

How accurate is Beer's Law in the analysis of NIR Data?

R. S. Anderssen¹

F. R. de Hoog²

I. J. Wesley³

A. B. Zwart⁴

(Received 12 February 2015; revised 11 November 2015)

Abstract

To construct a predictor of the proportional presence of some key component in a material, we assume that the near infrared (NIR) responses of the various component proportions are given by Beer's law. Using a set of NIR spectra of milk powder spiked with different amounts of casein, this assumption is tested by comparing different properties of this set of spectra with the equivalent properties of simulated spectra obtained by combining the spectrum of unspiked milk powder with the spectrum of casein, in the same proportions. The latter set of spectra corresponds to the situation where Beer's law holds exactly with no cross-interaction between the two linearly independent components, and, thereby, described by a rank-two row matrix. The

<http://journal.austms.org.au/ojs/index.php/ANZIAMJ/article/view/9291>

gives this article, © Austral. Mathematical Soc. 2016. Published January 11, 2016, as part of the Proceedings of the 17th Biennial Computational Techniques and Applications Conference. ISSN 1446-8735. (Print two pages per sheet of paper.) Copies of this article must not be made otherwise available on the internet; instead link directly to this URL for this article.

former corresponds to what happens in practice with the possibility of cross-NIR-interaction occurring between the components. The degree to which Beer's law is likely to fail is examined. It is shown how spiking allows the identification of the casein wavelengths that are strongly independent of the NIR spectral responses of the other components in the milk powder. However, there are slight deviations from Beer's law and these are, in part, explained by diffraction effects.

Contents

1 Introduction	C229
2 Beer's Law	C233
3 The calibration	C238
4 The single wavelength analysis	C239
5 Conclusions	C241
References	C244

1 Introduction

In NIR spectroscopy corrected for scattering effects, a key assumption is that the individual spectral responses of the various components in the material combine linearly in accordance with Beer's law. Essentially, Beer's law says that the NIR spectrum of a mixture of q components is a weighted sum of the spectra of the components, with the weights being the proportions of the components in the mixture. That is, if the spectra of the q components are given by the vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_q$, say, then the spectrum of a mixture, with

Table 1: The casein spiking fractions.

Sample Number k	1	2	3	4	5	6	7	8
Casein fraction	0	.0503	.1055	.2137	.4005	.5995	.6612	1.

components in the proportions given by the vector $\mathbf{p} = (p_1, p_2, \dots, p_q)$ with $p_i \geq 0$, is

$$\mathbf{v} = \sum_{i=1}^q p_i \mathbf{v}_i, \quad \text{with} \quad \sum_{i=1}^q p_i = 1. \quad (1)$$

The current investigation studies the spectra of milk powder spiked with different proportions of casein, with casein the dominant protein found in whole milk powder. In whole milk powder, the proportion of casein is about 30% and makes up about 80% of the protein-content. In this way, an analysis of Beer's law is reduced to analysing the spectra of the $q = 2$ component systems of milk powder and casein. The motivation is the need to check, for different materials, the extent to which Beer's law holds [6]. As explained by Dahm [6], the effect of the diffraction scattering is not purely linear and consequently Beer's law will only hold approximately for the second derivative spectra. Therefore, using samples spiked with the components of interest (casein in milk powder in the present situation) represents a natural way to check the extent to which Beer's law holds for combinations of different materials.

Specifically, in this article, the spectra analyzed are of spiked samples with the proportions of milk powder and casein given in Table 1.

Scattering has a major effect on the measurement of an NIR spectrum and is often modelled as a linear trend that is removed from the spectral data by taking second or higher differences [1, 2, 4, 5]. The spectra is often then smoothed as an aid to visualisation, which is also done here. For $n = 8$ samples and m the number of wavelengths at which the spectra are recorded, let the vector $\mathbf{s}_k \in \mathbb{R}^m$ ($k = 1, 2, \dots, n$) denote the second derivative spectrum of the k th sample. The associated smoothing aids the visualization of differences

between the spectra. However, such smoothing has minimal quantitative effects on the results and conclusions.

Let

$$\mathbf{S}^T = (\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n), \quad (2)$$

with the rows of \mathbf{S} corresponding to the spectra of the two-component mixtures. The smoothed second derivatives of the eight spectra are plotted in Figure 1. Since each spectrum \mathbf{s}_k , with $k = 1, 2, \dots, n$, is a mixture of just two components (namely, milk powder and casein), equation (1) is used to construct an approximation to the spectrum \mathbf{s}_k . Let $\mathbf{f} \in \mathbb{R}^n$ denote the vector of the entries which are the fractions of casein added (Table 1) and let $\mathbf{e} = (1, 1, \dots, 1)^T \in \mathbb{R}^n$. Then the theoretical NIR spectra of the two-component mixtures, satisfying Beer's law, are

$$\widehat{\mathbf{s}}_k = s_1(1 - f_k) + s_n f_k, \quad s_1 = \text{pure milk powder}, \quad s_n = \text{pure casein}. \quad (3)$$

It follows that the Beer's law counterpart of \mathbf{S} is given by

$$\widehat{\mathbf{S}}^T = (\widehat{\mathbf{s}}_1, \widehat{\mathbf{s}}_2, \dots, \widehat{\mathbf{s}}_n) = s_1(\mathbf{e} - \mathbf{f})^T + s_n \mathbf{f}^T, \quad (4)$$

on taking account of $\widehat{\mathbf{s}}_1 = s_1$ and $\widehat{\mathbf{s}}_n = s_n$. The rows of $\widehat{\mathbf{S}}$ are plotted in Figure 2.

Although Figure 1 and Figure 2 look qualitatively similar, this is in part because the differences in the row spectra are small compared to their amplitudes, which makes a visual comparison difficult.

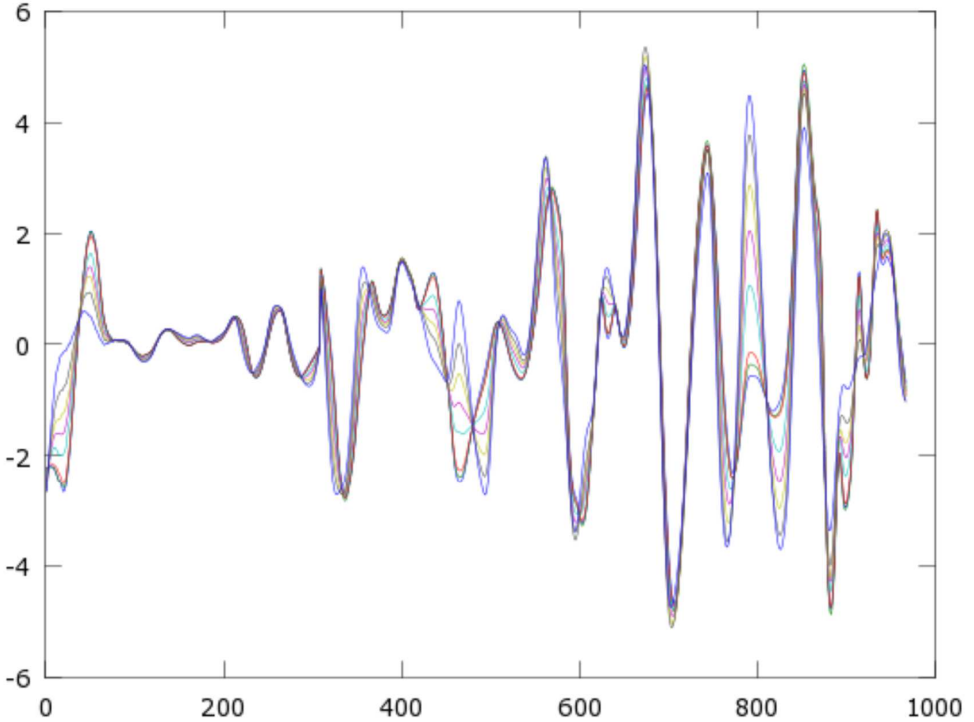
In the analysis of NIR data, it is standard practice to mean-centre the columns of \mathbf{S} [7]. We therefore replace \mathbf{S} and $\widehat{\mathbf{S}}$ with their corresponding mean-centred counterparts

$$\mathbf{C} = \left(\mathbf{I} - \frac{1}{n} \mathbf{e} \mathbf{e}^T\right) \mathbf{S}, \quad (5)$$

and

$$\begin{aligned} \widehat{\mathbf{C}} &= \left(\mathbf{I} - \frac{1}{n} \mathbf{e} \mathbf{e}^T\right) \widehat{\mathbf{S}} \\ &= \left(\mathbf{I} - \frac{1}{n} \mathbf{e} \mathbf{e}^T\right) [(\mathbf{e} - \mathbf{f}) s_1^T + \mathbf{f} s_n^T] \\ &= \left(\mathbf{I} - \frac{1}{n} \mathbf{e} \mathbf{e}^T\right) \mathbf{f} (s_n - s_1)^T. \end{aligned} \quad (6)$$

Figure 1: A plot of the second derivative of the measured spectra, for different wavelengths $[600 + 2 \times (\text{x-axis value})]$ nm, for the eight casein spiked milk powder samples.



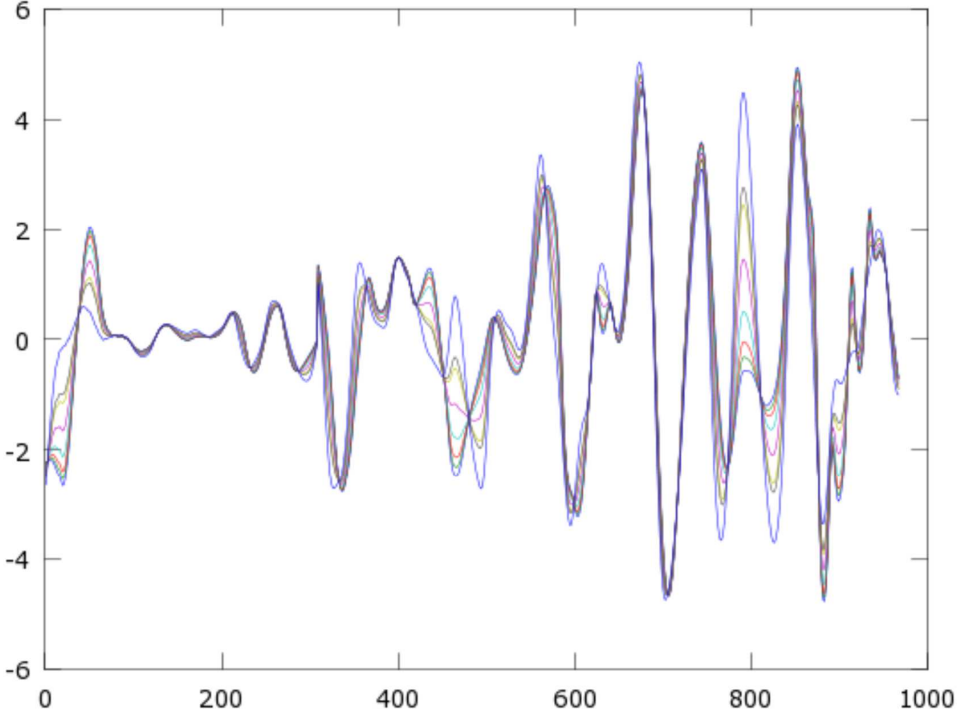
The column vectors of \mathbf{C} and $\widehat{\mathbf{C}}$ are denoted by \mathbf{c}_i and $\widehat{\mathbf{c}}_i$, respectively, with $i = 1, 2, \dots, m$.

Equation (6) is a direct consequence of the mean-centering relationship

$$\left(\mathbf{I} - \frac{1}{n}\mathbf{e}\mathbf{e}^T\right)\mathbf{e} = \mathbf{0}.$$

The benefit of performing the mean-centering is that $\widehat{\mathbf{S}}$ is replaced with the rank one matrix $\widehat{\mathbf{C}}$, thus simplifying the subsequent analysis.

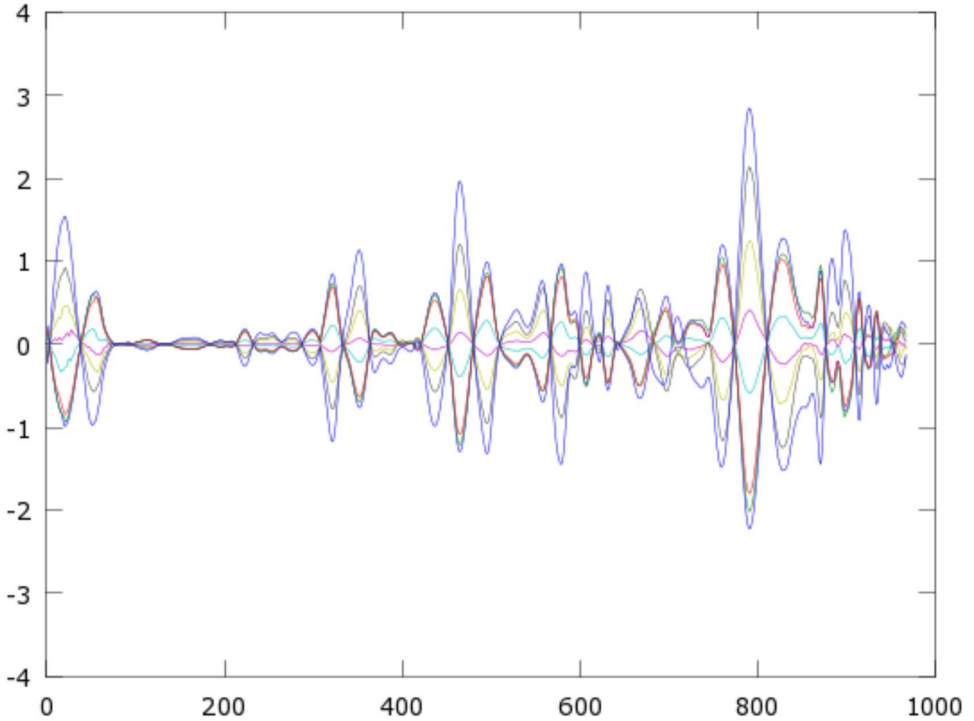
Figure 2: A plot of the second derivative Beer's law spectra for the eight casein spiked milk powder samples, as a function of wavelength $[600 + 2 \times (\text{x-axis value})]$ nm, generated using equation (3).



2 Beer's Law

Plots of the rows of \mathbf{C} and $\hat{\mathbf{C}}$ are given in Figures 3 and 4, respectively. The similarities and differences between real and simulated spectra are now much more apparent. For example, a comparison of the structure of peaks and troughs in Figure 3 (Figure 4) with their counterparts in Figure 1 (Figure 2) shows clearly that differences in spectra have become more pronounced by being more widely separated about the zero line of the mean-centering.

Figure 3: Plots of the rows of C : mean-centered second derivatives of the measured spectra for the eight spiked samples, as a function of wavelength $[600 + 2 \times (\text{x-axis value})]$ nm.



It follows from (6) that, for a two-component mixture, \hat{C} is a rank one approximation of C . This rank one approximation is a basis for assessing the extent to which Beer's law holds for the spectra S of the spiked data. For example, it implies that, if Beer's law is a good approximation for the structure in the observed spectra S , then the matrix C will be similar to \hat{C} ; a plot of the values of the matrix elements of C against the corresponding values of \hat{C} will approximate a straight line, as illustrated and confirmed in Figure 5.

Figure 4: Plots of the rows of $\widehat{\mathbf{C}}$: mean-centered second derivatives of the Beer's law spectra for the eight spiked samples, as a function of wavelength $[600 + 2 \times (\text{x-axis value})]$ nm.

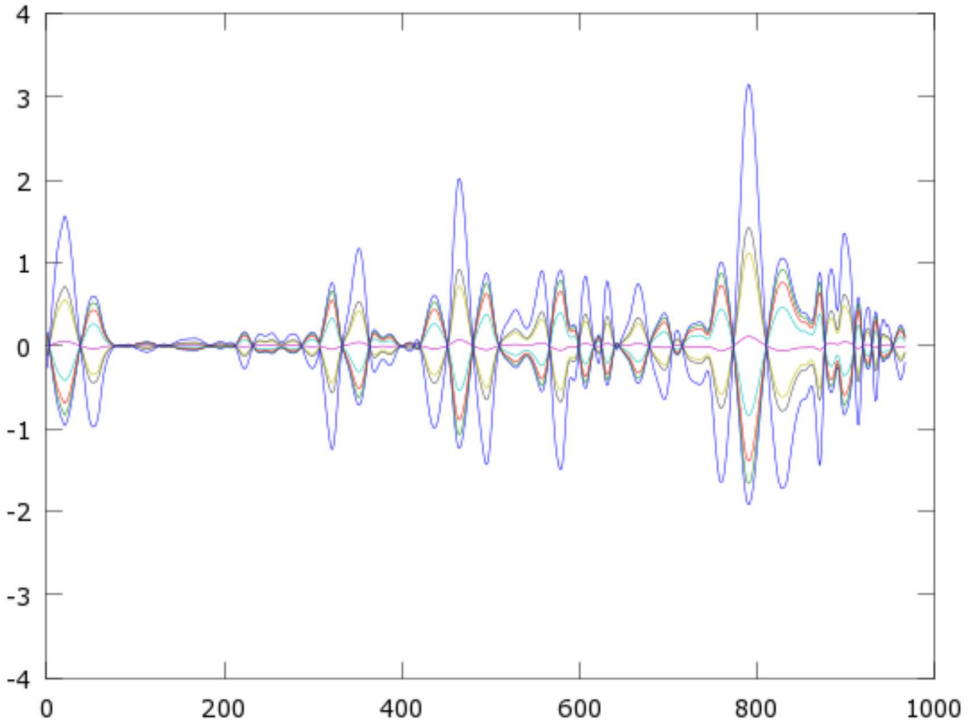
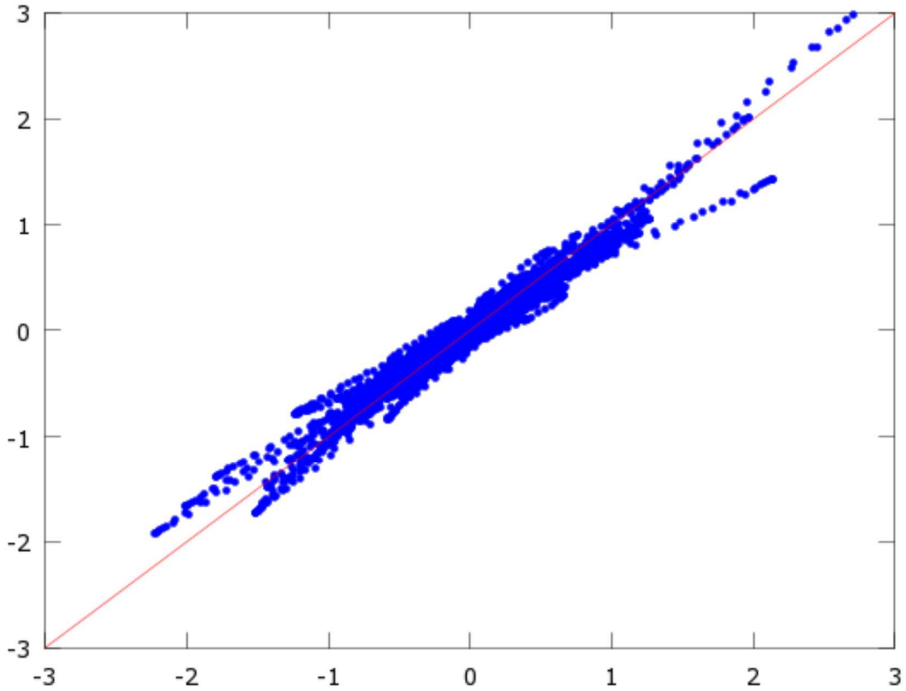


Figure 5, because of the clustering of the points about the line of equality (high correlation), yields qualitative support for the approximate validity of Beer's law for the casein spiked milk powder. Further insight is obtained by considering the relationship between \mathbf{C} and $\widehat{\mathbf{C}}$ from a number of independent perspectives.

Figure 6 plots the values of $\mathbf{c}_i^T \mathbf{c}_i$, $i = 1, 2, \dots, m$, as a function of wavelength. The magnitudes of $\mathbf{c}_i^T \mathbf{c}_i$ are a measure of the NIR molecular response to the spiking of the milk powder with the casein, with the peaks identifying

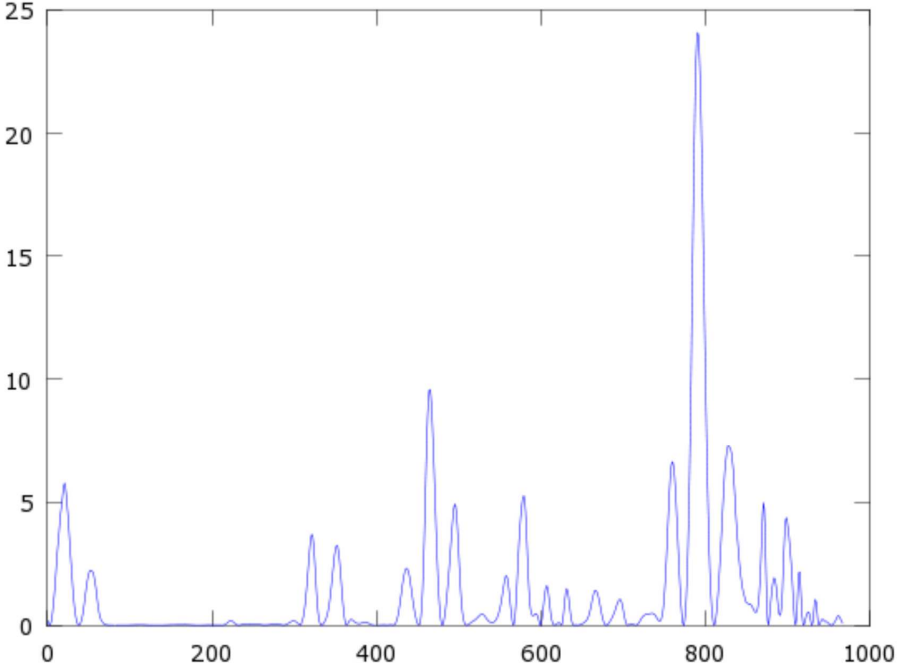
Figure 5: The scatter plot of the matrix elements of C plotted against the corresponding elements of \hat{C} .



the wavelengths of the molecular components in the casein most sensitive to the changing levels of the spiking. Identifying the strong peaks is an alternative way of identifying the wavelengths of the molecular components in the casein which are most sensitive to an NIR stimulus, which is consistent with the information recorded in an NIR spectrum of some given material, such as casein in milk powder. As is clear in Figures 1 and 2, the spectra will change only slightly at wavelengths not sensitive to the presence of the casein, whereas there will be proportional changes at the wavelengths which are strongly sensitive to the presence of the casein.

The identification of sensitive wavelengths represents indirect measurements

Figure 6: The plot of $\mathbf{c}_i^T \mathbf{c}_i$, $i = 1, 2, \dots, m$, as a function of wavelength $[600 + 2 \times (\text{x-axis value})]$ nm.



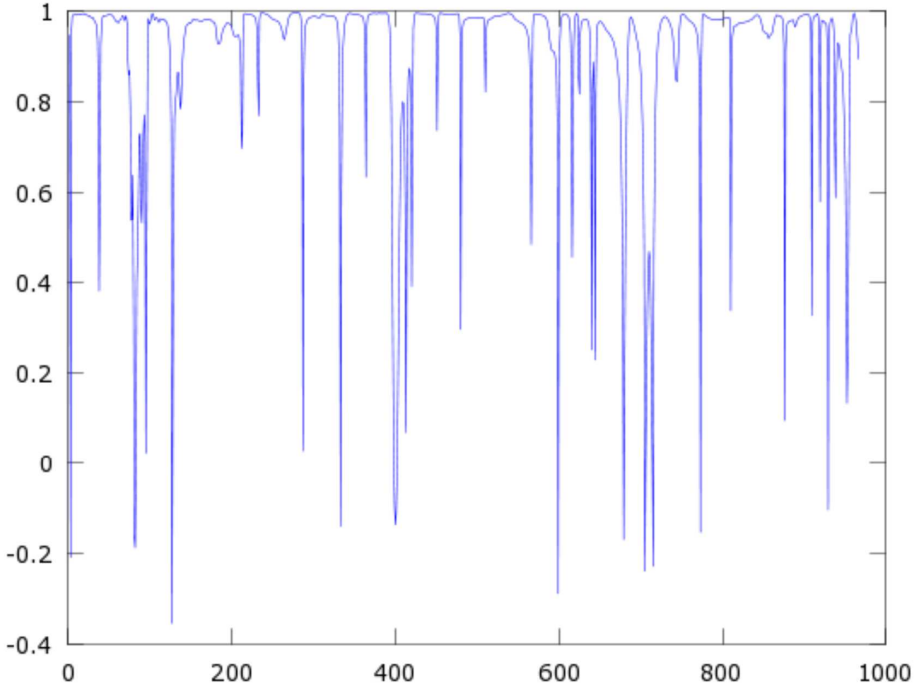
of the molecular components of a material most sensitive to an NIR stimulus. In addition, the peaks of $\mathbf{c}_i^T \mathbf{c}_i$, $i = 1, 2, \dots, m$, represent further support for the validity of Beer's law as the peaks occur where there is a strong correlation between spectra of Figure 3 and Figure 4.

An alternative way of highlighting the relationship between \mathbf{C} and $\hat{\mathbf{C}}$ is given in Figure 7, where we plot

$$\mathbf{c}_i^T \hat{\mathbf{c}}_i / \sqrt{\|\mathbf{c}_i\| \cdot \|\hat{\mathbf{c}}_i\|} = \cos(\theta_i), \quad i = 1, 2, \dots, m,$$

where the θ_i denote the angles between the vectors \mathbf{c}_i and $\hat{\mathbf{c}}_i$. The importance of this result is that it shows that the peaks in Figure 6 occur when the

Figure 7: A plot of the $\cos(\theta_i)$ for different wavelengths $[600 + 2 \times (\text{x-axis value})]$ nm.



vectors \mathbf{c}_i and $\hat{\mathbf{c}}_i$ are strongly aligned with $\cos \theta_i \sim 1$ (approximately parallel), whereas the other regions correspond to situations where the vectors \mathbf{c}_i and $\hat{\mathbf{c}}_i$ are not strongly aligned.

3 The calibration

The popularity of NIR spectroscopy arises from the ease and speed with which it can be utilized, once the calibration step is performed, to predict such things as the protein content in wheat, the casein content in milk and

whether pepper is adulterated. Various methods were proposed for performing the calibration step including the widely used partial least squares (PLS) [7] procedure. For the mean-centred matrix of second derivative spectra of the casein spiked milk powder, the calibration step corresponds to estimating an appropriate solution of the under-determined system of equations

$$\mathbf{C}\mathbf{x} = \left(\mathbf{I} - \frac{1}{n}\mathbf{e}\mathbf{e}^T\right)\mathbf{f} = \tilde{\mathbf{f}}, \quad (7)$$

such as the PLS solution \mathbf{x}_{PLS} . The corresponding PLS predictions of the added casein fractions are then determined as

$$\mathbf{f}_{\text{PLS}}(\mathbf{k}) = \mathbf{x}_{\text{PLS}}^T \mathbf{s}_{\mathbf{k}}, \quad \mathbf{k} = 1, 2, \dots, 8. \quad (8)$$

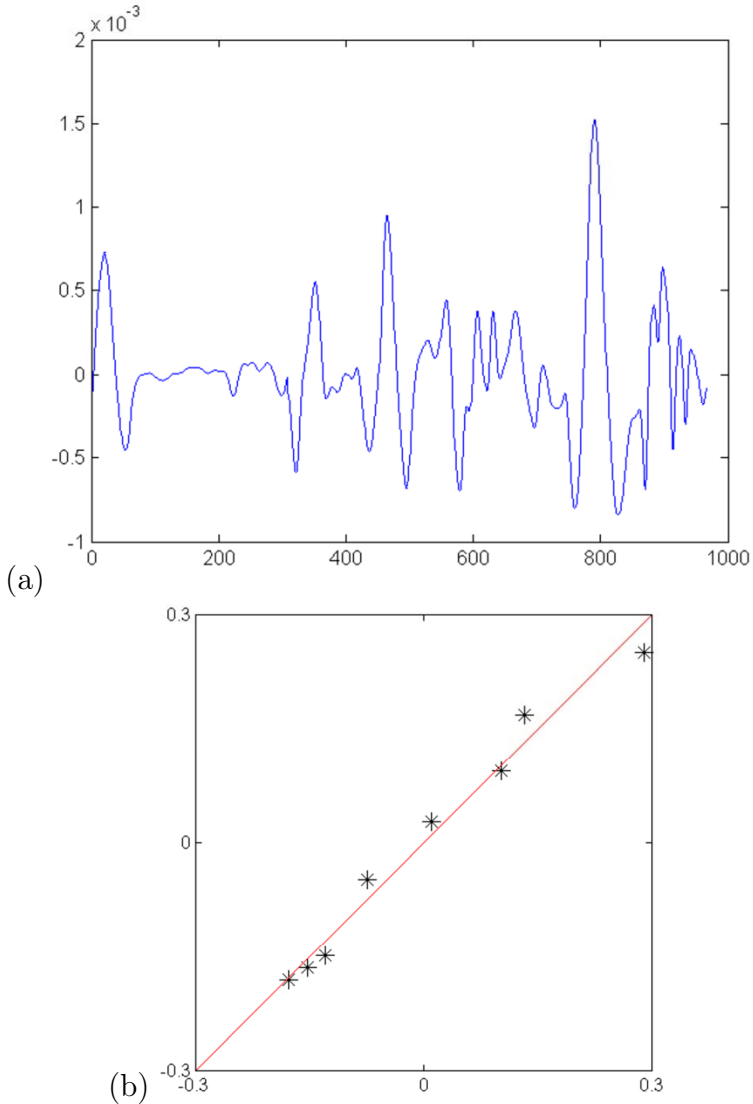
The PLS solution \mathbf{x}_{PLS} is plotted in Figure 8(a) along with a scatter plot in Figure 8(b) of the PLS predicted fractions of added casein $\mathbf{f}_{\text{PLS}}(\mathbf{k})$ as a function of the actual added casein fractions $\mathbf{f}_{\mathbf{k}}$, $\mathbf{k} = 1, 2, \dots, 8$.

A visual comparison of the peaks $\mathbf{c}_i^T \hat{\mathbf{c}}_i$ in Figure 7 with those in the PLS solution \mathbf{x}_{PLS} in Figure 8(a) shows a very strong correlation. The correlation provides support for the conclusion that the information recovered by PLS is consistent with Beer's law and identifies the wavelength components in the spectra which have a strong correlation with the measured proportions of the component being calibrated. It also gives support for the methodology proposed by Anderssen et al. [3] which uses derivative spectroscopy to recover information about the relevant wavelengths.

4 The single wavelength analysis

An interesting related question is the extent to which a single wavelength might be a good predictor of the added casein fractions. The motivation for exploring such a question in the current situation is the strong dominant peak in $\mathbf{c}_i^T \mathbf{c}_i$, at the wavelength $\approx 2198\text{nm}$ (at $i = 799$) in Figure 6, and in the PLS solution \mathbf{x}_{PLS} in Figure 8(a).

Figure 8: (a) The PLS solution of the under-determined system of equations (7). (b) Scatterplot of the PLS predictions of the fractions of added casein as a function of the actual added casein fractions.



For the subsequent discussion, let \mathbf{C}_ℓ denote the column vector of \mathbf{C} at the wavelength 2198 nm. Least squares minimization, with $\tilde{\mathbf{f}}$ defined in equation (7), gives

$$\min_{\alpha_i} \|\alpha_i \mathbf{c}_i - \tilde{\mathbf{f}}\|^2, \quad i = 1, 2, \dots, m,$$

and determines the multipliers

$$\alpha_i = (\mathbf{c}_i^T \tilde{\mathbf{f}}) / (\mathbf{c}_i^T \mathbf{c}_i),$$

for solving this least squares problem for corresponding choice of \mathbf{c}_i .

Two possibilities are examined.

- For column vector \mathbf{C}_ℓ , the corresponding value of α_ℓ is plotted in Figure 9(a), while Figure 9(b) plots the values of the column vector $\alpha_\ell \mathbf{C}_\ell$, which predicts the fractions of added casein as a function of the actual casein fractions. The PLS predictions $f_{\text{PLS}}(\mathbf{k})$ of equation (8) are also plotted in 9(b). It is clear from this plot that the least squares predictions $\alpha_\ell \mathbf{C}_\ell$ compare favourably with the PLS predictions.
- Find the largest value of the α_i . Let it and its corresponding column be denoted by α_j and \mathbf{c}_j . In terms of the plot in Figure 7, \mathbf{c}_j corresponds to the column where $\cos(\theta_j)$ is closest to the value one, and this corresponds to the angle between \mathbf{c}_i and $\hat{\mathbf{c}}_i$ being the closest to zero. For column vector \mathbf{c}_j , the corresponding value of α_j is plotted in Figure 10(a), while Figure 10(b) plots the predicted fractions of added casein $\alpha_j \mathbf{c}_j$ as a function of the actual casein fractions.

It is clear from Figures 9 and 10 that the least squares $\alpha_j \mathbf{c}_j$ predictions compare favourably with the PLS predictions.

5 Conclusions

On the basis of the above considerations and comparisons, it appears that, globally, Beer's law holds for the spiked milk powder spectra of Figure 1.

Figure 9: (a) The value of α_ℓ at the wavelenth 2198nm (x-axis value 799) of the maximum response. (b) Scatterplot of the PLS (*) and least squares $\alpha_\ell C_\ell$ (+) predictions of the fractions of added casein as a function of the actual added casein fractions.

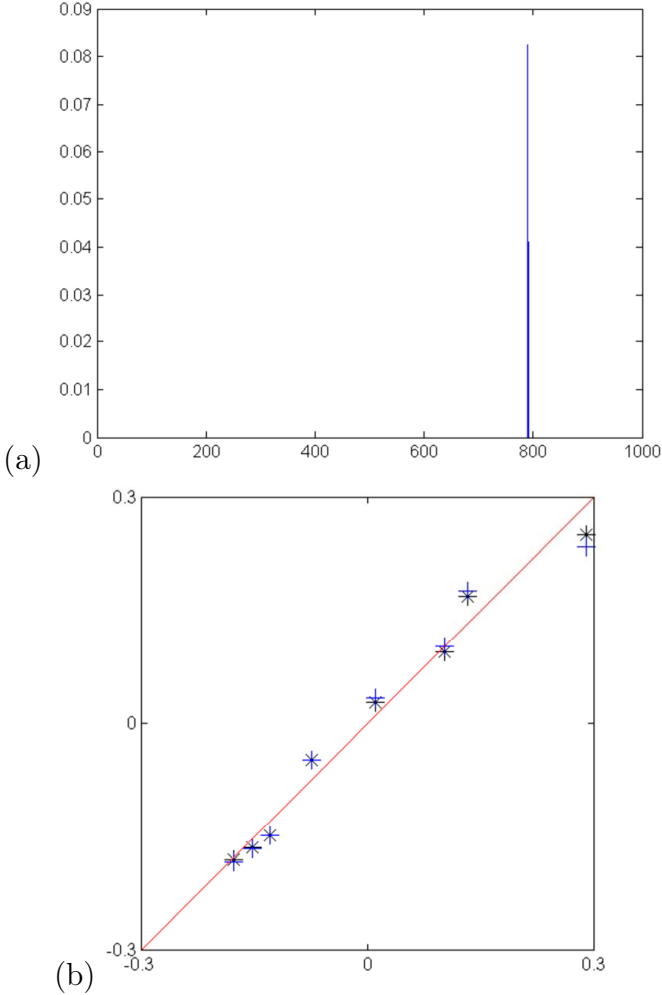
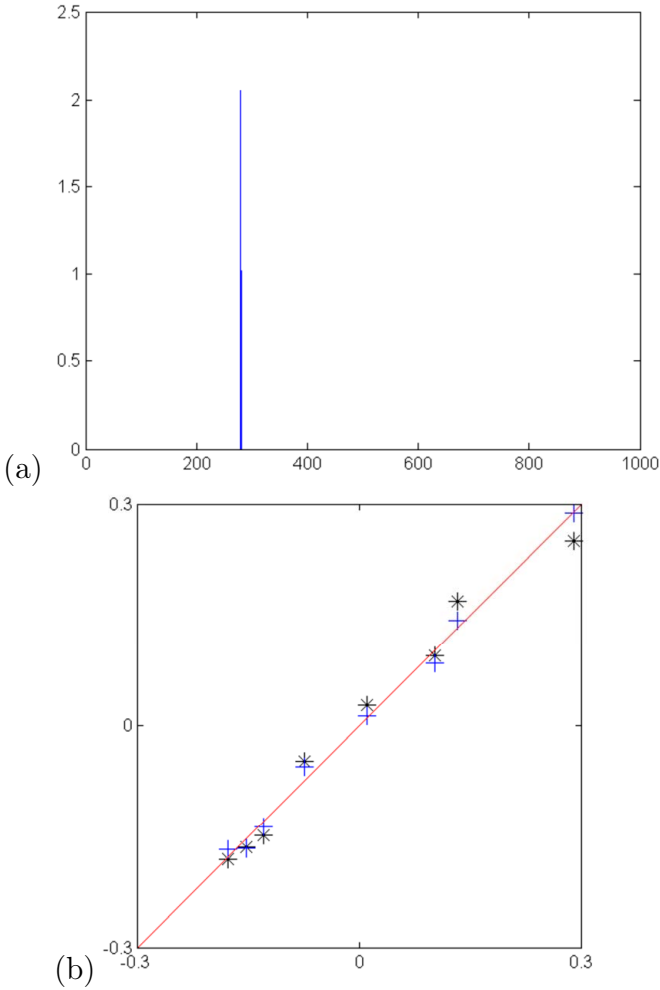


Figure 10: (a) The value of α_j which corresponds to the largest value of the α_j .
 (b) Scatterplot of the PLS (*) and $\alpha_j c_j$ (+) predictions of the fractions of added casein as a function of the actual added casein fractions.



However, as illustrated in Figures 6 and 7, when the NIR response is locally small, rounding error is likely to reduce confidence in Beer's law. Consequently, in practice, when making inferences based on comparing the NIR spectral responses of different property compositions of the same basic material, such as for different wheat varieties, only the regions of highest spectral intensity should be used. This is the essence of the approach outlined in Figures 9 and 10. In addition, in the utilization of the spectra in a calibration and prediction situation which assumes the validity of Beer's law, such a strategy minimizes the effect of second order non-linear diffraction scattering and circumvents the concerns raised by Dahm [6].

Not only does the current investigation give credence for the validity of Beer's law of NIR spectral analysis, it also gives support to the methodology proposed by Anderssen et al. [3] which is a more intuitive way of identifying the NIR spectral wavelengths which identify the spectral structures that provide the required information from the data.

References

- [1] B. Anderssen, F. de Hoog, and M. Hegland. A stable finite difference ansatz for higher order differentiation of non-exact data. *Bull. Austral. Math. Soc.* 58:223–232, 1998. doi:[10.1017/S0004972700032196](https://doi.org/10.1017/S0004972700032196) C230
- [2] R. S. Anderssen and F. R. de Hoog. Finite-difference methods for the numerical differentiation of non-exact data. *Computing* 33:259–267, 1984. doi:[10.1007/BF02242272](https://doi.org/10.1007/BF02242272) C230
- [3] R. S. Anderssen, F. R. de Hoog, I. J. Wesley, and A. B. Zwart. How much of a near infrared spectrum is useful? Sparse regularization—let the data decide! In S. McCue, T. Moroney, D. Mallet, and J. Bunder (Eds), *Proceedings of the 16th Biennial Computational Techniques and Applications Conference, CTAC-2012, ANZIAM J.* 54:C788–C808, 2012.

- <http://journal.austms.org.au/ojs/index.php/ANZIAMJ/article/view/6744> C239, C244
- [4] R. S. Anderssen and M Hegland. For numerical differentiation, dimensionality can be a blessing! *Math. Comput.* 68(227):1121–1141, 1999. doi:[10.1090/S0025-5718-99-01033-9](https://doi.org/10.1090/S0025-5718-99-01033-9) C230
- [5] R. S. Anderssen and M. Hegland. Derivative spectroscopy: An enhanced role for numerical differentiation. *J. Integral Equat. Appl.* 22:355–367, 2010. doi:[10.1216/JIE-2010-22-3-355](https://doi.org/10.1216/JIE-2010-22-3-355) C230
- [6] D. J. Dahm. Explaining some light scattering properties of milk using representative layer theory. *J. Near Infrared Spec.* 21:323–339, 2013. doi:[10.1255/jnirs.1071](https://doi.org/10.1255/jnirs.1071) C230, C244
- [7] T. Naes, T. Isaksson, T. Fearn, and T. Davies. *A User-Friendly Guide to Multivariate Calibration and Classification*. NIR Publications, Chichester, UK, 2002. doi:[10.1002/cem.815](https://doi.org/10.1002/cem.815) C231, C239

Author addresses

1. **R. S. Anderssen**, CSIRO Data61 GPO Box 664, Canberra, ACT 2601, Australia.
<mailto:Bob.Anderssen@csiro.au>
2. **F. R. de Hoog**, CSIRO Data61 GPO Box 664, Canberra, ACT 2601, Australia.
<mailto:Frank.deHoog@csiro.au>
3. **I. J. Wesley**, Grain Growers Limited, PO Box 7, North Ryde, NSW 1670, Australia.
4. **A. B. Zwart**, CSIRO Data61 GPO Box 664, Canberra, ACT 2601, Australia.
<mailto:Alec.Zwart@csiro.au>