

Always look at the spectrum: Part 2. 3-D Normalisation

A.M.C. Davies

Norwich Near Infrared Consultancy, 10 Aspen Way, Cringleford, Norwich NR4 6UA, UK. E-mail: td@nnirc.co.uk

Introduction

This is another old story but not so old as Part 1¹ and this one has not been published previously. I became an independent consultant in NIR spectroscopy in August 1989 and my first commission was to help Foss-Electric Development (Wheldrake, York, UK) with their development of a whole grain analyser "GrainSpecTM" (as it was later named).

The "spectrum"

GrainSpecTM (GS) made transmission measurements of whole grain, held in a variable pathlength cell, at 33 wavelengths in the range 800–1100 nm (see Box 1). The intention was to use these measurements to predict the protein and moisture content of the grain by Partial Least Squares (PLS) regression.² At the first meeting I asked to see what the spectrum looked like. Everyone looked puzzled; this was a team of mechanical, optical and electronic engineers and computer programmers, they knew quite a lot about measuring milk components from infrared measurements but none of them had previous experience

Box 1

According to IUPAC² the near infrared region extends from 780 nm to 2500 nm so measurements in the 800–1100 nm range are in the near infrared region. However, in order to differentiate this region from the more commonly used 1100–2500 nm range it has been given several different names. I prefer to call it the "Herschel" region as this was what Herschel discovered in 1800.^{3,4}

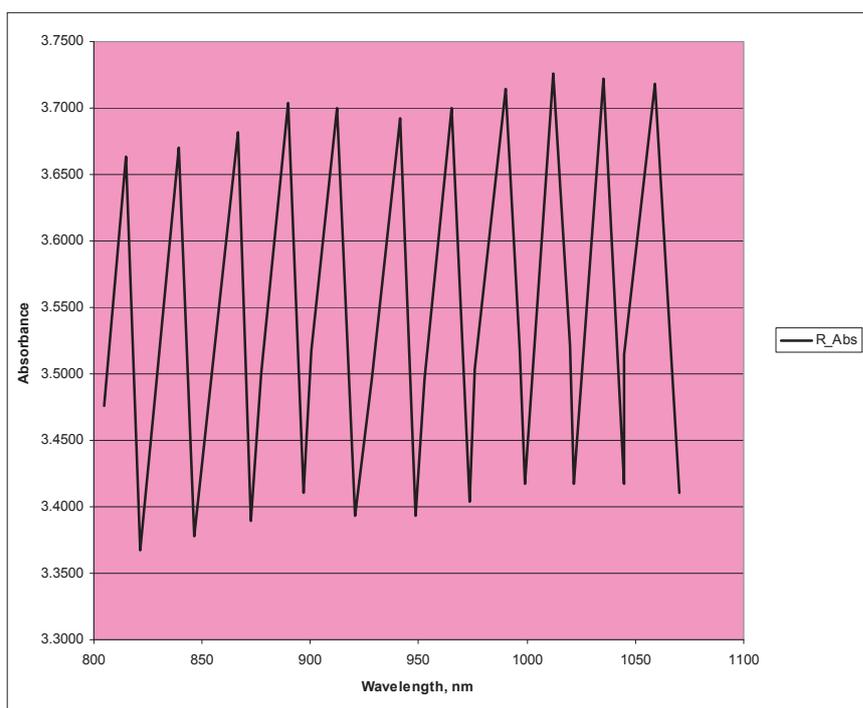


Figure 1. Original output of 33 readings.

of NIR spectroscopy or thought of themselves as "Spectroscopists". The response was: "Why would you want to see that—it's just data?" I explained that I needed to try to understand the measurements before it was put into the PLS "black-box". The data were not available and I had to wait some weeks for it to be produced in a form that I could use in ExcelTM. I was startled by the first plot; it was just a series of zigzag lines, Figure 1. In order to explain how these were produced I need to tell you a bit more about the workings of the GS.

GrainSpec

The instrument used a tungsten-halogen bulb as the energy source, NIR wave-

lengths were selected by use of a blocking filter to exclude visible radiation and then narrow bands of NIR wavelengths were selected by 11 interference filters held on a rotating wheel which brought each filter in turn into the optical path. The wavelengths selected by interference filters depend on the construction of the filter and the angle of the incident beam on the filter, thus varying the angle of the filter to the radiation will produce a range of output wavelengths. This principle was used in the low-cost "NeotecTM" range of NIR spectrometers. In GS the opposite occurred. A collimated band of energy that passed through the blocking filter illuminated a "tuning wheel". The tuning wheel, Figure 2, contained three

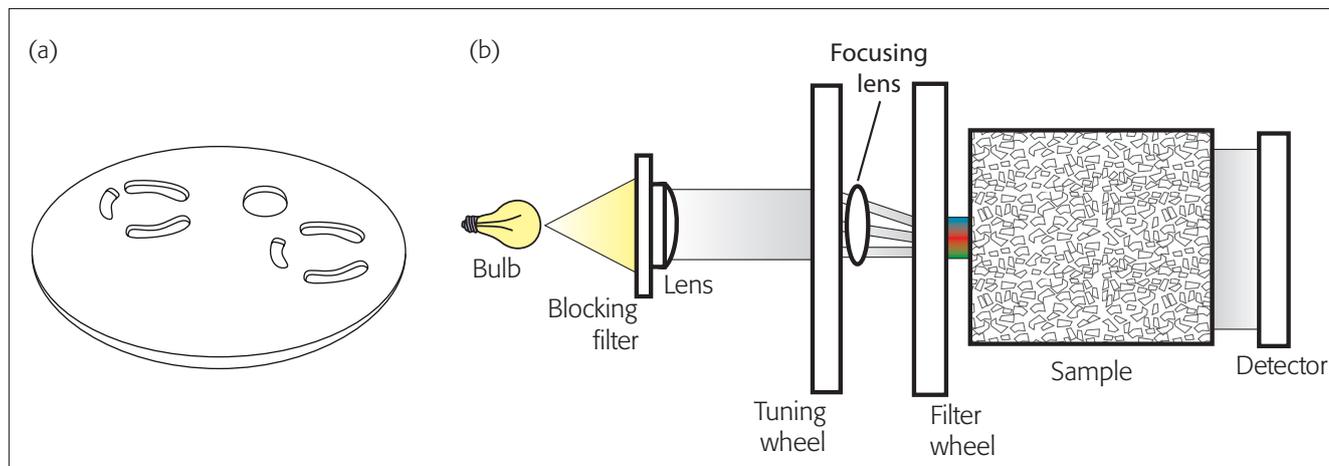


Figure 2. (a) The GrainSpec “Tuning Wheel” used to select different segments of radiation for focusing on to interference filters. (b) Schematic of the optical path.

series of cut-outs which could be positioned, in turn, in the optical path. The three cut outs selected energy at increasing distances from the axis of the optical path which when focused on to any of the filters, were at varying angles to the filter and thus produced different output radiation of approximately 15nm (TW1) and 8 nm (TW2) less than the nominal wavelength transmitted by incident radiation at right angles to the filter (the centre circle on the tuning wheel,

TW3). When I plotted absorptions separated by filter wheel position a very clear effect could be seen, Figure 3, and you can understand that each tuning wheel position is responsible for the each of the three parts of the zigzag line. My first thought was that there should be optical or mechanical improvements that could eliminate or reduce these variations but I was told that the instrument design was “frozen” so I would need to come up with chemometric ideas and

prove that they improved the calibration of the instrument. Grain was feed into the cell automatically from a hopper and a variable number (up to 10) of sub-samples was being used because there was a marked variability in results from individual sub-samples. When I looked at data from sets of sub-samples I was not surprised to find considerable variability in the zigzag lines. This had to be caused by physics rather than chemistry and (in my view) needed to be corrected.

I had several ideas about how to attempt to correct the spectra, my favourite was Fourier Transform processing as I had collaborated with Professor Fred McClure (North Carolina State University, USA) with applications of his original idea of applying FT processing to NIR data.^{5,6} FT processing could be used to produce a (relatively) smooth spectrum from the zigzag line but this idea failed when it was pointed out to me that GS would be a stand alone instrument and they were not going to be able to include FT processing on the small chip that was running the instrument. I started looking at simple normalisation techniques. The obvious one was to divide the spectrum by its average value. This did produce improvements which encouraged me to go a bit further. The additional step was to calculate an average for each of the three tuning wheel positions and use these to normalise each part of the set and then merge them into one spec-

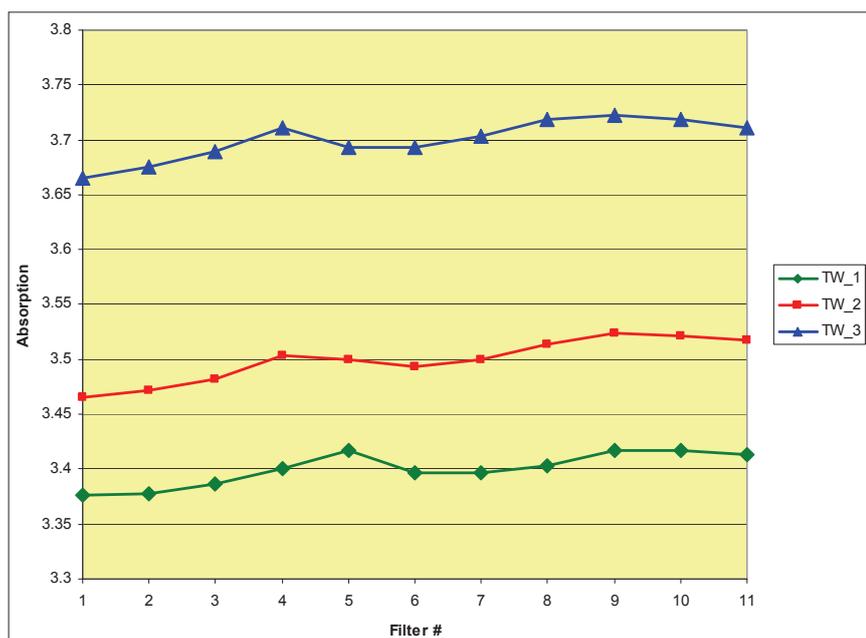


Figure 3. Readings separated by the tuning wheel used for 11 filter positions.

TONY DAVIES COLUMN

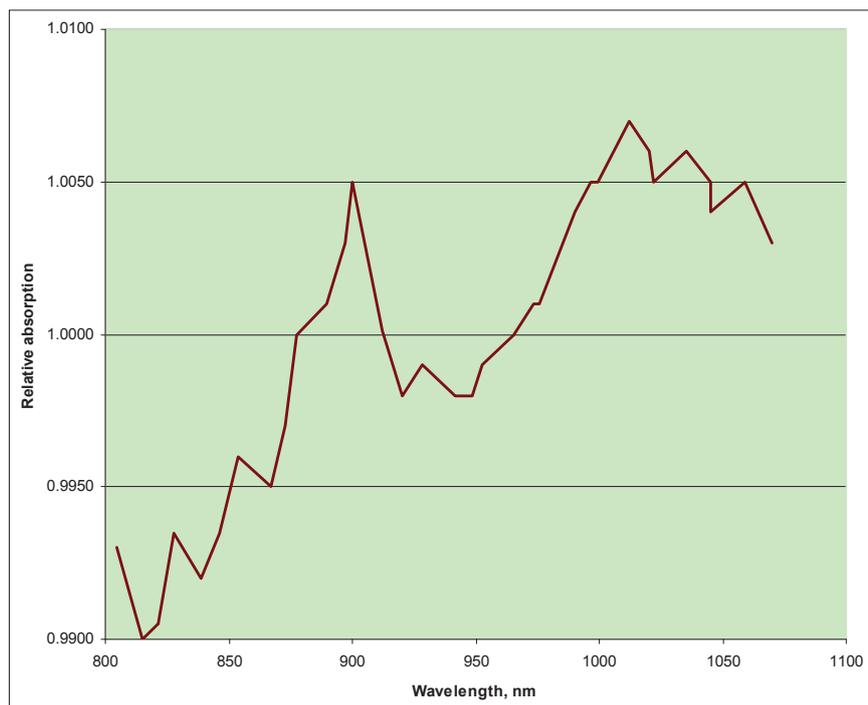


Figure 4. Output readings after division by the average value of the 11 readings for the appropriate tuning wheel position.



Figure 5. As for Figure 4 after each relative absorption has been multiplied by the average value from tuning wheel position 3.

trum. Figure 4 shows the result of this processing on the spectrum in Figure 1. In order to make the result look more like a real spectrum I multiplied the result by the averaged value of the absorptions by primary wavelengths, (TW3). This is just a change of scale, as shown in Figure 5.

Tests of the technique on sets of sub-samples gave an eight to ten fold reduction in the variability of sub-samples and it was agreed that this processing should become part of GS. I called the technique "3-D normalisation"; not a chemometrically sound

name (it is really all happening in a single dimension) but it was acceptable for internal use. After additional studies it was found that the first and last readings added more noise to the averages and so a refinement of the procedure was to use the results from filters 2–10.

As the importance of NIR in the Foss group grew, chemometricians joined the organisation. None of them liked 3-D normalisation and at least three members tested more chemometrically sophisticated techniques but much to my surprise none of them proved to be more useful than my naive normalisation! As far as I know, 3-D normalisation was incorporated into every GS, which was a successful instrument.

Conclusion

Let's return to the title; my important role in this story was to suggest that people should look at the output of the instrument before utilisation of PLS. Given the sight of the zigzag output almost every member of the team would have come-up with an improvement that was very similar if not the same as 3-D normalisation.

I hope that you will remember these examples and whenever someone tells you that the chemometrics is not working you will ask to see the spectrum.

Acknowledgement

I am very grateful to my friend Bo Büchmann for obtaining permission from the Foss Group for publication of this technique.

References

1. A.M.C. Davies, *Spectroscopy Europe* **23(2)**, 24 (2011).
2. N. Sheppard, H.A. Willis and J.C. Rigg, *Pure Appl. Chem.* **57**, 105 (1985).
3. A.M.C. Davies, "Letter to the Editor: 'Subdivisions of the Infrared Region'", *Appl. Spectrosc.* **44**, 14A (1990).
4. F.W. Herschel, *Phil. Trans. Roy. Soc. (London)* **90**, 284 (1800).
5. F.G. Giesbrecht, W.F. McClure and A. Hamid, *Appl. Spectrosc.* **35**, 210 (1981).
6. A.M.C. Davies and W.F. McClure, *Anal. Proc.* **22**, 321 (1985).