Classification of Wheat Kernels with Near-Infrared Hyperspectral Imaging

by

Alexia Irene Naudé

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

MASTER OF SCIENCE IN FOOD SCIENCE

in the Department of Food Science,

Faculty of AgriSciences

at Stellenbosch University

Supervisor: Dr Paul J. Williams

December 2019

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Alexia Irene Naudé

December 2019

ACKNOWLEDGEMENTS

Firstly I would like to thank our heavenly Father for granting me this opportunity and allowing me to prosper.

I would like to express my sincere gratitude to my supervisor, Dr Williams (Department of Food Science, Stellenbosch University). Thank you for the endless motivation and support. The past two years have been the most incredible experience, and I am so grateful to have worked under your supervision.

A big thank you to Dr Janine Colling, who help me with my data collection and allowed me to work in the laboratory occasionally when I needed a change in scenery. Thank you for all the lovely chats and for helping me stay positive.

Thank you to my incredible husband, Nico Naudé, for supporting me in my decision to pursue my Master's degree. Thank you for my cup of coffee every morning (the antidote) and for taking such good care of me. En dankie vir die hulp met my Afrikaanse uittreksel. Ek is so lief vir jou.

A special thank you to my father-in-law, Willem Naudé, and VKB who sourced wheat samples for me and exposed me to the wheat industry as much as possible. This Master's would not have been possible without your help and support. Dankie, Pa. En Dankie Ma vir al Ma se ondersteuning. Dit beteken baie vir my.

Thank you to my parents, Maria and Appie, and my sister, Stephanie, for always believing in me. All the love and support throughout this thesis means the world to me.

This work is based on the research supported wholly / in part by the National Research Foundation of South Africa (Grant Numbers: 116307). The National Research Fund (NRF) are hereby acknowledged for financial support (any opinion, findings and conclusions or recommendations expressed in this material are those of the author and therefore the NRF does not accept any liability in regard thereto).

ABSTRACT

Wheat (*Triticum spp.*) is a widely grown cereal crop and is one of the most important staple foods internationally. In South Africa, it is the second most produced cereal where the majority is used for the production of bread. Grading is an important step in wheat production to ensure that the grains are of an acceptable quality for food processing and consumption. The kernels are broadly assessed in terms of cleanliness, ripeness, damage and foreign materials. Current wheat grading practices are manual, tedious, time-consuming, and subjective. More modern analytical methods that can rapidly and accurately assess wheat quality are continually addressed. In this study, four common visual defects namely heat-damaged, *Fusarium*-damaged, sprout-damaged, and immature wheat kernels were considered. The study aimed to investigate the feasibility of hyperspectral imaging and chemometrics to discriminate sound wheat from the four defective categories. Dichotomous and multiclass classifications were performed, where the objectives were to identify the optimal pre-processing technique and classification algorithm. For the classification between sound and heat-damaged wheat, the logistic regression classifier, applied to data pre-processed with a combination of 2nd derivatives and Standard Normal Variate corrected spectra, provided the highest classification accuracy (99.2%). For sprouting analysis, Support Vector Machines applied to 2nd derivative and Standard Normal Variate corrected spectra achieved the highest classification accuracy (98.6%). *Fusarium*-damaged wheat scored 100% classification accuracy with the Random Forests, decision trees or k-Nearest Neighbours classifier applied to the spectra pre-processed with 2nd derivatives. For immature wheat, the best classification was accomplished with 2nd derivatives and Partial Least Squares Discriminant Analysis, where 99.3% accuracy was achieved. The multiclass analyses had a decreased performance compared to the dichotomous analyses. The highest classification accuracy was achieved by pre-processing with Standard Normal Variate, de-trending and 2nd derivatives, and classifying with Support Vector Machines (C: 1000; γ: 1). This model attained an accuracy of 84.6%. Given that wheat is an agricultural product that varies significantly in inherent characteristics due to genetics, cultivation practices, handling and storage, the overall results achieved are highly successful. Hyperspectral imaging proved to be capable of effectively discriminating sound wheat from common occurring defects. This technology, therefore, has the potential to asses wheat quality, offering a rapid, accurate and objective alternative grading method to the cereal industry.

UITTREKSEL

Koring (*Triticum spp.*) is 'n wydverboude graangewas en is internasionaal een van die belangrikste stapelvoedsels. Dit is in Suid-Afrika die tweede meeste geproduseerde graan, en die meerderheid word vir die produksie van brood gebruik. Gradering is 'n belangrike stap in koringproduksie om te verseker dat die monster 'n aanvaarbare gehalte vir voedselverwerking en -verbruik is. Die monster word breedweg beoordeel in terme van netheid (skoon), rypheid (onryp), skade en vreemde materiale. Huidige koring gradeer praktyke is afhanklik van menslike oordeel en word per hand en oog gradeer wat tydrowend is en wat ook tot gebrek aan konsentrasie lei. Moderne analitiese metodes wat koring kwaliteit vinnig en akkuraat kan bepaal, word deurlopend ontwikkel. In hierdie studie is vier algemene visuele defekte, naamlik hittebeskadigde, *Fusarium*-beskadigde, uitgeloopte en onvolwasse koringpitte, oorweeg. Die studie het ten doel gehad om die uitvoerbaarheid van hiperspektrale beelding en chemometrie te ondersoek om gesonde koring uit die vier gebrekkige kategorieë te onderskei. Dichotome en multiklas klassifikasies is uitgevoer, met die doel om die optimale voorverwerkingstegniek en klassifikasie-algoritme te identifiseer. Vir die klassifikasie tussen gesonde en hittebeskadigde koring, het die logistieke regressie-klassifiseerder, toegepas op data wat vooraf verwerk was met 'n kombinasie van tweede orde afgeleides en “Standard Normal Variate” gekorrigeerde spektra, die hoogste klassifikasie akkuraatheid (99.2%) behaal. Vir ontkiemingsanalise, het “Support Vector Machines” toegepas op die tweede orde afgeleide en “Standard Normal Variate” gekorrigeerde spektra, die hoogste klassifikasie akkuraatheid (98.6%) behaal. Koring met *Fusarium*-beskadiging is 100% ge-identifiseer met “Random Forests”, “Decision Trees” of “k-nearest neighbours” wat toegepas was op die spektra wat vooraf met die tweede orde afgeleides verwerk is. Onvolwasse koring was die beste ge-identifiseer met die tweede orde afgeleide metode en “Partial Least Squares Discriminant Analysis”, waar 99.3% akkuraatheid behaal was. Die multiklas ontledingsmetode het swakker gevaar in vergelyking met die digotome ontledingsmetode. Die hoogste klassifikasie-akkuraatheid was behaal deur voorafverwerking met “Standard Normal Variate”, “de-trending” en tweede orde afgeleides, en te klassifiseer met “Support Vector Machines” (C: 1000; γ: 1). Hierdie model het 'n akkuraatheid van 84.6% behaal. Aangesien koring 'n landbouproduk is wat baie verskil in inherente eienskappe as gevolg van genetika, verbouingspraktyke, hantering en opberging, is die algehele resultate wat behaal is baie suksesvol. Hiperspektrale beelding het die vermoë om goeie gehalte koring effektief te onderskei van algemene koring defekte. Hierdie tegnologie het dus die potensiaal om koringkwaliteit te beoordeel en bied 'n vinnige, akkurate en objektiewe alternatiewe graderingsmetode vir die graanbedryf.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS USED

3D	three dimensional
AI	artificial intelligence
ANN	artificial neural networks
CCD	charged coupled device
CNN	convolutional neural networks
DT	de-trending
HSI	hyperspectral imaging
KNN	k-nearest neighbours
LDA	linear discriminant analysis
QDA	quadratic discriminant analysis
MSC	multiplicative scatter correction
NIR	near-infrared
NW	norris-williams
PC	principal component
PCA	principal component analysis
PLS	partial least squares
PLS-DA	partial least squares discriminant analysis
RF	random forests
RFE	recursive feature elimination
SG	savitzky-golay
SNV	standard normal variate
SVM	support vectors machine
SWIR	short wave infrared
SIMCA	soft independent modeling of class analogy
LOOCV	leave-one-out cross-validation
NB	naïve bayes
Vis	visible
m/m	mass per mass
DON	deoxynivalenol

Fs	fumonisins
ZEN	zearalenone
AFs	aflatoxins
OTA	ochratoxin a
HFN	hagberg falling number
ICC	international association for cereal science and technology
ISO	international organization for standardization
AACCI	american association of cereal chemists international
LEDs	light-emitting diodes
PGP	prism-grating-prism
LCTF	liquid crystal tunable filter
InGaAs	indium-gallium-arsenide
PbS	lead-sulfide
InSb	indium-antimonide
HgCdTe	mercury-cadmium-telluride
MCR	multivariate curve resolution
RBF	radial basis function
SEP	standard error of prediction
FHB	*fusarium* head blight
FI	*fusarium* index
MNF	minimum noise fraction.
BPNN	back-propagation neural networks
FDA	factorial discriminant analysis
PLSR	partial least squares regression
PCR	principal component regression
FOV	field of view
MCT	mercury cadmium telluride

CHAPTER 1
INTRODUCTION

Wheat (*Triticum spp.*) is a widely grown cereal crop and forms an integral part of the human diet internationally (Mahesh *et al.*, 2008; Caporaso *et al.*, 2018a). One of the main reasons for the cereal's popularity is that it is a versatile and reliable food source. The grains are rarely consumed raw and are instead converted into flour to produce a myriad of food products such as bread, pasta, cakes, cookies, pastries, and crackers (Hernández-Espinosa *et al.*, 2018). For nutritional purposes, wheat is primarily considered a source of carbohydrates, but it also provides other vital nutrients such as protein, fibre, lipids, vitamins, minerals, and phytochemicals (Shewry & Hey, 2015). Because wheat is such a valuable commodity, the processing of these grains is of high importance. The processing ability and final product quality are highly dependent on the quality of grain used (Mahesh *et al.*, 2008). Thus, there is a need to monitor the quality of these grains.

Wheat quality is assessed through grading that involves evaluating the grain against a prescribed standard to allocate a grade. Grading provides a system to ensure standardised quality throughout the wheat industry. The grade ultimately indicates the wheat's processing ability and final product quality. Grading is incredibly useful to raw material producers (farmers), primary processors (millers) and secondary processors (bakers). It also helps facilitate fair trade as the market price of wheat is based on the allocated grade (Mahesh *et al.*, 2008). Ideally, wheat kernels should be plump, bright in colour and free of any defects and contamination (Department of Agriculture, 2016; Hernández-Espinosa *et al.*, 2018). Kernel hardness, protein content, moisture content, falling number and kernel appearance are the parameters incorporated in most wheat grading systems across all large-scale wheat-producing countries (Nuttall *et al.*, 2017; Caporaso *et al.*, 2018b; Erkinbaev *et al.*, 2019). For kernel appearance, signs of immaturity, heat damage, frost damage, insect damage, sprouting, and fungal deterioration are evaluated. All these grading parameters are necessary to maintain the palatability, sanitation, shelf-life and chemical profile of the grains; vital to processing and final product quality. Besides quality, there is also a safety aspect to grading. The presence of plant fungal pathogens poses a serious health risk for humans and animals as the pathogens produce toxic compounds, known as mycotoxins (Singh *et al.*, 2012; Femenias *et al.*, 2020). When ingested, the mycotoxins can cause numerous harmful effects such as headaches, throat irritation, diarrhoea, nausea, vomiting, and gastrointestinal haemorrhage (Femenias *et al.*, 2020). Excessive exposures to mycotoxins may even result in death. Accurate grading must be performed to prevent food-borne illnesses, customer dissatisfaction, and financial losses.

Currently, grading is still predominantly performed by human visual inspection (Shahin & Symons, 2008; Delwiche *et al.*, 2011; Elmasry *et al.*, 2012; Zhang *et al.*, 2012; Barbedo *et al.*, 2015; Sendin *et al.*, 2018). Human visual inspection is subjective, time-consuming, tedious and labour-intensive (Elmasry *et al.*, 2012). In South Africa, the inspection is done by drawing a representative sample from a consignment and spreading it out onto a flat surface. Each kernel is individually inspected to identify foreign materials and defective

wheat kernels. The weight of grains for each defective category is tallied as a percentage of the sample drawn, and referenced to the levels stipulated in the grading standard (Department of Agriculture, 2016). This procedure is very tedious, even on small samples, as a sample of merely 25 g amounts to approximately 600 kernels. Alternative methods are also used in some cases, such as for the detection of toxigenic fungi by using microbiological or immunological methods in a laboratory (Shen *et al.*, 2019). These methods are however technically complicated, time-consuming, and labour-intensive.

A system in which the visual inspection process can be improved is continually addressed. An improved method to rapidly, accurately and non-destructively assess the quality of wheat is highly sought-after. Moreover, routine analyses are progressively moving away from manual techniques to improved automatic systems (Elmasry *et al.*, 2012). Automation offers an optimised system that removes human subjectivity and inconsistency. An ideal approach is to employ a non-contact system to rapidly and accurately practice quality control in a non-destructive and non-invasive manner. Such a system can also be implemented on-line or in-line to increase efficiency and productivity by allowing higher throughput. A technique that offers a considerable amount of potential is near-infrared (NIR) hyperspectral imaging (HSI).

Hyperspectral imaging (HSI) is a powerful spectroscopic technique that incorporates conventional spectroscopy and imaging (Gowen *et al.*, 2007). It has already demonstrated its capabilities in a number of applications on cereals such as maize (Berardo *et al.*, 2005; Nansen *et al.*, 2008; Williams *et al.*, 2009; Fox & Manley, 2014; Wang *et al.*, 2015b), wheat (Gorretta *et al.*, 2006; Mahesh *et al.*, 2010; Singh *et al.*, 2012; Barbedo *et al.*, 2015; Ravikanth *et al.*, 2015; Senthilkumar *et al.*, 2017; Barbedo *et al.*, 2018; Caporaso *et al.*, 2018b), barely (Arngren *et al.*, 2011; McGoverin *et al.*, 2011), oats (Serranti *et al.*, 2013; Erkinbaev *et al.*, 2017) and rice (Sun *et al.*, 2015; Wang *et al.*, 2015a; Lu *et al.*, 2017; Sun *et al.*, 2017; Qiu *et al.*, 2018). As with conventional spectroscopy, HSI is rapid, non-destructive and non-invasive. HSI, however, has the additional benefit of acquiring spatial information, allowing chemical information to be spatially resolved. The spatial resolution gives rise to the possibility of integrating spectral imaging with automatic sorting systems (Delwiche *et al.*, 2019). Computer vision systems, such as digital cameras, are already prevalent in the food industry for quality control (Elmasry *et al.*, 2012). These systems are used to detect various visual defects such as surface blemishes; however, digital cameras are limited to the visible range (380 to 700 nm) of the electromagnetic spectrum (Manley, 2014). The significance of utilising a spectroscopic technique for food inspection is to reveal characteristics that cannot be seen by the naked eye. This allows defective grains with little to no visible damage to be identified.

Conventional NIR spectroscopy is also limited in that it acquires spectral data in a single point scan which only covers a small area of the grain. This makes it challenging to obtain spectral data that is representative of heterogeneous samples, such as whole wheat kernels. Single point scans are also unable to capture the spectral data of an entire bulk sample, which is necessary to analyse multiple kernels at the same time. Combining spectroscopy with imaging in an HSI system creates an ideal solution to assess the quality of multiple whole wheat kernels simultaneously. Whole grain applications are perfect for breeding

programmes and the wheat industry as the samples can be scanned in a non-destructive and rapid manner, with the possibility of on-line or in-line implementation (Caporaso *et al.*, 2018b).

A challenge encountered with HSI is that it generates a substantial amount of data (Burger & Gowen, 2011). This causes the analysis of the images to be more complicated than that of conventional spectroscopy. The data is presented in a three-dimensional matrix, known as a *hypercube*, which holds the spatial (x x y) and spectral information (λ). Each pixel in the image holds the spectral data collected at that exact location. Hypercubes also contain a lot of redundant information and need to undergo data cleaning to remove unwanted information. For instance, information relating to the background that the sample was imaged on is also captured in the hypercube which is not relevant to the sample. After data cleaning, adequate pre-processing methods need to be performed. Pre-processing is the most important step in chemometric modelling (Rinnan *et al.*, 2009). Pre-processing methods are useful to de-noise, smooth and compress images with minimal loss of information (Rinnan *et al.*, 2009; Amigo *et al.*, 2013). Fortunately, well-known pre-processing techniques used for classical spectroscopy are transferable for the use in hyperspectral imaging. These pre-processing techniques can be divided into two categories, scatter-correction methods and spectral derivatives (Rinnan *et al.*, 2009). Scatter-correction methods include Multiplicative Scatter Correction (MSC) (Esbensen & Geladi, 1989), de-trending (DT) (Barnes *et al.*, 1989), Standard Normal Variate (SNV) (Barnes *et al.*, 1989) and normalization. Spectral derivative methods include Norris-Williams (NW) (Norris, 1983) and Savitzky-Golay (SG) (Savitzky & Golay, 1964). Different pre-processing techniques will emphasise different details about the sample. Thus the proper pre-processing technique needs to be applied to obtain the desired outcome (Amigo *et al.*, 2013).

After pre-processing, the hyperspectral data is used in chemometric modelling to perform quantitative or qualitative analyses. To develop classification models, techniques such as Logistic Regression (McCullagh & Nelder, 1989), Partial Least Squares Discriminant Analysis (PLS-DA) (Wold *et al.*, 1987), Soft Independent Modeling of Class Analogy (SIMCA) (Wold & Sjöström, 1977), Linear Discriminant Analysis (LDA) (Fisher, 1936), K-nearest neighbours (KNN) (Sebestyen, 1962), Decision Trees (Quinlan, 1986), Random Forests (RF) (Breiman, 2001), Support Vector Machines (SVM) (Cortes & Vapnik, 1995), Artificial Neural Networks (ANN) and Convolutional Neural Networks (CNN) are available. In a study by Ravikanth *et al.* (2015), contaminants in wheat were classified using NIR HSI and chemometric modelling. Five different pre-processing techniques [first derivative, second derivative, Savitzky-Golay (SG) smoothing and differentiation, multiplicative scatter correction (MSC), and standard normal variate (SNV)] were investigated. The pre-processed and raw data were then subjected to different classification algorithms, namely SVM, Naïve Bayes (NB), and KNN. It was concluded that data pre-processed with SNV and classified using KNN provided the highest classification accuracy. Delwiche *et al.* (2011) assessed *Fusarium*-damaged wheat kernels using Vis/NIR HSI. The data was applied to an LDA classifier, and an average accuracy of approximately 95% was achieved. It was also concluded that NIR was better at identifying *Fusarium*-damaged kernels than the visible region, reiterating that the visible region of the spectrum is not sufficient to perform grain inspection. The

NIR region is highly populated with combination and overtone vibrations of organic molecules, namely C–H, N–H, and O–H, and therefore has the advantage to perform constituent analysis in the kernel, such as protein content, moisture, vitreousness, and kernel appearance.

This study aimed to investigate the potential of hyperspectral imaging to classify wheat. The specific objectives were to (1) investigate and identify effective pre-processing methods, (2) build robust classification models and (3) compare the efficacy of the models.

References

Amigo, J.M., Martí, I. & Gowen, A. (2013). Hyperspectral imaging and chemometrics: a perfect combination for the analysis of food structure, composition and quality. In: *Data Handling in Science and Technology*. Pp. 343-370. Amsterdam, The Netherlands: Elsevier Science.

Arngren, M., Hansen, P., Eriksen, B., Larsen, J. & Larsen, R. (2011). Analysis of pregerminated barley using hyperspectral image analysis. *Journal of Agricultural and Food Chemistry*, **59**, 11385-11394.

Barbedo, J.G.A., Guarienti, E.M. & Tibola, C.S. (2018). Detection of sprout damage in wheat kernels using NIR hyperspectral imaging. *Biosystems Engineering*, **175**, 124-132.

Barbedo, J.G.A., Tibola, C.S. & Fernandes, J.M.C. (2015). Detecting *Fusarium* head blight in wheat kernels using hyperspectral imaging. *Biosystems Engineering*, **131**, 65-76.

Barnes, R.J., Dhanoa, M.S. & Lister, S.J. (1989). Standard Normal Variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, **43**, 772-777.

Berardo, N., Pisacane, V., Battilani, P., Scandolara, A., Pietri, A. & Marocco, A. (2005). Rapid detection of kernel rots and mycotoxins in maize by near-infrared reflectance spectroscopy. *Journal of Agricultural and Food Chemistry*, **53**, 8128-8134.

Breiman, L. (2001). Random Forests. *Machine Learning*, **45**, 5-32.

Burger, J. & Gowen, A. (2011). Data handling in hyperspectral image analysis. *Chemometrics and Intelligent Laboratory Systems*, **108**, 13-22.

Caporaso, N., Whitworth, M.B. & Fisk, I.D. (2018a). Near-Infrared spectroscopy and hyperspectral imaging for non-destructive quality assessment of cereal grains. *Applied Spectroscopy Reviews*, **53**, 667-687.

Caporaso, N., Whitworth, M.B. & Fisk, I.D. (2018b). Protein content prediction in single wheat kernels using hyperspectral imaging. *Food Chemistry*, **240**, 32-42.

Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, **20**, 273-297.

Delwiche, S.R., Kim, M.S. & Dong, Y. (2011). *Fusarium* damage assessment in wheat kernels by Vis/NIR hyperspectral imaging. *Sensing and Instrumentation for Food Quality and Safety*, **5**, 63-71.

Delwiche, S.R., Rodriguez, I.T., Rausch, S.R. & Graybosch, R.A. (2019). Estimating percentages of fusarium-damaged kernels in hard wheat by near-infrared hyperspectral imaging. *Journal of Cereal Science*, **87**, 18-24.

Department of Agriculture (2016). Regulations relating to the grading, packing and marking of bread wheat intended for sale in the Republic of South Africa. In: *Agricultural Product Standards Act (Act No. 119 of 1990)*.

Elmasry, G., Kamruzzaman, M., Sun, D.W. & Allen, P. (2012). Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: a review. *Critical Reviews in Food Science and Nutrition*, **52**, 999-1023.

Erkinbaev, C., Derksen, K. & Paliwal, J. (2019). Single kernel wheat hardness estimation using near infrared hyperspectral imaging. *Infrared Physics & Technology*, **98**, 250-255.

Erkinbaev, C., Henderson, K. & Paliwal, J. (2017). Discrimination of gluten-free oats from contaminants using near infrared hyperspectral imaging technique. *Food Control*, **80**, 197-203.

Esbensen, K. & Geladi, P. (1989). Strategy of multivariate image analysis (MIA). *Chemometrics and Intelligent Laboratory Systems*, **7**, 67-86.

Femenias, A., Gatius, F., Ramos, A.J., Sanchis, V. & Marín, S. (2020). Use of hyperspectral imaging as a tool for *Fusarium* and deoxynivalenol risk management in cereals: A review. *Food Control*, **108**, 106819.

Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, **7**, 179-188.

Fox, G. & Manley, M. (2014). Applications of single kernel conventional and hyperspectral imaging near infrared spectroscopy in cereals. *Journal of the Science of Food and Agriculture*, **94**, 174-179.

Gorretta, N., Roger, J.-M., Aubert, M., bellon maurel, V., Campan, F. & Roumet, P. (2006). Determining vitreouness of durum wheat kernels using near infrared hyperspectral imaging. *Journal of Near Infrared Spectroscopy*, **14**, 231-239.

Gowen, A., Odonnell, C., Cullen, P., Downey, G. & Frias, J. (2007). Hyperspectral imaging – an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology*, **18**, 590-598.

Hernández-Espinosa, N., Mondal, S., Autrique, E., Gonzalez-Santoyo, H., Crossa, J., Huerta-Espino, J., Singh, R.P. & Guzmán, C. (2018). Milling, processing and end-use quality traits of CIMMYT spring bread wheat germplasm under drought and heat stress. *Field Crops Research*, **215**, 104-112.

Lu, X., Sun, J., Mao, H., Wu, X. & Gao, H. (2017). Quantitative determination of rice starch based on hyperspectral imaging technology. *International Journal of Food Properties*, **20**, S1037-S1044.

Mahesh, S., Jayas, D.S., Paliwal, J. & White, N.D.G. (2010). Identification of wheat classes at different moisture levels using near-infrared hyperspectral images of bulk samples. *Sensing and Instrumentation for Food Quality and Safety*, **5**, 1-9.

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. (2014). Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. *Chemical Society Reviews*, **43**, 8200-8214.

McCullagh, P. & Nelder, J.A. (1989). *Generalized linear models.* London: Chapman & Hall/CRC.

McGoverin, C.M., Engelbrecht, P., Geladi, P. & Manley, M. (2011). Characterisation of non-viable whole barley, wheat and sorghum grains using near-infrared hyperspectral data and chemometrics. *Analytical and Bioanalytical Chemistry*, **401**, 2283-2289.

Nansen, C., Kolomiets, M. & Gao, X. (2008). Considerations regarding the use of hyperspectral imaging data in classifications of food products, exemplified by analysis of maize kernels. *Journal of Agricultural and Food Chemistry*, **56**, 2933-2938.

Norris, K.H. (1983). Extracting information from spectrophotometric curves. Predicting chemical composition from visible and near-infrared spectra. In: *Food Research and Data Analysis* (edited by H. Martens; & H. Russwurm). Pp. 95-114. London: Applied Science Publishers.

Nuttall, J.G., O'Leary, G.J., Panozzo, J.F., Walker, C.K., Barlow, K.M. & Fitzgerald, G.J. (2017). Models of grain quality in wheat—A review. *Field Crops Research*, **202**, 136-145.

Qiu, Z., Chen, J., Zhao, Y., Zhu, S., He, L. & Zhang, C. (2018). Variety identification of single rice seed using hyperspectral imaging combined with convolutional neural network. *Applied Sciences*, **8**, 1-12.

Quinlan, J.R. (1986). Induction of decision trees. *Machine Learning*, **1**, 81-106.

Ravikanth, L., Singh, C.B., Jayas, D.S. & White, N.D.G. (2015). Classification of contaminants from wheat using near-infrared hyperspectral imaging. *Biosystems Engineering*, **135**, 73-86.

Rinnan, Å., Berg, F.v.d. & Engelsen, S.B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry*, **28**, 1201-1222.

Savitzky, A. & Golay, M.J.E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, **36**, 1627-1639.

Sebestyen, G. (1962). Pattern recognition by an adaptive process of sample set construction. *IRE Transactions on Information Theory*, **8**, 82-91.

Sendin, K., Manley, M. & Williams, P.J. (2018). Classification of white maize defects with multispectral imaging. *Food Chemistry*, **243**, 311-318.

Senthilkumar, T., Jayas, D.S., White, N.D.G., Fields, P.G. & Gräfenhan, T. (2017). Detection of ochratoxin A contamination in stored wheat using near-infrared hyperspectral imaging. *Infrared Physics & Technology*, **81**, 228-235.

Serranti, S., Cesare, D., Marini, F. & Bonifazi, G. (2013). Classification of oat and goat kernels using NIR hyperspectral imaging. *Talanta*, **103**, 276-284.

Shahin, M.A. & Symons, S.J. (2008). Detection of hard vitreous and starchy kernels in amber durum wheat samples using hyperspectral imaging (GRL Number M306). *NIR news*, **19**, 16-18.

Shen, F., Zhao, T., Jiang, X., Liu, X., Fang, Y., Liu, Q., Hu, Q. & Liu, X. (2019). On-line detection of toxigenic fungal infection in wheat by visible/near infrared spectroscopy. *Lwt - Food Science and Technology*, **109**, 216-224.

Shewry, P.R. & Hey, S.J. (2015). The contribution of wheat to human diet and health. *Food Energy Security*, **4**, 178-202.

Singh, C.B., Jayas, D.S., Paliwal, J. & White, N.D.G. (2012). Fungal damage detection in wheat using short-wave near-infrared hyperspectral and digital colour imaging. *International Journal of Food Properties*, **15**, 11-24.

Sun, J., Lu, X., Mao, H., Jin, X. & Wu, X. (2015). A method for rapid identification of rice origin by hyperspectral imaging technology. *Journal of Food Process Engineering*, **40**, 1-9.

Sun, J., Lu, X., Mao, H., Wu, X. & Gao, H. (2017). Quantitative determination of rice moisture based on hyperspectral imaging technology and BCC-LS-SVR algorithm. *Journal of Food Process Engineering*, **40**, 1-8.

Wang, L., Liu, D., Pu, H., Sun, D.-W., Gao, W.H. & Xiong, Z.J. (2015a). Use of hyperspectral imaging to discriminate the variety and quality of rice. *Food and Bioprocess Technology*, **8**, 515-523.

Wang, L., Sun, D.-W., Pu, H. & Zhu, Z. (2015b). Application of hyperspectral imaging to discriminate the variety of maize seeds. *Food Analytical Methods*, **9**, 225–234.

Williams, P., Geladi, P., Fox, G. & Manley, M. (2009). Maize kernel hardness classification by near infrared (NIR) hyperspectral imaging and multivariate data analysis. *Analytica Chimica Acta*, **653**, 121-130.

Wold, S., Geladi, P., Esbensen, K. & Öhman, J. (1987). Multi-way principal components-and PLS-analysis. *Journal of Chemometrics*, **1**, 41-56.

Wold, S. & Sjöström, M. (1977). SIMCA: A method for analyzing chemical data in terms of similarity and analogy. In: *Chemometrics: Theory and Application*. Pp. 243-282. AMERICAN CHEMICAL SOCIETY.

Zhang, X., Liu, F., He, Y. & Li, X. (2012). Application of hyperspectral imaging and chemometric calibrations for variety discrimination of maize seeds. *Sensors*, **12**, 17234-17246.

CHAPTER 2
LITERATURE REVIEW

Introduction

Wheat is a cereal crop that is extensively grown and consumed worldwide (Barbedo *et al.*, 2015; Caporaso *et al.*, 2018b). Since 10 000 B.C., wheat has been a staple commodity (Shewry, 2009), and it is no different in the world today. It remains one of the leading cereals, where its worldwide annual production volume amounts to over 600 million tonnes (Barlow *et al.*, 2015; Ravikanth *et al.*, 2016; Nuttall *et al.*, 2017). Wheat belongs to the grass family where the seeds, or wheat kernels, are harvested (Frølich & Åman, 2010). To be botanically correct, a wheat kernel is a type of dry one-seeded fruit called a *caryopsis*. Reasons for wheat's popularity are its ease of processing and nutritional value. The grains are transformed into a variety of all-time favourite baked goods such as bread, cake, pasta, crackers, pastries, breakfast cereals and biscuits (Mahesh *et al.*, 2008). Wheat is also used as livestock feed (Barbedo *et al.*, 2015) and even have applications in non-food related products such as cosmetics, soaps, plastics, varnishes, paper, rubber and biodegradable packaging (Lawton, 2000).

In South Africa, wheat is one of the main cereals produced and is mostly used for the production of bread, a staple food for the majority of South Africans (de Wet & Liebenberg, 2018). On average, South Africa has an annual wheat production volume of 1.7 million tonnes. As the total commercial demand amounts to just over 3 million tonnes, wheat is in short supply. To fulfil the need, South Africa is dependent on wheat imports (Nhemachena & Kirsten, 2017). As a lot of foreign wheat producers receive agricultural subsidies, cheap, subsidised wheat is often blended with high-quality South African wheat to make up for the deficit (de Wet & Liebenberg, 2018). Since the quality of grain used as raw material directly influences the quality of the final product, it is essential to determine the quality of wheat beforehand to be able to predict its processing and end-use quality. Quality determination is especially important with foreign wheat as the variety, environmental conditions and managerial practices are not always known.

Across all wheat-producing countries, the visual appearance of wheat plays a significant role in wheat grading systems (Erkinbaev *et al.*, 2019) and traditional manual methods are still being used to assess these qualities. The most common method used is human visual inspection (Shahin & Symons, 2008; Delwiche *et al.*, 2011; Zhang *et al.*, 2012; Barbedo *et al.*, 2015; Sendin *et al.*, 2018a). A representative wheat sample is taken from a consignment and evaluated against a standard that describes the criteria required to achieve a particular class and grade. The assigned class and grade ultimately indicates the performance of the wheat as raw material. Additionally, the market price of wheat is based on the allocated grade (Mahesh *et al.*, 2008; Sendin *et al.*, 2018b). Although the graders are trained, studies have shown that human visual inspection does not provide a consistent estimation of quality (Elmasry *et al.*, 2012b). It is subjective, time-consuming and becomes increasingly inaccurate with time as fatigue, external distractions and optical illusions increase (Barbedo *et al.*, 2015; Ravikanth *et al.*, 2015). To optimise food quality and safety operations, the grading

methods used in the industry need to be improved. Routine analyses are progressively moving away from manual labour to more modern alternatives, such as optimised automated systems, that can perform mundane tasks more efficiently (Lorente *et al.*, 2011; Siche *et al.*, 2015). The latest spectroscopy approach, hyperspectral imaging (HSI), shows a lot of promise for food inspection purposes.

HSI is an analytical technique that incorporates conventional spectroscopy and imaging (Gowen *et al.*, 2007). Attractive attributes of HSI are that the analysis is simple, fast and non-destructive. It is also suitable for online applications as it has automation potential and can handle a large throughput (Vermeulen *et al.*, 2018). While conventional spectroscopy has proven to be a highly successful analytical technique, it is limited to only scanning a small portion of the sample, which is not necessarily representative of the entire sample, especially heterogeneous samples (Manley, 2014). HSI acquires spectral data over a large surface area, providing spectral data that is more representative of the sample and spatially resolved. With the addition of spatial information, objects of interest can be located, giving HSI the ability to be coupled to an automatic sorting system (Vermeulen *et al.*, 2018). The most suitable form of HSI technique for food analysis is near-infrared (NIR) HSI, as the energy of NIR typically excites the major bonds of organic molecules; the building blocks of food (Manley, 2014). With the numerous positives traits of NIR-HSI stated, it is clear that HSI is a technique that meets all of the requirements to inspect wheat quality in such a manner that will optimise grading processes.

Since HSI scans a larger surface area than conventional NIR spectroscopy, acquiring both spectral and spatial data, a vast amount of data is captured (Burger & Gowen, 2011). This is advantageous as there is a large volume of data available for model development. Not only can the spectral information related to chemistry be used, but physical and textural features can also be extracted to improve models further. It is however unsurprising that the increased volume of data adds extra complexity to data analysis. Further steps are necessary, in addition to conventional spectroscopic data analysis steps. Data cleaning is a crucial step in HSI data analysis to remove unnecessary information such as the image background, shading, curvature errors, dead pixels and outliers (Caporaso *et al.*, 2018b). To further reduce redundant information, data reduction methods such as Principal Component Analysis (PCA) (Wold *et al.*, 1987a) are frequently employed to reduce the dimensionality of hyperspectral images (Burger & Gowen, 2011). Data reduction is especially necessary in the development of real-time applications to reduce computational times. As with conventional spectroscopy, spectral pre-processing techniques are required to correct variations and interferences within the data that result from the physical properties of the sample such as morphology, particle distribution and particle size and density variances (Rinnan *et al.*, 2009). However, in instances where the physical properties contribute to the classification between two categories, it may useful to retain the light scattering information. Spectral pre-processing also facilitates data extraction and interpretation by emphasising different details about the sample to get the desired outcome (Amigo *et al.*, 2013). Standard spectral pre-processing techniques include multiplicative scatter correction (MSC) (Esbensen & Geladi, 1989), standard normal variate (SNV) (Barnes *et al.*, 1989), linear de-trending (DT) (Barnes *et al.*, 1989) and pre-processing

methods based on the application of derivatives, such as the Savitzky-Golay smoothing (SG) method (Savitzky & Golay, 1964) and Norris-Williams (NW) (Norris, 1983) (Rinnan *et al.*, 2009).

Once the data has been cleaned and pre-processed, advanced image processing and multivariate statistical methods can be applied (Vidal & Amigo, 2012). In this study, qualitative multivariate statistical methods are of interest to perform classification analyses. Classification algorithms available, amongst others, are Logistic Regression (McCullagh & Nelder, 1989), Partial Least Squares Discriminant Analysis (PLS-DA) (Wold *et al.*, 1987b), Soft independent modeling of class analogy (SIMCA) (Wold & Sjöström, 1977), Linear Discriminant Analysis (LDA) (Fisher, 1936), K-Nearest Neighbours (KNN) (Sebestyen, 1962), Decision Trees (Quinlan, 1986), Random Forests (RF) (Breiman, 2001), Support Vector Machines (SVM) (Cortes & Vapnik, 1995), artificial neural networks (ANN) and convolutional neural networks (CNN). These techniques make use of machine learning, a type of artificial intelligence (AI), that provides computers with the ability to learn without being explicitly programmed (Mutihac & Mutihac, 2008).

This review addresses the parameters involved in wheat quality and safety. The current wheat grading practices are detailed while addressing the benefits and shortcomings. The technique of hyperspectral imaging is described with regards to its history, basic spectroscopic principles and data acquisition methods. The type of data obtained from hyperspectral imaging is discussed along with the spectral pre-treatments and classification algorithms needed to build accurate and precise models. Finally, the numerous applications of hyperspectral imaging for wheat quality and safety control are discussed to exemplify the potential that NIR hyperspectral imaging has to offer.

Wheat Quality and Safety

The wheat used as raw material directly impacts the outcome of the final product (Caporaso *et al.*, 2018a). Wheat that is in excellent condition will facilitate effective processing and produce products with high-quality attributes. The opposite is also true where defective and unhealthy kernels will impede the manufacturing and baking performance, resulting in inferior products. Therefore, the quality level of wheat is important as it is indicative of the overall performance of wheat used as raw material.

The quality of wheat is determined by a range of physical and chemical properties. In general, high-quality wheat is considered to be plump, bright in colour and devoid of any defects and contamination (Hernández-Espinosa *et al.*, 2018). Some properties result from genetics, while others are influenced by managerial practices and environmental conditions (Nuttall *et al.*, 2017). The most important quality parameters are kernel hardness, protein content, moisture content, falling number and the physical condition of the grain (Nuttall *et al.*, 2017; Caporaso *et al.*, 2018b).

Kernel Hardness and Protein Content

Wheat kernels are quite complex, despite their small appearance. They consist of endosperm, a bran layer and a germ, roughly in amounts of 80–85%, 10–14% and 2.5–3.0% respectively (Fardet, 2010; Ropelewska &

Zapotoczny, 2018). Since the endosperm makes up the majority of the wheat kernel, it is evident that starch and protein are the two main components of wheat. The remaining components, arabinoxylan, β-glucan and cellulose, are present in smaller quantities (Ropelewska & Zapotoczny, 2018). The chemical composition of the kernels affects their processing ability and nutritional value. If their processing ability and nutritional value are poor, then the final product quality cannot be expected to be any better.

Kernel hardness is a quality parameter that refers to the texture of the endosperm (Ramirez *et al.*, 2003). The endosperm consists of cells that are made up of starch granules in a protein matrix. The strength of the interaction between the starch granules and protein matrix determines the kernel hardness (Gorretta *et al.*, 2006). Kernel hardness is one of the most important criteria used to assess the quality of wheat as it significantly influences the milling process, the particle size of flour and milling yield (Mahesh *et al.*, 2015; Erkinbaev *et al.*, 2019). When hard textured kernels are subjected to milling, they require more grinding energy compared to that of softly textured grains. Hard textured kernels are also more susceptible to starch damage during milling. It has been shown that the presence of proteins with a low molecular weight contributes to kernel hardness. The main protein responsible in this regard is friabilin (Ramirez *et al.*, 2003). Kernels that have higher friabilin contents tend to be softer in texture than kernels with lower friabilin contents. At the same time, the opposite can be said about the protein gluten. Harder wheat kernels typically have higher gluten contents, and softer wheat kernels have lower gluten contents (Mahesh *et al.*, 2015). Gluten is an important protein since the amount and composition of the gluten present in wheat flour controls the viscoelastic properties of the dough (Shewry, 2009). It is made up of storage proteins called gliadin and glutenin where gliadins are responsible for the cohesive nature of the dough and glutenins are responsible for the elasticity. The ratio of these storage proteins affects the rheological performance of dough which translates to its end-uses, allowing the production of a wide range of products to be possible.

Protein content also varies between different classes of wheat, where the class is indicative of the wheat variety (Mahesh *et al.*, 2015). Wheat varieties are therefore typically divided into three hardness categories described as soft, hard and very hard. In South Africa, the classes are termed “Biscuit Wheat”, “Bread Wheat” and “Durum Wheat”, respectively (Department of Agriculture, 2016). As the names suggest, harder wheat kernels are ideal for producing products such as bread and pasta, and softer wheat kernels are used to make cakes, cookies and pastries. Kernel hardness and protein content determination are therefore essential to determine the most suitable manufacturing processes.

Moisture Content

The moisture content of grains is highly important to ensure long-term storage (Karunakaran *et al.*, 2001). As water is vital for all known life forms, having an increased amount of available water will inevitably initiate and facilitate unwanted growth. The two main problems that arise with high moisture contents are microbial growth and seed germination; both of which are detrimental to the quality of wheat (Ravikanth *et al.*, 2015;

Barbedo *et al.*, 2018). To preserve the quality of grain in storage, the moisture content should not exceed 14% [mass per mass (m/m)] (Department of Agriculture, 2016). In industry, the moisture content is determined using a whole grain analyser.

Typically, it is advised to harvest relatively dry wheat, with a moisture content that is between 13–15% m/m. This practice is not always possible, especially if the grain was harvested during wet weather conditions. Wet wheat needs to be dried as quickly as possible to minimise or avoid any damage altogether (Karunakaran *et al.*, 2001; Jayas & Ghosh, 2006). Conventional methods used to dry wheat include basic aeration or in-bin air drying. In cases where the grain is too wet, heat is typically added to increase the drying rate. Heat, however, needs to be used with caution as it can also negatively impact the quality of wheat. A good postharvest practice would be to rapidly dry the grain while avoiding high temperatures to preserve the grain quality and reduce the risk of fungal growth and potential toxin contamination. Failure to exercise this practice will lead to sprouting, fungal decay and mycotoxin production, causing deterioration in the quality, safety and nutritional value.

Fungal Damage

Fungal infestations cause a substantial amount of damage to wheat grains (Delwiche *et al.*, 2011). They affect the appearance and nutritional value of the grain, but most concernedly, they are a severe health threat to animals and humans. The fungi that invade wheat are broadly divided into two groups namely field fungi and storage fungi (Sadhasivam *et al.*, 2017). Field fungi invade the wheat grains before they are harvested and will typically not increase in stored grains if the storage conditions such as moisture content and temperature are controlled. *Fusarium spp*. is a field fungi commonly found in wheat (Delwiche *et al.*, 2019). These fungi invade the grain kernel and destroy the starch granules, storage proteins and cell walls, leaving the kernels shrivelled and discoloured. The damaged kernels can visually be identified as thin and shrunken chalk-like kernels, which give rise to the name “tombstone”; an alternative name for the fungal disease “*Fusarium* head blight” (Delwiche *et al.*, 2011). *Fusarium*-damaged kernels are also of low mass and have high amounts of reducing sugars and fat rancidity. The damage caused by the fungus has a detrimental effect on flour colour, ash content and the baking performance (Shahin & Symons, 2012). The biggest concern of a fungal infestation is that fungi produce toxic compounds, known as mycotoxins, which pose serious health hazards when consumed (Delwiche *et al.*, 2011; Shahin & Symons, 2012; Barbedo *et al.*, 2015; Ropelewska & Zapotoczny, 2018).

Storage fungi invade grains during storage and are not usually present to any severe extent before harvest (Sadhasivam *et al.*, 2017). The fungi typically come about through spores (Flannigan, 1978). Bulk grains could become contaminated with spores resulting from the environment, unclean storage equipment or harvest and handling practices. Under improper storage conditions such as increased heat and moisture, the fungal spores will germinate and rapidly grow, leading to significant problems. *Penicillium spp*. and *Aspergillus spp*. are the two most common storage fungi found in stored grain (Singh *et al.*, 2012; Sadhasivam

et al., 2017). *Fusarium, Penicillium* and *Aspergillus* are collectively responsible for the production of over 300 different mycotoxins (Singh *et al.*, 2012). Deoxynivalenol (DON), or vomitoxin, is the most prevalent mycotoxin that is produced by *Fusarium*, *Fusarium graminearum* specifically, which has strong emetic effects after consumption (Dowell *et al.*, 1999; Femenias *et al.*, 2020). DON also has high stability and is difficult to eliminate during processing (Femenias *et al.*, 2020). Adequate methods are therefore necessary to detect DON in cereals upfront. Other frequently detected mycotoxins produced by *Fusarium* are fumonisins (Fs) and zearalenone (ZEN) (Senthilkumar *et al.*, 2017). The most commonly perceived mycotoxins produced by *Aspergillus* and *Penicillium* species are aflatoxins (AFs) and ochratoxin A (OTA), respectively (Sadhasivam *et al.*, 2017).

Claviceps purpurea is another fungal species that grows parasitically and affects the wheat plant during the flowering period (Vermeulen *et al.*, 2012). The fungus produces small banana-shaped blackish resting stages, known as sclerotia or ergot bodies, which replace the head of wheat grain. The ergot bodies present a high toxicity risk for animals and humans due to the alkaloid content. Ergotism, a human disease, is caused by consuming products produced from ergot-contaminated grains. Ergot alkaloids also affect animals, causing bovine miscarriages and intestinal inflammation in sheep.

To prevent potential health risks, fungal infected grains must be identified and removed to avoid them from being used in human food and animal feed. During grading, wheat kernels are visually inspected for any evidence of a fungal infection. When only slight fungal damage is present, a magnifying lens is necessary to examine each suspect kernel to confirm fungal growth. Although severe fungal damaged kernels are easily detectable by human inspectors, the process is slow and challenging when detecting fungal damage at the early stages of fungal development.

Sprout Damage

Sprout damage occurs when wheat kernels undergo germination (Singh *et al.*, 2009). This process is instigated through the presence of water. Sprouting instances commonly occur when harvests take place soon after wet weather conditions and the wheat is not dried promptly. Wet storage conditions also facilitate germination. The process of germination inflicts mechanical damage on the kernels, making them more susceptible to deterioration in storage. The additional cracks and crevices of sprouted kernels make excellent hiding spots which trap microorganisms and promote fungal growth (Singh *et al.*, 2009; Ropelewska & Zapotoczny, 2018). These microorganisms, under the right conditions, grow and eventually spread to neighbouring kernels. Insect infestations also become a problem since they tend to feed on the exposed germs of sprouted kernels (Singh *et al.*, 2009).

Sprouted kernels are not chemically suitable for baking (Singh *et al.*, 2009; Barbedo *et al.*, 2018). The kernels have increased enzymatic activity (Barbedo *et al.*, 2018), which alters their chemical composition to accommodate plant growth. The main enzyme of interest is α-amylase. It actively degrades starch into simple sugars, reducing the overall starch content of the wheat kernel. The protein content and amino acid

profile are also affected due to increased proteolytic activity. For these reasons, the use of sprouted kernels in flour production results in flour of low yield and inferior chemical properties. The subsequent dough is wet, sticky and results in bread with poor characteristics such as a darker crumb colour and inferior texture. Sprout-damaged kernels, therefore, need to be minimised in production.

When performing grading, wheat consignments are inspected for sprouted kernels by visually inspecting a sample either using a magnifying glass or merely with the naked eye. An indication of sprout damage is when the germ significantly emerges from the kernel, breaking the bran that overlays the germ. However, some wheat varieties characteristically have an emerging germ that gives a false impression of sprouting (Singh *et al.*, 2009; Barbedo *et al.*, 2018). There are also wheat varieties that have a characteristic dark brown area across the germ, which are often mistaken for sprouting (Figure 2.1)

Figure 2.1 A digital image demonstrating the differences between sprout-damaged wheat (bottom), a common bread wheat variety in South Africa with a dark brown area across the germ (middle), and other typical bread wheat kernels in sound condition (top).

Visual inspection is, therefore, not always accurate. Another method to determine sprout damage is by determining the Hagberg Falling Number, which is correlated with α-amylase activity (Singh *et al.*, 2009; Barbedo *et al.*, 2018). The Hagberg falling number (HFN) method is the international standard [ICC (International Association for Cereal Science and Technology) 107/1] and is used in the South African wheat grading regulations. Measuring the Falling Number is, however, both destructive and time-consuming, as it uses wheat meal, which cannot be used for online inspection (Barbedo *et al.*, 2018).

Heat Damage

Heat-damaged kernels are defined as kernels that are materially discoloured and damaged by heat (Wang *et al.*, 2001). The leading causes of heat damage are prolonged hot weather, excessive heat from artificial drying practices and improper storage conditions. Besides discolouration, heat also causes severe damage to the chemical components of wheat kernels. Exposure to high temperatures causes proteins to denature which reduces its quality. Since protein, especially gluten, plays a crucial technological role in the formation

of baked products, heat damage has a detrimental effect on the ease of processing, dough formation and final product quality.

Often when heat damage results from poor storage practices, fungal infestations are present too (Ravikanth *et al.*, 2015). Poor storage practices include the presence of undesirable contaminants such as foreign materials, dockage and animal excreta. Dockage is often lighter than the wheat itself and has a higher moisture content. The lighter dockage tends to accumulate along the walls of the storage unit, allowing for hotspot development. Not only does this lead to heat-damaged kernels but promotes fungal growth too, accelerating spoilage.

Heat damage is an important grading factor. Heat-damaged kernels are visually identified by discolouration. The kernels are typically darker in colour, exhibiting shades of reddish-brown or mahogany (Wang *et al.*, 2001; Department of Agriculture, 2016). For visual inspection, heat-damage kernels, unfortunately, do not always exhibit discolouration. Even though the internal wheat protein may be damaged, there could be no visible indication of such, making visual grading unreliable. Other methods considered to assess heat damage in wheat include the use of a mixograph, dye-binding capacity, enzyme inactivation, turbidity and protein extractability (Wang *et al.*, 2001). All these methods are time-consuming, relatively complicated and most definitely not suited for online applications.

Frost Damage and Immature Kernels

The exposure of wheat plants to freezing temperatures has detrimental effects on the quality of the wheat kernels (Dexter *et al.*, 1985). The most damage is inflicted in the milk to the soft dough growth stage of the wheat plant, where extreme weather events, such as frost, causes damage to wheat kernels (Department of Agriculture, 2016). The cold temperatures affect the kernels' physiological processes, such as protein synthesis, which prevent the seeds from reaching full maturity (Preston *et al.*, 1991). The cold temperatures also bring about ice formation which disrupts the cell structure of the wheat kernels. The presence of frost events in the later, more mature stages of the wheat plant, has less of an effect on its physiological development, but may still cause the bran to wrinkle (Department of Agriculture, 2016). In some cases, the bran coat could be rubbed off due to handling, causing a false indication of frost damage.

Frost-damaged kernels are characterised as being somewhat plump with small blisters on the bran layer. The bran layer may even be flaked-off (Department of Agriculture, 2016). The kernels have increased kernel hardness, causing higher milling energy costs and starch damage levels (Dexter *et al.*, 1985). The flour produced from frost-damaged wheat has a higher proportion of shorts (branny material) as it is more difficult to separate the endosperm from the bran of frost-damaged wheat (Preston *et al.*, 1991). The resulting flour is of a decreased yield, darker colour and increased ash content. The baking performance is also affected where reduced dough strength, poorer loaf crumb colour and inferior bread quality is observed.

Immature wheat may be a consequence of a frost event, but they can also occur due to late rain. Immature kernels are characterised by their grass green colour due to incomplete kernel maturation

(Department of Agriculture, 2016). As with frost-damaged grains, immature wheat has increased kernel hardness, causing milling to be more difficult (Preston *et al.*, 1991). Since immature wheat kernels are underdeveloped, the chemical composition and physical properties are different from that of mature wheat. These properties change during the maturing process and depend on the degree of maturity. In a study by (Petrovska-avramenko *et al.*, 2016), it was found that the amount of sugars present in immature grains is higher than that of mature wheat. Therefore, the endosperm texture, which is the most critical compound in wheat flour, is chemically unfit for bread production. As with frost-damaged kernels, immature kernels result in poor loaf crumb colour and reduced dough strength (Tipples, 1980). It was however determined that immature grains have higher levels of B group vitamins and could be a valuable ingredient for the production of other functional foods (Petrovska-avramenko *et al.*, 2016).

Wheat Grading in South Africa

Wheat grading systems are used to evaluate the condition of whole wheat kernels. The primary purpose of grading is to determine the grains' suitability for long-term storage, processing and human consumption. It also facilitates fair trade since the market price is based on the allocated grade. Grading parameters are set out in rigid specifications, which cover a wide range of parameters. It ensures that wheat with reduced quality is assigned to lower grades. A non-subjective grading system is essential to accurately allocate grades, which in turn promotes fairness and eliminates bias tendencies.

In South Africa, a wheat classification system was introduced in 1989 (van der Merwe & Cloete, 2017). Wheat is classified into three hardness categories termed "Bread Wheat", "Soft Wheat" and "Durum Wheat". Bread wheat is also referred to as "Commercial Wheat" or "Common Wheat". The standard according to which bread wheat is graded is stipulated in the South African Agricultural Product Standards Act (Act No. 119 of 1990).

The quality of the wheat is analysed before storage placement in order to sort the grain accordingly. The grading procedure is simple and performed manually by trained personnel. Wheat grading is primarily based on visual characteristics; however, there are additional measurements, which include hectolitre mass, falling number, moisture content and protein content measurements (Department of Agriculture, 2016). In general, wheat consignments are inspected in terms of odour, taste and colour to ensure that it is typical of healthy wheat.

On arrival of a wheat consignment, a representative sample is extracted per the prescribed regulation (Addendum 1) (Department of Agriculture, 2016). For consignments delivered in bags, a bag probe is used to sample a minimum of 10% of the total number of bags. The composite sample must consist of wheat from at least 25 bags, regardless of the total number of bags. For consignments delivered in bulk, a bulk sampling apparatus is used to make up the sample. Wheat is taken throughout the entire depth of the consignment in at least six different places chosen at random. The composite sample must have a total mass of at least 10 kg and be thoroughly mixed.

To determine the hectolitre mass, moisture content, protein content and falling number, any suitable instrument is allowed provided that the apparatus complies with the specifications detailed in either the ISO (International Organization for Standardization), the AACCI (American Association of Cereal Chemists International) or the ICC (International Association for Cereal Science and Technology) (Addendum 1) (Department of Agriculture, 2016). In industry, a whole grain analyser is typically used to provide the hectolitre mass, moisture content and protein content.

A falling number instrument is used to determine the falling number that indicates α-amylase activity. It operates by placing a heated mixture of ground wheat and water in a precision bore glass tube with a plunger. The time it takes (in seconds) for the plunger to reach the bottom of the container is known as the falling number. A rapid visco-analyser operates in a similar manner where the ground wheat and water mixture is subjected to a heating and stirring protocol, and a final viscosity measurement is determined. The results of these methods indicate that high-quality wheat makes a thicker paste and has a higher falling number or final viscosity. A high falling number or final viscosity, therefore, means a low amount of α-amylase activity, suggesting no germination. The South African Bread Wheat Standard indicates that a falling number greater than 250 is suitable for bread making (Addendum 1) (Department of Agriculture, 2016).

To perform the visual inspection section of the standard, smaller working samples are sourced from the bulk representative sample for each deviation (Addendum 1) (Department of Agriculture, 2016). The different deviations are summarised in Table 1 in Addendum 1, where the nature of the deviation is accompanied by the maximum percentage permissible to obtain a specific grade. The deviations include screenings, frost-damaged wheat, foreign matter, mechanically damaged wheat, heat-damaged wheat, field fungi infected wheat, storage fungi infected wheat and the presence of other grains and unthreshed ears. For each of these deviations, a working sample of screened wheat is used. Screened wheat is obtained by sieving 500 g of the consignment sample using a standard sieve (see definition in Addendum 1). The wheat that remains in the sieve is known as “screened wheat”, and the material that passes through the sieve is known as “screenings”. Each wheat kernel in the sample is individually inspected to identify any content other than sound wheat, such as defective kernels or foreign material. Only one type of defect may be analysed at a time. Therefore, if the sample is inspected for frost damage and heat-damaged kernels are present, the heat-damaged kernels must be ignored for the time being. The mass of the defective kernels for each defect is determined as a percentage of the mass of the working sample.

The South African Bread Wheat Standard does not provide the information regarding ergot. Instead, the standard refers to the Foodstuffs, Cosmetics and Disinfectants Act (Act 54 of 1972) which stipulates the number of noxious seeds and ergot sclerotic infected kernels tolerable (Department of Agriculture, 2016). Wheat is deemed as contaminated when it contains more than 0.05% (m/m) of Ergot sclerotia.

Wheat grading is vital to every entity in the production chain, from the farmers to the final consumers. It is, therefore, a considerable responsibility and should be executed as accurately and thoroughly as possible, yet outdated traditional methods are still being utilised. Human visual inspection is

the most common grading technique around the world and has been proven to be inaccurate, time-consuming and labour intensive. There is a need to develop a rapid, reliable, robust and simple method to replace the traditional methods. Spectroscopic methods, such as NIR spectroscopy and NIR hyperspectral imaging, could be the answer and lead to superior quality control.

Near-Infrared Spectroscopy

The potential of NIR spectroscopy technique was first recognised by Karl Norris in the 1960s (Pasquini, 2018). Spectroscopy is the study of the interaction between matter and electromagnetic radiation (Osborne, 2006). In other words, spectroscopy utilises light to analyse a sample by describing the energy transfer between the two. This is based on the principle that all matter absorbs, reflects, scatters, or emits electromagnetic radiation in a characteristic manner, representative of its physical and chemical nature. Near-infrared (NIR) spectroscopy specifically utilises the NIR region of the electromagnetic spectrum (located between the visible and infrared region) which ranges from 780 to 2500 nm (Osborne, 2006). This region was discovered by William Herschel and is the first part of the electromagnetic spectrum found that is invisible to the naked eye (Pasquini, 2018). The photon energy from this region typically excites the major X–H bonds of organic molecules, namely C–H, O–H, N–H and S–H (Osborne, 2006; Manley, 2014). The energy of these bonds exists in the form of stretching and bending vibrations. Stretching vibrations are described as a constant change in bond length between two atoms, and bending vibrations are described as a change in the bond angle. The combination of these vibrations causes the molecules to exhibit anharmonic behaviour (Osborne, 2006). When an object is subjected to NIR radiation, the molecules experience changes in their vibrational energy where they selectively absorb the energy that corresponds to their vibrational energy. Therefore, absorption only occurs when the electromagnetic radiation is of a specific frequency that coincides with that of a molecular bond in the sample. The energy that is not absorbed is transmitted or reflected. The outcome is a spectrum that consists of absorption bands which correspond to overtones and combination bands caused by the anharmonicity of the vibrating atoms (Osborne, 2006). The resulting spectrum is seen as a spectral fingerprint that contains information about the components of the analyte investigated (ElMasry & Nakauchi, 2016)

The uses and applications of NIR spectroscopy in food science research are abundant (Porep *et al.*, 2015; Fu & Ying, 2016). The first notable work of NIR spectroscopy in cereal science was performed by Williams (1975), who measured protein and moisture content of wheat. NIR spectroscopy has since demonstrated to be one of the most useful techniques for monitoring and controlling product quality and safety (Fox & Manley, 2014). NIR spectroscopy, as mentioned, is ideal for the analysis of organic molecules, making it suitable for the study of food materials. It can also measure several constituents simultaneously (Osborne, 2006). The technology is fast, easy to use, inexpensive, non-destructive and usually requires no sample preparation. These advantages make NIR spectroscopy a practical solution for routine analysis in the food industry. It has shown to be the most promising technique for the inspection of cereals (Manley, 2014;

Roberts *et al.*, 2018), where it has successfully been applied in wheat studies to determine the geographical origin (Zhao *et al.*, 2013); the presence of mycotoxins (Pettersson & Åberg, 2003); the presence of external and internal insect infestations (Ridgway & Chambers, 1996); different wheat varieties (Bertrand *et al.*, 1985) and to predict various components such as the protein and moisture content (Dowell *et al.*, 2006). This indicates that NIR spectroscopy is a useful technique for detecting and analysing various properties that influence the quality and safety of wheat.

Despite the high success rate and the advantages that NIR spectroscopy has to offer, there is one significant shortfall which is the absence of spatial information. With additional spatial information, the chemical distribution of the sample can be visualised, and chemicals of interest can be located. This gives rise to the creation of prediction maps that can be used for smart technology such as automatic sorting systems. Spatial information also serves as a source of additional features that can be included when creating classification models. Features from the sample such as size, shape and texture can be extracted, further contributing to the performance of the model. The combination of NIR spectroscopy and digital imaging, known as NIR hyperspectral imaging, proposes an ideal technique for superior food inspection.

Hyperspectral Imaging

Introduction to NIR Hyperspectral Imaging

Hyperspectral imaging (HSI) is an advanced analytical technique used to perform multi-constituent analyses on practically any sample. It integrates spectroscopy and computer vision which provides spectral and spatial information simultaneously (Gowen *et al.*, 2007). A hyperspectral image contains a wealth of information where each pixel in the image contains a complete spectrum; therefore, each pixel has a spectral fingerprint. Alternatively, the spatial dimensions can be viewed in conjunction with a specific waveband, creating a chemical image or chemical map. Since hundreds of consecutive wavebands within a designated range are used, an equivalent amount of chemical images are created for the same sample (Figure 2.2).

The concept of HSI emerged in the 1980s (Goetz *et al.*, 1985; Goetz, 2009). The invention of a sensor for recording images, the charge-coupled device (CCD), was the key development in hyperspectral imaging. In fact, the CCD was a major breakthrough for all digital camera technology. Willard Boyle and George Smith are the great minds behind this invention and in 2009 shared half of the Nobel Prize in Physics "for the invention of an imaging semiconductor circuit – the CCD sensor" (Boyle, 2010; Smith, 2010). The other half, for interest sake, was awarded to Charles Kuen Kao "for groundbreaking achievements concerning the transmission of light in fibres for optical communication". HSI was initially developed for the field of remote sensing, but with recent advances in computer technology, it gained impressive applications in many areas such as astronomy, medicine, agriculture, pharmaceuticals and food science (Caporaso *et al.*, 2018a). Computer vision systems imitate the principle of human vision using red, green and blue wavebands. These systems are already utilised in food applications such as automatic apple sorting systems or potato inspection systems (ElMasry *et al.*, 2012a; Sofu *et al.*, 2016). The systems detect various visual defects based on shape,

size and colour such as surface blemishes; however, the cameras are limited to the visible range (380 to 700 nm) of the electromagnetic spectrum and are insufficient for detecting chemical and biological parameters (Razmjooy *et al.*, 2012; Manley, 2014; Sofu *et al.*, 2016). On the other hand, conventional NIR spectroscopy studies the interaction of matter and light extending beyond the visible region of the electromagnetic spectrum to wavelengths between 780 to 2500 nm (Osborne, 2006; Manley, 2014). Characteristics that cannot be seen by the human eye can be revealed when using NIR spectroscopy, however without spatial resolution. Another downside is that NIR spectroscopy only provides an average spectrum from randomised points measured on a bulk sample. Combining NIR spectroscopy and computer vision in a NIR-HSI system counteracts these limitations by providing spatially resolved spectral information of an entire bulk sample (Feng & Sun, 2012; Moghaddam *et al.*, 2013). This makes HSI ideal of the analysis of heterogeneous samples and for analysing multiple objects at a time. Furthermore, HSI possesses the same advantages of spectroscopy such as being rapid, accurate, non-destructive and requiring little sample preparation, allowing it to be highly suited for the non-invasive evaluation of food.

NIR Hyperspectral Data Acquisition and Data Structure

HSI generates a substantial amount of data (Gowen *et al.*, 2007). The data is presented in a three-dimensional cube, known as a *hypercube*, which holds spatial (x and y) and spectral information (λ) (Figure 2.2).

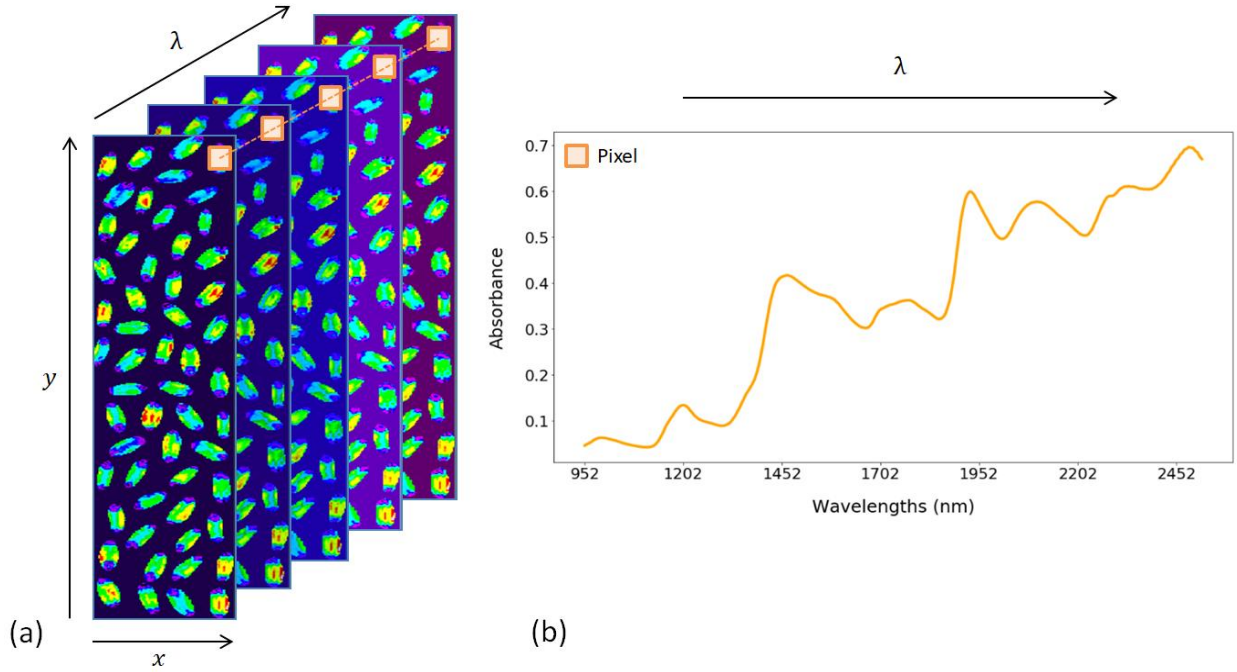


Figure 2.2 (a) A NIR hyperspectral image, hypercube, comprised of two spatial ($\boldsymbol{x}$ and $\boldsymbol{y}$) and spectral ($\boldsymbol{\lambda}$) dimension. The image plane at each respective wavelength can be visualised. Different colours indicate different absorbances; **(b)** The spectrum of a single pixel in the sample.

It is currently not possible to attain the data in all three dimensions simultaneously (Gowen *et al.*, 2007). The three-dimensional cube is assembled by concatenating two-dimensional images in sequence. To generate the hyperspectral image, three main approaches are available, namely: the focal plane scan, the point scan and the line scan approach (Boldrini *et al.*, 2012). These approaches are also referred to the staring imager, whisker-broom and push-broom configuration, respectively. The focal plane scan approach compiles a hypercube by acquiring an image of the entire field of view at each wavelength. Each image is then sequentially stacked to form the hypercube. During the whole acquisition period, all the components of the instrument remain stationary. The point scan approach compiles a hyperspectral image by collecting spatial and spectral information from the sample pixel per pixel. This configuration is ideal for samples that require high spatial resolution, such as microscopic imaging. A drawback of this approach is that it is quite time-consuming. The line scan approach compiles a hyperspectral image by moving the sample into the camera's field of view using a linear translation stage. All the spectral information is captured across one spatial dimension (*x*) and repeated line-for-line for the length of the sample, gradually capturing the second spatial dimension (*y*). This configuration is the most suited for online applications as it could be installed on a

conveyor belt system to rapidly and non-destructively perform quality control functions in real-time (Boldrini *et al.*, 2012).

A line-scan instrument is depicted in Figure 2.3. The instrument generally consists of an illumination source, an objective lens, spectrograph, camera and computer equipped with image acquisition software. A tungsten-halogen lamp is the most common source of illumination, which provides homogenous lighting (Elmasry & Sun, 2010). These lamps are also durable, have good stability and are capable of emitting a broad spectral range of light. Quartz-halogen lamps, light-emitting diodes (LEDs), tunable lasers and heated xenon lamps are other sources of illumination. The camera consists of a spectrograph that is comprised of either a prism-grating-prism (PGP) monochromator, liquid crystal tunable filter (LCTF) or acousto-optic tunable filter, which is responsible for wavelength separation and selection. The camera also is contains a detector made of either indium-gallium-arsenide (InGaAs), lead-sulfide (PbS), indium-antimonide (InSb) or mercury-cadmium-telluride (HgCdTe).

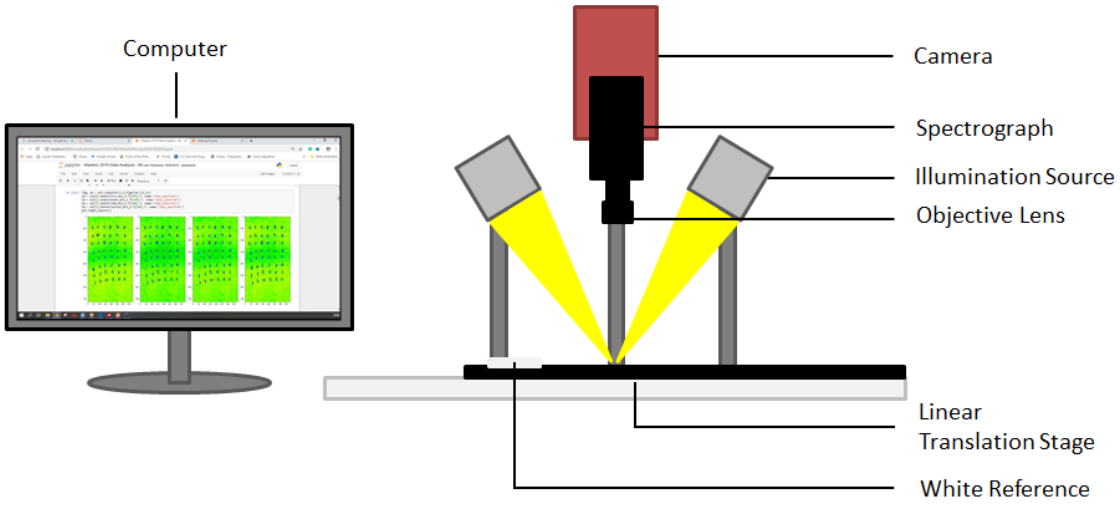


Figure 2.3 Schematic of a pushbroom NIR hyperspectral imaging system

The data is collected, stored and analysed on a computer coupled to the HSI system. A calibration step using a white reference should be regularly performed while imaging to ensure data standardisation (Boldrini *et al.*, 2012). To effectively use the hyperspectral data captured, chemometrics is needed to extract useful chemical information from the hypercube (Amigo *et al.*, 2013).

Chemometrics

Chemometrics is a term first introduced by Svante and Kowalski in 1972 (Mutihac & Mutihac, 2008). In essence, chemometrics is a form of data mining. Data mining is a term used in computer science and is the analytical process of exploring large amounts of data to find consistent patterns. The findings are used to create models that are then applied to new data to predict or estimate an outcome. The application of data mining in chemical fields gave rise to the term "Chemometrics". Massart *et al.* (1988) defined chemometrics as "*the chemical discipline that uses mathematical, statistical, and other methods employing formal logic to*

design or select optimal measurement procedures and experiments, and to provide maximum relevant chemical information by analysing chemical data". Simply stated, chemometrics is a discipline where statistical algorithms are created or "borrowed" from other mathematical fields and applied to chemical data (Mutihac & Mutihac, 2008). The general purpose of these algorithms is to extract meaningful information from a dataset and identify significant patterns to perform qualitative or quantitative analysis. The performance of each algorithm differs, depending on the dataset, the desired outcome and, to some extent, the instrumentation (Burger & Gowen, 2011).

Data Exploration

Data exploration is the first phase in data analysis. Before any attempts can be made to transform the data, it should first be thoroughly examined. A practical approach is by creating visual interpretations of the data. The construction of an image for each wavelength is useful to view the chemical variation. The absorbance spectrum for each pixel or group of pixels is useful to identify functional groups, which provides more information about the chemical constituents of the analyte. The absorbance values also relate to the concentration of the different components through the Beer-Lambert law (Rinnan *et al.*, 2009; Ravikanth *et al.*, 2015). A typical NIR spectrum of sound wheat is depicted in Figure 2.4 (Manley, 2014).

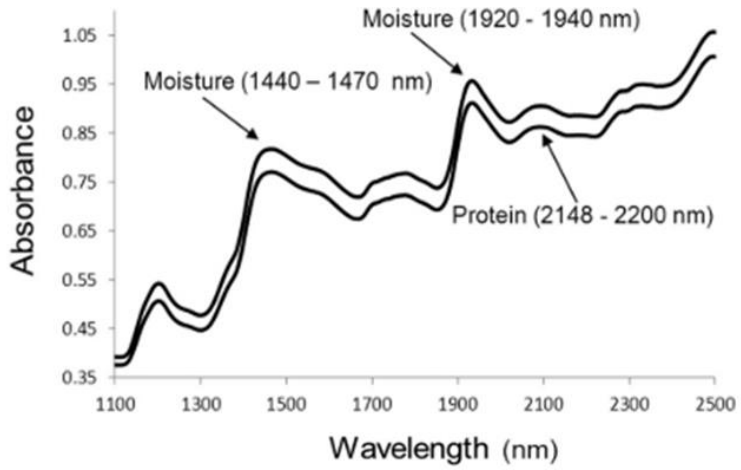


Figure 2.4 NIR spectra of a whole wheat kernel with moisture (1440-1470 nm and 1920-1940 nm), protein (2148 to 2200 nm) and starch (1130-1190 nm) absorption bands indicated (Manley, 2014).

Principal Component Analysis

Principal component analysis, first proposed by Pearson in 1901 (Pearson, 1901), is one of the most useful data exploratory techniques today (Amigo *et al.*, 2013). It finds the directions of maximum variance in high-dimensional datasets and projects it onto a lower-dimensional subspace while retaining only relevant information (Burger & Gowen, 2011). The data is condensed into fewer parameters, known as principal components (PCs), which describe the variance within the data. The first PC is in the direction of the observations that vary the most. Each succeeding PC is calculated orthogonally to the preceding one.

Because the PCs are orthogonal, they are uncorrelated. Multivariate curve resolution (MCR) is an alternative method that can be applied (Burger & Gowen, 2011; Boldrini *et al.*, 2012). This method decomposes spectra of multicomponent mixtures into the basic qualitative and quantitative responses given by each of the components in the sample analysed. It can determine the number of constituents in a mixture, their associated spectra, and their estimated concentrations based on Beer-Lambert's Law. However, MCR methods have only recently been introduced to image analysis, and PCA remains the most popular tool in overviewing the primary sources of variance in the hypercubes (Boldrini *et al.*, 2012).

To perform PCA on the hypercube, a three-dimensional array denoted as $D(X, Y, \lambda)$, is unfolded into the matrix $D(X*Y, \lambda)$. PCA can then be applied to decompose the unfolded matrix into a set of two sub-matrices known as the score (T) and loading (P') matrices (Amigo *et al.*, 2013). The main and relevant spatial information is preserved in the score images, and the spectral information is preserved in the loadings. The noisy information is omitted in the residuals. The loading and score vectors are arranged in decreasing order of variance. The equation for PCA is displayed in Eq. 2.1.

$$D = TP' + E \tag{2.1}$$

Pre-processing

Pre-processing is the most critical step in model development (Rinnan *et al.*, 2009). It is required to remove physical phenomena in the spectra to improve the subsequent exploratory analysis, multivariate regression or classification model. Numerous physical variations in a sample such as the morphology, particle distribution, particle size differences and density variances lead to light scattering that affects the measured spectra. A suitable pre-processing technique is dependent on the nature of the data and is difficult to determine prior to model validation (Rinnan *et al.*, 2009). The optimal pre-processing technique can only be determined by identifying the model that delivers the lowest statistical error during model validation. The pre-processing techniques used in classical spectroscopy are also relevant for hyperspectral imaging. To correct for particle size or other scattering effects, standard normal variate (SNV) (Barnes *et al.*, 1989) and multiplicative scatter correction (MSC) (Esbensen & Geladi, 1989) are often used (Rinnan *et al.*, 2009). MSC and SNV provide similar results, however, SNV is simpler than MSC (Boldrini *et al.*, 2012). The equation for SNV is provided in Eq. 2.2, where $A_{(SNV)}$ is the corrected absorbance value; A_{ij} is the measured absorbance value; I is the spectrum counter; J is the absorbance value counter of i^{th} spectrum; μ is the mean absorbance value of the uncorrected i^{th} spectrum; and σ is the standard deviation of the absorbance values of i^{th} spectrum. Spectra treated with SNV always have a zero mean value and a variance equal to one. They are therefore independent of the original absorbance values.

$$A_{snv} = \frac{A_{ij} - \mu}{\sigma} \tag{2.2}$$

To correct for baseline shifts or other non-linear effects in spectra, de-trending (Barnes *et al.*, 1989) can be applied. Linear detrending is done by estimating a polynomial trend, the least-squares fit of a straight line $y_{v,s}^{m}(j)$, and then subtracting it the from the original spectra. The function is displayed in Eq. 2.3. The *m* variable equal to one (*m* = 1) is the degree of the polynomial to eliminate the linear trend.

$$\tilde{Y}_s(j) = Y(j) - y_{v,s}^{m}(j) \tag{2.3}$$

Removing the linear trend from the data allows the analysis to be focussed to the variations in the data around the trend. Other useful pre-treatments are spectral derivatives such as Norris-Williams (NW) (Norris, 1983) and Savitzky-Golay (SG) (Savitzky & Golay, 1964) (Rinnan *et al.*, 2009). These methods help resolve overlapping peaks (Boldrini et al., 2012; Manley, 2014). However, derivatives, especially 2nd derivatives, should be used with caution as they tend to amplify noise. SG has a smoothing algorithm that can remove the amplification of random noise, but care should be taken not to smooth away pursued information.

Multivariate Data Analysis

NIR spectra are very complex and consist of many overlapping peaks and broad bands associated with overtones and combinations of vibrational bonds (Manley, 2014). The interpretation thereof is, therefore, challenging and requires chemometrics to extract relevant information (Mobaraki & Amigo, 2018). Chemometrics is directly linked to the target pursued and refers to the application of multivariate data analysis. Multivariate data analysis is the chemometric tool used to analyse the data. Developments in multivariate data analysis date back to the beginning of the nineteenth century (James *et al.*, 2014). Legendre and Gauss developed the *least-squares method,* which became a standard approach in statistical learning (Stigler, 1981). The earliest implementation of this method is now known as *linear regression*. Linear regression is a quantitative approach and was first successfully applied in astronomy. To produce a qualitative response, Fisher proposed *linear discriminant analysis* in 1936 (Fisher, 1936). After that, in the 1940s, *logistic regression (McCullagh & Nelder, 1989)* was developed as an alternative approach to linear discriminant analysis. These statistical learning methods were termed *generalised linear models* by Nelder and Wedderburn in the 1970s, and many more linear models were developed thereafter. The development of non-linear methods was simply not computationally possible in those years. With the advancements in computing technology, non-linear methods were no longer computationally restricted, and Breiman *et al*. introduced classification and regression trees (Breiman, 2001). Ever since then, machine learning and other statistical learning disciplines have been inspired, creating a broader statistical community.

For this study, supervised classification analyses were of interest. Supervised learning is when a model is trained using known input and output data so that it can predict future outputs (Mutihac & Mutihac, 2008; Brereton, 2015). Unsupervised learning is based on clustering and merely finds intrinsic structures or hidden patterns in input data without being trained. Within the realm of classification algorithms, the choice

of algorithm depends on the nature of the problem, the size of data set, the ease of implementation and the required accuracy of prediction. Popular supervised classification methods used to perform discriminant analysis include Logistic Regression (McCullagh & Nelder, 1989), Partial Least Squares Discriminant Analysis (PLS-DA) (Wold *et al.*, 1987b), Linear Discriminant Analysis (LDA) (Fisher, 1936), K-Nearest Neighbours (KNN) (Sebestyen, 1962), Decision Trees (Quinlan, 1986), Random Forests (RF) (Breiman, 2001) and Support Vector Machines (SVM) (Cortes & Vapnik, 1995). These techniques make use of machine learning, a broad subfield in artificial intelligence (AI), that automatically extracts information from data using computational and statistical methods embedded in algorithms that instruct computers how to learn (Mutihac & Mutihac, 2008). Machine learning methods are capable of revealing underlying patterns in large data sets, and its application is, therefore, necessary for hypercubes.

Logistic regression and LDA are two of the most important classical classification methods (James *et al.*, 2014). Logistic regression is used to determine the probability of *Y* belonging to a particular category where the dependent variable is binary (two-class response). It uses a sigmoidal function that takes in a value and outputs a *probability* between zero and one. Using a threshold, typically 0.5, *Y* will either be predicted as one class or the other. A simplified equation for logistic regression is presented in Eq. 2.4, where p = predictive outcome; X = single predictor variable; β_0 = intercept; β_1 = slope.

$$p(X) = \frac{e^{-(\beta_0 + \beta_1 X)}}{1 + e^{-(\beta_0 + \beta_1 X)}} \tag{2.4}$$

Linear Discriminant Analysis (LDA) is a linear technique that calculates an optimal linear projection that maximises the variance between classes (Fisher, 1936). LDA is similar to PCA in that they both reduce dimensions and find the components that maximise the variance within a data set (James *et al.*, 2014). What makes LDA different from PCA is that LDA does not just find the components with the highest variation but also uses the variation to maximise the separability of the different classes. LDA creates a new axis in the multidimensional space and projects the data with the highest variation onto the new axis in a way that maximises separation. To create the new axis, two important criteria need to be addressed. Firstly, the distance between the means of the classes needs to be maximised. Secondly, the variation within the class needs to be minimised. The variation is also known as "scatter". The mean values and the scatter are considered simultaneously in Eq. 2.5, where μ_1 is the mean of class 1, μ_2 is the mean of class 2, s_1 is the scatter of class 1 and s_2 is the scatter of class 2. The numerator is squared to account for not knowing which mean value is larger. Linear surfaces, known as hyperplanes, are used to separate the classes in the multidimensional space. To assign an unknown observation to a class, the distance between the unknown observation and each class is calculated. The class "nearest" to the unknown sample is allocated. LDA does, however, depend on restrictive assumptions such as linearity, normality or independence among predictors, which could result in collinearity problems.

$$\frac{(\mu_1 - \mu_2)^2}{s_1 + s_2} \tag{2.5}$$

For more than one class, Eq. 2.5 is adjusted, as seen in Eq. 2.6.

$$\frac{(\mu_1 - \mu)^2 + (\mu_2 - \mu)^2 + (\mu_3 - \mu)^2}{s_1 + s_2 + s_3} \tag{2.6}$$

PLS-DA is another method similar to LDA and PCA (James *et al.*, 2014). It is a dimension reduction method that tries to improve the separation between classes, by maximising the covariance between two matrices (*X* and *Y*). It does this by identifying the components which best account for the variances between the two categories, considering the independent variables (*X*) and the dependent variable (*Y*) (Figure 2.5). PLS-DA scores are then calculated to maximise the relationship between *X* and *Y*.

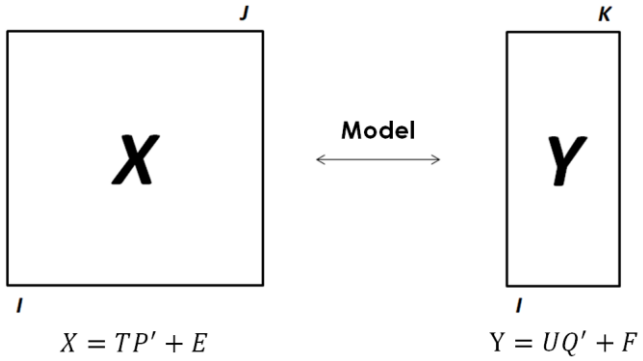


Figure 2.5 Basic structure of the PLS-DA method.

If PCA were to be performed on the two matrices (*X* and *Y*), the matrices could be described as the scores multiplied by the loadings, however, in PLS-DA, the scores are calculated in such a way that the first score in X (u_1) has maximum covariance (r) with the first score in Y (t_1) (Eq. 2.7).

$$u_1 = r_1 t_1 \tag{2.7}$$

Using the first score in *X* (T), the first score in *Y* (U) can be calculated, which can then be used to predict *Y*. Therefore, the key in PLS-DA is to identify the components in such a way that the score values have maximum covariance. PLS-DA is also able to overcome collinearity problems.

k-Nearest Neighbours (KNN) estimates the distance between an unknown observation and the training data points (James *et al.*, 2014). The *k*-value determines how many of the trained observations closest to the unknown observation must be taken into consideration. If *k* is equal to three, then the three

closest observations to the unknown observation will be evaluated. The unknown observation is classified to the class to which the majority of the *k* objects belong. An optimal value for *k* can be determined experimentally by evaluating the error rate with different *k* values in the form of a scree plot. It is crucial to select a good value for *k* as it has a strong effect on the KNN model. KNN is mathematically presented in Eq. 2.8, where K is a positive value; x_0 is the test observation; N_0 is the K points in the training data that are closest to x_0. The KNN classifier, therefore, estimates the conditional probability for class j as the fraction of points in N_0 whose response values equal j.

$$Pr(Y = j|X = x_0) = \frac{1}{K}\sum_{i \in N_0} I(y_i = j) \tag{2.8}$$

The decision tree technique is currently one of the most popular classification algorithms (Mutihac & Mutihac, 2008). The classifier is in the structure of a tree (Figure 2.6), where each split is known as a node.

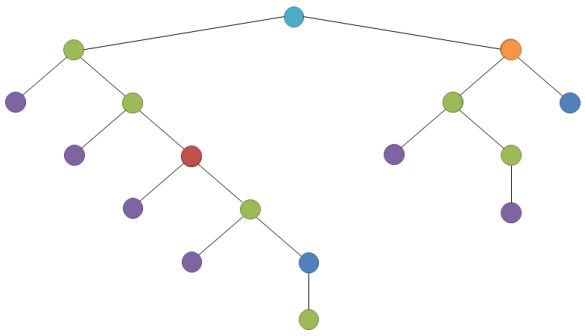


Figure 2.6 General structure of a decision tree.

The basic process of developing a decision tree is to split the predictor space into distinct, non-overlapping regions for all the possible outcomes (James *et al.*, 2014). This is done by recursive binary splitting, where the classification error rate is used as criteria for creating the splits. The classification error rate is the number of training observations in that region that do not belong to the most common class (James *et al.*, 2013). An unknown observation in a given area is assigned to the most frequently occurring class in that region.

Random forests (RF) is an ensemble of decision trees. How the trees are constructed for decision trees and random forests is different. For decision trees, each node in the tree must be selected in such a manner so that maximum information is gained. Random forests only consider a small subset of features rather than all of the features of the model when constructing a node. This is important because if a data set contains a few strong predictors, these predictors will consistently be chosen to be the top nodes in the random forest trees. All the trees will have a similar structure, resulting in highly correlated trees, which will cause overfitting. The number of features considered at every split is approximately equal to the square root of the total number of features.

Support Vector Machines (SVM) is a widely used statistical learning algorithm (James *et al.*, 2014). The algorithm aims to find a margin, known as a hyperplane, which separates different categories. The best hyperplane creates the most substantial separation between the two classes. The support vectors are the input vectors that touch the boundary of the margin. A visual representation of SVM is depicted in Figure 2.7.

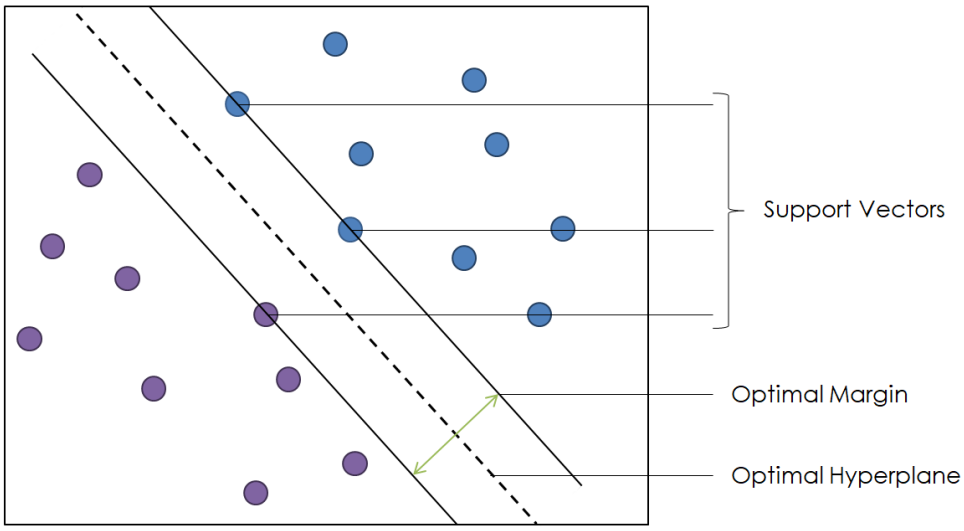


Figure 2.7 An example of SVM and linear separating hyperplanes.

If the data is significantly non-linear, SVM is ideal as different configurations and transformations are possible using linear, nonlinear, polynomial, radial basis function (RBF) and sigmoid kernels. The radial basis function is the most commonly used kernel function. To achieve optimal performances when conducting SVM using the RBF kernel, the parameters *penalty coefficient C* (cost of misclassification) and the *gamma* should be optimised. The choice of *C* is important as it determines the extent to which the model under-fits or over-fits the data. The margins of the hyperplane can be described as “soft margins” or “hard margins”. A "soft margin" allows some of the training samples to be misclassified, which often leads to a better overall fit. A “hard margin” could result in overfitting, leading to more reduced classification performance. The process of choosing a C value involves trading error penalty for stability. The gamma parameter defines how far the influence of a single training example reaches. A small value means 'far', and a large value means 'close'. A large gamma generally leads to high bias and low variance models, which is undesirable. A visual representation of the *penalty coefficient C* and the *gamma* parameter is depicted in Figure 2.8. A grid-search procedure can be applied to achieve the optimal values for these parameters.

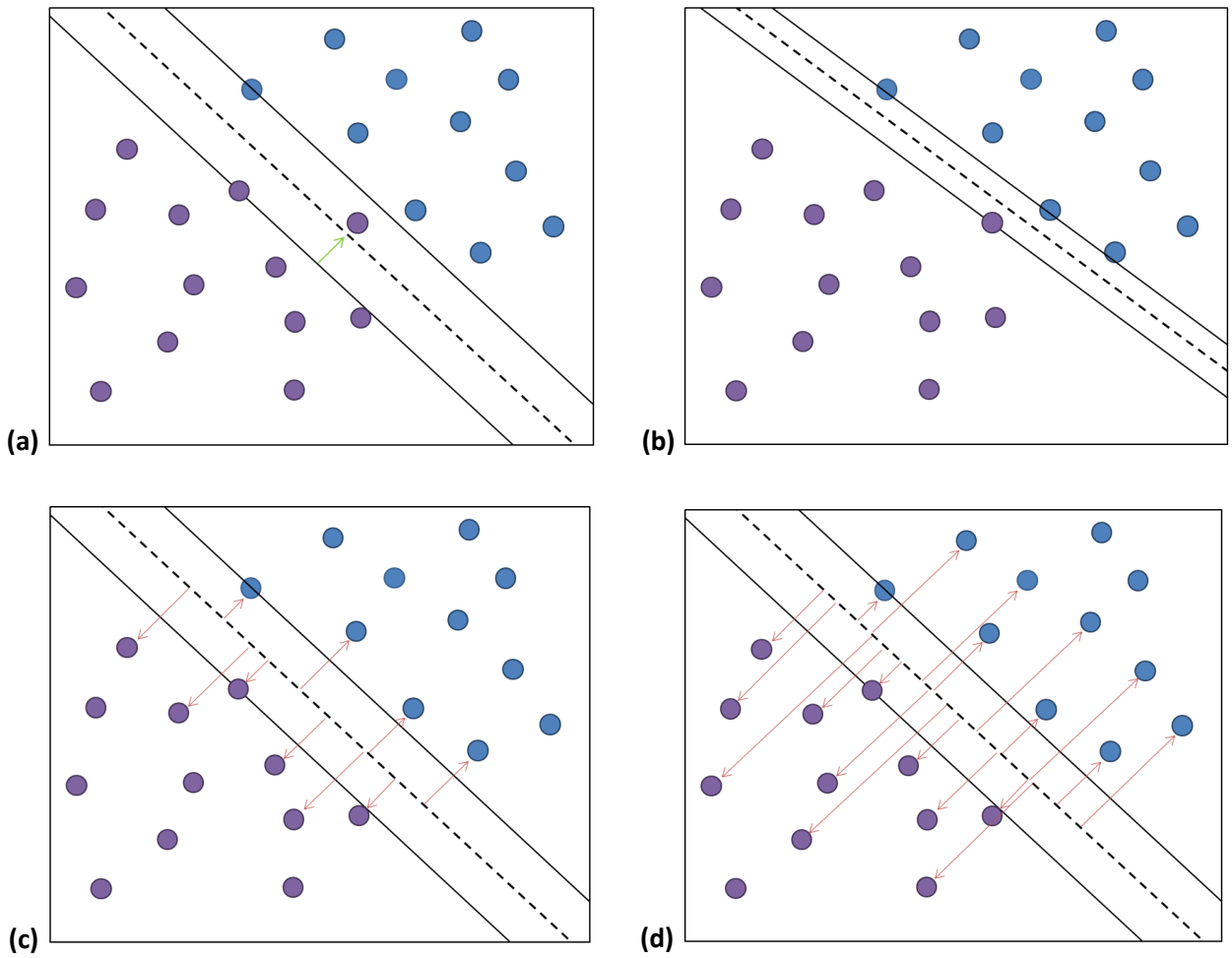


Figure 2.8 An example of **(a)** soft margin SVM with one sample violating the constraints; **(b)** hard margin SVM with no samples violating the constraints; **(c)** a high gamma only considering nearby samples; **(d)** a low gamma only considering far away samples.

Cross-Validation

Bias-Variance Trade-Off is the fundamental understanding of the model's performance and has to do with under or overfitting (James *et al.*, 2014). It is the point where the training error goes down, as it should, but the test error starts to increase. Often it is due to adding too much complexity to the model. At this point, the bias trade-off begins to over-fit. A model with low bias and high variance is desired.

Cross-validation is a statistical technique used to assess the performance of a classification model on a limited sample size. A portion of the data is used as training data, while the remaining part is reserved for testing data. The model is then evaluated by testing its performance on the testing data. The reason different data is necessary for training and testing is to prevent overfitting. For a model to have a good performance, a sensible cross-validation strategy is needed. The most common cross-validation techniques are k-fold cross-validation and Leave-one-out cross-validation (LOOCV).

k-fold cross-validation has a single parameter called *K* that indicates the number of groups a data set is to be split. The number of folds chosen depends on the size of the data. Lower *K*-values results in low

variance and high bias, while higher *K*-values results in high variance and low bias. The general procedure of k-fold is to shuffle the dataset randomly, split the dataset into K-groups and reserve each unique group as testing data while the remaining data (K-1) is used to fit the model. The model is then evaluated on the test set.

In the Leave-one-out cross-validation approach, only one data point is reserved from the dataset, and the model is trained on the rest of the data. This process iterates for each data point. Since all the data points are used, the bias will be low. However, if a data point turns out to be an outlier, it can lead to a higher variation resulting in overfitting.

There are a few requirements that need to be considered when performing cross-validation. Firstly, a large portion of the dataset should be used to train the model, or else the underlying trend in the data will be lost and will lead to an increase in bias. The ratio of training and testing data is also essential as a smaller amount of data points can lead to a variance error. Lastly, the cross-validation should be performed multiple times by changing the train and test dataset distribution. This will help validate the model's performance more thoroughly. The k-fold cross-validation technique satisfies all these requirements. It is also simple to understand and typically results in a less biased estimate of the model's performance compared to that of a simple train/test split.

In addition to estimating the performance of the model, cross-validation is also used to optimise or tune the hyperparameters for a learning algorithm. Hyperparameter tuning is the process of selecting the values for a model's parameters that maximise the accuracy of the model. By testing many different hyperparameter combinations and evaluating the performance of the resulting model, the optimal settings can be determined. Therefore hyperparameter tuning relies more on experimental results rather than theory. To iterate through the different combinations of hyperparameters effectively, a grid search can be performed. A grid search performs hyperparameter optimisation for a given model (such as SVM) by evaluating the model with a different combination of hyperparameters. The models are evaluated using cross-validation and the one that outperforms the others is selected.

Applications of Hyperspectral Imaging and Chemometrics for Wheat Evaluation

The applications of hyperspectral imaging for food analysis are numerous, especially for the inspection of cereal grains. Hyperspectral imaging has been applied to determine the kernel hardness (Gorretta *et al.*, 2006; Manley *et al.*, 2009; Williams *et al.*, 2009; Mahesh *et al.*, 2015; Erkinbaev *et al.*, 2019), chemical composition (Cogdill *et al.*, 2004; Mahesh *et al.*, 2015; Caporaso *et al.*, 2018b) and variety (Mahesh *et al.*, 2008; Wang *et al.*, 2015a; Wang *et al.*, 2015b) of the grain. Physical defects that affect the quality and safety of the grain has also been investigated, where most studies focus on fungal contamination (Delwiche *et al.*, 2011; Shahin & Symons, 2012; Vermeulen *et al.*, 2012; Williams *et al.*, 2012; Barbedo *et al.*, 2015; Ropelewska & Zapotoczny, 2018; Delwiche *et al.*, 2019; Shen *et al.*, 2019), sprouting (Singh *et al.*, 2009; Xing *et al.*, 2010; Arngren *et al.*, 2011; McGoverin *et al.*, 2011; Barbedo *et al.*, 2018), insect damage (Singh *et al.*, 2010) and the

presence of foreign materials (Ravikanth *et al.*, 2015). The general information of these studies is displayed in Table 2.1. The common aim amongst most of these studies is to improve on current grading practices, especially in the areas that rely on human visual inspection. Studies have reported that visual inspection does not provide an accurate estimation of quality as the damage is not always visible to the naked eye. In cases of slight damage, a magnifying glass is necessary to confirm whether a kernel is defective or not. This process is slow, tedious and often inaccurate due to human error and subjectivity. Furthermore, a safety risk is involved if grading is not appropriately performed. More rapid and accurate methods are necessary to meet the needs of the cereal industry.

The detection of *Fusarium*-damaged kernels is one of the most researched defects in cereals. *Fusarium* causes a decrease in yield, has a severe economic impact and most importantly has harmful effects on human and animal health (Ropelewska & Zapotoczny, 2018). Because it is a considerable safety risk, it is understandable that *Fusarium* infected wheat is exhaustively studied. A study was done by Delwiche *et al.* (2011), where sound and *Fusarium*-damaged kernels were imaged at wavelengths 1000 to 1700 nm as well as in the visible range of 400 to 1700 nm. A linear discriminant analysis (LDA) classifier was used to develop classification models. The models were based on the work of Polder *et al,* 2005, where the authors determined that using a ratio of two wavelengths was just as effective as using the entire spectrum. Wavelength selection is useful as hyperspectral imaging instrumentation is costly. A multispectral imaging system offers a more financially feasible option for the industry, where only a few select wavelengths are necessary to build accurate, parsimonious models. Multispectral imaging not only addresses the high instrumental costs but the lengthy computational times as well.

The study by Delwiche *et al.* (2011) also concluded that considering the spectral information from the entire kernel yielded better results than only using a portion of the grain, such as the endosperm or the germ tip separately. This finding also supports the notion that the bulk scanning ability of hyperspectral imaging offers superior analysis compared to that of conventional spectroscopy. The model accuracies ranged from 90 to 96%, but an average accuracy of approximately 95% was achieved. It was also discovered that the spectral absorption near 1200 nm was useful in spectral recognition of *Fusarium* damage as this wavelength is linked to ergosterol; a primary constituent in fungi cell membranes.

In a recent study by Delwiche *et al.* (2019), percentages of *Fusarium*-damage in hard wheat was estimated. LDA was the classifier used to develop models. The reflectance values of the interior pixels of each grain at four wavelengths (1100, 1197, 1308, and 1394 nm) were used as input. For pre-processing, only mean-centring was applied except in one case where SNV was used. The results displayed percentages of correct classification of cross-validation higher than 95% for sound and *Fusarium*-damaged kernels. An independent validation set of external samples was used to establish the standard error of prediction (SEP) which ranged between 4.9 and 6.6%, indicating that the percentages of *Fusarium*-damaged kernels determined were reasonably accurate. This study proves that only using as little as four wavelengths (1100, 1197, 1308, and 1394 nm) is sufficient for differentiating sound and *Fusarium*-damaged kernels.

Shahin and Symons (2012) performed a study to detect *Fusarium* damage in Canadian wheat samples using visible and near-infrared hyperspectral imaging. An incredible sample size of 5,221 kernels was imaged using the wavelength range of 400-950 nm. All the kernels were individually inspected and assessed by trained grain inspectors to sort the kernels into their respective categories. The samples ranged from slightly-damaged to severely-damaged to sound. PLS-DA models were developed to discriminate between sound and fungal-damaged kernels, based on their mean spectra. Models were developed using all 181 wavebands, 13 select bands, five bands, four bands and three bands. Overall an average accuracy of 90% was achieved. It was concluded that a set of only four wavebands (494, 578, 639, 678 nm) was sufficient to achieve similar accuracies to those with the entire range of 450–950 nm.

Hyperspectral imaging offers more than just spectral information. Spatial and textural information can also be extracted to improve models further. In a recent study by Ropelewska and Zapotoczny (2018), classification models were developed to distinguish between *Fusarium*-infected and sound wheat kernels, based on textural parameters. Images were acquired in the spectral range of 400-1100 nm, and the kernels were imaged on both the ventral and dorsal side. The spectral range of 550-850 nm was selected for further analysis as the pixels in the images in this region were the most clearly visible. The classifiers used were Bayes (Bayes Net), Function (LDA), Lazy (K Star), Rules (PART) and Decision Tree (LMT). The analyses were executed for every selected wavelength, namely 550, 710 and 850 nm. For kernels placed on the ventral side, classification accuracies of 78-100% were achieved. For the dorsal side classification accuracies of 76-98% were achieved. It was concluded that the position of the wheat kernels influenced the classification accuracy. A higher accuracy was achieved for kernels placed on the ventral than the dorsal side. Alternatively, Singh *et al.* (2009) found that kernels imaged crease-up was helpful for sprouted and midge-damage detection. This is interesting since the germ that displays signs of sprouting is more visible on the dorsal side. Since placing the kernels in a certain way for grading is impractical, accurate classifications should be achieved irrespective of orientation.

A concern with hyperspectral imaging is that for object-wise analyses, which is required for whole wheat grading, the placement of kernels plays a role. Often it is necessary to place the grains neatly, without touching, as kernel clusters could be identified as one object, affecting the results. Barbedo *et al.* (2015) proved that this hurdle need not be of concern. The authors developed an algorithm to detect kernels infected with *Fusarium* head blight (FHB) that was robust to factors such as shape, orientation, shadows and kernels touching one another. The algorithm returns a *Fusarium* index (FI) that is indicative of the likelihood of the kernel being infected. The higher the index, the more likely the grain has been infected by FHB. The algorithm was based on the proportion of pixels that delivered high reflections at the waveband 1411 nm. The 1411 nm band was selected as it provided the most significant separation, confirmed using PCA. In the research by Delwiche *et al.* (2011), it was noted that chitin, a component found in the cell walls of fungi, displayed absorption bands near 1480 nm, which is in the region of 1411 nm. The algorithm achieved a

classification accuracy of approximately 91% and shows promising capabilities of estimating DON concentrations in wheat kernels.

Vermeulen *et al.* (2012) created PLS-DA and SVM classification models to detect and quantify ergot bodies in cereals. While the models performed well, it was observed that some false positives arose due to wheat germ pixels being detected as ergot pixels. This aspect could most likely be overcome using an object-wise approach instead of pixel-wise analysis, where the mean spectrum for each kernel is calculated and used as input. False negatives also resulted from the non-distinction of two ergot bodies that were to close together. This problem could be minimised or eliminated by ensuring that no kernels are touching when the hyperspectral images are captured. However algorithms can be created to be more robust towards kernel clustering as seen in the study by Barbedo *et al.* (2015).

Many studies have been conducted on sprout-damage as sprouting has a detrimental effect on baking quality. Xing *et al.* (2010) applied *HSI* for the detection of moderate sprout damage and severe sprout damage in Canada Western Red Spring wheat. Both the visible and NIR range was used. The authors noticed that the average spectra of sprouted kernels had higher reflectance responses compared to that of sound kernels in the spectral bands above 720 nm. Morphological features such as smoothness and particle size were especially useful for the distinction between severely sprouted and sound kernels. A classification accuracy of 94% for the moderately sprouted kernels and 98% for the severely sprouted kernels was achieved.

A study by Barbedo *et al.* (2018) also investigated sprout damage. The samples were first subjected to HSI and then ground to determine the falling number. The images were captured from 528–1785 where the regions of 844–1140 nm and 1386–1700 nm displayed clear spectral differences between the two categories. The wavelengths 918 nm and 1411 nm were selected from the spectral ranges and combined into a sprouting index equation. The reflectance values at those select wavelengths were used to calculate the sprouting index. If the sprouting index was small, the kernel was more likely to have sprouted. A global threshold of 0.30 was used, as well as cultivar specific thresholds, where the cultivar specific threshold provided a higher accuracy. Imaging with the germ side up also produced superior results with accuracies of 100%. The authors concluded that the use of wavelength selection is useful as most of the information needed is usually contained in just a few specific bands.

For the classification of heat damage, there have been no studies using hyperspectral imaging. However, a study was done using conventional NIR spectroscopy by Wang *et al.* (2001). Reflectance spectra were measured from 400 to 1700 nm. Two separate partial least squares (PLS) models were developed using the wavelength region of 750–1700 nm and 400–750 nm. The wavelength range of 750–1700 nm provided the highest classification accuracy of 100%. The visible region, 400–750 nm, gave the lowest classification accuracy, emphasising that visual inspection in the visible range is not a reliable method to grade wheat. Singh *et al.* (2010) did a study to compare the efficacy of using the NIR and visible region for the analysis of insect damage in wheat. Three separate analyses were performed where the NIR and visible range were

investigated individually and thereafter, the two regions were combined. Overall, the combination of NIR and visible feature provided the highest classification accuracy where 96.4% of sound kernels and 91.0–100.0% of insect-damaged were correctly predicted. The combination of NIR and colour image features, therefore, has the potential of accurately detecting insect damage in wheat.

Ravikanth *et al.* (2015) created models to differentiate between seven foreign material types. Since pre-processing is the most critical step in model development, one of the goals in the study was to investigate different pre-processing techniques to determine the optimal pre-processing and classifier combination. Raw NIR reflectance spectra of the contaminants and sound wheat were collected in the range of 1000–1600 nm. These spectra were processed using the 1st derivative, 2nd derivative, Savitzky-Golay smoothing and differentiation, MSC and SNV. The raw and pre-processed data were classified using SVM, Naïve Bayes, and KNN. Overall, the highest accuracy was achieved with SNV followed by KNN.

Table 2.1 Applications of hyperspectral imaging and chemometrics for quality and safety evaluation of cereals.

Application	Key Wavelengths	Classifier	R^2 or Classification Accuracy	Reference
Wheat				
Kernel Hardness	910; 990; 1030	PLS-FDA	94%	(Gorretta *et al.*, 2006)
Kernel Hardness	1180–1220; 1320–1400; 1460–1490	PLSR; PCR	0.82; 0.75	(Mahesh *et al.*, 2015)
Kernel Hardness	-	PLS; ANN	0.80; 0.90	(Erkinbaev *et al.*, 2019)
Protein Content	960–1030; 1180– 1200; 1460–1500; 1670–1700	PLSR; PCR	0.68; 0.62	(Mahesh *et al.*, 2015)
Protein Content	1918	PLS	0.80	(Caporaso *et al.*, 2018)
Fusarium head blight	502, 678, 1198, 1496	LDA	95%	(Delwiche *et al.*, 2011)
Fusarium-damaged kernels	494, 578, 639, 678	PLS-DA	90%	(Shahin & Symons, 2012)
Toxigenic fungi	-	LDA	91.7%; 88.3%	(Shen *et al.*, 2019)
Fusarium graminearum	550, 710, 850	LDA, K Star, PART, LMT	76–100%	(Ropelewska & Zapotoczny, 2018)
Fusarium head blight	1000, 1197, 1394, 1308	PLS-DA	>92%	(Delwiche *et al.*, 2019)
		LDA	0.772; 0.811	
Fusarium head blight	647, 1411	*Fusarium* index	91%	(Barbedo *et al.*, 2015)
Ergot	-	PLS-DA; SVM	0.99	(Vermeulen *et al.*, 2012)
Sprouting	1920–1940	PLS-DA	0.67	(McGoverin *et al.*, 2011)
Sprouting	1102; 1132; 1305	LDA	94%	(Singh *et al.*, 2009)
		QDA	94%	

		Mahalanobis DA	95%	
Sprouting	918; 1411	-	-	(Barbedo *et al.*, 2018)
Sprouting	-	PCA	94%; 98%	(Xing *et al.*, 2010)
		LDA	79%	
Insect	870	QDA	86%	(Singh *et al.*, 2010)
		Mahalanobis DA	74%	
Variety	1130; 1310	LDA; QDA; ANN	>94; >86; >80%	(Mahesh *et al.*, 2008)
Foreign Matter	-	KNN	100.0%	(Ravikanth *et al.*, 2016)
Maize				
Kernel Hardness	1400	PLS-DA	0.85; 0.88; 0.76	(Williams *et al.*, 2009)
Kernel Hardness	1220; 1405; 1690; 1870	PLS-DA	0.85	(Manley *et al.*, 2009)
Moisture Content	910; 960-970	PLS; PCR	0.87	(Cogdill *et al.*, 2004)
Oil Content	910; 960-970	PLS; PCR	0.85	(Cogdill *et al.*, 2004)
Variety	438; 487; 511; 516; 700; 897	PCA; SVM	89; 90%	(Wang *et al.*, 2015)
Fusarium	1405; 1660–1668; 1900; 2136	PLSR	0.71; 0.63	(Williams *et al.*, 2012)
Rice				
Variety	419; 452; 593; 613; 743; 784; 966	PCA; BPNN	89; 94%	(Wang *et al.*, 2014)
Barely				
Sprouting	1430	MNF	97%	(Arngren *et al.*, 2011)
Sprouting	1920–1940	PLS-DA	0.68	(McGoverin *et al.*, 2011)
Sorghum				
Sprouting	1920–1940	PLS-DA	0.53	(McGoverin *et al.*, 2011)

(PLS) partial least squares; (FDA) factorial discriminant analysis; (PLSR) partial least squares regression; (PCR) principal component regression; (ANN) artificial neural networks; (LDA) linear discriminant analysis; (PLS-DA) partial least squares discriminant analysis; (SVM) support vector machines; (QDA) quadratic discriminant analysis; (PCA) principal component analysis; (KNN) k-nearest neighbours; (BPNN) back-propagation neural networks; (MNF) minimum noise fraction.

Conclusion

Wheat is a valuable agricultural commodity and a system to improve the visual grading process thereof is addressed on a continual basis. Currently, human visual inspection is still being used in South Africa, and other major wheat-producing countries, to assess the quality of wheat. From this review, it is clear that human visual inspection is an inadequate method to grade wheat. Human visual inspection has proven to be subjective, tedious and time-consuming. Even spectral imaging performed soley in the visible range provided inferior results to that in the NIR range. A method to rapidly and accurately evaluate the quality of wheat is highly sought-after, and NIR hyperspectral imaging is a viable alternative solution. Research on the use of hyperspectral imaging to evaluate grain characteristics has been extensively studied and delivered promising results. The studies mainly focused on the detection of fungal-infected wheat, sprout-damage, insect damage and the presence foreign matter. Areas which have not been researched are heat damage, frost damage and immature wheat. Heat- and frost-damaged wheat are of the most challenging defects to visually identify as they have minimal visual distinctive properties. Therefore, the use of a spectroscopic technique such as hyperspectral imaging is ideal to reveal characteristics that cannot be seen by the naked eye. With powerful programming languages being easily accessible and many statistical learning algorithms available, there is no limit to what can be done with the vast amount of data that hyperspectral imaging produces. Although hyperspectral instruments are currently more suited to laboratory work, they show great promise of being incorporated on-line or in-line in the industry to provide real-time classification and sorting of wheat.

References

Amigo, J.M., Martí, I. & Gowen, A. (2013). Hyperspectral imaging and chemometrics: a perfect combination for the analysis of food structure, composition and quality. In: *Data Handling in Science and Technology*. Pp. 343-370. Amsterdam, The Netherlands: Elsevier Science.

Arngren, M., Hansen, P.W., Eriksen, B., Larsen, J. & Larsen, R. (2011). Analysis of pregerminated barley using hyperspectral image analysis. *Journal of agricultural and food chemistry*, **59**, 11385-11394.

Barbedo, J.G.A., Guarienti, E.M. & Tibola, C.S. (2018). Detection of sprout damage in wheat kernels using NIR hyperspectral imaging. *Biosystems Engineering*, **175**, 124-132.

Barbedo, J.G.A., Tibola, C.S. & Fernandes, J.M.C. (2015). Detecting *Fusarium* head blight in wheat kernels using hyperspectral imaging. *Biosystems Engineering*, **131**, 65-76.

Barlow, K.M., Christy, B.P., O'Leary, G.J., Riffkin, P.A. & Nuttall, J.G. (2015). Simulating the impact of extreme heat and frost events on wheat crop production: A review. *Field Crops Research*, **171**, 109-119.

Barnes, R.J., Dhanoa, M.S. & Lister, S.J. (1989). Standard Normal Variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, **43**, 772-777.

Bertrand, D., Robert, P. & Loisel, W. (1985). Identification of some wheat varieties by near infrared reflectance spectroscopy. *Journal of the Science of Food and Agriculture*, **36**, 1120-1124.

Boldrini, B., Kessler, W., Rebner, K. & Kessler, R. (2012). Hyperspectral imaging: a review of best practice, performance and pitfalls for inline and online applications. *Journal of Near Infrared Spectroscopy*, **20**, 438.

Boyle, W.S. (2010). Nobel Lecture: CCD---An extension of man's view. *Reviews of Modern Physics*, **82**, 2305-2306.

Breiman, L. (2001). Random Forests. *Machine Learning*, **45**, 5-32.

Brereton, R.G. (2015). Pattern recognition in chemometrics. *Chemometrics and Intelligent Laboratory Systems*, **149**, 90-96.

Burger, J. & Gowen, A. (2011). Data handling in hyperspectral image analysis. *Chemometrics and Intelligent Laboratory Systems*, **108**, 13-22.

Caporaso, N., Whitworth, M.B. & Fisk, I.D. (2018a). Near-Infrared spectroscopy and hyperspectral imaging for non-destructive quality assessment of cereal grains. *Applied Spectroscopy Reviews*, **53**, 667-687.

Caporaso, N., Whitworth, M.B. & Fisk, I.D. (2018b). Protein content prediction in single wheat kernels using hyperspectral imaging. *Food Chemistry*, **240**, 32-42.

Cogdill, R.P., Hurburgh, C.R., Rippke, G.R., Bajic, S., Jones, R.W., McClelland, J.F., Jensen, T.C. & Liu, J. (2004). Single-kernel maize analysis by near-infrared hyperspectral imaging. *Transactions of the ASAE*, **47**, 311-320.

Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, **20**, 273-297.

de Wet, F. & Liebenberg, I. (2018). Food security, wheat production and policy in South Africa: Reflections on food sustainability and challenges for a market economy. *The Journal for Transdisciplinary Research in Southern Africa*, **14**, 1-11.

Delwiche, S.R., Kim, M.S. & Dong, Y. (2011). *Fusarium* damage assessment in wheat kernels by Vis/NIR hyperspectral imaging. *Sensing and Instrumentation for Food Quality and Safety*, **5**, 63-71.

Delwiche, S.R., Rodriguez, I.T., Rausch, S.R. & Graybosch, R.A. (2019). Estimating percentages of fusarium-damaged kernels in hard wheat by near-infrared hyperspectral imaging. *Journal of Cereal Science*, **87**, 18-24.

Department of Agriculture (2016). Regulations relating to the grading, packing and marking of bread wheat intended for sale in the Republic of South Africa. In: *Agricultural Product Standards Act (Act No. 119 of 1990)*.

Dexter, J.E., Martin, D.G., Preston, K.R., Tipples, K.H. & MacGregor, A.W. (1985). The effect of frost damage on the milling and baking quality of red spring wheat. *Cereal Chemistry*, **62**, 75-80.

Dowell, F.E., Maghirang, E.B., Xie, F., Lookhart, G.L., Pierce, R.O., Seabourn, B.W., Bean, S.R., Wilson, J.D. & Chung, O.K. (2006). Predicting wheat quality characteristics and functionality using near-infrared spectroscopy. *Cereal Chemistry*, **83**, 529-536.

Dowell, F.E., Ram, M.S. & Seitz, L.M. (1999). Predicting scab, vomitoxin, and ergosterol in single wheat kernels using near-infrared spectroscopy. *Cereal Chemistry*, **76**, 573-576.

ElMasry, G., Cubero, S., Moltó, E. & Blasco, J. (2012a). In-line sorting of irregular potatoes by using automated computer-based machine vision system. *Journal of Food Engineering*, **112**, 60-68.

Elmasry, G., Kamruzzaman, M., Sun, D.W. & Allen, P. (2012b). Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: a review. *Critical Reviews in Food Science and Nutrition*, **52**, 999-1023.

Elmasry, G. & Sun, D.W. (2010). Principles of hyperspectral imaging technology. In: *Hyperspectral Imaging for Food Quality Analysis and Control*. Pp. 3-44. London, UK: Elsevier.

ElMasry, G.M. & Nakauchi, S. (2016). Image analysis operations applied to hyperspectral images for non-invasive sensing of food quality – A comprehensive review. *Biosystems Engineering*, **142**, 53-82.

Erkinbaev, C., Derksen, K. & Paliwal, J. (2019). Single kernel wheat hardness estimation using near infrared hyperspectral imaging. *Infrared Physics & Technology*, **98**, 250-255.

Esbensen, K. & Geladi, P. (1989). Strategy of multivariate image analysis (MIA). *Chemometrics and Intelligent Laboratory Systems*, **7**, 67-86.

Fardet, A. (2010). New hypotheses for the health-protective mechanisms of whole-grain cereals: what is beyond fibre? *Nutrition Research Reviews*, **23**, 65-134.

Femenias, A., Gatius, F., Ramos, A.J., Sanchis, V. & Marín, S. (2020). Use of hyperspectral imaging as a tool for *Fusarium* and deoxynivalenol risk management in cereals: A review. *Food Control*, **108**, 106819.

Feng, Y.Z. & Sun, D.W. (2012). Application of hyperspectral imaging in food safety inspection and control: a review. *Critical Reviews in Food Science and Nutrition*, **52**, 1039-1058.

Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems *Annals of Eugenics*, **7**, 179-188.

Flannigan, B. (1978). Primary contamination of barley and wheat grain storage fungi. *Transactions of the British Mycological Society*, **71**, 37-42.

Fox, G. & Manley, M. (2014). Applications of single kernel conventional and hyperspectral imaging near infrared spectroscopy in cereals. *Journal of the Science of Food and Agriculture*, **94**, 174-179.

Frølich, W. & Åman, P. (2010). Whole grain for whom and why? *Food & Nutrition Research*, **54**.

Fu, X. & Ying, Y. (2016). Food safety evaluation based on near infrared spectroscopy and imaging: A review. *Critical Reviews in Food Science and Nutrition*, **56**, 1913-1924.

Goetz, A.F.H. (2009). Three decades of hyperspectral remote sensing of the Earth: A personal view. *Remote Sensing of Environment*, **113**, S5-S16.

Goetz, A.F.H., Vane, G., Solomon, J.E. & Rock, B.N. (1985). Imaging spectrometry for earth remote sensing. *Science*, **228**, 1147.

Gorretta, N., Roger, J.-M., Aubert, M., bellon maurel, V., Campan, F. & Roumet, P. (2006). Determining vitreouness of durum wheat kernels using near infrared hyperspectral imaging. *Journal of Near Infrared Spectroscopy*, **14**, 231-239.

Gowen, A., Odonnell, C., Cullen, P., Downey, G. & Frias, J. (2007). Hyperspectral imaging – an emerging process analytical tool for food quality and safety control. *Trends in Food Science & Technology*, **18**, 590-598.

Hernández-Espinosa, N., Mondal, S., Autrique, E., Gonzalez-Santoyo, H., Crossa, J., Huerta-Espino, J., Singh, R.P. & Guzmán, C. (2018). Milling, processing and end-use quality traits of CIMMYT spring bread wheat germplasm under drought and heat stress. *Field Crops Research*, **215**, 104-112.

James, G., Witten, D., Hastie, T. & Tibshirani, R. (2014). *An Introduction to Statistical Learning: with Applications in R.* Pp. 430. Springer Publishing Company, Incorporated.

Jayas, D.S. & Ghosh, P.K. (2006). Preserving quality during grain drying and techniques for measuring grain quality. In: *Proceedings of the 9th International Working Conference on Stored Product Protection*. Pp. 969-981. Campinas, Brazil: Brazilian Post-harvest Association.

Karunakaran, C., Muir, W.E., Jayas, D.S., White, N.D.G. & Abramson, D. (2001). Safe storage time of high moisture wheat. *Journal of Stored Products Research*, **37**, 303-312.

Lawton, J.W. (2000). Non-food uses of cereals. In: *Handbook of Cereal Science and Technology* (edited by K. Kulp; & J.G. Ponte). Pp. 725-738. New York: Marcel Dekker, Inc.

Lorente, D., Aleixos, N., Gómez-Sanchis, J., Cubero, S., García-Navarrete, O.L. & Blasco, J. (2011). Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment. *Food and Bioprocess Technology*, **5**, 1121-1142.

Mahesh, S., Jayas, D.S., Paliwal, J. & White, N.D.G. (2015). Comparison of Partial Least Squares Regression (PLSR) and Principal Components Regression (PCR) methods for protein and hardness predictions using the near-infrared (NIR) hyperspectral images of bulk samples of Canadian wheat. *Food and Bioprocess Technology*, **8**, 31-40.

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. (2014). Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. *Chemical Society Reviews*, **43**, 8200-8214.

Manley, M., Williams, P., Nilsson, D. & Geladi, P. (2009). Near infrared hyperspectral imaging for the evaluation of endosperm texture in whole yellow maize (Zea maize L.) kernels. *Journal of Agricultural and Food Chemistry*, **57**, 8761-8769.

Massart, D., Vandeginste, B., Deming, S., Michotte, Y. & Kaufman, L. (1988). *Chemometrics: A textbook.* Amsterdam: Elsevier Science Publishers.

McCullagh, P. & Nelder, J.A. (1989). *Generalized linear models.* London: Chapman & Hall/CRC.

McGoverin, C.M., Engelbrecht, P., Geladi, P. & Manley, M. (2011). Characterisation of non-viable whole barley, wheat and sorghum grains using near-infrared hyperspectral data and chemometrics. *Analytical and Bioanalytical Chemistry*, **401**, 2283-2289.

Mobaraki, N. & Amigo, J.M. (2018). HYPER-Tools. A graphical user-friendly interface for hyperspectral image analysis. *Chemometrics and Intelligent Laboratory Systems*, **172**, 174-187.

Moghaddam, T.M., Razavi, S.M.A. & Taghizadeh, M. (2013). Applications of hyperspectral imaging in grains and nuts quality and safety assessment: a review. *Journal of Food Measurement and Characterization*, **7**, 129-140.

Mutihac, L. & Mutihac, R. (2008). Mining in chemometrics. *Analytica Chimica Acta*, **612**, 1-18.

Nhemachena, C.R. & Kirsten, J. (2017). A historical assessment of sources and uses of wheat varietal innovations in South Africa. *South African Journal of Science*, **113**, 1-8.

Norris, K.H. (1983). Extracting information from spectrophotometric curves. Predicting chemical composition from visible and near-infrared spectra. In: *Food Research and Data Analysis* (edited by H. Martens; & H. Russwurm). Pp. 95-114. London: Applied Science Publishers.

Nuttall, J.G., O'Leary, G.J., Panozzo, J.F., Walker, C.K., Barlow, K.M. & Fitzgerald, G.J. (2017). Models of grain quality in wheat—A review. *Field Crops Research*, **202**, 136-145.

Osborne, B.G. (2006). NIR spectroscopy in food analysis. In: *Encyclopedia of Analytical Chemistry* (edited by M. R.A.). Pp. 1–14. Chichester, UK: John Wiley and Sons Ltd.

Pasquini, C. (2018). Near infrared spectroscopy: A mature analytical technique with new perspectives - A review. *Analytica Chimica Acta*, **1026**, 8-36.

Pearson, K. (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **2**, 559-572.

Petrovska-avramenko, N., Karklina, D. & Gedrovica, I. (2016). Investigation of immature wheat grain chemical composition. *Research for Rural Development*, **1**, 102-105.

Pettersson, H. & Åberg, L. (2003). Near infrared spectroscopy for determination of mycotoxins in cereals. *Food Control*, **14**, 229-232.

Porep, J.U., Kammerer, D.R. & Carle, R. (2015). On-line application of near infrared (NIR) spectroscopy in food production. *Trends in Food Science & Technology*, **46**, 211-230.

Preston, K.R., Kilborn, R.H., Morgan, B.C. & Babb, J.C. (1991). Effects of frost and immaturity on the quality of a canadian hard red spring wheat. *Cereal Chemistry*, **68**, 133-138.

Quinlan, J.R. (1986). Induction of decision trees. *Machine Learning*, **1**, 81-106.

Ramirez, A., Perez, G.T., Ribotta, P.D. & Leon, A.E. (2003). The occurrence of friabilins in triticale and their relationship with grain hardness and baking quality. *Journal of Agricultural and Food Chemistry*, **51**, 7176-7181.

Ravikanth, L., Singh, C.B., Jayas, D.S. & White, N.D.G. (2015). Classification of contaminants from wheat using near-infrared hyperspectral imaging. *Biosystems Engineering*, **135**, 73-86.

Ravikanth, L., Singh, C.B., Jayas, D.S. & White, N.D.G. (2016). Performance evaluation of a model for the classification of contaminants from wheat using near-infrared hyperspectral imaging. *Biosystems Engineering*, **147**, 248-258.

Razmjooy, N., Mousavi, B.S. & Soleymani, F. (2012). A real-time mathematical computer method for potato inspection using machine vision. *Computers & Mathematics with Applications*, **63**, 268-279.

Ridgway, C. & Chambers, J. (1996). Detection of external and internal insect infestation in wheat by near-infrared reflectance spectroscopy. *Journal of the Science of Food and Agriculture*, **71**, 251-264.

Rinnan, Å., Berg, F.v.d. & Engelsen, S.B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry*, **28**, 1201-1222.

Roberts, J., Power, A., Chapman, J., Chandra, S. & Cozzolino, D. (2018). A short update on the advantages, applications and limitations of hyperspectral and chemical imaging in food authentication. *Applied Sciences*, **8**, 505.

Ropelewska, E. & Zapotoczny, P. (2018). Classification of Fusarium-infected and healthy wheat kernels based on features from hyperspectral images and flatbed scanner images: a comparative analysis. *European Food Research and Technology*, **244**, 1453-1462.

Sadhasivam, S., Britzi, M., Zakin, V., Kostyukovsky, M., Trostanetsky, A., Quinn, E. & Sionov, E. (2017). Rapid detection and identification of mycotoxigenic fungi and mycotoxins in stored wheat grain. *Toxins*, **9**, 302.

Savitzky, A. & Golay, M.J.E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, **36**, 1627-1639.

Sebestyen, G. (1962). Pattern recognition by an adaptive process of sample set construction. *IRE Transactions on Information Theory*, **8**, 82-91.

Sendin, K., Manley, M. & Williams, P.J. (2018a). Classification of white maize defects with multispectral imaging. *Food Chemistry*, **243**, 311-318.

Sendin, K., Williams, P.J. & Manley, M. (2018b). Near infrared hyperspectral imaging in quality and safety evaluation of cereals. *Critical Reviews in Food Science and Nutrition*, **58**, 575-590.

Senthilkumar, T., Jayas, D.S., White, N.D.G., Fields, P.G. & Gräfenhan, T. (2017). Detection of ochratoxin A contamination in stored wheat using near-infrared hyperspectral imaging. *Infrared Physics & Technology*, **81**, 228-235.

Shahin, M.A. & Symons, S.J. (2008). Detection of hard vitreous and starchy kernels in amber durum wheat samples using hyperspectral imaging (GRL Number M306). *NIR news*, **19**, 16-18.

Shahin, M.A. & Symons, S.J. (2012). Detection of *Fusarium* damage in Canadian wheat using visible/near-infrared hyperspectral imaging. *Journal of Food Measurement & Characterization*, **6**, 3-11.

Shen, F., Zhao, T., Jiang, X., Liu, X., Fang, Y., Liu, Q., Hu, Q. & Liu, X. (2019). On-line detection of toxigenic fungal infection in wheat by visible/near infrared spectroscopy. *Lwt - Food Science and Technology*, **109**, 216-224.

Shewry, P.R. (2009). Wheat. *Journal of Experimental Botany*, **60**, 1537-1553.

Siche, R., Vejarano, R., Aredo, V., Velasquez, L., Saldaña, E. & Quevedo, R. (2015). Evaluation of food quality and safety with hyperspectral imaging (HSI). *Food Engineering Reviews*, **8**, 306-322.

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.

Singh, C.B., Jayas, D.S., Paliwal, J. & White, N.D.G. (2010). Identification of insect-damaged wheat kernels using short-wave near-infrared hyperspectral and digital colour imaging. *Computers and Electronics in Agriculture*, **73**, 118-125.

Singh, C.B., Jayas, D.S., Paliwal, J. & White, N.D.G. (2012). Fungal damage detection in wheat using short-wave near-infrared hyperspectral and digital colour imaging. *International Journal of Food Properties*, **15**, 11-24.

Smith, G.E. (2010). Nobel Lecture: The invention and early history of the CCD. *Reviews of Modern Physics*, **82**, 2307-2312.

Sofu, M.M., Er, O., Kayacan, M.C. & Cetişli, B. (2016). Design of an automatic apple sorting system using machine vision. *Computers and Electronics in Agriculture*, **127**, 395-405.

Stigler, S.M. (1981). Gauss and the invention of least squares. *The Annals of Statistics*, **9**, 465-474.

Tipples, K.H. (1980). Effect of immaturity on the milling and baking quality of red spring wheat. *Canadian Journal of Plant Science*, **60**, 357-369.

van der Merwe, J.D. & Cloete, P.C. (2017). Financial impact of wheat quality standards on South African wheat producers: A dynamic linear programming (DLP) approach. *Development Southern Africa*, **35**, 53-69.

Vermeulen, P., Pierna, J.A., Egmond, H.P., Dardenne, P. & Baeten, V. (2012). Online detection and quantification of ergot bodies in cereals using near infrared hyperspectral imaging. *Food Additives and Contaminants*, **29**, 232-240.

Vermeulen, P., Suman, M., Fernández Pierna, J.A. & Baeten, V. (2018). Discrimination between durum and common wheat kernels using near infrared hyperspectral imaging. *Journal of Cereal Science*, **84**, 74-82.

Vidal, M. & Amigo, J.M. (2012). Pre-processing of hyperspectral images. Essential steps before image analysis. *Chemometrics and Intelligent Laboratory Systems*, **117**, 138-148.

Wang, D., Dowell, F.E. & Chung, D.S. (2001). Assessment of heat-damaged wheat kernels using near-infrared spectroscopy. *Cereal Chemistry*, **78**, 625-628.

Wang, L., Liu, D., Pu, H., Sun, D.-W., Gao, W.H. & Xiong, Z.J. (2015a). Use of hyperspectral imaging to discriminate the variety and quality of rice. *Food and Bioprocess Technology*, **8**, 515-523.

Wang, L., Sun, D.-W., Pu, H. & Zhu, Z. (2015b). Application of hyperspectral imaging to discriminate the variety of maize seeds. *Food Analytical Methods*, **9**, 225–234.

Williams, P., Geladi, P., Fox, G. & Manley, M. (2009). Maize kernel hardness classification by near infrared (NIR) hyperspectral imaging and multivariate data analysis. *Analytica Chimica Acta*, **653**, 121-130.

Williams, P.C. (1975). Application of near infrared reflectance spectroscopy to analysis of cereal grains and oilseeds. *Cereal Chemistry*, **52**, 561-567.

Williams, P.J., Geladi, P., Britz, T.J. & Manley, M. (2012). Investigation of fungal development in maize kernels using NIR hyperspectral imaging and multivariate data analysis. *Journal of Cereal Science*, **55**, 272-278.

Wold, S., Esbensen, K. & Geladi, P. (1987a). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, **2**, 37-52.

Wold, S., Geladi, P., Esbensen, K. & Öhman, J. (1987b). Multi-way principal components-and PLS-analysis. *Journal of Chemometrics*, **1**, 41-56.

Wold, S. & Sjöström, M. (1977). SIMCA: A method for analyzing chemical data in terms of similarity and analogy. In: *Chemometrics: Theory and Application*. Pp. 243-282. AMERICAN CHEMICAL SOCIETY.

Xing, J., Symons, S., Shahin, M. & Hatcher, D. (2010). Detection of sprout damage in Canada Western Red Spring wheat with multiple wavebands using visible/near-infrared hyperspectral imaging. *Biosystems Engineering*, **106**, 188-194.

Zhang, X., Liu, F., He, Y. & Li, X. (2012). Application of hyperspectral imaging and chemometric calibrations for variety discrimination of maize seeds. *Sensors*, **12**, 17234-17246.

Zhao, H., Guo, B., Wei, Y. & Zhang, B. (2013). Near infrared reflectance spectroscopy for determination of the geographical origin of wheat. *Food Chemistry*, **138**, 1902-1907.

CHAPTER 3
MATERIALS AND METHODS

Grain Samples

For this study, only "Bread Wheat" was considered as it is the most produced wheat in South Africa. Bulk samples of bread wheat (*Triticum aestivum*) were obtained from VKB Agriculture (VKB, Reitz, South Africa) and the South African Grain Laboratory (SAGL, Pretoria, South Africa). The samples were visually analysed by skilled industry graders and the kernels were separated into different categories, based on their health status. The categories selected were in accordance with the South African grading regulations (Department of Agriculture, 2016). Only a subset of defective categories was chosen due to sample availability. At the time of this study, only a limited number of defective categories manifested in the seasons' harvest. Five categories in total were selected: sound, heat-damaged, *Fusarium*-damaged, sprout-damaged, and immature wheat kernels. A visual representation of sound and defective wheat kernels is displayed in Figure 3.1. Approximately 1100 wheat kernels were imaged in total.

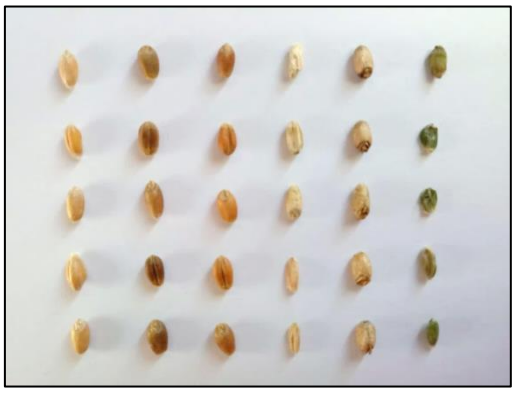


Figure 3.1 Digital image of the different categories, where the first column is sound wheat, followed by highly heat-damaged, moderately heat-damaged, *Fusarium*-damaged, sprout-damaged and immature wheat kernels.

Near-Infrared Hyperspectral Imaging System

All hyperspectral images were acquired with a HySpex SWIR-384 (short wave infrared) pushbroom imaging system (Norsk Elektro Optikk, Norway) (Figure 3.2) using Breeze® (Prediktera) acquisition software version 2018.17.1. The cameras were mounted on a laboratory rack with a translation stage, fitted with a 30 cm focal length lens giving a Field of View (FOV) of 9.447 cm. The system consisted of an imaging spectrograph coupled to an MCT (mercury cadmium telluride) sensor with a spectral range of 950 to 2500 nm and a spectral resolution of 5.45 nm. This resulted in a total of 288 spectral bands. The spatial resolution of the 30 cm lens resulted in a pixel size of 0.247 mm. Each image consisted of 384 pixels in width (x) and varied in the number of lines (y). The samples were illuminated with two 150 W halogen lamps (Ushio lighting Inc., Japan) with

the capacity to emit light in the 400 – 2500 nm wavelength range. A 50% grey reflectance standard was used as the white reference and a dark reference was recorded with the shutter closed. Radiometric calibration was performed in the Breeze software package.

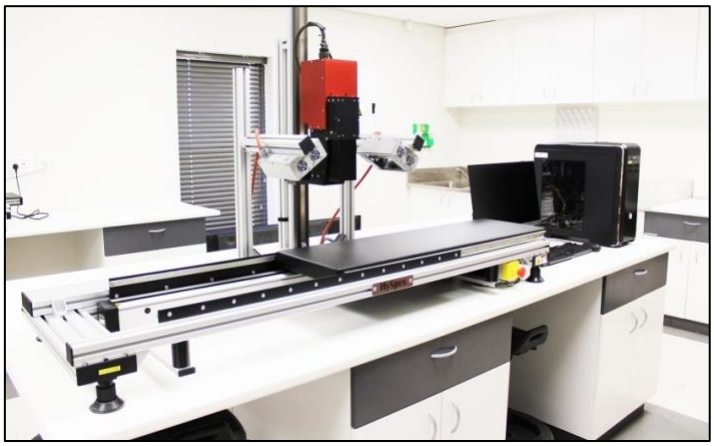


Figure 3.2 Photograph of the pushbroom NIR HySpex SWIR hyperspectral imaging system.

Image Acquisition

Each category (sound, heat-damaged, *Fusarium*-damaged, sprout-damaged, and immature kernels) varied in the number of kernels, depending on the availability. Approximately 30 to 60 kernels were imaged at a time until the entire sample had been captured. Since all the data of the kernels belonging to the same class were concatenated together, the number of kernels per image was irrelevant as long as the entire batch of samples were captured. The kernels were imaged on a smooth, black background to provide a distinct contrast between the foreground and background. A visual representation of using a dark versus a light background is displayed in Figure 3.3. Acquiring the data in this manner was necessary to increase the effectiveness of Otsu's thresholding method (Otsu, 1979), the technique used to perform image segmentation. Using a dark background minimised the visibility of shadows cast by the kernels while a smooth background was necessary to reduce light scattering effects. Rough-surfaced materials contain an increased surface area that increases the amount of light scattering, which in turn increases the amount of noise in the data.

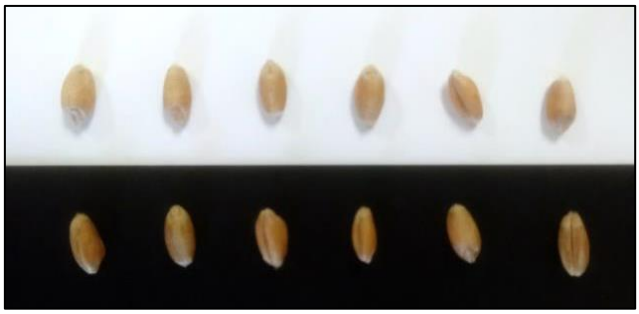


Figure 3.3 Digital image illustrating the difference between a light and dark background.

The kernels were randomly placed crease-up or crease-down to obtain data of both orientations. In industry, it would be impractical for the graders to orientate the kernels either crease-up or crease-down to perform the analysis, therefore, both orientations were acquired.

Image Correction

Image correction on the raw spectral images (*I*) was achieved with a radiometric calibration. For this correction, two reference standards are required: a dark current image (I_{dark}) and a white reference image (I_{white}). The white reference image is needed to set-up the maximum reflectance conditions. This is obtained by a white Teflon surface that is subjected to the same lighting conditions as the samples. A dark current image defines the absence of reflectance, which is acquired by turning the light source off and covering the lens with its opaque cap. The calibrated image (I_c) was then calculated using Eq. 3.1.

$$I_c = \frac{(I - I_{dark})}{(I_{white} - I_{dark})} \tag{3.1}$$

The corrected reflectance spectra were then converted to pseudo-absorbance by taking the logarithm of the reflectance values (Eq. 3.2). The absorbance images were exported as comma-separated values (CSV) files for further analysis.

$$A = log\frac{1}{R} \tag{3.2}$$

Hyperspectral Data Analyses

The hypercubes were analysed using Python v 3.6.5 (G. van Rossum, Python tutorial, Technical Report CS-R9526, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, May 1995). The raw hypercubes were explored to gain a better understanding of the data. Methods used to explore the data included spatial mapping, histograms, spectral graphs, and principal component analysis (PCA). Image segmentation was performed with Otsu's thresholding (Otsu, 1979). A mean spectrum was calculated for each kernel to perform object-wise analyses. The resultant spectra were pre-processed using different combinations of pre-processing techniques (Table 3.1) to correct for variations and interferences within the data. The cleaned and pre-processed data was used to build classification models (Table 3.2). The models were evaluated using cross-validation and further optimised to achieve a better performing model. Once the optimal model was identified, it was applied to external samples for final validation.

Image Segmentation

The background of the image was removed using Otsu's thresholding method (Otsu, 1979). It is an image thresholding algorithm where a greyscale image is reduced to a binary image by finding an optimal threshold based on the distribution of pixel values (Otsu, 1979). The method assumes that the image only consists of two classes, foreground and background, where the heterogeneity of the background is ignored. The method iterates through all the possible threshold values and calculates the foreground and background variances (a measure of spread) for the pixel levels each side of the threshold. The algorithm selects the threshold by minimizing intra-class variances of the two groups of pixels, which in turn maximises the inter-class variation. The goal is to find the optimal threshold value where the sum of foreground and background spreads is at its minimum.

After the application of the Otsu's thresholding algorithm, a binary image was created where the background was set to 0 while the foreground was set to 1. Therefore, the pixel value was replaced with either a 0 or 1, depending on whether it was background or foreground. This binary image was then used as a mask (Figure 3.4). The background of the original hyperspectral image was removed by multiplying the mask across the hypercube along the λ-dimension. Once only the foreground of the image remained, the kernels were individually identified and indexed according to its position. The mean absorbance spectrum for each kernel was calculated and used as input for statistical analysis.

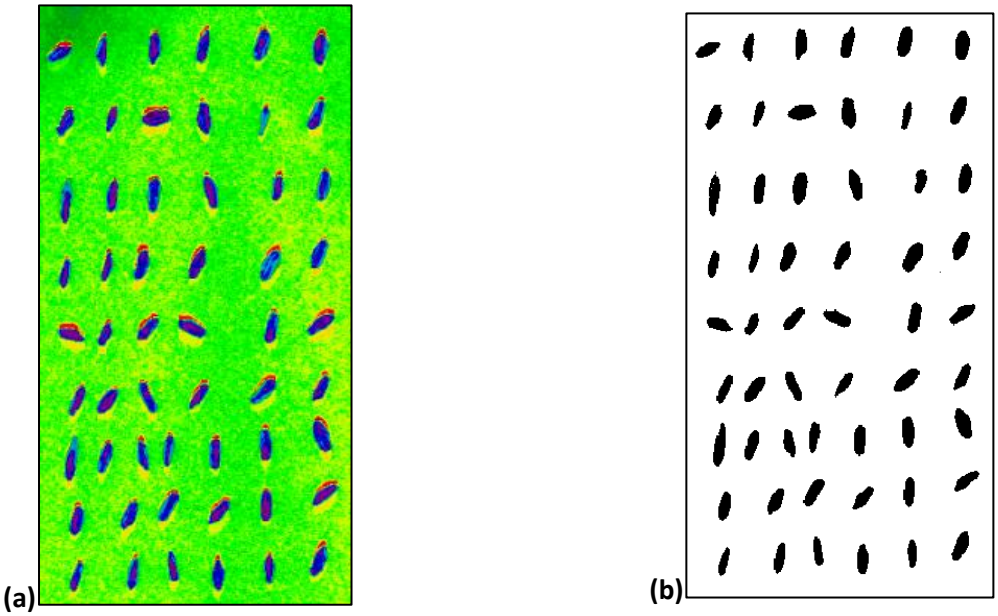


Figure 3.4 (a) A hyperspectral image of wheat kernels; **(b)** A mask created using Otsu's thresholding method.

Input Data

Hyperspectral images contain thousands of spectra. Sometimes, most of the spectra are similar as they correspond to the same measured variable. The standard approach to overcome this problem is to calculate a single spectrum that represents that portion of the sample (Caporaso *et al.*, 2018). This is typically done by calculating the mean or median spectrum of the measured variable.

Pre-processing Methods

Several pre-processing techniques were evaluated to identify the optimal combination to develop robust models. The techniques included were standard normal variate (SNV) (Barnes *et al.*, 1989), linear de-trending (DT) (Barnes *et al.*, 1989) and Savitzky-Golay (SG) 2nd derivatives (Savitzky & Golay, 1964). The different combinations of pre-processing methods, denoted as P, are presented in Table 3.1.

Table 3.1 The different combinations of pre-processing methods.

Label	Pre-processing Combination
P1	No pre-processing
P2	SNV
P3	SNV + DT (SNV-DT)
P4	DT + SNV (DT-SNV)
P5	SNV-DT + Savitzky-Golay (2nd derivative, 2nd polynomial order, 9 smoothing points)
P6	Savitzky-Golay (2nd derivative, 2nd polynomial order, 9 smoothing points)
P7	Savitzky-Golay (2nd derivative, 2nd polynomial order, 9 smoothing points) + SNV

Multivariate Data Analysis

A wide range of supervised classification algorithms was applied to the hyperspectral data namely, logistic regression, Partial Least Squares-Discriminant Analysis (PLS-DA), Linear Discriminant Analysis (LDA), k-Nearest Neighbours (KNN), Decision Trees, Random Forests (RF), and Support Vector Machines (SVM) (Table 3.2). These classification algorithms make use of machine learning that is a type of artificial intelligence (AI) providing computers with the ability to learn without being explicitly programmed (Mutihac & Mutihac, 2008). Machine learning is not to be confused with chemometrics. Chemometrics is the science of utilising mathematics, statistics and computational devices for chemical analyses (Massart *et al.*, 1988). Therefore, machine learning techniques can be applied in the field of chemometrics (Mutihac & Mutihac, 2008).

Table 3.2 Supervised classification algorithms used for model development.

Model	Statistical Classification Algorithm	Reference
CM1	Logistic Regression	(McCullagh & Nelder, 1989)
CM2	Partial Least Squares-Discriminant Analysis	(Wold *et al.*, 1987)
CM3	Linear Discriminant Analysis	(Fisher, 1936)
CM4	K-Nearest Neighbours	(Sebestyen, 1962)
CM5	Decision Trees	(Quinlan, 1986)
CM6	Random Forests	(Breiman, 2001)
CM7	Support Vector Machines (RBF)	(Cortes & Vapnik, 1995)

Firstly, a dichotomous analysis was performed where each classification algorithm was applied to a dataset that contained sound wheat and one defect. Each target category (sound vs. defect) was assigned a dummy variable, zero or one. The dataset was then split into a training set and a test set using k-fold cross-validation. The classification algorithms were trained on the training data and then applied to the test data. The predictions returned by the classification algorithm were compared to the known status of the kernel, and the performance of the model was evaluated using the performance measures. The models were further optimised by adjusting the hyperparameters. Grid search was the method used to perform hyper-parameter optimisation, by finding the best combination of hyperparameters, through k-folds cross-validation, for a given model and test dataset.

After all the dichotomous analyses were complete, a multiclass analysis was performed for the top-performing classification algorithms on all of the classes (sound and defective) at the same time. Once again, the performance measures were calculated to determine the performance of the model. Finally, the best pre-processing method and classification algorithm combination were selected as the final model for each case (dichotomous and multiclass).

Performance Measures

When using classifiers, two types of errors could occur. It could either incorrectly classify a sound kernel as defective (false negative) or incorrectly classify a defective kernel as sound (false positive). A confusion matrix, shown in Figure 3.5, is a convenient way to display this information.

n	**Predicted: NO**	**Predicted: YES**	
Actual: NO	TN	FP	Total of Row
Actual: YES	FN	TP	Total of Row
	Total of Column	Total of Column	

Figure 3.5 A confusion matrix where n is the number of samples, true negatives (TN) is a negative response correctly classified as a negative response; false positives (FP) is a negative response incorrectly classified as a positive response; false negative (FN) is a positive response incorrectly classified as a negative response; true positives (TP) is a positive response correctly classified as a positive response.

The overall performance of the classification models in combination with the various pre-processing techniques was attained by calculating the classification accuracy, false positive error and false negative error presented in Eq. 3.3, Eq. 3.4, and Eq. 3.5 respectively. The efficacy of the model was calculated using the classification accuracy. The sensitivity, specificity, and misclassification rate were additional performance measures calculated. The sensitivity, also known as recall or true positive rate, is the proportion of observed positives that were correctly classified to be positive (Eq. 3.6). The specificity is the exact opposite where it describes the percentage of observed negatives that were correctly predicted to be negatives (Eq. 3.7). The misclassification rate, or classification error, specifies how often the model was incorrect (Eq. 3.8).

$$Classification\ Accuracy\ (\%) = \frac{TN + TP}{TN + TP + FP + FN} \times 100\% \tag{3.3}$$

$$False\ Positive\ Error\ (\%) = \frac{FP}{TN + TP + FP + FN} \times 100\% \tag{3.4}$$

$$False\ Negative\ Error\ \ (\%) = \frac{FN}{TN + TP + FP + FN} \times 100\% \tag{3.5}$$

$$Sensitivity\ or\ Recall\ (\%) = \frac{TP}{TP + FN} \times 100\% \tag{3.6}$$

$$Specificity\ (\%) = \frac{TN}{TN\ + FP} \times 100\% \tag{3.7}$$

$$Misclassification\ Rate\ (\%) = \frac{FP + FN}{TN + TP + FP + FN} \times 100\% \tag{3.8}$$

Where:

True Positives (TP) = Positive response correctly classified as a positive response

True Negatives (TN) = Negative response correctly classified as a negative response

False Positives (FP) = Negative response incorrectly classified as a positive response

False Negative (FN) = Positive response incorrectly classified as a negative response

Optimal Wavelength Selection

Wavelength selection is necessary to reduce collinearity and feature redundancy. Fewer input variables also promote faster computational times and contribute to the development of multispectral systems and portable devices that are more cost-efficient. Recursive Feature Elimination (RFE) was used to identify optimal wavelengths. RFE is *wrapper method* which means it works with one machine learning algorithm and uses its performance as evaluation criteria (You *et al.*, 2014). The method recursively prunes the least important wavelengths and builds a model with the features that remain.

Prediction Maps

A script was created to remove the image background, identify each object/kernel in the image, remove pixel clusters smaller than 30 pixels, pre-process the data and perform classification analysis using the prescribed classifier. The kernels predicted as *sound* were rendered in green and the kernels predicted as *defective*, in a different colour. An example of the figure generated from the hyperspectral image is displayed in Figure 3.6.

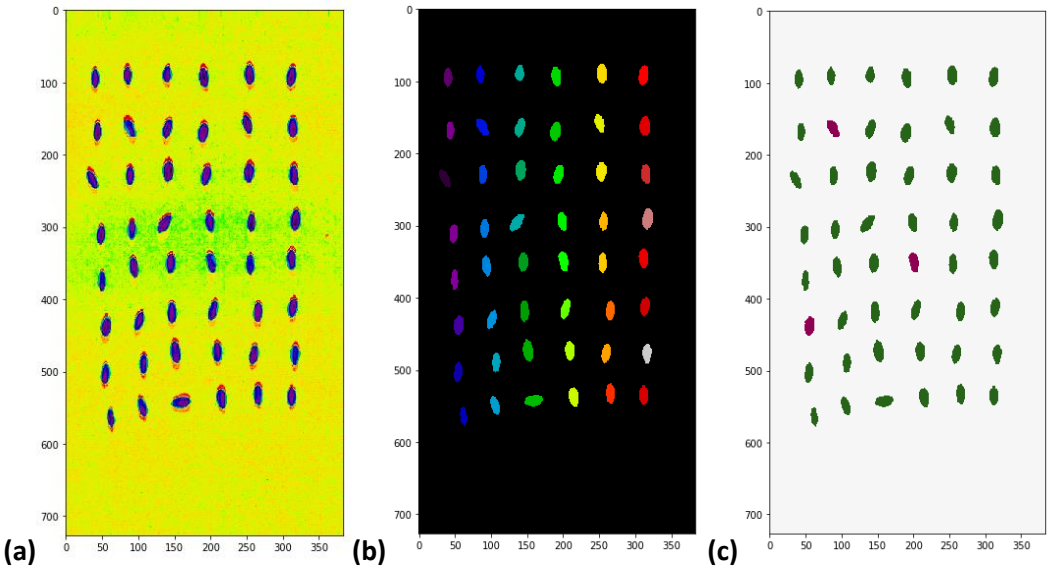


Figure 3.6 (a) Unprocessed hypercube containing sound and defective kernels; **(b)** Counting algorithm identifying each individual kernel; **(c)** Final image generated after the classification algorithm was applied, with the defective kernels in purple and the sound kernels in green.

References

Barnes, R.J., Dhanoa, M.S. & Lister, S.J. (1989). Standard Normal Variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, **43**, 772-777.

Breiman, L. (2001). Random Forests. *Machine Learning*, **45**, 5-32.

Caporaso, N., Whitworth, M.B. & Fisk, I.D. (2018). Protein content prediction in single wheat kernels using hyperspectral imaging. *Food Chemistry*, **240**, 32-42.

Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, **20**, 273-297.

Department of Agriculture (2016). Regulations relating to the grading, packing and marking of bread wheat intended for sale in the Republic of South Africa. In: *Agricultural Product Standards Act (Act No. 119 of 1990)*.

Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems *Annals of Eugenics*, **7**, 179-188.

Massart, D., Vandeginste, B., Deming, S., Michotte, Y. & Kaufman, L. (1988). *Chemometrics: A textbook.* Amsterdam: Elsevier Science Publishers.

McCullagh, P. & Nelder, J.A. (1989). *Generalized linear models.* London: Chapman & Hall/CRC.

Mutihac, L. & Mutihac, R. (2008). Mining in chemometrics. *Analytica Chimica Acta*, **612**, 1-18.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, **9**, 62-66.

Quinlan, J.R. (1986). Induction of decision trees. *Machine Learning*, **1**, 81-106.

Savitzky, A. & Golay, M.J.E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, **36**, 1627-1639.

Sebestyen, G. (1962). Pattern recognition by an adaptive process of sample set construction. *IRE Transactions on Information Theory*, **8**, 82-91.

Wold, S., Geladi, P., Esbensen, K. & Öhman, J. (1987). Multi-way principal components-and PLS-analysis. *Journal of Chemometrics*, **1**, 41-56.

You, W., Yang, Z. & Ji, G. (2014). PLS-based recursive feature elimination for high-dimensional small sample. *Knowledge-Based Systems*, **55**, 15-28.

CHAPTER 4
RESULTS AND DISCUSSION

Raw Spectral Analysis of Wheat Kernels

The mean spectrum of each wheat category was computed to investigate and compare their chemical properties. The spectral curves of the different categories (Figure 4.1) appeared to follow a similar trend, although they varied in relative intensity. The intensity variations are not necessarily associated with the chemical composition of the kernels as these spectra are unprocessed. These variations are more likely caused by the physical properties of the kernels such as morphology, particle distribution, particle size, and density variances. Additionally, these are the mean spectra, therefore a direct chemical interpretation is challenging as the spectra contain the combined information from all the components of the kernel.

Overall, five prominent absorption bands were observed at 1200, 1460, 1750, 1930 and 2100 nm for all the categories. These regions are associated with the general composition of wheat kernels and are highlighted in Figure 4.1.

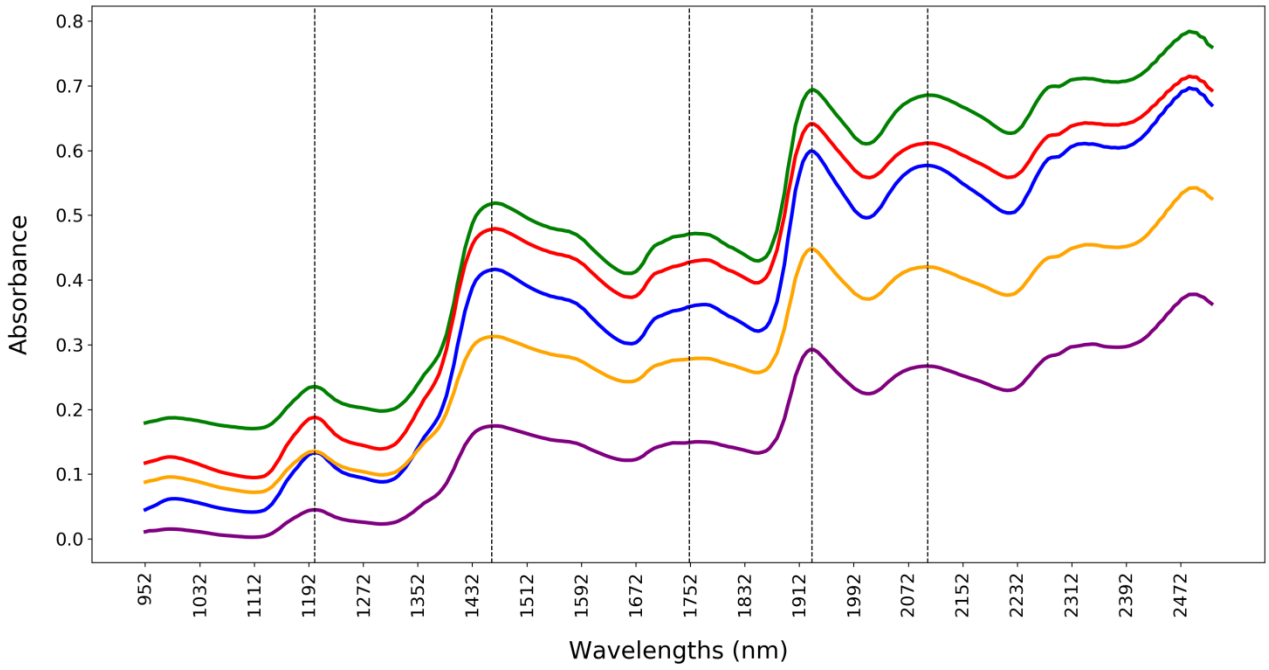


Figure 4.1 Unprocessed average pseudo-absorbance spectra for sound wheat (**blue**) and wheat defects [heat-damaged (**red**), sprout-damaged (**sprout**), *Fusarium*-damaged (**purple**), and immature wheat kernels (**green**)]; Five prominent absorption bands observed at 1200, 1460, 1750, 1930 and 2100 nm.

The two most prominent absorption bands observed at 1430 nm and 1930 nm are associated with the presence of moisture (Manley *et al.*, 2009). A less prominent moisture absorption band is observed at 960 nm (Murray & Williams, 1987; Delwiche, 1996; Wang *et al.*, 1999). Since starch molecules contain many hydroxyl groups, they too exhibit strong absorption in these areas of the spectrum. The absorbance band at

1460 nm, resulting from first overtone O–H stretching, is therefore attributed to starch. Other absorption bands that correlate to starch are 1200 nm (C–H stretch second overtone), 1750 nm (C-H first overtone region), and 2100 nm (O-H bend, C-O stretch combination) (Manley *et al.*, 2009; Joe & Gopal, 2017). The starch band at 2100 nm dominates the area of the spectrum to such an extent that the region where protein absorbs (2180 nm) (Fernández-Ibañez *et al.*, 2009) merely takes the form of a slight shoulder on the starch band. This is primarily due to the starch being significantly higher in concentration compared to that of protein. The relative sizes of these bands fluctuate depending on the ratio of the components in the kernels.

Principal Component Analysis

Principal Component Analysis (PCA) was applied to all spectral data (950 to 2500 nm) acquired from all the samples to observe the total variance among the wheat kernels. The explained variance for the first three principal components (PCs) were 94.84%, 4.49% and 0.53% respectively. Three-dimensional (3D) score plots of the first three PCs are displayed in Figure 4.2 from two different angles. While slight clustering was observed between the different categories, further pre-processing steps and classifiers were necessary to discriminate between the different categories.

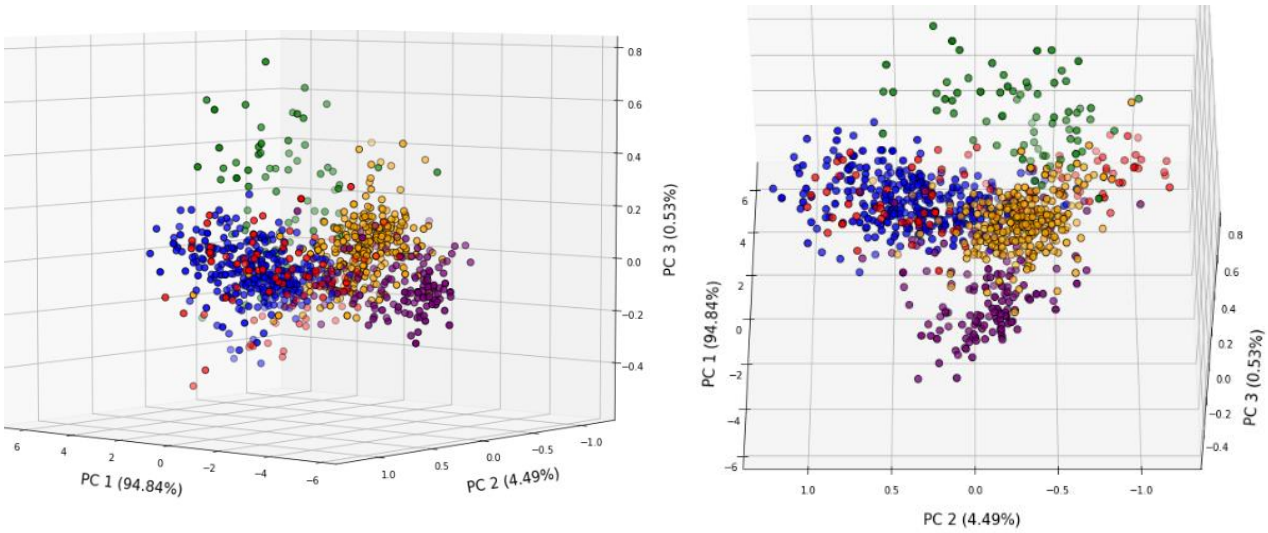


Figure 4.2 PCA 3D scores plots from two different angles for sound wheat (**blue**) and defective wheat [heat-damaged (**red**), sprout-damaged (**sprout**), *Fusarium*-damaged (**purple**), and immature wheat kernels (**green**)].

PCA was also applied to sound wheat and each defective category individually. The score plot of sound and heat-damage, Figure 4.3a, displayed no class separation, revealing that there was little difference in terms of spectral variation. It was thus evident that sound and heat-damaged wheat have minimal chemical and physical differences. The remaining score plots showed improved class separation, indicating a higher amount of variation between the two groups. The 3D score plots are displayed in Figure 4.3b – Figure 4.3d.

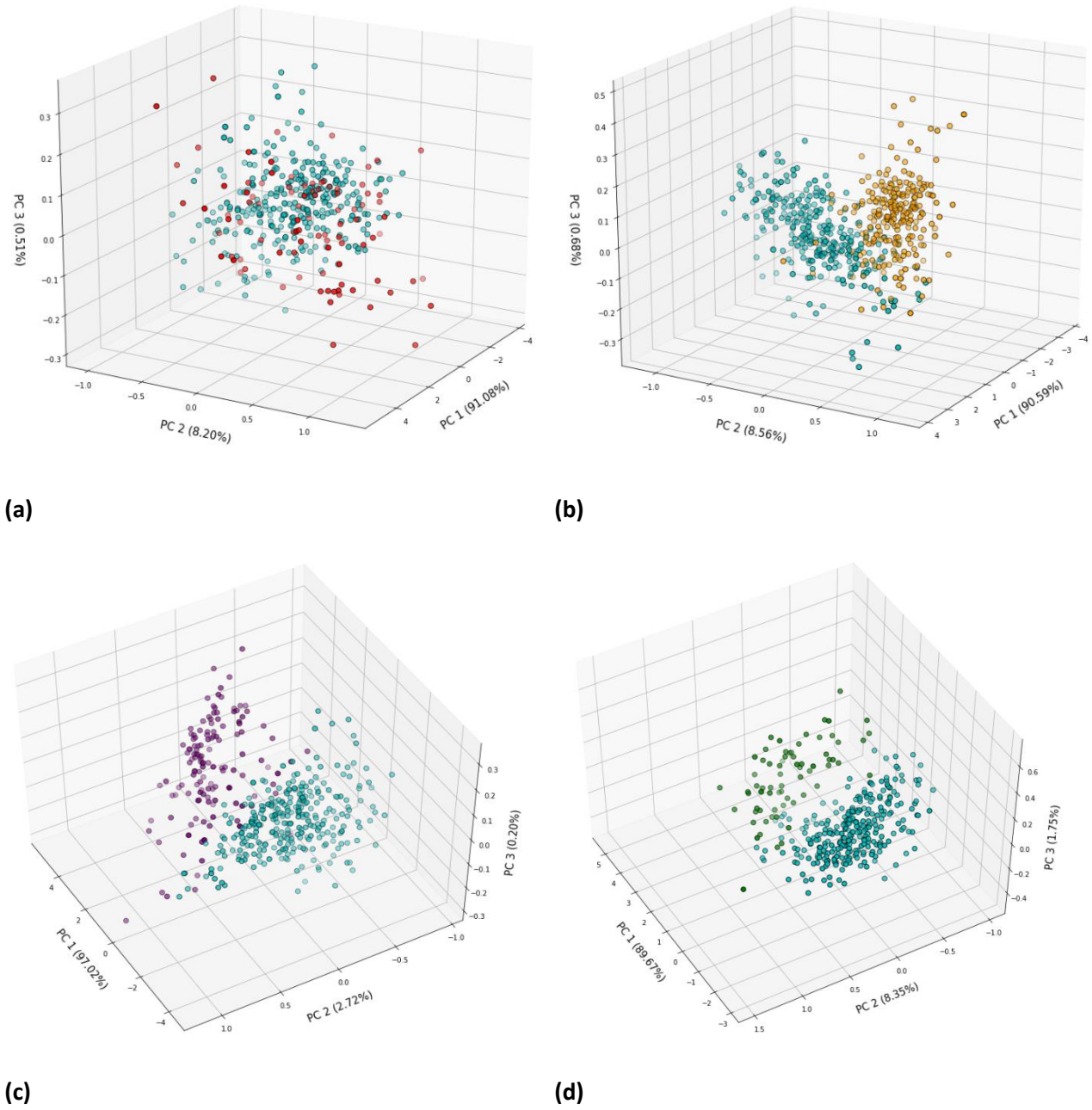


Figure 4.3 Three component PCA 3D scores plot for: **(a)** sound (**blue**) and heat-damaged wheat (**red**); **(b)** sound (**blue**) and sprout-damaged wheat (**sprout**); **(c)** sound (**blue**) and *Fusarium*-damaged wheat (**purple**); **(d)** sound (**blue**) and immature wheat (**green**).

Dichotomous Classifications

Supervised classification algorithms were used on the raw and pre-processed NIR spectra to differentiate between sound wheat and each of the defective categories. The performance measures claculated from the different classification models are given in Table 4.1 – Table 4.56. The dichotomous classifications are discussed in the arbitrary order of heat damage, sprout damage, *Fusarium* damage, and immature kernels.

Logistic Regression

Heat Damage

The results obtained by the logistic regression models for the classification between sound and heat-damaged wheat are displayed in Table 4.1 and Table 4.2. The results from the calibration set alone suggested that pre-processing the data with linear de-trending (DT) followed by standard normal variate (SNV) (P4) provides optimal results. However, after the model was verified on an external data set, it was discovered that the P4 pre-processing combination was highly inadequate. The classification accuracy decreased from 95.5% to 74.4%. This severe decrease in classification accuracy is indicative of over-fitting. In the paper by Barnes *et al.* (1989), it was recommended that de-trending should not be performed before SNV. These results reiterate the importance of the sequence of pre-treatments.

The optimal pre-processing technique for this classification was identified as 2nd derivatives followed by SNV (P7). A classification accuracy of 93.6% was achieved for the calibration set and 99.2% for the validation set. The improved accuracy is indicative of a model that was not over-fitted. The sensitivity and specificity of the validation set was 99.2% and 87.5%. The higher sensitivity indicates that the model was more sensitive at predicting a true positive (sound wheat) as correct but is not as specific when predicting a true negative (defective wheat) as correct. A total of 2.6% false positives and 0.6% false negatives were predicted for the validation set.

The pre-processing techniques, P5 and P6, provided the poorest results. The specificities for these pre-processing methods were 0%, indicating that no true negatives or false negatives were predicted. These two pre-processing combinations are therefore incompatible with the logistic regression classifier.

Table 4.1 The logistic regression calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	84.0	16.0	89.3	10.7
P2	92.5	7.5	96.0	4.0
P3	92.5	7.5	96.6	3.4
P4	95.5	4.5	74.4	25.6
P5	75.2	24.8	78.6	21.4
P6	75.2	24.8	78.6	21.4
P7	93.6	6.4	99.2	0.8

Table 4.2 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the logistic regression models on the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration

	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	15.4	0.6	99.2	38.5
P2	7.5	0.0	100.0	70.5
P3	7.5	0.0	100.0	70.5
P4	4.5	0.0	100.0	82.5
P5	24.8	0.0	100.0	0.0
P6	24.8	0.0	100.0	0.0
P7	6.4	0.0	100.0	75.0
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	10.0	0.7	99.2	53.1
P2	4.0	0.0	100.0	81.3
P3	3.3	0.0	100.0	84.4
P4	0.8	24.8	66.3	96.9
P5	21.3	0.0	100.0	0.0
P6	21.3	0.0	100.0	0.0
P7	2.6	0.6	99.2	87.5

Sprout Damage

The results for sound and sprout-damaged wheat are displayed in Table 4.3 and Table 4.4.

Table 4.3 The logistic regression calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus sprout-damaged wheat.

Model	**Calibration**		**Validation**	
	Classification Accuracy (%)	**Error Rate (%)**	**Classification Accuracy (%)**	**Error Rate (%)**
P1	94.4	5.6	91.8	8.2
P2	98.6	1.4	85.7	14.3
P3	98.5	1.5	85.7	14.3
P4	99.6	0.5	82.2	17.8
P5	53.2	46.8	80.3	19.7
P6	53.2	46.8	80.3	19.7
P7	99.6	0.5	96.6	3.4

Table 4.4 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the logistic regression models on the pre-processed data for sound versus sprout-damaged wheat.

Model	**Calibration**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**

P1	2.0	3.6	93.1	95.8
P2	1.4	0.0	100.0	97.1
P3	1.5	0.0	100.0	96.8
P4	0.5	0.0	100.0	99.0
P5	46.8	0.0	100.0	0.0
P6	46.8	0.0	100.0	0.0
P7	0.3	0.2	99.7	99.3

Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	8.2	0.0	100.0	58.6
P2	14.3	0.0	100.0	27.6
P3	14.3	0.0	100.0	27.6
P4	11.0	6.8	91.0	55.2
P5	19.7	0.0	100.0	0.0
P6	19.7	0.0	100.0	0.0
P7	2.7	0.7	91.0	86.2

The highest classification accuracy was also achieved with the P7 pre-processing method. The calibration set delivered an accuracy of 99.6% and the validation set, a accuracy of 96.6%. The validation model had a sensitivity and specificity of 91.0% and 86.2% respectively and a false positive rate of 2.7% and a false negative rate of 0.7%.

Fusarium Damage

The results for sound and *Fusarium*-damaged wheat are displayed in Table 4.5 and Table 4.6. The highest classification accuracy was achieved with the P7 pre-processing method. The calibration and validation model delivered a classification accuracy of 98.1% and 99.4% respectively. The model had a false positive rate of 0.6% and a false negative rate of 0.0%. The sensitivity and specificity was 100.0% and 98.1% respectively.

Table 4.5 The logistic regression calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	**Calibration**		**Validation**	
	Classification Accuracy (%)	**Error Rate (%)**	**Classification Accuracy (%)**	**Error Rate (%)**
P1	94.6	5.4	97.1	2.9

P2	96.7	3.3	91.8	8.2
P3	96.5	3.5	91.8	8.2
P4	98.6	1.4	92.9	7.1
P5	66.5	33.5	69.4	30.6
P6	66.5	33.5	69.4	30.6
P7	98.1	1.9	99.4	0.6

Table 4.6 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the logistic regression models on the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	**Calibration**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	5.4	0.0	100.0	83.5
P2	3.3	0.0	100.0	90.0
P3	3.5	0.0	100.0	89.4
P4	1.4	0.0	100.0	96.1
P5	33.5	0.0	100.0	0.0
P6	33.5	0.0	100.0	0.0
P7	1.6	0.4	99.4	95.5
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	2.9	0.0	100.0	90.4
P2	8.2	0.0	100.0	73.1
P3	8.2	0.0	100.0	73.1
P4	4.3	2.8	95.5	88.5
P5	30.6	0.0	100.0	0.0
P6	30.6	0.0	100.0	0.0
P7	0.6	0.0	100.0	98.1

Immature Kernels

In Table 4.7 and Table 4.8, the results for the classification between sound and immature wheat are displayed. The highest classification accuracy for both the calibration and validation set was achieved with the P7 pre-processing method. The calibration set achieved 99.8% accuracy and the validation set achieved 98.5% accuracy. The P2 pre-processing approach performed just as well in the external validation and achieved equivalent performance measures. For the calibration models, the P7 method outperformed the P2 method, although the difference was marginal. The P7 and P2 pre-processing method provided the lowest amount of false positives in the external validation, amounting to 1.5%. In the calibration model, the P2 pre-processing method predicted a much higher number of false positives (3.4%) compared to that of the P7 pre-processing method (0.2%). No false negatives were predicted with the P2 and P7 pre-processed validation sets and a sensitivity and specificity of 100.0% and 87.5% were achieved respectively. P7 and P2 pre-processing methods were the most favourable for classifying sound versus immature wheat using the logistic regression algorithm.

Table 4.7 The logistic regression calibration and validation results to assess the overall performance of the pre-processed data for sound versus immature wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	96.6	3.4	94.8	5.2
P2	96.6	3.4	98.5	1.5
P3	96.4	3.6	97.8	2.2
P4	99.6	0.5	46.7	53.3
P5	80.7	19.3	88.0	12.0
P6	80.7	19.3	88.0	12.0
P7	99.8	0.2	98.5	1.5

Table 4.8 The performance measures for the calibration and validation logistic regression model on the pre-processed data for sound versus immature wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	2.5	0.9	87.5	98.9
P2	3.4	0.0	83.1	100.0
P3	3.6	0.0	81.8	100.0
P4	0.5	0.0	97.9	100.0
P5	19.3	0.0	0.0	100.0
P6	19.3	0.0	0.0	100.0
P7	0.2	0.0	99.0	100.0

Model	Validation			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	5.2	0.0	100.0	56.3
P2	1.5	0.0	100.0	87.5
P3	2.2	0.0	100.0	81.3
P4	48.3	5.1	89.9	3.4
P5	11.9	0.0	100.0	0.0
P6	11.9	0.0	100.0	0.0
P7	1.5	0.0	100.0	87.5

Overall, the performance of the logistic regression models was satisfactory. In each dichotomous classification, the top achieving model delivered classification accuracies above 80%. The sensitivities and specificities delivered were also indicative of effective models. Across all logistic regression classifications, the P4, P5 and P6 pre-processing methods provided the most unsatisfactory results. For further investigation into the poor performance of the P5 and P6 pre-processing methods, the pre-processed spectrum for P5 was plotted and the ten critical wavelengths identified for each model by recursive feature elimination (RFE). In Figure 4.4 it can be seen that the key wavelengths are mainly attributed to moisture. Barbedo *et al.* (2018) found that sprouted kernels had a higher reflectance in the 1386 – 1700 nm spectral range than sound kernels. The authors rationalised that just before wheat kernels undergo germination, the endosperm is imbibed with water, which reduces the recorded reflectance. As germination progresses, the embryo absorbs the water, which results in the moisture content of the endosperm to decrease. This ultimately increases the reflectance in the 1386 – 1700 nm spectral range. The study by Barbedo *et al.* (2018) differed from the current study in that the authors subjected the samples to a heating chamber where the samples were dried for 24 h to reduce the water content to 10%. The samples used in the current study were not conditioned and therefore exhibited a considerable variation of moisture content. The kernels were also at various stages in the germination process; some had a visible root, while others merely had an exposed germ. In this instance, moisture content was not a reliable indicator as to whether sprouting had occurred or not. The P5 approach, however, emphasized moisture related wavelengths as key features, which in turn created an ineffective model.

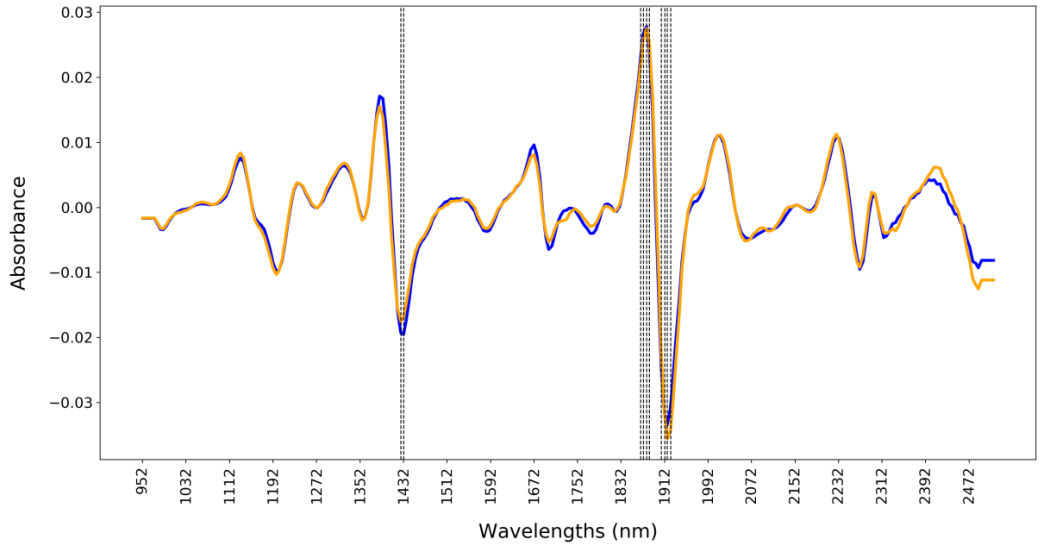


Figure 4.4 Average pseudo-absorbance spectra for P5 pre-processed sound wheat (**blue**) versus sprout-damaged wheat (**orange**). Ten key wavelengths indicated at 1427, 1432, 1868, 1873, 1879, 1884, 1906, 1912, 1917, 1923 nm.

Furthermore, 2nd derivative is useful to resolve overlapping peaks, however, the method tends to de-emphasise lower frequencies and emphasise higher frequencies, which amplifies noise more distinctly than other pre-processing methods (Buddenbaum & Steffens, 2012). This proved to be problematic in a logistic model as each data point contributes to the calculation of the regression line, making the model sensitive to outliers and noise.

Partial Least Squares Discriminant Analysis (PLS-DA)

Heat Damage

In Table 4.9, the PLS-DA classification accuracies for sound and heat-damaged wheat are displayed. The most common number of latent variables used was nine. The highest classification accuracy for the calibration model was accomplished by pre-processing the data with SNV (P2). A classification accuracy of 97.4% and a sensitivity and specificity of 100.0% and 90.1%, respectively, were achieved (Table 4.10). However, after validation, the best classification results were obtained by applying 2nd derivatives prior to SNV (P7). The classification accuracy for the validation set was 98.0%. The increase in classification accuracy from the calibration to validation set is a good indication that the model was not over-fitted. The performance measures in Table 4.10 were the true indicators that the P7 method was the optimal pre-processing method, as no false positives and only 2.0% false negatives were predicted in the validation set. The validation set also achieved a high sensitivity and specificity of 97.5% and 100% respectively, representing a useful model.

For the pre-processing methods P1 to P4, a drastic decrease in accuracy was exhibited from the calibration to validation set, signifying that these pre-processing methods were not ideal for this particular analysis with PLS-DA. Pre-processing the data with P5 to P7 produced the highest classification accuracies, where 2nd derivative was the common technique amongst the three pre-processing methods. Second

derivatives are useful to resolve overlapping peaks and correct baseline shifts. In Figure 4.5, a visual representation of the PLS-DA separation between the two categories is shown. The data was pre-processed with 2nd derivatives followed by SNV. A good amount of separation can be seen.

Table 4.9 The PLS-DA calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration			Validation	
	LV	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	9	96.1	3.9	80.7	19.3
P2	9	97.4	2.6	79.9	20.1
P3	9	96.8	3.3	78.6	21.4
P4	9	96.9	3.1	71.9	28.1
P5	8	95.4	4.6	98.7	1.3
P6	6	95.9	4.1	96.7	3.3
P7	7	95.9	4.1	98.0	2.0

Table 4.10 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the PLS-DA models on the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	3.9	0.0	100.0	84.7
P2	2.6	0.0	100.0	90.1
P3	3.3	0.0	100.0	87.6
P4	3.1	0.0	100.0	88.3
P5	4.6	0.0	100.0	82.4
P6	4.1	0.0	100.0	84.3
P7	4.1	0.0	100.0	84.7
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	19.3	0.0	100.0	9.4
P2	20.7	0.0	100.0	3.1
P3	21.3	0.0	100.0	0.0
P4	26.4	1.7	97.5	0.0
P5	1.3	0.0	100.0	93.8
P6	3.3	0.0	100.0	84.4
P7	0.0	2.0	97.5	100

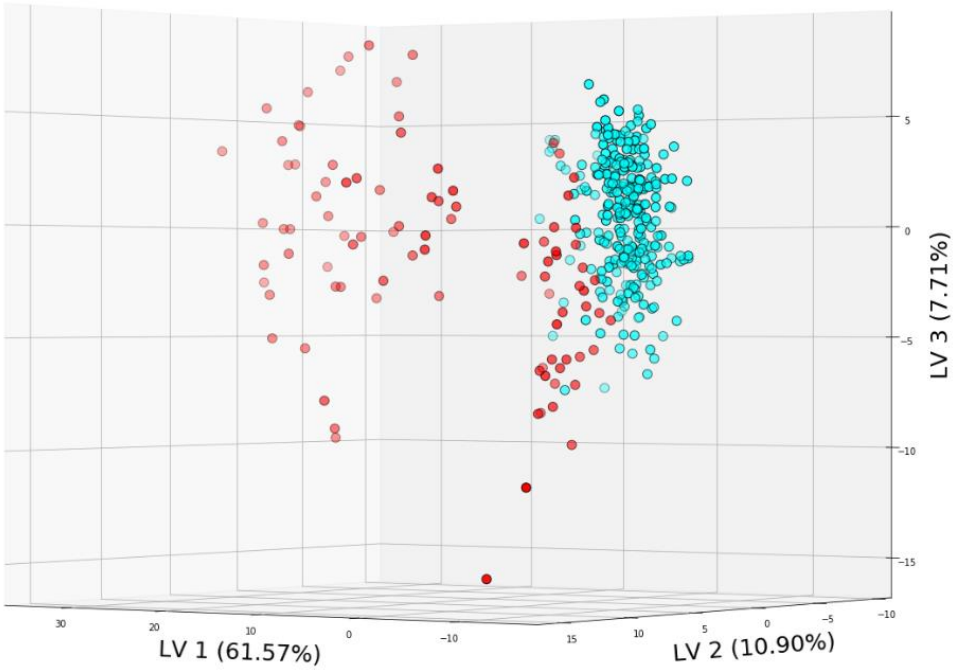


Figure 4.5 (a) PLS-DA 3D scores plot of sound (**blue**) and heat-damaged wheat (**red**); Pre-processed with 2nd Derivatives followed by SNV; LV1 (61.57%) vs LV2 (10.90%) vs LV3 (7.71%).

Sprout Damage

For classifying sound and sprout-damaged wheat, pre-processing the data with 2nd derivatives followed by SNV (P7) provided the highest performance (Table 4.11). A classification accuracy of 97.8% was achieved for the calibration set and 98.0% for the validation set. The discrimination ability of the P7 model can be seen in Figure 4.6, where a relatively clear separation between the classes is noticeable.

Table 4.11 The PLS-DA calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus sprout-damaged wheat.

Model	Calibration			Validation	
	LV	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	9	99.6	0.4	87.8	12.2
P2	9	99.2	0.8	87.8	12.2
P3	9	99.2	0.8	87.8	12.2
P4	9	98.7	1.3	82.2	17.8
P5	9	97.8	2.2	96.6	3.4
P6	9	99.6	0.4	95.9	4.1
P7	8	97.8	2.2	98.0	2.0

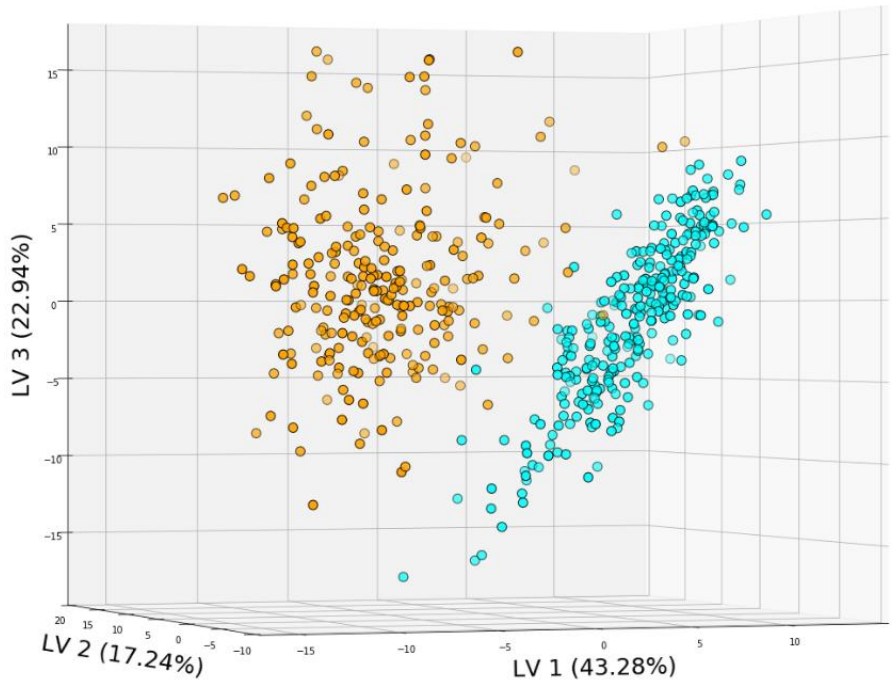


Figure 4.6 PLS-DA 3D scores plot of sound (**blue**) and sprout-damaged wheat (**orange**); Pre-processed by 2nd derivatives followed by SNV; LV1 (43.28%) vs LV2 (17.24%) vs LV3 (22.94%).

The results in Table 4.12 further substantiate the suitability of the P7 pre-processing technique. For the calibration model, the pre-processing method resulted in 2.2% false positives and zero false negatives. After validation, these results were further improved, resulting in 1.4% false positives. A high sensitivity and specificity of 99.2% and 93.3% respectively were also achieved.

Table 4.12 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the PLS-DA models on the pre-processed data for sound versus sprout-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	0.4	0.0	100.0	99.2
P2	0.9	0.0	100.0	98.2
P3	0.9	0.0	100.0	98.2
P4	1.3	0.0	100.0	97.1
P5	2.2	0.0	100.0	95.3
P6	0.4	0.0	100.0	99.2
P7	2.2	0.0	100.0	95.3

Model	Validation			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	6.8	5.4	93.2	65.5
P2	9.5	2.7	96.6	51.7
P3	8.8	3.4	95.8	55.2
P4	12.7	5.1	93.3	48.3
P5	2.0	1.4	98.3	89.7
P6	1.4	2.7	96.6	93.1
P7	1.4	0.7	99.2	93.3

In the calibration model, both the raw data (P1) and the data pre-processed with 2nd derivatives (P6) provided the highest classification accuracy of 99.6%. Since sprouted wheat kernels differ from sound wheat kernels in a structural and textural manner, it was speculated that the light scattering contributed to the discrimination of the two categories. Light scattering is due to the kernel shape and particle size differences amongst samples (Barnes *et al.*, 1989) and sprout-damaged kernels have a broken bran layer and an exposed germ. A digital image of sound wheat and sprout-damaged wheat is displayed in Figure 4.7 to visualise the difference in kernel surfaces. However, after external validation, it was clear that scatter correction was necessary. Because noise and background signals caused by interferences from the instrument and imaging conditions are random, they do not constructively contribute to the discrimination of the two categories (Siedliska *et al.*, 2018). The results from the calibration set did, however, reveal that extracting physical and textural features would be beneficial to further improve on the classification models.

Figure 4.7 Digital images of sound wheat and sprout-damaged wheat (*Photos from Agbiz Grain*).

Fusarium Damage

For classifying sound and *Fusarium*-damaged wheat, pre-processing the data with 2nd derivatives (P6) delivered the highest performance (Table 4.13). For the calibration set, an accuracy of 100% was achieved and for the validation set, an accuracy of 99.4% was achieved. The raw results from calibration to validation were undoubtedly bad, especially in comparison to the P6 pre-processed results, which emphasises the significance of pre-processing in model development. In Figure 4.15, it can be seen that P6 pre-processing resulted in no false positives or negatives for the calibration set. For the validation set, only 0.6% false

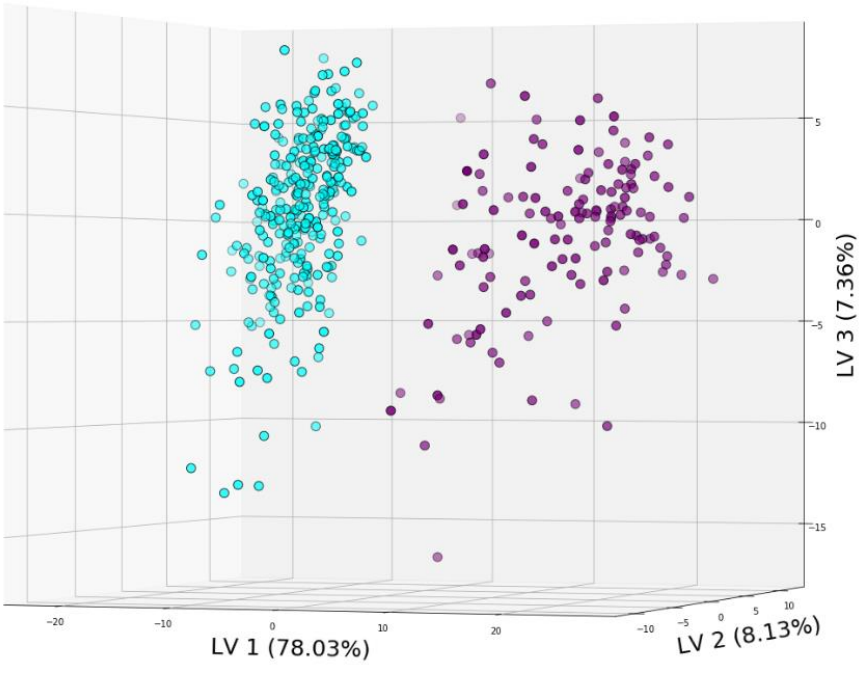
positives were predicted. A sensitivity and specificity of 100.0% and 92.3%, respectively, were achieved. The discrimination ability of the P6 model can be seen in Figure 4.8 where a clear separation between the two classes is visible.

Table 4.13 The PLS-DA calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	Calibration			Validation	
	LV	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	7	100.0	0.0	57.6	42.4
P2	8	98.5	1.6	94.7	5.3
P3	9	98.6	1.4	97.1	2.9
P4	9	98.8	1.2	90.1	9.9
P5	7	99.1	0.9	96.5	3.5
P6	9	100.0	0.0	99.4	0.6
P7	9	99.4	0.6	97.6	2.4

Table 4.14 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the PLS-DA models on the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	0.0	0.0	100.0	100.0
P2	1.6	0.0	100.0	95.6
P3	1.4	0.0	100.0	96.1
P4	1.2	0.0	100.0	96.5
P5	0.9	0.0	100.0	97.0
P6	0.0	0.0	100.0	100.0
P7	0.6	0.0	100.0	98.1
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	0.6	41.8	39.8	98.1
P2	5.3	0.0	100.0	82.7
P3	2.9	0.0	100.0	90.4
P4	3.5	6.4	94.4	82.7
P5	3.5	0.0	100.0	88.5
P6	0.6	0.0	100.0	98.1
P7	2.4	0.0	100.0	92.3



7

Figure 4.8 PLS-DA 3D scores plot of sound (**blue**) and *Fusarium*-damaged wheat (**purple**); Pre-processed by 2nd derivatives; LV1 (78.03%) vs. LV2 (8.13%) vs. LV3 (7.36%).

When analysing the calibration set results in isolation, the highest classification accuracy was accomplished by P1 and P6. Likewise with sprout-damage, *Fusarium*-damaged wheat kernels differ from sound kernels in a physical manner. The grains are small and shrunken, with a shrivelled surface area (Delwiche *et al.*, 2011) compared to sound wheat (Figure 4.9). Physical and textural features would, therefore, be favourable for the discrimination of the two categories.

Figure 4.9 Digital images of sound wheat and *Fusarium*-damaged wheat (*Photos from Agbiz Grain*).

In a study by Lim *et al.* (2018), PLS-DA was used to discriminate *Fusarium*-damaged wheat from sound wheat. High performing results were achieved overall (from 90.8% to 94.1%) and the optimal pre-processing method was identified as 2nd derivative, which corresponded to the results in the current study.

Immature Kernels

In Table 4.15 and Table 4.16, the results for sound and immature wheat are displayed. For the calibration set, 100% classification accuracy was accomplished by all pre-processing methods, except for P5. The P5

approach achieved a classification accuracy of 99.64% and a specificity of 100.00%. Since P5 consists of SNV, DT and second derivative, it is possible that the data is over pre-processed and essential information that would contribute to the discrimination between the two categories was lost. Rinnan *et al.* (2009) stated that pre-processing the data too severely could result in the loss of relevant information. For the validation set, the P5 pre-processed data did not perform as poorly and achieved a classification accuracy of 99.3%, but was not the optimal pre-processing technique. The top-performing model was achieved when pre-processing the data with 2nd derivate (P6). For the validation set, the model accomplished a classification accuracy of 100%. The preference for this model is further substantiated with the results in Table 4.16. No false positives or false negatives were predicted, producing a sensitivity and specificity of 100%.

Table 4.15 The PLS-DA calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus immature wheat.

Model	Calibration			Validation	
	LV	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	9	100.0	0.0	96.3	3.7
P2	9	100.0	0.0	98.5	1.5
P3	9	100.0	0.0	98.5	1.5
P4	9	100.0	0.0	51.7	48.3
P5	7	99.6	0.4	99.3	0.7
P6	9	100.0	0.0	100.0	0.0
P7	9	100.0	0.0	99.3	0.7

Table 4.16 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the PLS-DA models on the pre-processed data for sound versus immature wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	0.0	0.0	100.0	100.0
P2	0.0	0.0	100.0	100.0
P3	0.0	0.0	100.0	100.0
P4	0.0	0.0	100.0	100.0
P5	0.36	0.0	100.0	98.09
P6	0.0	0.0	100.0	100.0
P7	0.0	0.0	100.0	100.0

Model	Validation

	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	0.0	3.73	95.8	100.0
P2	1.5	0.0	100.0	87.5
P3	1.5	0.0	100.0	87.5
P4	46.6	1.7	96.6	6.7
P5	0.7	0.0	100.0	93.8
P6	0.0	0.0	100.0	100.0
P7	0.7	0.0	100.0	93.8

In Figure 4.10, the PLS-DA scores plot of sound and immature wheat, pre-processed with 2nd derivatives is displayed. A clear separation between the two categories can be seen, which reflects the high classification results delivered.

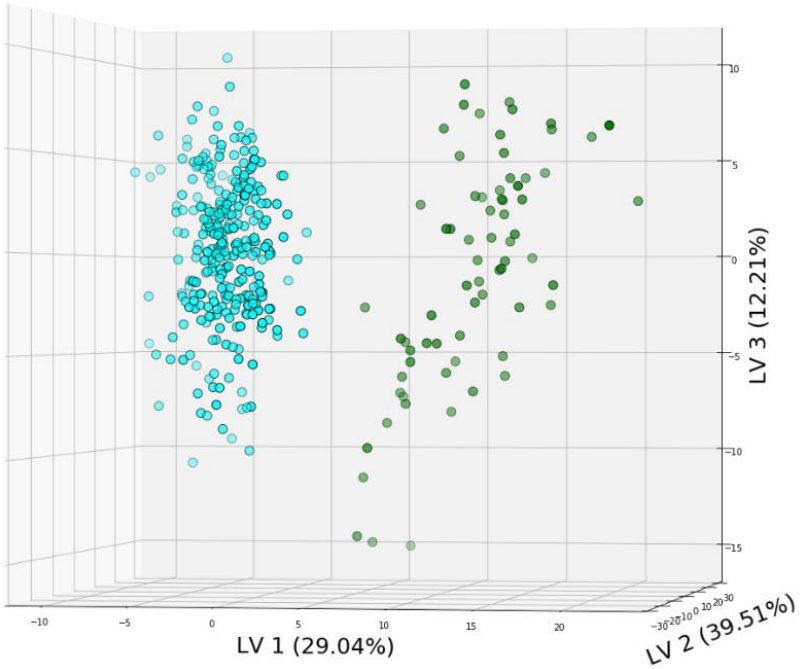


Figure 4.10 PLS-DA 3D scores plot of sound (**blue**) and immature wheat (**green**); Pre-processed by 2nd derivatives; LV1 (29.04%) vs LV2 (39.51%) vs LV3 (12.21%).

Linear Discriminant Analysis (LDA)

All LDA models only retained one component as further components gave no substantial improvement of the models. LDA is similar to PLS-DA in that it maximises the separability between two categories. However, PLS-DA outperformed LDA since PLS-DA overcomes the collinearity problems associated with LDA. Multicollinearity occurs when the predictor variables in a regression model are highly linearly related (Burger & Geladi, 2006). This correlation is a problem because the predictor variables should all be independent.

The data from NIR hyperspectral imaging is highly collinear, which creates redundant information and skews the results in a regression model.

Heat Damage

The pre-processing method that achieved the highest classification accuracy for both calibration and validation models was SNV (P2) (Table 4.17). The calibration and validation set achieved a classification accuracy of 82.1% and 90.7%, respectively. For the validation set a sensitivity and specificity of 71.2% and 93.8% were achieved (Table 4.18). Relatively low false positives and false negatives of 4.0% and 5.3%, respectively, were predicted. Overall, LDA is not fit to classify heat-damaged kernels.

Table 4.17 The LDA calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	79.8	20.2	79.3	20.7
P2	82.1	17.9	90.7	9.4
P3	79.5	20.5	89.9	10.1
P4	80.0	20.0	72.7	27.3
P5	70.1	29.9	88.0	12.0
P6	71.3	28.7	80.0	20.0
P7	70.4	29.6	76.0	24.0

Table 4.18 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the LDA models on the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	7.2	13.00	82.6	72.2
P2	6.8	11.1	85.2	73.5
P3	7.4	13.2	82.4	71.7
P4	5.1	14.9	80.0	79.9
P5	8.2	21.7	70.9	67.6
P6	8.2	20.5	72.6	68.0
P7	8.9	20.7	72.3	64.8

Model	Validation			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	2.6	18.0	77.1	87.5
P2	4.0	5.3	93.2	81.3
P3	4.0	6.7	91.5	81.3
P4	5.0	22.3	69.7	81.3
P5	4.0	8.0	89.8	81.3
P6	3.3	16.7	78.8	84.4
P7	1.3	22.7	71.2	93.8

Sprout Damage

For the calibration set, the raw data provided the highest classification accuracy (Table 4.19). The model achieved a classification accuracy of 98.2% and a sensitivity and specificity of 98.0% and 98.4% respectively (Table 4.20). This model was closely followed by the P6 model which delivered a classification accuracy of 98.1% and a sensitivity and specificity of 97.9% and 98.5% respectively. These results are similar to the PLS-DA models in that the physical properties of the kernels contributed to the efficacy of the model. Esteve Agelet *et al.* (2012) stated that scatter correction methods such as SNV removes information related to kernel structure which decreases the ability of the model to discriminate between categories that have prominent structural features. This finding reiterates that further studies can be done to combine spectral, spatial and textural features to distinguish between sound and sprouted wheat.

Table 4.19 The LDA calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus sprout-damaged wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	98.2	1.8	78.2	21.8
P2	95.3	4.7	98.0	2.0
P3	95.4	4.6	97.2	2.8
P4	95.2	4.8	79.6	20.4
P5	94.9	5.1	94.6	5.4
P6	98.1	1.9	87.8	12.2
P7	96.0	4.0	98.6	1.4

The classification accuracies of the validation set indicated that the P7 pre-processing method was more suitbale. A classification accuracy of 98.6% was achieved (Table 4.19), with a sensitivity and specificity of 100.0% and 93.3%, respectively (Table 4.20). The model predicted 1.4% false positives and no false negatives. These results are indicative of a well-performing model. The classification accuracy of the SNV (P2) pre-

processed model also performed well, where the classification accuracy of the calibration set was 95.3% and 98.0% for the validation set. The validation set had a sensitivity and specificity of 96.6% and 98.3% respectively. The model also only resulted in 1.4% false positives and 0.7% false negatives.

Table 4.20 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the LDA models on the pre-processed data for sound versus sprout-damaged wheat.

Model	Calibration			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	0.7	1.1	98.0	98.4
P2	3.5	1.2	97.8	92.5
P3	3.3	1.3	97.6	93.0
P4	3.6	1.2	97.8	92.2
P5	3.5	1.6	97.1	92.4
P6	0.7	1.2	97.9	98.5
P7	2.8	1.2	97.8	94.0
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	21.8	0.0	100.0	72.9
P2	1.4	0.7	96.6	98.3
P3	0.7	2.0	89.7	99.2
P4	16.9	3.4	86.2	77.5
P5	0.7	7.8	75.9	99.2
P6	12.2	0.0	100.0	84.7
P7	1.4	0.0	100.0	98.3

Fusarium Damage

As expected, the dichotomous classification between sound wheat and *Fusarium*-damaged behaved similarly to the sprout damage LDA classification where P1 and P6 achieved the highest classification for the calibration set. However, the external validation proved that scatter and base-line correction methods were needed to attain a well-performing model. In Table 4.21, the classification accuracies are displayed. The data pre-processed with 2nd derivatives and SNV (P7) provided the highest classification accuracy of 96.5% in the validation model. In the calibration model, the classification accuracy was much lower, amounting to 88.8%. The increase in accuracy from calibration to validation is a good indication that the model was not over-fitted. SNV also provided high results where a classification accuracy of 95.9% was achieved after validation. This is an improvement on the calibration model that delivered a classification accuracy of 84.7%. The drastic change in classification accuracies from calibration to validation emphasises the importance of

model validation on new external samples. This challenges the model and gives a better representation of the model's robustness.

Table 4.21 The LDA calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	93.3	6.7	58.2	41.8
P2	84.7	15.4	95.9	4.1
P3	84.3	15.7	95.3	4.7
P4	87.6	12.4	83.7	16.3
P5	87.0	13.0	95.3	4.7
P6	93.5	6.5	56.5	43.5
P7	88.8	11.2	96.5	3.5

In Table 4.22, it can be seen that pre-processing the data with P7 resulted in 3.5% false positives and no false negatives. A sensitivity and specificity of 100.0% and 88.5% were achieved respectively. These results are satisfactory.

Table 4.22 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the LDA models on the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	2.2	4.5	93.2	93.4
P2	8.2	7.1	89.5	75.3
P3	8.2	7.4	89.1	75.2
P4	7.9	4.5	93.4	76.5
P5	7.6	5.4	91.8	76.9
P6	2.6	0.7	94.1	92.2
P7	7.1	4.0	93.8	78.4

Model	Validation			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	0.0	41.8	39.8	100.0
P2	4.1	0.0	100.0	86.5
P3	4.1	0.6	99.2	86.5
P4	5.0	11.3	82.0	86.5
P5	4.1	0.6	99.2	86.5
P6	0.0	43.5	37.3	100.0
P7	3.5	0.0	100.0	88.5

Immature Kernels

For sound and immature wheat, two pre-processing methods, P2 and P7, delivered the best results. The classification results are displayed in Table 4.23. For the calibration models, P2 achieved 99.1%, and P7 achieved 96.6% classification accuracy. In the external validation, the accuracies increased to 99.2% and 99.3% respectively. P1, P4 and P6 performed well for the calibration set, but poorly for the validation set, resulting in misclassifications of 40.3%, 53.9% and 20.9% respectively.

Table 4.23 The LDA calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus immature wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	98.9	1.1	59.7	40.3
P2	99.1	0.9	99.2	0.8
P3	98.2	1.8	87.3	12.7
P4	98.7	1.3	46.1	53.9
P5	93.6	6.4	95.5	4.5
P6	95.5	4.6	79.1	20.9
P7	96.6	3.5	99.3	0.7

In Table 4.24 it can be seen that pre-processing the data with P2 resulted in 0.9% false positives and no false negatives for the calibration set. A sensitivity and specificity of 100.0% and 95.5% were achieved respectively. The P7 pre-processing method did not perform as well and predicted 2.7% false positives and 0.7% false negatives. A sensitivity and specificity of 99.1% and 86.7% were achieved respectively. For the external validation, the results improved where both pre-processing methods resulted in 0.7% false positives and no false negatives. A sensitivity and specificity of 100.0% and 93.8% respectively was also achieved by

both models. These results are satisfactory, and P2 was the preferred pre-processing method due to the higher performance in the calibration set.

Table 4.24 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the LDA models on the pre-processed data for sound versus immature wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	0.9	0.2	99.8	95.7
P2	0.9	0.0	100.0	95.5
P3	1.5	0.4	99.5	92.8
P4	1.1	0.2	99.8	94.1
P5	3.5	2.9	96.4	83.1
P6	2.4	2.2	97.3	87.4
P7	2.7	0.7	99.1	86.7
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	0.7	39.6	55.1	93.8
P2	0.7	0.0	100.0	93.8
P3	0.7	11.9	86.4	93.8
P4	42.7	11.2	77.5	14.6
P5	0.7	3.7	95.8	93.8
P6	0.7	20.1	77.1	93.8
P7	0.7	0.0	100.0	93.8

K-Nearest Neighbours (KNN)

Tables 4.25-4.32 display the classification results obtained with the K-Nearest Neighbours (KNN) algorithm. In a KNN model, the number of nearest neighbours (k) is crucial. A grid search with fivefold cross-validation was used to establish the amount of nearest neighbours that would provide the best classification. In this study, the resulting k values ranged between 1 and 4.

For all the dichotomous classifications, the data in raw form provided the most unsatisfactory results. KNN is a classification algorithm that relies on “majority voting” to assign an unknown sample to a category. It does this by measuring the distance (commonly Euclidean distance) of the unknown test sample to the training data. Thus, for classification algorithms where distance plays a role, standardisation is crucial to set the data to a common scale. This is reflected in the results as the application of pre-treatments involving data standardisation provided the best results across all dichotomous classifications.

Heat Damage

For sound versus heat-damaged wheat, the calibration set achieved accuracies ranging from 80.7% to 95.6% (Table 4.25). The highest classification accuracy was accomplished by pre-processing the data with SNV (P2) or SNV-DT (P3). Both these models (P2 and P3) had a k of four. A classification accuracy of 95.6% and a sensitivity and specificity of 98.6% and 87.2%, respectively, were achieved for both models (Table 4.26). The results indicated that these models are less specific and more sensitive when classifying objects. The model was thus better at predicting a true positive than a true negative.

For the validation model, the classification accuracies ranged from 78.7– 85.3% (Table 4.25). This is excluding the P4 pre-processing technique. The highest classification accuracy was accomplished by pre-processing with 2nd derivatives (85.3%), but at the same time, the model delivered a specificity of 37.5%. The pre-processing methods P1, P5 and P7 also had a low specificity of 34.4%. While P2 and P3 had higher specificities, their classification rates low, resulting in misclassifications of 21.3% and 20.7%.

Table 4.25 The KNN calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration			Validation	
	k	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	**3**	80.7	19.3	82.7	17.3
P2	**4**	95.6	4.4	78.7	21.3
P3	**4**	95.6	4.4	79.3	20.7
P4	**1**	95.2	4.8	66.9	33.1
P5	**2**	89.7	10.3	82.7	17.3
P6	**4**	88.2	11.8	85.3	14.7
P7	**2**	89.1	10.9	82.7	17.3

KNN was not suitable to classify heat-damaged wheat. The performance measures are deplorable compared to the other classification algorithms. The inability of the classifier was highlighted by the specificity parameter, which ranged between 34.4% and 68.8% for the external validation, and the high amount of false positives and negatives predicted (Table 4.26). The classification accuracy alone did not provide a proper perception of the model's performance. This emphasises the importance of additional parameters such as false positives, false negatives, sensitivity, and specificity to provide more insight and better represent a model's performance.

Table 4.26 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the KNN models on the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	13.9	5.5	92.6	45.4
P2	3.4	1.0	98.6	87.2
P3	3.4	1.0	98.6	87.2
P4	3.6	1.2	98.4	86.6
P5	8.9	1.4	98.2	66.1
P6	9.7	2.1	97.3	62.8
P7	9.1	1.9	97.3	64.9
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	14.0	3.3	95.8	34.4
P2	6.7	14.7	81.4	68.8
P3	6.7	14.0	82.2	68.8
P4	10.7	22.3	69.7	59.4
P5	14.0	3.3	95.8	34.4
P6	13.3	1.3	98.3	37.5
P7	14.0	3.3	95.8	34.4

Sprout Damage

For the calibration set, P2 and P6 pre-processing resulted in the highest performing KNN models (Table 4.27). These models had a k value of one and two respectively. The P6 pre-processing method produced the highest classification accuracy (97.5%), followed by P2 which had a classification accuracy of 97.33%. For the validation set, P6 remained the top-performing pre-processing technique, delivering a classification accuracy of 91.8%. A sensitivity and specificity of 65.5% and 98.3% were achieved respectively (Table 4.28). The model predicted 1.4% false positives and 6.8% false negatives. Overall, the KNN models were certainly challenged by the external validation set, delivering inferior results to the calibration set. The error rate of the KNN classifier is much larger compared to the other classification algorithms and is therefore not suitable for the classification between sound and sprouted wheat.

Table 4.27 The KNN calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus sprout-damaged wheat.

Model	Calibration			Validation	
	k	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	**3**	94.9	5.1	91.2	8.8
P2	**1**	97.3	2.7	81.0	19.0
P3	**2**	97.2	2.8	83.0	17.0
P4	**2**	97.3	2.7	77.1	22.9
P5	**2**	95.5	4.5	87.8	12.2
P6	**2**	97.5	2.6	91.8	8.2
P7	**2**	96.0	4.0	85.0	15.0

Table 4.28 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the KNN models on the pre-processed data for sound versus sprout-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	1.7	3.4	93.6	96.4
P2	1.9	0.7	98.7	95.8
P3	2.1	0.7	98.7	95.6
P4	2.3	0.4	99.3	95.0
P5	3.8	0.7	98.6	91.9
P6	1.5	1.1	97.9	96.9
P7	3.0	1.0	98.2	93.5
Model	**Validation**			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	4.8	4.1	79.3	94.1
P2	7.5	11.6	41.4	90.7
P3	0.0	17.0	13.8	100.0
P4	2.5	20.3	17.2	96.6
P5	1.4	10.9	44.8	98.3
P6	1.4	6.8	65.5	98.3
P7	2.0	12.9	34.5	97.5

Fusarium Damage

For sound versus *Fusarium*-damaged wheat, P6 delivered the highest classification accuracy for the calibration set, amounting to 98.3% (Table 4.29). Unlike for heat- and sprout-damaged kernels, KNN

performed excellently for the classification of *Fusarium*-damaged wheat. In the validation model, P6 pre-processing delivered the highest classification accuracy of 100.0%. Naturally, no false positives or false negatives were predicted and the sensitivity and specificity were 100.0% (Table 4.30).

Table 4.29 The KNN calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	Calibration		Validation	
	k	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	**1**	95.5	4.5	92.4	7.6
P2	**2**	97.5	2.5	94.1	5.9
P3	**1**	97.5	2.5	94.1	5.9
P4	**1**	96.9	3.1	94.3	5.7
P5	**2**	96.9	3.1	95.9	4.1
P6	**2**	98.3	1.7	100.0	0.0
P7	**2**	97.2	2.8	97.1	2.9

Table 4.30 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the KNN models on the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	3.1	1.4	97.9	90.2
P2	2.2	0.3	99.6	93.3
P3	2.2	0.3	99.6	93.3
P4	3.0	0.2	99.8	90.8
P5	3.1	0.0	100.0	90.7
P6	1.6	0.2	99.8	95.4
P7	2.8	0.0	100.0	91.6
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	0.6	7.1	89.8	98.1
P2	5.9	0.0	100.0	80.8
P3	5.9	0.0	100.0	80.8
P4	5.0	0.7	98.9	86.5
P5	3.5	0.6	99.2	88.5
P6	0.0	0.0	100.0	100.0
P7	2.4	0.6	99.2	92.3

Immature Kernels

The results for sound and immature wheat are displayed in Table 4.31 and Table 4.32 The P7 and P5 pre-processing techniques delivered the top results.

Table 4.31 The KNN calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus immature wheat.

Model		Calibration		Validation	
	k	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	**1**	96.6	3.5	94.0	6.0
P2	**2**	98.2	1.8	87.3	12.7
P3	**2**	98.2	1.8	88.1	11.9
P4	**2**	98.6	1.5	51.7	48.3
P5	**2**	97.1	2.9	95.5	4.5
P6	**2**	96.7	3.3	94.8	5.2
P7	**2**	96.6	3.5	95.5	4.5

Table 4.32 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the KNN models on the pre-processed data for sound versus immature wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	2.6	0.9	98.9	86.5
P2	1.6	0.2	99.8	92.3
P3	1.6	0.2	99.8	92.3
P4	1.5	0.0	100.0	92.9
P5	2.7	0.2	99.8	86.6
P6	2.7	0.6	99.3	86.3
P7	3.3	0.2	99.8	83.8
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	6.0	0.0	100.0	50.0
P2	1.5	11.2	87.3	87.5
P3	1.5	10.4	88.1	87.5
P4	38.2	10.1	79.8	23.6
P5	1.5	3.0	96.6	87.5
P6	0.7	4.5	94.9	93.8
P7	1.5	3.0	96.6	87.5

For the calibration set, P7 achieved 96.6% in classification accuracy and for the validation set, 95.5% accuracy was achieved. P5 also performed similarly, where the calibration set attained an accuracy of 97.1% and the validation set an accuracy of 95.5%. The k values for both P7 and P5 remained two.

In Table 4.32, it can be seen that pre-processing the data with P7 resulted in 1.5% false positives and 3.0% false negatives. A sensitivity and specificity of 96.6% and 87.5% were achieved respectively. The P5 pre-processing method achieved identical results. Once again, the classification accuracy for the calibration and validation model of the raw data gave a false impression of performance. The calibration and validation accuracies were 96.6% and 94.0%, however the validation set achieved a specificity of 50%, indicating that the model was not able to effectively classify true negatives.

Decision Trees

Heat Damage

The identification of the optimal pre-processing method for this application was redundant as all the pre-processing combinations performed very poorly. Initially, in the calibration models, the performance measures were adequate, where the error rate ranged from 7.4%–14.8% (Table 4.33). For the external validation, the error rates severely increased to 12.0%–42.7%, rendering the models unfit for classification purposes. The accompanying sensitivities and specificities were also very poor (Table 4.34), where models were either unable to predict true positives (low sensitivity) or true negatives (low specificity). The only pre-treatment that performed relatively well was P7 that delivered a classification accuracy of 92.7% for the calibration model and 87.3% for the validation model. The model had a sensitivity of 92.4% and a specificity of 71.9%. False positives and false negatives both amounted to 6.0%. In this case, decision trees were not the ideal classifier to discriminate sound and heat-damaged wheat.

Table 4.33 The Decision Trees calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	85.2	14.8	84.0	16.0
P2	91.5	8.6	57.3	42.7
P3	91.7	8.3	61.3	38.7
P4	91.7	8.3	57.9	42.1
P5	92.0	8.0	86.7	13.3
P6	93.6	6.4	88.0	12.0
P7	92.7	7.3	87.3	12.7

Table 4.34 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the Decision Trees models on the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	9.5	5.3	92.9	63.5
P2	4.7	3.9	94.9	82.4
P3	3.9	4.4	94.0	85.8
P4	3.6	4.7	93.6	86.7
P5	3.7	4.3	94.2	86.2
P6	3.3	3.1	95.8	87.8
P7	4.0	3.3	95.6	85.0
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	10.6	10.0	87.3	50.0
P2	4.7	36.7	53.4	78.1
P3	0.0	34.0	56.8	100.0
P4	0.0	42.1	42.7	100.0
P5	8.7	3.3	95.8	59.4
P6	10.0	3.3	95.8	53.1
P7	6.0	6.0	92.4	71.9

Sprout Damage

As with heat-damaged wheat, the identification of an optimal pre-processing method for sprout-damaged wheat was senseless as all the pre-processing combinations performed poorly. The results achieved are displayed in Table 4.35 and Table 4.36.

Table 4.35 The Decision Trees calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus sprout-damaged wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	94.4	5.6	75.5	24.5
P2	94.1	5.9	56.5	43.5
P3	94.9	5.2	50.3	49.7
P4	94.6	5.5	70.3	29.7
P5	90.8	9.2	79.6	20.4
P6	96.5	3.5	84.4	15.6
P7	92.7	7.3	79.6	20.4

Initially, for the calibration set, the performance measures were sufficient, where the error rate ranged from 3.5%–7.3% (Table 4.35). For the external validation, the error rates drastically increased to 15.6%–49.7%. The only pre-processing technique that gave relatively adequate results was P6. The P6 pre-processing method allowed for a classification accuracy of 96.5% for the calibration sets and 84.4% for the validation set. This resulted in a misclassification rate of 3.5% for the calibration model and 15.6% for the validation model. The model had a sensitivity of 89.7% and a specificity of 84.7%. The false positives and false negatives both amounted to 12.2% and 2.0% respectively.

Overall, the models were poor at classifying sprout-damaged wheat. The P3 pre-processing technique did not predict any true negatives at all. This resulted in the 100.0% sensitivity which gives a false perception of a high performing model. However, a specificity of 45.8% was delivered emphasising the model's inability to predict true negatives. The decision tree classifier was not suitable to discriminate sound and sprout-damaged wheat.

Table 4.36 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the Decision Trees models on the pre-processed data for sound versus sprout-damaged wheat.

Model	Calibration			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	2.4	3.1	94.2	94.8
P2	2.7	3.1	94.2	94.2
P3	2.4	2.7	95.0	94.8
P4	2.3	3.1	94.1	95.0
P5	4.3	4.9	90.9	90.6
P6	1.8	1.7	96.8	96.0
P7	3.3	3.9	92.6	92.9
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	20.4	2.7	86.2	74.6
P2	42.9	0.7	96.6	46.6
P3	43.5	0.0	100.0	45.8
P4	16.1	8.5	65.5	78.7
P5	17.0	4.1	79.3	78.8
P6	12.2	2.0	89.7	84.7
P7	12.2	8.2	58.6	84.7

Fusarium Damage

The decision trees performed better for the classification between sound and *Fusarium* damage compared to the other defective categories. The results are displayed in Table 4.37 and Table 4.38.

Table 4.37 The Decision Trees calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	95.7	4.3	85.3	14.7
P2	95.7	4.3	61.8	38.2
P3	95.9	4.1	93.5	6.5
P4	94.8	5.2	78.0	22.0
P5	96.9	3.1	90.0	10.0
P6	98.6	1.4	100.0	0.0
P7	96.6	3.4	98.8	1.2

Table 4.38 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the Decision Trees models on the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	2.7	1.6	97.7	91.7
P2	2.3	1.9	97.1	93.2
P3	2.7	1.4	97.9	92.1
P4	3.6	1.6	97.7	88.7
P5	1.6	1.6	97.6	95.6
P6	0.7	0.8	98.8	97.9
P7	1.8	1.6	97.7	94.8
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	0.6	18.2	76.7	98.1
P2	11.2	25.9	62.7	63.5
P3	12.4	7.1	89.8	59.6
P4	2.8	22.0	65.2	92.3
P5	1.8	9.4	86.4	94.2
P6	0.0	0.0	100.0	100.0
P7	0.0	10.0	85.6	100.0

Second derivative (P6) was the optimal pre-processing method and provided the highest classification accuracy for the calibration and validation model. The calibration set achieved 98.6% in accuracy and a sensitivity and specificity of 98.8% and 97.9% respectively. The results improved for the validation set, where 100% accuracy was achieved.

Immature Kernels

The decision trees classifier also performed better for the discrimination between sound and immature kernels. The classification accuracies are displayed in Table 4.39.

Table 4.39 The Decision Trees calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus immature wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	97.3	2.7	96.3	3.7
P2	98.2	1.8	96.3	3.7
P3	97.6	2.4	93.3	6.7
P4	98.8	1.2	48.3	51.7
P5	96.8	3.2	90.3	9.7
P6	97.3	2.7	97.7	2.3
P7	97.7	2.3	94.8	5.2

The P6 pre-processing method provided the highest classification accuracy for the calibration and validation model. The calibration model achieved 97.3% and the validation model achieved 97.7% in classification accuracy. The results in Table 4.40 further substantiate the model's performance. The calibration model delivered a sensitivity of 98.5% and a specificity of 92.1%. The validation model, a sensitivity and specificity of 100% and 81.3% were achieved respectively. The performance of the SNV (P2) was also satisfactory where the model achieved 97.3% accuracy in the calibration model and 97.7% accuracy in the validation model. The validation model predicted 1.5% false positives and false negatives and delivered a sensitivity and specificity of 98.3% and 97.41% respectively. Since SNV (P2) produced a higher specificity than second derivatives (P6), P2 was perceived to be the preferred pre-processing technique as it resulted in less false positives. However, a greater sample size and more validations are necessary to gain conclusive results.

Table 4.40 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the Decision Trees models on the pre-processed data for sound versus immature wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	1.5	1.2	98.5	92.5
P2	0.6	1.2	98.5	97.4
P3	1.7	0.8	99.1	92.4
P4	0.6	0.6	99.3	97.0
P5	2.0	1.2	98.5	89.9
P6	1.5	1.2	98.5	92.1
P7	1.5	0.8	99.1	92.3
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	3.0	0.0	100.0	75.0
P2	1.5	1.5	98.3	87.5
P3	1.5	5.2	94.1	87.5
P4	43.3	2.2	95.5	13.5
P5	2.2	7.5	91.5	81.3
P6	2.2	0.0	100.0	81.3
P7	2.2	3.0	96.6	81.3

Random Forests (RF)

The results indicated that it was possible to distinguish defective wheat from sound wheat. As expected, the random forest models performed better than the decision tree models. Random forests is a collection of decision trees where the number of trees is a parameter that needs to be tuned. An advantage of building a random forest model is that a higher number of trees does not cause overfitting. Generally, the more decision trees incorporated, the better the performance of the model. A higher number of trees merely increases the computational time. For the random forest models in the present study, an arbitrary number of 100 trees were used.

Heat Damage

For the dichotomous classification of sound wheat and heat-damaged wheat, the results are displayed in Table 4.41 and Table 4.42. The highest performance was achieved with P5 and P6. For the calibration set, the P5 pre-processing combination resulted in a classification accuracy of 95.2%. The sensitivity and specificity of the model was 99.1% and 84.4% respectively. The number of false positives and false negatives amounted to 4.1% and 0.7%. In the validation model, P5 resulted in a slightly lower classification accuracy of 92.0%. The sensitivity and specificity of the model was 98.3% and 75.0% respectively. The number of false

positives and false negatives amounted to 5.3% and 3.1% respectively. The P6 pre-processed data produced minor differences compared to P5. The classification accuracy of both the calibration and validation model was slightly higher, amounting to 95.7% and 94.0%, however the P5 pre-processing method allowed for a more sensitive model than P6. In the calibration model, the P6 pre-processing method had a sensitivity of 98.6% and a specificity of 87.5%. In the validation model, a sensitivity of 98.3% and a specificity of 71.9% were achieved. The differences in validation results are better understood using the confusion matrix depicted in Figure 4.11.

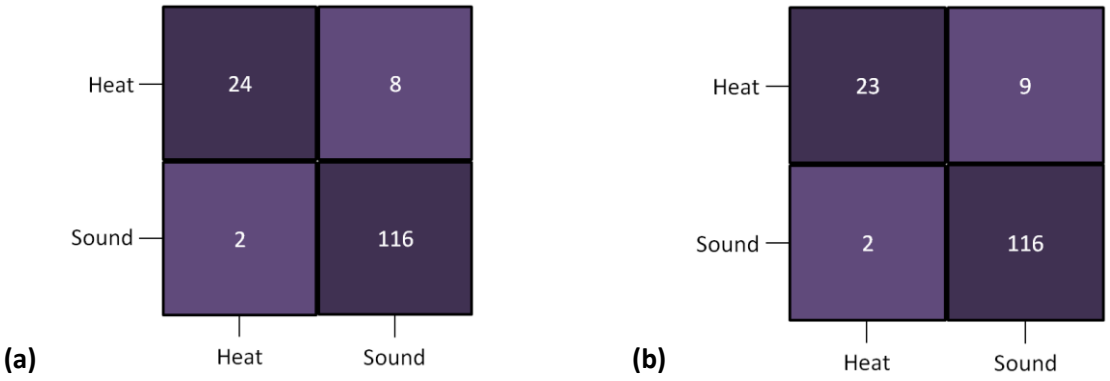


Figure 4.11 Confusion matrices for the random forest classifier using **(a)** P5 pre-processed data and **(b)** P6 pre-processed data.

In Figure 4.11b it can be seen that one true negative value was incorrectly predicted positive, resulting in a false positive. That is the only difference in results between the two pre-processing techniques. These pre-processing methods performed similarly, with insignificant differences. More conclusive results would be achieved with more validation tests and larger sample sizes.

Table 4.41 The RF calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	86.2	13.9	86.0	14.0
P2	94.5	5.5	76.7	23.3
P3	94.9	5.1	80.0	20.0
P4	96.1	3.9	65.3	34.7
P5	95.2	4.8	92.0	8.0
P6	95.7	4.3	94.0	6.0
P7	95.2	4.8	91.3	8.7

Table 4.42 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the RF models on the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	12.5	1.4	98.2	51.6
P2	3.8	1.7	97.7	85.7
P3	3.6	1.5	97.9	86.5
P4	3.1	0.9	98.9	88.4
P5	4.1	0.7	99.1	84.4
P6	3.3	1.0	98.6	87.5
P7	3.9	0.9	98.8	85.1
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	12.7	2.7	96.6	40.6
P2	0.7	16.0	79.7	96.9
P3	0.7	20.7	73.7	96.9
P4	0.0	35.5	64.5	51.7
P5	5.3	1.3	98.3	75.0
P6	6.0	1.3	98.3	71.9
P7	6.7	1.3	98.3	68.8

Sprout Damage

For sound and heat-damaged wheat, the performance measures are displayed in Table 4.43 and Table 4.44. The highest performance was achieved by P1 and P6. For the calibration set, the raw data resulted in a classification accuracy of 95.5%. The sensitivity and specificity of the model was 95.6% and 95.3% respectively. The number of false positives and false negatives amounted to 2.2% and 2.3%. In the validation set, P1 resulted in a slightly lower classification accuracy of 93.2%. The sensitivity and specificity of the model was 86.2% and 94.5% respectively. The number of false positives and false negatives amounted to 4.1% and 2.7% respectively. The results of the P6 varied slightly to P1. The classification accuracy of the calibration set was slightly higher than the P1 model, amounting to 97.3%. The classification accuracy of the validation set was 92.5% which was lower than the P1 model. In the calibration set, P6 had a sensitivity of 97.6% and a specificity of 96.9%. The validation set had a sensitivity of 72.4% and a specificity of 98.3%. While the raw data seemed to perform better, more validation tests are necessary to establish whether this finding is consistent. Raw data contains many variations that are unrelated to the chemical response. For this RF model, due to only one external validation test performed, the results are inconclusive as to whether a robust model is attainable using the raw spectra. Further tests need to be done to confirm these findings. Since P6

is preferred for all the other RF dichotomous classifications in this study, a larger sample size may prove this to be true for sound and sprout discrimination too.

Table 4.43 The RF calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus sprout-damaged wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	95.5	4.5	93.2	6.8
P2	97.8	2.2	91.8	8.2
P3	97.5	2.6	87.8	12.2
P4	98.1	1.9	79.7	20.3
P5	96.0	4.0	89.1	10.9
P6	97.3	2.7	92.5	7.5
P7	96.4	3.6	88.4	11.6

Table 4.44 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the RF models on the pre-processed data for sound versus sprout-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	2.2	2.3	95.6	95.3
P2	1.2	1.0	98.2	97.3
P3	1.5	1.1	98.0	96.8
P4	0.9	1.1	98.0	98.2
P5	1.6	2.4	95.4	96.6
P6	1.5	1.2	97.6	96.9
P7	1.9	1.7	96.7	95.8
Model	**Validation**			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	4.1	2.7	86.2	94.9
P2	4.1	5.4	72.4	94.9
P3	2.7	7.5	62.1	96.6
P4	11.0	5.1	79.3	85.4
P5	4.1	5..4	72.4	94.9
P6	1.4	5.4	72.4	98.3
P7	3.4	6.8	65.5	95.8

Fusarium Damage

For the classification of sound and *Fusarium*-damaged wheat, the results are displayed in Table 4.45 and Table 4.46.

Table 4.45 The RF calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	95.5	4.5	92.4	7.6
P2	96.6	3.4	93.5	6.5
P3	96.1	3.9	92.3	7.7
P4	96.7	3.3	87.9	12.1
P5	98.6	1.4	99.4	0.6
P6	99.2	0.8	100.0	0.0
P7	98.0	2.0	97.6	2.4

Table 4.46 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the RF models on the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	2.8	1.7	97.4	91.2
P2	3.0	0.5	99.3	90.8
P3	3.3	0.6	99.1	89.9
P4	3.0	0.3	99.5	90.8
P5	1.2	0.2	99.8	96.1
P6	0.8	0.0	100.0	97.5
P7	2.0	0.0	100.0	93.9
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	0.6	6.5	90.7	98.1
P2	7.6	0.0	100.0	75.0
P3	7.6	0.0	100.0	75.0
P4	5.7	4.3	93.3	84.6
P5	0.6	0.0	100.0	98.1
P6	0.0	0.0	100.0	100.0
P7	1.2	0.0	100.0	96.2

The P6 pre-processing technique resulted in the best performance measures, where 100% classification accuracy was achieved for the validation model. In the calibration model, P6 also had a high classification accuracy, amounting to 99.2%. Naturally, the validation set had a sensitivity and specificity of 100%, resulting in a complete classification.

Immature Kernels

Excellent results were achieved for sound versus immature wheat. The results are displayed in Table 4.47 and Table 4.48.

Table 4.47 The RF calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus immature wheat.

Model	Calibration		Validation	
	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	97.6	2.4	96.3	3.7
P2	98.0	2.0	97.8	2.2
P3	97.1	2.9	94.0	6.0
P4	99.5	0.5	52.2	47.8
P5	98.0	2.0	97.8	2.2
P6	99.3	0.7	99.3	0.7
P7	98.6	1.5	97.8	2.2

Table 4.48 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the RF models on the pre-processed data for sound versus immature wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	1.1	1.3	98.4	94.8
P2	1.8	0.2	99.8	91.8
P3	2.9	0.0	100.0	86.2
P4	0.4	0.2	99.8	98.2
P5	1.6	0.4	99.5	91.9
P6	0.7	0.0	100.0	96.2
P7	1.3	0.2	99.8	93.9

Model	Validation			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	3.0	0.0	100.0	75.0
P2	1.5	2.2	97.5	87.5
P3	2.2	3.7	95.8	81.3
P4	46.1	3.4	93.3	7.9
P5	1.55	0.0	100.0	87.5
P6	0.7	0.0	100.0	93.8
P7	1.5	0.0	100.0	87.5

The P6 combination was the top-performing pre-processing method. The calibration and validation models achieved identical classification accuracies of 99.3%. The sensitivity and specificity of the P6 calibration set was 100.0% and 96.2% respectively. For the validation model, a sensitivity and specificity of 100.0% and 93.8% were achieved respectively. Only 0.7% false positives were predicted for both the calibration and validation set.

Support Vector Machines (SVM)

The overall performance of the SVM models was satisfactory. Across all the models, Savitzky-Golay 2nd derivative (P6) alone was not effective, and a subsequent scatter correction method, SNV, was required to increase the performance. The SVM models performed well due to the use of the "kernel trick". When the training data is not linearly separable, the data is projected into a higher dimensional space to optimise the model's ability to identify a hyperplane that separates the categories.

To identify the optimal hyperparameters for each model, a grid search (GridSearchCV) was performed. GridSearchCV does an exhaustive search over all the specified parameter values for the classification algorithm. The "C" and "gamma" parameters needed to be defined for the RBF SVM. The parameters varied, depending on the pre-processing methods used and the categories analysed. The optimal parameters are displayed in Table 4.49 – Table 4.56.

Heat Damage

The results achieved by the SVM models for the classification between sound and heat-damaged wheat are displayed in Table 4.49 and Table 4.50.

Table 4.49 The SVM calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration				Validation	
	C	γ	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	1000	0.1	98.6	1.4	96.0	4.0
P2	1000	0.1	99.0	1.0	81.3	18.7
P3	1000	0.1	99.0	1.0	82.0	18.0
P4	1000	0.01	99.2	0.9	67.8	32.2
P5	1000	1	89.7	10.3	92.0	8.0
P6	1000	1	85.6	14.4	82.0	18.0
P7	100	0.01	97.4	2.6	97.3	2.7

Table 4.50 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the SVM models on the pre-processed data for sound versus heat-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	1.4	0.0	100.0	94.9
P2	0.9	0.2	99.8	96.8
P3	0.9	0.2	99.8	96.8
P4	0.7	0.2	99.8	97.4
P5	10.3	0.0	100.0	60.3
P6	14.4	0.0	100.0	44.4
P7	2.4	0.2	99.8	90.8
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	0.6	3.3	95.8	96.9
P2	2.0	16.7	78.8	90.6
P3	2.7	15.3	80.5	87.5
P4	0.8	31.4	57.3	96.9
P5	8.0	0.0	100.0	62.5
P6	18.0	0.0	100.0	15.6
P7	1.3	1.3	98.3	93.8

After the external validation, it was concluded that 2nd derivative followed by SNV (P7) provided the highest classification accuracy. The model had a "C" of 100 and a "gamma" of 0.01. The calibration classification accuracy was 97.4% and the validation classification accuracy 97.3%. The P7 model delivered a sensitivity and specificity of 99.8% and 90.8% respectively for the calibration set. The validation set delivered a

sensitivity and specificity of 98.3% and 93.8% respectively. The false positives and false negatives amounted to 1.3% each. These results are excellent especially given that heat-damaged and sound wheat were the most challenging categories to separate.

Sprout Damage

The results achieved with the SVM models for the classification between sound and sprout-damaged wheat are displayed in Table 4.51. The P7 model also provided the highest classification accuracy. The model had a "C" of 100 and a "gamma" of 0.01 and achieved a classification accuracy of 99.3% for the calibration set and 98.6% for the validation set.

Table 4.51 The SVM calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus sprout-damaged wheat.

Model	Calibration				Validation	
	C	γ	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	1000	1	99.5	0.5	68.7	31.3
P2	100	0.1	99.9	0.1	95.2	4.8
P3	100	0.1	99.9	0.1	95.2	4.8
P4	100	0.01	99.6	0.4	77.1	22.9
P5	1000	1	99.3	0.7	96.6	3.4
P6	1000	1	94.9	5.1	89.1	10.9
P7	100	0.01	99.3	0.7	98.6	1.4

In Table 4.52, it can be seen that the P7 model delivered a sensitivity and specificity of 99.8% and 98.7% respectively for the calibration set. The model predicted 0.6% false positive and 0.1% false negatives. The validation set delivered a sensitivity and specificity of 96.6% and 99.2% respectively. The false positives and false negatives amounted to 0.7% each.

Table 4.52 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the SVM models on the pre-processed data for sound versus sprout-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	0.5	0.0	100.0	99.0
P2	0.1	0.0	100.0	99.8
P3	0.1	0.0	100.0	99.8
P4	0.4	0.0	100.0	99.2
P5	0.7	0.0	100.0	98.4
P6	1.2	3.9	92.7	97.4
P7	0.6	0.1	99.8	98.7

Model	Validation			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	30.6	0.7	96.6	61.9
P2	2.0	2.7	86.2	97.5
P3	2.0	2.7	86.2	97.5
P4	18.6	4.2	82.8	75.3
P5	0.7	2.7	86.2	99.2
P6	4.1	6.8	65.5	94.9
P7	0.7	0.7	96.6	99.2

Fusarium Damage

The results achieved with the SVM models for the classification between sound and *Fusarium*-damaged wheat are displayed in Table 4.53 and Table 4.54. The highest performance was achieved with P5 and P6. The model had a "c" of 100 and a "gamma" of 0.01. For the calibration model, the P5 pre-processing combination resulted in a classification accuracy of 99.9%. The sensitivity and specificity of the model was 100.0% and 99.5% respectively. The number of false positives and false negatives amounted to 0.2% and 0.0%. In the validation model, P5 resulted in a slightly lower classification accuracy of 99.4%. The sensitivity and specificity of the model was 100.0% and 98.1% respectively. The number of false positives and false negatives amounted to 0.6% and 0.0% respectively. The P6 pre-processed data produced identical results in the validation model compared to P5. The classification accuracies were both 99.4%, and the sensitivities and specificities 100.0% and 98.1% respectively. Both these models are suitable for the discrimination of sound and *Fusarium*-damaged wheat; however, with further validations, one pre-processing technique will eventually outperform the other.

Table 4.53 The SVM calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	Calibration				Validation	
	C	γ	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	1000	0.1	99.5	0.5	68.8	31.2
P2	10	1	100.0	0.0	97.6	2.4
P3	10	1	100.0	0.0	97.6	2.4
P4	10	0.1	100.0	0.0	94.3	5.7
P5	1000	1	99.9	0.2	99.4	0.6
P6	1000	1	96.3	3.7	99.4	0.6
P7	10	0.1	99.4	0.6	98.8	1.2

Table 4.54 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the SVM models on the pre-processed data for sound versus *Fusarium*-damaged wheat.

Model	Calibration			
	False Positives (%)	False Negatives (%)	Sensitivity (%)	Specificity (%)
P1	0.5	0.0	100.0	98.5
P2	0.0	0.0	100.0	100.0
P3	0.0	0.0	100.0	100.0
P4	0.0	0.0	100.0	100.0
P5	0.2	0.0	100.0	99.5
P6	3.7	0.0	100.0	88.5
P7	0.6	0.0	100.0	98.0
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	0.0	31.2	55.1	100.0
P2	0.6	1.8	97.5	98.1
P3	0.6	1.8	97.5	98.1
P4	0.7	5.0	92.1	98.1
P5	0.6	0.0	100.0	98.1
P6	0.6	0.0	100.0	98.1
P7	0.6	0.6	99.2	98.1

Immature Kernels

The results for sound and immature wheat are displayed in Table 4.55 and Table 4.56.

Table 4.55 The SVM calibration and validation model accuracies to assess the overall performance of the pre-processed data for sound versus immature wheat.

Model	Calibration				Validation	
	C	γ	Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
P1	100	0.1	99.5	0.5	97.8	2.2
P2	10	1	100.0	0.0	92.5	7.5
P3	10	1	99.8	0.2	92.5	7.5
P4	1	1	99.6	0.4	51.1	48.9
P5	1000	1	100.0	0.0	99.3	0.7
P6	1000	1	79.8	20.2	91.8	8.2
P7	1	0.1	100.0	0.0	95.5	4.5

Table 4.56 The performance measures (false positives, false negatives, sensitivity and specificity) to assess the SVM models on the pre-processed data for sound versus immature wheat.

Model	Calibration			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	0.6	0.0	100.0	97.0
P2	0.0	0.0	100.0	100.0
P3	0.2	0.0	100.0	99.2
P4	0.4	0.0	100.0	98.3
P5	0.0	0.0	100.0	100.0
P6	20.2	0.0	100.0	0.0
P7	0.0	0.0	100.0	100.0
Model	**Validation**			
	False Positives (%)	**False Negatives (%)**	**Sensitivity (%)**	**Specificity (%)**
P1	2.2	0.0	100.0	81.3
P2	1.5	6.0	93.2	87.5
P3	1.5	6.0	93.2	87.5
P4	31.5	17.4	65.2	37.1
P5	0.7	0.0	100.0	93.8
P6	8.2	0.0	100.0	31.3
P7	1.5	3.0	96.6	87.5

The highest performance was achieved with the P5 pre-processing combination. For the calibration set, P5 resulted in a classification accuracy of 100.0%. The sensitivity and specificity of the model were both 100.0%, and the number of false positives and false negatives amounted to 0.0%. In the validation model, P5 resulted in a slightly lower classification accuracy of 99.3%. The sensitivity and specificity of the model was 100.0% and 93.8% respectively. The number of false positives and false negatives amounted to 0.7% and 0.0% respectively.

Optimal Model Overview

For each dichotomous classification, each classifier along with the optimal pre-processing technique was compared to identify the optimal model. An overview of all the top-performing models is given in Table 4.57 to Table 4.60. The models are arranged from best to worst performing.

Table 4.57 An overview of the various classification models for sound versus heat-damaged wheat.

Classification Algorithm	Pre-processing	Calibration		Validation	
		Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
Log. Reg.	P7	93.6	6.4	99.2	0.8
PLS-DA	P7	95.9	4.1	98.0	2.0
SVM	P7	97.4	2.6	97.3	2.7
RF	P5	95.2	4.8	92.0	8.0
LDA	P2	82.1	17.9	90.7	9.4
KNN	P3	95.6	4.4	79.3	20.7
D. Trees	NA	-	-	-	-

Table 4.58 An overview of the various classification models for sound versus sprout-damaged wheat.

Classification Algorithm	Pre-processing	Calibration		Validation	
		Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
SVM	P7	99.3	0.7	98.6	1.4
PLS-DA	P7	97.8	2.2	98.0	2.0
Log. Reg.	P7	99.6	0.5	96.6	3.4
RF	P6	97.3	2.7	92.5	7.5
KNN	P6	97.5	2.6	91.8	8.2
LDA	P1	98.2	1.8	78.2	21.8
D. Trees	NA	-	-	-	-

Table 4.59 An overview of the various classification models for sound versus *Fusarium*-damaged wheat.

Classification Algorithm	Pre-processing	Calibration		Validation	
		Classification Accuracy (%)	Misclassification Rate (%)	Classification Accuracy (%)	Misclassification Rate (%)
RF	P6	99.2	0.8	100.0	0.0
D. Trees	P6	98.6	1.4	100.0	0.0
KNN	P6	98.3	1.7	100.0	0.0
Log. Reg.	P7	98.1	1.9	99.4	0.6
SVM	P5	99.9	0.2	99.4	0.6
PLS-DA	P6	99.4	0.6	97.6	2.4
LDA	P7	88.8	11.2	96.5	3.5

Table 4.60 An overview of the various classification models for sound versus immature wheat.

Classification Algorithm	Pre-processing	Calibration		Validation	
		Classification Accuracy (%)	Misclassification Rate (%)	Classification Accuracy (%)	Misclassification Rate (%)
PLS-DA	P6	100.0	0.0	99.3	0.7
RF	P6	99.3	0.7	99.3	0.7
LDA	P2	99.1	0.9	99.2	0.8
Log. Reg.	P7	99.8	0.2	98.5	1.5
D. Trees	P2	98.2	1.8	96.3	3.7
KNN	P7	96.6	3.5	95.5	4.5
SVM	P5	100.0	0.0	95.5	4.5

Wavelength Selection

Since logistic regression performed relatively well for all dichotomous classifications, it was selected to demonstrate the effects of wavelength selection. Recursive Feature Elimination (RFE) was applied to the pre-processed spectra to identify an optimal number of wavelengths. The purpose of selecting a subset of wavelengths is to create a more straightforward, rapid and less expensive multispectral imaging system that is more suitable for on-line industry applications. A multispectral approach is also beneficial in that it addresses the multicollinearity problem since a smaller subset of wavelengths are less likely to be correlated. The wavelengths are selected in such a manner that there is little to no performance deterioration of the model. The results for reduced wavelengths are displayed in Table 4.61.

From Table 4.61, it is clear that using a subset of wavelengths did not have a drastic effect on the model's performance. For sprout-damaged wheat, using 50 wavelengths which is less than a quarter of the total number of wavelengths (total = 288) did not have any effect on the model's performance, and an accuracy of 96.6% was still attainable. Immature wheat maintained a classification accuracy of 97.8% throughout, which is only 0.7% less than the classification accuracy achieved using the full spectrum. For *Fusarium*-damaged, using as little as five wavelengths still produced a classification accuracy of 97.1%. Heat-damaged wheat did not perform as well with the reduced number of wavelengths. However that could also be as a result of the specific classifier. With a different classifier, these results could be improved. Nevertheless, a high classification accuracy of 92.7% was possible with as little as 20 wavelengths.

Table 4.61 The effect of using the full spectral range and a few select wavelengths on the performance of the Logistic Regression classifier on the validation set.

Defect	No. of Wavelengths	Wavelength Selection		Full Spectrum	
		Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
Heat	100	96.0	4.0	99.2	0.8
	50	92.7	7.3		
	20	92.7	7.3		
	10	86.7	13.3		
	5	84.7	15.3		
Sprout	100	96.6	3.4	96.6	3.4
	50	96.6	3.4		
	20	94.6	5.4		
	10	92.5	7.5		
	5	88.4	11.6		
Fusarium	100	98.8	1.2	99.4	0.6
	50	98.2	1.8		
	20	97.6	2.4		
	10	97.6	2.4		
	5	97.1	2.9		
Immature	100	97.8	2.2	98.5	1.5
	50	97.8	2.2		
	20	97.8	3.0		
	10	97.8	2.2		
	5	97.8	2.2		

To investigate the particular wavelengths that contribute to a high classification accuracy, the key wavelengths for each classifier were compared. The wavelengths identified for discerning sound and heat-damaged wheat correlate to starch and protein content (958-985 nm; 1225-1410 nm; 1933-2250 nm). This finding was unsurprising as wheat kernels that were exposed to heat sustain starch damage and protein denaturation. Wang *et al.* (2001) stated that during periods of excessive heat, wheat kernels might also have physical damage where stress-cracks develop in the endosperm and propagate outwards, exposing parts of the endosperm and aleurone layer. However, studies conducted with NIR spectroscopy to measure heat-damaged kernels are minimal, and there are no suitable studies to which these results can be compared.

For sound wheat versus sprout-damaged wheat, the important wavebands identified were 1209 nm to 1683 nm. These wavelengths correlate to starch. An association with starch is consistent with what is expected, as germinated kernels have an increase in enzymatic activity which results in a decrease in starch, increase in sugar, increase in total protein, change in amino acid composition, and dry matter loss (Singh *et al.*, 2009). Singh *et al*. (2009) found 1102 nm, 1132 nm and 1305 to be essential wavelengths for the classification of sprout-damaged wheat. These wavelengths were said to be associated with starch, emphasising that starch content is a key factor when discerning sound and sprout-damaged wheat. Since Singh *et al*. (2009) acquired the hyperspectral images in the range of 900-1700 nm, a direct comparison of wavelengths was not possible. Another study on the detection of sprout damage was done by (Barbedo *et al.*, 2018) and it was found that 918 nm and 1411 nm were the most relevant. The 1411 nm band is associated with moisture content. While all the kernels were conditioned to have similar moisture contents, the authors speculated that the spectral differences indicated by this band were likely due to chemical alterations caused by the sprouting. However, since the kernels in the present study were not conditioned, the wavelengths in the area of 2100 nm (attributed to starch content) were the most useful.

The essential wavelengths identified for *Fusarium*-damaged wheat were 1198 nm, 1203 nm, 1214 nm, 1219 nm, 1399 nm and 1443 nm. Research by Delwiche *et al.* (2011) indicated that ergosterol and chitin displayed absorption bands near 1200 nm and 1480 nm, respectively. Barbedo *et al.* (2015) found that wavelengths in the region of 1411 nm provided the best separation between sound and *Fusarium*-damaged kernels. Another study by Delwiche *et al.* (2019) found that the wavelengths 1197 nm (attributed to ergosterol) and 1394 nm (attributed to chitin) were the two best wavelengths in detecting *Fusarium*-damaged wheat kernels. In the present study, similar results were achieved. Wavelengths in the range of 1198 – 1219 nm were identified as key wavelengths and are attributed to ergosterol. The wavelengths 1399 nm and 1443 nm were also determined as key wavelengths, which are attributed to chitin. The other key wavelengths (996 nm, 980 nm, 1121 nm, 1132 nm and 1672 nm) are attributed to the macronutrients of wheat kernels since *Fusarium* damages the starch granules, storage proteins and cell wall, altering the chemical composition. This was also seen in the spectral signatures where the *Fusarium*-damaged wheat kernels tended to have lower absorbances than that of sound wheat (Figure 4.12).

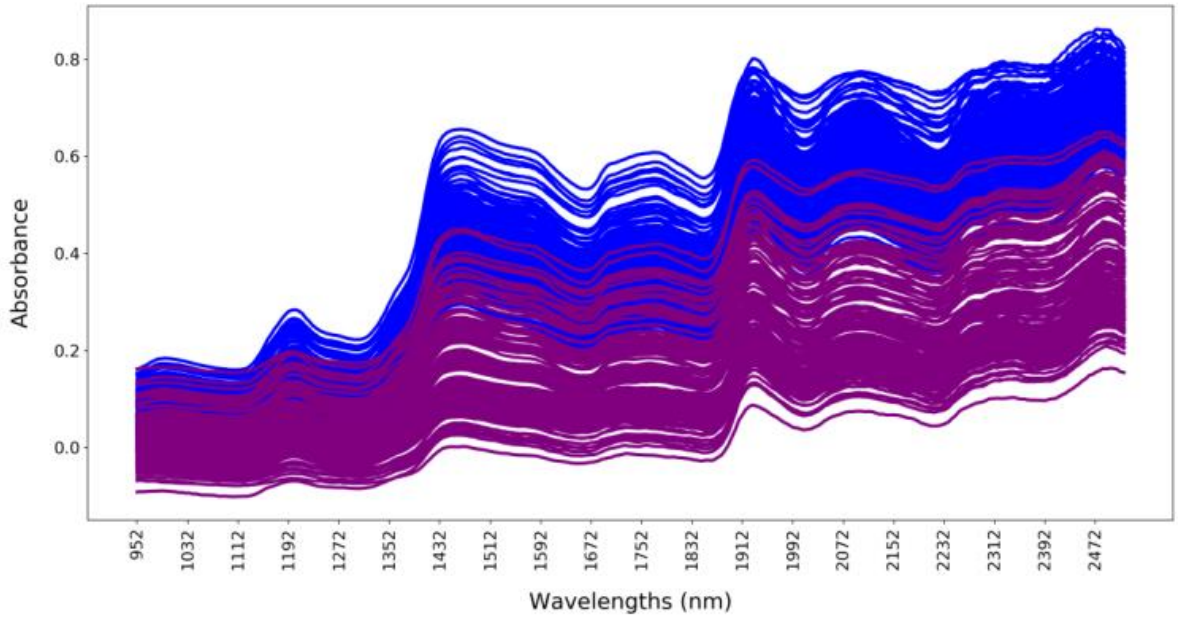


Figure 4.12 Unprocessed average pseudo-absorbance spectra for sound wheat (blue) and *Fusarium*-damaged wheat (purple).

For immature wheat, it was found that the wavelength region of 1323 – 1388 nm played a critical role in this classification. This region is likely attributed to starch. Immature kernels have an increased kernel hardness, and the amount of sugars present in immature grains is higher than that of mature wheat (Petrovska-avramenko *et al.*, 2016). Mahesh *et al.* (2015) predicted the hardness of bulk wheat samples and found the wavelength regions 1180 – 1220; 1320 – 1400; and 1460 – 1490 nm to be useful. They also stated that the 1320 – 1400 nm region, similar to the current study, is likely attributed to starch. In a similar study on maize kernel hardness, McGoverin and Manley (2012) identified 1367 nm as a key wavelength to classify maize kernel hardness. Absorbance in this region is ascribed to C–H vibrations (C–H stretch and C–H deformation), associated with starch.

Visualisation of Predicted Results

Visual prediction maps on a validation set for each dichotomous classification is displayed Figure 4.13. These results demonstrate that HSI can be used to visualise the prediction. This can be implemented in a graphical user interface (GUI) to create an interactive and user-friendly tool for cereal graders.

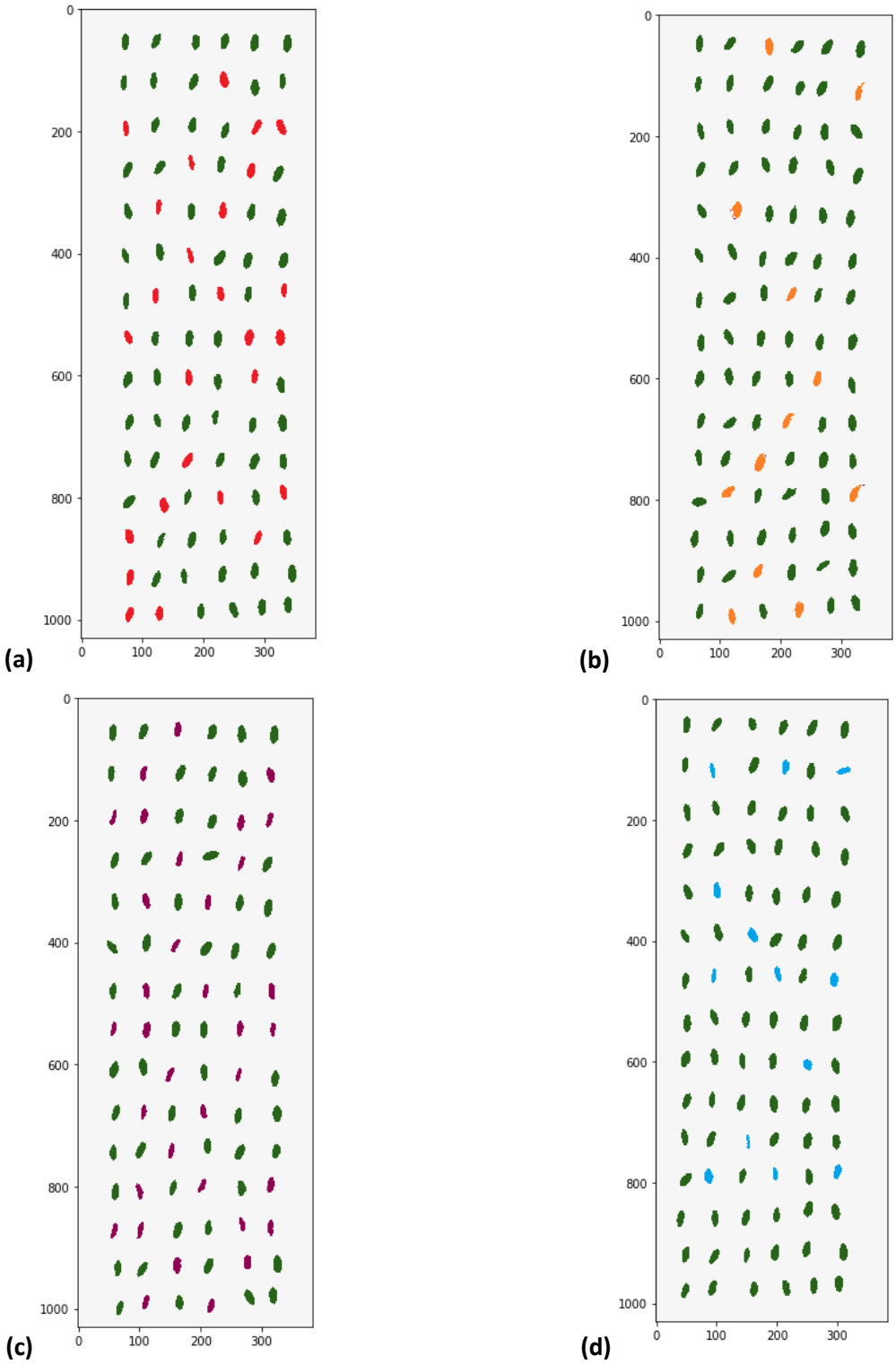


Figure 4.13 Prediction maps of sound and **(a)** heat-damaged; **(b)** sprout-damaged; **(c)** *Fusarium*-damaged and **(d)** immature wheat.

Multiclass Classification

The results obtained using multiclass classification algorithms on pre-processed NIR spectra is displayed in Table 4.62. The calibration and validation set consisted of 255 and 247 kernels respectively.

Table 4.62 The classification results for the multiclass analyses.

Model	Pre-processing	Hyperparameters	Calibration		Validation	
			Classification Accuracy (%)	Error Rate (%)	Classification Accuracy (%)	Error Rate (%)
Logistic Regression	P2	-	87.5	14.7	77.7	22.3
	P3	-	86.3	13.7	76.1	23.9
	P7	-	89.4	10.6	80.2	19.8
PLS-DA	P2	LV:9	86.7	13.3	76.1	23.9
	P3	LV:9	86.7	13.3	72.5	27.5
	P5	LV:9	85.8	14.2	78.5	21.5
	P6	LV:9	87.5	12.5	78.5	21.5
	P7	LV:9	86.3	13.7	76.9	23.1
LDA	P2	LD:1	92.1	7.9	66.4	33.6
	P3	LD:1	91.4	8.6	66.8	33.2
	P5	LD:1	90.6	9.4	68.8	31.2
	P6	LD:1	93.3	6.7	83.0	17.0
	P7	LD:1	92.9	7.1	73.3	26.7
RF	P2	Trees: 100	88.6	11.4	71.3	28.7
	P3	Trees: 100	87.1	12.9	70.4	29.6
	P5	Trees: 100	86.7	13.3	75.7	24.3
	P6	Trees: 100	89.8	10.2	78.1	21.9
	P7	Trees: 100	85.1	14.9	73.7	26.3
SVM	P2	C: 100; γ: 0.1	97.6	2.4	79.4	20.6
	P3	C: 100; γ: 0.1	97.6	2.4	79.8	20.2
	P5	C: 1000; γ: 1	89.8	10.2	84.6	15.4
	P6	C: 1000; γ: 1	76.1	23.9	68.8	31.2
	P7	C: 1000; γ: 0.001	92.5	7.5	81.0	19.0

Logistic Regression

The pre-processing methods P4, P5 and P6 were not included as they had proven to be ill-fitted for classification analyses using the logistic regression classifier. The best results were achieved by pre-processing the data with 2nd derivatives followed by SNV (P7). A classification accuracy of 89.4% was achieved for the calibration set, and 80.2% for the validation set. Inspecting the confusion matrices revealed more

detailed information on the predictions. Figure 4.14a displays the outcome of the calibration set. Out of the 44 *Fusarium*-damaged kernels, only 36 were correctly predicted as such. No false negatives were predicted as all 86 sound kernels were correctly classified. The defective categories had fairly similar percentages of kernels misclassified. In Figure 4.14b it can be seen that the validation set really challenged the model with heat-damaged and sprouted samples. Of the 29 sprout-damaged kernels, more than half were misclassified as *Fusarium*-damaged and five, two and three kernels were misclassified as heat, immature and sound respectively. The sensitivity of the model was satisfactory as only two sound kernels were incorrectly classified. These kernels were predicted as sprout-damaged.

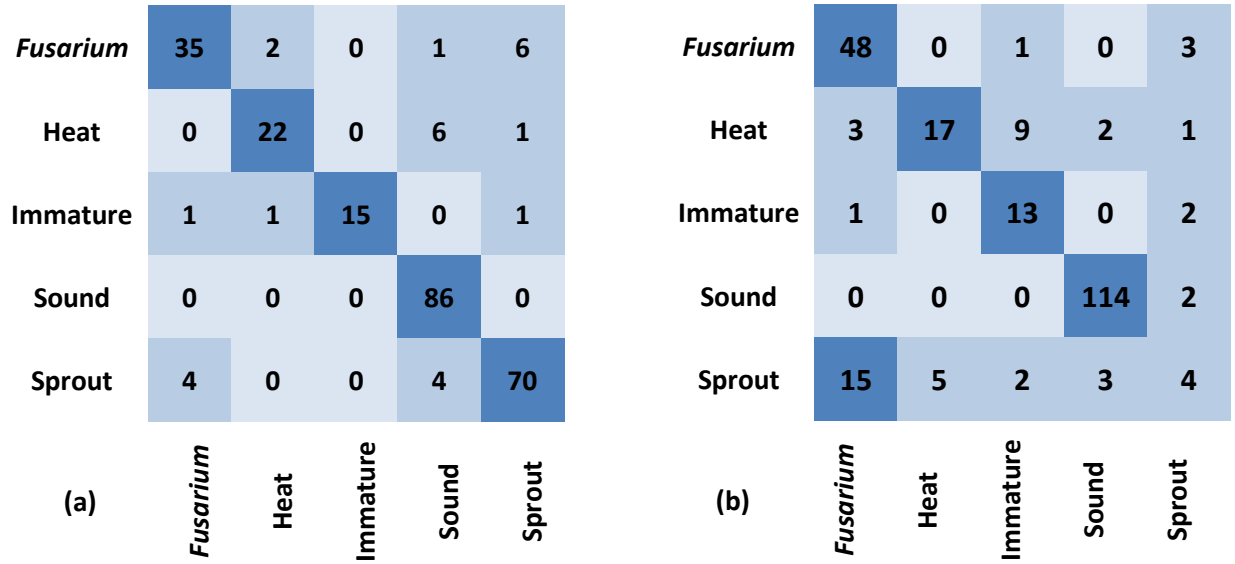


Figure 4.14 A confusion matrix of the Logistic Regression model on the **(a)** calibration set; pre-processed with 2nd derivatives followed by SNV (P7); **(b)** validation set; pre-processed with 2nd derivatives followed by SNV (P7). The categories are in alphabetical order.

Partial Least Squares Discriminant Analysis

From the classification accuracies, it appears that the P5 and P6 pre-processing methods performed well and achieved the highest performance. For P5, a classification accuracy of 85.8% was achieved for the calibration set, and 78.5% for the validation set. For P6, a classification accuracy of 87.5% was achieved for the calibration set, and 78.5% for the validation set. The PLS-DA classifier however performed poorly overall in the sense that it could not predict heat- and sprout-damaged kernels accurately in the validation set. Figure 4.15a shows the outcome of the P5 pre-processing method on the validation set. Out of the 32 heat-damaged kernels, none were correctly predicted. The majority of the heat-damaged kernels, 24, were predicted as *Fusarium*-damaged. Likewise, almost half of the sprout-damaged kernels were incorrectly classified. The model was, however, able to classify sound, immature and *Fusarium*-damaged wheat with minimal error.

Figure 4.15b shows the outcome of the P6 pre-processing method on the validation set. This approach performed worse than the P5 pre-processing method in terms of predicting sprout-damaged kernels. Twenty of the sprout-damaged kernels were predicted as *Fusarium*-damaged. Of the 32 heat-damaged kernels, 18 were incorrectly predicted as *Fusarium*-damaged. The PLS-DA classifier was therefore not effective in the multiclass classification despite the high performances in the dichotomous classifications.

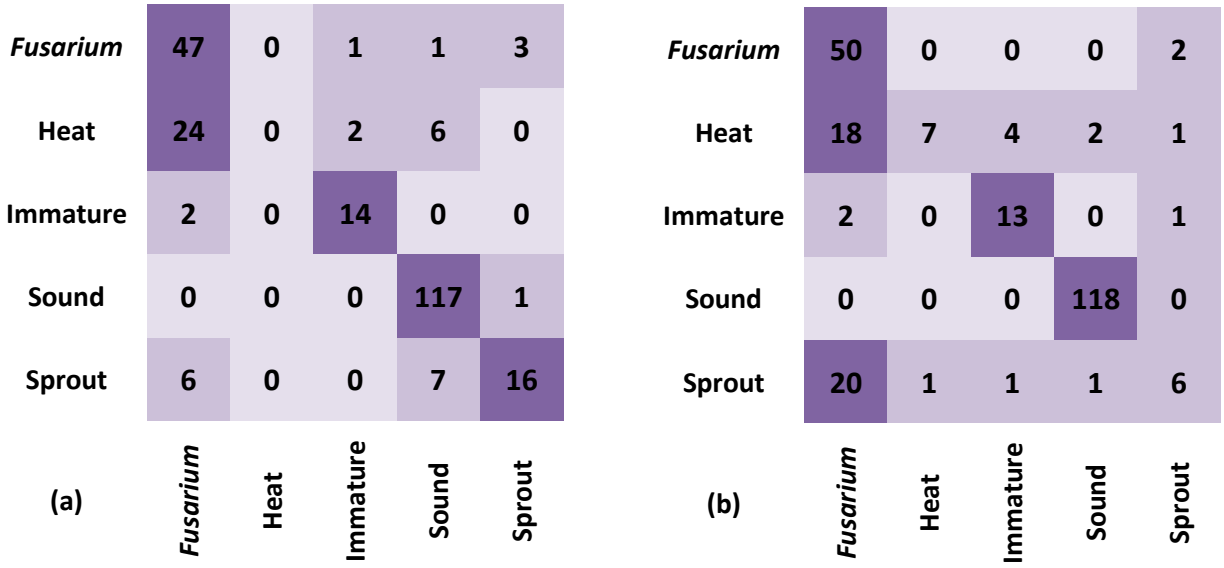


Figure 4.15 A confusion matrix of the PLS-DA model on the validation set containing all five categories. **(a)** Pre-processed with SNV, de-trending and 2nd derivatives (P5); **(b)** pre-processed 2nd derivatives (P6); the categories are in alphabetical order.

Linear Discriminant Analysis

For the LDA classifier, the data pre-processed with 2nd derivatives (P6) provided the highest results. A classification accuracy of 85.8% was achieved for the calibration set, and 78.5% for the validation set. The LDA classifier also had difficulty accurately predicting heat- and sprout-damaged kernels in the validation set. More than half of both these categories were misclassified. Once again, sound, immature and *Fusarium*-damaged kernels were predicted more accurately with a marginal error. The confusion matrix for the validation set is displayed in Figure 4.16.

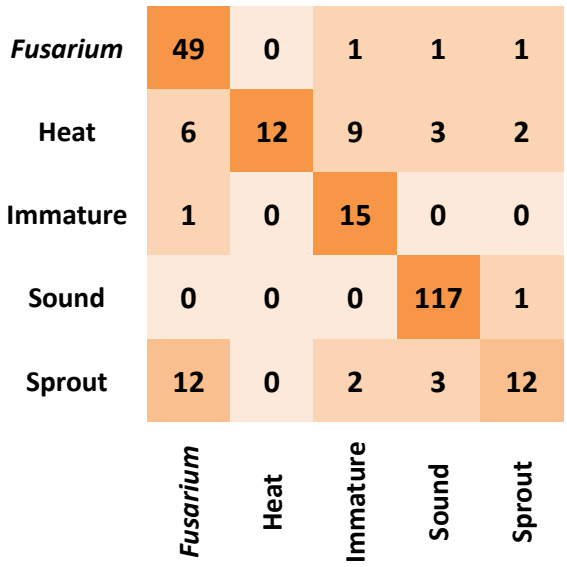


Figure 4.16 A confusion matrix of the LDA model on the validation set containing all five categories; pre-processed with 2nd derivatives (P6); the categories are in alphabetical order.

Random Forests

The RF classifier achieved the highest classification accuracy when performed on 2nd derivatives (P6) pre-processed data. The confusion matrix is displayed in Figure 4.17.

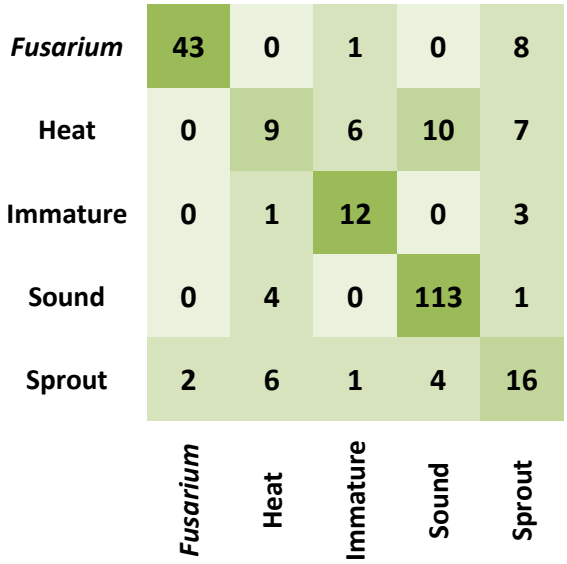


Figure 4.17 A confusion matrix of the RF model on the validation set containing all five categories; pre-processed with 2nd derivatives (P6); the categories are in alphabetical order.

A classification accuracy of 89.8% was achieved for the calibration set, and 78.1% for the validation set. The model performed poorly with regards to heat- and sprout-damage. The model also had an inferior overall performance compared to the previously discussed classifiers. Of the 32 heat-damaged kernels, 23 were

misclassified as either sound, sprout or immature. For sprout-damage, 13 of the 29 kernels were misclassified.

Support Vector Machines

The SVM classifier performed the best out of all the classification algorithms. This was expected as SVM makes use of kernels to better separate data that is linearly non-separable. Nevertheless, the classifier still had difficulty accurately predicting heat-damaged kernels. The prediction accuracy of sprout-damaged kernels was better than the previous models discussed; however, the prediction accuracy of *Fusarium*-damage was inferior. The pre-processing method that provided the highest performance was SNV, de-trending and 2nd derivatives (P5). A classification accuracy of 89.9% was achieved for the calibration set, and 84.6% for the validation set. This model had the hyperparameters defined as C = 1000 and gamma = 1. The confusion matrix for the SVM classifier on the validation set is displayed in Figure 4.18.

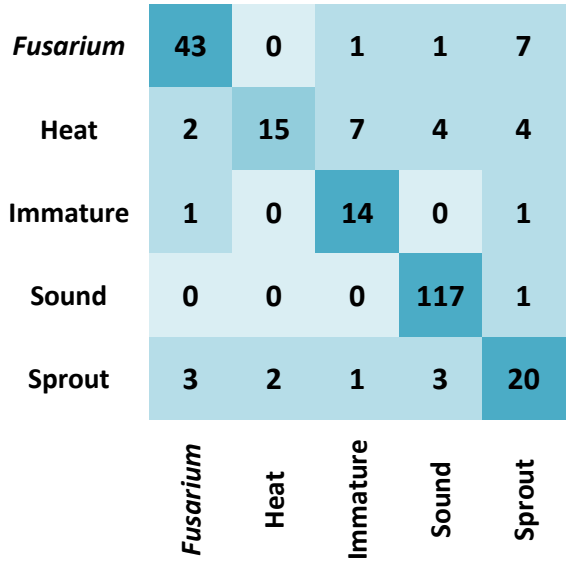


Figure 4.18 A confusion matrix of the SVM model on the validation set containing all five categories; pre-processed with SNV, de-trending and 2nd derivatives (P5); the categories are in alphabetical order.

Conclusion

Overall, the dichotomous classifications performed superiorly to the multiclass classifications. The results from the dichotomous classifications are satisfactory and prove that the discrimination of sound and defective wheat is possible. For the classification between sound and heat damaged wheat, the logistic regression classifier, applied to 2nd derivatives followed by SNV pre-processed spectra, provided the highest classification accuracy (99.2%). For sprouting analysis, SVM applied to 2nd derivatives followed by SNV pre-processed spectra achieved the highest classification accuracy (98.6%). *Fusarium*-damaged wheat scored 100% classification accuracy with the RF, decision trees or KNN classifier applied to the spectra pre-processed with 2nd derivatives. For immature wheat, the best classification was achieved with 2nd derivatives and PLS-

DA, where 99.3% accuracy was achieved. The models generally had a higher sensitivity than specificity, indicating that the models were more effective at recognising true positives than true negatives.

The multiclass classifications did not perform as well, and mainly had difficulty with heat- and sprout-damaged wheat. Due to the similarities in spectral variation, displayed in the PCA score plots, it is understandable that more errors would occur with these categories. The highest classification accuracy was achieved by pre-processing with SNV, de-trending and 2nd derivatives, and classifying with SVM (C: 1000; γ: 1). Alternatively, multiclass classification could be performed using the optimal dichotomous classifications in a hierarchical model, since they had a better performance.

References

Barbedo, J.G.A., Guarienti, E.M. & Tibola, C.S. (2018). Detection of sprout damage in wheat kernels using NIR hyperspectral imaging. *Biosystems Engineering*, **175**, 124-132.

Barbedo, J.G.A., Tibola, C.S. & Fernandes, J.M.C. (2015). Detecting *Fusarium* head blight in wheat kernels using hyperspectral imaging. *Biosystems Engineering*, **131**, 65-76.

Barnes, R.J., Dhanoa, M.S. & Lister, S.J. (1989). Standard Normal Variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, **43**, 772-777.

Buddenbaum, H. & Steffens, M. (2012). The effects of spectral pretreatments on chemometric analyses of soil profiles using laboratory imaging spectroscopy. *Applied and Environmental Soil Science*, **2012**, 1-12.

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. *The Analyst*, **131**, 1152-1160.

Delwiche, S.R. (1996). Classification of wheat by visible and near-infrared reflectance from single kernels. *Cereal Chemistry*, **73**, 399-405.

Delwiche, S.R., Kim, M.S. & Dong, Y. (2011). *Fusarium* damage assessment in wheat kernels by Vis/NIR hyperspectral imaging. *Sensing and Instrumentation for Food Quality and Safety*, **5**, 63-71.

Delwiche, S.R., Rodriguez, I.T., Rausch, S.R. & Graybosch, R.A. (2019). Estimating percentages of fusarium-damaged kernels in hard wheat by near-infrared hyperspectral imaging. *Journal of Cereal Science*, **87**, 18-24.

Esteve Agelet, L., Ellis, D.D., Duvick, S., Goggi, A.S., Hurburgh, C.R. & Gardner, C.A. (2012). Feasibility of near infrared spectroscopy for analyzing corn kernel damage and viability of soybean and corn kernels. *Journal of Cereal Science*, **55**, 160-165.

Fernández-Ibañez, V., Soldado, A., Martínez-Fernández, A. & de la Roza-Delgado, B. (2009). Application of near infrared spectroscopy for rapid detection of aflatoxin B1 in maize and barley as analytical quality assessment. *Food Chemistry*, **113**, 629-634.

Joe, A.A.F. & Gopal, A. (Year). Identification of spectral regions of the key components in the near infrared spectrum of wheat grain. In: *2017 International Conference on Circuit ,Power and Computing Technologies (ICCPCT)*. Pp. 1-5.

Lim, J., Kim, G., Mo, C., Oh, K., Kim, G., Ham, H., Kim, S. & Kim, M. (2018). Application of Near Infrared Reflectance Spectroscopy for Rapid and Non-Destructive Discrimination of Hulled Barley, Naked Barley, and Wheat Contaminated with Fusarium. *Sensors*, **18**, 113.

Mahesh, S., Jayas, D.S., Paliwal, J. & White, N.D.G. (2015). Comparison of Partial Least Squares Regression (PLSR) and Principal Components Regression (PCR) methods for protein and hardness predictions using the near-infrared (NIR) hyperspectral images of bulk samples of Canadian wheat. *Food and Bioprocess Technology*, **8**, 31-40.

Manley, M., Williams, P., Nilsson, D. & Geladi, P. (2009). Near infrared hyperspectral imaging for the evaluation of endosperm texture in whole yellow maize (Zea maize L.) kernels. *Journal of Agricultural and Food Chemistry*, **57**, 8761-8769.

McGoverin, C. & Manley, M. (2012). Classification of maize kernel hardness using near infrared hyperspectral imaging. *Journal of Near Infrared Spectroscopy*, **20**, 529.

Murray, I. & Williams, P.C. (1987). Chemical principles of near-infrared technology. In: *Near-infrared Technology in the Agricultural and Food Industries* (edited by P.C. Williams & K.H. Norris). Pp. 17-34. St. Paul, Minnesota: American Association of Cereal Chemists Inc.

Rinnan, Å., Berg, F.v.d. & Engelsen, S.B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *Trends in Analytical Chemistry*, **28**, 1201-1222.

Siedliska, A., Baranowski, P., Zubik, M., Mazurek, W. & Sosnowska, B. (2018). Detection of fungal infections in strawberry fruit by VNIR/SWIR hyperspectral imaging. *Postharvest Biology and Technology*, **139**, 115-126.

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.

Wang, D., Dowell, F. & Lacey, R. (1999). Single wheat kernel color classification using neural networks. *Transactions of the ASAE*, **42**, 233-240.

Wang, D., Dowell, F.E. & Chung, D.S. (2001). Assessment of heat-damaged wheat kernels using near-infrared spectroscopy. *Cereal Chemistry*, **78**, 625-628.

CHAPTER 5
GENERAL DISCUSSION AND CONCLUSION

In this study, it was possible to classify wheat kernels using hyperspectral imaging and chemometrics. Dichotomous and multiclass evaluations were executed on five categories: sound, heat-damaged, *Fusarium*-damaged, sprout-damaged, and immature wheat kernels. Of the four defective categories, heat and sprout damage were the most challenging to correctly predict while *Fusarium*-damaged and immature wheat were the simplest. Overall, the dichotomous classifications performed well, where prediction accuracies above 98% were achieved, given that the optimal pre-processing technique and classifier were used. When all the categories were combined for multiclass analyses, the overall performance decreased. The performance of the multiclass models can, however, be improved by increasing the sample size. Because the samples used in this study varied in terms of variety, growing region and growing season, a larger, broader sample size is necessary to incorporate all these factors. A larger sample size will better equip the model to make correct predictions, as it will provide more accurate mean values, and minimise the effect of outliers. Additional features can also be incorporated into the model, specifically wavelengths from the visible region, spatial features and textural elements. Immature, *Fusarium*- and heat-damaged kernels differ in colour, where immature grains are often green, *Fusarium*- damaged kernels have a chalky colour and heat-damaged kernels are often slightly darker than sound kernels. These distinguishing colours can help minimise incorrect predictions in multiclass classifications, especially since the spectral responses in the NIR region are very similar. Furthermore, sprout- and *Fusarium*-damaged kernels are physically distinguishable from sound kernels. Incorporating spatial features such as the kernel size and shape will help strengthen the classification. Image textural elements will also benefit the model, as defective kernels generally have a wrinkly surface area.

The most significant factor that affected the performance of the models was pre-processing. Investigating multiple pre-treatments and combinations thereof revealed that pre-processing is vital, and the selection of an appropriate pre-treatment must be prioritised. Incorrect or insufficient pre-processing leads to inaccurate results and misinterpretation. For instance, 2nd derivatives provided poor results with the logistic regression classifier. With this pre-processing method alone, it could have been interpreted that the sound and defective categories were inseparable, when in fact logistic regression delivered excellent predictions when 2nd derivatives was followed by SNV. The sequence in which the pre-treatments were applied was also crucial. The implementation of de-tending before SNV is an example of an incorrect application. The effects of this pre-processing combination was observed in the dichotomous classifications where the models performed well for the calibration set, but failed with the validation set. Overall, the optimal pre-processing techniques largely depended on the categories analysed and the classifier used. Therefore, pre-processing methods were not one-size-fits-all. Moreover, the optimal pre-processing technique could only be established after external validation. Due to the calibration and validation data

being captured on different days and at different times, it was speculated that the difference in environmental conditions during imaging affected the spectra. However, if the training set was comprised of data captured at random, non-specified times, the effect of the environmental imaging conditions would have been factored in during cross-validation. The optimal pre-processing technique was majorly 2nd derivatives or 2nd derivatives followed by SNV.

The different classifiers selected for this study predominantly performed well. For the dichotomous classification between heat-damaged and sound wheat, logistic regression, PLS-DA and SVM provided the best results, in that order. For sprout-damaged wheat, SVM was the top-performing classifier, followed by PLS-DA and logistic regression. Three classifiers for *Fusarium*-damage wheat achieved 100% accuracy, namely RF, Decision Trees and KNN. For immature wheat, all the models achieved an accuracy above 95%, where the top three performing classifiers were PLS-DA, RF and LDA. There are many other classification algorithms available. Soft independent modelling of class analogy (SIMCA) is another classification algorithm commonly used in the chemometrics literature, however since poor separation was observed in the PCA score plots due to spectral similarities, SIMCA was not considered in this study. Nevertheless, further research can be done to explore the efficacy of SIMCA for this classification.

For the multiclass analyses, logistic regression, LDA and SVM were the only classifiers that achieved above 80% for the validation set. To obtain a better multiclass analysis, a multilevel hierarchical model could be designed to accommodate the varying optimal pre-processing techniques and classifiers for each wheat defect. The optimal models for the dichotomous categories can be constructed in a hierarchical model where the models are organised into a multilevel structure. This enables multiple categories to be classified simultaneously without needing to use separate individual models for each classification. In future, to allow the results to be more comparable, each category should contain the same amount of kernels. The information obtained from confusion matrices would have been more informative if the number of kernels in each category were constant. However, it is understandable that this is not always possible due to sample availability.

Since hyperspectral imaging produces redundant information, variable selection is an excellent approach to create a multispectral system. A multispectral system will facilitate faster computational times and allow for accurate, parsimonious models (Calvini *et al.*, 2017). Multispectral systems are also a much cheaper alternative to hyperspectral imaging systems and are ideal for real-time application purposes. Waveband selection also addresses the problems associated with multicollinearity. Since the goal of regression analyses is to determine the relationship between each independent variable and the target variables, highly correlated variables lead to a biased estimation. By selecting a subset of wavelengths, the chances of the wavelengths being correlated are less likely. For sound wheat versus sprout-damaged wheat, the important wavelength regions identified were 2070 nm to 2146 nm. These wavelengths are highly correlated to starch which is consistent with what is seen in the literature (Singh *et al.*, 2009). The most important wavelengths identified for sound versus *Fusarium*-damaged wheat were 1198 nm, 1203 nm, 1214

nm, 1219 nm, 1399 nm and 1443 nm. These wavelengths are attributed to ergosterol and chitin which display absorption bands near 1200 nm and 1480 nm respectively (Delwiche *et al.*, 2011). For immature wheat, it was found that the wavelength region of 1323 – 1388 nm played a critical role which corresponds to results discovered by Mahesh *et al.* (2015). This region is attributed to starch which is valid as immature kernels have an increased kernel hardness.

Food quality and safety are of broad current interest. Consumers are becoming increasingly more conscious about food quality and safety, and there is little tolerance for low-quality products in the market. In addition to quality, the wheat supply and demand ratio is also a problem which is not only confined to the South African market. The human population is rapidly increasing and with this, so is the demand for faster grading and inspection processes. Currently, traditional manual methods are still being used to assess grain quality. Human visual inspection is still the most common practice, which is proven to be time-consuming, tedious and subjective. This study proved that NIR hyperspectral imaging is a feasible alternative to provide the wheat industry with a superior quality control system in the near future. The entire NIR spectral range need not be utilised as models based on a few key wavelengths also provided high classification results.

To ensure that the results obtained in this study were commercially relevant, the conditions of the experimental procedure were designed to imitate the conditions of current wheat grading practices. This meant that no kernel conditioning was performed and the kernels were imaged in the same format (as is) as an industry professional would grade. The kernels were also imaged in both crease-up and crease-down orientations as it would be impractical for graders to position the grains before imaging. Overall, the perfect classification method did not exist, however it was possible to classify the wheat with relatively high accuracies, despite the different wheat categories having many spectral similarities. This work should be furthered by extending the defective categories to include all parameters addressed in the South African grading regulations. The grading criteria regarding protein content, falling number and hectolitre mass could also be incorporated to create one system capable of grading wheat in its entirety.

NIR hyperspectral imaging combined with chemometrics offers an accurate, non-destructive and reliable technique for the rapid identification of defective wheat kernels. This study provides the South African cereal industry with an alternative technology to assess the quality and health status of cereal grains and contributes to improving food quality and safety monitoring practices both locally and internationally.

References

Calvini, R., Amigo, J.M. & Ulrici, A. (2017). Transferring results from NIR-hyperspectral to NIR-multispectral imaging systems: A filter-based simulation applied to the classification of Arabica and Robusta green coffee. *Analytica Chimica Acta*, **967**, 33-41.

Delwiche, S.R., Kim, M.S. & Dong, Y. (2011). *Fusarium* damage assessment in wheat kernels by Vis/NIR hyperspectral imaging. *Sensing and Instrumentation for Food Quality and Safety*, **5**, 63-71.

Mahesh, S., Jayas, D.S., Paliwal, J. & White, N.D.G. (2015). Comparison of Partial Least Squares Regression (PLSR) and Principal Components Regression (PCR) methods for protein and hardness predictions using the near-infrared (NIR) hyperspectral images of bulk samples of Canadian wheat. *Food and Bioprocess Technology*, **8**, 31-40.

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.

ADDENDUM 1

SOUTH AFRICAN WHEAT GRADING REGULATIONS

REGULATIONS RELATING TO THE GRADING, PACKING AND MARKING OF BREAD WHEAT INTENDED FOR SALE IN THE REPUBLIC OF SOUTH AFRICA

DEPARTMENT OF AGRICULTURE, FORESTRY AND FISHERIES

NO. R. 64 **29 JANUARY 2016**

AGRICULTURAL PRODUCT STANDARDS ACT, 1990
(ACT No.119 OF 1990)

REGULATIONS RELATING TO THE GRADING, PACKING AND MARKING OF BREAD WHEAT INTENDED FOR SALE IN THE REPUBLIC OF SOUTH AFRICA

The Minister of Agriculture, Forestry and Fisheries, acting under section 15 of the Agricultural Product Standards Act 119 of 1990, has

(a) made the regulations in the Schedule;

(b) determined that the said regulations shall come into operations on the date of publication; and

(c) read together with section 3(1) of the said Act, repealed the Regulations published by Government Notice No. R1186 of December 2010.

SCHEDULE

Definitions

1. Unless the context otherwise indicates, any word or expression in these regulations to which a meaning has been assigned in the Act shall have that meaning, and--

"**animal filth**" means dead rodents, dead birds and dung;

"bag" means a bag manufactured from --

(a) jute or phormium or a mixture of jute and phormium; or

(b) polypropylene that complies with SANS specification CKS632 1246:2012;

"bulk container" means any vehicle or container in which bulk wheat is stored or transported;

"**consignment**" means --

(a) a quantity of wheat of the same class, which belongs to the same owner, delivered at any one time under cover of the same consignment note, delivery note or receipt note, or delivered by the same vehicle or bulk container, or loaded from the same bin of a grain elevator or from a ship's hold; or

(b) in the case where a quantity referred to in paragraph (a), is subdivided into different grades, each such quantity of each of the different grades.

"container" means a bag or bulk container;

"**damaged wheat**" means wheat--

(a) which have been damaged by insects;

(b) which have been distinctly discoloured (orange-brown, dark brown or black) by external heat or as a result of heating caused by internal fermentation in wheat with an excessive moisture content, excluding wheat kernels in respect of which the discolouration is confined to the germ end;

(c) which are immature and have a distinctly green colour; and

(d) in which germination has proceeded to such an extent that the skin covering the embryo has been broken or the developing sprouts and/or rootlets are clearly visible.

"**ergot sclerotia**" means the sclerotia of the fungus *Claviceps purpurea*; and "ergot" has a corresponding meaning;

"**falling number**" means the time in seconds according to Hagberg-Perten as a measure of the degree of Alpha-Amylase activity in grain and flour;

"**field fungi-infected wheat**" means wheat of which the kernels are visibly infected with fungi, and that--

(a) clearly have greyish brush-ends that are discoloured as a whole; or where field fungi growth is present from the brush-ends into the crease; and

(b) have a dull, lifeless, chalky or pinkish and shrunken appearance as a result of *Fusarium* infection.

"**foreign matter**" means all matter excluding wheat, other grain and unthreshed ears. Coal, dung, glass and metal shall not be present in the consignment concerned;

"**heavily frost-damaged wheat**" means --

(a) wheat which have been damaged by severe frost during the milk to soft dough stage and which is characterised by the kernels being fairly plump, but covered entirely with small blisters extending into the crease, excluding --

(i) kernels in which blistering is confined to the back of the kernel; and

(ii) immature wrinkled kernels in which wrinkling has been caused by frost while the kernels were still immature; and

(b) kernels which have a slightly flaked-off bran coat due to frost: Provided that evidence of frost damage is present and that the bran coat had not been rubbed off due to handling.

"**hectolitre mass**" means the mass in kilogram per hectolitre;

"**insect**" means any live grain insect that is injurious to stored grain irrespective of the stage of development of that insect;

"**other grain**" means the kernels or pieces of kernels of barley, oats, triticale, maize, rye and sorghum;

"**poisonous seeds**" means the seeds or bits of seeds of plant species that may in terms of the Foodstuffs, Cosmetics and Disinfectants Act 54 of 1972 represent a hazard to human or animal health when consumed, including seeds of *Argemone mexicana, Convolvulus spp.*, *Crotalaria spp., Datura spp., Ipomoea purpurea, Lolium temulentum, Ricinus communis* or *Xanthium spp.*;

"**protein content**" means the percentage protein in wheat on a 12% moisture basis;

"**screenings**" means all material that passes through the standard sieve;

"standard sieve" means a slotted sieve --

(a) with a flat bottom of metal sheet of 1,0 mm thickness with apertures 12,7 mm long and 1,8 mm wide with rounded ends. The spacing between the slots in the same row must be 2, 43 mm wide and the spacing between the rows of slots must be 2,0 mm wide. The slots must be alternately orientated with a slot always opposite the solid inter segment of the next row of slots;

(b) of which the upper surface of the sieve is smooth;

(c) with a round frame of suitable material with an inner diameter of between 300 mm and 310 mm maximum and at least 50 mm high; and

(d) that fits onto a tray with a solid bottom and must be at least 20 mm above the bottom of the tray.

"**stinking smut infection**" means wheat that is infected with *Tilletia spp.* with the exception of wheat infected with *Tilletia indica* (karnal bunt). Wheat is considered to be infected by stinking smut infected if one or more of the following characteristics are present--

(a) an unmistakable stinking smut odour; or

(b) wheat kernels that are smeared with stinking smut; or

(c) more than four stinking smut balls (or pieces of balls equal to four stinking smut balls) per 100 g of wheat.

"**storage fungi infected wheat**" means wheat that are visibly infected with fungi, and that show --

(a) blue, green, blackish or yellow fungal growth anywhere on the kernel; or

(b) visible mould beneath the bran.

"**the Act**" means the Agricultural Product Standards Act 119 of 1990;

"**unthreshed ears**" means ears and bits of ears of wheat, barley, triticale and rye that still contain seeds that are completely covered with glumes; and

"**wheat**" means the kernels and pieces of kernels of the species *Triticum aestivum*.

Restrictions on sale of wheat

2. (1) No person shall sell a consignment of wheat in the Republic of South Africa --

(a) unless the wheat is sold according to the classes set out in regulation 3;

(b) unless the wheat complies with the standards for the classes set out in regulation 4;

(c) unless the wheat, where applicable, complies with the grades of wheat and the standards for grades set out in regulations 5 and 6 respectively;

(d) unless the wheat is packed in accordance with the packing requirements set out in regulation 7;

(e) unless the containers or sale documents, as the case may be, are marked in accordance with the marking requirements set out in regulation 8; and

(f) if such wheat contains a substance that renders it unfit for human consumption or for processing into or utilisation thereof as food or feed.

(2) The Executive Officer may grant written exemption, entirely or partially, to any person on such conditions as he or she may deem necessary, from the provisions of sub-regulation (1).

PART I

QUALITY STANDARDS

Classes of wheat

3. The classes of wheat are --

(a) Bread Wheat; and

(b) Other Wheat.

Standards for classes

4. (1) Notwithstanding the provisions of sub-regulations (2) and (3), a consignment of wheat shall --

(a) be free from any toxin, chemical or any other substance that renders it unsuitable for human consumption or for processing into or utilisation thereof as food or feed and may not exceed the permissible deviations regarding aflatoxin in terms of the Foodstuffs, Cosmetics and Disinfectants Act 54 of 1972;

(b) not contain more poisonous seeds or ergot sclerotia than permitted in terms of the Foodstuffs, Cosmetics and Disinfectants Act 54 of 1972;

(c) be free from organisms of phytosanitary importance as determined in terms of the Agricultural Pest Act 36 of 1983;

(d) be free from mould infected, sour and rancid other grain and foreign matter;

(e) be free from any undesired odour, taste or colour not typical of undamaged and sound wheat;

(f) be free from animal filth;

(g) be free from stones, glass, metal, coal or dung;

(h) with the exception of Class Other Wheat, be free from grain insects;

(i) with the exception of Class Other Wheat, be free from stinking smut infection; and

(j) with the exception of Class Other Wheat, have a moisture content not exceeding 13 percent.

(2) A consignment shall be classified as Bread Wheat if --

(a) the wheat in the consignment consists of at least 95 percent (m/m) of one or more of the bread wheat seeds; and

(b) it complies with the standards for Grade 1, Grade 2, Grade 3, Grade 4 or Utility Grade set out in regulation 6.

(3) A consignment of wheat shall be classified as Class Other Wheat if it does not comply with the standards for Bread Wheat.

Grades of wheat

5. (1) The grades for Bread Wheat shall be as follows:

(a) Grade 1.

(b) Grade 2.

(c) Grade 3.

(d) Grade 4; and

(e) Utility grade.

(2) No grades are determined for Class Other Wheat.

Standards for grades of wheat

6. (1) Subject to the provisions of subregulations (2), (3) and (4), a consignment of wheat shall be graded as --

(a) Grade 1 if the nature of deviation, specified in column 1 of Table 1 of the Annexure, in that consignment does not exceed the percentage specified in column 2 of the said table opposite the deviation concerned;

(b) Grade 2 if the nature of deviation, specified in column 1 of Table 1 of the Annexure, in that consignment does not exceed the percentage specified in column 3 of the said table opposite the deviation concerned;

(c) Grade 3 if the nature of deviation, specified in column 1 of Table 1 of the Annexure, in that consignment does not exceed the percentage specified in column 4 of the said table opposite the deviation concerned;

(d) Grade 4 if the nature of deviation, specified in column 1 of Table 1 of the Annexure, in that consignment does not exceed the percentage specified in column 5 of the said table opposite the deviation concerned; and

(e) Utility Grade if the nature of deviation, specified in column 1 of Table 1 of the Annexure, in that consignment does not exceed the percentage specified in column 6 of the said table opposite the deviation concerned.

(2) The minimum hectolitre masses for the different grades are as follows:

(a) Grade 1 - 77 kg.

(b) Grade 2 - 76 kg.

(c) Grade 3 - 74 kg.

(d) Grade 4 - 72 kg; and

(e) Utility Grade - 70 kg.

(3) (a) Grade 1, Grade 2 and Grade 3 shall have a minimum falling number value of not less than 250 seconds.

(b) Grade 4 shall have a minimum falling number value of not less than 200 seconds.

(c) Utility Grade shall have a minimum falling number value of not less than 150 seconds.

(d) Notwithstanding the provision of paragraph (a), wheat shall be deemed to comply with the requirements of the paragraph concerned if it deviates with not more than 30 seconds lower than the minimum prescribed for Grade 1, Grade 2 and Grade 3, as the case may be.

(4) The minimum protein content (on a 12 percent moisture basis) for the different grades shall be as follows:

(a) Grade 1 - 12 percent.

(b) Grade 2 - 11 percent.

(c) Grade 3 - 10 percent.

(d) Grade 4 - 9 percent; and

(e) Utility Grade - 8 percent.

PART II

PACKING AND MARKING REQUIREMENTS

Packing requirements

7. Wheat of different grades shall be packed in different containers, or stored separately.

Marking requirements

8. (1) Every container or the accompanying sale documents of a consignment of wheat shall be marked or endorsed by means of appropriate symbols specified in sub-regulation (2), with --

(a) the class of the wheat; and

(b) the grade.

(2) The symbols referred to in sub-regulation (1) shall appear in the order of class and grade.

(3) The symbols used to indicate the different --

(a) classes shall be --

(i) B in the case of Bread Wheat; and

(ii) O in the case of Other Wheat.

(b) grades shall be --

(i) 1 in the case of Grade 1;

(ii) 2 in the case of Grade 2;

(iii) 3 in the case of Grade 3;

(iv) 4 in the case of Grade 4; and

(v) UT in the case of Utility Grade.

PART III

SAMPLING

Taking of sample

9. (1) A sample of a consignment of wheat shall --

(a) in the case of wheat delivered in bags and subject to regulation 10, be obtained by sampling at least ten percent of the bags, chosen from that consignment at random, with a bag probe: Provided that at least 25 bags in a consignment shall be sampled and where a consignment consists of less than 25 bags, all the bags in that consignment shall be sampled; and

(b) in the case of wheat delivered in bulk and subject to regulation 10, be obtained by sampling that consignment throughout the whole depth of the layer, in at least six different places, chosen at random in that bulk quantity, with a bulk sampling apparatus.

(2) The collective sample obtained in sub-regulation (1) (a) or (b) shall --

(a) have a total mass of at least 10 kg; and

(b) be thoroughly mixed by means of dividing before further examination.

(3) If it is suspected that the sample referred to in sub-regulation (1)(a) is not representative of that consignment, an additional five percent of the remaining bags, chosen from that consignment at random, shall be emptied into a suitable bulk container and sampled in the manner contemplated in sub-regulation (1)(b).

(4) If it is suspected that the sample referred to in sub-regulation (1)(b) is not representative of that consignment, an additional representative sample shall be obtained by using an alternative sampling pattern, apparatus or method.

(5) A sample taken in terms of these regulations shall be deemed to be representative of the consignment from which it was taken.

Sampling if contents differ

10. (1) If, after an examination of the wheat taken from different bags in a consignment in terms of regulation 9(1) (a), it appears that the contents of those bags differ substantially --

(a) the bags concerned shall be placed separately;

(b) all the bags in the consignment concerned shall be sampled with a bag probe in order to do such separation; and

(c) each group of bags with similar contents in that consignment shall for the purposes of these regulations be deemed to be a separate consignment.

(2) If, after the discharge of a consignment of wheat in bulk has commenced, it is suspected that the consignment could be of a class or grade other than that determined by means of the initial sampling, the discharge shall immediately be stopped and the part of the consignment remaining in the bulk container as well as the wheat already in the hopper shall be sampled anew with a bulk sampling apparatus or by catching at least 20 samples, by means of a suitable container, at regular intervals throughout the whole offloading period from the stream of wheat flowing in bulk.

Working sample

11. A working sample is obtained by dividing the representative sample of the consignment according to the latest revision of the ICC (International Association for Cereal Science and Technology) 101/1 method.

PART IV

DETERMINATION OF OTHER SUBSTANCES

Determination of undesirable odours and harmful substances

12. A consignment of wheat or a sample of a consignment of wheat shall be sensorially assessed or chemically analysed in order to determine whether--

(a) it contains a substance that renders the wheat unfit for human consumption or for processing into or for utilisation as food or feed such as poisonous seeds, stones, glass, metal, coal or dung; and

(b) it has a musty, sour, rancid or other undesirable odour: Provided that a working sample of unscreened wheat that is ground in a grain mill to a fine meal may be used for the determination concerned.

PART V

DETERMINATION OF CLASS, HECTOLITRE MASS, MOISTURE CONTENT, PROTEIN CONTENT AND FALLING NUMBER

Determination of class

13. The class of a consignment of wheat shall be determined as follows:

(a) Obtain a working sample of at least 500 g and screen the working sample in the manner prescribed in regulation 18.

(b) Take at least 100 g of the screened wheat and remove all other grain, un threshed ears and foreign matter by hand.

(c) Obtain a working sample of at least 25 g each after all other grain, unthreshed ears and foreign matter have been removed and separate the different cultivars.

(d) Determine the combined mass of all of the cultivars that belong to the same class and express the mass thus determined as a percentage of the mass of the working sample.

(e) Such percentage represents the percentage of all the cultivars that belong to the same class in the consignment.

Determination of the hectolitre mass

14. The hectolitre mass of a consignment of unscreened wheat may be determined by any suitable instrument: Provided that the instrument complies with and has been calibrated to the specifications detailed in ISO (International Organization for Standardization) 7971-3.

Determination of moisture content

15. The moisture content of a consignment wheat may be determined by any suitable method: Provided that the results thus obtained is in accordance with the maximum permissible deviation for a class 1 moisture meter as detailed in ISO (International Organization for Standardization) 7700/1 based on the results of the 72 hour, 103°C oven dried method [the latest revision of the AACCI (American Association of Cereal Chemists International) Method 44-15A].

Determination of protein content

16. The percentage of protein of a consignment of wheat may be determined according to any suitable method: Provided that --

(a) the determination shall be conducted on a sample which had been sifted using a screen with the same apertures as the standard sieve and from which other grain, un threshed ears and foreign matter had been removed by hand; and

(b) the results thus obtained are in accordance (± 0,3 percent) with the results obtained by the Dumas Combustion Analysis Method [the latest revision of the AACCI (American Association of Cereal Chemists International) Method 46-30].

Determination of falling number in wheat

17. (1) The falling number of a consignment of wheat may be determined according to any suitable method: Provided that --

(a) the determination shall be conducted on a sample which had been sifted using a screen with the same apertures as the standard sieve and from which other grain, unthreshed ears and foreign matter had been removed by hand; and

(b) the results thus obtained are in accordance (± 5 percent) with the results obtained by the latest revision of the ICC (International Association for Cereal Science and Technology) 107/1 method.

(2) If the falling number of a consignment of wheat is determined according to the latest revision of the ICC (International Association for Cereal Science and Technology) 107/1 method --

(a) the sampling in the mentioned method shall be replaced with the manner prescribed in regulation 9; and

(b) only the altitude corrected value shall be used.

PART VI

DETERMINATION OF PERCENTAGE DEVIATIONS

Determination of percentage screenings

18. (1) The percentage screenings in a consignment of wheat shall be determined as follows:

(a) Obtain a working sample of at least 500 g.

(b) Place the sample on the standard sieve and screen the sample by moving the sieve 50 strokes to and fro, alternately away from and towards the operator of the sieve, in the same direction as the long axes of the slots of the sieve. Move the sieve, which rests on a table or other suitable smooth surface, 250 mm to 460 mm away from and towards the operator with each stroke. The prescribed 50 strokes must be completed within 50 to 60 seconds: Provided that the screening process may also be performed with the standard sieve in some or other container or an automatic sieving apparatus.

(c) Determine the mass of the material that has passed through the sieve and express it as a percentage of the mass of the working sample.

(d) Such percentage represents the percentage screenings in the consignment.

Determination of the percentage heavily frost-damaged wheat

19. The percentage heavily frost-damaged wheat in a consignment of wheat shall be determined as follows:

(a) Obtain a working sample of at least 25 g of a screened sample.

(b) Remove all heavily frost-damaged kernels by hand and determine the mass thereof.

(c) Express the mass thus determined as a percentage of the mass of the working sample.

(d) Such percentage represents the percentage heavily frost-damaged wheat in the consignment concerned.

Determination of the percentages other grain and unthreshed ears

20. The percentage other grain and unthreshed ears in a consignment of wheat shall be determined as follows:

(a) Obtain a working sample of at least 50 g from a screened sample.

(b) Remove all other grain and unthreshed ears by hand and determine the mass thereof.

(c) Express the mass thus determined as a percentage of the mass of the working sample.

(d) Such percentage represents the percentage other grain and unthreshed ears in the consignment concerned.

Determination of the percentage foreign matter

21. The percentage foreign matter in a consignment of wheat is determined as follows:

(a) Obtain a working sample of at least 100 g from a screened sample.

(b) Remove all foreign matter by hand and determine the mass thereof.

(c) Express the mass thus determined as a percentage of the mass of the working sample.

(d) Such percentage represents the percentage foreign matter in the consignment concerned.

Determination of the percentage damaged wheat

22. The percentage damaged wheat in a consignment of wheat shall be determined as follows:

(a) Obtain a working sample of at least 25 g of a screened sample.

(b) Remove all damaged kernels by hand and determines the mass thereof.

(c) Express the mass thus determined as a percentage of the mass of the working sample.

(d) Such percentage represents the percentage damaged wheat in the consignment concerned.

Determination of the percentage heat-damaged wheat

23. The percentage heat-damaged wheat in a consignment of wheat shall be determined as follows:

(a) Obtain a working sample of at least 100 g from a screened sample.

(b) Remove all heat-damaged kernels by hand and determine the mass thereof. Kernels from an additional working sample may also be sensorially assessed (by smelling and tasting the kernels) to confirm suspicion of heat damage.

(c) Express the mass thus determined as a percentage of the mass of the working sample.

(d) Such percentage represents the percentage heat-damaged wheat in the consignment concerned.

Determination of percentage field fungi infected wheat

24. The percentage field fungi infected wheat in a consignment of wheat shall be determined as follows:

(a) Obtain a working sample of at least 25 g from a screened sample.

(b) Remove all field fungi infected kernels by hand and determine the mass thereof.

(c) Express the mass thus determined as a percentage of the mass of the working sample.

(d) Such percentage represents the percentage of field fungi infected wheat in the consignment concerned.

Determination of percentage storage fungi infected wheat

25. The percentage storage fungi infected wheat in a consignment of wheat shall be determined as follows:

(a) Obtain a working sample of at least 100 g from a screened sample.

(b) Remove all storage fungi infected kernels by hand and determine the mass thereof.

(c) Express the mass thus obtained as a percentage of the mass of the working sample.

(d) Such percentage represents the percentage storage fungi infected wheat in the consignment concerned.

PART VII

Offence and penalties

26. Any person who contravenes or fails to comply with any provision of these regulations shall be guilty of an offence and upon conviction be liable to a fine of not exceeding R50 000 or to imprisonment for a period not exceeding two years, or to both that fine or imprisonment.

ANNEXURE

TABLE 1

STANDARDS FOR GRADES OF BREAD WHEAT

Nature of deviation	Maximum percentage permissible deviation (m/m)				
	Grade 1	Grade 2	Grade 3	Grade 4	Utility Grade
1	3	4	5	6	7
(a) Heavily frost-damaged kernel	5	5	5	5	10
(b) Field fungi infected kernels	2	2	2	2	2
(c) Storage fungi infected kernels	0,5	0,5	0,5	0,5	0,5
(d) Screenings	3	3	3	4	10
(e) Other grain and unthreshed ears	1	1	1	1	4
(f) Gravel, stones and turf.	0,5	0,5	0,5	0,5	0,5
(g) Foreign matter including gravel, stones and turf: Provided that such deviations are individually within the limits specified in item (f).	1	1	1	1	3
(h) Heat-damaged kernels	0,5	0,5	0,5	0,5	0,5
(i) Damaged kernels, including heat-damaged kernels: Provided that such deviations are individually within the limit specified in item (h) and provided further that the minimum falling number value prescribed in regulation 6(3) for the grade concerned is at least complied with.	2	2	2	2	5
(j) Deviations in items (d), (e), (g) and (i) collectively: Provided that such deviations are individually within the limits of the said items.	5	5	5	5	10