Monitoring bacterial culture growth by calibration models built on bulk optical properties

Elitsa Dzhongova[1], Colin Harwood[2] and Suresh Thennadil[3]*

[1]Newcastle University, School of Chemical Engineering and Advanced Materials, Newcastle upon Tyne, NE1 7RU, United Kingdom
[2]Newcastle University, Institute for Cell and Molecular Biosciences, Newcastle upon Tyne, NE1 7RU, United Kingdom
[3]University of Strathclyde, Department of Chemical and Process Engineering, Glasgow, G1 1XJ, United Kingdom
*Corresponding author: suresh.thennadil@strath.ac.uk

Introduction

In the pharmaceutical industry, the emphasis on 'quality by design' recognises the importance of process analytical technology (PAT) for efficient and safe manufacture of pharmaceutical products. The success of the PAT initiative hinges on the development of robust measurement and calibration methodologies which can be implemented on-line for monitoring important parameters such as concentrations of nutrients, products and biomass. Such information is useful both for ensuring the process is proceeding normally by taking timely corrective actions as well as being part of the input for purposes such as end-point prediction. Near infrared (NIR) optical sensor technology has the potential to deliver a platform for rapid physical and chemical characterisation of process streams. The major barrier has been the lack of calibration methodologies that can deliver robust models which can be constructed using small datasets and can be used across different sites and process scales. Multiple scattering of light by cells poses a significant challenge in the development of NIR-based calibration models to reliably extract physical and chemical information contained in the spectra collected during a bacterial growth cycle. The extent of information that can be obtained from NIR spectra could, in principle, be vastly improved if the scattering and absorption effects can be effectively decoupled. This work focuses on a methodology which uses the radiative transfer theory to decouple the absorption and scattering effects through the extraction of bulk absorption and scattering coefficients.[1,2] The feasibility of this approach to deliver better performing calibration models was studied by applying the method to a simple system consisting of *Bacillus subtilis* growing in an aqueous solution.

Materials and Methods

Bacterial growth experiments

A *Bacillus subtilis* culture growing in an aqueous solution was considered in this study. Five growth cycle experiments were carried out using *Bacillus subtilis*, obtained from the Institut Pasteur (Paris, France). Data were collected from the growing culture of the strain which was cultivated in 100 mL Spizizen's minimal medium containing trace element solution in a 250 mL Erlenmeyer flask.[3] All the bacterial growth cycle experiments were carried out with the medium at a pH of 7.0 ± 0.5 and the temperature controlled at 37°C ± 0.5. The cultivation was carried out in an orbital incubator with agitation rate set at 220 rpm. Five cultivation runs were carried out with each of the runs having different initial glucose concentrations of 8, 16, 25, 34 and 43 mg mL^{-1}. 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 medium. Also, by changing the initial glucose concentrations, the time profile of the cultivation was significantly changed in the sense that cultures had different time lengths during the growth phase. This is important since if all the cultivation runs have the same batch length, and thus the same time profile, there is a danger that PLS models will spuriously fit the profile of biomass change to non-relevant time-related factors rather than the actual changes that could be attributed to changes in biomass or glucose. Therefore, if the time profiles of biomass were to be the same for all the runs, the model could falsely indicate very good model performance even if it did not contain information relevant to changes in biomass. Figure 1 shows the profiles for the five cultivation runs; it may be seen that the cultivation run time was extended by increasing the initial glucose concentrations. The growth period varied such that it was about 10 h for cultivation runs 1 and 2, and about 30 h for cultivation runs 3 to 5. Thus, overall profiles were relatively the same but the difference in cultivation time ensured that, to some degree, process samples were taken at different slopes in the time profile.

Reference paper as:
Dzhongova, E,. Harwood, C. and Thennadil, S. (2012). Monitoring bacterial culture growth by calibration models built on bulk optical properties, in: Proceedings of the 15th International Conference on Near Infrared Spectroscopy, Edited by M. Manley, C.M. McGoverin, D.B. Thomas and G. Downey, Cape Town, South Africa, pp. 107-111.

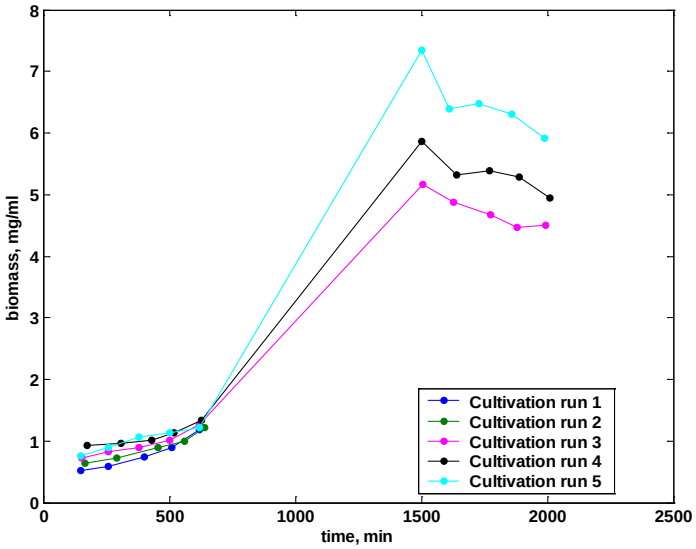

Figure 1. Time profile of the biomass with measurements taken every two hours except during night time when no samples were taken.

Measurements

For each of the five cultures, data were collected with samples taken at approximately 2 h intervals during the growth phase. Due to the different culture run times and variation in glucose, 5 samples were taken from runs 1 and 2 and 10 samples from runs 3 to 5. For runs 3 to 5 no samples were collected during the night. As a result, there is a gap in the concentration profile which will be seen in the results. For each sample, three spectroscopic measurements were made, namely total diffuse reflectance (R_d), total diffuse transmittance (T_d) and collimated transmittance (T_c). Each of these measurements was made using cuvettes with different path lengths (2, 4 and 10 mm) in order to collect data with different sample thicknesses. A UV-Vis-NIR spectrophotometer (Cary 5000) equipped with an external integrating sphere was used for the measurements and spectra were collected over the wavelength range 950 to1850nm at 4 nm intervals. The average integration time was set to 0.4 s with an average signal bandwidth of about 15 nm. The biomass of all samples was measured gravimetrically. The protocols followed in this study for spectroscopic data collection and reference method for biomass measurement were the same as in earlier work.[4]

Methodology

Determination of bulk optical properties

Bulk optical properties are determined by inverting measurements using the radiative transfer equation (RTE) which is given by,[5]

$$\frac{dI(\lambda,\vec{r},\hat{s})}{ds} = -\mu_t(\lambda)I(\lambda,\vec{r},\hat{s}) + \frac{\mu_t(\lambda)I(\lambda,\vec{r},\hat{s})}{4\pi}\int_{4\pi} p(\hat{s},\hat{s}')I(\lambda,\vec{r},\hat{s}')d\omega' \quad (1)$$

where $I(\lambda,\vec{r},\hat{s})$ is the specific intensity of light of wavelength λ at point $\vec{r}$ with radiation incident along direction $\hat{s}$; $\mu_t(\lambda)$ is the bulk extinction coefficient (mm^{-1}) and is given by $\mu_t(\lambda) = \mu_s(\lambda) + \mu_a(\lambda)$ where $\mu_s(\lambda)$ is the bulk scattering coefficient, $\mu_a(\lambda)$ is the bulk absorption coefficient and $p(\hat{s},\hat{s}')$ is the phase function which is a measure of the angular distribution of scattered light. This is usually approximated as a function of the anisotropy factor $g(\lambda)$ using the Henyey-Greenstein function.[6] To extract the bulk optical properties $\mu_s(\lambda)$, $\mu_a(\lambda)$ and $g(\lambda)$ through the inversion of (1) we need a minimum of 3 measurements at each wavelength. In this study, an integrating sphere set-up was used which provided the three measurements $R_d(\lambda)$, $T_d(\lambda)$ and $T_c(\lambda)$ over the wavelength range considered in this study. For each sample, these measurements were inverted using the adding-doubling method in an iterative manner. The iteration was started with guessed values of the bulk optical properties and using the adding-doubling method to calculate $\hat{R}_d(\lambda)$ and $\hat{T}_d(\lambda)$. $\hat{T}_c(\lambda)$ was calculated using Beer's law. The specular reflectances at the air-glass and glass-sample boundaries were included in the calculations through the use of Fresnel equations and

Reference paper as:
Dzhongova, E,. Harwood, C. and Thennadil, S. (2012). Monitoring bacterial culture growth by calibration models built on bulk optical properties, in: Proceedings of the 15th International Conference on Near Infrared Spectroscopy, Edited by M. Manley, C.M. McGoverin, D.B. Thomas and G. Downey, Cape Town, South Africa, pp. 107-111.

"adding" the sample and glass layers to obtain total diffuse reflectance and transmittance from the entire glass-sample-glass entity. Calculated values were compared with measured values, the guessed values were then updated and the process repeated until convergence was achieved. The iterations were carried out to minimise the objective function:

$$\xi = abs(R_d - \hat{R}_d) + abs(T_d + \hat{T}_d) + abs(T_c - \hat{T}_c) \quad (2)$$

Tolerance for convergence was set to $\xi \leq 1.0e-7$; minimisation was carried out using the function "fmincon" in the Matlab® optimisation toolbox. Further information regarding the details of the inversion and the adding-doubling method can be found in Dzhongova *et al*.[4] and Prahl.[6]

Multivariate calibration models

Models for predicting biomass were built using partial least squares (PLS) regression. A number of pre-processing methods such as first and second derivatives, multiplicative scatter correction (MSC), inverse scatter correction (ISC), standard normal variate (SNV) and extended multiplicative signal correction (EMSC) were tried on the data. The spectra were mean-centred. PLS calibration models were built using the measurements directly i.e. using $\boldsymbol{T}_c$, $\boldsymbol{T}_d$ and $\boldsymbol{R}_d$ to build the models for predicting biomass concentrations. In all cases, these measurements were converted to absorbance units before using them for model building. Models were also built using the extracted bulk optical properties, namely $\boldsymbol{\mu}_a$, $\boldsymbol{\mu}_s$ and $\boldsymbol{\mu}_t$ (the bold symbols are used to represent a vector of values at multiple wavelengths). The performances of these models were compared to those obtained using the measurements directly in order to examine if the extraction of bulk optical properties could provide better model performance. The reason for using $\boldsymbol{\mu}_t$ was to investigate if using a quantity, which is the sum of the bulk absorption and scattering coefficient, will lead to better performing models than using those quantities individually. In principle the combination of the two quantities can enhance the information content available for the predictive model.

The model parameters and the model performance were evaluated using leave-one-out (one cultivation at a time) cross-validation. Essentially this consisted of leaving out all the measurements from one of the five cultivation runs. The PLS model built was thus applied to the "left-out" cultivation and the error in the prediction for that cultivation calculated. This process was continued until all the cultivations were left out once. The root mean square error of cross-validation (RMSECV) was then calculated in the usual way. The models were built for measurements from all the three sample thicknesses, namely 2, 4 and 10 mm, in order to examine the effect of sample thickness on the performance of calibration models.

Results and Discussion

Table 1 reports the performance of PLS models for biomass prediction. The table shows only the best performing models for each of the six types of "measurements". While all the pre-processing techniques mentioned earlier were tried out, only results from techniques providing the best results are included in the table. Sample thickness indicated in the table gives the cuvette path length, the measurements from which gave the best performing model.

Table 1. Best cross-validation results for the 6 types of measurements used for building models **.**

Input spectral data	Sample thickness (mm)	Pre-processing	Number of LVs	RMSECV (mg.mL^{-1})
$\boldsymbol{T}_c$	4	Ordinary MSC	7	0.22
$\boldsymbol{T}_d$	10	none	3	0.51
$\boldsymbol{R}_d$	10	ISC	2	0.59
$\boldsymbol{\mu}_a$	10	Ordinary MSC	5	0.73
$\boldsymbol{\mu}_s$	2	none	3	0.36
$\boldsymbol{\mu}_t$	10	Ordinary MSC	7	0.22

From Table 1 it can be seen that the best models using different types of measurements were obtained with different sample thicknesses though in most cases the 10 mm sample thickness appears to the best

Reference paper as:
Dzhongova, E,. Harwood, C. and Thennadil, S. (2012). Monitoring bacterial culture growth by calibration models built on bulk optical properties, in: Proceedings of the 15th International Conference on Near Infrared Spectroscopy, Edited by M. Manley, C.M. McGoverin, D.B. Thomas and G. Downey, Cape Town, South Africa, pp. 107-111.

choice for the measurements. The best model using T_C is obtained using a 4 mm sample thickness. Increasing the sample thickness leads to lower signal reaching the detector when a collimated transmittance measurement is made. This is offset by the increased number of cells "seen" by the transmitted light. Thus there is a trade-off when sample thickness is increased and, for the samples considered here, 4 mm sample thickness appears to be optimal for providing the best performing model.

Diffuse reflectance and diffuse transmittance measurements gave the best models for the largest sample thickness. Due to this we would expect the extracted optical properties, which rely on these measurements, to give the best performance for this sample thickness. While this is certainly the case for μ_a and μ_t, the bulk scattering coefficient μ_s actually shows the opposite trend with models built using μ_s leading to lower RMSECV with increasing sample thickness.

It can be seen that the best model is obtained by combining the information present in both the bulk scattering and bulk absorption coefficients, i.e. by using μ_t obtained from 10 mm sample thickness as well as the collimated transmittance measurements with a smaller sample thickness of 4 mm. From Beer's law,[5] the collimated transmission through a sample of thickness l is given by:

$$-\ln T_C(\lambda) = \mu_t(\lambda) l \qquad (3)$$

From (3) it can be seen that if μ_t is accurately extracted it would provide the same information content as T_C. While collimated transmittance can be measured with very small errors, such measurements are only suitable for characterising suspensions in which the extent of scattering is small, as would be the case for suspensions which have low particle concentrations. As particle concentrations increase, in order to measure T_C accurately, progressively smaller sample thicknesses would have to be used. At moderate to high particle concentrations, which are normal levels encountered in industrial slurries, the sample thickness required to make accurate T_C measurements would be practically infeasible to achieve. However, by using a suitable set of measurement configurations such as spatially resolved fibre-optic diffuse reflectance measurements[7] it may be possible to extract μ_t for samples with high particle concentrations. Provided we can accurately extract μ_t, it will be possible to obtain model performance as good as that which would be obtained using collimated transmittance. Figures 2 and 3 show the cross-validation curves for models built using measurements from 4 mm and 10 mm sample thicknesses respectively to illustrate model performances of the various measurements used to build the models and the effect of sample thickness on model performance.

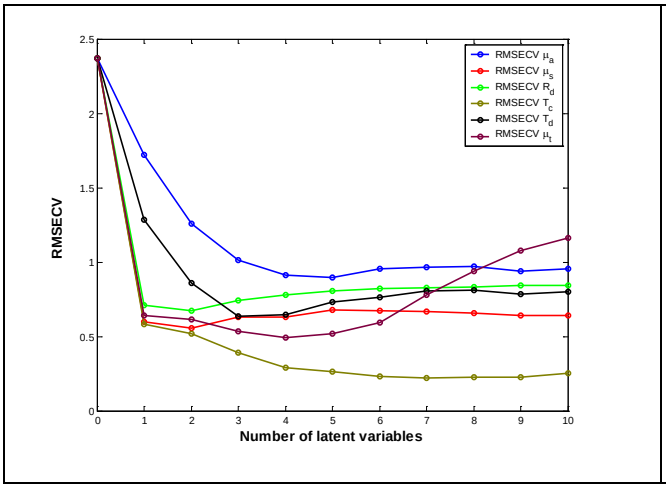

Figure 2. Cross-validation curve for models built, using measurements from samples of 4 mm thickness.

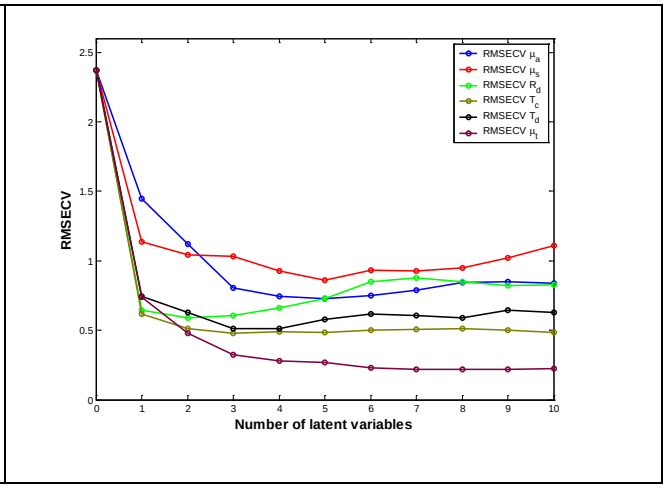

Figure 3. Cross-validation curve for models built, using measurements from samples of 10 mm sample thickness.

An interesting aspect of using the bulk optical properties is the fact that this property should be independent of sample thickness; Figures 4 and 5 illustrate this feature. In Figure 4, diffuse reflectance measurements collected at 1050 nm on a culture using 3 different sample thicknesses (2 mm, 4 mm and 10 mm) are shown in absorbance units. We see that there is a significant difference in the level of absorbance based on sample thickness. This absorbance does not "line up" with sample thickness, instead we see a drop in absorbance when the sample thickness is increased from 2 to 4 mm and then an appreciable increase in

Reference paper as:
Dzhongova, E,. Harwood, C. and Thennadil, S. (2012). Monitoring bacterial culture growth by calibration models built on bulk optical properties, in: Proceedings of the 15th International Conference on Near Infrared Spectroscopy, Edited by M. Manley, C.M. McGoverin, D.B. Thomas and G. Downey, Cape Town, South Africa, pp. 107-111.

absorbance when a 10 mm sample thickness is used. This absorbance level is higher than the 2 mm or 4 mm thicknesses. Further, a closer look at the relationship between absorbance measured in diffuse reflectance mode and biomass shows different functional relationships with biomass for different sample thickness. This was clear when the measurements were plotted individually (not shown here) and can be discerned even on

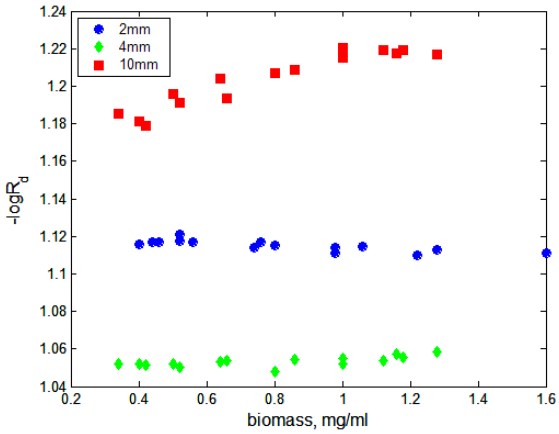

Figure 4. Cross-validation curve for models built using measurements from samples of 4 mm thickness.

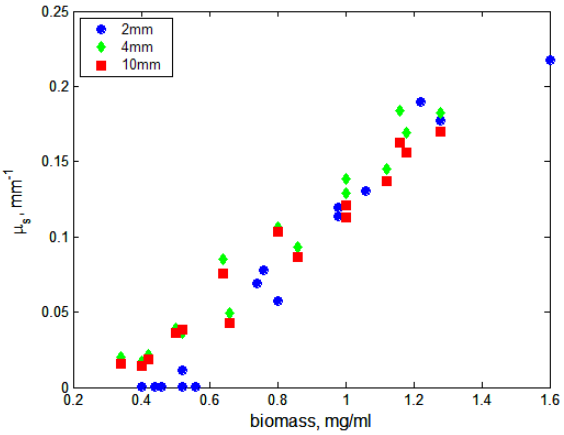

Figure 5. Cross-validation curve for models built using measurements from samples of 10 mm sample thickness.

the scale used in Figure 4. For the corresponding points in Figure 4, when the bulk scattering coefficient was plotted against biomass, we see (Figure 5) that the values of μ_s are the same within experimental error regardless of the sample thickness used. Furthermore, the bulk scattering coefficient varied linearly with biomass for all three sample thicknesses. A similar result was found with the bulk absorption coefficient (not shown here). This result has implications regarding calibration transfer. Since μ_a and μ_s are properties of the sample itself and are basically path length normalised, changes in path length due to changes in instrument or cuvette will not affect the model as long as these properties can be accurately extracted. The challenge lies in developing instrument configurations which will lead to accurate estimation of the bulk optical parameters.

Conclusion

This work indicates that calibration models based on optical properties have the potential to provide robust and accurate calibration models. The fact that optical properties are invariant to sample thickness has potential advantages regarding calibration transfer. For the method to be applied in practical situations, the challenge lies in developing methods that will provide accurate estimation of the bulk optical parameters and developing measurement configurations such as fibre optic spatially-resolved diffuse reflectance spectrometers which can be used for collecting accurate measurements at high particle concentrations

Acknowledgements

This work was funded by Marie Curie FP6 (INTROSPECT) and by EPSRC grants GR/S50441/01 and GR/S50458/01.

References

1. R. Steponavicius and S.N. Thennadil, *Anal. Chem.* **81**, 7713 (2009).
2. R. Steponavicius and S.N. Thennadil, *Anal. Chem.* **83**, 1931 (2011).
3. C.R. Harwood and S.M. Cutting, *Molecular biology methods for Bacillus*, John Wiley & Sons Ltd., UK (1990).
4. E. Dzhongova, C.R. Harwood and S.N. Thennadil, *Appl. Spectrosc.* **63**, 25 (2009).
5. A. Ishimaru, *Wave propagation and scattering in random media*. IEEE Press (1997).
6. S.A. Prahl, "The Adding-Doubling Method", in *Optical thermal response of laser irradiated tissue*, Ed by A.J. Welch and M.J.C. van Gemert, Plenum Press, New York, pp. 101–129 (1995).
7. P. Hjalmarsson and S.N. Thennadil, "Spatially-resolved in-vivo measurement system for estimating the optical properties of tissue in the wavelength range 1000-1700nm", in *Proceedings of SPIE, Diagnostic Optical Spectroscopy in Biomedicine IV*, Ed by D. Schweitzer and M. Fitzmaurice, 6628, paper 662805 (2007).

Reference paper as:
Dzhongova, E,. Harwood, C. and Thennadil, S. (2012). Monitoring bacterial culture growth by calibration models built on bulk optical properties, in: Proceedings of the 15th International Conference on Near Infrared Spectroscopy, Edited by M. Manley, C.M. McGoverin, D.B. Thomas and G. Downey, Cape Town, South Africa, pp. 107-111.