

Hyper-spectral imaging for the discrimination of milk powder

M. T. Munir^{1*}, Nick Depree¹, Brent R. Young¹ and David I. Wilson²

¹ Industrial Information & Control Centre (I2C2), University of Auckland, New Zealand

² Industrial Information & Control Centre (I2C2), Auckland University of Technology, New Zealand

* Corresponding author. Email: tajammal.munir@auckland.ac.nz

Abstract: Hyper-spectral imaging (HSI) is an emerging, hybrid process analytical technology, combining imaging and spectroscopic techniques for food quality monitoring and assessment. While this technique has recently proved popular for food quality assessment in the fruit and seafood industries, there are only a few reported applications of HSI in the dairy industry. The interest in HSI is due to its ability to process a considerable amount of spectral data over a spatial dimension. In this work we analysed three plants all making a specific valuable milk powder. However the milk powder produced by each plant is different and each plant has different key equipment types such as the dryer. It is hypothesised that there is a causal relationship here. In this paper, the potential application of HSI to discriminate between the milk powders produced at the three different plants is presented, specifically with respect to the prediction and monitoring of functional properties such as dispersibility and solubility. Principal component analysis (PCA) was applied on hyper-spectral data extracted from milk powder samples from the three plants. The results showed that the major discrimination between milk powders produced by the different factories occur in principal components (PC) 2 and 3, and not in the first PC as this component correlates to milk powder morphology. Furthermore, the potential of the HSI technique to classify the powder as either on or off-spec at close to real time speeds is explored. The current limitations of this process analytical technique and potential future developments involving HSI in the dairy industry are also discussed.

Keywords: hyper-spectral imaging, milk powder, principal component analysis.

1 Real-time quality of milk powder

Real-time quality (RTQ) is an initiative to actively manage product quality in situ during processing without the need to rely on post-manufacture testing [1]. This paper aims to evaluate the potential of hyper-spectral imaging technique to facilitate in achieving real-time quality of milk powder. Given that most of large-scale milk powder plants are continuous and require a narrow range of product specification or quality, it is prudent to consider non-invasive and spatial mapping technique such as hyper-spectral imaging for a RTQ initiative. Hyper-spectral imaging procedure integrated with a multivariate approach (mainly Principle Component Analysis (PCA)) was preferred over other sophisticated process analysers due to fact that it generates a spatial map of spectral variations, which is expected to be closely related to milk powder morphology.

This paper describes a robust strategy to distinguish between milk powders produced at different factories using hyperspectral imaging combined with some novel post-processing. This is the first step in a comprehensive real-time quality program for Dairy and will form the basis of work to establish the correlation between hyper-spectral analysis and functional properties of milk powder. Functional properties include things such as colour, taste, flavour, dispersibility and solubility for example. Prior work on this topic Munir et al. 2015 [2] has focussed on reviewing current status of process analytical technology (PAT) in the dairy industry. Process analytical technology is a widely used term in place of real-time quality.

This paper outlines the basic operations for a multivariate operation of milk powder using a hyper-spectral imaging analysis. Multivariate data analysis (mainly PCA) of hyperspectral data was employed to detect milk powder discrimination.

2 Introduction to Hyper-spectral Imaging

Hyper-spectral Imaging (HSI) is an emerging hybrid (imaging and spectroscopy) technology that generates a spatial map of spectral variations. It combines advantages of two-dimensional imaging and spectroscopy technologies to minimise the limitations of both technologies. The advantages of HSI include its non-invasive nature, modest sampling requirements; it potentially generates a large volume of information-rich data, and it generates spatial map of spectral variations simultaneously [3].

The procedure collects images as a function of wavelength. The images map the intensity of reflected light from the surface of the sample at different wavelengths. In other words, hyperspectral data contains both spatial and spectral information from product within a given image as shown in Figure 1.

The hyper-spectral image data, also called “hypercube”, is comprised of two spatial and one wavelength dimension as shown in Figure 1. Each slice or image is an absorbance of a sample image at a particular wavelength. The hyperspectral image shown in Figure 1 can be viewed as a spatial image $I(x, y)$ at a particular wavelength (λ) or as a spectrum $I(\lambda)$ at every pixel (x, y). An image at a particular wavelength is composed of smaller picture elements known as pixels. Each pixel represents single spectra, and therefore HSI acquires spatially distributed spectral responses at pixel levels. For instance, a single spectrum at a single pixel point (shown as red square) is shown in Figure 1.

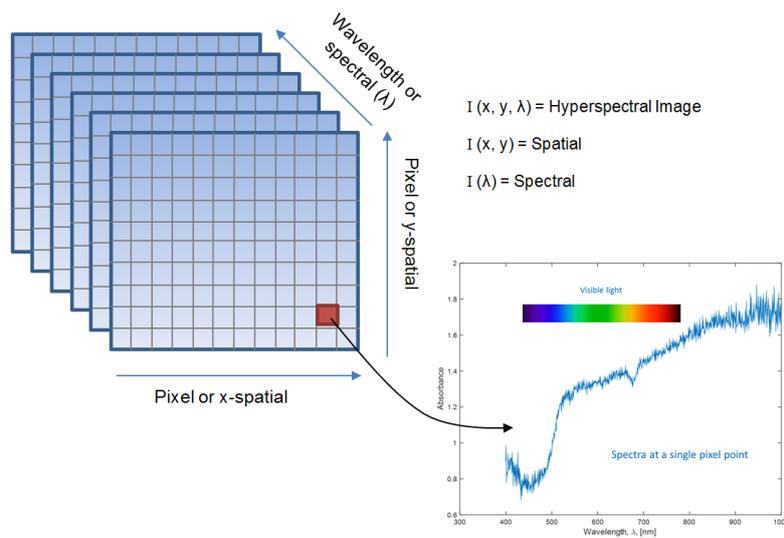


Figure 1: HSI images as a function of wavelength

Hyper-spectral imaging has been previously used in a wide array of applications such as: remote sensing, airborne hyper-spectral surveys, astronomy, agriculture, biomedical, mineralogy and food applications. For example, Kruse, 2012 [4] used HSI for mapping surface mineralogy, Giardino et al. 2015 [5] employed airborne hyper-spectral data to assess aquatic vegetation and particulate matter in a shallow and turbid lake, Fabricius et al. 2014 [6] applied hyper-spectral imaging for biomedical images, and Feng et al. 2012 [7] applied hyper-spectral imaging for food safety inspection and control. However in most of the above applications, the imaging exhibited clearly defined areas of different absorbances whereas in this milk powder application the visible differences in the images are far more subtle and not immediately apparent.

To date the use of hyper-spectral imaging has been mostly limited to remote sensing, agriculture, mineralogy or meat industries. Unlike remote sensing, agriculture, mineralogy or meat industries, the milk powder industry normally produces a product of consistent properties and texture. This seems to be a major reason of limited hyper-spectral imaging applications in dairy industry. However, there are a few case studies where hyper-spectral imaging has been applied in dairy industry as well. For example, Fu et al. 2014 [8] applied hyper-spectral imaging for the detection of melamine in milk powders, Woodcock et al. 2008 [9] applied hyper-spectral imaging to determine cheese quality and authenticity, and Qin et al. 2013 [10] used Raman hyper-spectral imaging to detect multiple adulterants in milk powder.

In other words, hyper-spectral imaging in milk powder industry has been mainly used to detect adulteration or materials of distinct properties such as melamine. However, detecting milk powder adulteration is not the focus of this article.

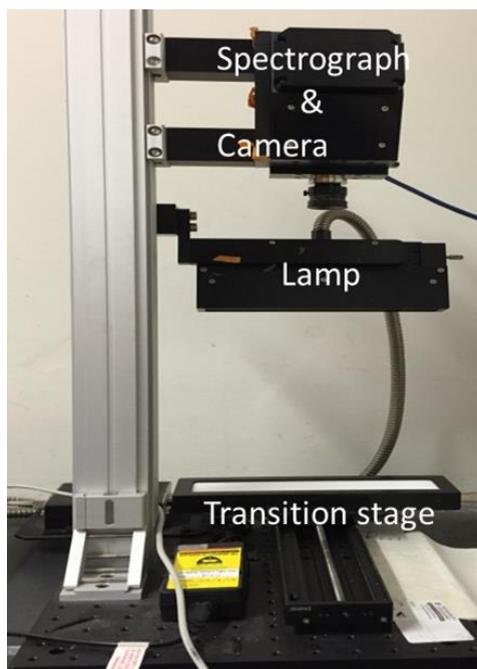
The main focus of this research is the discrimination of milk powder produced at different factories and of different quality. Quality of milk powder here is gauged by milk powder dissolution properties. Unlike detecting milk powder adulteration or materials of distinct properties, milk powder produced at different factories has not distinct properties, that makes a “hard problem” more difficult. The novelty of this research is discrimination of different milk powder of similar properties. It also includes the development of a robust hyper-spectral imaging analysis strategy integrated with multivariate data analysis to discriminate different milk powders. It is also interesting to note that Gowen et al. 2009 [11] acknowledges the importance of HSI technique for milk powder quality.

The manuscript is organised as follows. Section 3 explains HSI experimental set-up and pre-processing of hyper-spectral data. An overview of multivariate data analysis and basics of the principle component analysis are presented in sections 4 and 5 respectively. Section 6 presents two case studies: the first case study is milk powder discrimination by plant location; and the second is about testing hypothesis that PC1 is related to surface morphology. Finally, section 7 concludes this study and discusses future of real-time quality by using hyper-spectral imaging.

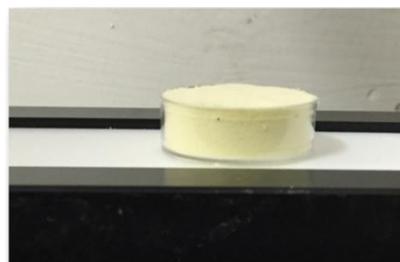
3 Materials and Methods

3.1 Experimental setup to collect the data

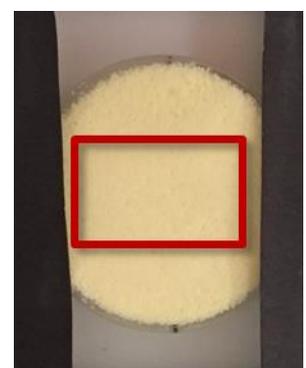
A Headwall Photonics Hyperspec™ - VNIR, C series, model no. 1003A - 10140, and with a 400 - 1000nm wavelength range, hyperspectral imaging instrument was used in this work, as shown in Figure 2. The hyperspectral imaging instrument consists of four major components: a spectrograph, camera (Schneider Kreuznach Xenoplan 1.4/23), lamp and transition stage as shown in Figure 2 (a). The spectrograph generates a spectrum for each pixel on the picture taken by camera. The lamp acts as lighting source and transition stage is used to place sample for analysis as shown in Figure 2 (b).



(a) HSI instrument and stage



(b) Powder sample on stage



(c) Sample viewed from above (red square is area of interest)

Figure 2: Hyperspectral imaging experimental setup

A milk powder sample was poured into a Petri dish and the surface was levelled as shown in Figure 2 (b). Figure 2 (c) shows the milk powder sample viewed from above. A red square in Figure 2 (c) shows

the approx. area of interest (AOI) for the HSI analysis. A square region in the centre of sample and black papers covering both edges of were used to avoid edges effect on the subsequent analysis.

The hyper-spectral data is collected using the HSI instrument. The data cube is stored and is augmented with metadata containing experiment specific information such as sample number, data and operator notes in a Matlab data structure. A standard data format here stands for a data format where row vectors represent pixels or spatial and column vectors represent wavelength or spectral variations.

3.2 Pre-processing of the hyper-spectral data

The pre-processing and analysis of the data was performed in Matlab. Due to circumstances that have yet to be satisfactorily explained, the HSI instrument exhibited some severe artefacts between 650 and 740 nm which show up as vertical lines irrespective of sample orientation and external lighting conditions. These image planes were removed from the data in a pre-processing step prior to a subsequent multivariate analysis. An image processing technique called grey level co-occurrence matrix (GLCM) [12] analysis was used to automatically detect those corrupted images or images containing artefacts.

Before doing any case studies, we evaluated HSI instrument repeatability and reproducibility. The effect of rotation angle of milk powder sample place on the transition stage of HSI instrument and effect of ambient light tests to analyse robustness and sensitivity of HSI instrument were also studied. The results showed that angle of sample rotation do not affect HSI analysis. The results also showed that an ambient light has a notable effect on HSI analysis, which means that HSI instrument covered with a black box, must be used to avoid any ambient light effects. Repeatability and reproducibility of HSI analysis results were also satisfactory.

4 Multivariate data analysis procedure

The first step in the Multivariate Data Analysis (MVDA) of pre-processed data is loading it into Matlab. The data cube and associated wavelength vector was also loaded using Matlab routines. The wavelength vector is spaced from 400 nm to 933 nm, although there are missing image planes discarded as mentioned in section 2.2. The typical spacing is $\Delta\lambda = 600/933 = 0.6431\text{nm}$.

The image dimensions are around 150 by 220 pixels measured at around 933 wavelengths. This equates to a datacube dimension of

$$S = 150\text{rows} \times 200\text{col} \times 933\text{wavelengths} = 27,990,000$$

so the corresponding memory size for double and single precision is

$$S_{\text{double}} \approx 233\text{ Mb} \quad \text{and} \quad S_{\text{single}} \approx 116\text{ Mb}$$

4.1 Plotting a pseudo colour map of a sample at one particular frequency

A pseudo colour map of a given milk powder sample at single particular frequency was first plotted, as shown in Figure 3. The slice in Figure 3 is an absorbance at a 488.95 nm wavelength. Note that the powder image is reasonably uniform with no apparent evidence of inclusions or large-scale structure.

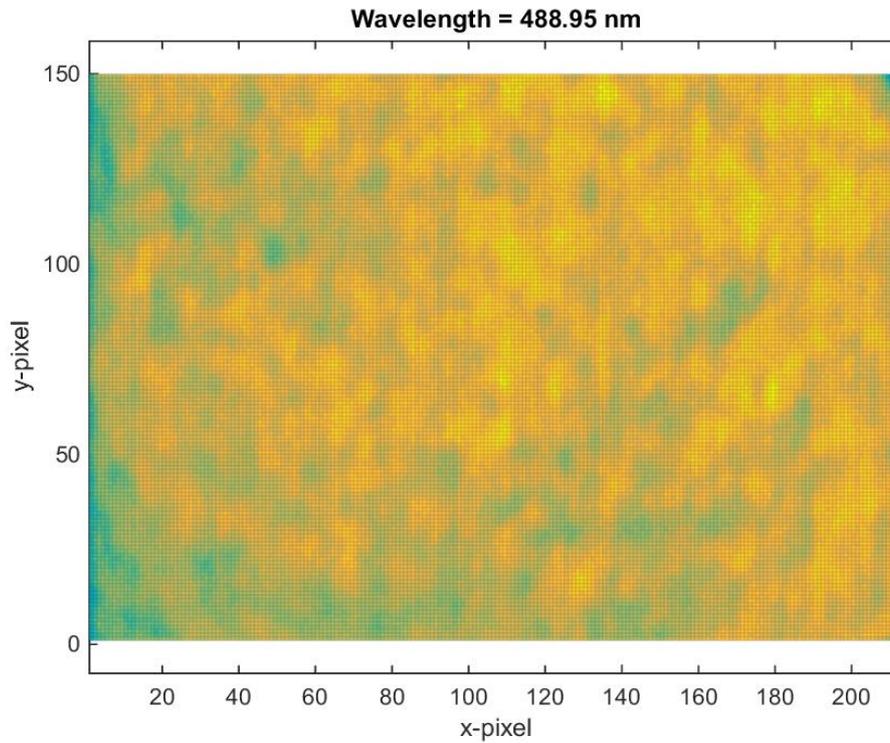
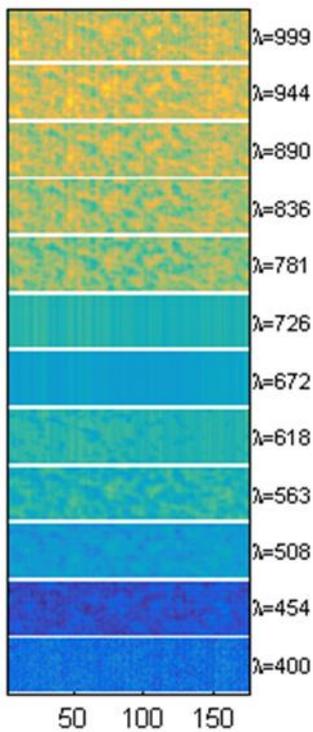
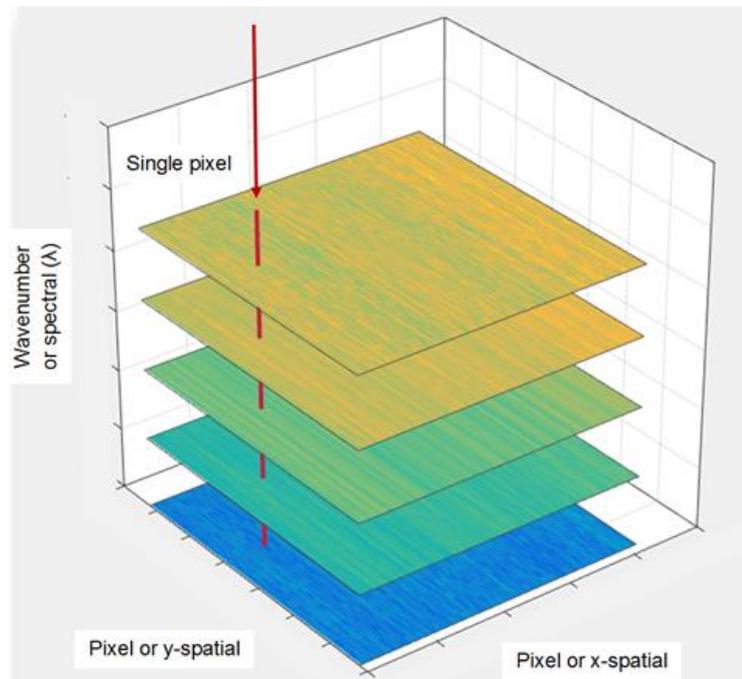


Figure 3: An image of the recorded spectra at a single wavelength, $\lambda = 488.95$ nm

At different wavelengths, a stack of such plots is given as shown in Figure 4. Each slice in a stack of such plots represents an absorbance at a particular frequency or wavelength. The number of plots in Figure 4 can be increased or decreased. A red arrow points out a single pixel or observation on a stack of hyperspectral images as shown in Figure 4 (b).



(a) A stack of images of the recorded spectra at different wavelengths



(b) Images at selected wavelengths, 3D

Figure 4: Images at selected wavelengths

If we ignore any spatial information, and simply view the data matrix as many observations, $n_r n_c \approx 150 \times 200 \approx 30,000$.

Now we have a block of normalised data \mathbf{X} where each row is an observation consisting of 933 absorbance at 933 wavelengths,

$$\mathbf{X} \in \mathfrak{R}^{30,000 \times 933}$$

Now we can plot all these individual spectra as shown in Figure 5.

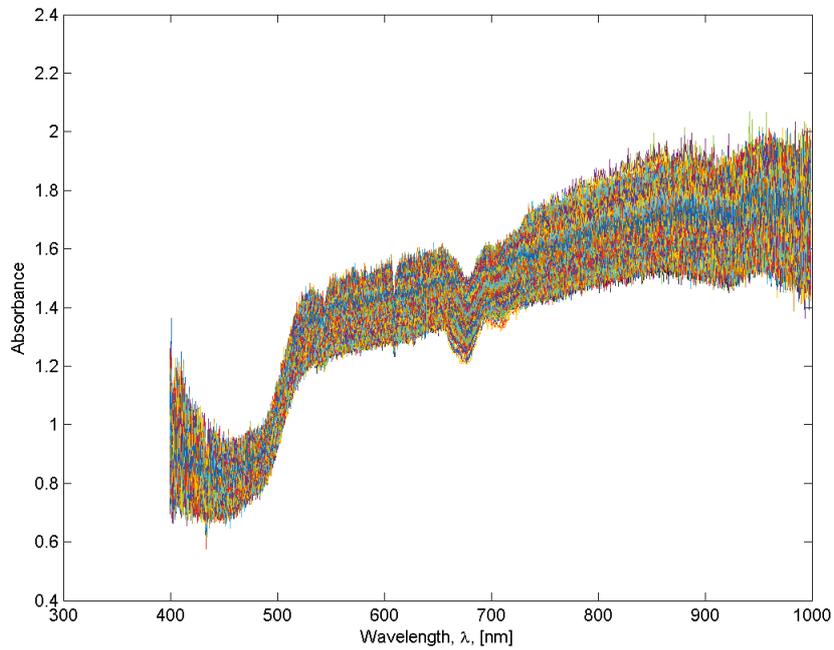


Figure 5: All 30,000 spectra

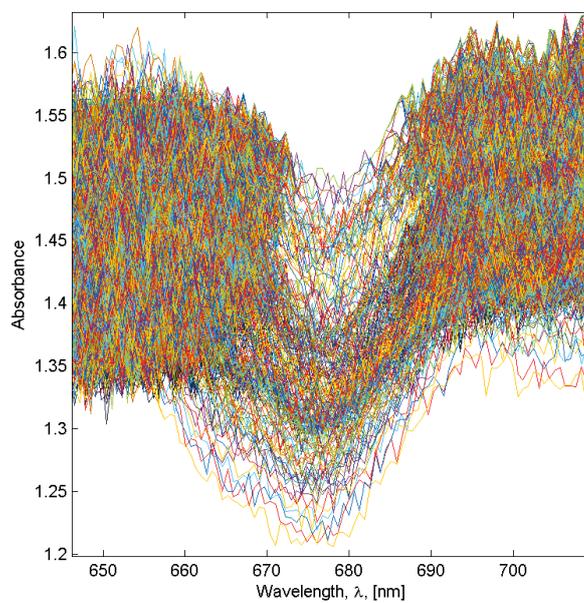


Figure 6: All spectra centered around 680nm. See also Figure 5

Note in Figure 5 the original data is the image is a strip much longer than high, $n_r < n_c$ in the pixel image, there is a clear vertical correlation in the data. This is an artefact of the instrument scanning procedure. There is a 'collapse' of data around 680 nm region, not evident in Figure 5, but is evident in the zoomed version shown in Figure 6. These were the image planes removed in the pre-processing step using the GLCM method.

4.2 Sub-sampling and a multivariate analysis

In many cases we may not want to find the principle components of all the $n_r \times n_c \times n_\lambda$ observations, simply to reduce the post-processing data analysis. For example we may just wish to select a random set of pixels, or we might want to just select pixels near the centre of the image.

4.2.1 Pre-processing and scaling

The spectral data pre-processing and scaling of hyper-spectral data is recommended before analysis to avoid any spectral variation created by ambient light, sample curvature or other effects. For example Gowen et al. 2008 [13], used reflectance calibration to account for background spectral response of the HSI instrument and Kandpal et al. 2013 [14] applied second derivative pre-process spectra of chicken samples. These methodologies were not applied in this study as we evaluated background or light effects and other effects before analysis. For example, ambient light effects and angle of sample rotation were evaluated. It seems consensus that ambient light has effect on HSI analysis, which means that HSI instrument covered with a black box must be used to avoid any ambient light effects. On the other hand, the analysis of same milk powder oriented at different angles at transition stage of HSI instrument show negligible differences.

5 Principle component analysis

Now that we have a (possibly reduced) data set \mathbf{X} , we can compute the scores and loadings of the observations. We wish to decompose the $(n \times m)$ data matrix \mathbf{X} of rank a into a sum of rank 1 matrices

$$\mathbf{X} = \mathbf{t}_1 \mathbf{p}_1^T + \mathbf{t}_2 \mathbf{p}_2^T + \dots + \mathbf{t}_a \mathbf{p}_a^T \quad (1)$$

where the vectors \mathbf{t}_i are scores and row vectors \mathbf{p}_i^T are loadings.

Name	Symbol	Dimension
# of observations	n	
# of measurements (wavelength)	m	
Data matrix	\mathbf{X}	$n \times m$
Scores	\mathbf{t}_i	$n \times 1$
Loadings, row vectors	\mathbf{p}_i^T	$1 \times m$
Component variance		$n \times 1$

The choice of the number of rank 1 matrices to keep, a , is a design parameter and is usually small, such as 2 or 3. We can concatenate all the vectors in Equation (1) together to get

$$\mathbf{X} = [\mathbf{t}_1 \ : \ \mathbf{t}_2 \ : \ \dots \ : \ \mathbf{t}_a] \begin{bmatrix} \mathbf{p}_1^T \\ \dots \\ \mathbf{p}_2^T \\ \dots \\ \vdots \\ \dots \\ \mathbf{p}_a^T \end{bmatrix} \quad (2)$$

$$\mathbf{X} = \mathbf{TP}^T \quad (3)$$

where, the $(n \times a)$ scores matrix T is made up of the first a t_i column vectors concatenated, and likewise the $(a \times m)$ loadings P^T matrix is made of the first a rows vectors p_i^T stacked on top of each other.

Further details on the PCA procedure are given in the tutorial paper, [15], although in this study we used the PCA routines from the Statistics toolbox for Matlab.

6 Case studies

6.1 Case study 1: milk powder discrimination by plant location

This first case study tries to distinguish between powder produced at three different plants (factory 'A', factory 'B' and factory 'C') located in different geographical locations, all making the same valuable milk powder. While each plant has slightly different key equipment such as dryers and evaporator configuration, they are required to produce a consistent product with the same functionals. Offline milk powder functional tests have shown that the products are different, and these differences are the motivation for this quality study.

Samples from three different factories 'A', 'B', and 'C' were taken and analysed following procedure discussed in sections 2 and 3.

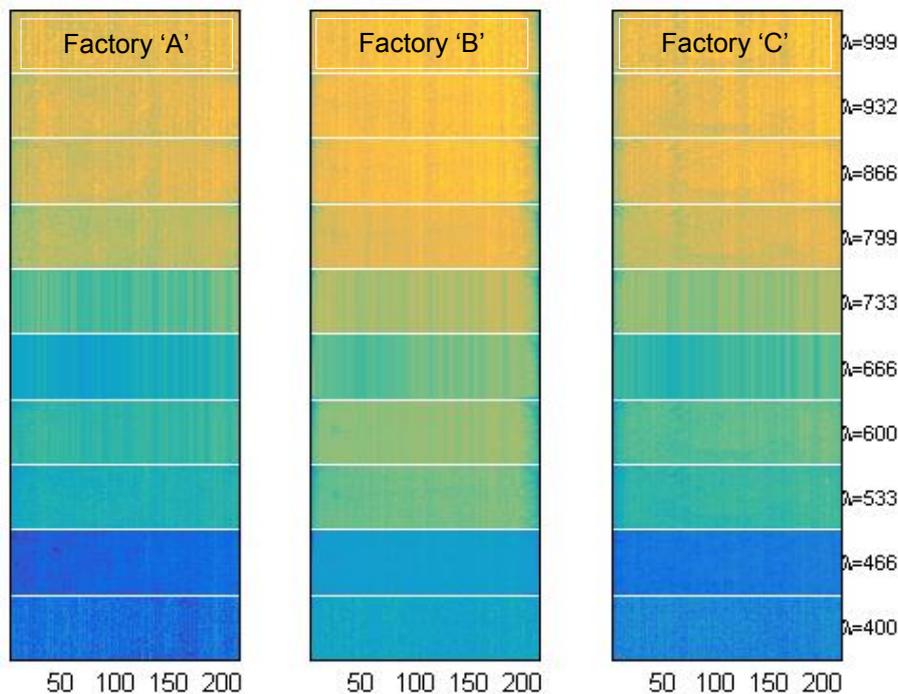


Figure 7: Comparison of 3 powders made at 3 different factories

Figure 7 compares 3 milk powder samples made from 3 different factories: 'A', 'B', and 'C'. It's a stack of such absorbance plots at different wavelength for three different powders. Notwithstanding the minor discrimination across different powders, the major discrimination across different powders can't be done with naked eye.

After performing a principle component analysis, a scores plot of the data is shown in Figure 8 where the powders are coloured according to factory (factory 'A' = blue, factory 'B' = red, factory 'C' = green). Around 10^3 randomly chosen samples (from the 2,000) were used for the PCA model building. The most obvious differences are in PC2 vs PC3 (as oppose do data including PC1) which is shown more clearly in Figure 9. Here three confidence ellipses are superimposed onto the data. Figures 8 and 9 clearly show discrimination between powders of different factories. Figure 8 is also an attempt to standardise how we present multivariate data in restricted dimensions. One method is to use a 3rd angle projection of the 3D data cube, and present the 2D 'shadows' of that data.

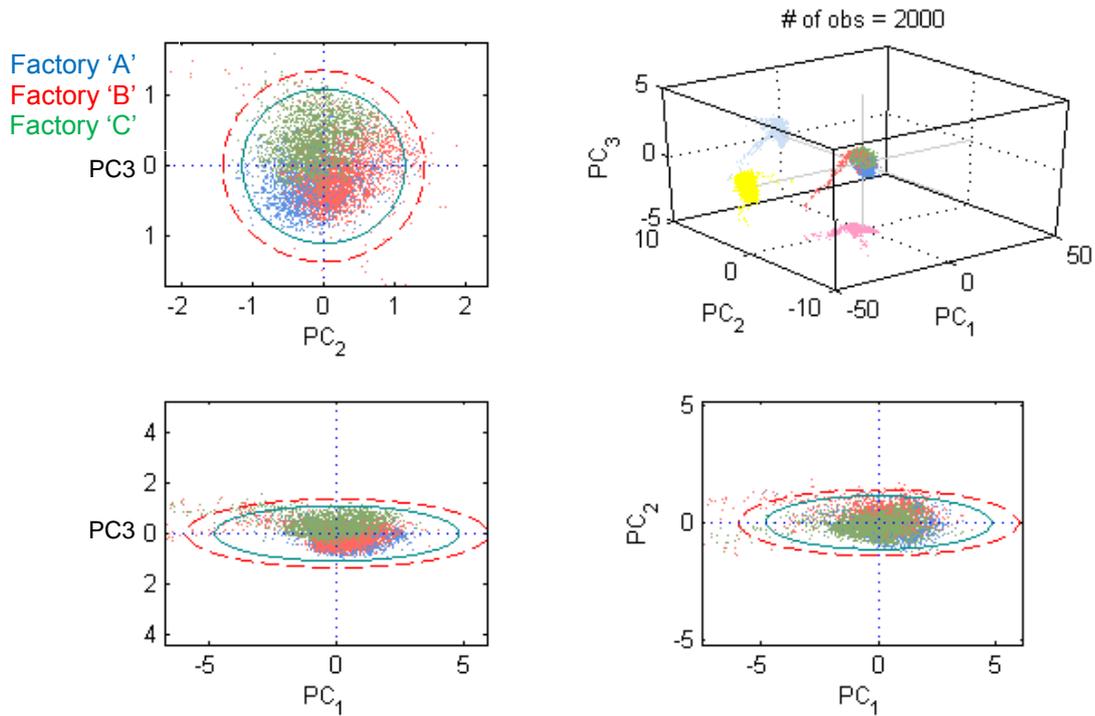


Figure 8: The first 3 principle components of 2000 spectra observations. (Factory 'A'=blue, factory 'B'=red, factory 'C'=green)

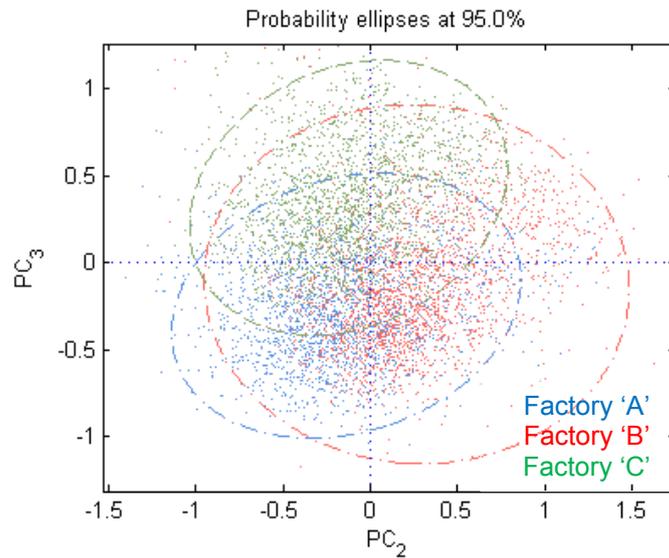


Figure 9: A zoomed plot of PC2 vs PC3 with individual confidence ellipses

The analysis of milk powder from three different factories given in this section showed differences primarily in PC2 and PC3 with very little difference in PC1. This is consistent with a previously published hypothesis that PC1 was related to surface morphology and higher scores were more related to the component's chemistry. Notwithstanding milk powder discrimination from three factories, more experimentation and analysis of milk powder from three different factories are required to develop a consensus.

6.2 Case study 2: testing hypothesis that PC1 is related to surface morphology

A hypothesis previously mooted in [16, 17] that the first principle component, PC1, was related to surface morphology was tested using same milk powder, but prepared differently. The difference in powder preparation methods means here is that sample 'A' was 'un-flattened' on the sample stage prior to the HSI spectral reading, while sample 'B' was flattened as shown in Figure 10. Two samples of powder from the same factory were used.

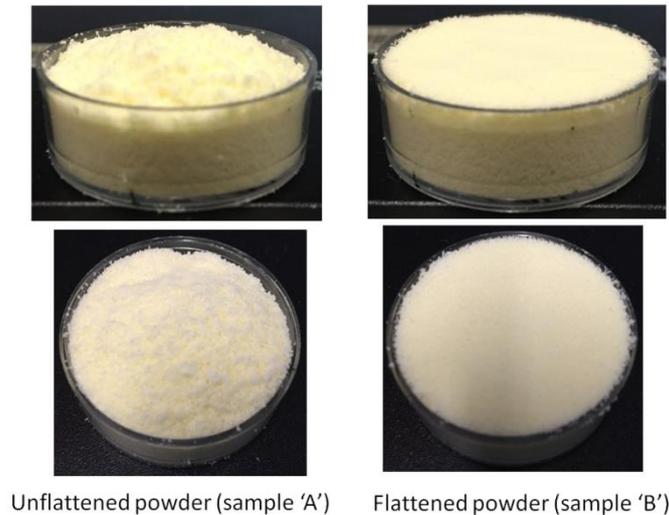


Figure 10: Hypothesis test that PC1 was related to surface morphology

Figure 11 shows the scores plot in three dimensions (for samples 'A' (un-flattened) and 'B' (flattened)). It is interesting to note that as opposed to section 5.1, in this case the differences are in PC1 and there is no difference between the samples in the PC2 vs PC3 space as highlighted in the zoomed portion re-plotted in Figure 11.

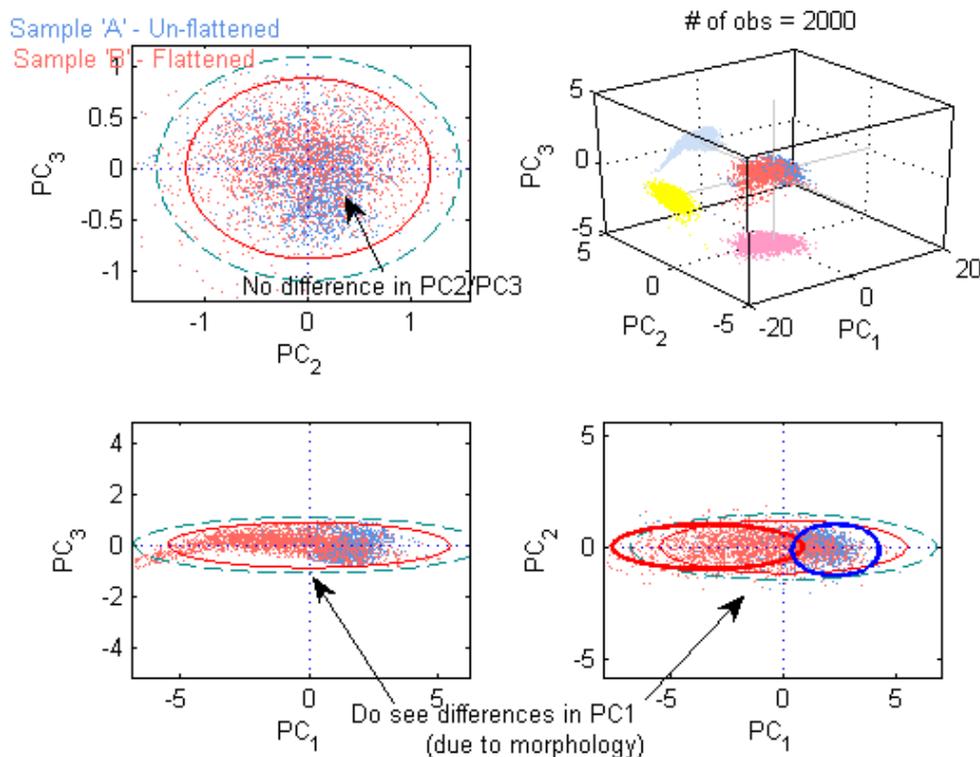


Figure 11: Scores plot of flat and un-flattened samples. Note that in this example the PC1 is different

The published hypothesis that PC1 was related to surface morphology is tested in this case study. The analysis of the same powder, but prepared differently given in this section shows differences in PC1, but not differences in PC2 or PC3. This also reinforces the hypothesis that PC1 was related to milk powder surface morphology.

7 Conclusions and future of real-time quality

This paper demonstrates the usefulness of HSI applied to milk powder. The analysis of milk powder from three different factories given in section 5.1 show differences primarily in PC2 and PC3 with very little difference in PC1. This is consistent with a previously published hypothesis that PC1 was related to surface morphology and the higher scores (PC2 and PC3) were more related to the component's chemistry. The analysis of the same powder, but prepared, and presented differently given in section 5.2 shows differences in PC1, but not differences in PC2 or PC3. This also reinforces the hypothesis.

The analysis given in sections 5 indicates that the hyper-spectral imaging technique looks to have promising potential to become a non-invasive and on-line monitoring technique for large scale milk powder plants. Notwithstanding, we acknowledge that RTQ by using hyper-spectral imaging technique has not fully been achieved yet, a robust strategy to distinguish different milk powder using hyper-spectral imaging has been developed. Building correlations to functional properties or quality attributes also yet to be made to close the loop of real-time milk powder quality.

Acknowledgements

The authors would like to acknowledge the Primary Growth Partnership program (PGP) from the New Zealand Ministry of Primary Industries for funding the project, and would also like to thank Fonterra for providing resources and support throughout this ongoing project.

References

1. Hunter, T.J., Russell, N. T, Wilson, D. I, Young, B. R, Munir, T. M, Depree, N., (2012), *Developing a real-time quality programme for dairy*, in *Chemeca: Quality of life through chemical engineering*: 23-26 September, Wellington, New Zealand.
2. Tajammal Munir, M., W. Yu, B.R. Young, and D.I. Wilson. (2015), *The current status of process analytical technologies in the dairy industry*. Trends in Food Science & Technology, **43**(2): p. 205-218.
3. Chang, C.I. (2003), *Hyperspectral imaging: techniques for spectral detection and classification*, Vol. 1. Springer Science & Business Media.
4. Kruse, F.A. (2012), *Mapping surface mineralogy using imaging spectrometry*. Geomorphology, **137**(1): p. 41-56.
5. Giardino, C., M. Bresciani, E. Valentini, L. Gasperini, R. Bolpagni, and V.E. Brando. (2015), *Airborne hyperspectral data to assess suspended particulate matter and aquatic vegetation in a shallow and turbid lake*. Remote Sensing of Environment, **157**(2015): p. 48-57.
6. Fabricius, H. and O. Pust. (2014), *Linear Variable Filters (LVF) for Biomedical and Hyperspectral Imaging Applications*. in *Biomedical Optics 2014*, Miami, Florida: Optical Society of America.
7. Feng, Y.-Z. and D.-W. Sun. (2012), *Application of Hyperspectral Imaging in Food Safety Inspection and Control: A Review*. Critical Reviews in Food Science and Nutrition, **52**(11): p. 1039-1058.
8. Fu, X., M.S. Kim, K. Chao, J. Qin, J. Lim, H. Lee, A. Garrido-Varo, D. Pérez-Marín, and Y. Ying. (2014), *Detection of melamine in milk powders based on NIR hyperspectral imaging and spectral similarity analyses*. Journal of Food Engineering, **124**(2014): p. 97-104.
9. Woodcock, T., C. Fagan, C. O'Donnell, and G. Downey. (2008), *Application of Near and Mid-Infrared Spectroscopy to Determine Cheese Quality and Authenticity*. Food and Bioprocess Technology, **1**(2): p. 117-129.
10. Qin, J., K. Chao, and M.S. Kim. (2013), *Simultaneous detection of multiple adulterants in dry milk using macro-scale Raman chemical imaging*. Food Chemistry, **138**(2-3): p. 998-1007.
11. Gowen, A.A., J. Burger, D. O'Callaghan, and C. O'Donnell. (2009), *Potential applications of hyperspectral imaging for quality control in dairy foods*. in *1st international workshop on computer image analysis in agriculture, Potsdam, Germany*.
12. Pons, M.N., H. Vivier, K. Belaroui, B. Bernard-Michel, F. Cordier, D. Oulhana, and J.A. Dodds. (1999), *Particle morphology: from visualisation to measurement*. Powder Technology, **103**(1): p. 44-57.

13. Gowen, A.A., C.P. O'Donnell, M. Taghizadeh, P.J. Cullen, J.M. Frias, and G. Downey. (2008), *Hyperspectral imaging combined with principal component analysis for bruise damage detection on white mushrooms (Agaricus bisporus)*. Journal of Chemometrics, **22**(3-4): p. 259-267.
14. Kandpal, L.M., H. Lee, M.S. Kim, C. Mo, and B.-K. Cho. (2013), *Hyperspectral reflectance imaging technique for visualization of moisture distribution in cooked chicken breast*. Sensors, **13**(10): p. 13289-13300.
15. Geladi, P. and B.R. Kowalski. (1986), *Partial least-squares regression: a tutorial*. Analytica Chimica Acta, **185**(1986): p. 1-17.
16. Otsuka, M. (2004), *Comparative particle size determination of phenacetin bulk powder by using Kubelka–Munk theory and principal component regression analysis based on near-infrared spectroscopy*. Powder Technology, **141**(3): p. 244-250.
17. Holroyd, S., (2014), *The Fonterra Research and Development Centre (FRDC) Meeting*: Palmerston north, New Zealand.

Presenting author biography



Dr Taj Munir received his PhD degree in the Department of Chemical and Materials Engineering at The University of Auckland, New Zealand in 2012. His PhD research involved designing controllable and eco-efficient plant wide control structures. Since 2012, Taj has been working with both Department of Chemical and Materials Engineering at The University of Auckland and Industrial Information & Control Centre (I2C2). He also contributes to a range of other projects including process plant simulation and process analytical technologies (PAT) for quality improvement.