

# NEAR INFRARED HYPERSPECTRAL IMAGING AND CHEMOMETRICS FOR EXPLORATION AND CLASSIFICATION OF WHOLE WHEAT KERNELS

**GERIDA DU TOIT**

Thesis presented in partial fulfilment of the requirements for the degree of

**MASTER OF SCIENCE IN FOOD SCIENCE**



Department of Food Science  
Faculty of AgriSciences  
Stellenbosch University

**Study Leader:** Dr Marena Manley  
**Co-study Leader:** Prof Paul Geladi

December 2009

## **Declaration**

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 26 November 2009

## Abstract

Near infrared (NIR) hyperspectral imaging together with multivariate image analysis was evaluated as a non-destructive method to distinguish between whole wheat kernels differing in hardness; and also to track the diffusion of conditioning water into whole wheat kernels of different hardness over a conditioning period of 36 hours.

Wheat kernels of varying hardness were imaged using a Spectral Dimensions MatrixNIR imaging system with a wavelength range of 960-1662 nm. Principal component analysis (PCA) was applied to clean the image, which entailed removal of bad pixels (background, shading, curvature errors, dead pixels and outliers). PCA also proved effective in the identification and classification of clusters in the score plot, relating to different hardness wheat endosperm (durum, hard and soft). PC 2 differentiated soft endosperm from hard and durum endosperm; while PC 3 distinguished durum endosperm from hard and soft endosperm. The loading line plot of PC 2 indicated absorbance peaks at 1195, 1450 and 1570 nm associated with starch, moisture and protein; while the loading line plot of PC 3 indicated absorbance peaks at 1195 and 1450 nm associated with starch and moisture. Partial least squares discriminant analysis (PLS-DA) was used to determine the ability to discriminate between different hardness endosperm classes using NIR hyperspectral imaging. The model of soft versus (vs) durum endosperm obtained a classification accuracy of 100%; the model of soft vs hard endosperm 98% classification accuracy; and the model of hard vs durum endosperm model classification accuracy up to 96%.

NIR hyperspectral images were acquired using the sisuChema SWIR (short wave infrared) imaging system with a wavelength range of 1000 to 2500 nm. Images of wheat conditioned with water ( $H_2O$ ) and deuterium oxide ( $D_2O$ ), respectively, were acquired at regular intervals between 0 and 36 hours. PCA proved effective in cleaning the image. The score images of PC 3 for wheat conditioned with  $H_2O$  indicated an increase in intensity over conditioning time. The loading line plots of PC 3 for wheat conditioned with  $H_2O$  indicated the variation in PC 3 due to bound moisture (1940 nm). Comparing the results from the score images and loading line plots, a conclusion could be made that the diffusion of conditioning water into soft wheat kernels reaches equilibrium after 18 hours, 24 hours for hard wheat and 36 hours for very hard wheat. The score images of wheat conditioned with  $D_2O$  indicated an increase in intensity within either PC 3 or PC 5; intensity increases were between 0 and 6 hours with no further increase up to 36 hours conditioning. The loading line plots of PC 3 and PC 5 indicated variation in these PCs due to  $D_2O$  (1954 nm). In contrast to results obtained with  $H_2O$ ,  $D_2O$  did not diffuse into the wheat endosperm as expected.

NIR hyperspectral imaging proved effective in differentiating between whole wheat kernels differing in hardness; and also in tracking the diffusion of conditioning water into whole wheat kernels.

## Uittreksel

Die gebruik van naby infrarooi (NIR) hiperspektrale beelding en veelvoudige beeldanalise is beoordeel as 'n nie-destruktiewe metode om onderskeid te tref tussen heel koringkorrels van verskillende hardhede, sowel as die volg van die diffusie van aanklammingswater in heel koringkorrels van verskillende hardhede oor 'n periode van 36 uur.

NIR hiperspektrale beelde is verkry van verskillende hardhede koringkorrels deur gebruik te maak van 'n *Spectral Dimensions Matrix* NIR kamera met 'n spektrale reikwydte van 960-1662 nm. Hoofkomponent-analise (HKA) is toegepas om die beeld skoon te maak. Skoonmaak van die beeld het ingesluit die verwydering van agtergrond, skadu, krommingsfoute en dooie piksels. HKA is doeltreffend gebruik vir die identifikasie en klassifikasie van histologiese klasse naamlik durum-endosperm, harde-endosperm en sagte-endosperm. Hoofkomponent (HK) 2 het duidelike onderskeid getref tussen sagte-endosperm en hard- en durum-endosperm; terwyl HK 3 onderskeid tussen durum-endosperm en harde- en sagte-endosperm aangedui het. Die HK lading-stip van HK 2 het absorpsiepieke by 1195, 1450 en 1570 nm aangedui wat met stysel, vog en proteïen geassosieer kon word. Die lading-stip van HK 3 het absorpsiepieke by 1195 en 1450 nm aangedui wat verband hou met stysel en vog. Parsiële kleinste waarde diskriminant-analise (PKW-DA) is gebruik om die moontlikheid van diskriminasie tussen verskillende hardhede koring vas te stel deur van NIR hiperspektrale beelding gebruik te maak. Die model van sagte- teenoor durum-endosperm het 'n klassifikasie koers van 100% bereik, die model van sagte- teenoor harde-endosperm 'n klassifikasie koers van 98%; en die model van harde- teenoor durum-endosperm 'n klassifikasie koers van 96%.

Beelde van koring, aangeklam met water ( $H_2O$ ) en deuteriumoksied ( $D_2O$ ), onderskeidelik, is verkry deur gebruik te maak van die *sisuChema SWIR (short wave infrared)* kamera met 'n spektrale reikwydte van 1000-2500 nm. Beelde van die aangeklamde koring is by gereelde intervalle oor 'n periode van 36 uur verkry. HKA kon effektief gebruik word om die beeld skoon te maak. Die telling-beeld van HK 3, vir koring met  $H_2O$  aangeklam, het 'n toename in intensiteit oor die aanklammingstydperk getoon, terwyl die lading-stip van HK 3 aangedui het dat die variasie in die HK aan gebonde vog (1940 nm) toegeskryf kon word. Deur die resultate van die lading-stip en telling-stip te vergelyk kon daar tot 'n gevolgtrekking kom dat die diffusie van aanklammingswater in sagte korrels na 18 uur 'n ewilibrum bereik het, terwyl die ewilibrum na 24 uur in harde en 36 uur in baie harde korrels bereik is. Die telling-beelde van koring, aangeklam met  $D_2O$ , het 'n toename in intensiteit aangedui in HK 3 of HK 5, onderskeidelik. Die intensiteit in die telling-beeld het toegeneem vanaf 0 tot 6 uur na aanklamming, waarna daar geen verdere toename in intensiteit tot en met 36 uur was nie. Die HK lading-stip van HK 3 en HK 5 het aangedui dat die variasie in hierdie hoofkomponente aan  $D_2O$  (1954 nm) toegeskryf kon word. In teenstelling met die resultate verkry van die aanklamming met  $H_2O$ , het  $D_2O$  geen diffusie binne die korrel aangedui nie.

Die gebruik van NIR hiperspektrale beelding was suksesvol om onderskeid te tref tussen heel koringkorrels van verskillende hardhede, en ook met die volg van die diffusie van aanklammingswater in heel koringkorrels.

## Acknowledgements

I recognise the following persons and institutions for their contribution to the successful completion of this thesis:

Dr Manley, my study leader, for all her help, support and advice during the last two years;

Prof Paul Geladi (Swedish University of Agricultural Sciences, Umeå, Sweden), my co-study leader, for help and advice, especially with the image analysis and chemometrics;

Dr David Nilsson and Oskar Jonsson (Umbio AB, Umeå, Sweden) for use of the sisuChema imaging system, assistance with the image acquisition, and also for support and supply of the Evince software;

Staff of Sasko Research and Development, Paarl for conditioning information and suggestions (Willem Koch, Divan September, Arie Wessels and Carien Roets);

The Winter Cereal Trust and National Research Foundation (NRF) for bursaries;

The Winter Cereal Trust for project funding;

The South African-Swedish Research Partnership Programme Bilateral Agreement, NRF, (UID 60958) for funding to visit the Swedish University of Agricultural Sciences for image acquisition purposes;

The Council for Near Infrared Spectroscopy for funding to attend the 14<sup>th</sup> International Diffuse Reflectance Conference 2008, Chambersburg, Pennsylvania, USA;

The Cereal Science and Technology Southern Africa for funding to attend the AACC International Annual Meeting 2008, Hawaii, USA;

Sensako (Pty) Ltd. for supplying wheat samples, and in particular Jan Cilliers for his assistance;

Dr Phil Williams (PDK Projects, Canada) for assistance and guidance with conditioning;

All the staff and postgraduate students of the Department of Food Science;

Paul Williams for all his help, support and advice the past two years;

To my parents and family for their love and support through all my studies, and especially the last year;

And above all I thank my Heavenly Father.

## Table of contents

Declaration	ii
Abstract	iii
Uittreksel	iv
Acknowledgements	vi
<b>Chapter 1:</b> Introduction	1
<b>Chapter 2:</b> Literature review	11
<b>Chapter 3:</b> Determination of wheat kernel hardness by near infrared hyperspectral imaging and hyperspectral image analysis	49
<b>Chapter 4:</b> Tracking diffusion of conditioning water in single wheat kernels differing in hardness by near infrared (NIR) hyperspectral imaging	70
<b>Chapter 5:</b> General discussion and conclusion	95

Language and style used in this thesis are in accordance with the requirements of the *International Journal of Food Science and Technology*. This thesis represents a compilation of manuscripts where each chapter is an individual entity and some repetition between chapters has, therefore, been unavoidable



# **CHAPTER 1**

## **Introduction**

## Introduction

Wheat used for food applications consist of mainly two species, i.e. bread wheat (*Triticum aestivum*) and durum wheat (*Triticum turgidum*). Wheat flour is characterised by its chemical and physical properties, which will ultimately affect functional quality, nutritional contribution and commercial value (Bietz, 1989; Kent & Evers, 1994; O'Brien & DePauw, 2004). Grain quality is important to the producer as it provides the best wheat grade and economic return. The miller requires optimum flour yield, while good baking performance is important to the baker to provide acceptable end products to the consumer (Sissons *et al.*, 2006). Breeders are constantly aspiring to improve the quality of newly developed cultivars. An important physical characteristic of wheat influencing quality, and considered in breeding programmes, is wheat hardness (Pomeranz & Williams, 1990).

Wheat hardness influences the milling process, including conditioning, as it determines the amount of water to be added and the time required for complete water diffusion. Conditioning prepares the grain for the dry milling process. Depending on the hardness of the wheat, the grain is conditioned up to 48 hours. The conditioning process allows the efficient separation of bran and endosperm, due to the mellowing of the endosperm and toughening of the bran layers. Mellowing of the endosperm prevents excessive starch damage during the milling process. Toughening of the bran layers prevents powdering of the bran, thus enabling a more thorough separation of bran and endosperm. By ensuring optimal separation of bran and endosperm the brightness (Irvine & Anderson, 1952; Prabhasankar *et al.*, 2000) and baking quality of the resulting flour are improved (Smith, 1956; Bass, 1988; Hosene, 1994; Dexter & Sarkar, 2004; Kweon *et al.*, 2009).

Wheat hardness has been studied extensively and its chemical basis is reasonably well understood (Simmonds *et al.*, 1973; Simmonds, 1974; Cobb, 1986; Greenwell & Schofield, 1986; Schofield & Greenwell, 1987; Bakhella *et al.*, 1990; Pomeranz & Williams, 1990; Hosene, 1994; Turnbull & Rahman, 2002). AACC International Approved Methods of Analysis for hardness determination include the single kernel characterisation system (SKCS) for whole grain (Method 55-31.01)(AACC, 2009c); and the particle size index (PSI) (Method 55-30.0)(AACC, 2009b) and near infrared (NIR) spectroscopy (Method 39-70.02)(AACC, 2009a) for flour. These and other methods are based on particle size (PSI and NIR spectroscopy); resistance to deformation (SKCS); grinding resistance (Stenvert hardness tester); and vitreousness (farinator) (Osborne, 1991).

Studies have been conducted to determine the mode of water penetration into the wheat endosperm (Campbell & Jones, 1957; Seckinger *et al.*, 1964; Butcher & Stenvert, 1973; Stenvert & Kingswood, 1976; Moss, 1977). According to these authors the water penetrates from the back (dorsal side) and the shoulders of the grain, with no entry near the crease. Moss (1977) showed that the bran and aleurone layer were saturated within an hour, followed by slower, irregular diffusion through the endosperm. The rate of moisture diffusion into the wheat kernel has also

been studied using an autoradiographic technique (Butcher & Stenvert, 1973; Stenvert & Kingswood, 1976; Hosney, 1994). Butcher and Stenvert (1973) showed that the rate of penetration differed between cultivars, but the mode of penetration remained the same; the study of Stenvert & Kingswood (1976) agreed with this. These studies were destructive and therefore the rate of diffusion could not be followed within the same kernel.

The AACC International Approved Methods of Analysis for hardness determination are destructive. There is, however, a method utilising NIR spectroscopy to determine hardness on bulk whole wheat samples (Williams, 1991), providing an average measurement of the entire sample scanned. It would be preferable in breeding programmes to determine kernel hardness on individual kernels using a non-destructive method. Single kernels with desired wheat hardness, and other wanted quality traits, could then be used further in the breeding trials.

The rate of diffusion of conditioning water is an important aspect. Currently fixed conditioning times are applied in the flour mill. Within breeding programmes kernel hardness is measured to determine conditioning requirements. Wheat that is in the process of conditioning takes up space in the flour mill, and a shorter period of conditioning time would thus have an enormous economic benefit for the milling industry. Determination of optimum conditioning time, taking kernel hardness in consideration, could thus be beneficial to the milling industry. To effectively determine the conditioning requirements for wheat of different hardness categories, it would be ideal to follow the rate of conditioning in a single kernel.

Analytical methods that employ the NIR spectroscopic region have the advantages of being rapid, non-destructive, and non-invasive. NIR spectroscopy also demands minimum sample preparation and can be applied to a range of samples with different shapes and textures; which makes it excellent for online use (Osborne *et al.*, 1993; Dahm & Dahm, 2001; Koehler IV *et al.*, 2002; Pasquini, 2003). NIR is not only useful for routine analysis, but has tremendous research potential providing unique information not accessible by other techniques (Siesler, 2002). NIR spectroscopy unfortunately only provides an average spectrum, obtained from point measurements of the bulk sample. The inability of bulk NIR spectroscopy to provide spatial information makes it undesirable for applications where location of a compound is important. NIR hyperspectral imaging poses a potential solution for the localisation of chemical compounds in a non-homogenous sample.

Near infrared (NIR) spectroscopy employs photon energy to collect information in the energy range of 750 to 2500 nm (Bokobza, 2002; Pasquini, 2003). Absorption bands of chemical constituents in a sample can be observed in the NIR region due to the consequence of molecular stretching and bending vibrations of O-H, C-H, N-H and S-H chemical bonds (Miller, 2001; Siesler, 2002). The sample measured is irradiated with NIR radiation; while radiation penetrates the product, the spectral characteristics of the incoming light change due to wavelength dependent scattering or absorption processes. These changes are dependent on the chemical composition of the sample as well as light scattering characteristics (Nicolai *et al.*, 2007). The anharmonicity of

chemical bonds such as O-H, C-H, N-H and S-H, causes overtones and combination bands (Osborne *et al.*, 1993; Siesler, 2002).

Initially bulk NIR spectroscopy has been used extensively for quantitative determinations in agricultural products (Williams *et al.*, 1982; Williams *et al.*, 1985; Williams & Sobering, 1993). Within cereals it has been applied to determine seed composition in maize (Eyherabide *et al.*, 1996; Baye *et al.*, 2006), wheat hardness (Downey *et al.*, 1986; Norris *et al.*, 1989; Osborne, 1991; Manley *et al.*, 2002) and change in carbohydrate and protein content of wheat during maturation (Gergely & Salgo, 2005; Gergely & Salgo, 2007). In recent years it has been used for more innovative applications such as the detection of faeces at slaughter plants (Liu *et al.*, 2006), detection of water change in sheep meat (McGlone *et al.*, 2005), determination of lipids in roasted coffee (Pizarro *et al.*, 2004), verification of adulteration in alcoholic beverages (Pontes *et al.*, 2006), monitoring of polymer extrusion processes (Rohe *et al.*, 1999), pharmaceutical applications (Quaresima *et al.*, 2003; Zhou *et al.*, 2003; Colón *et al.*, 2005; Blanco & Alcalá, 2006; Sakudo *et al.*, 2006), and a variety of other applications (Osborne, 2000).

Combining the physics of NIR spectroscopy and that of digital imaging results in an advanced analytical technique, i.e. NIR hyperspectral imaging, which allows both spatial and spectral information to be obtained from a sample (Koehler IV *et al.*, 2002; Reich, 2005; Roggo *et al.*, 2005; Burger, 2006; Burger & Geladi, 2006; Gowen *et al.*, 2007; Grahn & Geladi, 2007; Gowen *et al.*, 2008a). Hyperspectral images comprise hundreds of adjacent wavebands for each spatial position of a sample. Each pixel in a hyperspectral image therefore contains a full spectrum for that specific position in the sample. Hyperspectral images, commonly known as hypercubes, result in a three-dimensional block of data. The hypercube consist of two-dimensional images composed of pixels in the x, y direction and a wavelength dimension in the z direction (Koehler IV *et al.*, 2002; Reich, 2005; Burger, 2006; Burger & Geladi, 2006; Gowen *et al.*, 2007; Gowen *et al.*, 2008b; Shahin & Symons, 2008). NIR hyperspectral imaging not only has all of the advantages of bulk NIR spectroscopy, but also the added spatial dimensions and the possibility of parallel data collection (Koehler IV *et al.*, 2002). NIR hyperspectral imaging thus allows for the determination and localisation of chemical constituents in a sample.

Multivariate data analysis, using specific chemometrics techniques, is required to extract the relevant information buried in the data matrix resulting from NIR measurements. This involves extracting relevant information about the objects and variables, enabling reduction of the information into fewer compounds that could be more easily interpreted, and separation of a residual containing mainly noise (Geladi, 2003). This reduced number of terms will have increased stability, due to noise or less useful information being removed from the data (Geladi, 2003). Classification of hyperspectral data can be done through unsupervised classification techniques such as principal component analysis (PCA) if no information is available about the samples, or through supervised classification techniques such as partial least squares discriminant analysis (PLS-DA) when sufficient information about the samples is available.

NIR hyperspectral imaging has been applied to detect bruises on pickling cucumbers (Ariana *et al.*, 2006) and mushrooms (Gowen *et al.*, 2008a), inspection of poultry carcasses (Chao *et al.*, 2007), single kernel maize analysis (Cogdill *et al.*, 2004), detection of faecal contamination on apples (Kim *et al.*, 2005; Liu *et al.*, 2007), tenderness prediction of beef (Naganathan *et al.*, 2008) and estimation of the firmness of strawberries (Tallada *et al.*, 2006). Applications specifically in the wheat industry include the determination of vitreousness in durum wheat kernels (Gorretta *et al.*, 2006; Shahin & Symons, 2008), determination of wheat pre-germination (Smail *et al.*, 2006; Koç *et al.*, 2008; Singh *et al.*, 2009), differentiation of Canadian wheat classes (Mahesh *et al.*, 2008).

NIR hyperspectral imaging provides the opportunity to capture an image of multiple kernels, although analysis of single kernels can be performed. This is possible due to the enormous amount of information available in the spatial dimension of each kernel. Discrimination between single kernels differing in hardness could thus be possible using this technique. Similarly NIR hyperspectral imaging could potentially be applied in studying the diffusion of conditioning water. Due to its non-destructive nature NIR hyperspectral imaging would allow tracking the diffusion of conditioning water in the same wheat kernel at different time intervals during the conditioning period.

The aim of this study was therefore to evaluate the use of NIR hyperspectral imaging together with multivariate image analysis as a non-destructive method to a) distinguish between whole wheat kernels differing in wheat hardness; and b) track the diffusion of conditioning water into wheat samples of different hardness categories over a period of 36 hours.

## References

- AACC (2009a). AACC International Approved Methods of Analysis, 11th Ed. Method 39-70.02. Near-Infrared Reflectance Method for Hardness Determination in Wheat. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-39-70.02.
- AACC (2009b). AACC International Approved Methods of Analysis, 11th Ed. Method 55-30.01. Particle Size Index for Wheat Hardness. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-55-30.01.
- AACC (2009c). AACC International Approved Methods of Analysis, 11th Ed. Method 55-31.01. Single Kernel Characterisation System for Wheat Kernel Texture. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-55-31.01.
- Ariana, D.P., Lua, R. & Guyer, D.E. (2006). Near-infrared hyperspectral reflectance imaging for detection of bruises on pickling cucumbers. *Computers and Electronics in Agriculture*, **53**, 60-70.
- Bakhella, M., Hosene, R.C. & Lookhart, G.L. (1990). Hardness of Moroccan wheats. *Cereal Chemistry*, **67**, 246-250.
- Bass, E.J. (1988). Wheat flour milling. In: *Wheat Chemistry and Technology* (edited by Y. Pomeranz). Pp. 1-68. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Baye, T.M., Pearson, T.C. & Settles, M. (2006). Development of a calibration to predict maize seed composition using single kernel near infrared spectroscopy. *Journal of Cereal Science*, **43**, 236-243.

- Bietz, J.A. (1989). Telling differences among wheat proteins may make a difference in marketing. In: *Wheat is Unique. Structure, Composition, Processing, End-use Properties, and Products* (edited by Y. Pomeranz). Pp. 303-315. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Blanco, M. & Alcalá, M. (2006). Simultaneous quantitation of five active principles in a pharmaceutical preparation: development and validation of a near infrared spectroscopic method. *European Journal of Pharmaceutical Sciences*, **27**, 280-286.
- Bokobza, L. (2002). Origin of near infrared absorption bands. In: *Near-Infrared Spectroscopy: Principles, Instruments, Applications* (edited by H.W. Siesler, Y. Ozaki, S. Kawata & H.M. Heise). Weinheim, Germany: Wiley-VCH Verlag GmbH.
- Burger, J. (2006). Hyperspectral NIR image analysis: data exploration, correction and regression. PhD Thesis Unit of Biomass Technology and Chemistry, Swedish University of Agricultural Sciences.
- Burger, J. & Geladi, P. (2006). Hyperspectral NIR imaging for calibration and prediction: a comparison between image and spectrometer data for studying organic and biological samples. *Analyst*, **131**, 1152-1160.
- Butcher, J. & Stenvert, N.L. (1973). Conditioning studies on Australian wheat; III. The role of the rate of water penetration into wheat grain. *Journal of the Science of Food and Agriculture*, **24**, 1077-1084.
- Campbell, J.D. & Jones, C.R. (1957). The rates of penetration of moisture to different points in the central cross-section of the endosperm in damped Manitoba wheat grains. *Cereal Chemistry*, **34**, 110-116.
- Chao, K., Yang, C.C., Chen, Y.R., Kim, M.S. & Chan, D.E. (2007). Hyperspectral-multispectral line-scan imaging system for automated poultry carcass inspection applications for food safety. *Poultry Science*, **86**, 2450-2460.
- Cobb, N.A. (1986). The hardness of the grain in the principal varieties of wheat. *Agricultural Gazette New South Wales*, **7**, 279-298.
- Cogdill, R.P., Hurburgh, C.R. & Rippke, G.R. (2004). Single-kernel maize analysis by near-infrared hyperspectral imaging. *Transactions of the ASAE*, **47**, 311-320.
- Colón, J., Peroza, C., Caraballo, W., Conde, C., Li, T., Morris, K.R. & Románach, R.J. (2005). On line non-destructive determination of drug content in moving tablets using near infrared spectroscopy. *Journal of Process Analytical Technology*, **2**, 8-15.
- Dahm, D.J. & Dahm, K.D. (2001). The physics of near infrared scattering. In: *Near-Infrared Technology in the Agricultural and Food Industries* (edited by P. Williams & K. Norris). Pp. 1-18. St. Paul, Minnesota, USA: American Association of Cereal Chemists.
- Dexter, J.E. & Sarkar, A.K. (2004). Wheat dry milling. In: *Encyclopedia of Grain Science* (edited by C. Wrigley, H. Corke & C.E. Walker). **Vol. 3**. Pp. 363-374. Kidlington, Oxford, UK: Elsevier Ltd.
- Downey, G., Byrne, S. & Dwyer, E. (1986). Wheat trading in the republic of Ireland: the utility of a hardness index derived by near infrared reflectance spectroscopy. *Journal of the Science of Food and Agriculture*, **37**, 762-766.
- Eyherabide, G., Robutti, J. & Borrás, F. (1996). Effect of near infrared transmission-based selection on maize hardness and the composition of zeins. *Cereal Chemistry*, **73**, 775-778.
- Geladi, P. (2003). Chemometrics in spectroscopy. Part 1. Classical chemometrics. *Spectrochimica Acta Part B*, **58**, 767-782.
- Gergely, S. & Salgo, A. (2005). Changes in carbohydrate content during wheat maturation - what is measured by near infrared spectroscopy? *Journal of Near Infrared Spectroscopy*, **13**, 9-17.

- Gergely, S. & Salgo, A. (2007). Changes in protein content during wheat maturation - what is measured by near infrared spectroscopy? *Journal of Near Infrared Spectroscopy*, **15**, 49-58.
- Gorretta, N., Roger, J.M., Aubert, M., Bellon-Maurel, V., Campan, F. & Roumet, P. (2006). Determining vitreousness of durum wheat kernels using near infrared hyperspectral imaging. *Journal of Near Infrared Spectroscopy*, **14**, 231-239.
- Gowen, A.A., O'Donnell, C.P., Cullen, P.J., Downey, G. & Frias, J.M. (2007). Hyperspectral imaging - an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology*, **18**, 590-598.
- Gowen, A.A., O'Donnell, C.P., Taghizadeh, M., Cullen, P.J., Frias, J.M. & Downey, G. (2008a). Hyperspectral imaging combined with principal component analysis for bruise damage detection on white mushrooms (*Agaricus bisporus*). *Journal of Chemometrics*, **22**, 259-267.
- Gowen, A.A., O'Donnell, C.P., Cullen, P.J. & Bell, S.E.J. (2008b). Recent applications of chemical imaging to pharmaceutical process monitoring and quality control. *European Journal of Pharmaceutics and Biopharmaceutics*, **69**, 10-22.
- Grahn, H.F. & Geladi, P. (2007). *Techniques and Applications of Hyperspectral Image Analysis*. Pp. 1-15, 313-334. Chichester, England: John Wiley & Sons Ltd.
- Greenwell, P. & Schofield, J.D. (1986). A starch granule protein associated with endosperm softness in wheat. *Cereal Chemistry*, **63**, 379-380.
- Hoseney, R.C. (1994). Dry milling of cereals. In: *Principles of Cereal Science and Technology* (edited by R.C. Hoseney). Pp. 130-132. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Irvine, G.N. & Anderson, J.A. (1952). A note on the determination on the brightness in the flour. *Transactions of the American Association Cereal Chemistry*, **10**, 59-60.
- Kent, N.L. & Evers, A.D. (1994). *Kent's Technology of Cereals*. Oxford, UK: Pergamon.
- Kim, M.S., Lefcourt, A.M., Chen, Y.-R. & Tao, Y. (2005). Automated detection of fecal contamination of apples based on multispectral fluorescence image fusion. *Journal of Food Engineering*, **71**, 85-91.
- Koç, H., Smail, V.W. & Wentzel, D.L. (2008). Reliability of InGaAs focal plane array imaging of wheat germination at early stages. *Journal of Cereal Science*, **48**, 394-400.
- Koehler IV, F.W., Lee, E., Kidder, L.H. & Lewis, N.E. (2002). Near infrared spectroscopy: the practical chemical imaging solution. *Spectroscopy Europe*, **14**, 12-19.
- Kweon, M., Martin, R. & Souza, E. (2009). Effect of tempering conditions on milling performance and flour functionality. *Cereal Chemistry*, **86**, 12-17.
- Liu, Y., Chao, K., Chen, Y.R., Kim, M.S., Nou, X., Chan, D.E. & Yang, C. (2006). Comparison of visible and near infrared reflectance spectroscopy for the detection of faeces/ingesta contaminants for sanitation verification at slaughter plants. *Journal of Near Infrared Spectroscopy*, **14**, 325-331.
- Liu, Y., Chen, Y.R., Kim, M.S., Chan, D.E. & Lefcourt, A.M. (2007). Development of simple algorithms for the detection of fecal contaminants on apples from visible/near infrared hyperspectral reflectance imaging. *Journal of Food Engineering*, **81**, 412-418.
- Mahesh, S., Manickavasagan, A., Jayas, D.S., Paliwal, J. & White, N.D.G. (2008). Feasibility of near-infrared hyperspectral imaging to differentiate Canadian wheat classes. *Biosystems Engineering*, **101**, 50-57.
- Manley, M., Van Zyl, L. & Osborne, B.G. (2002). Using Fourier transform near infrared spectroscopy in determining kernel hardness, protein and moisture content of whole wheat flour. *Journal of Near Infrared Spectroscopy*, **10**, 71-76.

- McGlone, V.A., Devine, C.E. & Wells, R.W. (2005). Detection of tenderness, post-rigor age and water status changes in sheep meat using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, **13**, 277-285.
- Miller, C.E. (2001). Chemical principles of near infrared technology. In: *Near Infrared Technology in the Agricultural and Food Industries* (edited by K. Norris & P.C. Williams). St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Moss, R. (1977). An autoradiographic technique for the location of conditioning water in wheat at cellular level. *Journal of the Science of Food and Agriculture*, **28**, 23-33.
- Naganathan, G.K., Grimes, L.M., Subbiah, J., Calkins, C.R., Samal, A. & Meyer, G.E. (2008). Visible/near-infrared hyperspectral imaging for beef tenderness prediction. *Computers and Electronics in Agriculture*, **64**, 225-233.
- Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I. & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Postharvest Biology and Technology*, **46**, 99-118.
- Norris, K.H., Hruschka, W.R., Bean, M.M. & Slaughter, D.C. (1989). A definition of wheat hardness using near infrared reflectance spectroscopy. *Cereal Foods World*, **34**, 696-705.
- O'Brien, L. & DePauw, R. (2004). Wheat breeding. In: *Encyclopedia of Grain Science* (edited by C. Wrigley, H. Corke & C.E. Walker). **Vol. 3**. Pp. 330-336. Kidlington, Oxford, UK: Elsevier Ltd.
- Osborne, B. (1991). Measurement of the hardness of wheat endosperm by near infrared spectroscopy. *Postharvest News and Information*, **2**, 331-334.
- Osborne, B.G. (2000). Near-infrared spectroscopy in food analysis. In: *Encyclopedia of Analytical Chemistry* (edited by R.A. Meyers). Pp. 4069-4081. Chichester: John Wiley & Sons.
- Osborne, B.G., Fearn, T. & Hindle, P.H. (1993). *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*. Harlow, England: Longman Scientific & Technical.
- Pasquini, C. (2003). Near infrared spectroscopy: fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, **14**, 198-219.
- Pizarro, C., Esteban-Díez, I., Nistal, A.-J. & González-Sáiz, J.-M. (2004). Influence of data pre-processing on the quantitative determination of the ash content and lipids in roasted coffee by near infrared spectroscopy. *Analytica Chimica Acta*, **509**, 217-227.
- Pomeranz, Y. & Williams, P.C. (1990). Wheat hardness: its genetic, structural, and biochemical background, measurement, and significance. In: *Advances in Cereal Science and Technology* (edited by Y. Pomeranz). Pp. 471-529. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Pontes, M.J.C., Santos, S.R.B., Araújo, M.C.U., Almeida, L.F., Lima, R.A.C., Gaiaño, E.N. & Souto, U.T.C.P. (2006). Classification of distilled alcoholic beverages and verification of adulteration by near infrared spectrometry. *Food Research International*, **39**, 182-189.
- Prabhasankar, P., Sudha, M.L. & Haridas Rao, P. (2000). Quality characteristics of wheat flour milled streams. *Food Research International*, **33**, 381-386.
- Quaresima, V., Lepantor, R. & Ferrari, M. (2003). The use of near infrared spectroscopy in sports medicine. *Journal of Sports Medicine and Physical Fitness*, **43**, 1-13.
- Reich, G. (2005). Near-infrared spectroscopy and imaging: basic principles and pharmaceutical applications. *Advanced Drug Delivery Reviews*, **57**, 1109-1143.



- Roggo, Y., Edmond, A., Chalus, P. & Ulmschneider, M. (2005). Infrared hyperspectral imaging for qualitative analysis of pharmaceutical solid forms. *Analytica Chimica Acta*, **535**, 79-87.
- Rohe, T., Becker, W., Kölle, S., Eisenreich, N. & Eyerer, P. (1999). Near infrared (NIR) spectroscopy for in-line monitoring of polymer extrusion processes. *Talanta*, **50**, 283-290.
- Sakudo, A., Suganuma, Y., Kobayashi, T., Onodera, T. & Ikuta, K. (2006). Near-infrared spectroscopy: promising diagnostic tool for viral infections. *Biochemical and Biophysical Research Communications*, **341**, 279-284.
- Schofield, J.D. & Greenwell, P. (1987). Wheat starch granule proteins and their technological significance. In: European Conference on Food Science and Technology (edited by I.D. Morton). Pp. 407-420. Bournemouth, England: Ellis Horwood.
- Seckinger, H.L., Wolf, M.J. & Dimler, R.J. (1964). Micro method for determining moisture distribution in wheat kernels, based on iodine staining. *Cereal Chemistry*, **41**, 80-87.
- Shahin, M.A. & Symons, S.J. (2008). Detection of hard vitreous and starchy kernels in amber durum wheat samples using hyperspectral imaging. *NIR News*, **19**, 16-18.
- Siesler, H.W. (2002). Introduction. In: *Near Infrared Spectroscopy: Principles, Instruments, Applications* (edited by H.W. Siesler, Y. Ozaki, S. Kawata & H.M. Heise). Pp. 1-10. Weinheim, Germany: Wiley-VCH Verlag GmbH.
- Simmonds, D.H. (1974). Chemical basis of hardness and vitreosity in the wheat kernel. *Baker's Digest*, **48**, 16-29, 63.
- Simmonds, D.H., Barlow, K.K. & Wrigley, C.W. (1973). The biochemical basis of grain hardness in wheat. *Cereal Chemistry*, **50**, 553-562.
- Singh, C.B., Jayas, D.S., Paliwal, J. & White, N.D.G. (2009). Detection of sprouted and midge-damaged wheat kernels using near infrared hyperspectral imaging. *Cereal Chemistry*, **86**, 256-260.
- Sissons, M., Osborne, B. & Sissons, S. (2006). Application of near infrared reflectance spectroscopy to a durum wheat breeding programme. *Journal of Near Infrared Spectroscopy*, **14**, 17-25.
- Smail, V.W., Fritz, A.K. & Wetzels, D.L. (2006). Chemical imaging of intact seeds with NIR focal plane array assists plant breeding. *Vibrational Spectroscopy*, **42**, 215-221.
- Smith, L. (1956). Objects of conditioning. In: *Flour Milling Technology* (edited by L. Smith). Pp. 115. Liverpool, England: Northern Publishing Co. Ltd.
- Stenvert, N.L. & Kingswood, K. (1976). An autoradiographic demonstration of the penetration of water into wheat during tempering. *Cereal Chemistry*, **53**, 141-149.
- Tallada, J.G., Nagata, M. & Kobayashi, T. (2006). Non-destructive estimation of firmness of strawberries (*Fragaria x ananassa* Duch.) using NIR hyperspectral imaging. *Environment Control in Biology*, **44**, 245-255.
- Turnbull, K.M. & Rahman, S. (2002). Endosperm texture in wheat. *Journal of Cereal Science*, **36**, 327-337.
- Williams, P.C. (1991). Prediction of wheat kernel texture in whole grains by near-infrared transmittance. *Cereal Chemistry*, **68**, 112-114.
- Williams, P.C., Norris, K.H. & Sobering, D.C. (1985). Determination of protein and moisture in wheat and barley by near-infrared transmission. *Journal of Agricultural and Food Chemistry*, **33**, 239-244.
- Williams, P.C., Norris, K.H. & Zarowski, W.S. (1982). Influence of temperature on estimation of protein and moisture in wheat by near-infrared reflectance. *Cereal Chemistry*, **59**, 473-477.

- Williams, P.C. & Sobering, D.C. (1993). Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds. *Journal of Near Infrared Spectroscopy*, **1**, 25-32.
- Zhou, G.X., Ge, Z., Dorwart, J., Izzo, B., Kukura, J., Bicker, G. & Wyvratt, J. (2003). Determination and differentiation of surface and bound water in drug substances by near infrared spectroscopy. *Journal of Pharmaceutical Sciences*, **92**, 1058-1065.

## **CHAPTER 2**

### **Literature review**

## Table of contents

<b>1. Introduction</b>	14
<b>2. Wheat</b>	14
2.1 <i>Principle wheat species</i>	14
2.2 <i>Wheat kernel morphology</i>	15
2.2.1 External	15
2.2.2 Internal	16
<u>2.2.2.1 Bran</u>	16
<u>2.2.2.2 Endosperm</u>	17
<u>2.2.2.3 Germ</u>	18
2.3 <i>Wheat hardness</i>	18
2.3.1 Definition and significance	18
2.3.2 Hardness determination methods	20
<u>2.3.2.1 Particle size index</u>	20
<u>2.3.2.2 Near infrared spectroscopy</u>	21
<u>2.3.2.3 Single kernel characterisation system</u>	21
<u>2.3.2.4 Farinator or vitreousness cutter</u>	22
2.4 <i>Wheat Conditioning</i>	23
2.4.1 Definition and significance	23
2.4.2 Factors influencing conditioning requirements	25
<u>2.4.2.1 Wheat hardness</u>	25
<u>2.4.2.2 Initial moisture content</u>	25
<u>2.4.2.3 Temperature</u>	25
<b>3. Near infrared analysis</b>	25
3.1 <i>Near infrared spectroscopy</i>	25
3.1.1 Development and principles	25
3.1.2 Applications	26
3.1.3 Advantages and limitations	27
3.2 <i>Near infrared hyperspectral imaging</i>	27
3.2.1 Principles and applications	27
3.2.2 Image acquisition and instrument configurations	29
<u>3.2.2.1 Point scan imaging configuration</u>	30
<u>3.2.2.2 Focal plane scan imaging configuration</u>	30
<u>3.2.2.3 Line scan imaging configuration</u>	30
<b>4. Chemometrics and multivariate image analysis</b>	32
4.1 <i>Pretreatment</i>	32
4.1.1 Derivatives	33

4.1.2 Standard normal variate	33
4.1.3 Multiplicative scatter correction	33
4.2 <i>Principal component analysis</i>	34
4.3 <i>Partial least squares discriminant analysis</i>	36
4.4 <i>Multivariate image analysis</i>	37
<b>5. Conclusion</b>	38
<b>6. References</b>	39

# Literature review

## 1. Introduction

Cereals are the fruits of cultivated grasses, which are all included in the monocotyledonous family Gramineae (Kent & Evers, 1994). Cereals include wheat, rice, maize, barley, sorghum, oats and rye, as well as oilseeds and pulses (Beta, 2004). Wheat, maize and rice are the leading grains in terms of production and planting area occupied (Orth & Shellenberger, 1988; Beta, 2004). It is widely accepted that wheat has been grown as a food crop as early as 10000-8000 B.C. (Orth & Shellenberger, 1988). Wheat cultivation started particularly in Iran, Egypt, Greece and Europe (Kent & Evers, 1994). Bread wheat originated due to the hybridisation of a wild species grass together with an emmer type of grass (Percival, 1921). Bread wheat occupies 93% of the world's wheat growing area, whilst the remainder is devoted to soft and durum wheat (Beta, 2004).

Wheat is one of the world's most popular crops, as it provides more nutrition to humans than any other grain species. Not only is it used as food for humans, but also as livestock feed, composting material for mushrooms, making bricks, and producing starch for many industrial uses (Paulsen & Shroyer, 2004). Only a small proportion (ca. 12.5%) of wheat in Africa are being used as animal feed (Taylor, 2004).

In this literature study wheat morphology and quality characteristics with specific reference to wheat hardness and conditioning requirements will be reviewed. In addition near infrared (NIR) hyperspectral imaging with reference to imaging instruments, image acquisition and applications in food and agriculture will be reviewed. Multivariate data analysis will be discussed in terms of chemometrics techniques such as pretreatment, principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA).

## 2. Wheat

### 2.1 Principle wheat species

Wheat comprises a group of species that belong to the grass family (Gramineae) (Paulsen & Shroyer, 2004). The principle species of wheat, their genetic type and commercial uses are listed in **Table 2.1**. Common wheat or bread wheat (*Triticum aestivum*) are allohexaploids with 21 homologous pairs of chromosomes from three similar genomes (AABBDD). Bread wheat resulted from the natural hybridisation between *T. dicoccoides* (AABB) and *Aegilops squarrosa* (*T. tauchii*) (DD) (Orth & Shellenberger, 1988; O'Brien & DePauw, 2004). The D genome, donated to the hexaploid bread wheat by *Ae. squarrosa*, are directly implicated in the protein components responsible for the bread baking quality (Orth & Shellenberger, 1988). *Ae. squarrosa*, known to be well adapted to a wide range of environments, contributed to bread wheat being highly adaptable to many growth environments (Orth & Shellenberger, 1988). Durum wheat (*T. turgidum*) is an allotetraploid (28 chromosomes) that consist of two genomes (AABB). It resulted from the natural

hybridisation of two diploid grasses, *T. monococcum* (AA) and *Ae. speltoides* (BB) (O'Brien & DePauw, 2004).

**Table 2.1** The principle wheat species, genetic type and commercial uses (Percival, 1921; Kent & Evers, 1994; Paulsen & Shroyer, 2004)

Common name	Species name	Genetic type	Commercial uses
Bread wheat	<i>Triticum aestivum</i> L.	Hexaploid	Raised bread, buns, cakes, pastries
Durum wheat	<i>T. turgidum</i> L. var. <i>durum</i>	Tetraploid	Pasta, cous-cous, bulgur
Club wheat	<i>T. aestivum</i> L. var. <i>compactum</i>	Hexaploid	Raised bread, buns, cakes, pastries
Einkorn	<i>T. monococcum</i> L.	Diploid	Pearled for use in soups
Emmer	<i>T. turgidum</i> L. var. <i>dicoccum</i>	Tetraploid	Bread and porridge
Spelt	<i>T. aestivum</i> L. var. <i>spelt</i>	Hexaploid	Bread, pilaf, hot cereals

Different wheat species or types are used as primary ingredients for specific products (Mahesh *et al.*, 2008). Bread wheat (*T. aestivum*) and durum wheat (*T. turgidum*) are characterised by different chemical and physical properties (Kent & Evers, 1994; O'Brien & DePauw, 2004). Based on these different properties, wheat will differ in functional quality, nutritional contribution and consequently commercial value (Bietz, 1989). Durum wheat is grown in areas with considerable environmental stress, and nearly all have a spring growth habit. Durum wheat is extremely hard, has high protein content, and is commonly used for pasta, cous-cous and bulgur production (Paulsen & Shroyer, 2004).

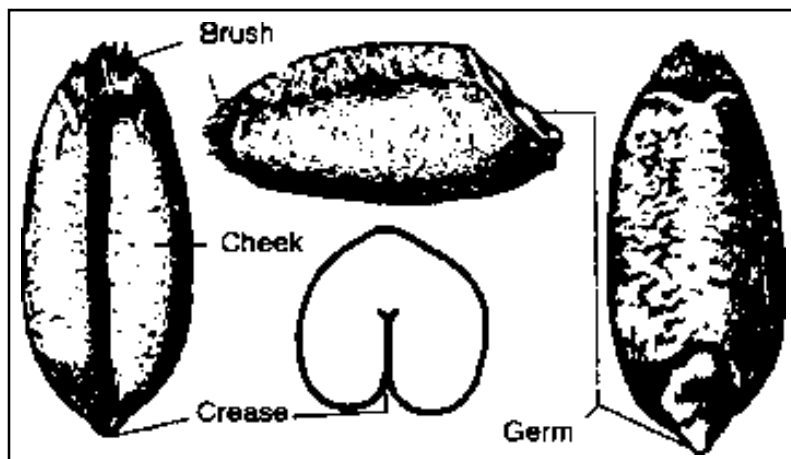
Bread wheat species include different classes with hard and soft endosperm, spring or winter growth habit and red or white pericarp appearance (Orth & Shellenberger, 1988; Paulsen & Shroyer, 2004). Different wheat classes could also differ in physical and chemical properties, which will ultimately affect the end product quality (Mahesh *et al.*, 2008). Wheat classes could be further distinguished by means of hectolitre mass, which measures bulk density and soundness of the grain, cleanliness, level of screenings, protein and moisture content and rheological properties (Orth & Shellenberger, 1988; O'Brien & DePauw, 2004). Some markets have more specific criteria relating to dough properties and the end use quality of wheat (O'Brien & DePauw, 2004). These requirements should all be considered in wheat breeding programmes (O'Brien & DePauw, 2004).

## 2.2 Wheat kernel morphology

### 2.2.1 External

The wheat kernel appears oval, elliptical, elongated or truncated from above (**Fig. 2.1**) depending on the cultivar (Halverson & Zeleny, 1988; Grundas & Wrigley, 2004). The wheat kernel has an average weight of 30-40 mg, with dimensions of 2.5-3.0 mm (thickness) by 3.0-3.5 mm (width) by 6.0-7.0 mm (length). The dorsal side of the kernel is rounded, with a characteristic longitudinal

crease on the ventral side (**Fig. 2.1**) (Halverson & Zeleny, 1988; Grundas & Wrigley, 2004). The embryo is situated on the dorsal side of the kernel, while the opposite end is covered with small hairs known as the “brush” (**Fig. 2.1**) (Halverson & Zeleny, 1988; Grundas & Wrigley, 2004). Characteristics such as these might be helpful to identify different wheat cultivars.



**Figure 2.1** Diagram of the external features of the wheat kernel (Kirby, 2002).

## 2.2.2 Internal

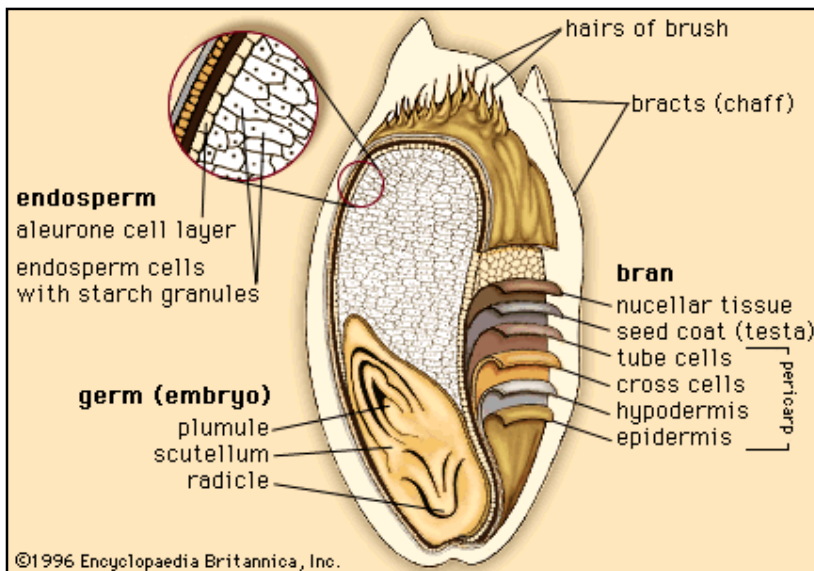
The wheat kernel (**Fig. 2.2**) consists of the bran (nucellar tissue, seed coat, and pericarp), endosperm (aleurone layer and endosperm cells) and germ (plumule, scutellum and radicle).

### 2.2.2.1 Bran

Nucellar tissue, the seed coat and the pericarp layers are collectively referred to as the bran. The pericarp, which constitutes several layers, surrounds the entire wheat kernel. The outer pericarp consists of the epidermis (beeswing) and hypodermis. The epidermis is a complete layer without intracellular spaces that completely covers the wheat kernel, except where the kernel were attached to the rachilla (Evers & Bechtel, 1988). The inner pericarp consists of cross cells and tube cells. The cross cells are tightly packed with little intracellular space. Their long axis is perpendicular to the long axis of the kernel, while tube cells, with many intracellular spaces, lie parallel to the long axis. The pericarp comprises 5% of the kernel, and consists of protein, minerals, cellulose and fat (Hoseney, 1994).

The seed coat and nucellar epidermis are situated between the pericarp and endosperm. The seed coat consists of cellulose tissue and contains the pigments responsible for the colour of the wheat, e.g. red (Hoseney, 1994).





**Figure 2.2** Diagram of the internal structures of a wheat kernel (Anonymous, 1996).

### 2.2.2.2 Endosperm

Botanically the aleurone layer is the outermost layer of the endosperm, and collectively covers the endosperm and germ. The aleurone layer is, however, removed during milling together with the bran. This layer is high in protein, minerals, phosphorous, fat, thiamine and riboflavin. In the region of the germ, the aleurone cells become thin walled, i.e. one third of the thickness found elsewhere (Hoseney, 1994).

The wheat endosperm comprises peripheral, prismatic and central cells, constituting hemicelluloses and  $\beta$ -glucan (Hoseney, 1994). Peripheral cells are located adjacent to the aleurone cells, followed by prismatic cells, which are located towards the centre of the cheeks or in the middle of the dorsal side; and the central cells in the middle of the cheeks (Evers & Bechtel, 1988). The thickness of these cell walls differs between hard and soft wheat cultivars, with hard cultivars having thicker cell walls than soft cultivars.

Inside the endosperm cells are a network of starch granules embedded in a protein matrix. The starch and protein contents differ between the different cell types. Peripheral cells have the lowest starch content, and thus the highest protein content. The starch content of the cells increases as they are situated more towards the centre of the grain (Evers & Bechtel, 1988). In low protein wheat the outer endosperm cells are filled with more starch granules, while in high protein wheat these cells contain the largest proportion of protein found in the kernel (Simmonds, 1974).

The starch granules occur as large lens-shaped granules (40  $\mu\text{m}$  lengthwise) and small spherical granules (2-8  $\mu\text{m}$  in diameter). The protein matrix consists mostly of gliadins and glutenins, which are responsible for the ability of wheat flour to be baked into leavened breads (Shewry *et al.*, 1987). Protein is deposited as protein bodies in the endosperm; as the wheat matures they are pressed together into a clay-like matrix that surrounds the starch granules

(Hoseney, 1994). The endosperm cells also contain lipid bodies, with the sub-aleurone region having the highest and the central region the lowest content (Evers & Bechtel, 1988).

The embryo, aleurone layer, pericarp and testa are removed during milling, thus leaving the starchy endosperm as the main contributor to white flour.

### 2.2.2.3 Germ

The germ comprises 2.5-3.5% of the wheat kernel and consists of the embryonic axis and the scutellum, which functions as a storage organ. The germ is high in protein (25%), sugar (18%), oil (48%), and minerals (5%) and contains enzymes and vitamin E (Hoseney, 1994).

## 2.3 *Wheat hardness*

### 2.3.1 Definition and significance

Wheat kernel hardness is a physical characteristic of wheat that determines the conditioning requirements of the wheat, flour yield, particle size and shape and ultimately the end use properties of the flour (Pomeranz & Williams, 1990).

Wheat kernel hardness is described in different ways; e.g. as the resistance of kernels to deformation by an outside source (Greenaway, 1969; Turnbull & Rahman, 2002) or not being easily penetrated or separated into parts (Pomeranz & Williams, 1990). Kent and Evers (1994) referred to wheat hardness and softness as milling characteristics related to the way the wheat kernel breaks during milling, while Downey *et al.* (1986) referred to it as a genetically linked trait which appears to be related in a general way to bread baking quality.

Wheat endosperm hardness or texture is determined by one genetic factor, i.e. the *Hardness (Ha)* locus on the short arm of the chromosome 5D (Symes, 1965). The genetic basis for variation in endosperm texture was established as a single major gene whose allelic expressions is associated with mutations in the proteins puroindoline a (*Pina*) and puroindoline b (*Pinb*) (Giroux & Morris, 1997; Lillemo & Morris, 2000). Failure to express *Pina* or point mutations in *Pinb* presents a hard phenotype (Lillemo & Morris, 2000).

In an attempt to explain the chemical difference in endosperm texture, Simmonds *et al.* (1973) investigated the water soluble proteins associated with the starch granules. Although the starch granules are entirely surrounded by proteins, the water soluble proteins were shown to be in direct contact with the starch granules (Barlow *et al.*, 1973; Simmonds *et al.*, 1973). They concluded that these proteins play the role of a “cementing” material between starch granules and the storage protein, and that the amount and composition of this material might express the genetic control of wheat hardness (Barlow *et al.*, 1973; Simmonds *et al.*, 1973; Simmonds, 1974).

Greenwell and Schofield (1986; 1987) further investigated these water soluble proteins assumingly responsible for wheat hardness, and found that a low molecular weight protein (15000 Da) were always present in soft wheat, but absent from hard cultivars (Greenwell & Schofield, 1986; Schofield & Greenwell, 1987). Similar to the findings by Simmonds *et al.* (1973), they also

suggested that this protein was associated with the starch granule surface. However, in contrast to these earlier studies (Simmonds *et al.*, 1973; Simmonds, 1974) they established that this protein reduced the adhesion between the starch granules and the protein matrix (Schofield & Greenwell, 1987). Schofield and Greenwell (1987) named this protein friabilin. Friabilin consists of the two major polypeptides, *Pina* and *Pinb* (Oda, 1994). The amount of friabilin extracted from wheat starch was shown to correlate perfectly with different endosperm textures (Greenwell & Schofield, 1986; Schofield & Greenwell, 1987; Bakhella *et al.*, 1990). Soft wheat had a high content of friabilin, hard bread wheat less friabilin, while in durum wheat friabilin was completely absent. The endosperm cell content, with regard to starch granules embedded in the protein matrix, of hard wheat are thus firmly bonded together even at low protein levels. Soft wheat kernels separates easily even in soft wheats that appear vitreous (Pomeranz & Williams, 1990).

Traditionally vitreousness has been associated with hardness and high protein content of the kernel and opacity with softness and low protein content of the kernel. Vitreousness of the kernel is related to the compactness of the kernel, and the presence of air spaces between starch granules in the endosperm (Yamazaki & Donelson, 1983). It is now known that wheat hardness is caused by the genetically controlled strength of the protein-starch bond in the wheat endosperm (Hoseney, 1994), while vitreous wheat develops under conditions of high nitrogen availability and high temperature during the maturation phase of wheat (Pomeranz & Williams, 1990). Some wheats maintain their distinctive type of endosperm under different circumstances, while others can be influenced to yield vitreous or mealy wheat by altering external growth conditions (Percival, 1921).

Although wheat hardness is primarily under genetic control, it can also be influenced by environmental factors. The presence of friabilin and the strength of the starch-protein bond will, however, not be affected; only the vitreous or mealy appearance of the wheat kernel. In spite of friabilin controlling wheat hardness, soft wheat cultivars in general have higher starch and lower protein contents than hard cultivars (Hopkins & Graham, 1935).

Endosperm texture is of considerable importance to the miller, because hard wheat gives a coarse flour with a high level of starch damage while soft wheat gives a finer flour with lower levels of starch damage (Bolling, 1987). The point of fracture during milling is influenced by the hardness of the wheat. In hard wheat kernels the first point of fracture occurs at the endosperm cell walls rather than through the cell contents. This is due to the cell contents of hard wheat being much more firmly bound, thus resulting in a point of relative weakness between the cell walls (Simmonds, 1974; Hoseney, 1994). On further reduction of endosperm to flour sized particles the cell contents would rather fracture through the starch granule than at the starch-protein interface. Hard wheat always fractures in the same way irrespective of vitreousness. In soft wheat the adhesion between the starch granules and the protein matrix is weaker and fracture of the endosperm tends to occur through the cell contents, at the starch-protein interface

(Simmonds, 1974; Schofield & Greenwell, 1987). Flour from soft wheat thus contains more intact starch granules, and in general smaller particle sizes than hard wheat.

The method of fracturing of the hard wheat results in flour with larger particles, a higher percentage of starch damage, and higher water absorption than soft wheat flours (Schofield & Greenwell, 1987; Tippels *et al.*, 1994).

### 2.3.2 Hardness determination methods

The measurement of wheat hardness dates back to the work of Cobb in 1896 who measured the force to cut a wheat kernel in half with a pair of pinchers; simulating the biting of the front teeth. Measuring wheat hardness became a major factor in differentiating between hard and soft wheat classes during the mid-1980's (Cobb, 1986; Halverson & Zeleny, 1988). Crossing of wheat classes in breeding programmes made it more difficult to visually classify the different hard and soft wheat classes (Halverson & Zeleny, 1988).

The two main reasons for measuring wheat hardness are price and functionality, due to wheat of different hardness classes differing in functionality during further processing. Kernel texture affects the way in which the specific wheat must be conditioned for milling; the flour yield and particle size, shape and density of flour particles; and the end use properties in milling and baking (Pomeranz & Williams, 1990). Wheat hardness is therefore an extremely important characteristic for both the milling and baking industry (Pomeranz & Williams, 1990). Acceptance of a load as well as payment to the producer depends on various grading factors, in some countries this includes wheat hardness, and it is therefore important that this factor should be determined accurately and rapidly (Osborne, 1991).

Methods for determining wheat hardness are based on sieving, grinding resistance, vitreousness and near infrared (NIR) methods dependent on particle size (Osborne, 1991). These methods include the particle size index (PSI) (Osborne *et al.*, 2001; AACC, 2009b), NIR spectroscopy on whole kernels (Williams, 1991) and on flour (Downey *et al.*, 1986; AACC, 2009a), single kernel characterisation system (SKCS) (Worzella & Cutler, 1939; AACC, 2009c), and in the South African cereal industry the farinator or hardness cutter. These methods will be discussed briefly, in terms of their principles and the results obtained using them.

#### 2.3.2.1 Particle size index

The particle size index (PSI) test is based on sieving, and performed with the Alpine air-jet sieve according to the AACC method 55-30.01 (AACC, 2009b). The PSI value is related to wheat hardness in that hard wheat produce flour with a bigger particle size and thus a lower percentage throughs; resulting in a lower PSI value. The percentage of flour that has moved through the sieve (throughs) is weighed and the PSI of the sample is presented as the percentage throughs. The average PSI values for the different wheat hardness classes (extra soft to very hard) can be seen

in **Table 2.2**. Although the PSI method is very precise, it is not a rapid test which makes it unsuitable for use in an industrial environment (Williams & Sobering, 1986).

**Table 2.2** The average particle size index values of different hardness wheat determined with the Alpine air-jet sieve (AACC, 2009b)

Hardness category	PSI (%)
Extra hard	Up to 7
Very hard	8-12
Hard	13-16
Medium hard	17-20
Medium soft	21-25
Soft	26-30
Very soft	31-35
Extra soft	Above 35

#### 2.3.2.2 Near infrared spectroscopy

Near infrared (NIR) reflectance spectroscopy is sensitive to variation in particle size and particle size distribution. This makes it an excellent method in differentiating between wheats of different hardness; as wheats of different hardness yield flours with different mean particle sizes. The particle size of ground wheat increase with increase in wheat hardness, therefore hard wheat flour has a higher apparent absorption value than soft wheat flour. Change in particle size causes a change in the amount of NIR radiation scattered in the sample; causing a baseline shift in the absorbance spectra obtained. Larger particles absorb more of the radiation before it leaves the sample and thus has an absorption spectrum with higher values than smaller particles would have (Pomeranz & Williams, 1990; Hruschka, 2001). Absorption values at 1680 nm and 2230 nm were chosen as these wavelengths were shown to be sensitive to different particle size distributions (Gaines & Windham, 1998). These wavelengths are of importance in the calculation of the wheat hardness score using NIR spectroscopy. The NIR reflectance technique applied to ground wheat samples is now a recognised and reliable technique which is used as a standard AACC method for hardness determination (Norris *et al.*, 1989; Windham *et al.*, 1993).

It is also possible to predict the hardness of whole wheat kernels, although this method is not AACC approved. This research has been performed using NIR transmittance spectroscopy, whereafter partial least squares (PLS) regression were used to determine the prediction accuracy for hardness prediction (Williams, 1991).

#### 2.3.2.3 Single kernel characterisation system

The SKCS test is based on the force required to crush a single wheat kernel, and is performed according to the AACC Method 55-31.01 (AACC, 2009c). Results obtained are given in terms of the hardness index (HI), which is related to wheat hardness in that hard wheat require a greater

force to be crushed, than soft wheat, and thus would have a higher HI value. The average HI values for different hardness categories wheat (extra soft to extra hard) are displayed in **Table 2.3** (Gaines *et al.*, 1996; AACC, 2009c).

**Table 2.3** The average hardness index values of different hardness categories wheat determined with the SKCS (AACC, 2009c)

Hardness category	Hardness Index (HI) value
Extra hard	Above 90
Very hard	81-90
Hard	65-80
Medium hard	45-64
Medium soft	35-44
Soft	25-34
Very soft	10-24
Extra soft	Up to 10

#### 2.3.2.4 Farinator or vitreousness cutter

Vitreous grains appear dark and translucent, while opaque grains appear yellow and starchy (Simmonds, 1974; Korkut *et al.*, 2007). The percentage of vitreous kernels in a wheat sample can be determined by examining the cross-section of the kernels. The farinator or vitreousness cutter is a device designed to hold 50 wheat kernels firmly while a blade cuts them transversely. The percentage vitreousness of the sample is then determined by giving each kernel a value depending on their translucent or opaque appearance. Grains that are completely translucent receive two points, those completely opaque 0 points, and those that are both translucent and opaque receive one point. By adding these points the percentage vitreousness of the grain is then obtained. This value is used in South Africa to determine the conditioning requirements of wheat in the laboratory (Mr D September, Pioneer Food Group Ltd., South Africa, personal communication, 2008). As discussed previously, vitreousness and hardness are not the same property, and therefore the use of the farinator to determine the hardness of a wheat sample is not recommended.



**Figure 2.3** The Farinator device used to determine the vitreousness percentage of wheat (Simmonds, 1974).

## 2.4 Wheat Conditioning

### 2.4.1 Definition and significance

The dry milling process is concerned with transforming whole grains into forms suitable for consumption, or for the conversion into consumable goods. This involves size reduction of the grain and separation of endosperm cells or cell fragments from the bran of the wheat kernel (Hoseney, 1994; Kent & Evers, 1994). The terms conditioning and tempering are interchangeable. Tempering consists of adding water to grain whereafter the grain rests for a certain amount of time to allow the diffusion of water through the endosperm. Conditioning is the same as tempering, but heat is also used in conjunction with the water. In South Africa, the addition of water to wheat before milling is referred to as conditioning, although no heat treatment is used. For the purpose of this thesis the addition of water, without the use of heat, will be referred to as conditioning.

Conditioning is the process of preparing wheat for the dry milling process, facilitating the best separation of bran from the endosperm; and thus improving the baking quality of the flour produced (Smith, 1956; Bass, 1988; Hoseney, 1994; Dexter & Sarkar, 2004). During conditioning water is added to the wheat, which is then left to rest for a period of time before it is milled. Conditioning of wheat has five objectives, namely

- The bran coat is toughened by the addition of water, making the different layers of the bran “stick” together, enabling an easier and more effective removal of the bran (Smith, 1956; Bass, 1988; Hoseney, 1994). Toughening of the bran prevents powdering during the milling process;
- Facilitating the physical separation of endosperm from the bran during milling (Bass, 1988);
- Mellowing or softening of the endosperm makes it easier to grind, and be broken into flour sized particles (Smith, 1956; Hoseney, 1994). Moisture that penetrates the bran layers, reaches the cellulose cell walls of the endosperm cells first, and with time penetrates to the starchy endosperm. During the milling process these cells easily fall apart with less pressure than normally required in comparison with its hardness, producing more granular and livelier flour with less starch damage (Smith, 1956);
- Ensuring that all material leaving the grinding rollers are in optimum condition for sifting (Bass, 1988);
- To ensure that grinding produces the optimum level of damaged starch consistent with the hardness of the wheat, ensuring the optimum end use qualities for the specific wheat type and hardness (Bass, 1988).

Studies, reviewed by Bradbury *et al.* (1960), indicated that the mode of water penetration is firstly through the germ of the kernel, and later through the bran and the brush area. Further studies performed to investigate this phenomenon included iodine staining (Seckinger *et al.*, 1964), change in kernel density (Campbell & Jones, 1957) and an autoradiographic technique (Butcher & Stenvert, 1973; Stenvert & Kingswood, 1976; Moss, 1977). The iodine staining technique

included, staining with iodine followed by cutting of the kernel in sections, whereafter the mode of penetration was determined by microscopic investigation (Seckinger *et al.*, 1964). The variation in endosperm density, with change in moisture content in different regions of the wheat kernel over a conditioning period, was used to determine the mode and rate of moisture penetration (Campbell & Jones, 1957). According to these authors (Campbell & Jones, 1957; Seckinger *et al.*, 1964; Butcher & Stenvert, 1973; Stenvert & Kingswood, 1976; Moss, 1977) water penetrates from the back and shoulders of the grain, with no entry near the crease. The studies performed with the iodine staining technique and the technique reliant on endosperm density only concentrated on the diffusion of water in the endosperm, and did not relate to movement of water in the bran.

An autoradiographic technique was also used to determine the rate of water penetration into kernels of English and Australian wheat cultivars (Butcher & Stenvert, 1973; Stenvert & Kingswood, 1976; Moss, 1977). This involved conditioning of wheat with tritiated water ( $T_2O$  or tritium oxide) whereafter x-ray images of wheat slices were acquired. These authors achieved similar results indicating that the rate of penetration differs between different wheat cultivars, while the mode of penetration remains the same.

Stenvert and Kingswood (1976) also showed that water initially binds to the bran and enhanced entry of water in the germ region occurs, especially near the top of the germ region. The germ as a whole readily absorbs water due to the characteristics of the surrounding layers of the germ. The thinnest portion of the outer testa covers the germ area, and the nucellar layer is absent. A modified aleurone layer with thin walled cells also extends over the scutellum and part of the germ (Stenvert & Kingswood, 1976). A guideline for conditioning time and final moisture content of different hardness classes to be used in the milling industry can be seen in **Table 2.4**.

**Table 2.4** Recommended conditioning time and final moisture content for different hardness classes of wheat to be used in the milling industry (Wahrenberger, 2004)

Hardness category	Final moisture content (%)	Recommended conditioning time (hours)
Very hard and vitreous	Above 16 %	36 – 48
Hard and vitreous	16 %	22 – 36
Semi-hard	15,5 %	18 – 24
Semi-soft	14,5 – 15 %	12 – 18
Soft	14,5 %	6 – 12

Conditioned wheat needs to rest to allow the penetration and uniform distribution of moisture through the wheat endosperm (Dexter & Sarkar, 2004). Optimum resting time, allowing moisture distribution, depends on the wheat moisture content and hardness, but temperature also plays a role in the rate of water penetration (Dexter & Sarkar, 2004).



## 2.4.2 Factors influencing conditioning requirements

### 2.4.2.1 Wheat hardness

The amount of water added to wheat depends on the moisture content and hardness of the grain. The time required to reach an even distribution of moisture in the grain varies from six hours for soft, opaque kernels with low protein content, to over 24 hours for hard, vitreous, kernels with high protein content. Time given for water to penetrate the grain kernel is much shorter in soft than in hard wheat, and it appears that compact endosperm (vitreous) slow water uptake into the grain (Bass, 1988; Hosney, 1994). The hardness of the grain is thus a very important factor to consider when determining conditioning requirements.

### 2.4.2.2 Initial moisture content

Low initial moisture content results in a slower uptake of added conditioning water. The optimum conditioning moisture should be applied to the wheat, as too much moisture reduces flour yield obtained. Flour yield is reduced because the complete separation of bran from endosperm is reduced. The sieving efficiency is also influenced; soft wheat flour is more “sticky” and harder to sift than hard wheat, and this dilemma is increased by high moisture content above the optimum (Dexter & Sarkar, 2004).

### 2.4.2.3 Temperature

The penetration of water into the wheat kernel is essentially due to the diffusion process. The diffusion of the water is increased by a temperature increase of the grain during conditioning. Temperatures higher than 50°C should be avoided as gluten damage could occur above this temperature (Hosney, 1994).

## **3. Near infrared analysis**

### *3.1 Near infrared spectroscopy*

#### 3.1.1 Development and principles

The near infrared (NIR) region refers to the spectral region ranging from 750 to 2500 nm in the electromagnetic region of light (Miller, 2001). NIR spectroscopy is a vibrational spectroscopy that employs photon energy to collect information in the range of 750 to 2500 nm (Pasquini, 2003). NIR data are acquired in the form of transmission or reflectance counts due to the stretch and bending vibrations of O-H, C-H, N-H and S-H chemical bonds in the sample (Miller, 2001; Siesler, 2002). The sample measured is irradiated with NIR radiation, whereafter the radiation may be absorbed, reflected or transmitted. The penetration of radiation causes spectral characteristics of the radiation light to change due to wavelength dependent scattering or absorption processes. These changes are dependent on the chemical composition of the sample, as well as, light scattering properties (Nicolai *et al.*, 2007). The anharmonicity of bonds involving the hydrogen atom, causes overtones and combination bands of fundamental frequencies of O-H, C-H, N-H and

S-H chemical bonds (Osborne *et al.*, 1993; Siesler, 2002). The fact that the anharmonicity of these bonds involving hydrogen are measured makes NIR particularly useful for studying hydrogen bonding in foods.

The first NIR spectra were recorded in 1800 by Herschel while measuring heat energy of solar emission (Herschel, 1800). In 1881 the NIR spectra of several organic liquids in the wavelength range of 700 to 1200 nm were measured, leading to the realisation of the importance of hydrogen bonds in NIR absorption (Abney & Festing, 1881). Functional group peak assignments began with the work of Coblenz (1905) who measured compounds from 800 to 2800 nm and noticed bands characteristic of C-H chemical bonds (Coblenz, 1905). Further band assignments were made between 1922 and 1929 at the University of California, Los Angeles and John Hopkins University (Ellis, 1929); between 1930 and 1945 at the United States of America (USA) Bureau of Standards; and in the 1980's in the United Kingdom at Aberdeen University and Reading University, and in the USA at the United States Department of Agriculture (USDA) in Athens.

The first NIR spectrophotometer (Cary 14) became available in 1954, after development of the photoelectric detector for NIR, during research into communication in the Second World War (Osborne *et al.*, 1993). The application of NIR in the food and agricultural industry began with Karl Norris at the USDA with the first published application of NIR in 1962, the determination of moisture in methanol extracts of seeds (Hart *et al.*, 1962).

The NIR spectrum of a sample is essentially composed of a large set of overtones or combination bands, and due to the complexity of most agricultural samples, extremely intricate. In general NIR spectra of food constituents show broad bands, containing envelopes of overlapping absorptions (Osborne *et al.*, 1993). The spectrum could further be complicated by wavelength dependant scattering effects, instrument noise, temperature effects and sample heterogeneities (Nicolai *et al.*, 2007). These influences make it difficult to assign specific absorption bands to specific functional groups and components. Therefore chemometrics techniques are required to extract the relevant information buried in the NIR spectroscopic data.

Chemometrics techniques most commonly used for NIR data analysis are multiple linear regression (MLR), principal component regression (PCR), partial least squares regression (PLS-R), partial least squares discriminant analysis (PLS-DA), linear discriminant analysis (LDA), factor analysis (FA), cluster analysis (CA) and principal component analysis (PCA) (Esbensen, 2006).

### 3.1.2 Applications

Since the work of Karl Norris, bulk NIR spectroscopy has been applied in many different fields of research. NIR has been used in the research of a variety of foodstuffs (Williams *et al.*, 1982; Perez *et al.*, 2001; Peirs *et al.*, 2002; Chao *et al.*, 2003; Delwiche, 2003; McGlone *et al.*, 2005; Liu *et al.*, 2006; Pontes *et al.*, 2006; Sissons *et al.*, 2006; Khodabux *et al.*, 2007; Yan, 2007; Sinellia *et al.*, 2008; Cho *et al.*, 2009), animal feed industry (Valdes *et al.*, 1985; Pazourek & M., 1988;

Murray, 1996), medicine (Shibata *et al.*, 1999; Quaresima *et al.*, 2003; Litscher, 2005; Sakudo *et al.*, 2006; Spahn *et al.*, 2008), the polymer and textile industry (Rohe *et al.*, 1999; Cleve *et al.*, 2000; Scherzera *et al.*, 2005), pharmaceuticals (Blanco *et al.*, 1998; Colón *et al.*, 2005; Blanco & Alcalá, 2006) and the petrochemical industry (Davidson *et al.*, 1992; Macho & Larrechi, 2002).

### 3.1.3 Advantages and limitations

Analytical methods resulting from the use of the NIR region have the advantage of being rapid, non-destructive to the sample, non-invasive and with high penetration of the radiation beam. Methods employing NIR is also suitable for online use as it demands minimum sample preparation; and can be nearly universally applied due to its detection of any molecule containing C-H, N-H, O-H and S-H chemical bonds (Koehler IV *et al.*, 2002; Pasquini, 2003). NIR spectra of intact, opaque, biological samples with a wide range of different shapes and textures, can be obtained using diffuse reflectance, without any sample preparation or a special sample cell (Osborne *et al.*, 1993; Dahm & Dahm, 2001). NIR analysis has many advantages over chemical and other instrumental methods, and is therefore not only a routine analysis tool, but has a tremendous research potential providing unique information not accessible by other techniques (Siesler, 2002).

The major limitation of NIR in food analysis is the dependence on chemical methods of analysis for the development of a specific calibration (Osborne *et al.*, 1993). Bulk NIR spectrometers are not able to detect compositional differences within the sample. The inability of predicting distribution of compounds in the sample might lead to differences between the predicted value using calibrations, and the measured composition of the sample. This limitation makes bulk NIR spectroscopy undesirable for applications where the localisation of compounds within the sample is desired.

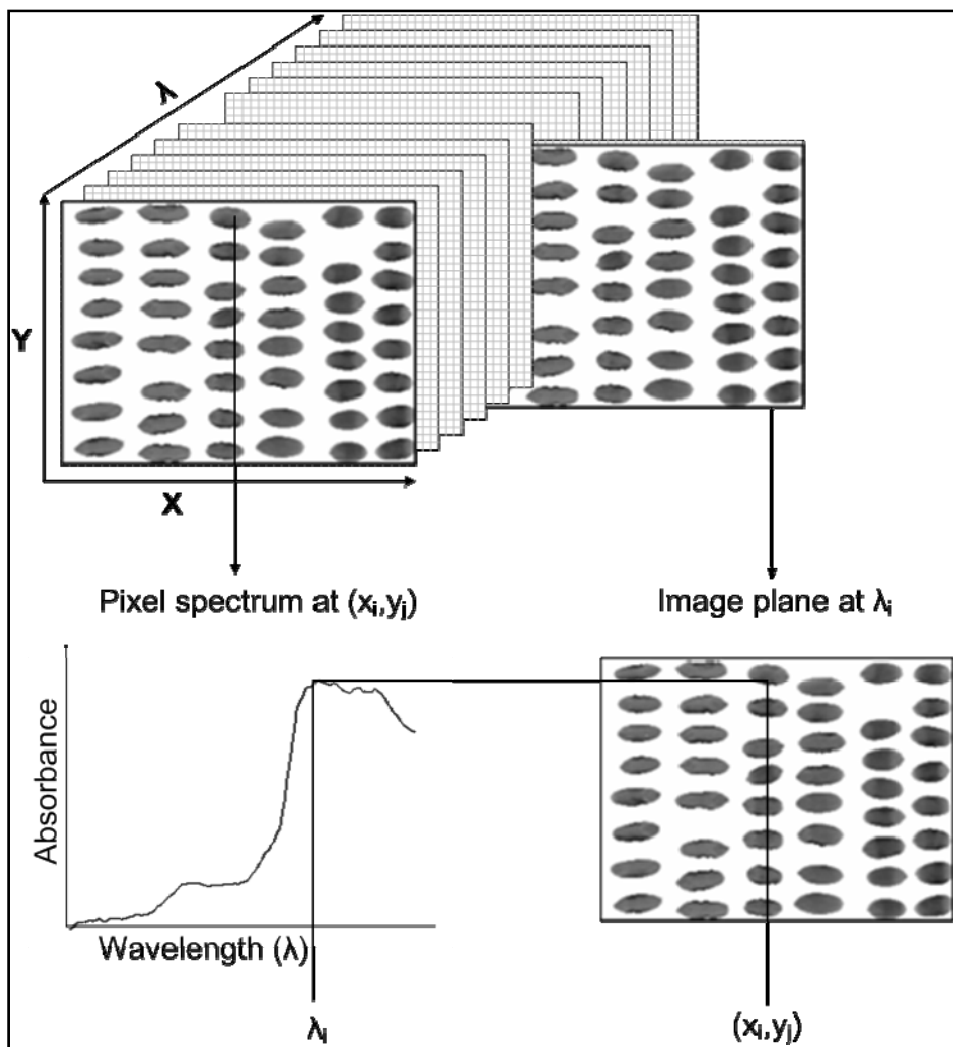
## 3.2 Near infrared hyperspectral imaging

### 3.2.1 Principles and applications

Hyperspectral imaging, chemical imaging or spectroscopic imaging is an advanced analytical technique that combines conventional digital imaging and the physics of NIR spectroscopy to obtain both spatial and spectral information of a sample (Koehler IV *et al.*, 2002; Reich, 2005; Roggo *et al.*, 2005; Burger, 2006; Burger & Geladi, 2006; Gowen *et al.*, 2007; Grahn & Geladi, 2007; Gowen *et al.*, 2008b). Hyperspectral imaging was originally developed for remote sensing purposes (Goetz *et al.*, 1985), and has in the last decade emerged as a powerful analytical tool in agriculture (Kazemi *et al.*, 2005; Fernández Pierna *et al.*, 2006; Gorretta *et al.*, 2006; Liu *et al.*, 2006; Weinstock *et al.*, 2006; Baeten *et al.*, 2007; Koç *et al.*, 2008; Mahesh *et al.*, 2008; Shahin & Symons, 2008), astronomy (Wood *et al.*, 2002; Hege *et al.*, 2003), pharmaceuticals (Hamilton & Lodder, 2002; Roggo *et al.*, 2005) and medicine (Vo-Dinh *et al.*, 2004; Chen *et al.*, 2005).

Hyperspectral images are composed of hundreds of adjacent wavebands for each spatial position of a sample (Koehler IV *et al.*, 2002; Cogdill *et al.*, 2004; Geladi *et al.*, 2004; Burger, 2006; Burger & Geladi, 2006; Gowen *et al.*, 2007). Each pixel in a hyperspectral image consists of a complete NIR spectrum for that specific position of the sample, and so does prove a fingerprint for that position.

Hyperspectral images are commonly known as hypercubes (Reich, 2005; Burger, 2006; Burger & Geladi, 2006; Gowen *et al.*, 2007; Shahin & Symons, 2008). Hypercubes are constructed of hundreds of single channel grayscale images, each representing a single band of spectral wavelength (Burger & Geladi, 2005; Burger, 2006; Burger & Geladi, 2006). The hypercube is a three-dimensional block of data, consisting of two-dimensional images composed of pixels in the x, y direction, and a wavelength dimension in the z direction (**Fig. 2.4**) (Koehler IV *et al.*, 2002; Reich, 2005; Gowen *et al.*, 2007; Shahin & Symons, 2008).



**Figure 2.4** Schematic representation of the hyperspectral imaging hypercube showing the relationship between spectral and spatial dimensions. Adapted from (Gowen *et al.*, 2008b).

A typical commercial instrument will produce hyperspectral images with 256 x 320 pixels, with 128 single channel images; or 81920 spectra each with 128 wavelength points (Burger & Geladi, 2006). However the amount of pixels and wavelengths may vary, depending on the type of camera used. A great amount of NIR spectra are thus captured in one NIR hyperspectral image, with both spectral and spatial information of the sample captured in the image (Koehler IV *et al.*, 2002). The spectral dimension provides information as to which chemical compounds, and what concentration, are found in the sample, while the spatial dimension provides information on the localisation of these chemical compound in the sample (Burger, 2006; Gowen *et al.*, 2007; Gowen *et al.*, 2008b).

NIR hyperspectral imaging not only has all of the advantages that NIR spectroscopy has, but also the added spatial dimensions and the possibility of parallel data collection (Koehler IV *et al.*, 2002). Chemical information of multiple constituents is obtained with a high sensitivity to low concentrations (Gowen *et al.*, 2007). Little or no sample preparation are required, and prediction can be performed from a single spectrum (Gowen *et al.*, 2008b). RGB colour imaging systems are already in use for grading and detection of surface defects (Daley *et al.*, 1993; Chao *et al.*, 1999; Luzuriaga & Balaban, 2002), but lack sensitivity to surface features detected by other wavebands than RGB. The need for online grading and quality control imaging systems thus exists that can detect wavebands other than RGB.

### 3.2.2 Image acquisition and instrument configurations

The NIR imaging system usually consists of the following general components: an illumination source, a camera, a spectrograph and a detector; which would all be coupled to a computer (Reich, 2005; Grahn & Geladi, 2007; Gowen *et al.*, 2008b). For macroscopic or microscopic images, a focusing lens or a microscope objective are also used. The most popular illumination sources are tungsten halogen lamps, or xenon gas plasma lamps. The spectrograph used most commonly contains a liquid crystal tuneable filter (LCTF), an acousto-optic tuneable filter (AOTF) or a prism-grating-prism filter (PGP). These filters are responsible for the selection of the wavelength bands to be measured. The camera unit used could contain an Indium Gallium Arsenide detector (InGaAs), a lead sulphide (PbS) detector or a HgCdTe (mercury-cadmium-telluride) detector (Geladi *et al.*, 2007).

It is currently not possible to obtain data in all three dimensions of a hypercube simultaneously, only two dimensions can be obtained at a time. The three dimensional image is constructed by stacking all these two dimensional images in a sequence (Gowen *et al.*, 2008b). The different configurations existing for acquiring a hyperspectral image are the point scan, focal plane scan and the line scan imaging configuration.

### 3.2.2.1 Point scan imaging configuration

The point scan configuration (**Fig. 2.5a**) is used to measure a complete spectrum on a small spot of the sample, whereafter the sample is then repositioned before obtaining a new spectrum (Burger, 2006; Gowen *et al.*, 2007; Grahn & Geladi, 2007). A complete hyperspectral image can be obtained by systematically moving the sample within the two spatial dimensions. Image calibration for wavelength and reflectance only needs to be performed once before starting to acquire the image. The spectra obtained are very stable with high resolution, but the acquisition process is very time consuming and places a high demand on the repositioning hardware to attain reproducible results. The spatial size dimensions of the image obtained is limited only by the sample positioning hardware (Burger, 2006; Grahn & Geladi, 2007). The point scan configuration is commonly used in a chemical laboratory at the microscopic level (Sahlin & Peppas, 1997).

### 3.2.2.2 Focal plane scan imaging configuration

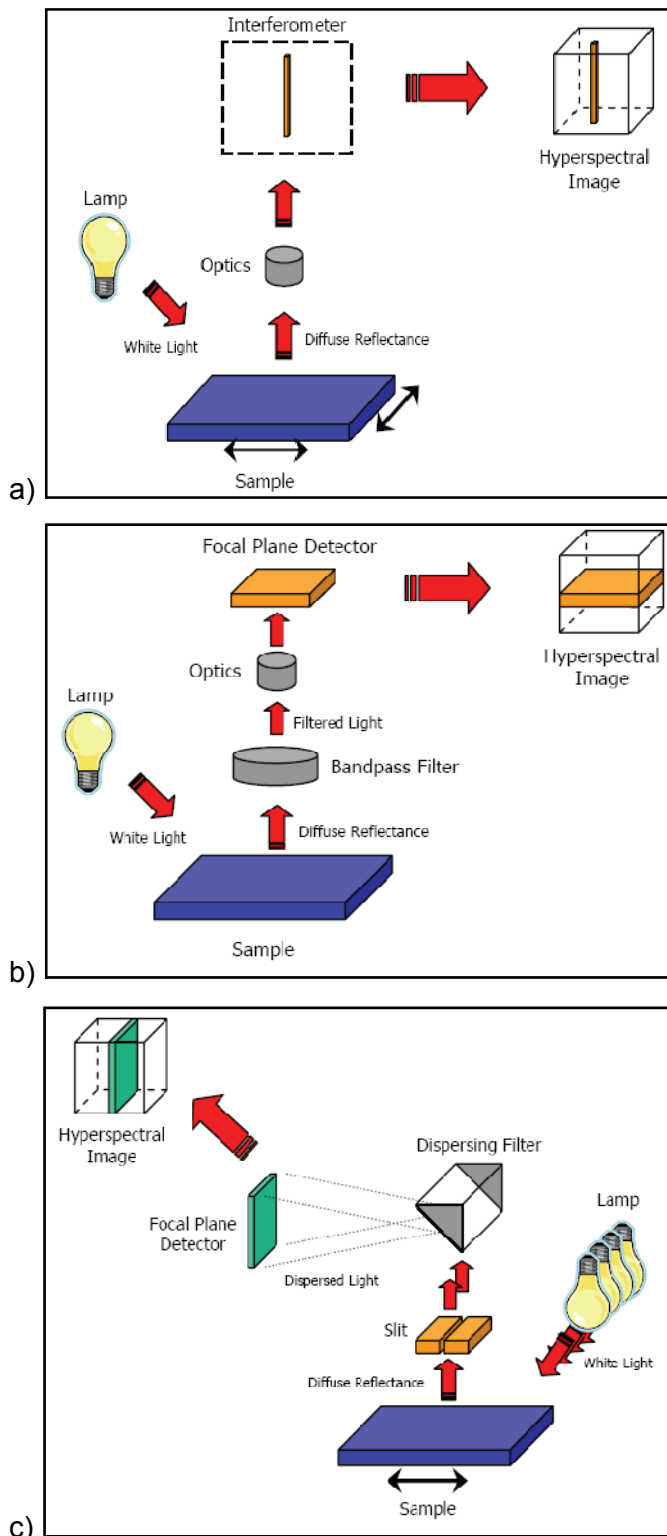
When using the focal plane scan, or staring imager configuration (**Fig. 2.5b**) the image field of view is fixed, with the detector situated parallel to it; one wavelength after another is acquired. It may also be that the illumination is scanned as in the videometer where the incoming light is filtered, or light emitting diodes (LEDs) of fixed wavelength in used. Hypercubes obtained, using this configuration, consist of a three dimensional stack of images, one image per wavelength measured (Burger, 2006; Gowen *et al.*, 2007; Grahn & Geladi, 2007). AOTFs and LCTFs are mostly employed in this configuration. AOTFs has the advantage of good transmission efficiency, fast scanning times and a large spectral range, while LCTFs have superior spectral band pass and image quality (Gowen *et al.*, 2007). Light illumination intensities may vary across the spatial axes of the sample area, thus requiring calibration at all pixel locations and all wavelengths performed with the use of dark and white reference standards (Burger, 2006).

The lack of moving parts in the focal plane scan is an advantage in many situations, as it may lead to a lower background noise level, providing the sample remain stationary (Burger, 2006; Gowen *et al.*, 2007). This configuration has been used in many pharmaceutical applications (Roggo *et al.*, 2005).

### 3.2.2.3 Line scan imaging configuration

The line scan imaging or pushbroom configuration (**Fig. 2.5c**) uses a two-dimensional detector perpendicular to the surface of the sample, and requires movement of the sample relative to the detector (Burger, 2006; Gowen *et al.*, 2007; Grahn & Geladi, 2007). The reflectance from the sample is passed through a "slit" and dispersed onto the detector (Burger, 2006). The spectrum of each pixel per line of the sample is simultaneously recorded by a detector, where a two-dimensional image (one spatial by wavelength dimension) is acquired (Gowen *et al.*, 2008b). The sample moves forward on a motorised stage, and the line adjacent to the one already acquired is then imaged (Gowen *et al.*, 2008b). No filter change is necessary in this configuration, enabling

very fast image acquisition time as the speed of acquisition is only limited by the camera speed. The camera can be operated in a continuous mode, permitting measurement of samples with an unlimited size and the ability of being implemented for online applications (Burger, 2006; Chao *et al.*, 2007).



**Fig 2.5** Schematic representation of the a) point scan, b) focal plane scan and c) the line scan configuration showing how images are acquired and the hyperspectral image obtained (Burger, 2006).

#### 4. Chemometrics and multivariate image analysis

Chemometrics can be defined as the use of statistical and mathematical techniques to analyse chemical data, and the entire process whereby data are transformed into information used for decision making purposes (Beebe *et al.*, 1998). Chemometrics involves taking a data matrix and extracting significant information about the objects and variables from it (Geladi, 2003). This enables the reduction of information contained in enormous data matrices to information that can be understood more easily, and separation of a residual containing mainly noise (Geladi, 2003).

In the multivariate data analysis (MDA) process, chemometrics is applied to a data set, whereby extracting relevant information from the data matrix. Chemometrics techniques are also applied in the form of multivariate image analysis (MIA) for analysing multivariate images, and lately hyperspectral image analysis for analysing hyperspectral images.

Different outcomes for the data, whether it be data exploration, classification or curve resolution, require different chemometrics techniques to be achieved. Data exploration is performed to find interesting phenomena about the data, often without prior expectations (Geladi, 2003). Classification of data into different groups can be performed through unsupervised classification techniques such as principal component analysis (PCA) if no information is known about the samples, or through supervised classification techniques such as partial least squares discriminant analysis (PLS-DA) when sufficient information is known about the sample.

In this section pretreatments used to prepare data for further analysis, as well as, PCA and PLS-DA with their applicability to hyperspectral image analysis will be briefly discussed.

##### 4.1 Pretreatment

Signal processing is used to transform spectral data prior to calibration, therefore referred to as data pretreatment or preprocessing methods (Brereton, 1990; Bro & Heimdal, 1996; Beebe *et al.*, 1998; Heise & Winzen, 2002). Pretreatment methods increase the signal-to-noise ratio in the data by reducing the noise level arising from random noise, baseline effects or spectral interferences (Heise & Winzen, 2002; Pizarro *et al.*, 2004; Nicolai *et al.*, 2007; Gowen *et al.*, 2008b; Hilden *et al.*, 2008; Qina & Lu, 2008). Sources of spectral variation could be the interaction of compounds, light scattering, path length variations and spectral distortions due to spectrometer hardware (Heise & Winzen, 2002; Pizarro *et al.*, 2004; Burger, 2006). The aim of signal processing is to reduce, eliminate or standardise these effects on the spectral data without influencing the spectroscopic information (Beebe *et al.*, 1998; Heise & Winzen, 2002; Pizarro *et al.*, 2004). Most pretreatments also bear the potential danger of influencing a useful part of the spectroscopic information, therefore it is essential that the spectroscopist has knowledge of the suitable technique including the proper optimisation and pretreatment parameters (Bro & Heimdal, 1996; Heise & Winzen, 2002).

Pretreatment methods commonly used include standardising, normalisation, sample weighting, smoothing, local filters, Savitzky-Golay smoothing, Fourier filtering, derivatives,



baseline correction methods, multiplicative scatter correction (MSC), standard normal variate (SNV), orthogonal signal correction (OSC), mean centering and variable weighting (Beebe *et al.*, 1998; Feudale *et al.*, 2002; Heise & Winzen, 2002; Nicolaï *et al.*, 2007). For the purpose of this thesis only derivatives, SNV and MSC will be discussed.

#### 4.1.1 Derivatives

Changes in the instrumentation used, lighting, temperature, detector response, sample presentation and orientation and sample particle size could result in background signal added uniformly as a function of wavelength, or equally through the spectra (Burger, 2006). These effects can be corrected for by applying first and second derivatives to the data (Norris & Williams, 1984; Hruschka, 2001). First and second derivatives are used to separate overlapping peaks and to correct for baseline shifts. Application of derivatives leads to an increase in spectral resolution, with a resultant increase in the signal-to-noise ratio (Luypaert *et al.*, 2004; Pizarro *et al.*, 2004). In order to decrease the signal-to-noise ratio, the Savitzky-Golay algorithm is applied to the data before the derivative. The Savitzky-Golay algorithm is a moving window averaging method, which smooths the spectra (Savitzky & Golay, 1964). A window is selected where the data are fitted by a polynomial; the central point in the window is replaced by the value calculated from the polynomial and then by the derivative of this polynomial (Luypaert *et al.*, 2004).

#### 4.1.2 Standard normal variate

Standard normal variate (SNV) is a scatter correction method, which normalizes spectra when the effective path length varies amongst the samples of a data set (Barnes *et al.*, 1989). The path length differences could be caused by granular samples, if the sample presentation is not fully reproducible or if the particle size differs between the samples (Pizarro *et al.*, 2004). SNV effectively creates a common baseline offset (zero) and variance for each spectrum (Burger, 2006).

This correction is performed according to **equation 2.1**, spectrum wise, by subtracting the mean spectrum ( $\bar{x}_i$ ) from the spectrum ( $x_{ij}$ ), whereafter dividing each element in the spectrum with its standard deviation ( $s_i$ ).

$$x_{i,SNV} = \left( \frac{x_{ij} - \bar{x}_i}{s_i} \right) \quad \text{equation 2.1}$$

#### 4.1.3 Multiplicative scatter correction

Multiplicative scatter correction (MSC) is based on the principle that light scatter has wavelength dependencies different from that of chemically based light absorption (Geladi *et al.*, 1985; Isaksson & Næs, 1988). Distinction between absorption and scatter is possible when data from many wavelengths are available. The scatter for each sample is estimated relative to that of the

ideal sample; resulting in each sample being corrected to have similar scatter levels as that of the ideal sample.

MSC uses linear regression of spectral variables vs the average spectrum; correcting for both multiplicative and additive scatter effects. The mean spectrum is calculated from all the spectra in the dataset. A least squares linear regression is performed on the sample spectrum versus (vs) those at corresponding wavelengths in the mean spectrum. A linear equation (**equation 2.2**) with defined intercept and slope is then obtained. The obtained value of the slope is subtracted from each data point in the spectrum, whereafter each absorbance value in the resulting spectrum is divided by the value of the slope (**equation 2.3**). The same operation is performed on every spectrum within the dataset. The aim of MSC is to separate the multiplicative and additive effects in NIR spectra, to minimize the spectral variations that are not due to the analyte concentration (Pizarro *et al.*, 2004).

$$x_{ij} = a_i + b_i \bar{x}_j + e_i \quad \text{equation 2.2}$$

$$x_{ij,MSC} = \left( \frac{x_{ij} - a_i}{b_i} \right) \quad \text{equation 2.3}$$

where  $x_{ij}$  is an individual spectrum,  $i$ ,  $\bar{x}_j$  the mean spectrum and  $e_i$  the residual. The fitted constants  $a_i$  (intercept) and  $b_i$  (slope) is used to correct each value of the spectrum  $i$ , to compute the corrected spectrum  $x_{ij,MSC}$ .

After pretreatment has assisted with increasing the signal-to-noise ratio, further regression and calibration techniques could be applied to the data. Regression techniques are necessary to extract information entangled in the overtones and combination bands of NIR spectra, as well as, important information captured in the hypercube.

#### 4.2 Principal component analysis

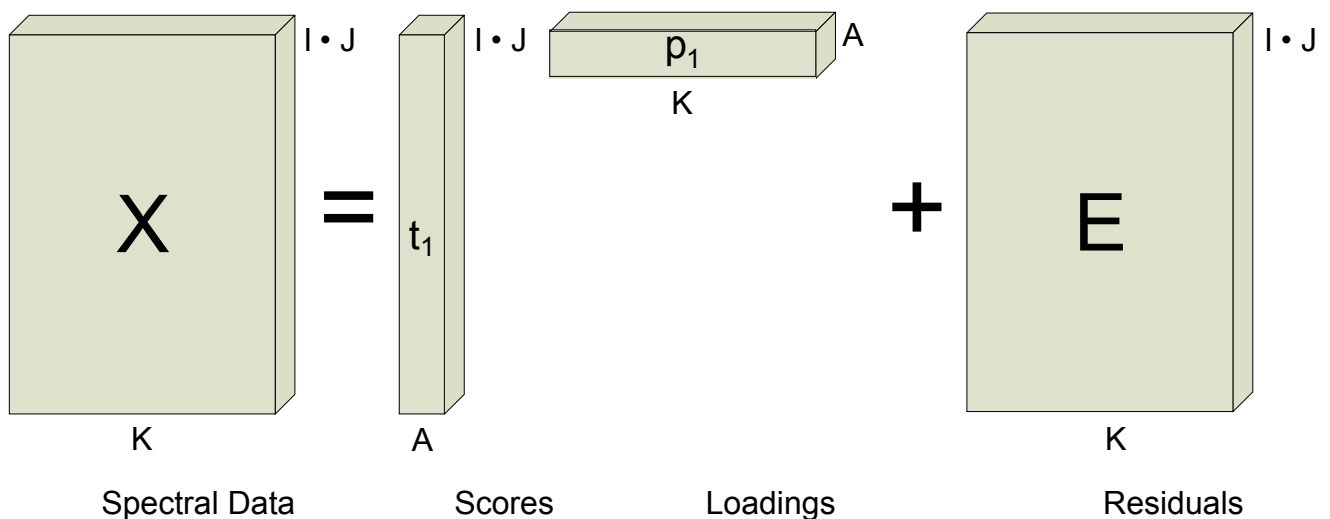
The hypercube obtained during hyperspectral image acquisition has three dimension, namely distance (I pixels) x distance (J pixels) x wavelength (K). This hypercube can be unfolded to create a two dimensional dataset, matrix  $\mathbf{X}$ , with the size  $K \times (I \cdot J)$  (Geladi *et al.*, 1989; Burger, 2006; Manley *et al.*, 2009). Principal component analysis (PCA) can now be applied to this unfolded matrix. PCA is a chemometrics technique employed to reduce dimensionality, as well as, the number of variables to be included in the chemometrics model, whilst retaining as much as possible of the variation present in the data set (Pearson, 1901; Jolliffe, 1986; Bucci *et al.*, 2002; Geladi, 2003; Wang *et al.*, 2006; Grahn & Geladi, 2007).

PCA projects data onto a lower orthogonal subspace, so that the first principal component (PC) accounts for the largest amount of variation originally present in the data set; the second PC for the largest amount of residual variance not explained by the first PC, and so on until all the sample variance are combined into component groups (Jolliffe, 1986; Bucci *et al.*, 2002; Geladi, 2003).

Each PC represents an independent source of variation in the data matrix (Wang *et al.*, 2006). PCA forms the basis for multivariate image analysis and enables the reduction of hundreds of image planes to a smaller amount (<10) that capture the maximum variation in the sample and are easier to interpret (Roggo *et al.*, 2005; Gowen *et al.*, 2008a). The bi-linear decomposition of the data is explained by **equation 2.4**:

$$X = TP' + E \quad \text{equation 2.4}$$

Where  $X$  is a  $(I \cdot J)$  matrix, mean centred or scaled in some way;  $T$  the score matrix with score vectors as columns;  $P$  the loading matrix with loading vectors as columns and  $E$  the  $(I \cdot J)$  residual matrix (also see **Fig. 2.6** for geometric representation). The components can also be referred to as latent variables or principal components (Geladi, 2003; Manley *et al.*, 2009).



**Figure 2.6** A data matrix of size  $K \times [I \cdot J]$  are reduced to smaller matrices of size  $A \times [I \cdot J]$  and  $K \times A$  that are easier to interpret and contain all the relevant information from the data matrix. Noise and other disturbances remain in the residual matrix of size  $K \times [I \cdot J]$ . Adapted from (Geladi, 2003; Burger, 2006).

The column vectors in  $P$  are called loadings, while those in  $T$  are called scores. These can be viewed as scatter plots and line plots to allow efficient interpretation of the data space. Scatter plots are frequently used to detect similarities, differences or other interesting phenomena between samples, while outliers, subgroups and clusters can also be detected. The plot of each loading vector against wavelength number is used to visualise the spectral pattern of each PC. Loadings are used to interpret the relationship between scores and the original variables (wavelength), which would indicate which variables are most important for the different components in the data set. When classifying clusters in the score plot, the loading lines are useful for identifying which variables characterise the different clusters (Næs *et al.*, 2002).

PCA has been successfully used in numerous applications of image analysis, e.g. for the detection of bruises on pickling cucumbers (Ariana *et al.*, 2006) and mushrooms (Gowen *et al.*, 2008a); the detection of cucumber chilling damage (Cheng *et al.*, 2004); the investigation of beef tenderness prediction (Naganathan *et al.*, 2008); the detection of defects on apple cultivars (Mehl *et al.*, 2002) and the detection of faecal contamination on apples (Liu *et al.*, 2007).

The advantage of PCA is that it reduces the number of variables to analyse. The drawback of PCA is the interpretation of the loadings, because several chemical components could contribute to one loading (Roggo *et al.*, 2005).

#### 4.3 Partial least squares discriminant analysis

The partial least squares (PLS) technique was first developed by Herman Wold in the 1970's and 1980's to deal with problems concerning econometric path-modelling (Wold, 1982).

Partial least squares regression (PLS-R) is a statistical regression method that bears some resemblance to principal component regression (PCR). Instead of finding hyperplanes of maximum variance in **X** as with PCR, it builds a linear regression model by estimating **B** in such a way as to maximise the covariance between **Y** and **X** (**equation 2.5**).

$$Y = XB + F \quad \text{equation 2.5}$$

Where **Y** is a vector of mean centred concentrations, **X** is a matrix with mean centred spectra, **B** is a vector of regression coefficients and **F** is a matrix of residuals.

The most important asset of PLS-R is that components referred to as latent variables are used. These latent variables are aligned with the maximum covariance of **X** and **Y**. The amount of latent variables to use are very important. Not using enough components results in underfitting and bad models; while using too many components result in overfitting and bad prediction (Burger, 2006; Esbensen, 2006; Grahn & Geladi, 2007).

Partial least squares discriminant analysis (PLS-DA) is a supervised pattern recognition technique that correlates variation in the data set with class membership. This technique provides classification and discrimination between different classes. The basis of PLS-DA consist of the application of a PLS-R model on variables, which are indicators of certain groups, on a set of **X** predictor variables (Perez-Enciso & Tenenhaus, 2003; Karp *et al.*, 2005). Classification of observations from the results of PLS-R is then performed on indicator variables (**equation 2.5**) (Chevallier *et al.*, 2006). Interaction takes place between **X** data (e.g. NIR spectra) and **Y** variables (e.g. wheat hardness, pH, protein) for discrimination and classification purposes. **Y** is a dummy matrix with ones and zeros representing classes, which is paired with **X**.

PLS-DA has been successfully implemented in research fields such as proteomics (Karp *et al.*, 2005; Marengo *et al.*, 2008), the prediction of tumour type (Perez-Enciso & Tenenhaus, 2003) and the discrimination between meat from adult and young cattle (Prieto *et al.*, 2008).

#### 4.4 Multivariate image analysis

A hyperspectral image is merely an image containing much more variables than a multivariate image (Geladi *et al.*, 2004). The approach for analysing hyperspectral images is thus similar to those used for analysing multivariate images.

The proposed approach for the analysis of hyperspectral images contains eight steps:

##### 1. Image correction

Images are corrected by using dark and white reference standards, and thereafter are converted from A/D converter values to pseudo absorbance values.

##### 2. Image cleaning

After the images are converted to absorbance, image analysis will be continued with, e.g. dedicated image analysis software such as Evince multivariate image analysis software (UmBio AB, Umeå, Sweden). The first step is generally to do a PCA with three PCs of the image. The largest amount of variation will be explained in these first PCs (i.e. variation between wheat kernels and background, shading or dead pixels) which will allow effective cleaning of the image. The PCA performed will generate score images, score plots and loading line plots. The interactive score images and score plots are used to detect unwanted regions of the image, such as background, shading errors, dead pixels, detector errors and detector saturation (Grahm & Geladi, 2007). These unwanted pixels are identified by selecting a certain region in the score plot, whereafter the corresponding region will be highlighted in the score image due to the interactive nature of the score plot and score image, indicating the localisation of the regions in the image (Grahm & Geladi, 2007). This is performed with various combinations of PCs plotted against each other to effectively remove all unwanted data from the image. The unwanted pixels can then be removed from the image and the PCA recalculated, to remove them from further calculations. By the removal of these unwanted pixels the remaining pixels in the score plot are allowed to spread out more, making previously hidden clusters visible and improving the overall visibility of data points in the score plot.

##### 3. Add additional principal components

Additional PCs are added; normally no more than six PCs are useful in analysis of images. The appearance of detector errors becomes much more apparent in the higher components, making it difficult to extract useful information from the higher components.

#### 4. Examination of the principal components

Various combinations of PCs are plotted against each other in the score plot and studied for the clustering of data points. By studying the score plots it can be determined which components contribute mostly to distinctly separating the clusters. Due to knowledge of your sample(s) one would know how many clusters to expect; i.e. if wheat with three different hardness classes are analysed, three clusters are expected to be visible in the score plot. Overlapping of clusters could occur due to chemical similarities between samples.

The interactivity of the score plot and score image are once again used to identify regions in the score image. Clusters in the score plot are selected; resulting in the regions where they originate in the score image to be highlighted.

#### 5. Identification of clusters in the score plot

Prior knowledge of the sample is used to identify the clusters. Observations from the interactivity of the score plot and score image with clustering in the score plot together with the knowledge about the samples are used to identify what chemical compounds each cluster represents.

#### 6. Classification of clusters in the score plot

Clusters, related to the differences or similarities in the samples, are identified and each cluster is classified accordingly. When clusters are classified in the score plot the score image changes to a classification image, whereby visualisation of the distribution of different chemical compounds in the sample is improved.

#### 7. Interpretation of loading line plots

Loading line plots are studied to determine if there are absorption peaks of interest present, and to identify the chemical components responsible for variation within the component. Chemical components present are identified by using absorption tables (Osborne *et al.*, 1993) and correlating them with absorption peaks in the loading line plots of each PC.

#### 8. Application of pretreatment techniques

Depending on the samples being analysed and the results obtained, pretreatment techniques may be applied (Beebe *et al.*, 1998). After pretreatment has been applied, the analysis process repeats from step 4.

If the analyst has prior knowledge of the samples analysed, supervised chemometrics techniques such as PLS-DA could further be applied.

## 5. Conclusion

The determination of wheat hardness is possible with a variety of destructive methods. There still remains a need for the non-destructive determination of wheat hardness during the breeding phase of wheat, where each kernel is vitally important. Due to the non-destructive nature of NIR hyperspectral imaging, this new technology could have promising outcomes for the detection of wheat hardness. Hyperspectral imaging technology are not yet directly implemented for online sorting of agricultural products, due to the extensive time required for image acquisition and analysis; however this could have useful application in the agricultural industry. Hyperspectral image analysis can at this stage be very useful as a research tool for determining spectral bands of interest as in the case of wheat hardness determination, which could then be applied to an online multispectral imaging system.

The diffusion of conditioning water into single wheat kernels has been studied in various ways. The use of NIR hyperspectral imaging would not only give a new perspective on the diffusion of the conditioning water, but also a quick look at novel applications of NIR hyperspectral imaging.

## 6. References

- AACC (2009a). AACC International Approved Methods of Analysis, 11th Ed. Method 39-70.02. Near-Infrared Reflectance Method for Hardness Determination in Wheat. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-39-70.02.
- AACC (2009b). AACC International Approved Methods of Analysis, 11th Ed. Method 55-30.01. Particle Size Index for Wheat Hardness. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-55-30.01.
- AACC (2009c). AACC International Approved Methods of Analysis, 11th Ed. Method 55-31.01. Single Kernel Characterisation System for Wheat Kernel Texture. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-55-31.01.
- Abney, W. & Festing, E.R. (1881). On the influence of the atomic grouping in the molecules of organic bodies on their absorption in the infra-red region of the spectrum. *Philosophical Transactions of the Royal Society*, **172**, 887-918.
- Anonymous. (1996). Wheat: wheat kernel. [WWW document]. <http://www.britannica.com/EBchecked/topic/103350/cereal-processing/50119/Milling>. 4 February 2009.
- Ariana, D.P., Lua, R. & Guyer, D.E. (2006). Near-infrared hyperspectral reflectance imaging for detection of bruises on pickling cucumbers. *Computers and Electronics in Agriculture*, **53**, 60-70.
- Baeten, V., Fernández Pierna, J.A. & Dardenne, P. (2007). Hyperspectral imaging techniques: an attractive solution for the analysis of biological and agricultural materials. In: *Techniques and applications of hyperspectral image analysis* (edited by H.F. Grahn & P. Geladi). Pp. 289-311. Chichester, England: John Wiley & Sons, Ltd.
- Bakhella, M., Hosene, R.C. & Lookhart, G.L. (1990). Hardness of Moroccan wheats. *Cereal Chemistry*, **67**, 246-250.
- Barlow, K.K., Buttrose, M.S., Simmonds, D.H. & Vesik, M. (1973). The nature of the starch-protein interface in wheat endosperm. *Cereal Chemistry*, **50**, 443-454.

- Barnes, R.J., Dhanoa, M.A. & Lister, S.J. (1989). Standard normal variate transform and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, **43**, 772-777.
- Bass, E.J. (1988). Wheat flour milling. In: *Wheat Chemistry and Technology* (edited by Y. Pomeranz). Pp. 1-68. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Beebe, K.R., Pell, R.J. & Seasholtz, M.B. (1998). *Chemometrics: a Practical Guide*. Pp. 1-8, 26-55. New York, USA: John Wiley & Sons, Inc.
- Beta, T. (2004). Grain production and consumption, overview. In: *Encyclopedia of Grain Science* (edited by C. Wrigley, H. Corke & C.E. Walker). **Vol. 3**. Pp. 61-70. Kidlington, Oxford, UK: Elsevier Ltd.
- Bietz, J.A. (1989). Telling differences among wheat proteins may make a difference in marketing. In: *Wheat is Unique. Structure, Composition, Processing, End-use Properties, and Products* (edited by Y. Pomeranz). Pp. 303-315. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Blanco, M. & Alcalá, M. (2006). Simultaneous quantitation of five active principles in a pharmaceutical preparation: development and validation of a near infrared spectroscopic method. *European Journal of Pharmaceutical Sciences*, **27**, 280-286.
- Blanco, M., Coello, J., Iturriaga, H., MasPOCH, S. & De la Pezuela, C. (1998). Near-infrared spectroscopy in the pharmaceutical industry. *Analyst*, **123**, 135-150.
- Bolling, H. (1987). Milling quality of wheat. In: *European Conference on Food Science and Technology* (edited by I.D. Morton). Pp. 259-284. Bournemouth, England: Ellis Horwood.
- Brereton, R.G. (1990). Pattern recognition. In: *Chemometrics: Applications of Mathematics and Statistics to Laboratory Systems*. Pp. 239-295. Chichester, West Sussex, England: Ellis Horwood Limited.
- Bro, R. & Heimdal, H. (1996). Enzymatic browning of vegetables: calibration and analysis of variance by multiway methods. *Chemometrics and Intelligent Laboratory Systems*, **34**, 85-102.
- Bucci, R., Magriä, A.D., Magriä, A.L., Marini, D. & Marini, F. (2002). Chemical authentication of extra virgin olive oil varieties by supervised chemometric procedures. *Journal of Agricultural and Food Chemistry*, **50**, 413-418.
- Burger, J. (2006). Hyperspectral NIR image analysis: data exploration, correction and regression. PhD Thesis Unit of Biomass Technology and Chemistry, Swedish University of Agricultural Sciences.
- Burger, J. & Geladi, P. (2005). Hyperspectral NIR image regression part I: calibration and correction. *Journal of Chemometrics*, **19**, 355-363.
- Burger, J. & Geladi, P. (2006). Hyperspectral NIR imaging for calibration and prediction: a comparison between image and spectrometer data for studying organic and biological samples. *Analyst*, **131**, 1152-1160.
- Butcher, J. & Stenvert, N.L. (1973). Conditioning studies on Australian wheat; III. The role of the rate of water penetration into wheat grain. *Journal of the Science of Food and Agriculture*, **24**, 1077-1084.
- Campbell, J.D. & Jones, C.R. (1957). The rates of penetration of moisture to different points in the central cross-section of the endosperm in damped Manitoba wheat grains. *Cereal Chemistry*, **34**, 110-116.
- Chao, K., Chen, Y.R. & Chan, D.E. (2003). Analysis of VIS/NIR spectral variations of wholesome, septicemia, and cadaver chicken samples. *Applied Engineering in Agriculture*, **19**, 453 - 458.
- Chao, K., Chen, Y.R., Early, H. & Park, B. (1999). Color image classification systems for poultry viscera inspection. *Applied Engineering in Agriculture*, **15**, 363-369.



- Chao, K., Yang, C.C., Chen, Y.R., Kim, M.S. & Chan, D.E. (2007). Hyperspectral-multispectral line-scan imaging system for automated poultry carcass inspection applications for food safety. *Poultry Science*, **86**, 2450-2460.
- Chen, M.H., Kerechanin, C.W., Greenspan, D.G., Criss, T.B., Franckowiak, S.C., Vincent, J.A. & Pattay, R.S. (2005). Development of a thermal and hyperspectral imaging system for wound characterization and metabolic correlation. *Johns Hopkins APL Technical Digest*, **26**, 67-74.
- Cheng, X., Chen, Y.R., Tao, Y., Wang, C.Y., Kim, M.S. & Lefcourt, A.M. (2004). A novel intergrated PCA and PLD method on hyperspectral image feature extraction for cucumber chilling damage inspection. *Transactions of the ASAE*, **47**, 1313 - 1320.
- Chevallier, S., Bertrand, D., Kohler, A. & Courcoux, P. (2006). Application of PLS-DA in multivariate image analysis. *Journal of Chemometrics*, **20**, 221-229.
- Cho, M.J., Morawicki, R. & Guerra, C. (2009). Ability of near-infrared spectroscopy to predict solid loss, volumetric expansion, and water uptake of cooked rice. *Cereal Chemistry*, **86**, 113-116.
- Cleve, E., Bach, E. & Schollmeyer, E. (2000). Using chemometric methods and NIR spectrophotometry in the textile industry. *Analytica Chimica Acta*, **420**, 163-167.
- Cobb, N.A. (1986). The hardness of the grain in the principal varieties of wheat. *Agricultural Gazette New South Wales*, **7**, 279-298.
- Coblentz, W. (1905). Investigations of infrared spectra, part 1. In: *Publication No. 35*. Carnegie Institute of Washington: Washington DC. (republished under the joint sponsorship of the Coblentz Society and the Perkin-Elmer Corporation, 1962)
- Cogdill, R.P., Hurburgh, C.R. & Rippke, G.R. (2004). Single-kernel maize analysis by near-infrared hyperspectral imaging. *Transactions of the ASAE*, **47**, 311-320.
- Colón, J., Peroza, C., Caraballo, W., Conde, C., Li, T., Morris, K.R. & Romañach, R.J. (2005). On line non-destructive determination of drug content in moving tablets using near infrared spectroscopy. *Journal of Process Analytical Technology*, **2**, 8-15.
- Dahm, D.J. & Dahm, K.D. (2001). The physics of near infrared scattering. In: *Near-Infrared Technology in the Agricultural and Food Industries* (edited by P. Williams & K. Norris). Pp. 1-18. St. Paul, Minnesota, USA: American Association of Cereal Chemists.
- Daley, W.D., Carey, R. & Thompson, C. (1993). Poultry grading/inspection using color imaging. *Proceedings of the SPIE*, **1907**, 124-132.
- Davidson, T.M., Deconde, K.T., Hake, R., Tracy, D.H., Gantz, A. & McDermott, L. (1992). Precision of the petrochemical process analysis using NIR spectroscopy. *Proceedings of the International Society for Optical Engineering*, **1681**, 231-235.
- Delwiche, S.R. (2003). Classifications of scab- and other mold-damaged wheat kernels by near-infrared reflectance spectroscopy. *Transactions of the ASAE*, **46**, 731-738.
- Dexter, J.E. & Sarkar, A.K. (2004). Wheat dry milling. In: *Encyclopedia of Grain Science* (edited by C. Wrigley, H. Corke & C.E. Walker). **Vol. 3**. Pp. 363-374. Kidlington, Oxford, UK: Elsevier Ltd.
- Downey, G., Byrne, S. & Dwyer, E. (1986). Wheat trading in the republic of Ireland: the utility of a hardness index derived by near infrared reflectance spectroscopy. *Journal of the Science of Food and Agriculture*, **37**, 762-766.
- Ellis, J.W. (1929). Molecular absorption spectra of liquids below 3 $\mu$ . *Transactions of the Faraday Society*, **25**, 888-898.

- Esbensen, K.H. (2006). *Multivariate Data Analysis in Practice*. Pp. 598. Norway: CAMO Software AS.
- Evers, A.D. & Bechtel, D.B. (1988). Microscopic structure of the wheat grain. In: *Wheat: Chemistry and Technology* (edited by Y. Pomeranz). Pp. 47-95. St. Paul, Minnesota, USA: American Association of Cereal Chemists.
- Fernández Pierna, J.A., Baeten, V. & Dardenne, P. (2006). Screening of compound feeds using NIR hyperspectral data. *Chemometrics and Intelligent Laboratory Systems*, **84**, 114-118.
- Feudale, R.N., Woody, N.A., Tan, H., Myles, A.J., Brown, S.D. & Ferré, J. (2002). Transfer of multivariate calibration models: a review. *Chemometrics and Intelligent Laboratory Systems*, **64**, 181- 192.
- Gaines, C.S., Finney, P.F., Fleege, L.M. & Andrews, L.C. (1996). Predicting a hardness measurement using the single-kernel characterization system. *Cereal Chemistry*, **73**, 278-283.
- Gaines, C.S. & Windham, W.R. (1998). Effect of wheat moisture content on meal apparent particle size and hardness scores determined by near-infrared reflectance spectroscopy. *Cereal Chemistry*, **75**, 386-391.
- Geladi, P. (2003). Chemometrics in spectroscopy. part 1. classical chemometrics. *Spectrochimica Acta Part B*, **58**, 767-782.
- Geladi, P., Burger, J. & Lestander, T. (2004). Hyperspectral imaging: calibration problems and solutions. *Chemometrics and Intelligent Laboratory Systems*, **72**, 209-217.
- Geladi, P., Isaksson, H., Lindqvist, L., Wold, S. & Esbensen, K. (1989). Principal component analysis and multivariate images. *Chemometrics and Intelligent Laboratory Systems*, **5**, 209-220.
- Geladi, P., MacDougall, D. & Martens, H. (1985). Linearization and scatter-correction for near-infrared reflectance spectra of meat. *Applied Spectroscopy*, **39**, 491-500.
- Geladi, P.L.M., Grahn, H.F. & Burger, J.E. (2007). Multivariate images, hyperspectral imaging: background and equipment. In: *Techniques and Applications of Hyperspectral Image Analysis* (edited by P.L.M. Geladi & H.F. Grahn). Pp. 1-16. Chichester, England.: John Wiley & Sons.
- Giroux, M.J. & Morris, C.F. (1997). A glycine to serine change in puroindoline b is associated with wheat grain hardness and low levels of starch-surface fraibilin. *Theoretical and Applied Genetics*, **95**, 857-864.
- Goetz, A.F.H., Vane, G., Solomon, T.E. & Rock, B.N. (1985). Imaging spectrometry for earth remote sensing. *Science*, **228**, 1147-1153.
- Gorretta, N., Roger, J.M., Aubert, M., Bellon-Maurel, V., Campan, F. & Roumet, P. (2006). Determining vitreousness of durum wheat kernels using near infrared hyperspectral imaging. *Journal of Near Infrared Spectroscopy*, **14**, 231-239.
- Gowen, A.A., O'Donnell, C.P., Cullen, P.J., Downey, G. & Frias, J.M. (2007). Hyperspectral imaging - an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology*, **18**, 590-598.
- Gowen, A.A., O'Donnell, C.P., Taghizadeh, M., Cullen, P.J., Frias, J.M. & Downey, G. (2008a). Hyperspectral imaging combined with principal component analysis for bruise damage detection on white mushrooms (*Agaricus bisporus*). *Journal of Chemometrics*, **22**, 259-267.
- Gowen, A.A., O'Donnell, C.P., Cullen, P.J. & Bell, S.E.J. (2008b). Recent applications of chemical imaging to pharmaceutical process monitoring and quality control. *European Journal of Pharmaceutics and Biopharmaceutics*, **69**, 10-22.
- Grahn, H.F. & Geladi, P. (2007). *Techniques and Applications of Hyperspectral Image Analysis*. Pp. 1-15, 313-334. Chichester, England: John Wiley & Sons Ltd.
- Greenaway, W.T. (1969). A wheat hardness index. *Cereal Science Today*, **12**, 4-7.

- Greenwell, P. & Schofield, J.D. (1986). A starch granule protein associated with endosperm softness in wheat. *Cereal Chemistry*, **63**, 379-380.
- Grundas, S.T. & Wrigley, C. (2004). Wheat, ultrastructure of the grain, flour, and dough. In: *Encyclopedia of Grain Science* (edited by C. Wrigley, H. Corke & C.E. Walker). **Vol. 3**. Pp. 391-400. Kidlington, Oxford, UK: Elsevier Ltd.
- Halverson, J. & Zeleny, L.W. (1988). Criteria of wheat quality. In: *Wheat Chemistry and Technology* (edited by Y. Pomeranz). Pp. 15-33. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Hamilton, S.J. & Lodder, R.A. (2002). Hyperspectral imaging technology for pharmaceutical analysis. *Proceedings of the SPIE*, **4626**, 136-147.
- Hart, J.R., Norris, K.H. & Golumbic, C. (1962). Determination of the moisture content of seeds by near-infrared spectrophotometry of their methanol extracts. *Cereal Chemistry*, **39**, 94-99.
- Hege, E., O'Connell, D., Johnson, W., Basty, S. & Dereniak, E. (2003). Hyperspectral imaging for astronomy and space surveillance. *Proceedings of the SPIE*, **5159**, 380-391.
- Heise, H.M. & Winzen, R. (2002). Chemometrics in near-infrared spectroscopy. In: *Near-Infrared Spectroscopy: Principles, Instruments, Applications* (edited by H.W. Siesler, Y. Ozaki, S. Kawata & H.M. Heise). Pp. 125-161. Weinheim, Germany: Wiley-VCH Verlag GmbH.
- Herschel, W. (1800). Investigation of the power of the prismatic colours to heat and illuminate objects; with remarks, that prove the different refrangibility of radiant heat. To which is added, an inquiry into the methods of viewing the sun advantageously with telescopes of large apertures and high magnifying powers. *Philosophical Transactions of the Royal Society*, **90**, 255-283.
- Hilden, L.R., Pommiera, C.J., Badawya, S.I.F. & Friedman, E.M. (2008). NIR chemical imaging to guide/support BMS-561389 tablet formulation development. *International Journal of Pharmaceutics*, **353**, 283-290.
- Hopkins, C.Y. & Graham, R.P. (1935). Starch content of some samples of Canadian wheat. *Canadian Journal of Research*, **12**, 820-824.
- Hoseney, R.C. (1994). *Principles of Cereal Science and Technology*. Pp. 378. St. Paul, Minnesota: American Association of Cereal Chemists, Inc.
- Hruschka, W.R. (2001). Data analysis: wavelength selection methods. In: *Near-infrared technology in the agriculture and food industries* (edited by P.C. Williams & K. Norris). Pp. 39-58. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Isaksson, T. & Næs, T. (1988). The effect of multiplicative scatter correction (MSC) and linearity improvement in NIR spectroscopy. *Applied Spectroscopy*, **42**, 1273-1284.
- Jolliffe, I.T. (1986). *Principal Component Analysis*. Pp. 271. New York, USA: Springer-Verlag.
- Karp, N.A., Griffin, J.L. & Lilley, K.S. (2005). Application of partial least squares discriminant analysis to two-dimensional difference gel studies in expression proteomics. *Proteomics*, **5**, 81-90.
- Kazemi, S., Wang, N., Ngadi, M. & Prasher, S.O. (2005). Evaluation of frying oil quality using VIS/NIR hyperspectral analysis. *Agricultural Engineering International: The CIGR Journal*, **VII**, 1-12.
- Kent, N.L. & Evers, A.D. (1994). *Kent's Technology of Cereals*. Oxford, UK: Pergamon.
- Khodabux, K., L'Omelette, M.S.S., Jhaumeer-Laulloo, S., Ramasami, P. & Rondeau, P. (2007). Chemical and near-infrared determination of moisture, fat and protein in tuna fishes. *Food Chemistry*, **102**, 669-675.

- Kirby, E.J.M. (2002). Botany of the wheat plant. In: *Bread Wheat Improvement and Production* (edited by B.C. Curtis, S. Rajaram & H.G. Macpherson). Food and Agriculture Organisation.
- Koç, H., Smail, V.W. & Wentzel, D.L. (2008). Reliability of InGaAs focal plane array imaging of wheat germination at early stages. *Journal of Cereal Science*, **48**, 394-400.
- Koehler IV, F.W., Lee, E., Kidder, L.H. & Lewis, N.E. (2002). Near infrared spectroscopy: the practical chemical imaging solution. *Spectroscopy Europe*, **14**, 12-19.
- Korkut, K.Z., Bilgin, O., Başer, I. & Sağlam, N. (2007). Stability of grain vitreousness in durum wheat (*Triticum durum* L.Desf.) genotypes in the north-western region of Turkey. *Turkish Journal of Agriculture and Forestry*, **31**, 313-318.
- Lillemo, M. & Morris, C.F. (2000). A leucine to proline mutation in puroindoline b is frequently present in hard wheats from Northern Europe. *Theoretical and Applied Genetics*, **100**, 1100-1107.
- Litscher, G. (2005). NIR spectroscopy explores the mystery of acupuncture. *Spectroscopy Europe*, **17**, 8-14.
- Liu, Y., Chao, K., Chen, Y.R., Kim, M.S., Nou, X., Chan, D.E. & Yang, C. (2006). Comparison of visible and near infrared reflectance spectroscopy for the detection of faeces/ingesta contaminants for sanitation verification at slaughter plants. *Journal of Near Infrared Spectroscopy*, **14**, 325-331.
- Liu, Y., Chen, Y.R., Kim, M.S., Chan, D.E. & Lefcourt, A.M. (2007). Development of simple algorithms for the detection of fecal contaminants on apples from visible/near infrared hyperspectral reflectance imaging. *Journal of Food Engineering*, **81**, 412-418.
- Luypaert, J., Heuerding, S., Van der Heyden, Y. & Massart, D.L. (2004). The effect of preprocessing methods in reducing interfering variability from near-infrared measurements of creams. *Journal of Pharmaceutical and Biomedical Analysis*, **36**, 495-503.
- Luzuriaga, D.A. & Balaban, M.O. (2002). Color machine vision system: an alternative for color measurement. In: *World Congress of Computers in Agriculture and Natural Resources*. Pp. 93-100. Iguacu Falls, Brazil.
- Macho, S. & Larrechi, M.S. (2002). Near-infrared spectroscopy and multivariate calibration for the quantitative determination of certain properties in the petrochemical industry. *Trends in analytical chemistry*, **21**, 799-806.
- Mahesh, S., Manickavasagan, A., Jayas, D.S., Paliwal, J. & White, N.D.G. (2008). Feasibility of near-infrared hyperspectral imaging to differentiate Canadian wheat classes. *Biosystems Engineering*, **101**, 50-57.
- Manley, M., Williams, P., Nilsson, D. & Geladi, P. (2009). Near infrared hyperspectral imaging for the evaluation of endosperm texture in whole yellow maize (*Zea mays* L.) kernels. *Journal of Agricultural and Food Chemistry*, DOI 10.1021/jf9018323.
- Marengo, E., Robotti, E., Bobba, M., Milli, A., Campostrini, N., Righetti, S.C., Cecconi, D. & Righetti, P.G. (2008). Application of partial least squares discriminant analysis and variable selection procedures: a 2D-PAGE proteomic study. *Analytical and Bioanalytical Chemistry*, **390**, 1327-1342.
- McGlone, V.A., Devine, C.E. & Wells, R.W. (2005). Detection of tenderness, post-rigor age and water status changes in sheep meat using near infrared spectroscopy. *Journal of Near Infrared Spectroscopy*, **13**, 277-285.
- Mehl, P.M., Chao, K., Kim, M. & Chen, Y.R. (2002). Detection of defects on selected apple cultivars using hyperspectral and multispectral image analysis. *Applied Engineering in Agriculture*, **18**, 219-226.

- Miller, C.E. (2001). Chemical principles of near infrared technology. In: *Near Infrared Technology in the Agricultural and Food Industries* (edited by K. Norris & P.C. Williams). St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Moss, R. (1977). An autoradiographic technique for the location of conditioning water in wheat at cellular level. *Journal of the Science of Food and Agriculture*, **28**, 23-33.
- Murray, I. (1996). The value of traditional analytical methods and near-infrared (NIR) spectroscopy to the feed industry. In: *Recent Advances in Animal Nutrition* (edited by P.C. Gamsworthy & D.J.A. Cole). Pp. 87-110. Nottingham: Nottingham University Press.
- Næs, T., Isaksson, T., Fearn, T. & Davies, T. (2002). *A User Friendly Guide to Multivariate Calibration and Classification*. Pp. 344. Chichester, UK: NIR Publications.
- Naganathan, G.K., Grimes, L.M., Subbiah, J., Calkins, C.R., Samal, A. & Meyer, G.E. (2008). Visible/near-infrared hyperspectral imaging for beef tenderness prediction. *Computers and Electronics in Agriculture*, **64**, 225-233.
- Nicolaï, B.M., Beullens, K., Bobelyn, E., Peirs, A., Saeys, W., Theron, K.I. & Lammertyn, J. (2007). Nondestructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Postharvest Biology and Technology*, **46**, 99-118.
- Norris, K.H., Hruschka, W.R., Bean, M.M. & Slaughter, D.C. (1989). A definition of wheat hardness using near infrared reflectance spectroscopy. *Cereal Foods World*, **34**, 696-705.
- Norris, K.H. & Williams, P.C. (1984). Optimisation of mathematical treatments of raw near-infrared signal in the measurement of protein in hard red spring wheat. I. Influence of particle size. *Cereal Chemistry*, **61**, 158.
- O'Brien, L. & DePauw, R. (2004). Wheat breeding. In: *Encyclopedia of Grain Science* (edited by C. Wrigley, H. Corke & C.E. Walker). **Vol. 3**. Pp. 330-336. Kidlington, Oxford, UK: Elsevier Ltd.
- Oda, S. (1994). Two dimensional electrophoretic analysis of friabilin. *Cereal Chemistry*, **71**, 394-395.
- Orth, R.A. & Shellenberger, J.A. (1988). Origin, production, and utilization of wheat. In: *Wheat Chemistry and Technology* (edited by Y. Pomeranz). Pp. 1-14. St. Paul, Minnesota, USA: American Association of Cereal Chemists.
- Osborne, B. (1991). Measurement of the hardness of wheat endosperm by near infrared spectroscopy. *Postharvest News and Information*, **2**, 331-334.
- Osborne, B., Turnbull, K.M., Anderssen, R.S., Rahman, S., Sharp, P.J. & Appels, R. (2001). The hardness locus in Australian wheat lines. *Australian Journal of Agricultural Research*, **52**, 1275-1286.
- Osborne, B.G., Fearn, T. & Hindle, P.H. (1993). *Practical NIR Spectroscopy with Applications in Food and Beverage Analysis*. Harlow, England: Longman Scientific & Technical.
- Pasquini, C. (2003). Near infrared spectroscopy: fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, **14**, 198-219.
- Paulsen, G.M. & Shroyer, J.P. (2004). Wheat agronomy. In: *Encyclopedia of Grain Science* (edited by C. Wrigley, H. Corke & C.E. Walker). **Vol. 3**. Pp. 337-347. Kidlington, Oxford, UK: Elsevier Ltd.
- Pazourek, K. & M., C. (1988). Application of near infrared reflectance technique in the analysis of selected pig feed mixes. *Acta Alimentaria*, **17**, 113-125.
- Pearson, K. (1901). On lines and planes of closest fit to systems and points in space. *Philosophical Magazine*, **6**, 559-572.

- Peirs, A., Scheerlinck, N., Touchant, K. & Nicolai, B.M. (2002). Comparison of fourier transform and dispersive near-infrared reflectance spectroscopy for apple quality measurements. *Biosystems Engineering*, **81**, 305-311.
- Percival, J. (1921). *The Wheat Plant: a Monograph*. London: Duckworth.
- Perez-Enciso, M. & Tenenhaus, M. (2003). Prediction of clinical outcome with microarray data: a partial least squares discriminant analysis (PLS-DA) approach. *Human Genetics*, **112**, 581-592.
- Perez, D.P., Sanchez, M.T., Cano, G. & Garrido, A. (2001). Authentication of green asparagus varieties by near-infrared reflectance spectroscopy. *Journal of Food Science*, **66**, 323-327.
- Pizarro, C., Esteban-Díez, I., Nistal, A.-J. & González-Sáiz, J.-M. (2004). Influence of data pre-processing on the quantitative determination of the ash content and lipids in roasted coffee by near infrared spectroscopy. *Analytica Chimica Acta*, **509**, 217-227.
- Pomeranz, Y. & Williams, P.C. (1990). Wheat hardness: its genetic, structural, and biochemical background, measurement, and significance. In: *Advances in Cereal Science and Technology* (edited by Y. Pomeranz). Pp. 471-529. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Pontes, M.J.C., Santos, S.R.B., Araujo, M.C.U., Almeida, L.F., Lima, R.A.C., Gaia, E.N. & Souto, U.T.C.P. (2006). Classification of distilled alcoholic beverages and verification of adulteration by near infrared spectrometry. *Food Research International*, **39**, 182-189.
- Prieto, N., Andrés, S., Giráldez, F.J., Mantecón, A.R. & Lavín, P. (2008). Discrimination of adult steers (oxen) and young cattle ground meat samples by near infrared reflectance spectroscopy (NIRS). *Meat Science*, **79**, 198-201.
- Qina, J. & Lu, R. (2008). Measurement of the optical properties of fruits and vegetables using spatially resolved hyperspectral diffuse reflectance imaging technique. *Postharvest Biology and Technology*, **49**, 355-365.
- Quaresima, V., Lepantor, R. & Ferrari, M. (2003). The use of near infrared spectroscopy in sports medicine. *Journal of Sports Medicine and Physical Fitness*, **43**, 1-13.
- Reich, G. (2005). Near-infrared spectroscopy and imaging: basic principles and pharmaceutical applications. *Advanced Drug Delivery Reviews*, **57**, 1109-1143.
- Roggo, Y., Edmond, A., Chalus, P. & Ulmschneider, M. (2005). Infrared hyperspectral imaging for qualitative analysis of pharmaceutical solid forms. *Analytica Chimica Acta*, **535**, 79-87.
- Rohe, T., Becker, W., Kölle, S., Eisenreich, N. & Eyerer, P. (1999). Near infrared (NIR) spectroscopy for in-line monitoring of polymer extrusion processes. *Talanta*, **50**, 283-290.
- Sahlin, J.J. & Peppas, N.A. (1997). Near-field FTIR imaging: a technique for enhancing spatial resolution in FTIR microscopy. *Journal of Applied Polymer Science*, **63**, 103-110.
- Sakudo, A., Suganuma, Y., Kobayashi, T., Onodera, T. & Ikuta, K. (2006). Near-infrared spectroscopy: promising diagnostic tool for viral infections. *Biochemical and Biophysical Research Communications*, **341**, 279-284.
- Savitzky, A. & Golay, M. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, **36**, 1627-1639.
- Scherzera, T., Müllera, S., Mehnerta, R., Vollandb, A. & Lucht, H. (2005). In-line monitoring of the conversion in photopolymerized acrylate coatings on polymer foils using NIR spectroscopy. *Polymer*, **46**, 7072-7081.

- Schofield, J.D. & Greenwell, P. (1987). Wheat starch granule proteins and their technological significance. In: European Conference on Food Science and Technology (edited by I.D. Morton). Pp. 407-420. Bournemouth, England: Ellis Horwood.
- Seckinger, H.L., Wolf, M.J. & Dimler, R.J. (1964). Micro method for determining moisture distribution in wheat kernels, based on iodine staining. *Cereal Chemistry*, **41**, 80-87.
- Shahin, M.A. & Symons, S.J. (2008). Detection of hard vitreous and starchy kernels in amber durum wheat samples using hyperspectral imaging. *NIR News*, **19**, 16-18.
- Shewry, P.R., Field, J.M. & Tatham, A.S. (1987). The structures of cereal seed storage proteins. In: European Conference of Food Science and Technology (edited by I.D. Morton). Pp. 421-437. Bournemouth, England: Ellis Horwood.
- Shibata, S., Ohdan, H., Noriyuki, T., Yoshioka, S., Asahara, T. & Dohi, K. (1999). Novel assessment of acute lung injury by in vivo near-infrared spectroscopy. *American Journal of Respiratory and Critical Care Medicine*, **160**, 317-323.
- Siesler, H.W. (2002). Introduction. In: *Near Infrared Spectroscopy: Principles, Instruments, Applications* (edited by H.W. Siesler, Y. Ozaki, S. Kawata & H.M. Heise). Pp. 1-10. Weinheim, Germany: Wiley-VCH Verlag GmbH.
- Simmonds, D.H. (1974). Chemical basis of hardness and vitreosity in the wheat kernel. *Baker's Digest*, **48**, 16-29, 63.
- Simmonds, D.H., Barlow, K.K. & Wrigley, C.W. (1973). The biochemical basis of grain hardness in wheat. *Cereal Chemistry*, **50**, 553-562.
- Sinellia, N., Spinardib, A., Di Egidioa, V., Mignanib, I. & Casiraghi, E. (2008). Evaluation of quality and nutraceutical content of blueberries (*Vaccinium corymbosum* L.) by near and mid-infrared spectroscopy. *Postharvest Biology and Technology*, **50**, 31-36.
- Sissons, M., Osborne, B. & Sissons, S. (2006). Application of near infrared reflectance spectroscopy to a durum wheat breeding programme. *Journal of Near Infrared Spectroscopy*, **14**, 17-25.
- Smith, L. (1956). Objects of conditioning. In: *Flour Milling Technology* (edited by L. Smith). Pp. 115. Liverpool, England: Northern Publishing Co. Ltd.
- Spahn, G., Plettenberg, H., Nagel, H., Kahl, E., Klinger, H.M., Mückley, T., Günther, M., Hofmanne, G.O. & Mollenhauer, J.A. (2008). Evaluation of cartilage defects with near-infrared spectroscopy (NIR): an ex vivo study. *Medical Engineering & Physics*, **30**, 285-292.
- Stenvert, N.L. & Kingswood, K. (1976). An autoradiographic demonstration of the penetration of water into wheat during tempering. *Cereal Chemistry*, **53**, 141-149.
- Symes, K.J. (1965). The inheritance of grain hardness in wheat as measured by the particle size index. *Australian Journal of Agricultural Research*, **16**, 113-123.
- Taylor, J.R.N. (2004). Grain production and consumption, Africa. In: *Encyclopedia of Grain Science* (edited by C. Wrigley, H. Corke & C.E. Walker). **Vol. 3**. Pp. 71-77. Kidlington, Oxford, UK: Elsevier Ltd.
- Tippels, K.H., Kilborn, R.H. & Preston, K.R. (1994). Bread-wheat quality defined. In: *Wheat production, properties and quality* (edited by W. Bushuk & V.F. Rasper). Pp. 26-27. London: Chapman & Hall.
- Turnbull, K.M. & Rahman, S. (2002). Endosperm texture in wheat. *Journal of Cereal Science*, **36**, 327-337.
- Valdes, E.V., Young, L.G., Leeson, S., McMillan, I., Portela, F. & Winch, J.E. (1985). Application of near infrared reflectance spectroscopy to analyses of poultry feeds. *Poultry Science*, **64**, 2136-2142.

- Vo-Dinh, T., Stokes, D.L., Wabuyele, M.B., Martin, M.E., Song, J.M., Jagannathan, R., Michaud, E., Lee, R.J. & Pan, X. (2004). A hyperspectral imaging system for in vivo optical diagnostics. *IEEE Engineering in Medicine and Biology Magazine*, **23**, 40-49.
- Wahrenberger, H. (2004). Advanced milling course. Uzwil, Switzerland: Buhler.
- Wang, L., Lee, F.S.C., Wang, X. & He, Y. (2006). Feasibility study of quantifying and discriminating soybean oil adulteration in camellia oils by attenuated total reflectance MIR and fiber optic diffuse reflectance NIR. *Food Chemistry*, **95**, 529-536.
- Weinstock, B.A., Janni, J., Hagen, L. & Wright, S. (2006). Prediction of oil and oleic acid concentrations in individual corn (*Zea mays* L.) kernels using near-infrared reflectance hyperspectral imaging and multivariate analysis. *Applied Spectroscopy*, **60**, 9-16.
- Williams, P.C. (1991). Prediction of wheat kernel texture in whole grains by near-infrared transmittance. *Cereal Chemistry*, **68**, 112-114.
- Williams, P.C., Norris, K.H. & Zarowski, W.S. (1982). Influence of temperature on estimation of protein and moisture in wheat by near-infrared reflectance. *Cereal Chemistry*, **59**, 473-477.
- Williams, P.C. & Sobering, D.C. (1986). Attempts at standardization of hardness testing of wheat. I. the grinding/sieving (particle size index) method. *Cereal Foods World*, **31**, 359.
- Windham, W.R., Gaines, C.S. & Leffler, R.G. (1993). Effect of wheat moisture content on hardness scores determined by near-infrared reflectance and on hardness score standardization. *Cereal Chemistry*, **70**, 662-666.
- Wold, H. (1982). Soft modeling. the basic design and some extensions. In: *Systems Under Indirect Observation* (edited by J. K.-G. & W. H.). Pp. 1-53. Amsterdam: New Holland Publishing.
- Wood, K.S., Gulian, A.M., Fritz, G.G. & Van Vechten, D. (2002). A QVD detector for focal plane hyperspectral imaging in astronomy. *Bulletin of the American Astronomical Society*, **34**, 1241.
- Worzella, W.W. & Cutler, G.H. (1939). A critical study of techniques for measuring granulation in wheat meal. *Journal of Agricultural Research*, **58**, 329-341.
- Yamazaki, W. & Donelson, J.R. (1983). Kernel hardness of some US wheats. *Cereal Chemistry*, **60**, 344.
- Yan, S.H. (2007). NIR evaluation of the quality of tea and its marketplace. *Spectroscopy Europe*, **19**, 16-19.



## **CHAPTER 3**

**Determination of wheat kernel hardness by near infrared hyperspectral imaging and hyperspectral image analysis**

# Determination of wheat kernel hardness by near infrared hyperspectral imaging and hyperspectral image analysis

## Abstract

The use of near infrared (NIR) hyperspectral imaging to distinguish between whole wheat kernels differing in hardness (durum, hard bread, soft white) were evaluated. NIR hyperspectral images were acquired with a Spectral Dimensions MatrixNIR focal plane array imaging system in the spectral range of 960 to 1662 nm.

Exploratory principal component analysis (PCA) was used to clean the image, which entailed removing background, dead pixels, as well as shading and curvature errors. PCA indicated a clear separation of soft endosperm from hard and durum endosperm in principal component (PC) 2, while a clear distinction of durum from hard and soft endosperm could explain the variation in PC 3. Studying PC loading line plots, the largest variation due to differences in wheat hardness could be explained in PC 2, based on the separation between the absorption peaks of starch (1450 nm) and protein (1570 nm) in different hardness endosperm classes.

Partial least squares discriminant analysis (PLS-DA) was performed to determine the extent to which different hardness endosperm could be predicted using NIR hyperspectral imaging. Two PLS components were adequate to explain 91% of the Y variation in the model of soft versus durum endosperm with 100% prediction accuracy. The model of soft versus hard endosperm required two PLS components to explain 75% of the Y variation, with 98% accuracy when predicting hard or soft endosperm. Three PLS components were adequate to explain 67% of the Y variation in the hard versus durum endosperm model, with 96% accuracy when predicting durum endosperm and 95% accuracy when predicting hard endosperm.

## Introduction

Wheat used for food applications consist mainly of two species, i.e. bread (*Triticum aestivum*) and durum wheat (*T. turgidum*), which are characterised by different chemical and physical properties (Kent & Evers, 1994; O'Brien & DePauw, 2004). Based on these different chemical and physical properties, wheat flour will differ in functional quality, nutritional contribution and consequently commercial value (Bietz, 1989). Bread wheat can further be divided into different hardness classes, which influence the baking quality thereof. Soft wheat is commonly used for biscuits, while hard bread wheat is commonly used for bread (Kent & Evers, 1994; Oleson, 1994).

Wheat kernel hardness is one of the physical characteristics of wheat that influences the conditioning process of the wheat, flour yield and particle size of the flour, the shape and size of the flour particles, and ultimately the end use properties of the flour (Pomeranz & Williams, 1990). Hard wheat produces flour with significantly higher water absorption than soft wheat flours at an

equivalent level of protein. This characteristic directly determines the amount of bread to be produced from a given weight of flour (Tippels *et al.*, 1994).

Wheat hardness is caused by a genetically controlled strength of the protein-starch bond in the wheat endosperm (Simmonds *et al.*, 1973; Simmonds, 1974; Greenwell & Schofield, 1986; Schofield & Greenwell, 1987; Hosney, 1994b). Friabilin is a low molecular weight protein (15000 Da) that is situated on the surface of starch granules, thus influencing the bond between starch and the protein matrix (Greenwell & Schofield, 1986). Friabilin is present in highest amounts in soft wheat, less in hard wheat and completely absent in durum wheat. Hard wheat has starch granules bound tighter in the protein matrix than soft wheat with a weaker bond between the protein matrix and starch granules (Greenwell & Schofield, 1986; Schofield & Greenwell, 1987). Many factors such as moisture content, lipid content, water soluble materials such as carbohydrates, protein and pentosans influence wheat hardness determinations, but are not the causing factor of wheat hardness (Turnbull & Rahman, 2002).

Acceptance of a wheat load as well as payment to the producer depends on various grading factors, including wheat hardness, it is therefore important that this factor be determined accurately and rapidly (Osborne, 1991). Wheat hardness determination is presently not one of the grading factors for wheat in South Africa, however taking into consideration the enormous effect it has on the dry milling process and the outcomes thereof, it would be beneficial to consider it for future applications.

Most methods for the determination of wheat hardness are based on sieving, grinding resistance, vitreousness and near infrared (NIR) methods and are dependent on particle size (Osborne, 1991). Methods currently used to measure hardness include the particle size index (PSI) (AACC Approved Method 55-30.01) (Osborne *et al.*, 2001; AACC, 2009b), NIR spectroscopy on ground wheat (AACC Approved Method 39-70.02) (Downey *et al.*, 1986; Norris *et al.*, 1989; AACC, 2009a), NIR spectroscopy on whole wheat (Downey *et al.*, 1986), the single kernel characterisation system (SKCS) (AACC Approved Method 55-31.01) (Worzella & Cutler, 1939; AACC, 2009c), and the vitreousness cutter or farinator. All of the methods currently used for hardness determination are destructive. There is, however, a NIR spectroscopy method on whole grain samples for determining wheat hardness, including a research study conducted using NIR transmittance spectroscopy to determine hardness in single kernels (Nielsen, 2002). The NIR method on whole grain samples (Downey *et al.*, 1986) does not allow for the hardness determination of single wheat kernels, neither for the visualisation of chemical distribution within a single wheat kernel. The study on hardness determination of single wheat kernels (Nielsen, 2002) did not provide good prediction of hardness using near infrared transmittance, thus leaving the possibility of NIR hyperspectral imaging to provide good results on single kernels.

The NIR region refers to the spectral region ranging from 750 to 2500 nm. NIR data are acquired in the form of transmission or reflectance intensity related to the stretching and bending vibrations of O-H, C-H, N-H and S-H chemical bonds in the sample being analysed (Miller, 2001).

A hyperspectral image is a three dimensional data cube consisting of a certain amount of pixels in the x and y space, with a certain amount of wavelengths in the z space, depending on the camera used for the image acquisition. An enormous amount of NIR spectra are thus captured in one NIR hyperspectral image, with both spectral and spatial information of the sample analysed present in the image (Koehler IV *et al.*, 2002). Each pixel of a hyperspectral image thus contains the spectral information for that position (Gowen *et al.*, 2007). Multivariate image analysis software packages are used to extract the appropriate and useful information from a hyperspectral image

NIR hyperspectral imaging has been used widely in the fields of food science and agriculture, of which the detection of faecal contamination on apples (Liu *et al.*, 2006); evaluation of frying oil quality (Kazemi *et al.*, 2005); prediction of oil and oleic acid content in single maize kernels (Weinstock *et al.*, 2006); and the control and monitoring of food and feed products (Fernández Pierna *et al.*, 2006) are but a few examples. Hyperspectral imaging has also been applied specifically to wheat in the detection of vitreous and starchy kernels in amber durum showing bands of difference corresponding to starch and/or protein between starchy and vitreous kernels (Shahin & Symons, 2008); classification of sound and stained wheat grains with overall classification accuracies of 95% (Berman *et al.*, 2007); determining vitreousness of durum wheat kernels with a classification rate reaching up to 94% for the discrimination of total and partially starchy kernel classes (Gorretta *et al.*, 2006); investigating the feasibility to differentiate between Canadian wheat classes with classification accuracies of 86 to 100% (Mahesh *et al.*, 2008); and the determination of wheat pre-germination (Smail *et al.*, 2006; Koç *et al.*, 2008; Singh *et al.*, 2009). It is important to use non-destructive methods, such as NIR spectroscopy and NIR hyperspectral imaging, to determine wheat hardness in breeding programmes, as wheat kernels that are of interest should remain intact to be used in further breeding trails.

Plant breeders are constantly trying to improve the yield, disease resistance and quality of new cultivars. The quality of the crop is important to the farmer as it provides the best wheat grade and economic return. The miller requires the best flour yield, while the baker requires good performance during processing to provide an acceptable end product to the consumer (Sissons *et al.*, 2006). Wheat hardness is thus an important physical characteristic influencing quality, and should be considered in wheat breeding programmes.

Due to the non-destructive nature of NIR hyperspectral imaging and the importance of determining wheat hardness in a non-destructive manner during the breeding phase of wheat for further breeding use, this new science could have promising outcomes for the detection of wheat hardness. The objective of this research was thus to evaluate the use of NIR hyperspectral imaging and multivariate image analysis as a non-destructive method of determining wheat hardness of whole wheat kernels of different hardness (soft, hard and durum) over the wavelength range of 960 to 1662 nm, and to determine to what extent endosperm hardness could be predicted by applying PLS-DA.

## Materials and Methods

### *Samples*

Six wheat cultivars varying in degree of hardness were obtained. These wheat cultivars comprised Canada Western Amber Durum (CWAD), South African Durum, Canada Western Red Spring (CWRS), South African Tankwa (hard bread cultivar), Canada Eastern White Winter (CEWW), and South African Wit Wol (soft white cultivar). The three Canadian wheat samples had known particle size index (PSI) values of 71.0 for CEWW, 52.0 for CWRS and 36.7 for CWAD, with PSI values determined according to the method described by (Manley *et al.*, 2002). The three samples with known PSI values thus served as reference samples for wheat hardness determinations. Six sets of eight whole wheat kernels were randomly selected from three hardness categories, i.e. soft, hard and durum wheat.

### *Near infrared hyperspectral imaging system*

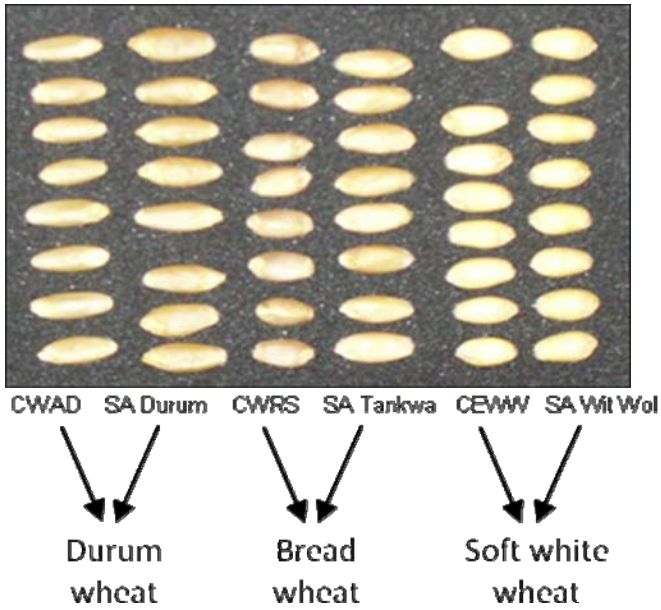
Images were acquired with the Spectral Dimensions MatrixNIR focal plane array imaging system (Malvern Instruments Ltd., Malvern, Worcestershire, UK). The MatrixNIR imaging system consisted of an Indium Gallium Arsenide (InGaAs) camera combined with a monochromator called liquid crystal tuneable band pass filter (LCTF). The field of view of the camera was 50 mm x 62 mm. Four quartz-halogen lamps with current control provided illumination. The spectral range varied from 960 to 1662 nm, with 6 nm intervals and 6 nm spectral resolution. Each channel was scanned 16 times and co-added to reduce noise. This produced hyperspectral images with 256 × 320 pixels at 118 wavelength channels (81920 pixels or spectra in total). The integration time and lamp strength were adjusted in order to avoid saturation of the A/D converter .

### *Image acquisition*

Images of wheat types differing in hardness were acquired on silicon carbide (SiC) sandpaper, at room temperature. Wheat kernels were placed on the sandpaper crease side down, to avoid possible shading errors and curvature effects.

Eight kernels of six different wheat cultivars were placed on the sandpaper; each cultivar formed a vertical column as can be seen in **Fig. 3.1**. Prior to image acquisition of the wheat samples, images of reflectance standards (Labsphere, USA) were acquired, i.e. a 2% background spectralon (dark reference) as well as a 50% white reference spectralon.

The image of different wheat samples was then acquired using MatrixAquire software (Malvern Instruments Ltd, Malvern, Worcestershire, UK). The total image acquisition time was 8 min per image or reference standard at an integration time of 128 ms, and the wheat samples were kept cool with a small box fan positioned just beyond the field of view.



**Figure 3.1** Digital image of whole wheat kernels of three different wheat types and six different cultivars imaged with the MatrixNIR. This digital image was acquired with a Samsung Digimax A40 camera, with 4.0 mega pixels and the flash switched on.

### *Hyperspectral image analysis*

#### Image correction

MatrixNIR images were transformed from instrument measurement counts (A/D converter counts) to absorbance in Isys 4.0 multivariate image analysis software (Malvern Instruments Ltd, UK) using **equation 3.1**. The transformation and standardisation involved correcting for dark counts with a 2% reflectance standard (dark reference image) subtracted from the raw image, and subsequently dividing by a total reflectance spectrum of a 2% reflectance standard subtracted from a 50% reflectance standard (white reference image) which served as white reference. Corrected images were saved as Matlab files for further analysis in Evince 2.3.0 (UmBio AB, Umeå, Sweden).

$$I_{\lambda,n} = -\log \left[ \left( \frac{S_{\lambda,n} - B_{\lambda,n}}{W_{\lambda,n} - B_{\lambda,n}} \right) \times 0.50 \right] \quad \text{equation 3.1}$$

Where:

$n$  = pixel index variable ( $n = 1 \dots N$ )

$I_{\lambda,n}$  = standardised absorbance intensity, pixel  $n$ , at wavelength  $\lambda$

$S_{\lambda,n}$  = sample image, pixel  $n$ , at wavelength  $\lambda$

$B_{\lambda,n}$  = dark reference image (2% reference standard), pixel  $n$ , at wavelength  $\lambda$

$W_{\lambda,n}$  = white reference image (50% reference standard), pixel  $n$ , at wavelength  $\lambda$

0.50 = total reflectance of the standard used.

In addition to the image standardisation, image transformation was also performed with a Fourier transform filter applied to the total reflectance spectrum. This was done with an exponential low pass filter to remove lines and lighting errors from the background to provide a more homogenous total reflectance image.

#### Image cleaning

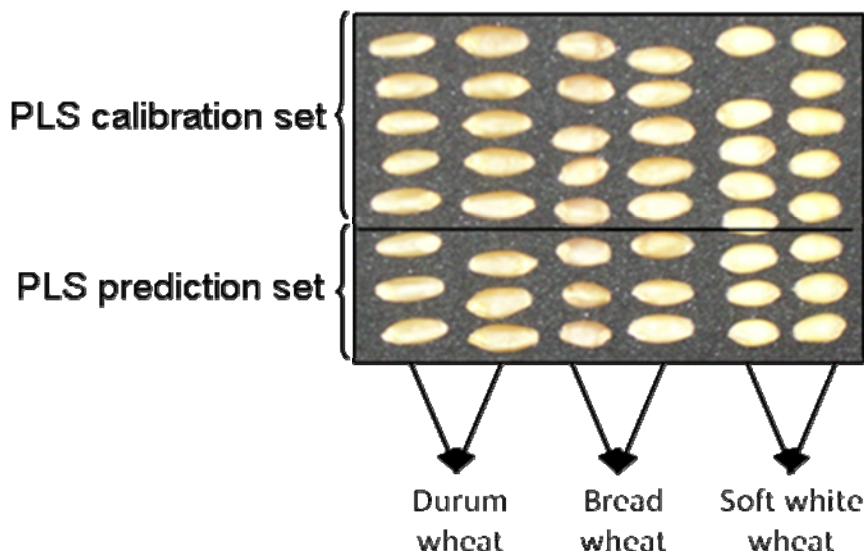
Absorbance images were cleaned using principal component analysis (PCA) with mean centering. Three principal components (PCs) were used for cleaning the image, which involved utilisation of PCA score plots and PCA score images interactively for the identification and classification of outlying pixels, dead pixels, shading errors, background and edge effects. Unwanted pixels were classified and removed. The cleaning was performed in Evince 2.3.0 multivariate image analysis software and after removing these unwanted pixels PCA was recalculated with the remaining pixels. The cleaned image was subjected to further image analysis.

#### Image analysis of cleaned image

After obtaining a cleaned absorbance image, supplementary PCs were added. Savitzky-Golay smoothing with polynomial order 3, derivative order 0 and 7 left and right points were applied to the dataset. Identification and classification of the prominent clusters in the score plot were done using the interactive nature of the score plot projected onto the score image to form a classification image. This process is called brushing (Esbensen & Geladi, 1989). Principal component (PC) loading line plots of the individual clusters were studied in Evince 2.3.0 to explain the variance in each PC contributed by the different clusters.

#### Partial least squares discriminant analysis of image

Partial least squares discriminant analysis (PLS-DA) was performed on the MatrixNIR data to determine the possibility of predicting endosperm hardness. When using a discriminant analysis method such as PLS-DA only two classes can be modelled simultaneously. Three combinations of the three hardness classes were thus modelled, namely soft and durum endosperm; soft and hard endosperm; and hard and durum endosperm. In each of these combinations the third class and all pixels not classified were excluded from the image before the discriminant analysis was performed. **Fig. 3.2** illustrates the top five rows of wheat samples used as the calibration set, and the bottom three rows as the prediction set for the PLS-DA performed. All the pixels allocated to the respective classes were used in the PLS\_DA models. It would have been preferable to perform prediction on a separate, new image. This, however, was not possible at the time of data analysis due to software limitations.



**Figure 3.2** Digital image of wheat kernels used for NIR hyperspectral imaging to distinguish between wheat types, and the sections used as calibration and prediction sets for PLS-DA.

## Results and discussion

### *Hyperspectral image analysis*

#### Image correction

Instrument signal transformation and standardisation are common in NIR spectroscopy, which can also be extended to hyperspectral imaging. The raw instrument signal obtained from image acquisition has to be transformed to ensure that a uniform response has been achieved for all pixel locations at all wavelengths in the hyperspectral image (Burger & Geladi, 2007). Variation between day-to-day image acquisition and image-to-image variation also needs to be corrected for by the use of spectralon standards. Spatial variation in illumination and other system responses were corrected for, using this independent reflectance transform model.

The A/D converter counts obtained during image acquisition are arbitrary units defined by the integration time and lamp strength, and has no meaning unless converted to reflectance (Geladi *et al.*, 2004). These reflectance values are converted to absorbance values to allow for the spectral comparison with known absorbance peaks assigned to chemical constituents.

#### Image cleaning

After PCA, with mean-centering, was performed, the three PCs calculated, amounted to a total of 97.97% of the total sum of squares (SS). The PCA score plot with three PCs were used in conjunction with the two dimensional score image to identify and remove bad pixels (background, dead pixels, shading errors, side effects and outlying pixels). **Fig. 3.3** illustrates how the score plot (**Fig. 3.3a**) of the absorbance image with 81920 pixels was used interactively with the score image (**Fig. 3.3b**) to identify unwanted regions in the image. The unwanted regions were identified in the score images and score plots as indicated on the classification plot (**Fig. 3.3c**) and classification



image (**Fig. 3.3d**). These regions were identified as background and bad pixels (green), shading and curvature effects (blue) and wheat kernels (red). Pixels with similar chemical or physical characteristics will originate from the same region in the PCA score plot, thus forming clusters. The connection between data points on the score plot with pixels in the score image is critical in multivariate image analysis (Geladi *et al.*, 2007). This interactivity allows for the identification and classification of biochemical constituents and subsequent removal of unwanted clusters from the image.

After removal of the background, bad pixels and shading errors pixels, the PCA was recalculated, and three additional PCs were added. The resulting cleaned image (**Fig. 3.4a**) could be used further to interpret chemical differences in the wheat samples.

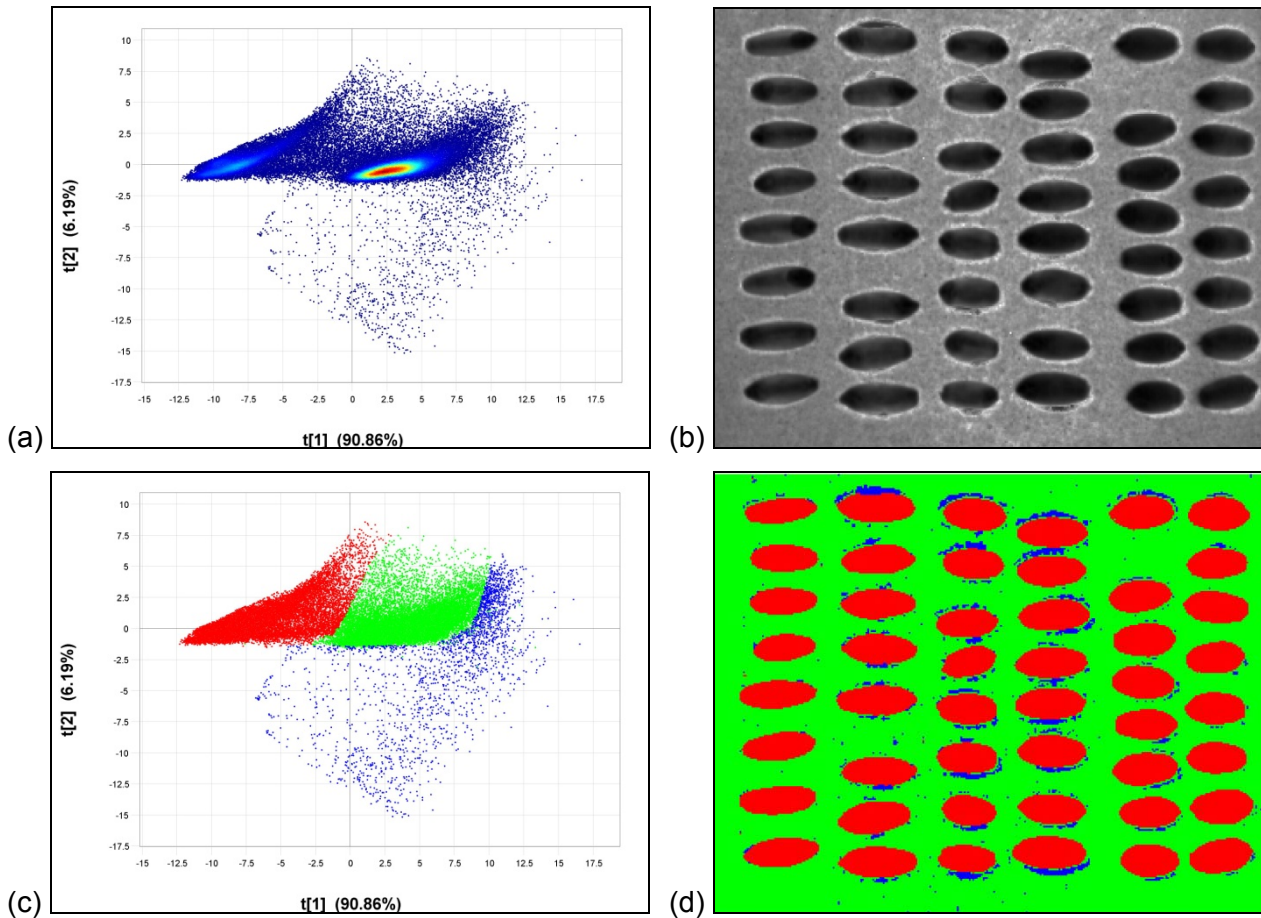
#### Image analysis of cleaned image

The cleaned score image (**Fig. 3.4a**) and score plots with 24058 pixels remaining were used for further classification of the wheat samples. All possible combinations of PCs were plotted in the score plot, to get an overview of the compound data structure (Esbensen, 2006). There was no separation in the score plot due to PC 1 (93.4%) when plotted against PC 2 (5.67%) or any of the other PCs (**Fig. 3.4b-d**). However there were clusters due to the variation in PC 2 which were classified as soft wheat endosperm, and durum and hard wheat endosperm. When using the brushing technique (Esbensen & Geladi, 1989; Geladi & Grahn, 1996; Esbensen & Lied, 2007; Manley *et al.*, 2009) in the score plot and score image, the classification of these clusters were possible.

When mean centered data are used, the variation in PC 1 is most often due to physical characteristics of the sample or a typical average spectrum of the sample, and not due to chemical composition explained in the PC 1 (Esbensen, 2006). Studying the intensity differences in the score image of PC 1 (**Fig. 3.4a**) it was apparent that the variation in PC 1 is attributed to the difference in scattering of light due to the wheat endosperm and the wheat germ. This is a physical characteristic not useful in distinguishing between wheat of different hardness, as wheat hardness is caused by the friabilin content in the endosperm (Greenwell & Schofield, 1986; Schofield & Greenwell, 1987). Combinations of higher order PCs were investigated. This was used for the interpretation of the chemical difference between the samples.

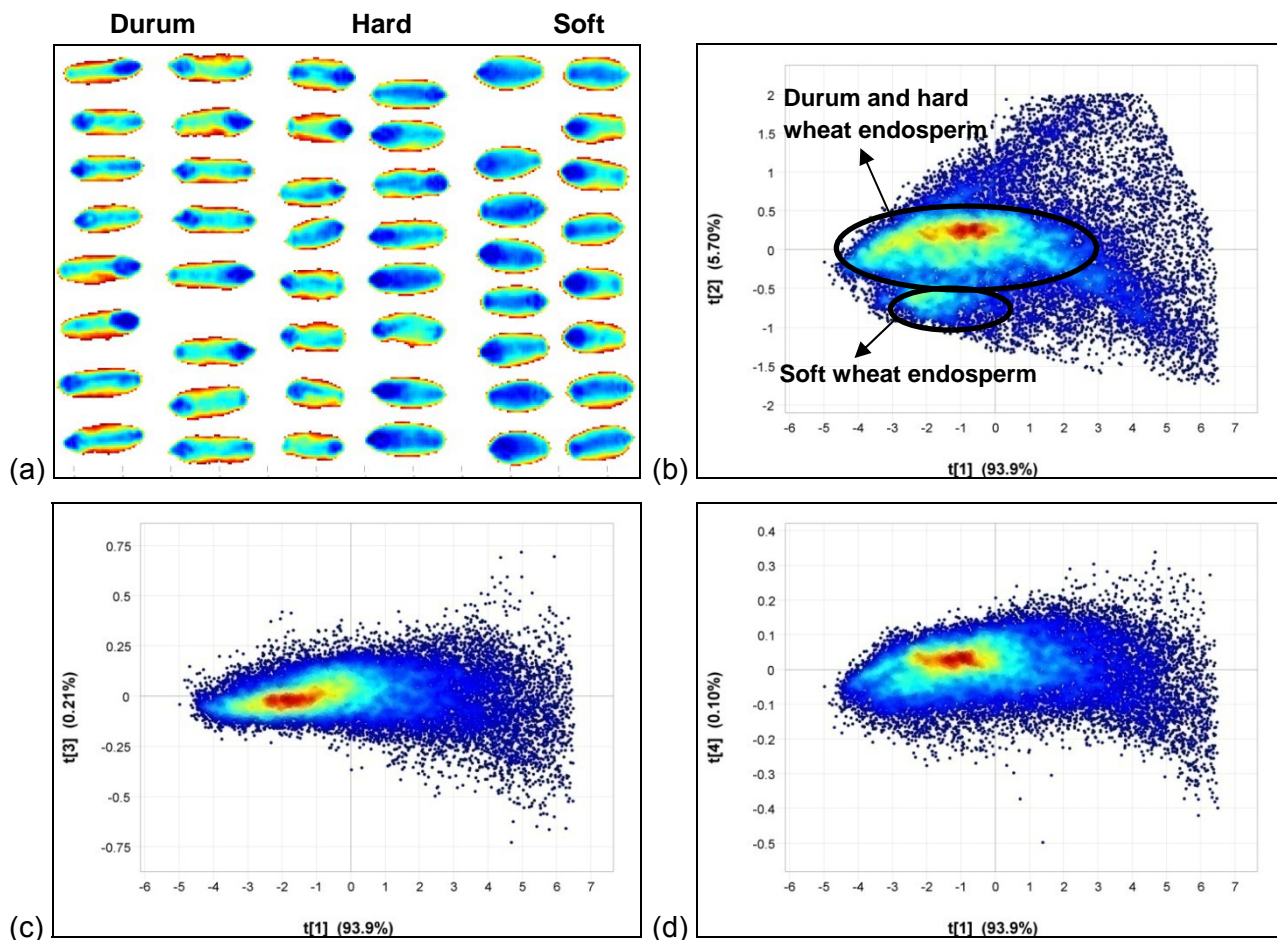
Investigating combinations of higher order PCs (i.e. PC 2 vs PC 3; PC 2 vs PC 4; ..., PC 5 vs PC 6) that were plotted against each other; only the score plot of PC 2 and PC 3 provided useful results. Three clusters could be clearly distinguished in the score plot (**Fig. 3.5a**). The clusters were identified with aid of the interactive nature of the score plot and score image, i.e. brushing (Manley *et al.*, 2009). Clusters selected in the score plot (**Fig. 3.5a**) would be highlighted in the score image (**Fig. 3.5b**), where the location of those pixels could be visualised. By using the information of the location of the pixels from the three clusters in the score image, and the known

PSI values of the Canadian wheat types, the three clusters could be identified as durum, hard and soft wheat.



**Figure 3.3** (a) PCA score plot (PC 1 vs PC 2) of an uncleaned absorbance image (81920 pixels); (b) Score image of PC 1 of the uncleaned image; (c) classification plot and (d) classification image (green = SiC sandpaper background and bad pixels; blue = shading errors and side effects, red = wheat kernels).

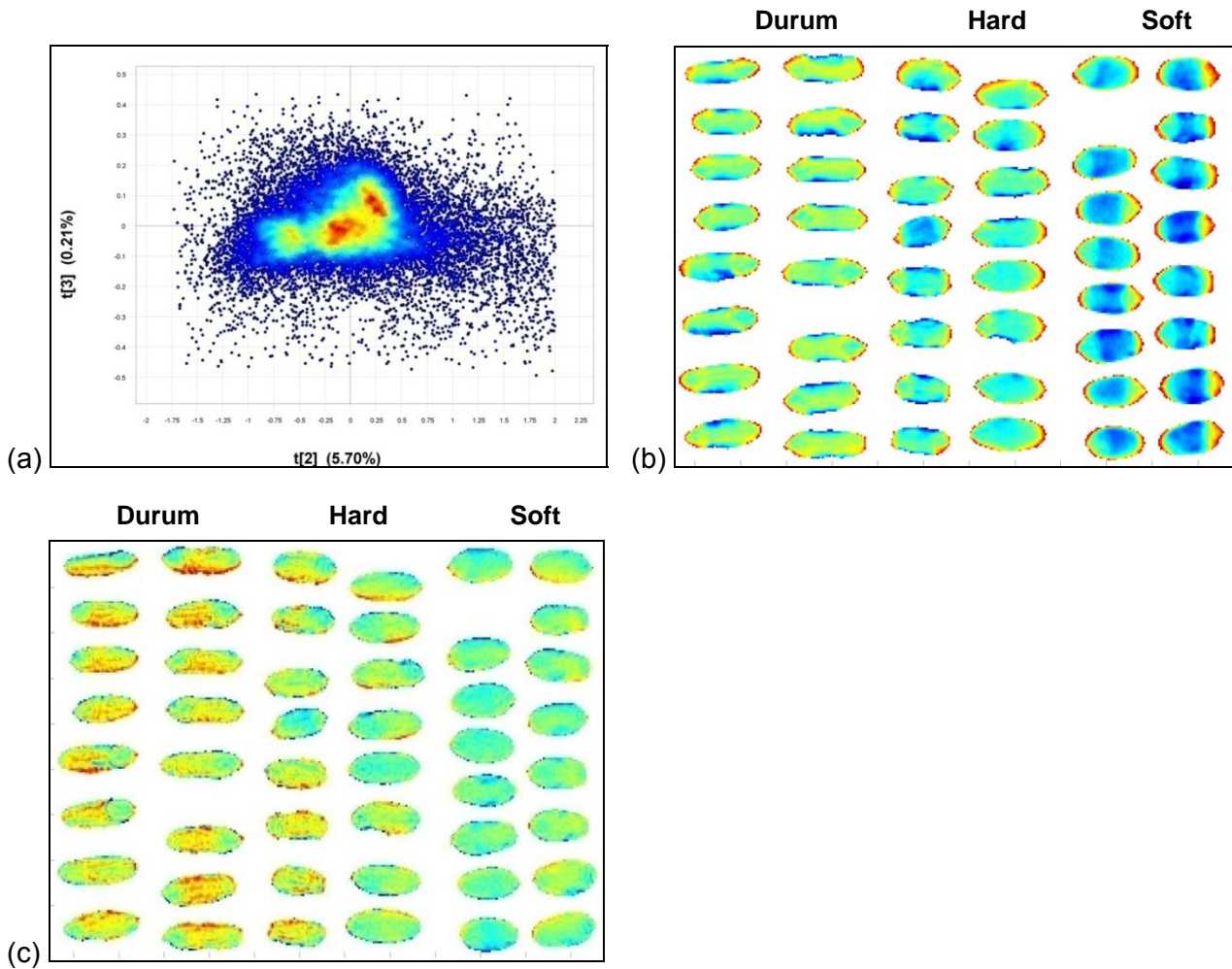
The three clusters were classified as durum, hard and soft wheat endosperm in the score plot (**Fig. 3.5a**) to form a classification plot (**Fig. 3.6a**). The classification plot was projected onto the score image (**Fig. 3.5b**) to form a classification image (**Fig. 3.6b**). Green indicates the soft endosperm, yellow the hard endosperm, blue the durum endosperm and red those pixels that were not assigned to a specific class. The durum endosperm could mainly be observed in the durum wheat, while hard endosperm could mainly be observed in the hard wheat and soft endosperm in the soft wheat. The chemical difference between some of the kernels, from the different hardness groups, could be closely related; therefore their regions of origination in the hypercube would be in close proximity (Gowen *et al.*, 2007). This resulted in an overlapping between the classes as can be seen in the classification image (**Fig. 3.6b**). The difference in wheat hardness between the durum, hard and soft wheat is the binding of starch to the protein matrix in the presence or absence of the protein associated with wheat hardness (friabilin) (Greenwell & Schofield, 1986; Hoseney, 1994b).



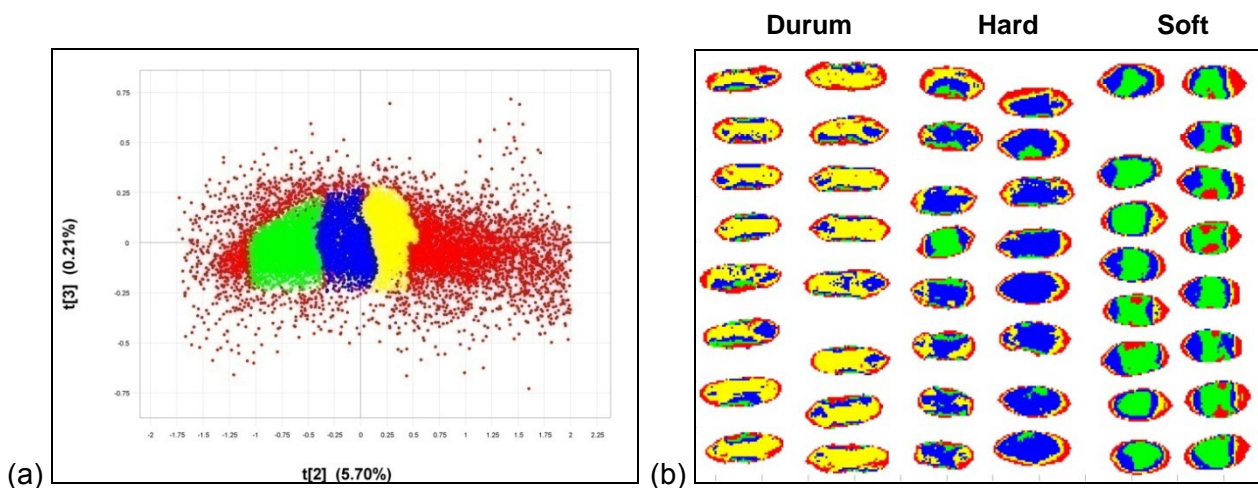
**Figure 3.4** (a) Cleaned score image for PC 1 (24 904 pixels remaining); (b) PCA score plot of PC 1 vs PC 2, (c) PC 1 vs PC 3, and (d) PC 1 vs PC 4.

The intensity difference of pixels in the score image indicated separation in the data swarm, which could be explained due to the chemical variation in the particular PC (Geladi & Grahn, 1996). When studying the intensity differences in the score image of PC 2 (**Fig. 3.5b**) it is clear that a separation of soft endosperm from hard and durum endosperm took place in this PC. A separation of durum endosperm from hard and soft endosperm was apparent in the score image of PC 3 (**Fig. 3.5c**). Investigation of the score plot of PC 2 vs PC 3 (**Fig. 3.5a**) also concurred with these observations from the score images. Separation of soft from hard and durum endosperm occurred in PC 2; and durum from hard and soft endosperm in PC 3.

Differences in the hardness of the wheat were expected to be ascribed to starch and protein. The binding of starch molecules, due to the presence or absence of friabilin, to the protein matrix in wheat endosperm distinguished hardness classes of wheat. The content of starch and protein in wheat is also inversely related, with soft wheat having high starch and low protein content.



**Figure 3.5** (a) Cleaned PCA score plot (PC 2 vs PC 3) (24058 remaining pixels) with accompanied score image of (b) PC 2 and (c) PC 3.



**Figure 3.6** (a) PCA score plot (PC 2 vs PC 3) with classes of soft, hard and durum endosperm; and (b) with the clusters projected onto the score image, forming a classification image (green = soft endosperm; blue = hard endosperm; yellow = durum endosperm and red = unclassified).

The differences between different classes of endosperm hardness were investigated by studying each of the classes individually. Analyses by PCA were applied to each of the three classes individually. The loading line plots of the three classes were compared to determine the chemical difference between the hardness classes.

The loading line plot of PC 1 (**Fig. 3.7a**) for all three hardness classes gave an average spectrum of wheat as expected in the first component (Esbensen & Lied, 2007). PC 1 did thus not contain any chemical information useful in differentiating between endosperm hardness of different wheat types. PC 1 for soft endosperm accounted for 99.38% of the total SS, hard endosperm for 99.43% and durum endosperm for 99.16%.

Loadings line plots of PC 2 for durum endosperm (4.88% of total SS), hard endosperm (3.85% of total SS) and soft endosperm (4.85% of total SS) illustrated variation ascribed to starch, moisture and protein. All three hardness classes in PC 2 (**Fig. 3.7b**) showed an absorption peak at 1195 nm (C-H stretching vibrations, second overtone) indicating the presence of a CH<sub>2</sub> bond which in turn is related to starch. Loading line plots for all three classes in PC 2 (**Fig. 3.7b**) indicated an absorption peak at 1450 nm (O-H stretching vibrations, second overtone) indicating the variation due to starch and moisture. A third absorption peak was present in the loading line plots of all hardness classes for PC 2 (**Fig. 3.7b**) at 1570 nm (N-H stretching vibrations, first overtone) which indicated the presence of a -CONH- bond related to protein. These results agree with literature, as the two major constituents of wheat are starch and protein with starch being present in the highest concentration (Orth & Shellenberger, 1988).

The starch peak at 1450 nm indicated a decrease in absorption with an increase in endosperm hardness, while the protein peak at 1570 nm indicated an increase in absorption with an increase in endosperm hardness. The protein content of wheat appears to be inversely related to the starch content (Hopkins & Graham, 1935). Soft wheat cultivars in general have a higher starch content than hard wheat cultivars, because of their lower protein content (Miller, 1974). This concurs with the observed decrease in starch absorption and increase in protein absorption, with an increase in endosperm hardness. The variation in starch and protein were expected in PC 2 since the score plot of PC 2 indicated separation between different hardness endosperm clusters. The results obtained corresponded favourably with the study of Shahin and Symons (2008) that determined the absorbance bands of starch and protein as regions of interest to differentiate between starchy and vitreous kernels.

The loading line plots of PC 3 (**Fig. 3.7c**) for soft endosperm (0.66% of total SS) and durum endosperm (1.68% of total SS) exhibited variation due to starch and/or moisture. This was observed with an absorption peak at 1450 nm (H<sub>2</sub>O and starch) due to an O-H stretch, first overtone in this region. The loading line plot of PC 3 for hard endosperm (0.89% of total SS) exhibited variation due to starch. This was indicated by an absorption peak at 1195 nm (C-H stretching vibrations, second overtone) and 1395 nm (2 x C-H stretch and C-H deformation).

The variation in PC 4 (**Fig. 3.7d**) was attributed to starch and moisture for soft and durum endosperm. This was indicated by an absorption peak at 1195 nm (C-H stretching vibrations, second overtone) and 1395 nm (2 x C-H stretch and C-H deformation) which indicated the presence of a CH<sub>2</sub> bond which in turn is related to starch. The variation in PC 4 for hard endosperm was due to a starch-moisture contrast. This was indicated by a starch peak at 1395 nm (2 x C-H stretch and C-H deformation) and a moisture peak at 1450 nm (O-H stretching vibrations, second overtone). For durum endosperm PC 4 accounted for 1.35% of the total SS, hard endosperm for 0.69% of the total SS and soft endosperm for 4.39% of the total SS. Detector errors were found in higher PCs, and they were thus not used in analysing the image.

The hardness of wheat depends on the absence or presence of friabilin in the wheat endosperm. The loading line plot of PC 2 clearly indicated that variation could be attributed due to protein and starch. It is not clear if this could be linked to the presence or absence of friabilin, and should be evaluated in future studies. Hyperspectral image analysis would have to be performed in conjunction with a reference technique, such as SDS-PAGE, to determine the absence or presence of friabilin. Nevertheless the results obtained indicated that an imaging instrument that covers the wavelength range of 1450 and 1570 nm be used to differentiate between wheat kernels differing in hardness.

#### Partial least squares discriminant analysis of image

Three different PLS-DA models were investigated to determine the possibility of predicting endosperm hardness. These three combinations were soft versus hard endosperm, soft versus durum endosperm, and hard versus durum endosperm.

#### Soft vs durum endosperm

The Y variation explained for the PLS-DA model was 21% in the first PLS component and 91% after two components, thereafter no further increase in Y variation explained was observed. Two components would thus be adequate to explain the variation in the model.

The amount of pixels in each class assigned to the training and test sets of soft and durum endosperm can be seen in **Table 3.1**. The PLS-DA test set comprised of 4444 pixels, of which 1771 were assigned as soft endosperm and 2673 pixels as durum endosperm. Of the 1771 pixels assigned as soft endosperm 1771 (100%) were correctly predicted. All of the pixels assigned as durum endosperm (2673) were correctly predicted. These results indicate that the classes were correctly assigned and that discrimination between soft and durum endosperm is possible using NIR hyperspectral imaging.

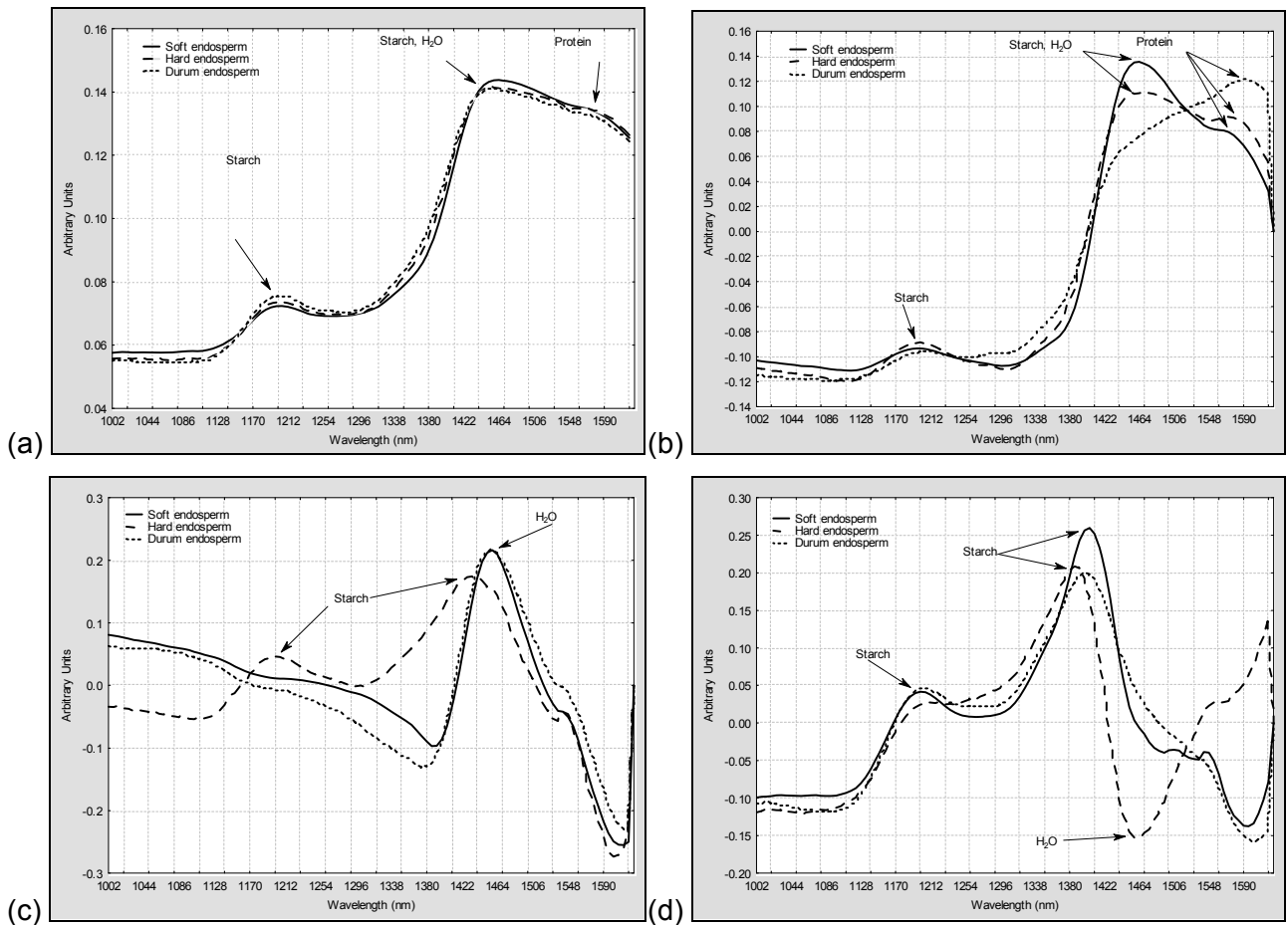
**Table 3.1** The amount of pixels included in the training and test sets of PLS-DA models using the three hardness classes biscuit, bread and durum endosperm

	Soft endosperm	Bread endosperm	Durum endosperm
Training set	2988	4838	3950
Test set	1746	2916	2673
Total	4734	7754	2673

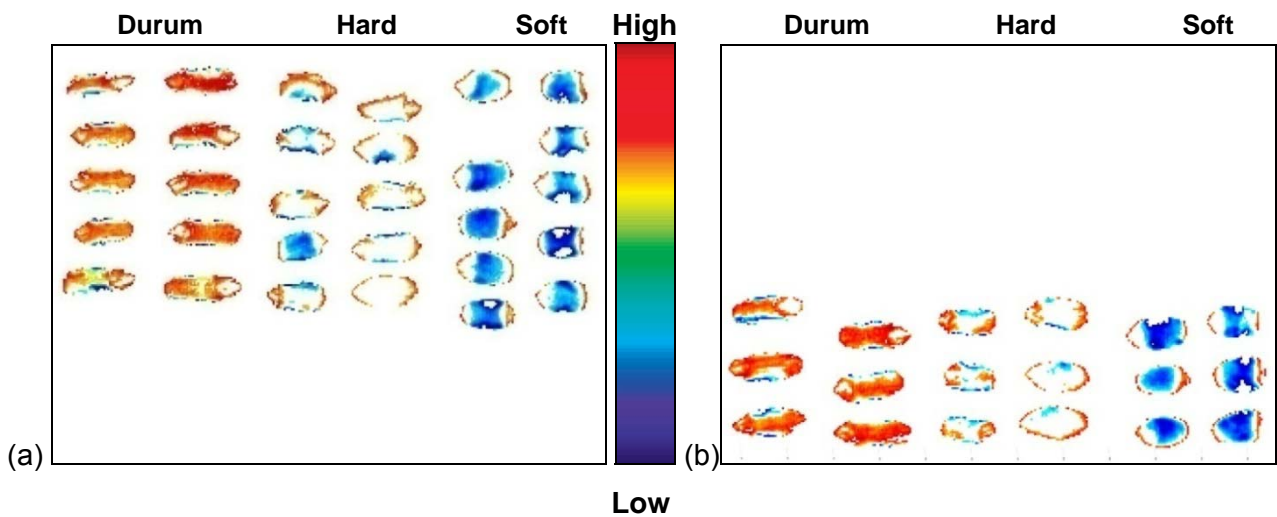
The calibration (**Fig. 3.8a**) and prediction images (**Fig. 3.8b**) of durum endosperm illustrates the localisation of the modelled and predicted classes after two PLS components. The intensity of the colour indicates how well the characteristic was modelled or predicted, in this instance durum endosperm. Thus in this instance the warmer the colour the better durum endosperm was modelled, while the cooler the colour the better soft endosperm was modelled or predicted. The calibration image (**Fig. 3.8a**) gives an indication of the model's ability to represent durum endosperm, while the prediction image (**Fig. 3.8b**) gives an indication of the model's ability to predict durum endosperm.

#### Soft vs hard endosperm

The variation in Y explained were 6% for the first component and 75% for components two to six. Two PLS components would thus be adequate for the model. The test set comprised of 4687 pixels, of which 1774 were assigned as soft endosperm and 2913 pixels as hard endosperm. Of the 1774 pixels assigned to soft endosperm, 1722 (97.1%) were correctly predicted and 52 (2.9%) incorrectly predicted. Of the 2913 pixels assigned to hard endosperm, 2864 (98.3%) were correctly predicted as hard endosperm, while 49 (1.7%) were incorrectly predicted. These results indicate that the classes were correctly assigned, although some overlapping between classes occurred that led to incorrect assignment of a few pixels. Images of calibration and prediction will not be shown; these were similar to those interpreted for the model of soft and durum endosperm. The amount of pixels in each class assigned to the training and test sets of soft and hard endosperm are indicated in **Table 3.1**.



**Figure 3.7** Loading line plots of (a) PC 1, (b) PC 2, (c) PC 3 and (d) PC 4 for soft, hard and durum endosperm.



**Figure 3.8** (a) PLS-DA calibration image and (b) PLS-DA prediction image for durum endosperm modelled after 2 PLS components.



### Hard vs durum endosperm

The variation in Y explained for the PLS-DA model were 4% after one component, 66% in the second component, 67% in the third component and 68% of the variation in Y explained for components four to six. Using two to four PLS components would be adequate for the model, as it would include 66-68% of the variation in Y (hard and durum endosperm).

The amount of pixels in each class assigned to the training and test sets of hard and durum endosperm can be seen in **Table 3.1**. The test set comprised of 5589 pixels, of which 2747 were assigned as hard endosperm and 2842 pixels as durum endosperm. Of the 2747 pixels assigned to hard endosperm, 2725 (99.2%) were correctly predicted and 22 (0.8%) incorrectly predicted. Of the 2842 pixels assigned to durum endosperm, 2651 (93.3%) were correctly predicted as durum endosperm, while 191 (6.7%) were incorrectly predicted. As with the soft vs hard model, some overlapping between classes could have occurred and caused incorrect assignment of a few pixels. Images of calibration and prediction will not be shown; these were similar to those interpreted for the model of soft and durum endosperm.

The use of up to two to three PLS components to model adequate Y variation could be related back to the variation between different hardness endosperm being explained in the loading line plot of PC 2, while a separation between clusters were observed in the score plot of PC 2 vs PC 3 (**Fig. 3.5a**). In the score plots of PC 1 vs PC 2 (**Fig. 3.4a**), and PC 2 vs PC 3 (**Fig. 3.5a**) it was clear that soft endosperm could be separated from hard and durum in PC 2, while PC 3 separated durum from hard or soft endosperm.

The results obtained from the PLS-DA models correspond well with classification rates obtained in previous studies. These include the differentiation between vitreousness of durum wheat kernels with a classification rate of 94% for starchy kernels (Gorretta *et al.*, 2006); and classification accuracies of 86 to 100% differentiating between Canadian wheat classes (Mahesh *et al.*, 2008).

Compared to methods currently used in industry to determine hardness, NIR hyperspectral imaging has the advantage of being a non-destructive technique. NIR hyperspectral imaging also has the advantage over bulk NIR spectroscopic methods used on single kernels, due to the enormous amount of spatial information per kernel that is acquired. The amount of spatial information acquired allows for visualisation of chemical compounds in the kernel. An image can be acquired of multiple kernels, while analysis on single kernels within that image is possible. Application of this analysis method in breeding trials would allow the breeder to analyse each kernel individually and non-destructively, using kernels with the required traits in further trials. The combination of spectral and spatial information acquired per kernel could be investigated as a future technique to determine the distribution of chemical compounds of interest in the wheat kernel.

## Conclusion

PCA proved to be an effective tool for data exploration as well as identification and classification of clusters of different wheat hardness. Even though three hardness classes could be clearly distinguished in the PCA score plot, some overlapping between classes occurred due to suspected small differences in chemical composition of some samples between these hardness classes. This could be observed in the PLS-DA models with an assigned class modelling as the other class.

The largest difference between hardness classes was explained in PC 2; most likely due to the difference in amounts of starch, moisture and protein present in the endosperm of different hardness wheat kernels. These wavelengths of interest where a difference between the hardness categories was observed could be used in a commercial instrument to separate hardness classes.

NIR hyperspectral imaging has the potential of being used in breeding trials to differentiate between single kernels differing in hardness. The selection of wheat kernels with required hardness in early generations would improve the efficacy of wheat quality selection in later generations.

## References

- AACC (2009a). AACC International Approved Methods of Analysis, 11th Ed. Method 39-70.02. Near-Infrared Reflectance Method for Hardness Determination in Wheat. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-39-70.02.
- AACC (2009b). AACC International Approved Methods of Analysis, 11th Ed. Method 55-30.01. Particle Size Index for Wheat Hardness. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-55-30.01.
- AACC (2009c). AACC International Approved Methods of Analysis, 11th Ed. Method 55-31.01. Single Kernel Characterisation System for Wheat Kernel Texture. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-55-31.01.
- Berman, M., Connor, P.M., Whitbourn, L.B., Coward, D.A., Osborne, B.G. & Southan, M.D. (2007). Classification of sound and stained wheat grains using visible and near infrared hyperspectral image analysis. *Journal of Near Infrared Spectroscopy*, **15**, 351 - 358.
- Bietz, J.A. (1989). Telling differences among wheat proteins may make a difference in marketing. In: *Wheat is Unique. Structure, Composition, Processing, End-use Properties, and Products* (edited by Y. Pomeranz). Pp. 303-315. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Burger, J.E. & Geladi, P.L.M. (2007). Hyperspectral image data conditioning and regression analysis. In: *Techniques and Applications of Hyperspectral Image Analysis* (edited by H.F. Grahn & P. Geladi). Pp. 127-153. Chichester, West Sussex, England: John Wiley & Sons, Ltd.
- Downey, G., Byrne, S. & Dwyer, E. (1986). Wheat trading in the republic of Ireland: the utility of a hardness index derived by near infrared reflectance spectroscopy. *Journal of the Science of Food and Agriculture*, **37**, 762-766.
- Esbensen, K. & Geladi, P. (1989). Strategy of multivariate image analysis. *Chemometrics and Intelligent Laboratory Systems*, **7**, 67-86.
- Esbensen, K.H. (2006). *Multivariate Data Analysis in Practice*. Pp. 598. Norway: CAMO Software AS.

- Esbensen, K.H. & Lied, T.T. (2007). Principles of multivariate image analysis (MIA) in remote sensing, technology and industry. In: *Techniques and Applications of Hyperspectral Image Analysis* (edited by H.F. Grahn & P. Geladi). Pp. 17-41. Chichester, West Sussex: John Wiley & Sons, Ltd.
- Fernández Pierna, J.A., Baeten, V. & Dardenne, P. (2006). Screening of compound feeds using NIR hyperspectral data. *Chemometrics and Intelligent Laboratory Systems*, **84**, 114-118.
- Geladi, P., Burger, J. & Lestander, T. (2004). Hyperspectral imaging: calibration problems and solutions. *Chemometrics and Intelligent Laboratory Systems*, **72**, 209-217.
- Geladi, P. & Grahn, H.F. (1996). *Multivariate Image Analysis*. Pp. 316. Chichester, West Sussex: John Wiley & Sons Ltd.
- Geladi, P., Grahn, H.F. & Burger, J. (2007). Multivariate images, hyperspectral imaging: background and equipment. In: *Techniques and applications of hyperspectral image analysis* (edited by H.F. Grahn & P. Geladi). Pp. 1-14. Chichester, West Sussex: John Wiley & Sons Ltd.
- Gorretta, N., Roger, J.M., Aubert, M., Bellon-Maurel, V., Campan, F. & Roumet, P. (2006). Determining vitreousness of durum wheat kernels using near infrared hyperspectral imaging. *Journal of Near Infrared Spectroscopy*, **14**, 231-239.
- Gowen, A.A., O'Donnell, C.P., Cullen, P.J., Downey, G. & Frias, J.M. (2007). Hyperspectral imaging - an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology*, **18**, 590-598.
- Greenwell, P. & Schofield, J.D. (1986). A starch granule protein associated with endosperm softness in wheat. *Cereal Chemistry*, **63**, 379-380.
- Hopkins, C.Y. & Graham, R.P. (1935). Starch content of some samples of Canadian wheat. *Canadian Journal of Research*, **12**, 820-824.
- Hoseney, R.C. (1994). Structure of cereals. In: *Principles of Cereal Science and Technology* (edited by R.C. Hoseney). Pp. 1-13. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Kazemi, S., Wang, N., Ngadi, M. & Prasher, S.O. (2005). Evaluation of frying oil quality using VIS/NIR hyperspectral analysis. *Agricultural Engineering International: The CIGR Journal*, **VII**, 1-12.
- Kent, N.L. & Evers, A.D. (1994). *Kent's Technology of Cereals*. Oxford, UK: Pergamon.
- Koç, H., Smail, V.W. & Wentzel, D.L. (2008). Reliability of InGaAs focal plane array imaging of wheat germination at early stages. *Journal of Cereal Science*, **48**, 394-400.
- Koehler IV, F.W., Lee, E., Kidder, L.H. & Lewis, N.E. (2002). Near infrared spectroscopy: the practical chemical imaging solution. *Spectroscopy Europe*, **14**, 12-19.
- Liu, Y., Chao, K., Chen, Y.R., Kim, M.S., Nou, X., Chan, D.E. & Yang, C. (2006). Comparison of visible and near infrared reflectance spectroscopy for the detection of faeces/ingesta contaminants for sanitation verification at slaughter plants. *Journal of Near Infrared Spectroscopy*, **14**, 325-331.
- Mahesh, S., Manickavasagan, A., Jayas, D.S., Paliwal, J. & White, N.D.G. (2008). Feasibility of near-infrared hyperspectral imaging to differentiate Canadian wheat classes. *Biosystems Engineering*, **101**, 50-57.
- Manley, M., Van Zyl, L. & Osborne, B.G. (2002). Using Fourier transform near infrared spectroscopy in determining kernel hardness, protein and moisture content of whole wheat flour. *Journal of Near Infrared Spectroscopy*, **10**, 71-76.
- Manley, M., Williams, P., Nilsson, D. & Geladi, P. (2009). Near infrared hyperspectral imaging for the evaluation of endosperm texture in whole yellow maize (*Zea mays* L.) kernels. *Journal of Agricultural and Food Chemistry*, DOI 10.1021/jf9018323.

- Miller, C.E. (2001). Chemical principles of near infrared technology. In: *Near Infrared Technology in the Agricultural and Food Industries* (edited by K. Norris & P.C. Williams). St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Miller, D.L. (1974). Industrial uses of flour. In: *Wheat: Production and Utilization* (edited by G.E. Inglett). Pp. 398-411. Westport, CT: Avi Publication Co.
- Nielsen, J.P. (2002). Fast quality assesment of barley and wheat: chemometric exploration of instrumental data with single seed applications. PhD Department of Dairy and Food Science, The Royal Veterinary and Agricultural University
- Norris, K.H., Hruschka, W.R., Bean, M.M. & Slaughter, D.C. (1989). A definition of wheat hardness using near infrared reflectance spectroscopy. *Cereal Foods World*, **34**, 696-705.
- O'Brien, L. & DePauw, R. (2004). Wheat breeding. In: *Encyclopedia of Grain Science* (edited by C. Wrigley, H. Corke & C.E. Walker). **Vol. 3**. Pp. 330-336. Kidlington, Oxford, UK: Elsevier Ltd.
- Oleson, B.T. (1994). World wheat production, utilisation and trade. In: *Wheat Production, Properties and Quality* (edited by W. Bushuk & V.F. Rasper). Pp. 2. London: Chapman & Hall.
- Orth, R.A. & Shellenberger, J.A. (1988). Origin, production, and utilization of wheat. In: *Wheat Chemistry and Technology* (edited by Y. Pomeranz). Pp. 1-14. St. Paul, Minnesota, USA: American Association of Cereal Chemists.
- Osborne, B. (1991). Measurement of the hardness of wheat endosperm by near infrared spectroscopy. *Postharvest News and Information*, **2**, 331-334.
- Osborne, B., Turnbull, K.M., Anderssen, R.S., Rahman, S., Sharp, P.J. & Appels, R. (2001). The hardness locus in Australian wheat lines. *Australian Journal of Agricultural Research*, **52**, 1275-1286.
- Pomeranz, Y. & Williams, P.C. (1990). Wheat hardness: its genetic, structural, and biochemical background, measurement, and significance. In: *Advances in Cereal Science and Technology* (edited by Y. Pomeranz). Pp. 471-529. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Schofield, J.D. & Greenwell, P. (1987). Wheat starch granule proteins and their technological significance. In: *European Conference on Food Science and Technology* (edited by I.D. Morton). Pp. 407-420. Bournemouth, England: Ellis Horwood.
- Shahin, M.A. & Symons, S.J. (2008). Detection of hard vitreous and starchy kernels in amber durum wheat samples using hyperspectral imaging. *NIR News*, **19**, 16-18.
- Simmonds, D.H. (1974). Chemical basis of hardness and vitreosity in the wheat kernel. *Baker's Digest*, **48**, 16-29, 63.
- Simmonds, D.H., Barlow, K.K. & Wrigley, C.W. (1973). The biochemical basis of grain hardness in wheat. *Cereal Chemistry*, **50**, 553-562.
- Singh, C.B., Jayas, D.S., Paliwal, J. & White, N.D.G. (2009). Detection of sprouted and midge-damaged wheat kernels using near infrared hyperspectral imaging. *Cereal Chemistry*, **86**, 256-260.
- Sissons, M., Osborne, B. & Sissons, S. (2006). Application of near infrared reflectance spectroscopy to a durum wheat breeding programme. *Journal of Near Infrared Spectroscopy*, **14**, 17-25.
- Smail, V.W., Fritz, A.K. & Wetzel, D.L. (2006). Chemical imaging of intact seeds with NIR focal plane array assists plant breeding. *Vibrational Spectroscopy*, **42**, 215-221.
- Tippels, K.H., Kilborn, R.H. & Preston, K.R. (1994). Bread-wheat quality defined. In: *Wheat production, properties and quality* (edited by W. Bushuk & V.F. Rasper). Pp. 26-27. London: Chapman & Hall.
- Turnbull, K.M. & Rahman, S. (2002). Endosperm texture in wheat. *Journal of Cereal Science*, **36**, 327-337.

- Weinstock, B.A., Janni, J., Hagen, L. & Wright, S. (2006). Prediction of oil and oleic acid concentrations in individual corn (*Zea mays* L.) kernels using near-infrared reflectance hyperspectral imaging and multivariate analysis. *Applied Spectroscopy*, **60**, 9-16.
- Worzella, W.W. & Cutler, G.H. (1939). A critical study of techniques for measuring granulation in wheat meal. *Journal of Agricultural Research*, **58**, 329-341.

## **CHAPTER 4**

**Tracking diffusion of conditioning water in single wheat kernels differing in hardness by near infrared hyperspectral imaging**

# Tracking diffusion of conditioning water in single wheat kernels differing in hardness by near infrared hyperspectral imaging

## Abstract

The use of near infrared (NIR) hyperspectral imaging was evaluated to determine the possibility of tracking the diffusion of conditioning water into single wheat kernels over time. NIR hyperspectral images of different hardness wheat samples were acquired at regular intervals between 0 and 36 hours. Images were acquired using a sisuChema SWIR (short wave infrared) hyperspectral, pushbroom imaging system. The images of the wheat conditioned with deionised water (dH<sub>2</sub>O) and deuterium oxide (D<sub>2</sub>O), respectively, were acquired in the wavelength range of 1000 to 2498 nm.

Exploratory principal component analysis was applied to clean the image. Image cleaning involved the removal of bad pixels and background. Standard normal variate and Savitzky-Golay smoothing were applied and score images and loading line plots were studied to determine which principal component (PC) contributed to variation because of moisture. The loading line plot of PC 3 portrayed a peak at 1940 nm which explained the variation due to moisture in wheat conditioned with H<sub>2</sub>O. The loading line plots of PC 3 and PC 5 portrayed an absorption peak at 1954 nm which explained variation due to D<sub>2</sub>O in wheat conditioned with D<sub>2</sub>O.

Comparison of difference spectra from 6 minus 0 to 36 minus 0 hrs after conditioning showed an increase in the peak at 1940 nm (bound moisture) for conditioning with H<sub>2</sub>O. In contrast to this, the comparison of difference spectra for wheat conditioned with D<sub>2</sub>O did not show an increase in the peak at 1954 nm as expected, indicating that absorption of D<sub>2</sub>O into the wheat endosperm did not occur.

## Introduction

Conditioning is one of the steps during dry milling that facilitate the best separation of bran from the wheat endosperm. By ensuring optimal separation of bran and endosperm the baking quality of the resulting flour is improved (Smith, 1956; Bass, 1988; Hosney, 1994; Dexter & Sarkar, 2004; Kweon *et al.*, 2009). The conditioning process allows the physical separation of bran and endosperm by mellowing or softening the endosperm and toughening the bran layers. Toughening of the bran layers prevent powdering of the bran during milling, thus enabling a more thorough separation of bran and endosperm. It also ensures that the flour, after milling, is in optimal condition for sifting and in the desired condition to produce optimal end use qualities (Smith, 1956; Bass, 1988; Hosney, 1994). The conditioning requirements are influenced by kernel hardness, initial moisture content, temperature of the wheat and the condition of the bran coat (Bass, 1988; Hosney, 1994; Dexter & Sarkar, 2004). The harder the wheat, the longer the time required for diffusion of conditioning water to evenly diffuse through the wheat endosperm.

Studies, reviewed by Bradbury (1960), indicated that the mode of water penetration into the kernel start with water diffusion into the bran layers, whereafter water enters the grain through the germ, and later from the bran and the brush area into the endosperm (Bradbury *et al.*, 1960). Studies performed to investigate the diffusion of conditioning water include iodine staining followed by cutting of the kernel in sections whereafter the mode of penetration was determined by microscopic investigation (Seckinger *et al.*, 1964). The measurement of the variation in endosperm density with change in moisture content in different regions of the wheat kernel over a conditioning period was used to determine the mode and rate of moisture penetration (Campbell & Jones, 1957). An autoradiographic technique was also used to determine the mode of water penetration during conditioning into English and Australian wheat cultivars (Butcher & Stenvert, 1973; Stenvert & Kingswood, 1976; Moss, 1977). This involved conditioning of wheat with tritiated water ( $T_2O$  or tritium oxide) whereafter x-ray images of wheat slices were acquired. According to these authors (Campbell & Jones, 1957; Seckinger *et al.*, 1964) the water penetrates from the back (dorsal side) and the shoulders of the grain, with no entry near the crease; while Moss (1977) showed that the bran and aleurone layer were saturated within an hour, followed by slower, irregular diffusion through the endosperm. A study conducted using the audiographic technique indicated that although the rate of water penetration differs between wheat types, the mode of penetration essentially remains the same (Stenvert & Kingswood, 1976; Hosenev, 1994).

No studies have been performed recently to study the rate and mode of water penetration into the wheat kernel. The development of new techniques of analysis, such as near infrared (NIR) hyperspectral imaging, allows investigations where this type of analysis could be performed non-destructively. This would allow investigation of the penetration of the water, in the same kernels, over time. The aim of this study was to evaluate NIR hyperspectral imaging to determine the rate of penetration of water into single wheat kernels; and the time necessary for the water to diffuse completely into the kernel.

NIR hyperspectral imaging is an advanced analytical technique which combines conventional digital imaging and NIR spectroscopy to obtain both spectral and spatial information from a sample (Koehler IV *et al.*, 2002; Reich, 2005; Roggo *et al.*, 2005; Burger, 2006; Burger & Geladi, 2006; Gowen *et al.*, 2007; Grahn & Geladi, 2007; Gowen *et al.*, 2008). Hyperspectral images are commonly known as hypercubes and are constructed of hundreds of single channel grayscale images, each representing a single band of spectral wavelengths. The hypercube consists of two-dimensional images composed of pixels in the x, y direction and a wavelength dimension in the z direction (Reich, 2005; Burger, 2006; Burger & Geladi, 2006; Gowen *et al.*, 2007; Shahin & Symons, 2008). The spectral dimension allows for identification of chemical constituents in the sample, while the spatial dimensions allow for the localisation of these components.

Water has strong absorption bands in the NIR region, providing the necessary sensitivity for accurate determination of moisture (Zhou *et al.*, 2003). NIR spectroscopy has been used to determine moisture differences in skin (Walling & Dabney, 1989) and free and bound moisture in a



pharmaceutical application (Zhou *et al.*, 2003). NIR hyperspectral imaging has been used widely in agriculture, amongst others, for the detection of pre-germination in wheat kernels (Smail *et al.*, 2006; Koç *et al.*, 2008; Singh *et al.*, 2009) and predicting composition of single maize kernels (Cogdill *et al.*, 2004; Baye *et al.*, 2006; Weinstock *et al.*, 2006).

The most important aspect of NIR hyperspectral imaging for this study, in contrast to previous methods used for determination of conditioning rate, is the non-destructive nature thereof and the enormous amount of spatial data that can be acquired per wheat kernel imaged. The rate of diffusion of conditioning water is an important aspect to industry. Wheat that is in the process of conditioning takes up space in the flour mill, and a shorter period of conditioning time would thus have an enormous economic benefit for the milling company.

The time whereafter diffusion of conditioning water has reached equilibrium in different hardness wheat samples were determined, and compared to guidelines used in industry. Conditioning water added to wheat was distinguished from water already present in the wheat kernel. Parallel experiments were done with D<sub>2</sub>O and H<sub>2</sub>O to allow this distinction. Deuterium is an isotope of hydrogen, and thus possesses one electron and one neutron. D<sub>2</sub>O is also referred to as “heavy water” since it has a higher molecular mass than H<sub>2</sub>O. Comparing data from the studies with D<sub>2</sub>O and H<sub>2</sub>O; D<sub>2</sub>O displayed an absorption peak in the same region as that of water added during conditioning, but at a slightly higher wavelength due to the higher molecular mass (Bokobza, 2002). Comparing conditioning with D<sub>2</sub>O and H<sub>2</sub>O enabled the assignment of wavelengths specific to water added during conditioning.

## **Materials and methods**

### *Samples*

Whole wheat kernels of varying degrees of hardness were kindly provided by Sensako (Pty) Ltd. RSA (P.O. Box 556, Bethlehem, 9700, South Africa). Whole wheat kernels were selected from bread wheat lines used in breeding trails. Selection of wheat lines were based on endosperm hardness as determined by the single kernel characterisation system (SKCS) (Method 55-31.01) (AACC, 2009). The wheat samples selected consisted of six lines in three hardness categories, i.e. soft, hard and very hard, with two lines in each of these categories.

### *Moisture determination*

The moisture determination was performed in a vacuum oven (Heraeus Model RVT 360, Henau, Germany) at 130°C until constant weight was achieved.

The wheat was milled for one minute with a coffee bean grinder (Mellerware Model 29105, Cape Town, South Africa). Moisture dishes were dried in the vacuum oven at 130°C for 30 min, whereafter it was allowed to cool in a desiccator for 40 min. The mass of each moisture dish with its lid was measured to the nearest 0.001 g and recorded ( $W_1$ ). The required amount of flour ( $5 \pm 0.001$  g) were weighed into the moisture dish and the new mass of moisture dish with sample and

lid were recorded ( $W_2$ ). Each moisture dish was placed in the oven uncovered, with the lid beneath the dish, for 1 hr. After being removed from the oven the lids were placed back on the dishes which were then allowed to cool in a desiccator for 40 min. The new mass of the moisture dish and dried sample were recorded to the nearest 0.001 g ( $W_3$ ). The moisture content of the sample was determined according to **equation 4.1**. The moisture dish was then placed back in the oven, and the process repeated until constant weight was reached. The results obtained were used to determine the amount of conditioning water to be added.

$$\% \text{ moisture} = [(W_2 - W_3) / (W_2 - W_1)] \times 100 \quad \text{equation 4.1}$$

Where:

$W_1$  = mass of moisture dish

$W_2$  = mass of moisture dish and sample before drying

$W_3$  = mass of moisture dish and sample after drying

#### *Hardness determination*

The hardness index (HI) of each wheat sample was determined using the SKCS according to AACC Approved Method 55-31.01 (Gaines *et al.*, 1996; AACC, 2009). The wheat kernel texture or hardness was determined by the instrumental force needed to crush single kernels. The SKCS model 4100 (Perten Instruments, Springfield IL) was set up according to the manufacturer's instructions; whereafter 12-16 g of sample was added to the access hopper of the instrument. The instrument counted 300 kernels and results for 1000 kernel weight, kernel diameter, HI and moisture content were obtained. The HI values as illustrated in **Table 4.1** were used to divide the wheat samples into different hardness categories for conditioning purposes.

**Table 4.1** The average hardness index values of different hardness categories wheat (AACC, 2009)

<b>Hardness category</b>	<b>Hardness Index (HI) value</b>
Extra hard	Above 90
Very hard	81-90
Hard	65-80
Medium hard	45-64
Medium soft	35-44
Soft	25-34
Very soft	10-24
Extra soft	Up to 10

#### *Wheat conditioning*

Two experiments were performed in parallel with deionised H<sub>2</sub>O and D<sub>2</sub>O, respectively. Nine wheat kernels, from each sample, were randomly chosen and imaged before conditioning. These

kernels were marked and returned to the rest of the corresponding sample. A 50 g sample of each sample was then conditioned to 17% moisture with H<sub>2</sub>O and D<sub>2</sub>O respectively. Conditioning of each sample was done in an airtight container. After water was added the wheat was stirred to distribute water through the wheat sample. The wheat was left to rest to allow the water to equilibrate. NIR hyperspectral images were acquired at 6, 12, 18, 24 and 36 hours after conditioning with the same marked kernel and the same layout was used at each of the time intervals used previously.

The starting point of the conditioning process was the determination of the initial moisture content of the sample. Normally the different hardness wheat would be conditioned to different target moisture contents (**Table 2.4**), but for the purpose of this study all six wheat samples were conditioned to 17% moisture. This was done to avoid the addition of another variable to the study.

The amount of water to add for conditioning was calculated according to **equation 4.2**.

$$\text{required water (L)} = \frac{(\text{target moisture (\%)} - \text{initial moisture (\%)})}{100} \times \text{mass (kg)} \quad \text{equation 4.2}$$

Where:

required water (L) = amount of water to be added to wheat to reach desired end moisture percentage after conditioning

target moisture (%) = desired moisture content of wheat after conditioning

initial moisture (%) = moisture content of wheat before conditioning

mass (kg) = mass of wheat to be conditioned

When this calculated amount of water was added to the wheat the total mass of the conditioned wheat increased, resulting in the total moisture in the wheat not adding up to the desired target moisture percentage. Adding an extra 8% (approximately) of the theoretical amount of water determined, corrected for this small discrepancy (Dr Phil Williams, PDK Projects Inc., Canada, personal communication).

#### *Near infrared spectroscopy measurements*

The spectral data of pure deionised water and 99.9% deuterium oxide (Sigma-Aldrich Chemie GmbH, Steinheim, Germany) were obtained using a Büchi NIRFlex N-500 Fourier transform NIR (FT-NIR) spectrophotometer (Büchi Labortechnik AG, Flawil, Switzerland). Average spectra were obtained in transmittance from 1000-2500 nm with 1501 data points per complete spectra. A resolution of 8 cm<sup>-1</sup> was used with 32 co-added scans and the sample was presented in a cuvette made of Quartz SUPRASIL<sup>®</sup> (Hellma GmbH, Müllheim, Germany) with a path length of 0.2 mm. The obtained spectra were converted to absorbance using NIRCal software version 5.1 (Büchi Labortechnik AG, Flawil, Switzerland)

### *Near infrared hyperspectral imaging system*

Near infrared hyperspectral images were acquired using a sisuChema SWIR (short wave infrared) hyperspectral, pushbroom imaging system (Specim, Spectral Imaging Ltd, Oulu, Finland). The sisuChema imaging system comprised of an imaging spectrograph coupled to a 2-D array Mercury-cadmium-telluride (HgCdTe) detector. A 73 mm lens was used for image acquisition, providing a 30  $\mu\text{m}$  spatial resolution with magnification 1. The field of view for the higher magnification images were 10 mm x 100 mm. Complete images were acquired in the spectral range of 1000-2498 nm with 6-7 nm spectral resolution. The spectrum of each pixel in a line of the sample was simultaneously recorded by the detector, where a two-dimensional image (one spatial x wavelength dimension) was acquired. The sample then moved forward on a motorised stage and the line adjacent to the one already acquired was then imaged. These individual images were stacked to form a hypercube (2 spatial (x;y) and 1 wavelength (z) dimension with the dimensions 320 (x) x 3225 (y) x 239 (z).

### *Image acquisition*

Nine wheat kernels were randomly selected from each of the samples, and were arranged in a cyclic permutation design (**Fig. 4.1a**) on silicon carbide (SiC) sandpaper background. Each higher magnification image consisted of one row in the design (**Fig. 4.1b**).

### *Hyperspectral image analysis*

#### Image correction

Correction of the raw image for white and dark reference, and transformation to pseudo-absorbance were automatically done in Evince version 2.3.0 multivariate image analysis software (UmBio AB, Umeå, Sweden). Image correction was done as discussed in Chapter 3.

#### Image cleaning

Images of the same kernels, imaged at the same sample position over the different time intervals after conditioning, were added next to each other in Evince version 2.3.0 multivariate image analysis software. The merged images of the different image acquisition times per wheat sample will be further referred to as mosaic images and had the layout as illustrated in **Fig. 4.2**. Mosaic images were imported in Evince version 2.3.0. Conditioning over time was studied per sample. An image of three kernels, of a particular sample, was merged at each of the imaging times, obtaining the layout as depicted in **Fig. 4.2**. The soft sample was studied at 0, 6, 12, 18, 24 and 36 hrs after conditioning while the hard and very hard samples was studied at 0, 6, 12, 24 and 36 hrs after conditioning.

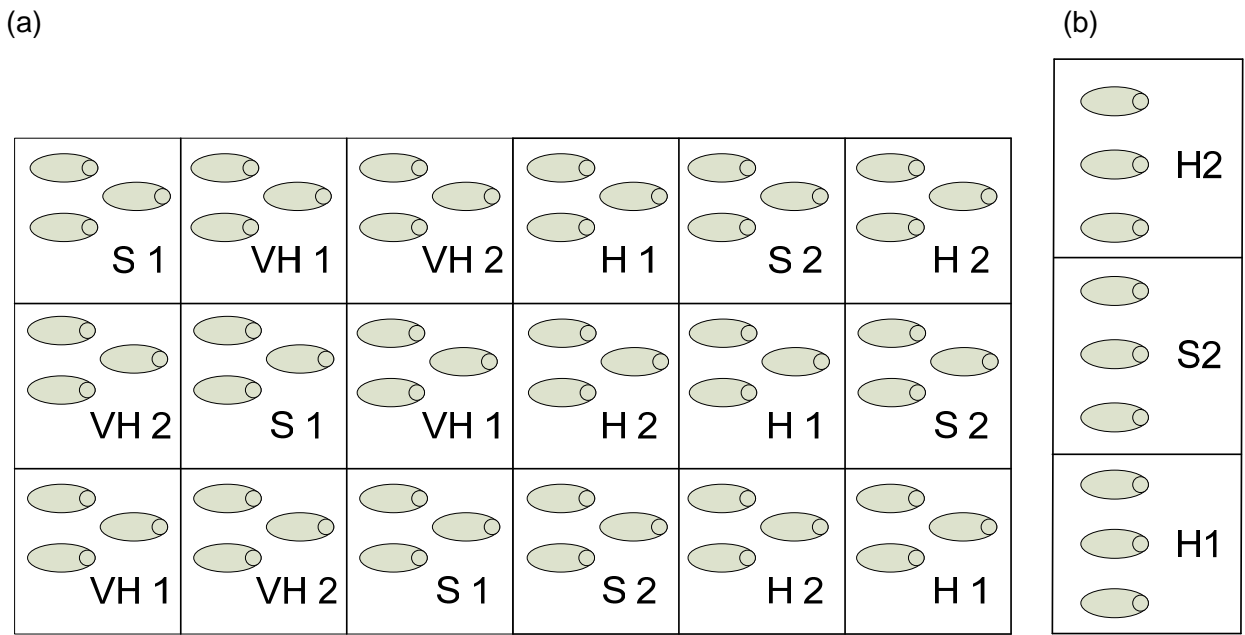
Principal component analysis (PCA) with mean centering was applied to the mosaic images in absorbance for image cleaning. Three principal components (PCs) were calculated, whereafter score plots and score images were used interactively for identification and classification of

unwanted regions in the image (outlying pixels, dead pixels, shading errors, background and edge effects). These unwanted regions were removed and PCA was recalculated with the remaining pixels which comprised data from only the wheat kernels. The cleaning was performed in Evince 2.3.0 multivariate image analysis software. Three additional PCs were calculated and the remaining cleaned image was subjected to further image analysis.

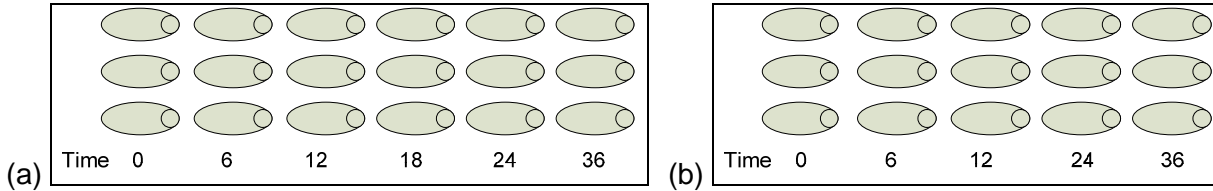
#### Image analysis of cleaned image

Pretreatment techniques were applied to the mosaic images, these included multiplicative scatter correction (MSC), standard normal variate (SNV) and Savitzky-Golay smoothing (polynomial order 3, derivative 0 and 7 left and right points). The different pretreatment techniques were evaluated based on the score plots, score images and loading line plots obtained. Score plots were evaluated for clustering due to moisture, while score images were evaluated for a clear intensity difference between different time intervals due to moisture equilibration in the whole wheat kernels. Principal component (PC) loading line plots of the individual clusters were evaluated to explain the variation in each principal component.

The average absorbance of samples at each time interval within a hardness class was compared for a change in absorption peaks due to the diffusion of conditioning water. The loading line plots and average spectra of the images after conditioning with D<sub>2</sub>O and H<sub>2</sub>O were compared. The average spectra of the three kernels in a mosaic image were acquired at each time interval within the image. The average spectra before conditioning (0 hrs) were subtracted from each of the other conditioning time's average spectra. This resulted in difference spectra that could be compared for each wheat sample to determine the change in spectra over time caused by the conditioning process.



**Figure 4.1** (a) Grid illustrating cyclic permutation design and sample presentation of the six wheat samples for normal image acquisition and (b) example of higher magnification image acquisition of one row for conditioning, with both  $D_2O$  and  $H_2O$  (S 1= soft sample one; S 2 = soft sample two; H 1 = hard sample one; H 2 = hard sample two; VH 1 = very hard sample one; VH 2 = very hard sample two).



**Figure 4.2** Images per line were merged to provide mosaic images for (a) soft wheat (0, 6, 12, 18, 24 and 36 hours), (b) hard wheat (0, 6, 12, 24 and 36 hours) and very hard wheat (0, 6, 12, 24 and 36 hours).

## Results and discussion

### *Moisture and wheat hardness determinations*

Results from the SKCS determinations allowed the distinction of wheat samples in hardness classes. This was done according to the guidelines in **Table 4.1**. The six samples were divided into soft, hard and very hard classes, with two samples in each class respectively. The moisture content determination was the first step in the conditioning process, allowing the determination of the amount of water to be added to obtain the required end moisture content.

The results of moisture content, the HI from SKCS determinations and the resultant hardness class of each wheat sample are portrayed in **Table 4.2**.

**Table 4.2** The HI results, with consequent hardness class and the results of moisture content for each of the six wheat samples

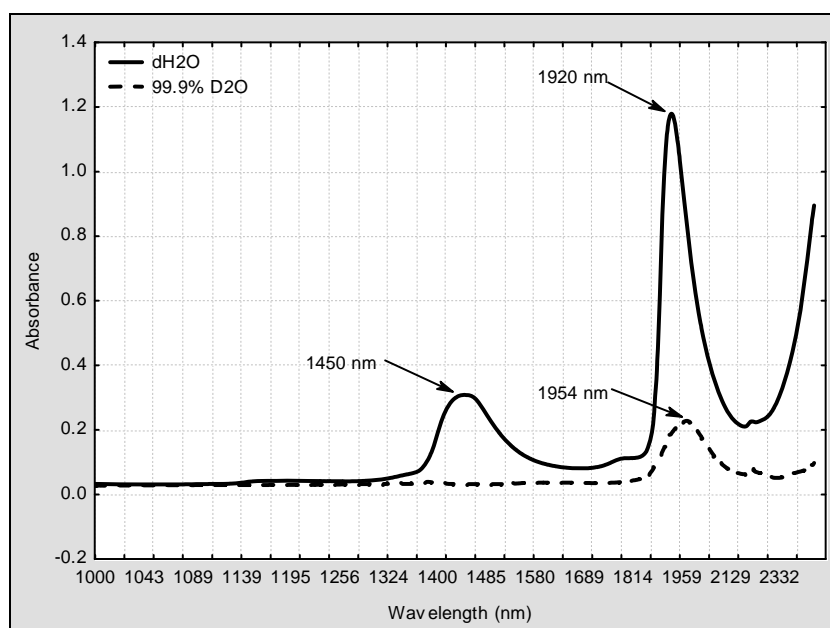
Sample	Hardness Index (HI)	Hardness Class	Moisture content (%)
1	34.0	Soft	11.90 ± 0.09*
2	44.0	Medium soft	12.86 ± 0.13
3	70.8	Hard	12.67 ± 0.00
4	72.9	Hard	12.45 ± 0.04
5	90.1	Very hard	12.74 ± 0.28
6	90.8	Very hard	12.58 ± 0.12

\* Mean ± standard deviation

### *Near infrared spectroscopy measurements*

The average spectra of water (**Fig. 4.3**) show absorption peaks at 1430 nm (O-H stretch first overtone) and 1920 nm (O-H stretch and O-H deformation). The water absorption band at 1920 nm can be described as a sum of bands arising from four different states of OH-groups: 1) 1886 nm, free water molecules with the hydrogens unbound; 2) 1899 nm, molecules with one hydrogen bonded and one unbound; 3) 1936 nm, hydrogen bonded OH-groups with energetically unfavoured bond angles; 4) 1984 nm, hydrogen bonded OH-groups with linear bonding (Luck, 1976).

The average spectra of D<sub>2</sub>O (**Fig. 4.3**) portrayed a peak at 1954 nm (O-D stretch and O-D deformation). The shift observed from 1920 nm for pure H<sub>2</sub>O and 1954 nm for D<sub>2</sub>O was due to the higher atomic mass of deuterium compared to that of hydrogen (Bokobza, 2002). The shift in frequency attained when deuterating XH groups aids in the assignment of absorption bands due to water added during conditioning. This was done by comparing the absorption peaks obtained with the two experiments. The absorbance peaks obtained with D<sub>2</sub>O conditioning were expected to be slightly shifted from those of H<sub>2</sub>O added during conditioning. This would enable us to determine the peak that can be ascribed to added water during conditioning, thus track the conditioning water through the wheat kernel over time.



**Figure 4.3** Typical spectra of dH<sub>2</sub>O and 99.9% D<sub>2</sub>O.

### *Hyperspectral image analysis*

#### Image cleaning

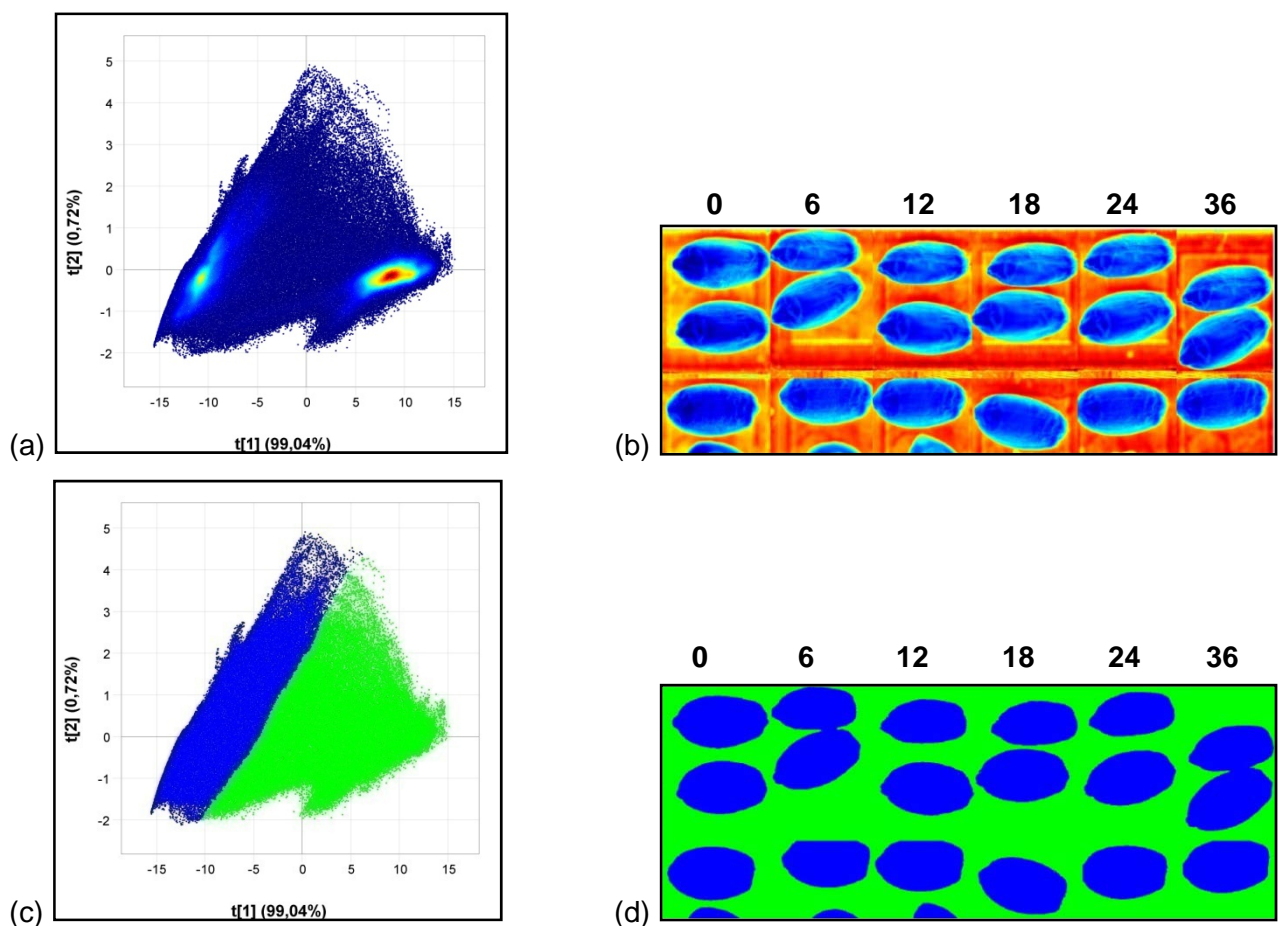
Principal component analysis with mean centering was applied to the mosaic image of soft, hard and very hard wheat for conditioning with both H<sub>2</sub>O and D<sub>2</sub>O. The interactive nature of score plots and score images in the Evince 2.3.0 image analysis software allowed for the identification and classification of classes (background and wheat) in the raw image. **Fig. 4.4** illustrates how the score plot (**Fig. 4.4a**) of the uncleaned absorbance image, of soft wheat conditioned with H<sub>2</sub>O (728973 pixels), was used interactively with the score image (**Fig. 4.4b**) to identify unwanted regions (background, shading errors, dead pixels) in the image. The unwanted regions were identified in the score image and score plot as indicated in the classification plot (**Fig. 4.4c**) and classification image (**Fig. 4.4d**). Regions defined in the classification plot and classification image were background and bad pixels (green) and wheat kernels (blue). Regions of the image with similar chemical properties have similar spectral properties, allowing for visualisation of the constituents of a sample separated into particular areas of the image (Gowen *et al.*, 2007).

Only three of the six samples studied will be included in this chapter as similar results were obtained with the other samples studied. One sample in each of the hardness classes (soft, hard and very hard) will be discussed. Information regarding all six of the mosaic images cleaned can be seen in **Table 4.3**. This includes each wheat type and conditioning treatment performed, the amount of pixels before and after cleaning, and the sum of squares (SS) for PCA with the first three PCs performed on the uncleaned image.



**Table 4.3** Indication of the amount of pixels before and after cleaning and the sum of squares for principal components used for cleaning for each wheat type and conditioning treatment performed

Wheat type	Conditioning treatment	# pixels before cleaning	# pixels after cleaning	Sum of Squares (SS)		
				PC 1	PC 2	PC 3
Soft	H <sub>2</sub> O	728973	313780	99.04	0.72	0.05
	D <sub>2</sub> O	748898	271880	99.06	0.07	0.01
Hard	H <sub>2</sub> O	685124	307305	98.39	1.33	0.05
	D <sub>2</sub> O	608379	304157	98.08	0.02	0.01
Very Hard	H <sub>2</sub> O	615595	231277	98.39	1.01	0.09
	D <sub>2</sub> O	628480	293953	98.66	0.01	0.01



**Figure 4.4** (a) PCA score plot (PC 1 vs PC 2) of an uncleaned absorbance image (728973 pixels); (b) PC 1 score image of the uncleaned image; (c) classification plot and (d) classification image (green = SiC sandpaper background ; blue = wheat kernels) of soft wheat conditioned with H<sub>2</sub>O.

## Image analysis of cleaned images

The results for different pretreatment methods were interpreted and compared based on the loading line plots obtained and clear intensity differences, over conditioning time, in the score images indicated variation due to moisture. Intensity differences in the score image were evaluated based on a colour gradient from cool (blue) to warm (red) colours. When the intensity of pixels in the score image increases it can be related to an increase of the chemical bond responsible for variation in that PC (Grahn & Geladi, 2007). A combination of SNV and Savitzky-Golay smoothing (polynomial order 3, derivative order 0 and 7 right and left points) portrayed the best results and will be used in further results and the discussion thereof.

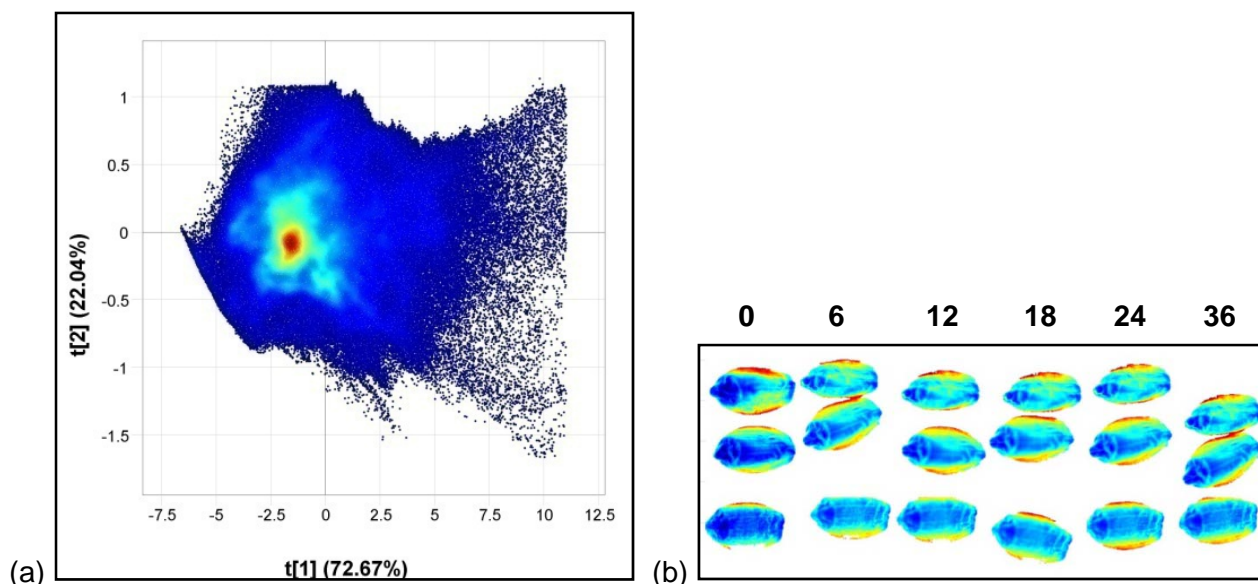
The % SS for each mosaic image with mean centering, PCA, SNV and Savitzky-Golay smoothing applied can be seen in **Table 4.4**.

**Table 4.4** Indication of the sum of squares for principal components after cleaning for each wheat type and conditioning treatment performed

Wheat type	Conditioning treatment	% Sum of Squares (SS)					
		PC 1	PC 2	PC 3	PC 4	PC 5	PC 6
Soft	H <sub>2</sub> O	72.67	22.04	1.54	0.89	0.82	0.55
	D <sub>2</sub> O	80.42	15.92	1.02	0.76	0.58	0.46
Hard	H <sub>2</sub> O	59.98	34.49	1.50	1.15	0.70	0.39
	D <sub>2</sub> O	52.04	42.84	1.18	1.13	0.72	0.38
Very Hard	H <sub>2</sub> O	58.99	33.72	2.40	1.53	0.80	0.53
	D <sub>2</sub> O	68.71	26.39	1.17	1.00	0.78	0.47

### *Conditioning with H<sub>2</sub>O*

The images of conditioning with H<sub>2</sub>O and pretreatment with SNV and Savitzky-Golay smoothing applied were used in further analysis of water penetration over time into the wheat samples. All possible combinations of PCs, in the score plot, were plotted against each other during analysis of the mosaic image. This allowed for an overview of the compound data structure (Esbensen, 2006). No distinct clusters were visible in the score plots of PC 1 (72.67%) plotted against PC 2 (22.04%) (**Fig. 4.5a**) or any other combinations of PCs plotted. Due to SNV being applied to the data, the scattering effects due to physical differences in the dataset had been removed (Burger & Geladi, 2007). The variation in PC 1 would thus not be attributed to physical differences between samples as would normally be observed in images with mean-centering and PCA without a pre-processing for scatter correction. When studying the intensity differences in the score image of PC 1 (**Fig. 4.5b**) it is clear that PC 1 could not be used to determine an increase in water content of the kernels over the conditioning period, as no intensity gradient was apparent between intervals. The results obtained for the soft sample were also similar with the hard and very hard samples when plotting PC 1 against the other five PCs.



**Figure 4.5** (a) Cleaned PCA score plot (PC 1 vs PC 2) (313780 remaining pixels) and (b) the accompanied score image of PC 1 for soft wheat conditioned with H<sub>2</sub>O.

The score images were studied for a change in intensity of the absorption per time interval over conditioning time, which could indicate an increase in chemical components responsible for the variation in a particular PC (Grahn & Geladi, 2007). An increase in intensity could be seen in the score image of PC 3 as time progressed for the soft sample (**Fig. 4.6a**). The loading line plots of the soft wheat sample were studied, indicating a very prominent water peak at 1940 nm (O-H stretch and O-H deformation) in the loading line plot of PC 3 for soft wheat (**Fig. 4.6b**). When referring to the score image with regard to intensity changes, it seemed that intensity increased until 18 hrs conditioning, whereafter no further increase in intensity occurred. This indicated that the water equilibration through the kernel has been achieved after 18 hours for soft samples, which corresponds with the time indication as applied in industry (**Table. 2.4**) (Wahrenberger, 2004). The loading line plot for PC 3 with its very prominent water peak confirmed that the variation in the score image of PC 3 were due to water.

The score image and loading line plots for hard and very hard wheat were also studied in the same manner as soft wheat. An increase in intensity of the score image for PC 3 of hard wheat was apparent until 24 hours after conditioning (**Fig. 4.7a**); there after no further changes were observed. The loading line plot of PC 3 (**Fig. 4.7b**) for hard wheat indicated that the variation in PC 3 was due to moisture, with an absorption peak at 1940 nm (O-H stretch and O-H deformation). Thus, it seemed that the diffusion of conditioning water reached equilibrium after 24 hours for hard wheat. This corresponds well with industry standards (Wahrenberger, 2004).

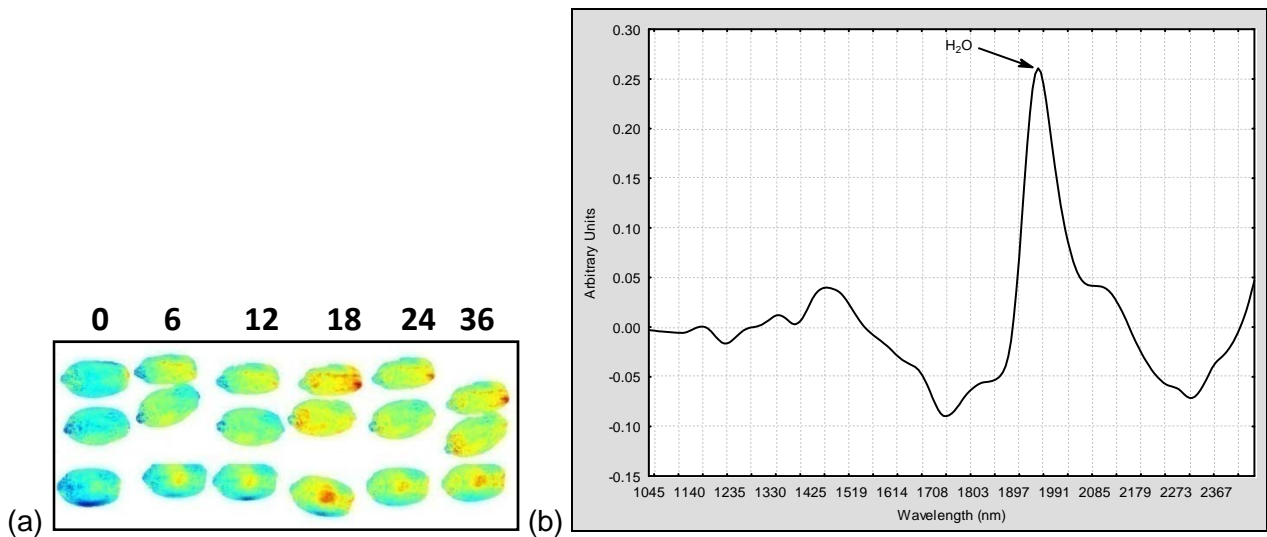
The score image of PC 3 for very hard wheat (**Fig. 4.8a**) indicated an increase in intensity till 36 hours. The loading line plot of PC 3 (**Fig. 4.8b**), with an absorption peak at 1940 nm (O-H stretch and O-H deformation), indicated that the variation in PC 3 was attributed to water. It was thus concluded that the diffusion of conditioning water reached equilibrium after 36 hours, or that

imaging over a longer conditioning time was needed to determine the optimum conditioning time for very hard wheat. Recommended time according to literature was 24 to 36 hours for very hard wheat (Wahrenberger, 2004).

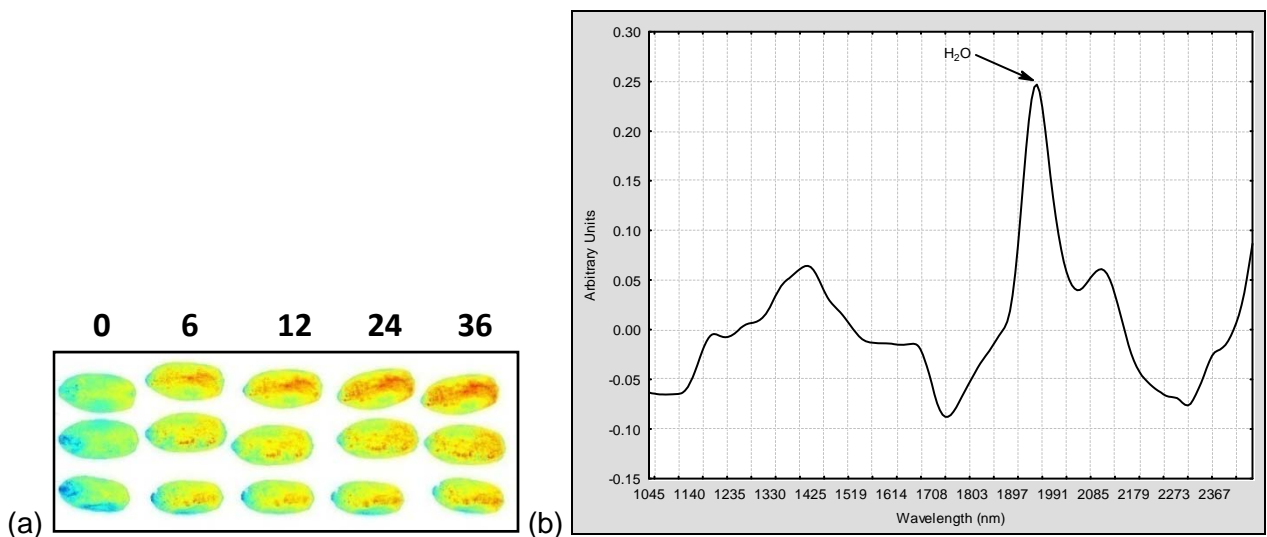
The average spectra of each sample at the different time intervals of conditioning were obtained to determine the chemical difference during the stages of conditioning. The average spectra before conditioning (0 hrs) of each sample were subtracted from the average spectra of subsequent conditioning times, to obtain difference spectra. Comparing the difference spectra would thus provide information of chemical changes during conditioning, without the presence of chemical information originally present in the sample. **Fig. 4.9** shows the difference spectra of the different samples. It was clear that the difference lies in the region of 1900 to 2100 nm for all three wheat samples, where peaks due to moisture could be observed. Both bound water and free water portray absorbance peaks in this region. The NIR band of H<sub>2</sub>O situated near 1900 nm represents water molecules with free OH groups due to bulk water (not bound to starch or protein in the wheat endosperm). This water with OH groups unbound will be referred to as free water. The bound water, however, has an absorption band closer to 1940 nm caused by less free OH groups due to the more tightly bound water molecules (Pyper *et al.*, 1985; Walling & Dabney, 1989; Zhou *et al.*, 2003). The bound water band at 1940 nm coincides with types 3 and 4 as described by Luck (1976), tightly bound water bonded by both hydrogens. The absorption of water is known to cause a shift in the peak of water from free water at 1920 nm to bound water at 1940 nm (Luck, 1976).

Comparing average spectra of each sample indicated a change in both the free (1910 nm) and the bound water (1940 nm) in the grain. The greatest difference, however, appeared to be in the region of 1940 nm, indicating that water added during conditioning was ultimately present as bound water in the kernel.

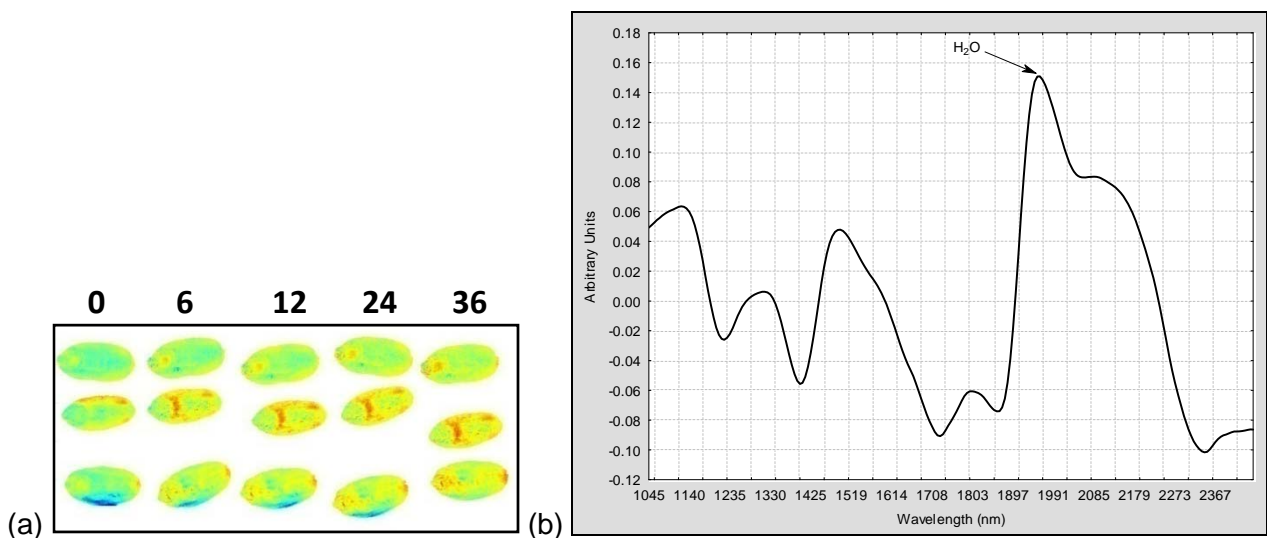
In the difference plots of the three samples, the biggest difference was situated between 6 minus 0 hours after conditioning and 36 minus 0 hours after conditioning. The conclusions made from the score images suggests that the greatest difference should have been between 18 minus 0 and 6 minus 0 hours for soft, 24 minus 0 and 6 minus 0 hours for hard, and 36 minus 0 and 6 minus 0 hours for very hard. This differs from the results obtained from the average spectra. We could speculate that the moisture distribution in the wheat endosperm has reached equilibrium in the binding state of free water, by the time intervals concluded from the score images, but that moisture change from free water to bound water still occurs up to 36 hours. This continual change from free to bound water causes the peak at 1940 nm to increase up to 36 hours.



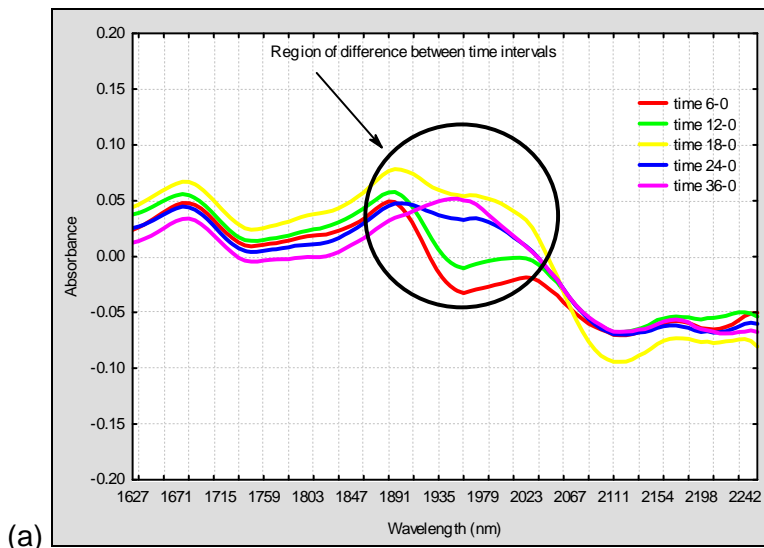
**Figure 4.6** (a) Score image of PC 3 for the soft sample and (b) the corresponding loading line plot.



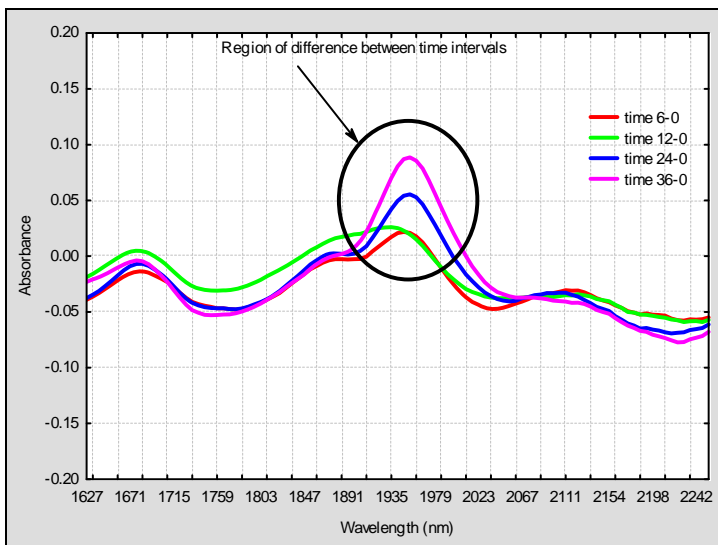
**Figure 4.7** (a) Score image of PC 3 for the hard sample and (b) the corresponding loading line plot.



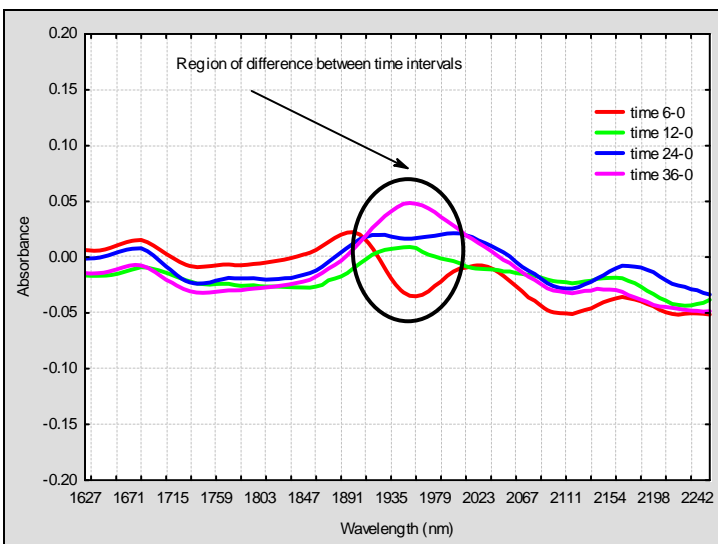
**Figure 4.8** (a) Score image of PC 3 for the very hard sample and (b) the corresponding loading line plot.



(a)



(b)



(c)

**Figure 4.9** Plot of difference spectra for the (a) soft sample, (b) hard sample, and (c) very hard sample.

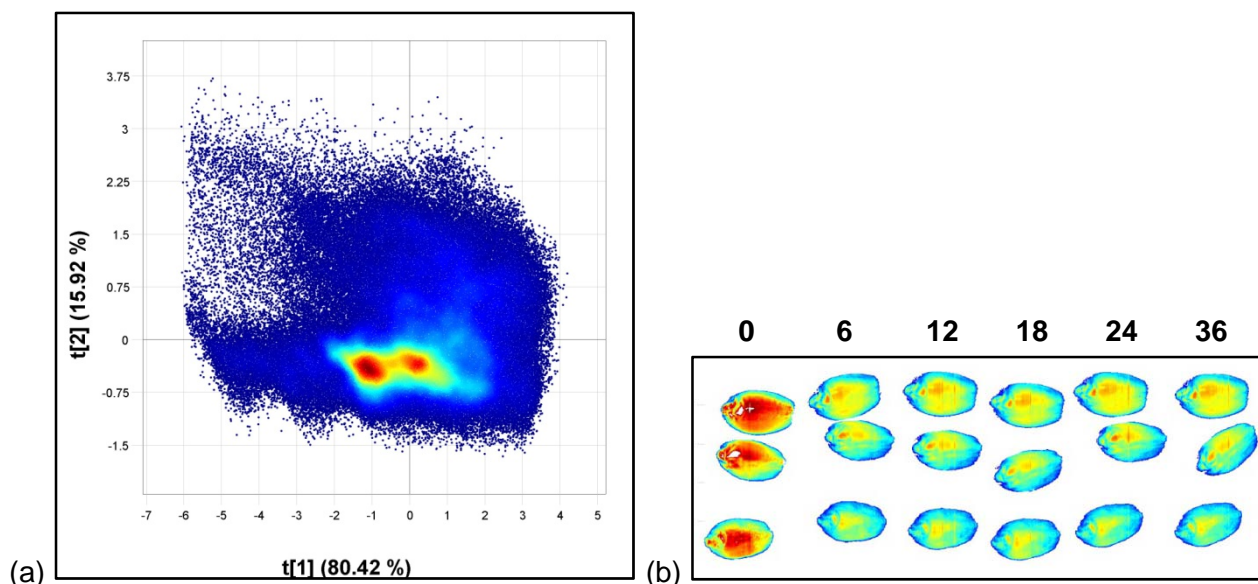
### *Conditioning with D<sub>2</sub>O*

The score plots and score images of the three samples conditioned with D<sub>2</sub>O were analysed in the same manner as with the wheat conditioned with H<sub>2</sub>O. PC 1 (80.42%) of soft wheat was plotted against PC 2 (15.92%) (**Fig. 4.10a**) and the other four PCs, however no distinct clusters relating to moisture diffusion were visible in the score plots. As with H<sub>2</sub>O conditioning images, SNV was applied to remove the scatter effects in the data set (Burger & Geladi, 2007). Savitzky-Golay smoothing (polynomial order 3, derivative order 0, 7 left and right points) was also applied to the images of all three samples. Variation in PC 1 (**Fig. 4.10b**) was not attributed to the physical differences between samples as the influence of this has been removed by applying SNV. The loading line plot of PC 1 (results not shown) indicated an average spectrum of wheat, with starch, protein and moisture contributing to the variation.

Similar results were obtained with the score plots and score images of the hard and very hard samples conditioned with D<sub>2</sub>O.

The score images for the soft wheat sample were studied for a change in intensity of the absorption per time interval over time, which could indicate an increase in chemical components responsible for the variation in a particular PC (Grahn & Geladi, 2007). The score image of PC 5 (**Fig. 4.11a**) showed an increase in intensity over time. Loading line plots of PC 5 were studied to determine if the variation was contributed due to D<sub>2</sub>O. The loading line of PC 5 (**Fig. 4.11b**) of soft wheat portrayed an absorbance peak at 1954 nm (O-D stretch and O-D deformation) which indicated that the variation in PC 5 could be attributed to D<sub>2</sub>O. The intensity of the score image portrayed a slight difference between 0 and 6 hours after conditioning, thereafter no definite change was noticeable. This most likely indicates absorption of D<sub>2</sub>O into the wheat kernel directly after addition of conditioning water, but no further diffusion of the water into the wheat endosperm.

The loading line plots of PC 3 for hard (**Fig. 4.12b**) and PC 5 for very hard (**Fig. 4.13b**) wheat samples showed a prominent absorption peak at 1954 nm (O-D stretch and O-D deformation) indicating that the variation in the PCs could be attributed to D<sub>2</sub>O. As with the soft sample only a slight increase in intensity between 0 and 6 hours after conditioning was apparent in the score image of hard (**Fig. 4.12a**) and very hard (**Fig. 4.13a**) wheat, indicating that diffusion of D<sub>2</sub>O did not occur between 6 hrs and 36 hrs.



**Figure 4.10** (a) Cleaned PCA score plot (PC 1 vs PC 2) (313780 remaining pixels) and (b) the accompanied score image of PC 1 for soft wheat conditioned with D<sub>2</sub>O.

The average spectra of each sample, at the different time intervals of conditioning, were obtained to determine the chemical difference caused by the addition of D<sub>2</sub>O during conditioning. As determined by the NIR spectroscopy spectra of 99.9% D<sub>2</sub>O, the absorbance peak for D<sub>2</sub>O is situated at 1954 nm. A shift in the absorbance peak of D<sub>2</sub>O due to binding to starch and protein was anticipated, as was seen with H<sub>2</sub>O which exhibited a shift from 1920 nm (free) to 1940 nm (bound). The discrepancies between difference spectra of the three samples with D<sub>2</sub>O conditioning (**Fig. 4.14a-c**) in the region of 1954 nm were not as prominent as those with H<sub>2</sub>O conditioning (**Fig. 4.9a-c**). It seems that the discrepancy observed between the difference spectra in **Fig. 4.14a-c** was also present at other absorption peaks in the sample spectrum, and that the difference was not caused by D<sub>2</sub>O absorbance into the wheat.

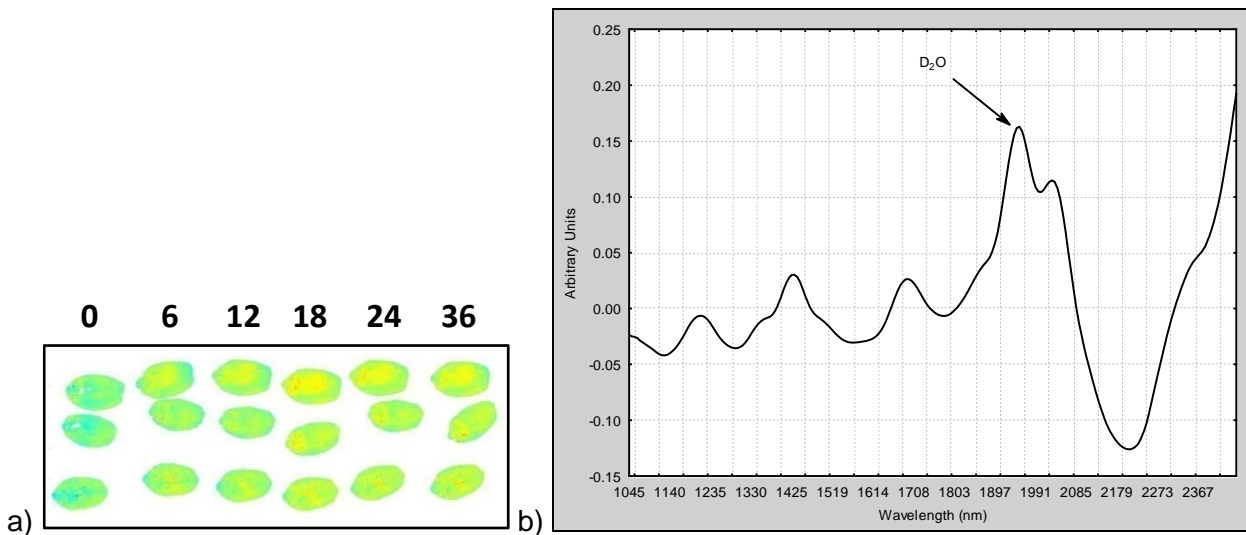
No shift in the peak was apparent from 1954 nm to a higher wavelength due to binding of D<sub>2</sub>O to starch or protein. Keeping the results from the score images in mind, it was concluded that the D<sub>2</sub>O was only absorbed into the bran layer of the wheat, or only present on the surface of the wheat with no diffusion into the wheat endosperm. Further studies regarding this theory would be necessary to determine the interaction of D<sub>2</sub>O with starch and protein in the wheat endosperm. In contrast to results obtained with conditioning with H<sub>2</sub>O, results of the score images and average spectra of conditioning with D<sub>2</sub>O did not give results as anticipated.

From a preliminary study using an imaging instrument with the wavelength range of 960 to 1662 nm to track the diffusion of conditioning water into the wheat kernel, an immense improvement in results has been observed using the sisuChema instrument with a higher wavelength range. Although the higher magnification only amplifies information already present in a data set, it allows regions of interest to be more easily identified. It is important to use an imaging instrument covering the wavelength range of 1910 to 1940 nm as this is the region where

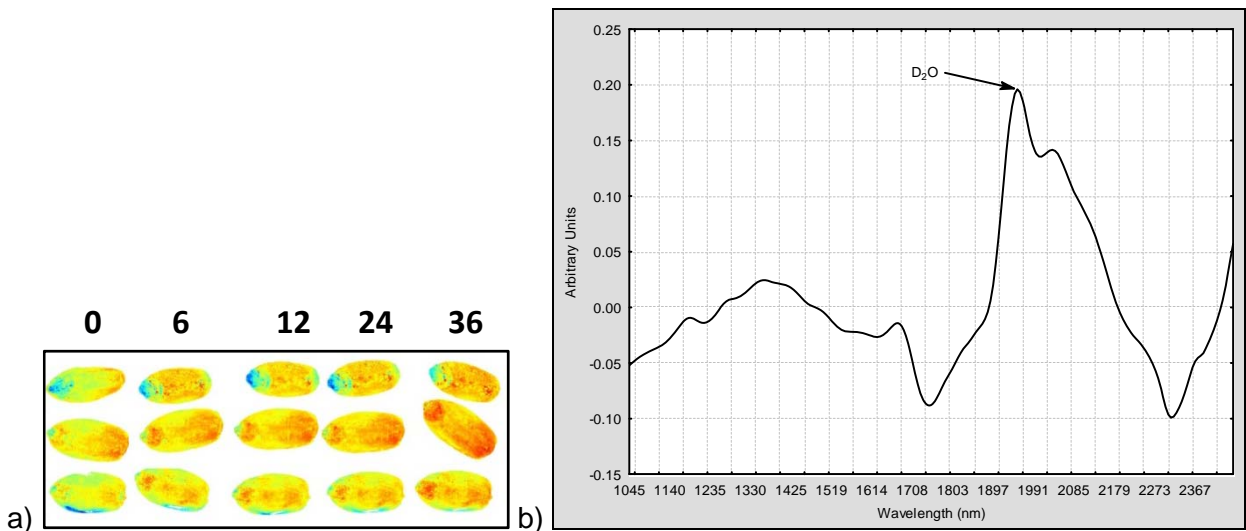


changes in free and bound conditioning water were apparent. It would be interesting to determine when free water (1910 nm) binds to the starch-protein-matrix when it is detected as bound water (1940 nm). This would improve the understanding of the diffusion of conditioning water into the wheat endosperm, and the rate and mode of penetration thereof.

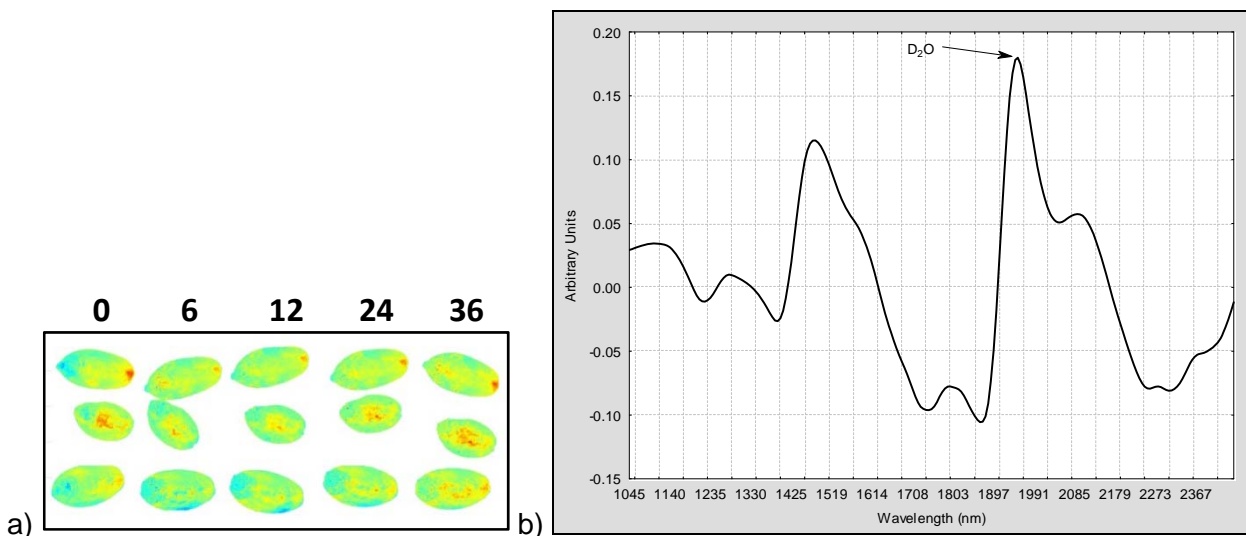
The non-destructive nature of NIR hyperspectral imaging allowed for the same wheat kernel to be used at every conditioning time interval. This enabled the variation of protein, starch and moisture originally present in the wheat kernel to be removed for the analysis of data. The changes due to the addition of conditioning water could thus be observed.



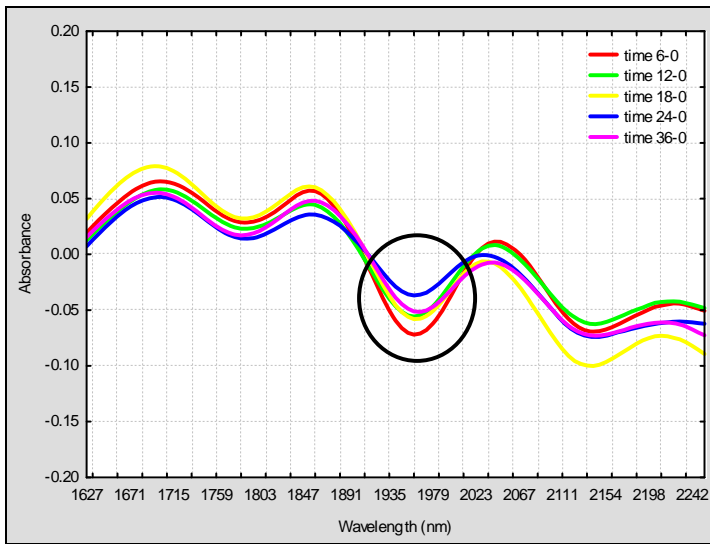
**Figure 4.11** (a) Score image of PC 5 for the soft sample and (b) the corresponding loading line plot.



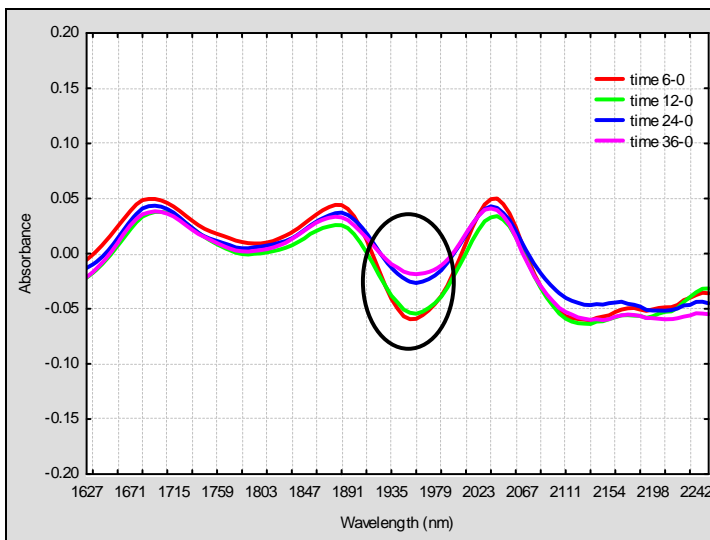
**Figure 4.12** (a) Score image of PC 3 for the hard sample and (b) the corresponding loading line plot.



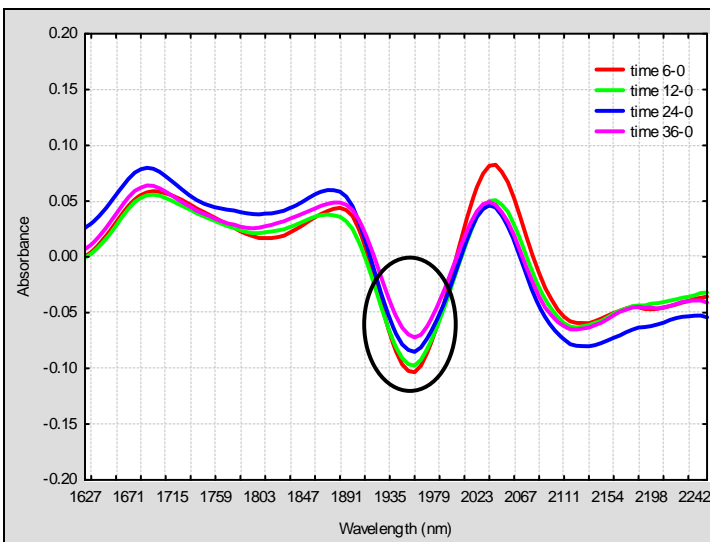
**Figure 4.13** (a) Score image of PC 5 for the very hard sample and (b) the corresponding loading line plot.



(a)



(b)



(c)

**Figure 4.14** Plot of difference spectra for the (a) soft wheat sample, (b) hard sample, and (c) the very hard sample conditioned with  $D_2O$ .

## Conclusion

This study allowed for the tracking of conditioning water from unconditioned wheat until 36 hrs after conditioning. Using the same kernel in every hyperspectral image and mean difference spectra, variation due to starch, protein and original moisture content was removed from the data. This was an improvement compared to previous techniques used to determine the mode and rate of water penetration during conditioning, since the non-destructive nature of NIR hyperspectral imaging allowed for the same kernels to be used at each time interval.

Conditioning with H<sub>2</sub>O, as opposed to results of conditioning with D<sub>2</sub>O, gave an indication of the time required for water to diffuse through the wheat kernel. Estimated conditioning times from interpretation of the score images seem to agree with those obtained in literature, and those currently used in industry. It would be interesting to further study the interaction of bound and free water during conditioning. This would improve the understanding of the interaction of water with wheat starch and protein during conditioning.

D<sub>2</sub>O did not penetrate into the wheat kernel as expected. This led to the conclusion that, although D<sub>2</sub>O had similar properties to H<sub>2</sub>O, it was not absorbed into the wheat endosperm in the same manner. It is suspected that D<sub>2</sub>O only absorbs into the wheat bran or is situated around the wheat kernel and does not penetrate the endosperm at all. This is most likely due to the heavier molecular mass of deuterium oxide compared to hydrogen, however further studies regarding this theory would be necessary.

The addition of another chemical (e.g. iodine), that has an assigned absorbance wavelength in the NIR region, to the conditioning water could aid in the tracking of conditioning water diffusing into the wheat endosperm.

## References

- AACC (2009). AACC International Approved Methods of Analysis, 11th Ed. Method 55-31.01. Single Kernel Characterisation System for Wheat Kernel Texture. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-55-31.01.
- Bass, E.J. (1988). Wheat flour milling. In: *Wheat Chemistry and Technology* (edited by Y. Pomeranz). Pp. 1-68. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Baye, T.M., Pearson, T.C. & Settles, M. (2006). Development of a calibration to predict maize seed composition using single kernel near infrared spectroscopy. *Journal of Cereal Science*, **43**, 236-243.
- Bokobza, L. (2002). Origin of near infrared absorption bands. In: *Near-Infrared Spectroscopy: Principles, Instruments, Applications* (edited by H.W. Siesler, Y. Ozaki, S. Kawata & H.M. Heise). Weinheim, Germany: Wiley-VCH Verlag GmbH.
- Bradbury, D., Hubbard, J.E., MacMasters, M.M. & Senti, F.R. (1960). Conditioning wheat for milling. A survey of the literature. *US Department of Agriculture*, **824**.
- Burger, J. (2006). Hyperspectral NIR image analysis: data exploration, correction and regression. PhD Thesis Unit of Biomass Technology and Chemistry, Swedish University of Agricultural Sciences.

- Burger, J. & Geladi, P. (2006). Hyperspectral NIR imaging for calibration and prediction: a comparison between image and spectrometer data for studying organic and biological samples. *Analyst*, **131**, 1152-1160.
- Burger, J.E. & Geladi, P.L.M. (2007). Hyperspectral image data conditioning and regression analysis. In: *Techniques and Applications of Hyperspectral Image Analysis* (edited by H.F. Grahn & P. Geladi). Pp. 127-153. Chichester, West Sussex, England: John Wiley & Sons, Ltd.
- Butcher, J. & Stenvert, N.L. (1973). Conditioning studies on Australian wheat; III. The role of the rate of water penetration into wheat grain. *Journal of the Science of Food and Agriculture*, **24**, 1077-1084.
- Campbell, J.D. & Jones, C.R. (1957). The rates of penetration of moisture to different points in the central cross-section of the endosperm in damped Manitoba wheat grains. *Cereal Chemistry*, **34**, 110-116.
- Cogdill, R.P., Hurburgh, C.R. & Rippke, G.R. (2004). Single-kernel maize analysis by near-infrared hyperspectral imaging. *Transactions of the ASAE*, **47**, 311-320.
- Dexter, J.E. & Sarkar, A.K. (2004). Wheat dry milling. In: *Encyclopedia of Grain Science* (edited by C. Wrigley, H. Corke & C.E. Walker). **Vol. 3**. Pp. 363-374. Kidlington, Oxford, UK: Elsevier Ltd.
- Esbensen, K.H. (2006). *Multivariate Data Analysis in Practice*. Pp. 598. Norway: CAMO Software AS.
- Gaines, C.S., Finney, P.F., Fleege, L.M. & Andrews, L.C. (1996). Predicting a hardness measurement using the single-kernel characterization system. *Cereal Chemistry*, **73**, 278-283.
- Gowen, A.A., O'Donnell, C.P., Cullen, P.J., Downey, G. & Frias, J.M. (2007). Hyperspectral imaging - an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology*, **18**, 590-598.
- Gowen, A.A., O'Donnell, C.P., Cullen, P.J. & Bell, S.E.J. (2008). Recent applications of chemical imaging to pharmaceutical process monitoring and quality control. *European Journal of Pharmaceutics and Biopharmaceutics*, **69**, 10-22.
- Grahn, H.F. & Geladi, P. (2007). *Techniques and Applications of Hyperspectral Image Analysis*. Pp. 1-15, 313-334. Chichester, England: John Wiley & Sons Ltd.
- Hoseney, R.C. (1994). Dry milling of cereals. In: *Principles of Cereal Science and Technology* (edited by R.C. Hoseney). Pp. 130-132. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Koç, H., Smail, V.W. & Wentzel, D.L. (2008). Reliability of InGaAs focal plane array imaging of wheat germination at early stages. *Journal of Cereal Science*, **48**, 394-400.
- Koehler IV, F.W., Lee, E., Kidder, L.H. & Lewis, N.E. (2002). Near infrared spectroscopy: the practical chemical imaging solution. *Spectroscopy Europe*, **14**, 12-19.
- Kweon, M., Martin, R. & Souza, E. (2009). Effect of tempering conditions on milling performance and flour functionality. *Cereal Chemistry*, **86**, 12-17.
- Luck, W.A. (1976). Hydrogen bonds in liquid water. In: *The hydrogen bond* (edited by P. Shuster, G. Zundel & C. Sandorty). Pp. 1369-1423. Amsterdam: New Holland Publishing.
- Moss, R. (1977). An autoradiographic technique for the location of conditioning water in wheat at cellular level. *Journal of the Science of Food and Agriculture*, **28**, 23-33.
- Pyper, J.W., Buettner, H.M., Cerjan, C.J., Hallam, J.S. & King, R.J. (1985). The measurement of bound and free moisture in organic materials by microwave methods. In: *Proceedings of the 1985 International Symposium on Moisture and Humidity*. Pp. 909-917. Washington, D.C.: ISA.
- Reich, G. (2005). Near-infrared spectroscopy and imaging: basic principles and pharmaceutical applications. *Advanced Drug Delivery Reviews*, **57**, 1109-1143.

- Roggo, Y., Edmond, A., Chalus, P. & Ulmschneider, M. (2005). Infrared hyperspectral imaging for qualitative analysis of pharmaceutical solid forms. *Analytica Chimica Acta*, **535**, 79-87.
- Seckinger, H.L., Wolf, M.J. & Dimler, R.J. (1964). Micro method for determining moisture distribution in wheat kernels, based on iodine staining. *Cereal Chemistry*, **41**, 80-87.
- Shahin, M.A. & Symons, S.J. (2008). Detection of hard vitreous and starchy kernels in amber durum wheat samples using hyperspectral imaging. *NIR News*, **19**, 16-18.
- Singh, C.B., Jayas, D.S., Paliwal, J. & White, N.D.G. (2009). Detection of sprouted and midge-damaged wheat kernels using near infrared hyperspectral imaging. *Cereal Chemistry*, **86**, 256-260.
- Smail, V.W., Fritz, A.K. & Wetzel, D.L. (2006). Chemical imaging of intact seeds with NIR focal plane array assists plant breeding. *Vibrational Spectroscopy*, **42**, 215-221.
- Smith, L. (1956). Objects of conditioning. In: *Flour Milling Technology* (edited by L. Smith). Pp. 115. Liverpool, England: Northern Publishing Co. Ltd.
- Stenvert, N.L. & Kingswood, K. (1976). An autoradiographic demonstration of the penetration of water into wheat during tempering. *Cereal Chemistry*, **53**, 141-149.
- Wahrenberger, H. (2004). Advanced milling course. Uzwil, Switzerland: Buhler.
- Walling, P.L. & Dabney, J.M. (1989). Moisture in skin by near-infrared reflectance spectroscopy. *Journal of the Society of Cosmetic Chemists*, **40**, 151-171.
- Weinstock, B.A., Janni, J., Hagen, L. & Wright, S. (2006). Prediction of oil and oleic acid concentrations in individual corn (*Zea mays* L.) kernels using near-infrared reflectance hyperspectral imaging and multivariate analysis. *Applied Spectroscopy*, **60**, 9-16.
- Zhou, G.X., Ge, Z., Dorwart, J., Izzo, B., Kukura, J., Bicker, G. & Wyvratt, J. (2003). Determination and differentiation of surface and bound water in drug substances by near infrared spectroscopy. *Journal of Pharmaceutical Sciences*, **92**, 1058-1065.

## **CHAPTER 5**

### **General discussion and conclusion**

## General discussion and conclusion

Near infrared (NIR) hyperspectral imaging is a new innovative method combining the physics of NIR spectroscopy and digital imaging. This advanced technique provides information of which compounds are present in a sample, as well as the distribution of those compounds in the sample. Two imaging systems, Spectral Dimensions MatrixNIR (960-1662 nm) and the sisuChema SWIR (short wave infrared) system (1000-2500 nm) were evaluated as non-destructive methods to distinguish between whole wheat kernels differing in wheat hardness; and to track the diffusion of conditioning water into wheat samples of different hardness categories over a period of 36 hrs. Preliminary work (not published) on wheat hardness included the analysis of images acquired on the sisuChema camera, while the MatrixNIR was used to study the diffusion of conditioning water. The results of the preliminary work will be briefly discussed in this chapter with reference to results obtained and discussed in Chapters 3 and 4.

Wheat kernel hardness is an important physical characteristic of interest to the whole wheat industry. This includes producers, millers, bakers and ultimately the consumer (Sissons *et al.*, 2006). Wheat hardness should thus be considered in breeding programmes as it has an important influence on quality (Pomeranz & Williams, 1990). Various methods are being used to determine wheat hardness; these include the particle size index (AACC, 2009b), single kernel characterisation system (AACC, 2009c), NIR spectroscopy on flour (Downey *et al.*, 1986; AACC, 2009a) and bulk whole grain (Williams, 1991). All these methods, except for bulk NIR spectroscopy applied to whole grain, are destructive. The latter method provides an average measurement of the sample. Hardness determination on single wheat kernels would however be preferable for the discrimination of samples in the breeding industry. This possibility has been studied using bulk NIR spectroscopy (Nielsen, 2002), but proved to be ineffective. The ineffectiveness could be ascribed to the single wheat kernel being a heterogenous sample, and bulk NIR spectroscopy only taking an average measurement of the kernel, and not taking every spatial dimension of the entire kernel into account.

The hardness of individual wheat kernels were studied using NIR hyperspectral imaging. Wheat kernels, varying in hardness (durum-, bread- and a soft white cultivar), were imaged using the MatrixNIR imaging system. In this study principal component analysis (PCA) for reducing and visualizing multivariate space was successfully applied to distinguish between wheat kernels varying in hardness. PCA, particularly exploratory PCA, is crucial for hyperspectral image analysis as a dimension reduction technique and as a means of image exploration. This was proved with the use of the brushing technique (Manley *et al.*, 2009a; Manley *et al.*, 2009b) in the score plot and score image. Effective cleaning, identification and classification of clusters were possible.

PC 1 attributed 93.4%, PC 2 5.67% and PC 3 0.21% of the variation. A clear separation of soft endosperm from hard and durum endosperm were apparent in PC 2; while PC 3 distinguished



durum endosperm from hard and soft endosperm for the MatrixNIR. The loading line plot of PC 2 showed absorbance peaks at 1195, 1450 and 1570 nm associated with starch and protein. The starch peak (1450 nm) in PC 2 indicated a decrease in absorption value, while the protein peak (1570 nm) increased in absorption value, with an increase in endosperm hardness. The loading line plot of PC 3 indicated absorbance peaks at 1195 and 1450 nm associated with starch and moisture. This relationship of starch and protein to the hardness of wheat was anticipated. The protein content of wheat appears to be inversely related to starch, with soft wheat having a higher starch and lower protein content than hard wheat (Hopkins & Graham, 1935; Miller, 1974). A study performed on the detection of vitreous and starchy kernels using NIR hyperspectral imaging also found differences relating to starch and protein content (Shahin & Symons, 2008).

The hardness of wheat is caused by the presence or absence of friabilin, a 15000 Da protein, in the wheat endosperm (Greenwell & Schofield, 1986; Schofield & Greenwell, 1987). Although it was apparent that the variation in PC 2 was attributed to protein, it is not clear what exactly was contributing to these differences observed, and if this could be linked to friabilin. This would have to be evaluated, in conjunction with a reference technique, using wheat samples with friabilin content known (based on the intensity of the protein band at 15000 Da determined with SDS-PAGE).

The study on wheat hardness was also performed using the sisuChema imaging system, to evaluate an instrument with a higher wavelength range to differentiate between hardness classes. A durum cultivar and three bread wheat cultivars (soft, hard and very hard) were imaged. PCA was effective in cleaning the images; SNV and Savitzky-Golay smoothing (polynomial order 3, derivative order 0 and 7 left and right points) were applied with PC 1 attributed 61.99% of the variation, PC 2 3.28% and PC 3 0.22%. No clustering could be observed in the score plots; however, soft bread wheat could be distinguished from the other samples in the score image of PC 1. Comparison of the loading line plots of the four samples could not explain the difference between the samples; PC 1 portrayed average spectra of wheat with slight differences in moisture at 1940 nm. This distinction of soft from the other samples in PC 1 could not be due to scattering or physical differences, since SNV removed these effects from the data. The samples imaged with the sisuChema were very closely related in terms of their hardness, thus their friabilin content was expected to vary slightly; while the samples imaged on the MatrixNIR instrument were extremes, thus their friabilin content was expected to vary considerably. The friabilin content of wheat has been directly linked with the hardness thereof, where soft wheat has the highest, hard wheat less, and durum wheat no friabilin (Greenwell & Schofield, 1986; Schofield & Greenwell, 1987). Since it is well known that each PC explains the largest spectral variation, in decreasing order i.e. PC 1 the largest, PC 2 the second largest, etc., we could speculate that PCA could not effectively differentiate between the small differences in friabilin content expected in the hard, very hard and durum samples. This study could be repeated on the sisuChema with wheat samples in the classes of soft, hard and durum wheat as with the MatrixNIR. This would provide information of

differences observed between these hardness classes in the higher wavelength region (i.e. 1663 to 2500 nm).

NIR analyses that could be performed on bulk wheat samples (Williams, 1991) to determine hardness, provides an average spectrum of the sample. Wheat kernel hardness differs between kernels of a cultivar. A method such as NIR hyperspectral imaging that captures information for every spatial dimension imaged in every kernel could thus be considered in differentiating between kernels of different hardness. It is not, as yet, known exactly which compounds in the kernel are responsible for differences observed with NIR hyperspectral imaging. To determine whether differences are due to friabilin, a reference technique, such as sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), for determining the presence of friabilin should be used in conjunction with NIR hyperspectral imaging. This could be used to determine the wavelength at which friabilin absorbs, the presence or absence of friabilin and can be correlated to determine wheat hardness. Nevertheless with the results obtained from the MatrixNIR images, it is recommended that an instrument that covers the wavelength range of 1450 and 1570 nm be used to distinguish between hardness classes.

Partial least squares discriminant analysis (PLS-DA) was used as a supervised classification technique to determine the ability to discriminate between different hardness endosperm classes using NIR hyperspectral imaging. PLS-DA allowed the results obtained of the MatrixNIR image to be evaluated as a prediction method for wheat hardness. For this study the image was divided into a training and test set. The classification and prediction was done on the same image, however, this is not preferable and prediction should rather be performed on a completely new image to display the true predictive potential of the model. Due to software limitations during analyses it was not possible to do this. The model of soft vs durum endosperm obtained 100% classification accuracy, the model of soft vs hard had 98% classification accuracy, while that of the hard vs durum endosperm model obtained classification rates of up to 96%. Two to three PLS components were adequate to model the Y variation for the prediction of endosperm hardness of wheat; and this concurs with predictive results previously obtained on whole grain maize kernels (Williams, 2008). These classification results compared well with results obtained differentiating between vitreous and starchy kernels with 94% classification accuracy (Gorretta *et al.*, 2006), and discriminating between wheat cultivars with classification accuracy of 84 to 100% (Mahesh *et al.*, 2008). This demonstrated the ability of NIR hyperspectral imaging to correctly classify single wheat kernels.

Conditioning requirements are also influenced by wheat hardness. Wheat that is conditioned takes up space in the flour mill, and a shorter period of conditioning time would have an enormous economic benefit for the milling industry. It would thus be beneficial to the milling industry if the optimum conditioning requirements for different hardness wheat could be determined. To effectively determine these conditioning requirements it would be ideal to use a non-destructive method, allowing the tracking of conditioning water in a single kernel. All of the methods previously

used to investigate conditioning were destructive. The tracking of diffusion of conditioning water in the wheat kernel has, to date, not been studied using NIR hyperspectral imaging; and would have the enormous benefit of tracking water movement in the same kernel over time.

NIR hyperspectral images were acquired, in high magnification with a spatial resolution of 30  $\mu\text{m}$ , using the sisuChema SWIR (short wave infrared) imaging system. PCA once again proved effective in cleaning the image, but the score plots did not indicate any clustering related to the difference in rate of diffusion of conditioning water. An increase in intensity over conditioning time was apparent in the score image of PC 3 for wheat conditioned with  $\text{H}_2\text{O}$ . The score image of soft wheat showed an increase in intensity up to 18 hours, whereafter intensity remained constant up to 36 hours, indicating optimum conditioning time at 18 hrs for soft wheat. The optimum conditioning time for hard wheat was 24 hrs and 36 hrs or more for very hard wheat. These conditioning times concurred with the times already implemented in the cereal industry (Wahrenberger, 2004). The variation in PC 3 could be explained due to free (1910 nm) and bound (1940 nm) moisture. The score images of wheat conditioned with deuterium oxide ( $\text{D}_2\text{O}$ ) indicated an increase in intensity between 0 and 6 hours, thereafter it remained unchanged. The variation was due to  $\text{D}_2\text{O}$  (1954 nm) as determined by studying the loading line plots of PC's 3 and 5. Comparison of difference spectra indicated an increase in the bound moisture peak (1940 nm) for  $\text{H}_2\text{O}$ . In contrast, difference spectra of  $\text{D}_2\text{O}$  did not increase at 1954 nm, indicating that absorption of  $\text{D}_2\text{O}$  into the wheat endosperm did not occur. It is most likely that  $\text{D}_2\text{O}$  does not penetrate the endosperm at all, and only absorbs into the wheat bran or is situated around the wheat kernel. The inability to absorb into the endosperm could be attributed to the heavier molecular weight of  $\text{D}_2\text{O}$  compared to that of the water molecule. Further studies regarding this theory would be necessary to determine the interaction of  $\text{D}_2\text{O}$  with starch and protein, which are the main constituents of wheat endosperm. It is apparent that the wavelength region of interest, when studying water diffusion into wheat, is 1900 to 1954 nm. This was anticipated since water displays absorbance peaks in the range of 1886 nm to 1984 nm depending on the different binding states of the OH-groups. Absorbance peaks of water are greatly influenced by the presence of starch or protein in the endosperm.

A preliminary study was also conducted to track conditioning water with images acquired using the MatrixNIR imaging system. Durum, bread and soft white wheat were conditioned and imaged at regular intervals between 0 and 48 hours. Clusters in the score plot of PC 2 (5.42%) and PC 3 (0.21%) were visible. Principal component 2 indicated differences in hardness classes due to starch and protein. Comparison between the time intervals of durum, bread and soft wheat in PC 3 indicated variance due to starch and moisture, although no distinct differences between the time intervals of these hardness classe were apparent. These results of the MatrixNIR images could not be used successfully to distinguish water diffusion in the samples. The sisuChema results indicated the wavelengths of interest are in the range 1900 to 1954 nm; thus explaining why the MatrixNIR data (spectral range 960 to 1662 nm) did not prove effective to determine diffusion

of conditioning water. The promising results obtained from the sisuChema images could also be ascribed to the higher magnification images acquired and thus the higher spatial resolution obtained. Although higher spatial magnification images only amplify the information already present in the data set, it allows specific regions of interest to be more easily distinguished from other pixels in the image.

To date no studies have been conducted where the interaction of bound and free water with starch and protein, respectively during conditioning was investigated. This would be an interesting study as it would lead to new information improving the understanding of the diffusion of conditioning water and its influence on the rate and mode of water penetration. Cross-sections of kernels could be imaged at each time interval to “see” if imaging results concur with research stating that water penetration occurs through the back and shoulder areas with no penetration through the crease.

Compared to the methods currently used in the cereal industry, NIR hyperspectral imaging has the advantage of being a non-destructive method for determining wheat hardness and for the tracking of conditioning water. The non-destructive nature of this technique allowed the use of the same kernel at each time interval during conditioning. This enabled the variation due to protein, starch and moisture already present in the grain to be removed in the analysis of data, by using difference spectra. To our knowledge this was the first study performed investigating the diffusion of conditioning water where the same kernel could be imaged at each time interval. Performing the measurement on the same kernel over conditioning time would not have been possible without the non-destructive nature of NIR hyperspectral imaging. The non-destructive determination of wheat hardness is also of great importance, as this would allow breeders to determine hardness of single kernels which could then be used in further breeding studies. NIR hyperspectral imaging could also be evaluated for other quality traits of importance in the wheat breeding industry especially where the localisation of constituents within the wheat kernel is required.

Compared to using bulk NIR spectroscopy on single kernels, an enormous amount of spectral data per kernel imaged is obtained; for every spatial position of a kernel imaged a NIR spectrum is obtained. This amount of data per kernel allow for the prediction on single kernels. NIR hyperspectral imaging provides the opportunity to capture an image of multiple kernels, although analysis of single kernels can be performed. The most important application of NIR hyperspectral imaging would be for plant breeding applications. Hardness prediction could be performed non-destructively on single kernels, and the kernels with required hardness used in further breeding trials. This application would require building an extensive calibration model, including wheat from a wide range of hardness values. NIR hyperspectral imaging has proved a useful technique for the non-destructive determination of wheat hardness, and also in tracking the diffusion of conditioning water in wheat. This technique not only has tremendous potential as an analytical technique, but could also as a technique for identification or discrimination of wheat samples based on quality characteristics such as hardness.

## References

- AACC (2009a). AACC International Approved Methods of Analysis, 11th Ed. Method 39-70.02. Near-Infrared Reflectance Method for Hardness Determination in Wheat. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-39-70.02.
- AACC (2009b). AACC International Approved Methods of Analysis, 11th Ed. Method 55-30.01. Particle Size Index for Wheat Hardness. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-55-30.01.
- AACC (2009c). AACC International Approved Methods of Analysis, 11th Ed. Method 55-31.01. Single Kernel Characterisation System for Wheat Kernel Texture. Approved November 3, 1999. AACC International. St. Paul, MN, USA. Doi: 10.1094/AACCIntMethod-55-31.01.
- Downey, G., Byrne, S. & Dwyer, E. (1986). Wheat trading in the republic of Ireland: the utility of a hardness index derived by near infrared reflectance spectroscopy. *Journal of the Science of Food and Agriculture*, **37**, 762-766.
- Gorretta, N., Roger, J.M., Aubert, M., Bellon-Maurel, V., Campan, F. & Roumet, P. (2006). Determining vitreousness of durum wheat kernels using near infrared hyperspectral imaging. *Journal of Near Infrared Spectroscopy*, **14**, 231-239.
- Greenwell, P. & Schofield, J.D. (1986). A starch granule protein associated with endosperm softness in wheat. *Cereal Chemistry*, **63**, 379-380.
- Hopkins, C.Y. & Graham, R.P. (1935). Starch content of some samples of Canadian wheat. *Canadian Journal of Research*, **12**, 820-824.
- Mahesh, S., Manickavasagan, A., Jayas, D.S., Paliwal, J. & White, N.D.G. (2008). Feasibility of near-infrared hyperspectral imaging to differentiate Canadian wheat classes. *Biosystems Engineering*, **101**, 50-57.
- Manley, M., Williams, P., Nilsson, D. & Geladi, P. (2009a). Near infrared hyperspectral imaging for the evaluation of endosperm texture in whole yellow maize (*Zea mays* L.) kernels. *Journal of Agricultural and Food Chemistry*, DOI 10.1021/jf9018323.
- Manley, M., Williams, P., Nilsson, D. & Geladi, P. (2009b). Near Infrared Hyperspectral Imaging for the Evaluation of Endosperm Texture in Whole Yellow Maize (*Zea mays* L.) Kernels. *Journal of Agricultural and Food Chemistry*.
- Miller, D.L. (1974). Industrial uses of flour. In: *Wheat: Production and Utilization* (edited by G.E. Inglett). Pp. 398-411. Westport, CT: Avi Publication Co.
- Nielsen, J.P. (2002). Fast quality assesment of barley and wheat: chemometric exploration of instrumental data with single seed applications. PhD Department of Dairy and Food Science, The Royal Veterinary and Agricultural University
- Pomeranz, Y. & Williams, P.C. (1990). Wheat hardness: its genetic, structural, and biochemical background, measurement, and significance. In: *Advances in Cereal Science and Technology* (edited by Y. Pomeranz). Pp. 471-529. St. Paul, Minnesota, USA: American Association of Cereal Chemists, Inc.
- Schofield, J.D. & Greenwell, P. (1987). Wheat starch granule proteins and their technological significance. In: *European Conference on Food Science and Technology* (edited by I.D. Morton). Pp. 407-420. Bournemouth, England: Ellis Horwood.
- Shahin, M.A. & Symons, S.J. (2008). Detection of hard vitreous and starchy kernels in amber durum wheat samples using hyperspectral imaging. *NIR News*, **19**, 16-18.

- Sissons, M., Osborne, B. & Sissons, S. (2006). Application of near infrared reflectance spectroscopy to a durum wheat breeding programme. *Journal of Near Infrared Spectroscopy*, **14**, 17-25.
- Wahrenberger, H. (2004). Advanced milling course. Uzwil, Switzerland: Buhler.
- Williams, P.C. (1991). Prediction of wheat kernel texture in whole grains by near-infrared transmittance. *Cereal Chemistry*, **68**, 112-114.
- Williams, P.J. (2008). Near infrared (NIR) hyperspectral imaging for evaluation of whole maize kernels: chemometrics for exploration and classification. MSc in Food Science, Stellenbosch University