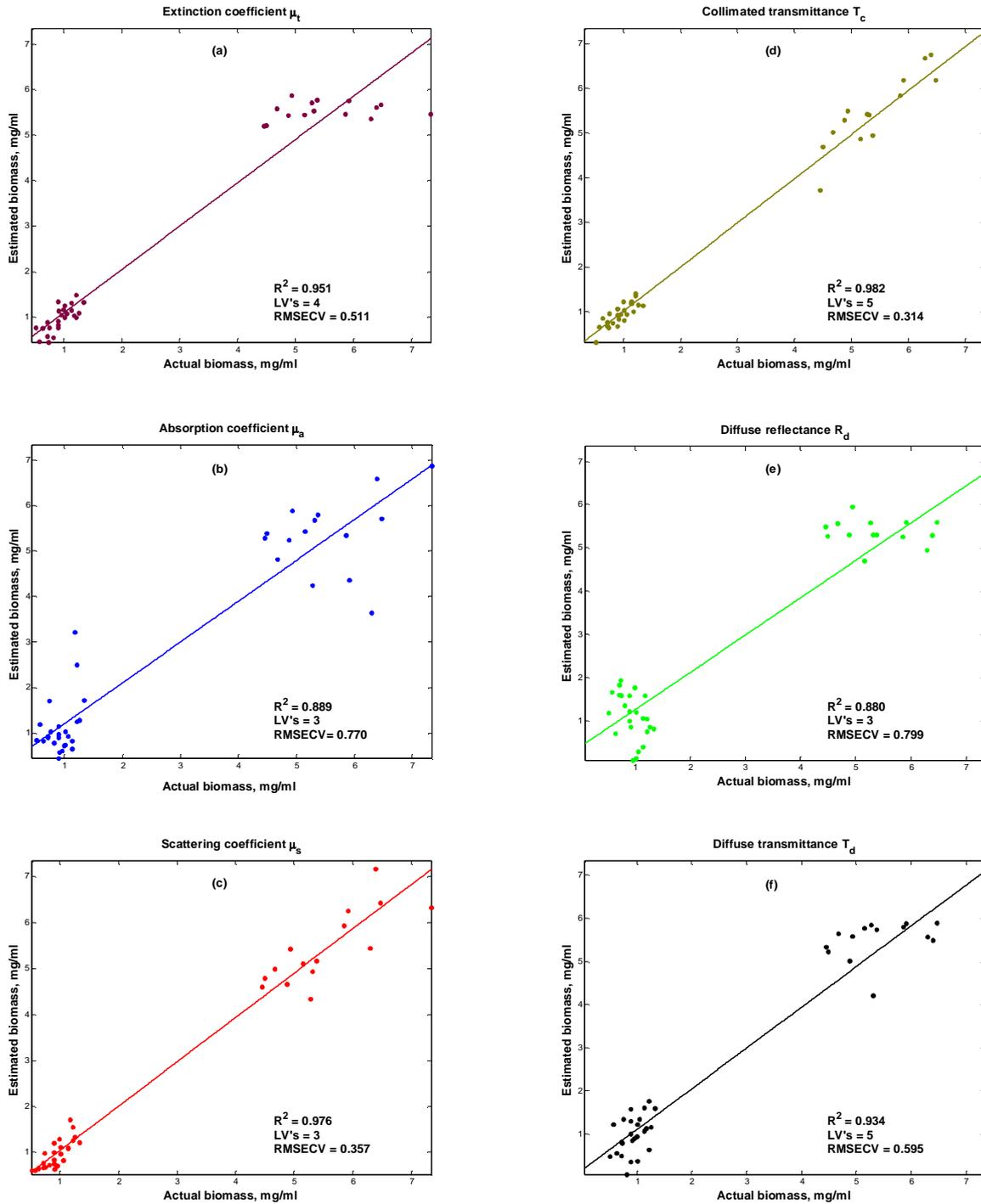
Fig. 6.3. PLS models performance based on $\mu_t$(a), $\mu_a$(b), $\mu_s$(c), Tc (d), Rd (e), and Td (f) variables for prediction of biomass while raw measurements were taken with 2mm sample thickness.

The best model based on the scattering coefficient was obtained without pre-processing and without applying derivatives but for the model based on Tc, the best results were obtained when the raw data were pre-processed with ordinary MSC.

While comparing the results from models obtained for Td (figure 6.3f), Rd (figure 6.3e), $\mu_t$ (figure 6.3a), and $\mu_a$ (figure 6.3b) with Tc and $\mu_s$, the last two have shown better performance.

The second section of table 6.1 shows the results obtained for measurements taken with samples of 4mm thickness. The results indicate better biomass predictions for models built on Tc (figure 6.5d). In this case, the RMSECV value is 0.220 mg/ml which is lower than that obtained using a 2mm sample thickness (0.314 mg/ml). However, the model built using the scattering coefficient (figure 6.5c) has a higher error (0.557 mg/ml) compared to the one built on measurements obtained with 2mm sample thickness (0.357 mg/ml). For this particular study, it has been noticed that with the increase in sample thickness, the bulk extinction coefficient shows better model performance in terms of RMSECV (0.492mg/ml) compared to bulk scattering coefficient (0.557mg/ml).

In the third section of table 6.1 results for models based on measurements obtained using 10mm sample thickness are shown. Biomass prediction (figure 6.6) for this sample thickness showed less accurate models with decline in the prediction ability, except the model built on the bulk extinction coefficient, where value for RMSECV was found 0.216 mg/ml, the lowest obtained while predicting biomass. This value placed bulk extinction coefficient to be the best prediction obtained using ordinary MSC and 7 LV's. Further analysis were completed in order to investigate whether this low error value obtained for 10mm sample thickness could be due to the fact that only the 950-1350nm wavelength region has been used for the construction of the model. For this purpose additional models were built using only this region for the other two sample thicknesses, namely 2mm and 4mm. The results are presented in table 6.2 where two different colours can be distinguished. Data in grey colour are obtained from models built on the entire region from 950nm to 1850nm that was presented earlier in table 6.1. The new

information in table 6.2 are data with black colour, and they were obtained from models

constructed on the wavelength region 950-1350nm.

Table 6.2 Results for PLS calibration models performance for prediction of biomass.

| Input spectral data | Pre-processing technique | Derivative choice | Number of LV's | RMSECV (mg/ml) | $R^2$ |
|---|---|---|---|---|---|
| **Cuvette pathlength – 2mm; grey colour\*; black colour\*\*** | | | | | |
| Td | none | none | 5 | 0.595 | 0.934 |
|  | EMSCL | none | 3 | 0.521 | 0.949 |
| Rd | EMSCW | none | 3 | 0.799 | 0.880 |
|  | ISC | none | 3 | 0.675 | 0.915 |
| Tc | Ordinary MSC | none | 5 | 0.314 | 0.982 |
|  | Ordinary MSC | none | 5 | 0.402 | 0.969 |
| $\mu_s$ | none | none | 3 | 0.357 | 0.976 |
|  | none | none | 2 | 0.576 | 0.938 |
| $\mu_a$ | none | none | 3 | 0.770 | 0.889 |
|  | ISC | none | 6 | 0.668 | 0.917 |
| $\mu_t$ | none | none | 4 | 0.511 | 0.951 |
|  | none | none | 6 | 0.397 | 0.970 |
| **Cuvette pathlength – 4mm** | | | | | |
| Td | ISC | none | 3 | 0.636 | 0.924 |
|  | ISC | none | 7 | 0.491 | 0.955 |
| Rd | none | none | 2 | 0.672 | 0.915 |
|  | none | none | 3 | 0.639 | 0.923 |
| Tc | Ordinary MSC | none | 7 | 0.219 | 0.991 |
|  | Ordinary MSC | none | 5 | 0.308 | 0.982 |
| $\mu_s$ | none | none | 2 | 0.557 | 0.942 |
|  | none | none | 5 | 0.516 | 0.950 |
| $\mu_a$ | EMSCL | none | 5 | 0.898 | 0.852 |
|  | ISC | none | 4 | 0.586 | 0.936 |
| $\mu_t$ | ISC | none | 4 | 0.492 | 0.955 |
|  | none | none | 6 | 0.419 | 0.967 |
| **Cuvette pathlength – 10mm** | | | | | |
| Td | none | none | 3 | 0.510 | 0.951 |
| Rd | ISC | none | 2 | 0.592 | 0.934 |
| Tc | Ordinary MSC | none | 3 | 0.480 | 0.957 |
| $\mu_s$ | ISC | none | 5 | 0.860 | 0.866 |
| $\mu_a$ | Ordinary MSC | none | 5 | 0.729 | 0.901 |
| $\mu_t$ | Ordinary MSC | none | 7 | 0.216 | 0.991 |

\* models were built on data from 950-1350nm&1550-1850nm

\*\* models were built on data from 950-1350nm

The new models have shown small changes with the modified set of data. For data using 2mm sample thickness, they led to a slight decrease in RMSECV for Td, Rd, $\mu_a$, and $\mu_t$, but the number of LV's for the parameters $\mu_a$, and $\mu_t$ have increased at the same time. Models based on $\mu_s$ and Tc have shown slight deterioration in terms of RMSECV. For the 4mm sample thickness, the change in model performance was similar to the 2mm sample thickness measurements. Models constructed on Td, Rd, $\mu_s$, $\mu_a$, and $\mu_t$ have demonstrated decrease in RMSECV while the model based on Tc demonstrated increase of RMSECV. LV's undergo small changes as well. It is apparent from the comparison between different models parameters shown in table 6.2 that limiting the wavelength region only to a scattering dominated region, could not lead to great improvement on the constructed models. This suggests that the improved model performance for biomass prediction built on $\mu_t$ obtained using a 10mm sample thickness was not due to the restriction in the wavelength range.
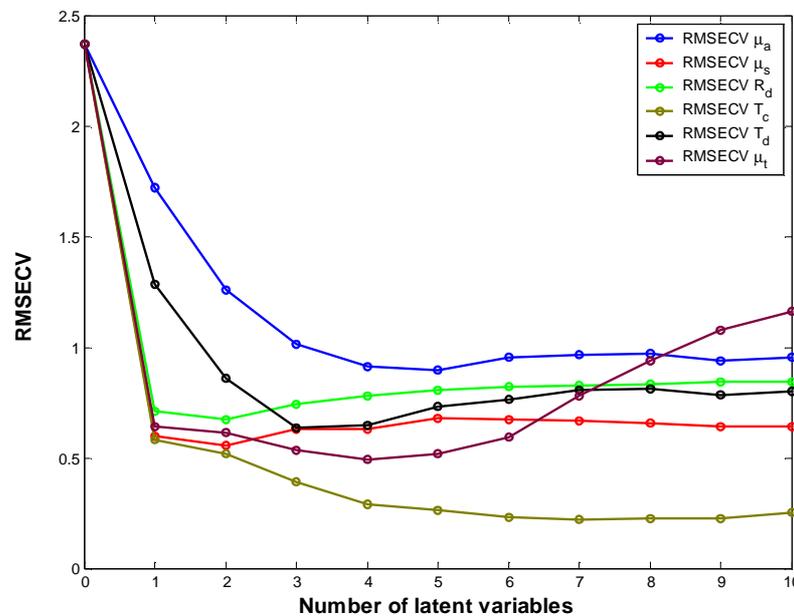


Fig. 6.4 RMSECV curves for $\mu_a$, $\mu_s$, Tc, Rd, and Td variables for prediction of biomass while raw measurements were taken with 4mm sample thickness.
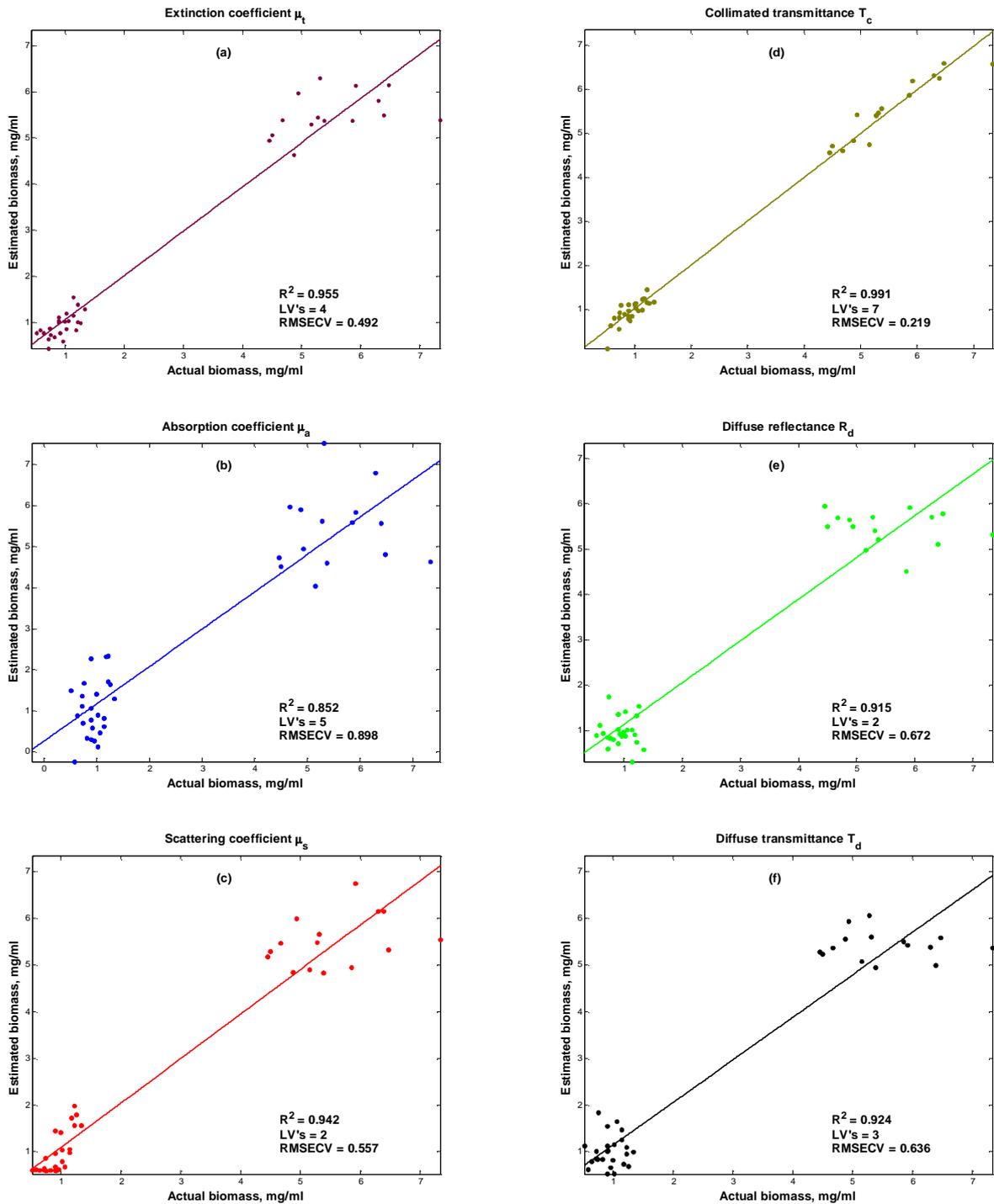
Fig. 6.5 PLS models performance based on $\mu_t$(a), $\mu_a$(b), $\mu_s$(c), Tc (d), Rd (e), and Td (f) variables for prediction of biomass while raw measurements were taken with 4mm sample thickness.
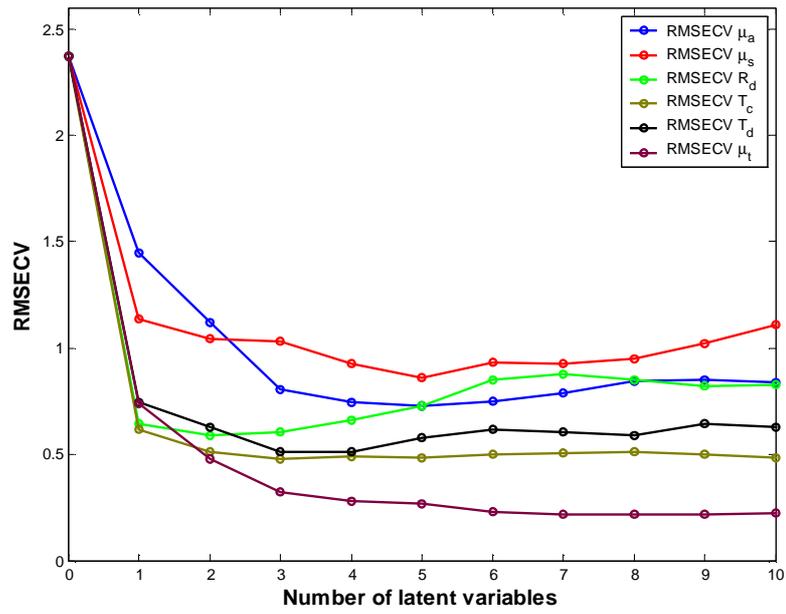
Fig. 6.6 RMSECV curves for $\mu_t$, $\mu_a$, $\mu_s$, Tc, Rd, and Td variables for prediction of biomass while raw measurements were taken with 10mm sample thickness.
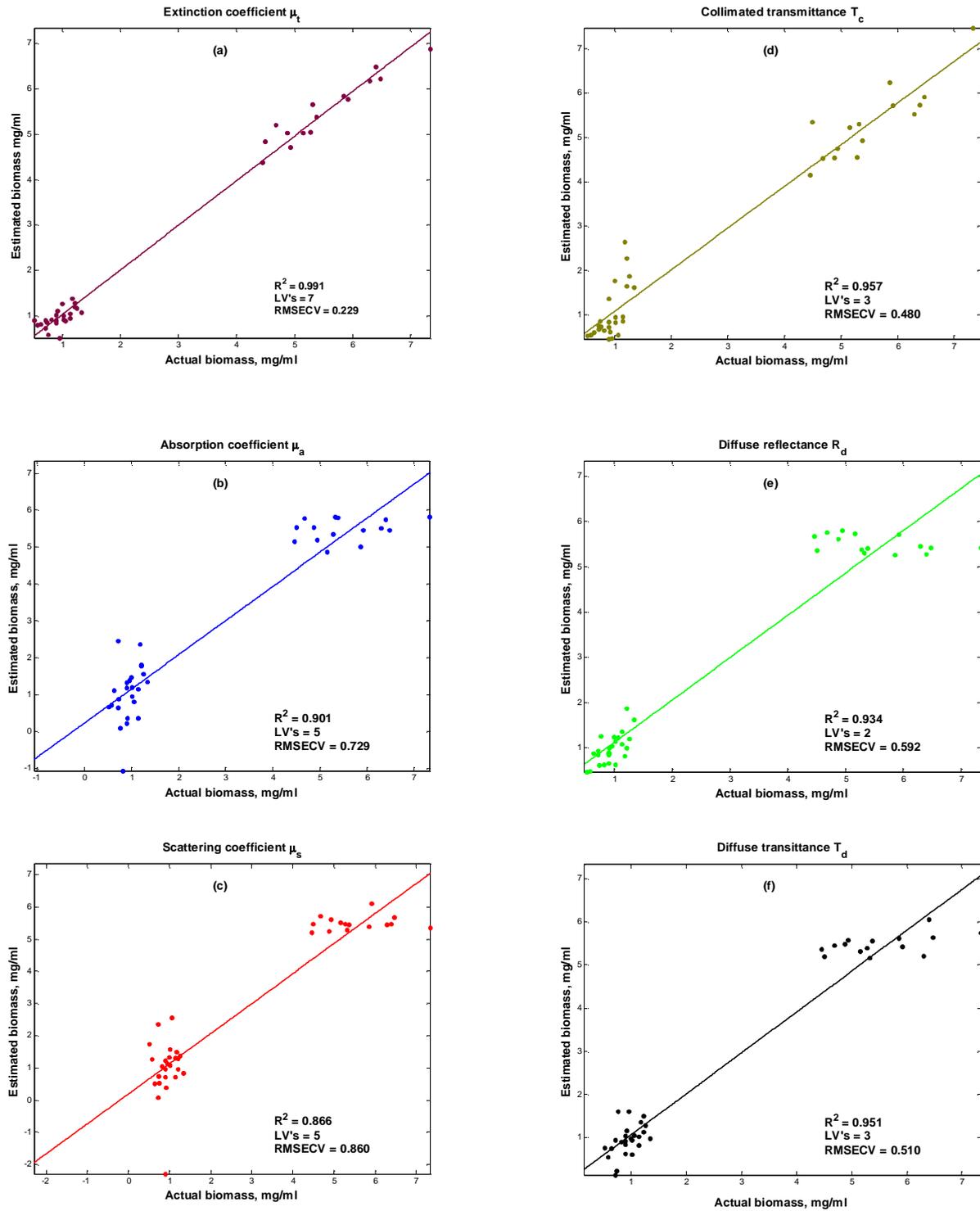
Fig. 6.7 PLS models performance based on $\mu_t$(a), $\mu_a$(b), $\mu_s$(c), Tc (d), Rd (e), and Td (f) variables for prediction of biomass while raw measurements were taken with 10mm sample thickness.

In general, reasonable (low RMSECV) prediction results were obtained for all three

sample thicknesses, and for each all pre-processing methods have been examined. The

following observations were made by examining table 6.1 in conjunction with the results from the RMSECV and the PLS models shown in figure 6.2, figure 6.4, and figure 6.6.

(a) Five different pre-processing techniques (figure 6.1) have been applied for the construction of models built on Rd, Td, Tc, $\mu_a$, $\mu_s$, and $\mu_t$ in order to compare differences in model performances and to determine the technique to be used for providing the best results. It was found that the effectiveness of the pre-processing method is dependent on the measurement configuration (when raw measurements are used) or on the optical parameter (when extracted optical properties are used) that have been used for the construction of the models. The ISC method, the Ordinary MSC method, and the EMSC (in its two forms EMSCW and EMSCL) method, which were found to be the best performing scatter correction techniques for predicting biomass concentration, in the present work, need a fewer number of latent variables (LV's) and, therefore, are reported.

(b) Application of first and second derivatives led to degradation in the model performance (higher number of LV's and higher RMSECV values).

(c) PLS models for biomass concentration were built using directly raw measurements (Td, Rd, and Tc) and extracted optical parameters ($\mu_a$, $\mu_s$, $\mu_t$). The performances of the constructed models have been compared and the effect of sample thickness was evaluated. Results from RMSECV and LV's have shown that models built on bulk extinction coefficient and scattering coefficient can provide better predictive models compared to models built on absorption coefficient.

(d) Models were built without pre-processing techniques and with pre-processing techniques however, the scenario with pre-processing techniques yielded slightly better prediction results in terms of RMSECV.

(e) Best performing models for biomass prediction were obtained with $\mu_t$ extracted from

measurements made using 10mm sample thickness (7 LV's and RMSECV=0.216 mg/ml), and models constructed on Tc measurements made using 4mm sample thickness (7 LV's and RMSECV=0.219 mg/ml). Models built on $\mu_s$, extracted from measurements obtained with 2mm sample thickness (3 LV's and RMSECV=0.357 mg/ml) and with 4mm sample thickness (2 LV's and RMSECV=0.557 mg/ml) were poorer then the models constructed with $\mu_t$ and Tc in terms of RMSECV but less LV's were needed to model the prediction.

From the above summary, it can be seen that development of models for prediction of biomass based on scattering and absorption coefficients is feasible, and both $\mu_a$ and $\mu_s$ are rich in information related to the biomass. However, results have shown that while the proposed method led to a $\mu_s$ based model that performed better then the ones built using Rd and Td, its performance was equal compared to when Tc measurements were used to construct the models. Even by adopting a variety of pre-processing techniques it was evident that it is not possible to generate more accurate predictive models built on $\mu_s$. In this case, improvement of the performance and greater accuracy of the $\mu_a$ and $\mu_s$ based models will depend on some conditions, including integrating sphere geometry (errors in Rd, Td and Tc measurement) and convergence issues. First, future efforts should focus on improving measurement configuration by introducing double integrating sphere configuration for the measurements instead of single integrating sphere, or even further, the sphere setup can be substituted by spatially-resolved spectroscopy systems [6, 7]. Double integrating sphere could provide near-simultaneous results for the three measurements (Tc, Td, and Rd) and to enable collection of unique information during the analysis of dynamic systems, such as growing biological suspensions. It helps also to avoid change over between the different measurements and to minimize the time necessary for their collection. Spatially-resolved systems on the other hand, enable the determination of two optical properties (absorption and effective scattering

coefficient) by means of relatively simple and inexpensive equipment. The quantity of diffusely reflected light is only monitored, as a function of the distance from the centre of the incident beam. Measurements are fast and could be used for monitoring optical properties of dense biological suspensions where collimated transmittance could no longer be obtained. As a second drawback it was underlined that convergence issues in the inversion techniques do not allow extraction of the full set of optical properties in the region of the water absorption peak. In this sense, inversion technique needs to be improved in order to provide more complete and accurate information of optical properties changes in the region 1350-1550nm.

### 6.3.2. Models for the prediction of glucose concentration

The raw spectral measurements (in the range 950nm-1850nm) and reference measurements used for the construction of predictive models for glucose concentration were collected in the same order as it was in the case of biomass prediction. Raw spectra was converted to absorbance units and pre-processed with the techniques presented in figure 6.1. Models were investigated for all five variables (Td, Tc, Rd, $\mu_a$, and $\mu_s$), however only two (Td and Tc) of the studied variables produced accurate models for prediction of glucose concentration. Table 6.3 shows the performance of the models and the pre-processing techniques that have been applied.

PLS models were developed using measurements from three different sample thicknesses: 2mm, 4mm, and 10mm. Table 6.3 includes only models constructed on Tc, Td, and $\mu_a$.

Table 6.3 Results for PLS calibration models performance for prediction of glucose concentration.

| Cuvette pathlength – 2mm | | | | | |
|---|---|---|---|---|---|
| Input spectral data | Pre-processing technique | Derivative choice | Number of LV's | RMSECV (mg/ml) | $R^2$ |
| Td | SNV | none | 14 | 3.142 | 0.938 |
| Tc | EMSCW | none | 8 | 2.802 | 0.951 |
| $\mu_a$ | none | none | 2 | 10.128 | 0.379 |
| **Cuvette pathlength – 4mm** | | | | | |
| Td | none | none | 13 | 6.257 | 0.769 |
| Tc | Ordinary MSC | none | 13 | 3.427 | 0.930 |
| $\mu_a$ | EMSCL | none | 2 | 9.949 | 0.390 |
| **Cuvette pathlength – 10mm** | | | | | |
| Td | none | none | 9 | 4.589 | 0.872 |
| Tc | none | none | 7 | 9.156 | 0.507 |
| $\mu_a$ | none | none | 2 | 12.508 | 0.115 |

Models built on other three variables, Rd/$\mu_s$/$\mu_t$ have shown low values for coefficient of determination and very high values for RMSECV, and the results were as follows:

(1) 2mm sample thickness: $R^2$ = 0.36/0.38/0.39 and RMSECV = 9.93/11.56/10.11 mg/ml respectively for Rd/ $\mu_s$/$\mu_t$;

(2) 4mm sample thickness: $R^2$ = 0.10/0.35/0.44 and RMSECV = 10.77/12.15/9.62 mg/ml respectively for Rd/ $\mu_s$/$\mu_t$;

(3) 10mm sample thickness: $R^2$ = 0.20/0.23/0.33 and RMSECV = 11.09/12.4/11.04mg/ml respectively for Rd/ $\mu_s$/$\mu_t$.

The results show that for all three sample thicknesses, the RMSECV values were found

high; however the best RMSECV value was calculated while building models with spectral data obtained with 2mm sample thickness (figure 6.8). The reasons for the high error values are possibly embedded in the following three reasons. First, the average error for glucose measurements (chapter 4, section 4.2.1) has been estimated close to 2% based on obtained measurements while 4mg/ml of glucose was considered as an initial concentration. In the present chapter 6, glucose concentration has been used in much higher ranges, between 8mg/ml and 45 mg/ml. This high level of glucose exceeds from 2 to 10 times the one that has been used for the error estimation in chapter 4, as well as exceeds the range of the test kit for glucose concentration determination, and a dilution factor of 500 folds has been applied. In practice it is well known that the higher the dilution factor, the higher the measurement error is expected to be. The second reason relates to the integrating sphere geometry and the Td measurement where the diffuse transmittance is affected slightly by introducing the sample at the entrance port. In this situation the sample possibly reflects some of the light inside the sphere instead to be collected all as a transmitted light, which in all cases will influence the measurement. The third influential error arises from lateral losses of light due to the port sizes, where some part of the initial beam reflects at wide angles and could be intercepted by the sphere wall. This leads to overestimation [8, 9]and large errors when extracting $\mu_a$ for the case of this biological system where glucose signal is very small. Previous studies with polystyrene particles [10] have shown that $\mu_a$ is much more accurate when samples with high concentration of particles were studied. The peak where polystyrene absorbs was expected to increase systematically with concentration but for samples with the lowest concentration the absorption spectra did not follow the right order.

From the obtained results it can be seen that models built on absorption coefficient show RMSECV values of 10.128 mg/ml (2mm sample thickness, figure 6.8a), 9.949 (4mm sample thickness, figure 6.9a), and 12.508 (10mm sample thickness, figure 6.10a),

whereas models built on Td/Tc show values respectively of 3.142/2.802mg/ml (2mm sample thickness), 6.257/3.427mg/ml (4mm sample thickness), and 4.589/9.156mg/ml (10mm sample thickness). Despite the fact of high RMSECV values, the absorption coefficient based models in all three cases have shown that only 2 LV's are necessary to model the glucose concentration, compared to the raw spectral measurements of Td/Tc where the number is 14/8 LV's (2mm sample thickness), 13/13 LV's (4mm sample thickness), and 9/7 LV's (10mm sample thickness).
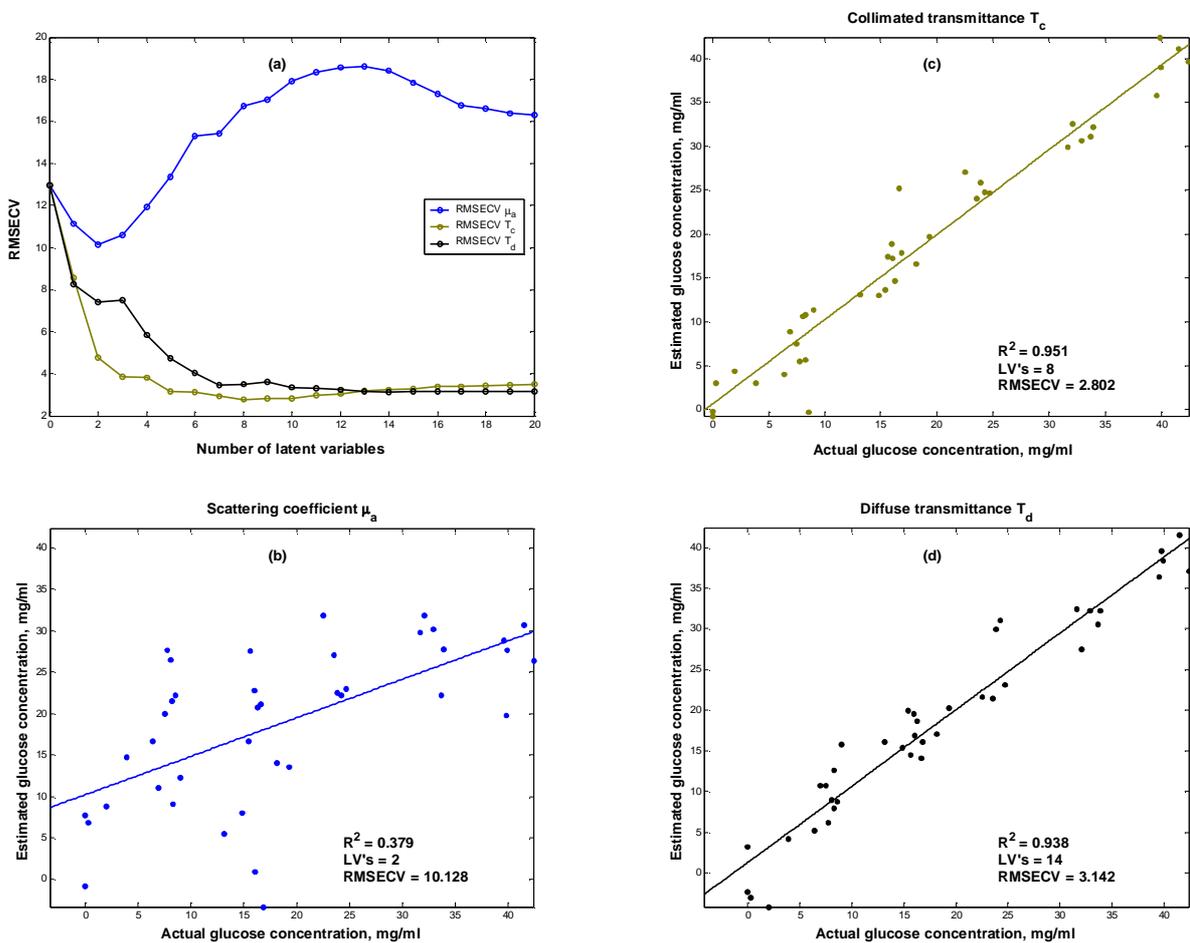


Fig. 6.8. RMSECV (a) curve and PLS models performance based on $\mu_a$(b), Tc (c), and Td (d) variables for prediction of glucose while raw measurements were taken with 2mm sample thickness.
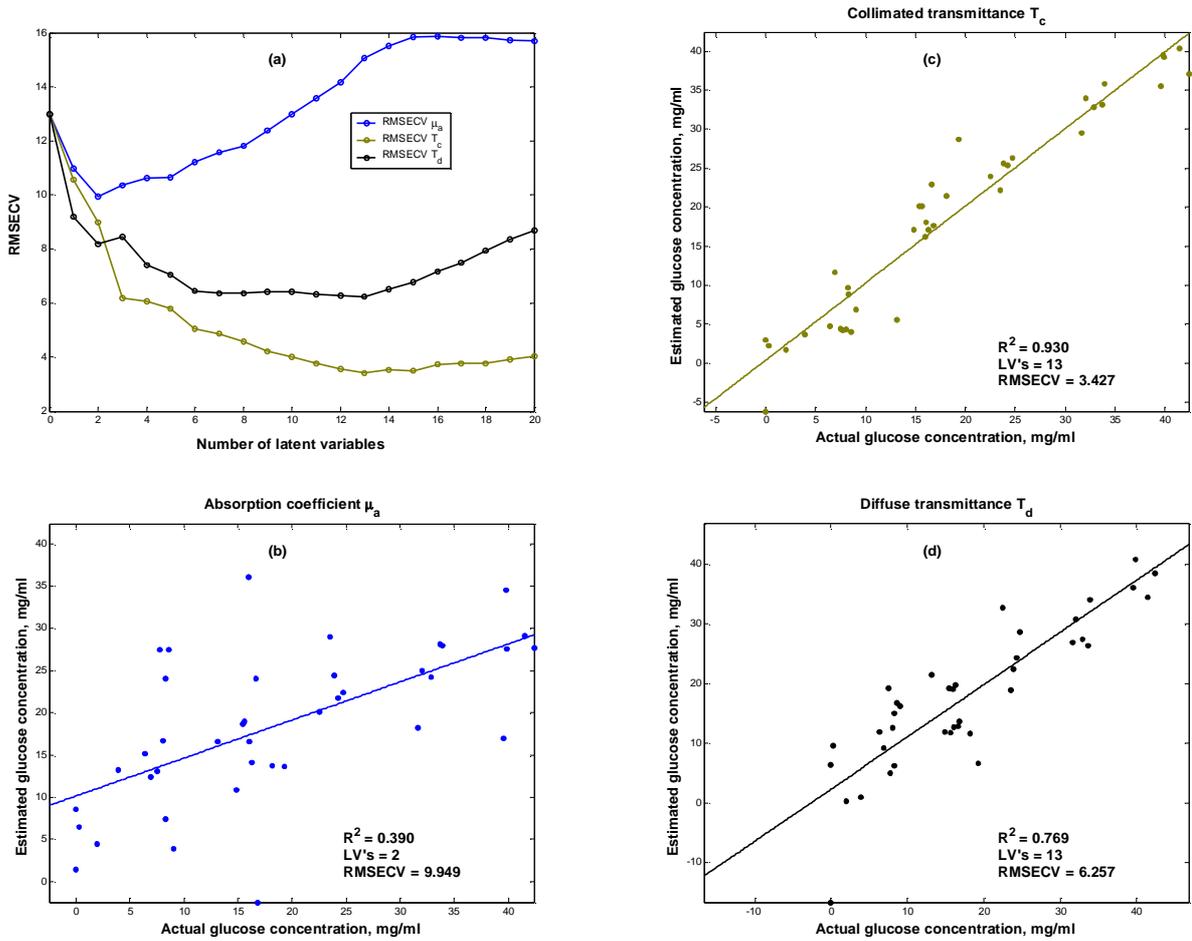
Fig. 6.9. RMSECV (a) curve and PLS models performance based on $\mu_a$(b), Tc (c), and Td (d) variables for prediction of glucose while raw measurements were taken with 4mm sample thickness.
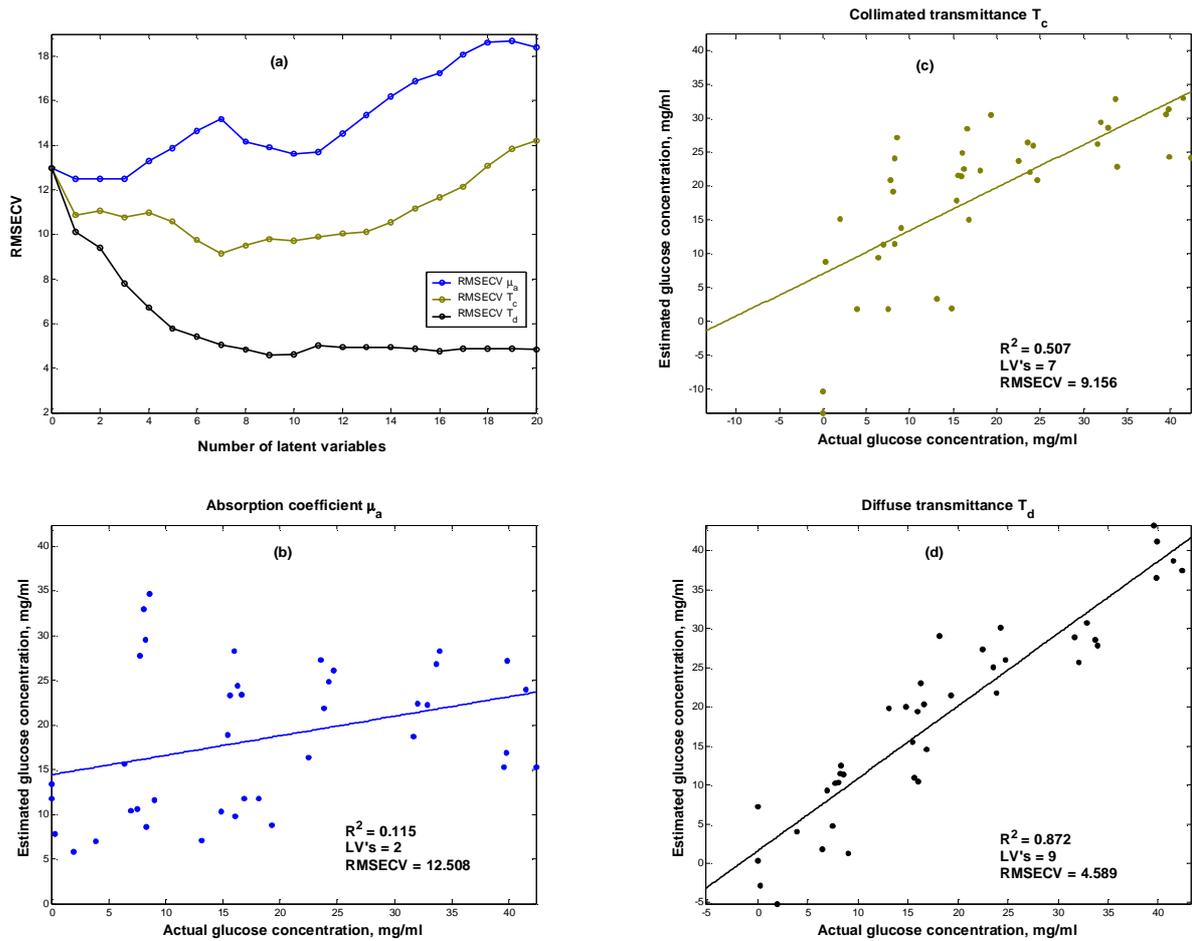
Fig. 6.10. RMSECV (a) curve and PLS models performance based on $\mu_a$(b), Tc (c), and Td (d) variables for prediction of glucose while raw measurements were taken with 10mm sample thickness.

If the error sources for absorption coefficient itself, associated with the measurement setup and the calculation routines could be successfully eliminated and the latent variables number is still 2, then a significant improvement in the glucose prediction could be expected. However, further studies are needed to investigate this effect. The multiple coefficient of determination for the models for glucose prediction deteriorate when thicker samples are used, which is the opposite effect to what was found with models built for biomass prediction, where increase of sample thickness led to better prediction. PLS models built on Td/Tc for 2mm sample thickness have shown to have the best results for coefficient of determination ($R^2$= 0.938/0.951).

The multiple coefficients of determination for models built on $\mu_a$ possess very low values for all three sample thicknesses and can not be considered as successful ($R^2$ = 0.379/0.390/0.115 respectively for 2mm/4mm/10mm sample thicknesses). However the model does distinguish between high and low glucose levels.

## 6.4. Conclusions

The near-infrared spectral information from *Bacillus subtilis* growing in a liquid culture was investigated in this chapter along with the possibility of predicting key analytes such as biomass and glucose concentration. The study was focused on the region 950-1850nm, and the performance of the predictive models was studied as a function of the sample thickness. This range was chosen because it includes the first and second overtone bands of proteins, sugars, and other organics. It has been demonstrated that the two new variables, namely absorption coefficient $\mu_a$ and scattering coefficient $\mu_s$ (introduced as indirect parameters) can be used to predict biomass and glucose concentration in a *Bacillus subtilis* fermentation process. PLS models were constructed on absorption and scattering coefficients and then compared with models constructed with traditional spectra (e.g. transmittance and reflectance). Improvements over the current calibration model approaches were sought.

As a result, PLS models that have been developed based on scattering coefficients demonstrated good response while predicting biomass concentration for the investigated biological system. In the present study, the obtained results have equal or greater accuracy when compared to results based on transmittance and reflectance measurements. It was shown that by separating scattering from absorption and then building calibration models offered increased accuracy on the prediction of biomass concentration. This was less accurate with the reflectance and transmittance

measurements since the raw spectra depend on the pathlength of the photons due to the multiple scattering of light. Meanwhile, models based on absorption coefficient have demonstrated limited success and were found to distinguish only between low and high level of glucose concentration. While discussing the low predictive ability of the absorption coefficient based models, the main point to be noted is that the error in extracting $\mu_a$ itself needs to be reduced in order to achieve greater accuracy in the predictive models. This can be done by improvements to measurements configuration and overcoming convergence problems related with the inversion technique.

To select measurements from only one sample thickness which can be used for the prediction of both glucose and biomass concentration could be erroneous given results that have been obtained during this investigation. Good models for biomass prediction were obtained when built on measurements taken with 10mm sample thickness for example, whereas the best glucose prediction model was obtained with 2mm sample thickness. Significant changes occur during the course of the fermentation and adopting a single sample thickness for the two constituents to predict can limit the accuracy of prediction. The presented results indicate that sample thicknesses should be chosen based on the specific analyte of interest in order to obtain the best performing models.

A number of issues still remain to be addressed. In order to ensure robustness of the method it is necessary to expand the data set. For this to be achieved, information from more cultivation runs has to be collected in order to increase the population of the calibration samples. These samples have to be again analyzed by the reference methods for the constituent of interest and by spectroscopic measurements. Sensitivity of the models and the effectiveness of the proposed methodology, need to be further investigated by reflecting real industrial system.

### References

1. Williams, P. and K. Norris, eds. *Near-Infrared Technology in the Agricultural and Food Industries*. Second ed. 2001, American Association of Cereal Chemists, Inc.: St. Paul, Minnesota, USA.

2. Geladi, P., D. MacDougall, and H. Martens, *Linearization and scatter-correction for near-infrared reflectance spectra of meat.* Applied Spectroscopy, 1985. **39**(3): p. 491-500.

3. Naes, T., T. Isaksson, T. Fearn, and T. Davies, *A user-friendly guide to multivariate calibration and classification.* 2004: NIR Publications.

4. Thennadil, S.N. and E.B. Martin, *Empirical preprocessing methods and their impact on NIR calibrations: a simulation study.* Journal of Chemometrics, 2005. **19**: p. 77-89.

5. Spiegelman, C., M.J. McShane, M.J. Goetz, M. Motamedi, Q.L. Yue, and G.L. Cote, *Theoretical justification of wavelength selection in PLS calibration: development of a new algorithm.* Analytical Chemistry, 1998. **70**(1): p. 35-44.

6. Hjalmarsson, P. and S.N. Thennadil, *Spatially-resolved in-Vivo measurement system for estimating the optical properties of tissue in the wavelength range 1000-1700nm - art. no. 662805.* Diagnostic Optical Spectroscopy in Biomedicine Iv, 2007. **6628**: p. 62805-62805.

7. Hjalmarsson, P. and S. Thennadil. *Determination of glucose concentration in tissue-like material using spatially resolved steady-state diffuse reflectance spectroscopy.* in *Proceedings of SPIE: Complex Dynamics and Fluctuations in Biomedical Photonics V* (2008).

8. Torres, J.H., A.J. Welch, I. Cilesiz, and M. Motamedi, *Tissue Optical Property Measurements - Overestimation of Absorption-Coefficient with Spectrophotometric Techniques.* Lasers in Surgery and Medicine, 1994. **14**(3): p. 249-257.

9. Vries, G.d., J.F. Beek, G.W. Lucassen, and M.J.C.v. Gemert, *The effect of light losses in double integrating spheres on optical properties estimation.* IEEE Journal of Selected Topics in Quantum Electronics, 1999. **5**(4): p. 944-947.

10. Steponavicius, R. and S.N. Thennadil, *Extraction of chemical information of suspensions using radiative transfer theory to remove multiple scattering effects: application to a model two-component system.* Analytical Chemistry, 2009. **81**(18): p. 7713-7723.

# Conclusions and Future Work

## 7.1 Conclusions

The ultimate success of bioprocesses and high quality of the end product heavily depend on well defined parameters and conditions which requires monitoring during the production process. For this, measurement techniques such as chromatographic methods, spectroscopic methods and biosensors were developed and widely applied for analysis of the cultivation media and its constituents. They were identified and discussed in Chapter 2 of this thesis. Although these techniques are potentially available, the reliability and sensitivity of some are still questionable. In addition, some analyses are time consuming, very complex and require the use of environmentally unacceptable solvents. From this perspective, Near-Infrared spectroscopy has the potential to improve bioprocesses because the measurements are rapid, non-destructive, compatible with fiber optics, and require minimal or no sample preparation. However, the interpretation of spectral information is not straightforward as a result of the chemical and physical complexity of the growth media, the highly overlapped absorption bands, and the presence of multiple light scattering effects, which leads to the degradation of chemical information due to variation in the pathlength travelled by the photons. Consequently, these variations have an impact on the robustness of the multivariate calibration techniques, usually employed to extract relevant information from a given sample's spectra. Keeping the above facts in mind, the investigation in this work has been initiated with the idea of removing multiple scattering effects from the raw measurements using physics of light propagation in order to improve robustness of PLS models built for prediction of given analytes.

For this, a method has been developed for the estimation of absorption and scattering spectra in the Near-Infrared region for *Bacillus subtilis* growing in a liquid culture where

multiple scattering is considered. By separating absorption from scattering effects, simpler and most robust models can be built. This can be achieved due to the fact that absorption and scattering data discriminates better between variations in the constituents of the growth medium and constituents of the biomass, by taking advantage of the different information content of the two coefficients ($\mu_a$ and $\mu_s$). The initial application focus of the thesis was on fermentation processes. However, due to their complexity, a simple system of *Bacillus subtilis* growing in liquid culture was chosen as a substitute in order to test repeatability, feasibility and robustness of the proposed methodology.

The investigation was started by collecting traditional spectroscopic measurements (Diffuse Transmittance-Td, Diffuse Reflectance-Rd, and Collimated Transmittance-Tc) for the given culture using a single integrating sphere setup, and then these raw data were used in an inversion method (Inverse Adding Doubling Method - IAD) in order to solve the radiative transfer equation (RTE) for separating absorption and scattering effects by extracting the bulk absorption and scattering coefficients and the anisotropy factor ($\mu_a$, $\mu_s$, and $g$). Previously, the method for extracting the optical properties and the measurement system was independently validated (Chapter 3) for the near-infrared region by using a suspension of low concentration polystyrene microspheres suspended in water. The validation procedure was carried out in two steps: (1) from the first step it was found that experimentally obtained values for bulk scattering coefficient ($\mu_s$) were in good agreement with theoretically computed values for $\mu_s$; and (2) from the second step, experimentally obtained values and calculated values for spectral measurements (Td, Rd, and Tc) were compared and it was shown that they are in a very good agreement over the wavelength region studied in this thesis (950-1850nm).

Continuing with the investigation, the next step (Chapter 4) of the study was to focus on the methodology for extracting and effectively separating the optical properties and to

investigate the nature and extent of their changes over growth cycle. For this, a simple culture system was chosen consisting of *Bacillus subtilis* growing in an aqueous solution (minimum medium). The changes in the optical properties during growth, stationary, and decline phases were studied using measurements in the wavelength region of 950-1850nm. During this investigation, it was shown that absorption and scattering property changes during cultivation of bacterial culture can be consistently extracted from measurements under multiple scattering conditions. In addition, it was found that the major changes monitored during the cultivation were due to biomass. This has been reflected by the fact that the greatest relative changes were seen during the growth phase, predominately in the scattering spectra. Meanwhile, biomass was also found to affect the absorption spectra. However, absorption due to biomass in the near-infrared region was found small compared to absorption due to water, therefore the relative changes due to an increase in biomass were small. The anisotropy factor was found to exhibit changes in the region 950-1350nm throughout the growth phase. In the region of the water absorption peak, effective extraction of optical properties was identified as a difficult step. With sample thickness of 4mm, optical properties results confirmed that the inversion using the adding-doubling method was not able to converge in this specific region (1350-1550nm) due to the high level of absorbance. One approach to improve the estimates was to identify the optimal sample thickness which would provide accurate estimates of the optical parameters. Further, since the ultimate goal of extracting the optical properties was to build robust calibration models for estimating the concentration of various constituents in the sample, it was important to know how the sample thickness will affect the estimation of the optical properties and thus the performance of the corresponding calibration models. For this, the accuracy and feasibility of the method were verified for three different sample thicknesses (2mm, 4mm, and 10mm). Optical properties, in addition to the raw spectral measurements obtained for all three

different samples were analyzed and compared. The extracted optical properties were found fairly consistent even though they were obtained from measurements from different sample thicknesses. Moreover, further analyses were also carried out to examine how the relationship between the biomass and the optical properties as well as the reflectance and transmittance are affected by the changes in the sample thickness. It was found that the nature of the relationship between reflectance measurements and biomass can vary depending on the sample thickness. Inspite of this, the estimates of optical properties remained stable. In addition, further evidence supporting the consistency in the extraction of the optical properties was found when measurements were collected with cuvettes having different refractive indexes leading to significant difference in the input data, did not appreciably affect the extracted values of $\mu_a$, $\mu_s$, and $g$.

In terms of optimal sample thickness for the extraction of the optical properties, analysis indicates that it will be difficult to recommend only one thickness of sample when the entire wavelength region (950-1850nm) is considered. Analyses have shown that measurements with 2mm sample thickness did partially improve the final results, especially in the water peak region, however only the absorption coefficient was fully extracted. Measurements with 10mm cuvette limited the optical properties extraction only in the region 950-1350nm. It was found that although different wavelengths regions tolerate different sample thicknesses, the appropriate sample thickness to be used for this case study was 4mm. This sample thickness allowed effective study of biomass in both scattering and absorbing region. If only scattering region needs to be studied, then 10mm sample would be better solution, whereas 2mm sample was most appropriate for the absorbing region.

Furthermore, variations in the optical properties and in the raw measurements over the course of cultivation were investigated by extracting values at 1050nm and 1602nm

single wavelengths and displaying them as a function of biomass. Accordingly, the knowledge gained during this investigation was encouraging, and repeatedly confirmed that linear relation could be established between the newly introduced variables and the analytes of interest. Given this fact and the consistency from the extraction, the next question was whether optical properties are sufficiently accurate such that models for predicting of glucose and biomass will perform better than when models are built using the reflectance or transmittance spectra directly.

For this, a new strategy from experimental point of view was designed, that combines broad range of glucose concentration and a two-step approach for the construction of the predictive models. First, the spectral measurements (Td, Rd, and Tc) were used to extract the optical properties. Second, PLS models were built based on spectral measurements and on the extracted optical properties for predicting glucose and biomass during bacterial growth cycle. The performances of these various models were compared. As a result, scattering coefficient based models demonstrated good performance while predicting biomass. The obtained results have shown equal or in some cases greater accuracy when compared to models constructed on transmittance and reflectance measurements. On the other hand, absorption coefficients based models demonstrated limited success, and were able to distinguish only between low and high level of glucose concentration. To improve the predictive ability of the absorption coefficients based models, first and foremost, the error in extracting $\mu_a$ itself needs to be reduced. This can be done by improvements to measurements configuration and overcoming convergence problems related with the inversion technique. If assessment of the models is needed, overall, the best model for biomass prediction was obtained when absorption and scattering coefficients were extracted and combined together as a total extinction coefficient (RMSECV=0.216mg/ml and $R^2$=0.991) from measurements taken with 10mm sample thickness, along with model constructed on collimated transmittance collected

with 4mm sample thickness (RMSECV=0.219mg/ml and $R^2$=0.991). Accordingly, glucose concentration was fairly well predicted from collimated transmittance measurements taken with sample of 2mm thickness (RMSECV=2.802mg/ml and $R^2$=0.951). Although $\mu_a$ was used for the development of models for glucose prediction, the final results (high RMSECV values and low $R^2$) have shown that the PLS models did not pick up the entire information for the analyte of interest.

In summary, the findings and the overall investigation provide evidence that by separating absorption from scattering for *Bacillus subtilis* culture growing in a minimum medium, it is possible to construct models that describes required biomass and glucose information based on new variables, namely absorption ($\mu_a$) and scattering coefficients ($\mu_s$). Once $\mu_a$ and $\mu_s$ are extracted, they are independent of the pathlength variations in the photons travelling through a medium, and as such, they can provide fundamental information while new calibration models are constructed. The present study also highlights the fact that optical properties information remain stable although $\mu_a$ and $\mu_s$ were extracted under various conditions, such as different sample thickness or refractive indexes. The overall investigation provides the evidence for the fact that, using NIR in a fermentation process, it is possible to obtain identical optical information through the approach proposed in this thesis. The estimated optical properties could be used for the construction of predictive models, and these models could be robust to sample variations even if there are general differences in the raw measurements. It provides fundamental information on the scattering spectra of biomass that could prove useful in bioprocess monitoring applications, as was demonstrated for predicting biomass in *Bacillus subtilis* fermentation. It also highlights the fact that if near-infrared spectra is collected from different instruments, the extracted optical properties information could still be identical.

This is an important step in the wider application of the method because it could be possible to aid in development of calibration equations and their transfer from one instrument to another. However, several issues remain to be addressed, such as upgrading of the measurement setup along with improvement of the inversion technique.

## 7.2 Recommendations for future work

On the future research, a number of issues have been arisen, opening a window for further investigation:

In terms of improving the measurement setup it is recommended double integrating sphere to be used for future studies instead of single integrating sphere in order to (1) minimize light losses from the port of the sphere resulting in overestimation of the absorption coefficient, and (2) to avoid delay between the different measurements which play an important role when dynamic systems such as biological suspensions are considered. If bacterial systems with industrial application will be studied, in this case increase of biomass concentration will be regarded as an obstacle because measurements of collimated transmittance can not be obtained with such a clear signal as in this work. For this kind of application time-resolved and spatially-resolved techniques can be used[1, 2].

In addition to the measurement setup, improvement of the inversion technique for extraction of the optical properties is much needed. It has been shown in Chapter 4 and Chapter 5 that reliable estimates for absorption coefficient in the region 1350-1550nm can be obtained only if measurements are collected with 2mm sample thickness. More work on the inversion technique need to be carried out in order to reach convergence and to obtain reliable information on the optical properties in the region of the water absorption peak.

On the goal for providing accurate and precise determination of the analytes of interest by the use of predictive PLS models built in this study, special attention needs to be given to the reference methods. In order to decrease the experimental error (average error for glucose concentration measurements $\sim$ 2% and for biomass concentration determination $\sim$ 8.2%), and as a result from this to improve the predictive ability of the models, it is extremely important that the chosen reference method is accurate, precise, and stable.

It would beneficial to the present study if growing cultures with different microorganisms can be investigated in order to study the changes and variations that could possible occur due to the various morphological forms. Since morphological changes could influence spectroscopic measurements it would be interesting to investigate optical properties behaviour of different structure of microorganisms, such as mycelia or yeast base system.

In order to test the sensitivity of the proposed methodology towards the scale of operation, the situation considered in this thesis needs to be extended to reflect real industrial system. The type and sensitivity of changes that could be captured in the absorption and scattering spectra on a wider scale need to be further investigated.

**References**

1.  Abrahamsson, C., J. Johansson, S. Andersson-Engels, S. Svanberg, and S. Folestad, *Time-Resolved NIR Spectroscopy for Quantitative Analysis if Intact Pharmaceutical Tablets.* Analitical Chemistry, 2005. **77**: p. 1055-1059.

2.  Hjalmarsson, P. and S.N. Thennadil. *Spatially-resolved in-Vivo measurement system for estimating the optical properties of tissue in the wavelength range 1000-1700nm - art. no. 662805.* in *Proceedings of SPIE-OSA Biomedical Optics, Diagnostic Optical Spectroscopy in Biomedicine IV.* (2007).
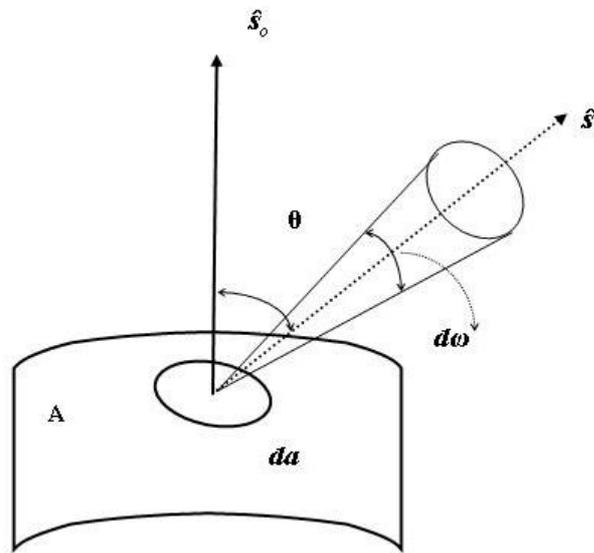
Fig. A.1. Total flux with direction ŝ passing through a small area *da* on a surface A.
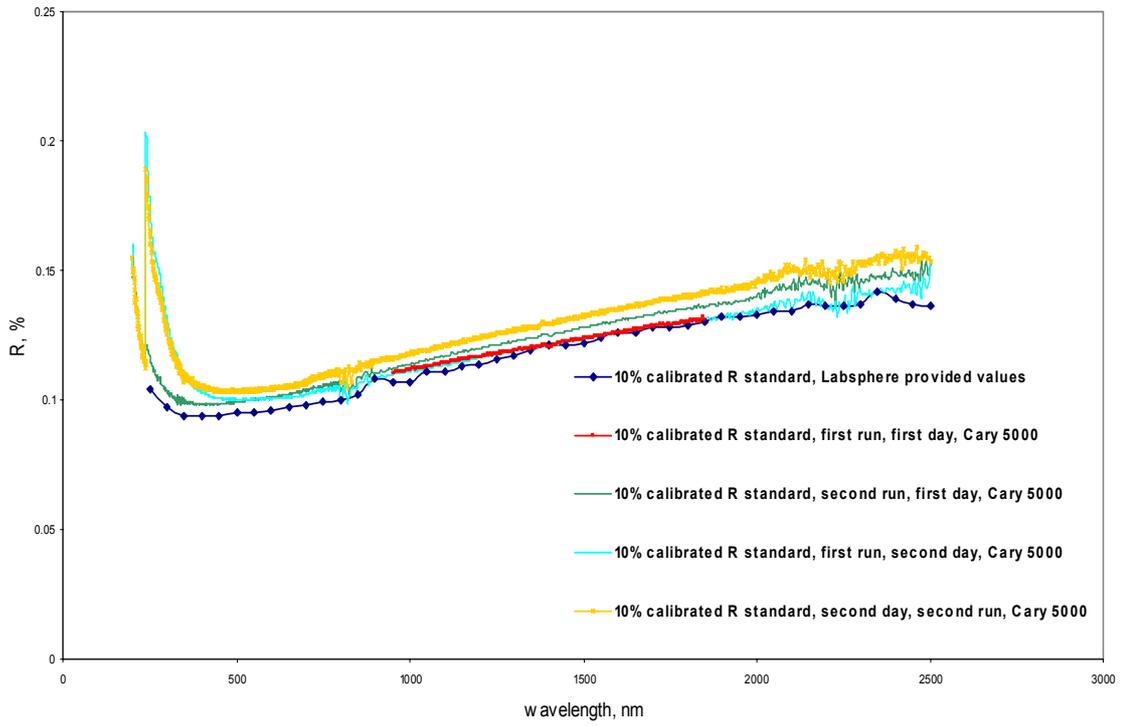
Figure A.2 Repeatability for diffuse reflectance measurements performed with Cary 5000
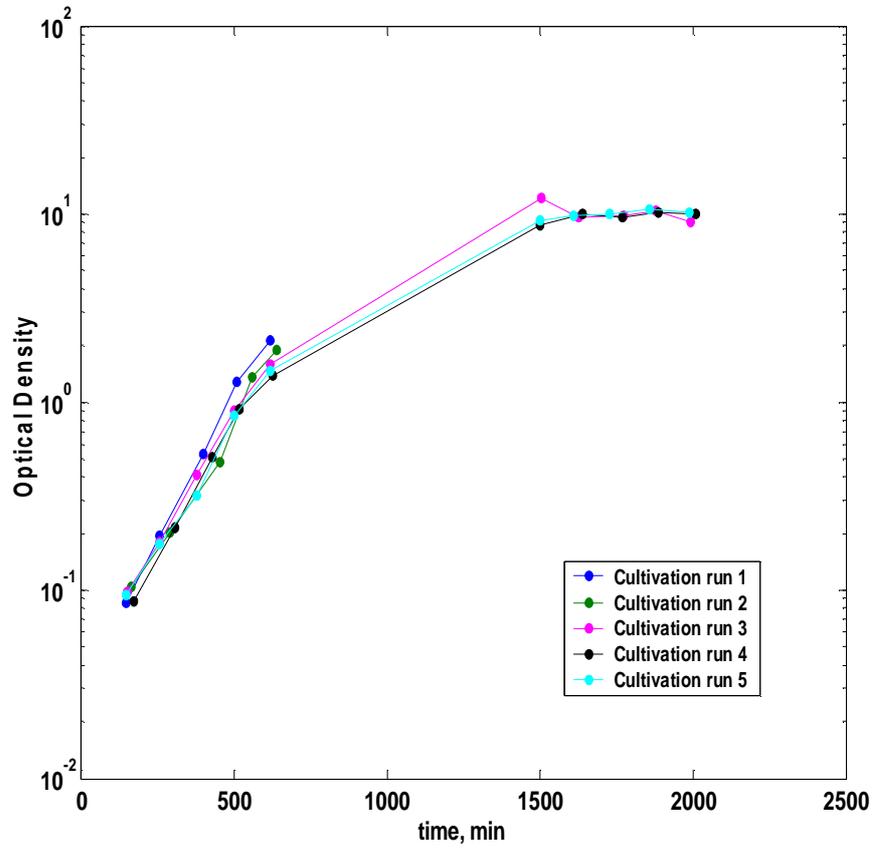
Figure B.1 Optical density
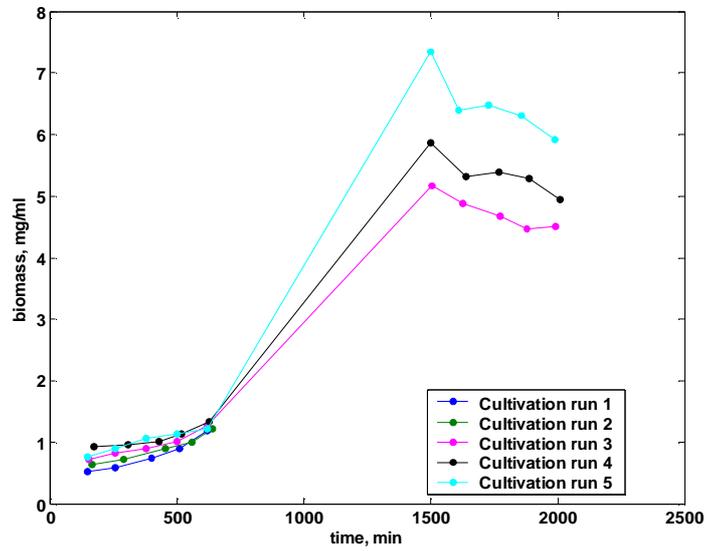
Figure C.1 Dry weight of cells profiles for cultivation runs (1) to (5)
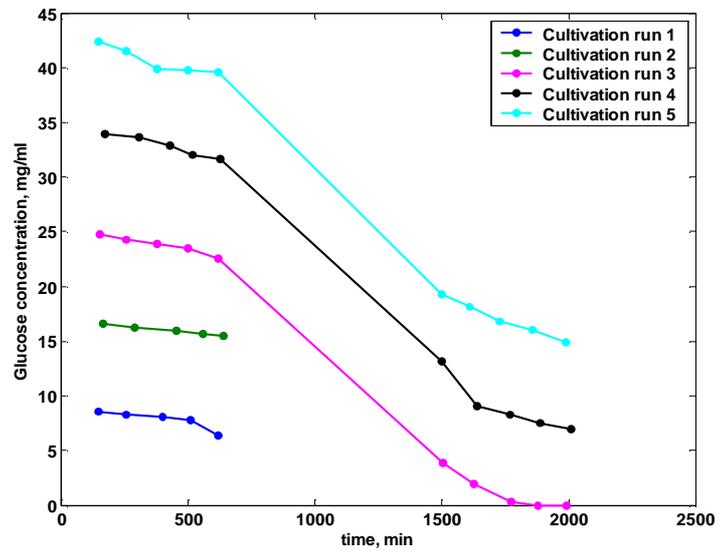


Figure C.2 Glucose concentration profiles for cultivation runs (1) to (5)
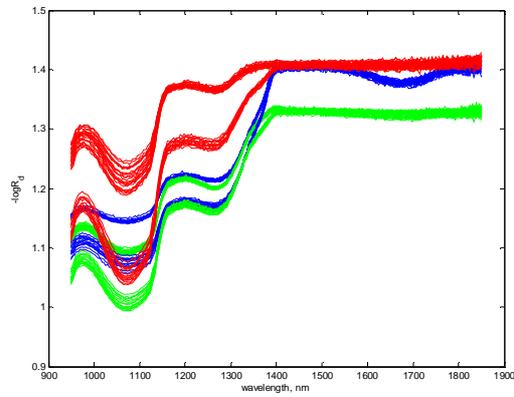
(a)
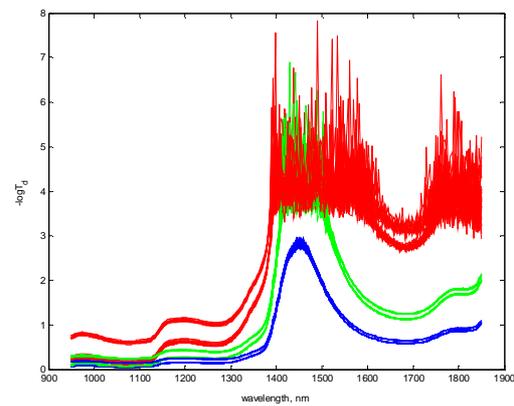


(b)
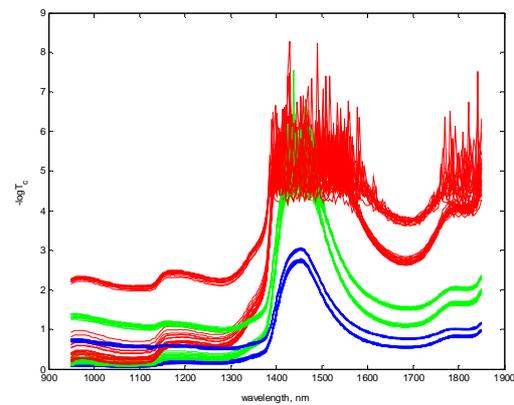


(c)

Figure D.1 Spectra of samples taken from cultivation run (1)-(5) with 2mm (blue lines), 4mm (green lines), and 10mm (red lines) cuvette pathlength. Shown are (a) diffuse reflectance spectra, (b) diffuse transmittance spectra, and for (c) collimated transmittance spectra
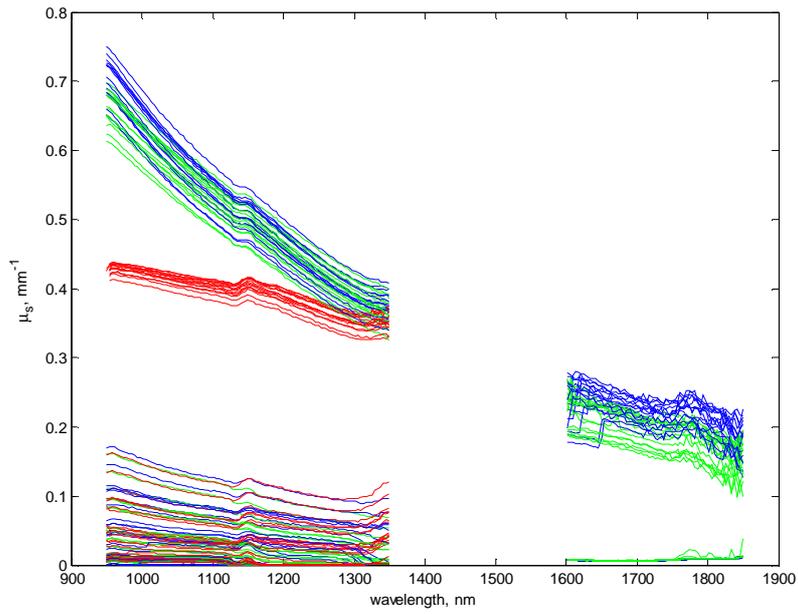
(a)
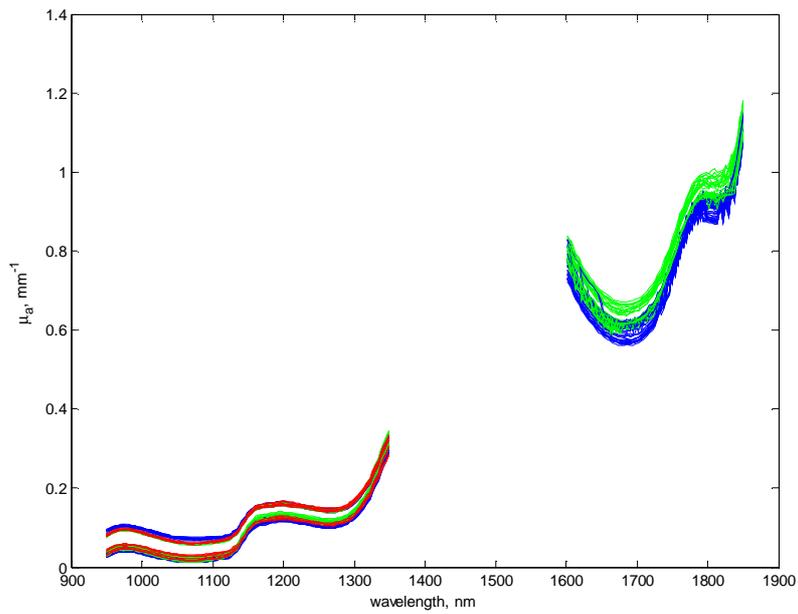


(b)

Figure D.2 Estimated optical properties and anisotropy factor from cultivation run (1)-(5) with 2mm (blue lines), 4mm (green lines), and 10mm (red lines) cuvette pathlength. Shown are (a) scattering coefficient and (b) absorption coefficient

**Data pre-processing techniques**

## E.1. First and second derivatives

First and second derivatives are the most commonly applied smoothing techniques when it comes to spectral pre-processing. By applying these two methods, the goal usually is to ensure robustness of a given calibration model if peak overlap needs to be reduced or baseline drift needs to be eliminated. The first derivative allows the additive background effect to be removed whilst the second derivative removes the baseline linear slope variations and additive effects.

## E.2. Ordinary multiplicative scatter correction method

Multiplicative scatter correction (MSC) also known as multiplicative signal correction has been developed for near-infrared spectral data[1], and the model for an individual spectrum is,

$$x_{ik} = a_i + b_i \bar{x}_k + e_{ik} \, (i = 1,...,N; k = 1,...,K) \qquad (6.1)$$

where $i$ is the sample number and $\kappa$ is the wavelength number. The constant $a_i$ represents the additive effect while $b_i$ represents the multiplicative effect for sample $i$. The mean $\bar{x}_k$

$$\bar{x}_k = \frac{1}{N} \sum_{i=1}^{N} x_{ik} \qquad (6.2)$$

is the average over samples at the $k$th wavelength. The $a_i$ and $b_i$ coefficients are unknown and must be estimated individually for each sample using all or a subset of the $k$ spectral measurements. The $e_{ik}$ is determined as an error, and in the model corresponds to all other effects in the spectrum that cannot be modelled by an additive and multiplicative constant. Once the constant $a_i$ and $b_i$ are estimated they are used in the MSC transform which subtracts $\hat{a}_i$ from $x_{ik}$ and divides the result by $\hat{b}_i$. In this equation $x_{ik}^*$ represents the corrected spectral data and $x_{ik}$ is the uncorrected spectral data.

$$x_{ik}^* = (x_{ik} - \hat{a}_i)/\hat{b}_i \qquad (6.3)$$

This transformation eliminates most of the non-chemical variations in the spectra including additive and multiplicative effects. Application of multiplicative scatter correction simplifies the calibration model and is expected to improve linearity.

### E.3. Piecewise multiplicative scatter correction method

Piecewise multiplicative scatter correction (PMSC) is a nonlinear method, which is an extension of the multiplicative scatter correction method (MSC), and is used most often to improve the predictive ability of NIR spectral data. The main aim while applying the method is to correct given spectral data for nonlinear additive and multiplicative

scatter effects. The main idea behind the method is to fit a linear regression in a local wavelength region. For continuous spectra equation 6.1 is fitted in a moving window of wavelengths or a fixed size of window can be used[2].

## E.4. Standard normal variate method

Standard normal variate method (SNV) removes from spectra multiplicative effects related to scatter and particle sizes[3]. Through this method individual spectra are centred and scaled, and the final result is similar to the one obtained with MSC[4]. If as it is in the case of MSC $x_{ik}$ is the spectral data at the $k$th wavelength for the $i$th sample, the transformed value is then given by

$$x_{ik}^* = (x_{ik} - m_i)/s_i \qquad (6.4)$$

where $m_i$ is the mean of the $k$ spectral measurements for sample $i$ and $s_i$ is the standard deviation of the $k$ measurements. The main difference between MSC and SNV is that SNV standardises each spectrum using data only from the same spectrum and does not use mean spectrum of any set.

## E.5. Extended multiplicative signal correction

Extended multiplicative signal correction (EMSC) is used to eliminate pathlength or scattering effects on a spectral data. It is en extension of MSC involving the addition of wavelength dependent terms in the formula to correct for scattering effects which are a function of wavelength. In this work two forms for pre-processing were considered:

EMSCW and EMSCL.

### E.5.1. EMSCW

The first form is presented below and has been named extended multiplicative signal correction for wavelength dependence (EMSCW)[5].

$$x_i = a_i 1 + b_i x_{m+} d_i \lambda + e_i \lambda^2 + \varepsilon_i \qquad (6.5)$$

The equation above relates the spectrum $x_i$ of the $i$th sample to the mean spectrum $x_m$ plus a quadratic equation for wavelength-dependent scattering effects.

### E.5.2. EMSCL

The second term has been introduced recently[5] and uses a different from EMSCW wavelength-dependent term referred to as EMSCL, where L stands for log term, and the expression is the following

$$x_i = a_i 1 + b_i x_{m+} d_i \log \lambda + \varepsilon_i \qquad (6.6)$$

The reason for using the log term is embedded in the Rayleigh approximation where scattering intensity from small size particles is proportional to $\lambda^{-4}$. It is expected that scattering by any particle has the form of $\alpha\lambda^{\beta}$, and the log of it leads to the log term of equation 6.6.

## E.6. Inverse signal correction

The main idea of inverse signal correction (ISC)[6] is to keep the transformed spectral values as much as possible close to the mean spectrum. ISC uses same regression relationship between measured spectral data and their mean as it is used in the case of ordinary MSC. The difference is embedded in the correlation coefficient which is taken in such a way that the mean spectrum is the dependent variable. The expression is the following,

$$(m_i)_a = \frac{\sum_k (x_{ik} - \bar{x}_{i.})(\bar{x}_{.k} - \bar{x})}{\sum_k (x_{ik} - x_{i.})^2} \tag{6.7}$$

The equation above includes, $(m_i)_a$ for multiplicative constant and $x_{ik}$ for spectral values (R or T) where $i$ is the sample index and $k$ gives the wavelength. Vector $x_i$ is the spectral response vector from sample $i$, $\bar{x}$ is the mean spectral response, and $\bar{x}_k$ is the spectral value of the mean spectrum.

## References

1. Geladi, P., D. MacDOUGALL, and H. MARTENS, *Linearization and scatter-correction for near-infrared reflectance spectra of meat.* Applied Spectroscopy, 1985. **39**(3): p. 491-500.
2. Isaksson, T. and B. Kowalski, *Piece-wise multiplicative scatter correction applied to near-infrared diffuse transmittance data from meat products.* Applied Spectroscopy, 1993. **47**(6): p. 702-709.
3. Barnes, R., M. Dhanoa, and S. Lister, *Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra.* Applied Spectroscopy, 1989. **43**(5): p. 772-777.
4. Naes, T., T. Isaksson, T. Fearn, and T. Davies, *A user-friendly guide to multivariate calibration and classification.* 2004: NIR Publications.
5. Thennadil, S.N. and E.B. Martin, *Empirical preprocessing methods and their impact on NIR calibrations: a simulation study.* Journal of Chemometrics, 2005. **19**: p. 77-89.
6. Helland, I., T. Naes, and T. Isaksson, *Related versions of the multiplicative scatter correction method for preprocessing spectroscopic data.* Chemometrics and Intelligent Laboratory Systems, 1995. **29**: p. 233-241.