Hierarchical classification pathway for white maize, defect and foreign material classification using spectral imaging

Kate Sendin[1], Marena Manley[1], Federico Marini[1,2] & Paul J. Williams[1]*

[1]*Department of Food Science, Stellenbosch University, Private Bag X1, Matieland (Stellenbosch) 7602, South Africa*

[2]*Department of Chemistry, University of Rome (La Sapienza), P. le Aldo Moro 5, Rome 00185, Italy*

*Corresponding author: Email: pauljw@sun.ac.za; Tel: +27 21 808 3155

ABSTRACT

This study aimed to present the South African maize industry with an accurate and affordable automated analytical technique for white maize grading using near infrared (NIR) spectral imaging. The 17 categories and sub-categories stipulated in South African maize grading legislation were simultaneously classified (1044 samples; 60 kernels of each class) using 25 partial least squares discriminant analysis (PLS-DA) models. The models were assembled in a hierarchical decision pathway that progressed from the most easily classified classes to the most difficult. The full NIR spectrum (288 wavebands) model performed with an overall accuracy of 93.3% for the main categories. Three waveband selection techniques were employed, waveband windows (48 wavebands), variable importance in projection (VIP) (21 wavebands) and covariance selection (CovSel) (13 wavebands). Overall, the VIP set based on only 7.3% of the original spectral variables was recommended as the best trade-off between performance and expected cost of a reduced waveband system.

Keywords: near infrared spectroscopy; spectral imaging; waveband optimization; chemometrics; covariance selection; maize

1. Introduction

Maize grading is conducted throughout the market value chain as maize is traded between the farmer, storage provider and miller. Grading ensures a fair market price per consignment and is based on the basic condition of a sub sample of the maize. South Africa currently uses an inspection method, where a grader manually sorts a 150 g (*c.* 1000 kernels that is representative of a consignment) sample to determine grade based on the presence of undesirable materials (e.g. damaged maize or foreign materials). During this grading process, defective kernels are not discarded they sorted, weight and used to determine the quality of the consignment. To increase throughput and decrease the error associated with this process, the industry is seeking an appropriate analytical method to replace manual inspection. A previous study [1] demonstrated the potential of using hyperspectral imaging for sorting 13 South African maize grading classes with an overall classification accuracy of 99.4% across the 804 kernels/objects. However, this study only considered two-way separations and did not offer a single system for evaluating all the classes simultaneously. While achieving separation of two classes at a time is relatively easy, separating multiple classes is a much more challenging endeavour.

Hyperspectral imaging has been used extensively in cereal science research to evaluate a large array of cereal properties, including hardness classification, chemical composition, variety identification, sprouting detection, physical quality classification, fungal contamination detection and parasitic contamination detection [2]. However, due to crucial drawbacks of the technique, it is seldom implemented for routine analysis in industry. These drawbacks include the high cost and relatively low speed of the hyperspectral imaging instruments found in research laboratories. A viable solution to these issues is the development of a multispectral imaging instrument that is tailor-made for one application. Waveband selection studies have been successfully conducted for separating grain from foreign material [3], identifying maize [4], rice [5] and black bean varieties [6], detecting genetically modified maize[7], tracking texture deterioration in fresh maize [8], and determining spelt flour authenticity [9]. Successive projection algorithm (SPA) is a popular waveband selection method that aims to minimise collinearity between spectral variables [10]. However, SPA only considers the **X**-data and selects wavebands without considering the class information (**y**-data) [11]. The class information should be considered in applications with closely related classes to identify wavebands that specifically highlight the differences between two classes. Inspired by SPA, covariance selection (CovSel) works in a similar way but accounts for the covariance between the **X**- and **y**-data. Simply put, the difference between the two is comparable with the differences between principal component analysis (PCA) and partial least squares (PLS).

Hierarchical or decision pathway modelling is a potential solution for multi-class classification problems [12]. Many studies with two, three or even four classes utilise a single globally optimised model to discriminate all classes. This approach is easily and widely accessible to perform, but one critical assumption must be satisfied, i.e. all classes must be fully separable using the selected set of spectral features. This assumption is often not fulfilled, especially when dealing with heterogeneous samples and closely related classes, as was observed in the partial least squares-discriminant analysis (PLS-DA) scores plots in Sendin, Manley, Baeten, Fernández Pierna and Williams [1]. Instead of performing multi-class classification (e.g. 13 class PLS-DA model for the abovementioned study), hierarchical modelling decomposes the problem into simpler binary classification steps (two or three class PLS-DA models) that are reassembled into a single hierarchical structure. To minimise the effects of error propagation through the successive steps of the decision pathway, the pathway must be carefully selected. A prudent approach is to handle the most easily classified classes first and work towards the most challenging [13]. A recent study demonstrated value of hierarchical modelling for the rapid detection of meat species, processing (fresh or frozen) and muscle type using a hand-held near infrared (NIR) spectrometer [14]. However, the use of hierarchical pathway modelling for multi-class problems in NIR hyperspectral imaging or for the classification of cereals remains limited.

The aim of this study was to simultaneously distinguish sound white maize kernels from common undesirable material types stipulated in the South African maize grading legislation using NIR spectral imaging. This was achieved through hierarchical assembly of PLS-DA classification models for the separation of 17 classes. One hierarchical model based on the full spectrum and three based on different waveband selection methods (waveband windows, waveband optimisation based on VIP scores, and waveband optimisation using the CovSel algorithm) were developed.

2. Materials and methods

2.1 Samples

White maize kernels and undesirable materials were obtained from the Southern African Grain Laboratory (SAGL, Pretoria, South Africa) and Pioneer Foods (Paarl, South Africa) in August 2018. These maize samples were silo samples (i.e. mixed origin, cultivar, and harvest date), and were graded visually by expert graders according to South African white maize grading regulations ([15]. This act stipulates five main categories, namely sound (healthy) white maize, defective white maize, pinked white maize, other colour (yellow) maize, and foreign materials (Table 1). The legislation also stipulates sub-categories for defective kernels and foreign materials. Of the nineteen defects stipulated in the grading regulation, twelve were evaluated during this study since these were prevalent during the 2018 season. These included *Fusarium* fungal, *Diplodia* fungal, heat, water, frost and pest (rodent and insect) damage, as well as broken (screenings), sprouted and immature kernels. Foreign materials included five common commodities, including soy, sorghum, sunflower seeds and wheat, as well as miscellaneous plant materials. See Fig. 1 for a digital image of all classes included in the study.

Calibration and validation sample sets were selected at random for each of the 17 classes, where one set of 60 kernels/objects was used for calibration and another set of 60 was used for validation. There were three exceptions, namely (1) pest damage, which included separate sets of 60 kernels for rodent damage and insect damage (total of 120 for calibration and 120 for validation); (2) sprouted kernels, where 30 kernels were used in each set due to limited availability; and (3) immature kernels, where 54 kernels were used in each set due to limited availability. Overall, 1044 samples were used for calibration and 1044 for validation, giving a total of 2088 samples.

2.2 NIR hyperspectral system

Hyperspectral images were acquired using a short-wave infrared (SWIR) camera (Hyspex SWIR-384 Norsk Elektro Optikk, Norway) in reflectance mode. The camera had a mercury–cadmium–telluride (HgCdTe) detector and operated in the range 953 to 2517 nm, with 5.45 nm between each of the 288 spectral bands. Images were 384 pixels wide, and varied in length of *ca.* 700 pixels. The frame period was 3800 µs and the integration time was 3600 µs, chosen by visually assessing the saturation images of the samples during test scans. Samples were illuminated with a halogen light source, which was switched on 10 min before imaging to avoid light source temperature drift and ensure spatial lighting uniformity. A 50% grey Zenith Allucore diffuse reflectance standard (SphereOptics GmbH, Germany) was used for image correction and calibration, and was scanned every 30 min during the imaging session.

2.3 Image acquisition

Unique calibration and validation images were captured for each of the 17 classes individually. As pest damage included rodent and insect damage, two image sets were taken for pest damage, giving a total of 18 calibration and validation image sets (36 images). Sixty kernels/objects of a single class were arranged in a grid of 6 ×

10. In applicable classes, the top three rows were placed with the maize germ facing up towards the camera, and the bottom three rows with germ facing down.

2.4 Hyperspectral image analysis

2.4.1 Image correction and cleaning

Radiometric calibration from irradiance to radiance to pseudo-absorbance was done in the HySpex Ground software v4.1 (HySpex, Norsk Elektro Optikk, Norway). The grey reference image acquired most recently preceding the acquisition of each image was used. There is no significant difference in the calculated reflectance values of the sample when using the 50% (grey) or 99% (white) reflectance target. The 50% target allows for using longer integration times that would typically saturate the 99% target. The longer integration times lead to increased SNR for these samples.

Cleaning was conducted using the interactive Evince v.2.7.0 (Prediktera AB, Umeå, Sweden) spectral image analysis software. On average, the raw hypercubes were 384 * 973 * 288 in dimension (373632 pixels). However some differed slightly in the y-dimension due to the number of samples or length of the sample holder imaged. PCA was applied to the mean-centred unfolded hypercubes (373632 * 288), and the score plots and score images were used interactively to identify unwanted pixels, e.g. outliers, sample stage background, dead pixels, shading errors and edge effects [16]. A notable issue in several images was either specular reflection or overexposure occurring in some small regions. All unwanted pixels were removed.

2.4.2 Particle analysis

The cleaned images were analysed further in PLS_Toolbox (Eigenvector Research Inc., Wenatchee, WA) software package and subjected to particle (object) analysis. Objects were identified as isolated contiguous regions of pixels with similar intensity values. Each pixel was assigned either 0 or 1 to indicate non-object pixels (deleted background) and potential object pixels (maize kernels), creating a binary image or image mask. The mean spectrum of each object was calculated based on the arithmetic mean of all pixel spectra within the object. Thus, an image of *ca.* 200000 pixel spectra was reduced to *ca.* 60 mean spectra while the retaining spatial information. A table of the 60 mean spectra of each calibration image was created to further reduce the data size from *ca.* 100 MB to 100 KB. As the objects were numbered when calculated (i.e. 1 – 60), the information from the table could be related back to the image mask at a later stage.

The mean spectra table of all 18 calibration images were combined to give one table with 288 wavebands as columns and 1044 calibration samples as rows. The class of each sample in the table was assigned. This was repeated for the validation data.

2.5 Optimal waveband selection

2.5.1 Reduced spectral channels (windows)

The number of spectral channels was reduced by dividing the 288 wavebands (953 – 2517 nm) into 48 windows of 6 wavebands. The third waveband in each window was chosen as the centre point. The 48 selected wavebands were: 964, 996, 1029 1062, 1095, 1127, 1159, 1193, 1225, 1258, 1291, 1323, 1356, 1388, 1421, 1454, 1487, 1520, 1552, 1585, 1618, 1651, 1683, 1716, 1749, 1781, 1814, 1847, 1879, 1912, 1945, 1978, 2010, 2043, 2076, 2108, 2141, 2174, 2206, 2239, 2272, 2305, 2337, 2370, 2403, 2435, 2468 and 2501 nm. The pre-processed mean spectrum (SNV transformation) marked with the windows (grey and red), VIP (red) and CovSel (green) waveband sets is shown in Fig. 2.

2.5.2 Variable importance in projection scores

Variable importance in projection (VIP) scores were calculated based on the PLS-DA models for each of the 15 levels or sub-levels of the full spectrum hierarchical model (see Section 2.6). VIP scores evaluate the importance of each waveband for separating the classes in a PLS-DA model, where the VIP score of waveband k was calculated according to Eq. 1:

$$VIP_k = \sum_{j=1}^{a} \left(w_{jk}^2 SSR_j \right) \frac{L}{SST} \quad (1)$$

where:

k is waveband, a is the number of latent variables (LVs) in the PLS-DA model, w is the PLS weight of waveband k, SSR is the residual sum-of-squares, L is the number of wavebands (288) and SST is the total sum-of-squares.

The VIP scores were calculated based on the PLS-DA calibration dataset, and thus pre-processed spectra (Savitzky-Golay (7 smoothing points; 3rd order polynomial; 1st derivative), SNV and mean-centring). A line chart was generated displaying the waveband and VIP score value for each PLS-DA model. Waveband windows or groupings were used to overcome multicollinearity issues. If a maximum value appeared at any waveband in this window, it was recorded and shaded as follows: below 0.99 – unshaded, 1 to 1.49 – green, 1.5 to 1.99 – yellow, 2 to 2.49 – orange, and above 2.5 – red (Fig. 3). A VIP score value greater than 1 indicated that a window was highly influential for the separation of a particular class. Any window scoring above 1 in 7 or more of the 15 PLS-DA models was chosen as part of the optimised waveband set. The 21 selected wavebands were: 964, 1127, 1159, 1323, 1356, 1388, 1421, 1716, 1847, 1879, 1912, 1945, 2043, 2239, 2272, 2305, 2337, 2403, 2435, 2468 and 2501 nm.

2.5.3 Covariance selection

CovSel was calculated based on methods described by Roger, Palagos, Bertrand and Fernandez-Ahumada [10] and Biancolillo, Marini and Roger [17]. The process takes place in two main steps: (i) identifying the variable with the highest covariance by calculating the covariance between all the **X**- and **y**-variables; and (ii) projecting all the **X**- and **y**-variables orthogonally to the identified variable until an optimal number of wavebands was selected. The 13 selected wavebands were: 953, 1122, 1340, 1416, 1574, 1721, 1869, 1901, 1939, 1994, 2097, 2250 and 2512 nm.

2.5.4 Mean spectrum pre-treatment

The following pre-treatments were considered: (1) mean-centring; (2) standard normal variate (SNV); (3)

Savitzky-Golay transformation (various polynomial, derivative, and smoothing parameters); and (4) detrending. Preliminary two-way PLS-DA models of sound maize vs. each class were calculated to evaluate the pre-treatment combinations based on cross-validated classification results (venetian blinds cross-validation). Pre-treatments yielding consistently good classification results were chosen. No noisy wavebands were observed in the mean spectra, thus all 288 wavebands were kept as variables.

Savitzky-Golay (7 smoothing points; 3rd order polynomial; 1st derivative), SNV and mean-centring was applied for the full spectra [18, 19]. Only SNV and mean-centring were applied to the reduced waveband spectra. Savitzky-Golay transformation was not applied to these discrete datasets because the transformation involves smoothing and the calculation of derivatives, thus requiring continuous points.

2.6 Hierarchical model development and calibration

A series of PLS-DA models were calculated and assembled in a hierarchical model that consisted of various levels and sub-levels. A detailed description of the model's architecture is given in the Supplementary Information. The mean spectra of all 18 calibration images were used to calibrate all PLS-DA models in the hierarchical model. Four separate hierarchical models were developed, for the full spectrum, window wavebands, VIP scores and CovSel waveband sets. The specific order of the sub-levels in each hierarchical model was optimised individually according to the performance of the PLS-DA models, where the structure of the hierarchical model based on the full spectrum data is given in Table S1.

Hierarchical pathway Level 1 classified each object as either a foreign material or a maize kernel. A two-class PLS-DA model of a grouped maize class vs. a grouped foreign material class was calculated. The grouped maize class consisted of the sound white maize, all defective white maize classes, pinked white maize and yellow maize classes, and the grouped foreign material class consisted of soy, sorghum, sunflower seeds, wheat and plant material. If classified as a maize kernel, the object proceeded to the hierarchical model branch for maize kernel classification (Level 2), and if classified as a foreign material, the object proceeded to the hierarchical model branch for foreign material classification (Level 3).

At Level 2, the maize hierarchical model was used to classify the following 12 classes: sound maize, screenings, *Fusarium* damage, *Diplodia* damage, heat damage, water damage, frost damage, pest damage, sprouted kernels, immature kernels, pinked maize, and yellow maize. The hierarchical model structure was designed based on separating the most easily separated class from the rest, with following steps working towards the most difficult class. The order was determined by calculating two-class PLS-DA models of one class vs. a grouped class of all other maize classes and evaluating the cross-validated classification result, where the classes were ordered according to descending model performance.

At Level 3, the foreign material hierarchical model first separated objects into two main categories of surface chemical composition, namely cellulose-rich and starchy. Thus, the first two-class PLS-DA model separated a grouped class of soy, sorghum and wheat (starchy) and a grouped class of sunflower seeds and plant material (cellulose-rich). Next, a three-class PLS-DA separated soy, sorghum and wheat, and a two-class PLS-DA model separated sunflower seeds and plant material.

A secondary classification step was added to the decision pathway for most classes, including both maize and foreign materials. This accounted for easily confused classes within the grouped classes. For example, heat damage and yellow maize were easily confused. When a classification result of 'heat damage' was generated, this object was classified by a second two-way PLS-DA of heat damage vs. yellow maize. The result of the secondary classification was used as the final result in all instances where implemented. The secondary classification steps included in the hierarchical model based on the full spectrum data are listed in Table S1.

2.7 Hierarchical model validation

The hierarchical model was tested using the mean spectra of the 18 validation images. The mean spectrum of each object was classified by the hierarchical model, beginning at the Level 1 classification of maize or foreign materials, and moving on to the relevant subsequent branches and models until a final classification was made.

The final classification was evaluated in two ways, namely whether the main category was correct and, where relevant, whether the sub-category was correct. For legislative purposes, only the main category is necessary, but a human grader (to which this study is comparing spectral imaging) is able to see the difference between sub-categories.

3. Results and discussion

3.1 Experimental design

This study aimed to emulate the current grading practices as closely as possible. Although based on legislative guidelines, human graders make logical decisions which cannot always be strictly defined. This is an important aspect of human decision-making that is difficult to replace using an automated analytical technique. An important decision-making step occurred when sorting the defective white maize samples used in the study. Many of the defects occur simultaneously, or one defect can make a kernel susceptible to another at a later stage. For example, a kernel may become sprouted, water and/or frost damaged during a bout of bad weather, while rodent damage may leave a kernel vulnerable to insect or fungal infestation. Further, some ear diseases, including *Diplodia*, induce sprouting of kernels. Kernels presenting symptoms of multiple defects were encountered regularly during the grading of samples for this study, in which case the grader determined the predominant defect. However, this is very subjective. Further, the legislation is based solely on levels of each main category (e.g. all defective kernels). Thus, misclassifications at the sub-category level (as in any of the examples given above) were not considered a major shortfall, and would have no effect on the overall accuracy of the assigned grade.

The aims of this study presented two challenges, namely separating closely related samples and separating an unusually large number of classes compared to other similar hyperspectral imaging studies. This application was a good candidate for hierarchical modelling, where multiple classes are classified stepwise, working from most easily separated to most closely related. An object was first classified as a maize kernel or foreign material, and these two groups were further classified separately. Separating the maize kernels was the most challenging task, where the twelve closely related classes were separated sequentially as sound maize, yellow maize, pinked maize and defective maize (nine sub-categories). The classification of foreign materials was less challenging, as this was separating different commodities, not classes of a single commodity.

The secondary classification step was introduced to minimise misclassification between closely related classes. Due to the large number of classes used, a clear separation between each class was not expected. During hierarchical model development, the classification results for one class vs. all remaining classes were examined. If a substantial number (*ca.* 10%) of class 1 was misclassified as class 2 (i.e. errors do not occur randomly), a secondary step was included, where all objects classified as class 2 were predicted by a simple two-class PLS-DA (class 1 vs. class 2) and this result was taken as the final classification. The secondary step classification models had high Q^2-values and excellent cross-validated classification accuracies (often 98 – 100%), and single classes were well-defined and easily separated. The number of errors between closely related classes was greatly reduced by including this step.

3.3 Full spectrum classification

The full spectrum hierarchical model performed well, with high classification accuracy (75 – 100%) considering the challenging task of separating 17 classes (Table 2). The sub-category and main category classification accuracies were both recorded, describing if an object was classified as the correct class, or as any class in the correct main category, respectively. The sub-category accuracy of the defective white maize classes appeared low (13 – 95%), but this should not be of concern if they are misclassified among themselves and not as other main categories. As previously mentioned, these defects either occur simultaneously or cause vulnerability to other defects. For instance, the water damaged kernels (23% sub-category accuracy) were almost exclusively misclassified for frost damage and sprouting. Frost damage is often viewed as a severe form of water damage (see digital image in Fig. 1) and sprouting occurs as a result of prolonged exposure to water. An experienced grader struggles to determine the sub-category of kernels with differing severity of multiple defects, thus the class determined by the reference method (human grading) in these cases was not necessarily more accurate than the hyperspectral imaging method under investigation.

The main category accuracy is the most important parameter for grading based on the current legislation. An overall classification accuracy of 93.3% was achieved across the 1044 validation samples. The individual accuracies were as follows: 88% for sound white maize, 93% for defective white maize, 83% for pinked white maize, 75% for yellow maize and 100% for foreign materials. There was a tremendous improvement in the accuracy of detecting defective maize kernels in the main category accuracy compared to the sub-categories, confirming that these classes were predominantly confused with sub-categories within the same main category.

Sound white maize was the most important class to classify accurately, as a normal grading sample is expected to contain *ca.* 95% sound white maize. If a large number of errors occur in this class, an inaccurate grade is likely to be assigned. The results were fair, but 7 of the 60 kernels were misclassified. These misclassifications were of an array of defects, pinked and yellow maize, with no clear links to a specific class. Conversely, very few objects of other classes were misclassified as sound. In other words, the model was sensitive for the detection of undesirable materials but less specific for the detection of sound maize. This shows promise for ensuring a system that does not allow defective or unsafe materials to enter the food chain.

Although visually distinct, the separation of yellow maize was the most challenging. This was also observed in an earlier study by Sendin, Manley, Baeten, Fernández Pierna and Williams [1], where the two-way PLS-DA separation of white and yellow maize in a similar spectral region also achieved 75% correct classification. There is only one specific difference between white and yellow maize and it is determined by the presence of a single gene [20]. This gene controls the production of yellow beta-carotene pigment in the maize endosperm. The presence of two recessive alleles results in no pigment formation (white) and the presence of a single dominant allele causes pigment formation (yellow). While the two commodities exhibit a distinct colour difference in the visible region, beta-carotene has an absorbance maximum at 440 nm and does not interact strongly with NIR radiation [21]. Thus, NIR spectroscopic techniques are not suited to detect this specific chemical constituent. Other differences in the chemical composition of maize samples, such as hardness, moisture content and oil content, can vary as greatly between cultivars of white maize as between white and yellow maize. The classification of yellow and white maize using NIR spectral imaging remains a challenge, and a possible solution is including spectral variables from the visible region, specifically a band at 440 nm. A similar phenomenon was observed for pinked maize, as this class is also highly related to sound white maize. The light pink superficial discolouration of these maize kernels is due to the production of a red pigment, anthocyanin [22]. The discolouration is limited to the pericarp and does not affect meal colour after milling. Furthermore, it does not cause any other internal changes. Certain white hybrids are simply prone to pinking under specific climatic conditions, such as sunlight exposure, and the defect is very leniently legislated with a maximum allowed content of 20%. However, the subtlety of this defect resulted in a classification accuracy of only 83%. The addition of a spectral variable at 550 nm is expected to aid the classification of pinked maize. It should also be noted that improved classification of yellow and pinked maize should lead to improved classification of sound white maize due to increased class separation.

3.3 Reduced spectral channels (windows) classification

Spectral features in NIR spectroscopy (e.g. harmonics and combination bands) are associated with broad peaks. Pure substances are often characterised by natural bandwidths larger than 10 nm, while mixtures are usually broader [23, 24]. Examples of these large bandwidths include 22.5 nm for sucrose (centred at ~2046 nm), 30.1 nm for maize oil (~2305 nm), 110.4 nm for moisture (~1928 nm) and 162 nm for wheat starch (~2103 nm) [25]. The spectral intervals of the hyperspectral instrument's full spectrum were 5.45 nm, and the interval of each window of 6 wavebands was 32.7 nm. Thus, many of the spectral features that play a role in classifying maize kernels spanned one or more windows (Fig. 2).

An instrument that measures a large number of wavebands is high in cost, as a more expensive sensor is required. When considering the requirements of an instrument for a specific application, a trade-off between performance and price is often unavoidable. The window-based hierarchical model was based on 16.7% of the original spectral variables. The model did not perform as accurately as the full spectrum model, with an overall loss of accuracy from 93.3 to 87.1% that affected all classes (Table 3). Furthermore, the windows waveband set did not offer better classification performance than the other two reduced waveband sets in a number of classes. While an instrument that acquires only 48 wavebands will cost less than the hyperspectral instrument with 288 wavebands, it would be the most expensive of the three optimised sets presented. The trade-off between instrument performance and cost associated with the windows waveband set was not favourable.

3.4 Variable Importance in Projection (VIP) classification

VIP scores reveal which wavebands are the biggest drivers of separation throughout all of the LVs calculated in a single PLS-DA model. This is an advantage over using loadings values, where it is difficult to assess the importance of a waveband when numerous components are calculated. Higher VIP scores are considered more important, where a score greater than 1 is considered as highly influential, between 0.8 and 1 as moderately influential, and less than 0.8 as less influential [26]. Due to the large number of analyses, only wavebands with a score greater than 1 were investigated. The VIP scores results in Fig. 3 were shaded according to increasing values as follows: below 0.99 – unshaded, 1 to 1.49 – green, 1.5 to 1.99 – yellow, 2 to 2.49 – orange, and above 2.5 – red. This allowed easy visual assessment of important spectral regions, appearing as hot zones, and uninformative regions which remained blank.

Cereals, including maize, comprise of several major chemical constituents, namely starch, protein, fat/oil and moisture. When interpreting the NIR spectrum of cereal samples, prominent regions are expected to be associated with chemical bonds present in these constituents. A total of 21 wavebands were selected based on VIP scores, of which a large number were attributed to starch, including 1879 nm (O–H stretch and C–O stretch), 2272 nm (O–H stretch and C–C stretch), 2435, 2468 and 2501 nm (all C–H stretch and C–C stretch) [27, 28]. Absorption bands related to cellulose were in close proximity to the starch associated wavebands, as the two chemical components are very similar, and included 1847 nm (O–H and C–O stretch) and 2337 nm (C–H stretch and deformation) [29]. Wavebands attributed to protein or amino acids included 1127 nm (N–H stretch), 2043 nm (N–H symmetrical stretch) and 2239 nm (N–H stretch and NH_3 deformation) [27]. While CH_2 and CH_3 groups are common in organic molecules, fats are the main chemical components in maize associated with these functional groups and were related to 1159 nm (C–H stretch), 1323 nm (C–H stretch and deformation), 1716 nm (C–H stretch) and 2305 nm (C–H stretch and deformation) [27, 28]. The wavebands linked to moisture were among the lower scoring significant VIP scores, which included 964 nm (O–H stretch), 1945 nm (O–H stretch and deformation) and 2403 nm (O–H deformation) [27]. The absorption bands at 1421 and 1912 nm (both O–H stretch) were associated with alcohol groups. Specifically, 1421 nm is attributed to absorption by an aromatic alcohol and may be related to the amino acid tyrosine, as maize is known to be rich in this minor component [30].

The performance of the hierarchical models based on the windows (87.1%) and VIP (84.5%) wavebands sets were comparable (Table 3). While the windows waveband included 16.7% of the spectral variables, this was further reduced to 7.3% with little further loss of classification accuracy. The classification accuracy of white maize was 78.3%, which is a 10% drop from the full spectrum classification. This was concerning, as the majority of a white maize grading sample is expected to belong to this class and a large error is likely to result in an unacceptably high rate of misclassification. However, of the three reduced waveband sets, the VIP set performed best for the sound maize class. The classification accuracy for the defective white maize main category also decreased by 10% compared to the full spectrum. This change was most notable in the subtle defects (e.g. screenings and *Diplodia* fungal damage). The 18% decrease in classification accuracy for yellow maize was related to the 18% decrease for heat damage, as a large proportion of the heat damaged kernels were misclassified as yellow maize and *vice versa*. Pinked maize was difficult to classify using the full spectrum, and while a 5% decrease was observed in comparison to the full spectrum, the VIP hierarchical model outperformed the windows hierarchical model by 3%. Lastly, a small number of the foreign materials were misclassified (4%), but the occurrence of these materials is very rare due to the use of dockage sorters early in the processing of maize. A classification accuracy of 95.7% is considered high in NIR hyperspectral imaging applications, and comparable to the performance of the windows hierarchical model (96.7%).

3.5 Covariance selection classification

CovSel has not been reported extensively in literature but is a hybrid of the popular SPA technique. It was prudent to use CovSel in this application, as the samples within each class were highly heterogeneous. An unsupervised technique (e.g. SPA) will identify sources of variation between all of the spectra, regardless of class, and would thus include intra-class variation. By considering the covariance between the **X**- and **y**-variables, only inter-class variation was considered. Nine of the thirteen CovSel wavebands were associated with the same spectral features as the VIP wavebands discussed in Section 3.4, including 953 nm (linked to 964 nm), 1122 nm (1127 nm), 1340 nm (1323 nm), 1721 nm (1716 nm), 1869 and 1901 nm (1879 nm), 1939 nm (1945 nm), 1994 nm (2043 nm) and 2512 nm (2501 nm). The waveband at 1416 nm was closely related to the aromatic alcohol band in the VIP set (1421 nm), however 1416 nm is attributed to C–H stretching and deformation in aromatic rings, and not to the alcohol group (O–H) [26]. Of the three wavebands unique to the CovSel, two were associated with starch, namely 2097 nm (O–H deformation and C–O stretch) and 2250 nm (O–H stretch and deformation) [26]. Lastly, the band at 1574 nm (N–H stretch) is specifically related to the peptide bonds (-CONH-) linking amino acids in proteins [27].

While all three of the reduced waveband sets resulted in decreased model performance, CovSel gave the poorest results, with an overall classification accuracy of 81.9% (Table 3). A notable difference between the VIP and CovSel sets lies in the range 2250 to 2512 nm (Fig. 2). Numerous high absorbance bands occurred in this part of the spectrum and the VIP scores clearly highlighted this region. These features were attributed to C–C and C–H bonds in starch and cellulose, two chemical components which contribute to large proportions of a maize kernel. The omission of this region from the CovSel waveband set likely contributed to its poor performance. The class of most concern was sound white maize, which was classified with a relatively low classification accuracy of 63% (25% decrease compared to full spectrum). Yellow maize was also classified poorly with an accuracy of 48% (27% decrease). Pinked maize did not exhibit the same dramatic decrease in classification accuracy (3% decrease), and was the only main category in which the windows and VIP hierarchical models were outperformed by CovSel.

While the capacity of a human to conduct manual classification reliably and consistently is limited [31], the performance of the CovSel hierarchical model would not be an attractive option for the industry. The cost of building an instrument based on this set would be the cheapest, as only 4.5% of the original spectral variables were used. However, the loss of classification ability by removing 95.5% of the spectral variables was too high.

4. Conclusion

NIR hyper- and multispectral imaging show promise as an automated analytical technique for white maize grading. The complex task of maize grading was broken down to simple binary steps that were assembled in a single hierarchical decision pathway. The hierarchical model classified the kernels according to their full mean spectrum with an overall accuracy of 93.3% for the main categories (5 classes) and 75.5% for the sub-categories (17 classes). While the accuracy of human grading is difficult to determine and accurately compare with an alternative method, the performance of the NIR hyperspectral imaging method was impressive.

Waveband reduction and optimisation was conducted to establish if a simpler spectral imaging solution could be provided to the South African maize industry at a lower cost. Three approaches to waveband selection were investigated, namely waveband windows, VIP waveband selection and CovSel waveband selection. The overall classification accuracy decreased from 93.3% for the full spectrum to 87.1%, 84.5% and 81.9% for the windows, VIP and CovSel waveband sets, respectively. The waveband windows approach simply reduced the number of spectral variables from 288 to 48 by selecting every sixth waveband from the full spectrum, thus preserving only 16.7% of the spectral variables. The decreased accuracy was consistent across the main categories, although yellow maize suffered a considerable drop. VIP scores highlighted the 21 wavebands with the highest weighting throughout the individual PLS-DA models in the hierarchical model. This hierarchical model performed with similar main category classification accuracies to the windows model, despite using only 7.3% of the spectral variables. CovSel is a sophisticated waveband optimisation algorithm that was used to select 13 wavebands (4.5% of the spectral variables wavebands) based on the co-variance between the **X**- and **y**-data. However, the overall main category classification was considered too low. Throughout the waveband selection trials, a trade-off between performance and price was unavoidable. Considering the results of all three waveband reduction and optimisation approaches, the 21 wavebands selected based on VIP scores (964, 1127, 1159, 1323, 1356, 1388, 1421, 1716, 1847, 1879, 1912, 1945, 2043, 2239, 2272, 2305, 2337, 2403, 2435, 2468 and 2501 nm) are recommended for white maize grading using reduced waveband spectral imaging.

The classification of sound white, pinked white and yellow maize should ideally be improved. This could be achieved by including visible wavebands, as pinked white maize and yellow maize are distinguishable due to the presence of anthocyanin (550 nm) and beta-carotene (440 nm), respectively, which do not interact with NIR radiation. The overall performance and robustness could be improved using a larger sample set collected over several harvest seasons. Furthermore, this larger number of samples would justify recalculation of the models using non-linear techniques such as locally weighted PLS-DA, kernel- or dissimilarity-PLSDA, or non-linear support vector machines (SVM). Overall, hierarchical modelling allowed for the classification of 17 classes and shows promise for extending the application of NIR hyperspectral imaging to more complex applications in the food and agro-product industries.

Acknowledgements

The authors also wish to acknowledge The South African Grain Laboratory and Pioneer Foods for their contribution of samples.

Funding

This work is based on the research supported in part by the National Research Foundation (NRF) of South Africa (Grant Number: 116307); The NRF is also acknowledged for financial support of Ms K Sendin (Freestanding Doctoral Bursary; Grant number 112569). The NRF's National Equipment Program is acknowledged for funding the NIR hyperspectral imaging instruments through the grant UID 105652. The hyperspectral imaging work performed in this article was conducted at the Vibrational Spectroscopy Unit, which belongs to the Central Analytical Facility at Stellenbosch University.

References

[1] K. Sendin, M. Manley, V. Baeten, J.A. Fernández Pierna, P.J. Williams, Near Infrared Hyperspectral Imaging for White Maize Classification According to Grading Regulations, Food Anal. Meth., 12 (2019) 1612-1624.

[2] K. Sendin, P.J. Williams, M. Manley, Near infrared hyperspectral imaging in quality and safety evaluation of cereals, Crit. Rev. Food Sci. Nutr., 58 (2018) 575-590.

[3] C. Wallays, B. Missotten, J. De Baerdemaeker, W. Saeys, Hyperspectral waveband selection for on-line measurement of grain cleanness, Biosys. Eng., 104 (2009) 1-7.

[4] L. Wang, D.-W. Sun, H. Pu, Z. Zhu, Application of hyperspectral imaging to discriminate the variety of maize seeds, Food Anal. Meth., 9 (2016) 225-234.

[5] L. Wang, D. Liu, H. Pu, D.-W. Sun, W. Gao, Z. Xiong, Use of Hyperspectral Imaging to Discriminate the Variety and Quality of Rice, Food Anal. Meth., 8 (2014) 515-523.

[6] J. Sun, S. Jiang, H. Mao, X. Wu, Q. Li, Classification of black beans using visible and near infrared hyperspectral imaging, Int. J. Food Prop., 19 (2016) 1687-1695.

[7] X. Feng, Y. Zhao, C. Zhang, P. Cheng, Y. He, Discrimination of transgenic maize kernel using NIR hyperspectral imaging and multivariate data analysis, Sensors, 17 (2017) 1894.

[8] L. Wang, H. Pu, D.-W. Sun, D. Liu, Q. Wang, Z. Xiong, Application of hyperspectral imaging for prediction of textural properties of maize seeds with different storage periods, Food Anal. Meth., 8 (2015) 1535-1545.

[9] W.-H. Su, D.-W. Sun, Facilitated wavelength selection and model development for rapid determination of the purity of organic spelt (Triticum spelta L.) flour using spectral imaging, Talanta, 155 (2016) 347-357.

[10] J. Roger, B. Palagos, D. Bertrand, E. Fernandez-Ahumada, CovSel: Variable selection for highly multivariate and multi-response calibration: Application to IR spectroscopy, Chemometrics Intellig. Lab. Syst., 106 (2011) 216-223.

[11] E. Vigneau, F. Thomas, Model calibration and feature selection for orange juice authentication by 1H NMR spectroscopy, Chemometrics Intellig. Lab. Syst., 117 (2012) 22-30.

[12] A.J. Myles, S.D. Brown, Decision pathway modeling, Journal of Chemometrics: A Journal of the Chemometrics Society, 18 (2004) 286-293.

[13] T. Arnalds, J. McElhinney, T. Fearn, G. Downey, A Hierarchical Discriminant Analysis for Species Identification in Raw Meat by Visible and near Infrared Spectroscopy, J. Near Infrared Spectrosc., 12 (2004) 183-188.

[14] K. Edwards, M. Manley, L.C. Hoffman, A. Beganovic, C.G. Kirchler, C.W. Huck, P.J. Williams, Differentiation of South African Game Meat Using Near-Infrared (NIR) Spectroscopy and Hierarchical Modelling, Molecules, 25 (2020) 1845.

[15] Regulations relating to the grading, packing and marking of maize intended for sale in the Republic of South Africa, in: D.o. Agriculture (Ed.) Agricultural Product Standards Act., 2009.

[16] K. Esbensen, P. Geladi, Strategy of multivariate image analysis (MIA), Chemometrics Intellig. Lab. Syst., 7 (1989) 67-86.

[17] A. Biancolillo, F. Marini, J.M. Roger, SO-CovSel: A novel method for variable selection in a multiblock framework, J. Chemometrics, 34 (2020) e3120.

[18] R.J. Barnes, M.S. Dhanoa, S.J. Lister, Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra, Appl. Spectrosc., 43 (1989) 772-777.

[19] A. Savitzky, M.J. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem., 36 (1964) 1627-1639.

[20] B. Buckner, T.L. Kelson, D.S. Robertson, Cloning of the y1 Locus of Maize, a Gene Involved in the Biosynthesis of Carotenoids, The Plant Cell, 2 (1990) 867-876.

[21] S. Sen Gupta, M. Ghosh, In Vitro Antioxidative Evaluation of-and-Carotene, Isolated from Crude Palm Oil, Journal of Analytical Methods in Chemistry, 2013.

[22] E.-S.M. Abdel-Aal, J.C. Young, I. Rabalski, Anthocyanin composition in black, blue, pink, purple, and red cereal grains, J. Agric. Food Chem., 54 (2006) 4696-4704.

[23] L.S. Galvão, Í. Vitorello, R. Almeida Filho, Effects of band positioning and bandwidth on NDVI measurements of tropical savannas, Remote Sens. Environ., 67 (1999) 181-193.

[24] M. Golic, K. Walsh, P. Lawson, Short-wavelength near-infrared spectra of sucrose, glucose, and fructose with respect to sugar concentration and temperature, Appl. Spectrosc., 57 (2003) 139-145.

[25] Metrohm, A Technology Comparison of Near-Infrared Spectroscopy. AZoM., 2019.

[26] G. ElMasry, N. Wang, C. Vigneault, J. Qiao, A. ElSayed, Early detection of apple bruises on different background colors using hyperspectral imaging, LWT-Food Science and Technology, 41 (2008) 337-345.

[27] B.G. Osborne, T. Fearn, P.H. Hindle, Practical NIR Spectroscopy with Applications in Food and Beverage Analysis, 2nd ed., Longman Scientific & Technical, Essex, England, 1993.

[28] L.G. Weyer, Near-infrared spectroscopy of organic substances, Applied Spectroscopy Reviews, 21 (1985) 1-43.

[29] B. Pérez-Vich, L. Velasco, J. Fernández-Martínez, Determination of seed oil content and fatty acid composition in sunflower through the analysis of intact seeds, husked seeds, meal and oil by near-infrared reflectance spectroscopy, J. Am. Oil Chem. Soc., 75 (1998) 547-555.

[30] Food, N. Agriculture Organization of the United, Maize in human nutrition, Food and Agriculture Organization of the United Nations, Rome, 1992.

[31] D. Lorente, N. Aleixos, J. Gómez-Sanchis, S. Cubero, O.L. García-Navarrete, J. Blasco, Recent advances and applications of hyperspectral imaging for fruit and vegetable quality assessment, Food. Bioprocess. Technol., 5 (2012) 1121-1142.

[32] Maize Crop Quality Report 2017-2018, The Southern African Grain Laboratory NPC, Pretoria, South Africa, 2018, pp. 117.

Figure 1 Digital image of all sample classes: (a) sound; (b) *Fusarium*; (c) *Diplodia*; (d) heat; (e) water; (f) frost; (g) pest (rodent); (h) pest (insect); (i) sprouted; (j) immature; (k) pinked; (l) screenings; (m) yellow; (n) sorghum; (o) soy; (p) wheat; (q) sunflower; and (r) plant material.

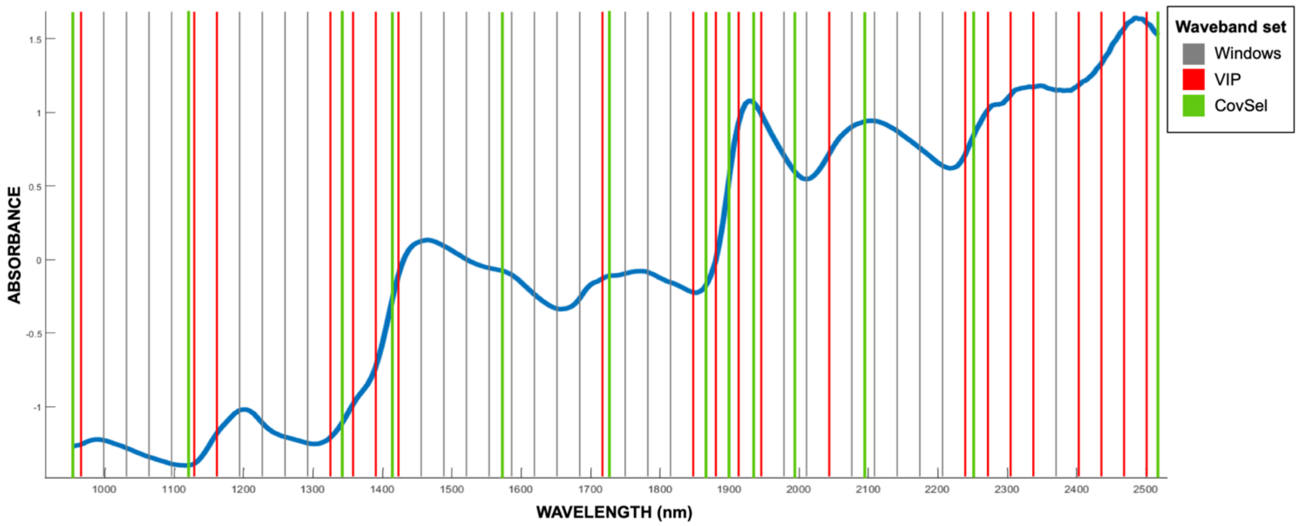

Figure 2 The pre-processed mean spectrum (SNV transformation) with the windows (grey and red), VIP (red) and CovSel (green) waveband sets indicated.

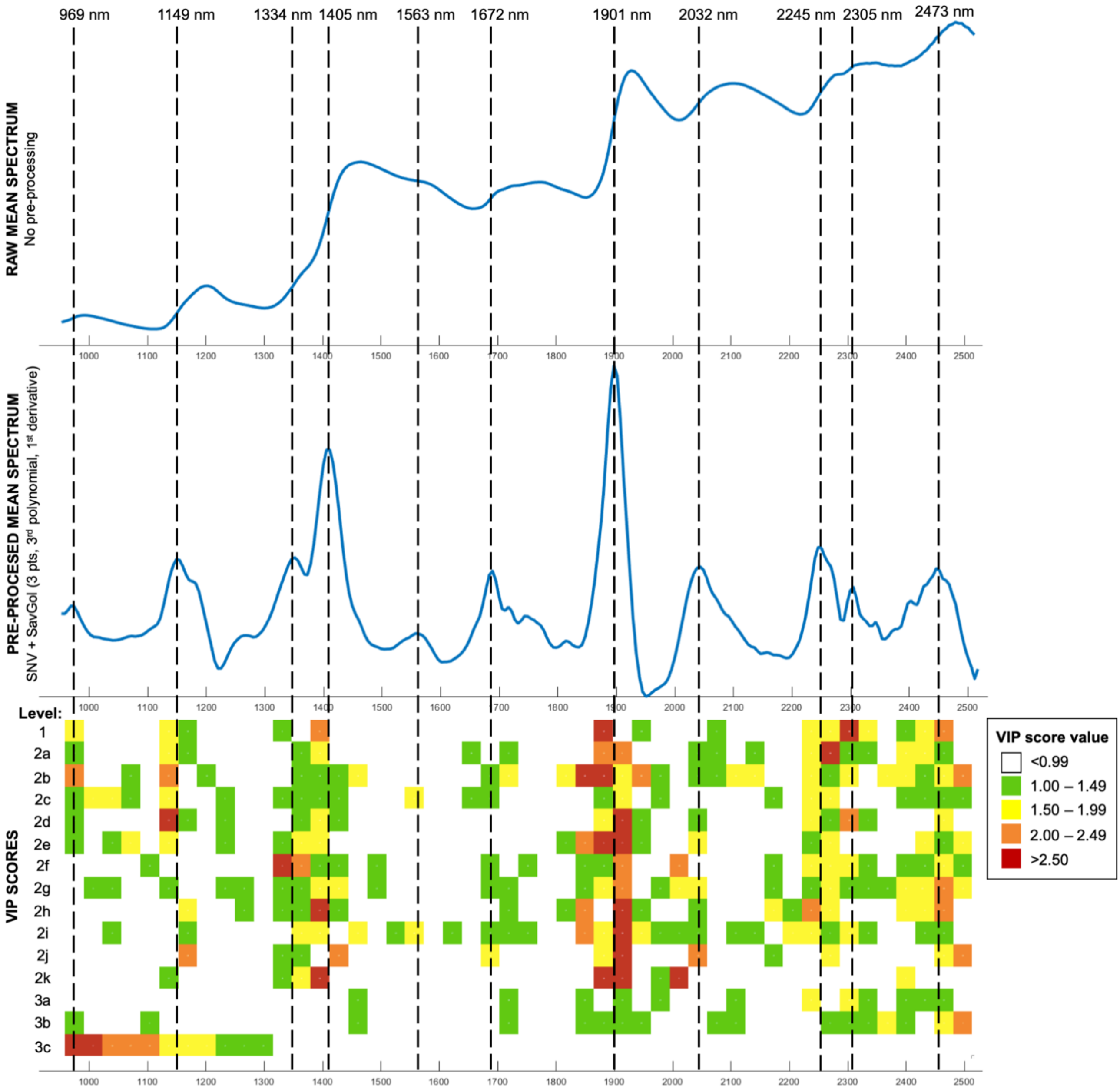

Figure 3 The raw mean spectrum of all 1044 calibration samples (top); the pre-processed mean spectrum with Savitzky-Golay (7 smoothing points; 3rd order polynomial; 1st derivative), SNV and mean-centring transformations (middle); and the VIP scores for 48 waveband groups (6 wavebands per group) in the PLS-DA models in classification models.

Table 1 White maize grading main categories and sub-categories, with their shorthand names used in this article, the maximum allowed levels for the best white maize grade (WM1) and the summarised levels for the 2017/2018 harvest season [only main categories are reported according regulations].

Main category	Sub-categories	Shorthand	Max. level (WM1)	Average level observed 2017/2018 harvest [32]
Sound white maize	-	Sound	N/A	
Defective white maize	Broken kernels (screenings)	Screenings		
	Fusarium fungal damage	*Fusarium*		
	Diplodia fungal damage	*Diplodia*		
	Heat damage	Heat		
	Water damage	Water	7%	4.4%
	Frost damage	Frost		
	Pest damage (rodent and insect)	Pest		
	Sprouted kernels	Sprouted		
	Immature kernels	Immature		
Pinked white maize	-	Pinked	12%	0.4%
Yellow maize	-	Yellow	3%	0.3%
Foreign materials	Soy	Soy		
	Sorghum	Sorghum		
	Sunflower seeds	Sunflower	0.3%	0.1%
	Wheat	Wheat		
	Plant material	Plant		

Table 2 Validation results for the classification of 1044 samples using the hierarchical models based on the full spectrum (288 wavebands). Sub-category classification indicates classification as the true class only and main category classification as either the true class or a class in the correct main category (where applicable)

Main category	Sub-categories	Sub-category classification accuracy	Main category classification accuracy
Sound white maize	-	**88.3%**	**88.3%**
Defective white maize	Average	**60.0%**	**93.3%**
	Screenings	86.7%	93.3%
	Fusarium	95.0%	100%
	Diplodia	65.0%	90.0%
	Heat	91.7%	95.0%
	Water	23.3%	90.0%
	Frost	45.0%	96.7%
	Pest	40.0%	95.8%
	Sprouted	13.3%	86.7%
	Immature	79.6%	92.6%
Pinked white maize	-	**83.3%**	**83.3%**
Yellow maize	-	**75.0%**	**75.0%**
Foreign materials	Average	**99.3%**	**100%**
	Soy	100%	100%
	Sorghum	100%	100%
	Sunflower	100%	100%
	Wheat	100%	100%
	Plant	96.7%	100%
OVERALL		**75.5%**	**93.3%**

Table 3 Validation results for the classification (main category) of 1044 samples using the hierarchical models based on the full spectrum (288 wavebands), windows wavebands (48), VIP wavebands (21) and CovSel wavebands (13)

Main category	Sub-categories	Full (288)	Windows (48)	VIP (21)	CovSel (13)
Sound white maize	-	**88.3%**	**76.7%**	**78.3%**	**63.3%**
Defective white maize	Average	**93.3%**	**86.7%**	**83.2%**	**81.0%**
	Screenings	93.3%	78.3%	78.3%	73.3%
	Fusarium	100%	88.3%	86.7%	83.3%
	Diplodia	90.0%	90.0%	71.7%	78.3%
	Heat	95.0%	76.7%	76.7%	61.7%
	Water	90.0%	88.3%	81.7%	76.7%
	Frost	96.7%	90.0%	86.7%	85.0%
	Pest	95.8%	93.3%	82.5%	88.3%
	Sprouted	86.7%	90.0%	86.7%	93.3%
	Immature	92.6%	85.2%	98.1%	88.9%
Pinked white maize	-	**83.3%**	**75.0%**	**78.3%**	**80.0%**
Yellow maize	-	**75.0%**	**60.0%**	**56.7%**	**48.3%**
Foreign materials	Average	**100%**	**96.7%**	**95.7%**	**94.32%**
	Soy	100%	100%	100%	100.0%
	Sorghum	100%	98.3%	100%	98.3%
	Sunflower	100%	100%	100%	100%
	Wheat	100%	85.0%	83.3%	83.3%
	Plant	100%	100%	95.0%	90.0%
OVERALL		**93.3%**	**87.1%**	**84.5%**	**81.9%**

Declaration of interests

▯ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

CRediT author statement

Kate Sendin - Writing – conceptualization; Original Draft; formal analysis, validation, methodology; investigation

Marena Manley – Supervision; validation; Writing - Review & Editing

Federico Marini - Writing - Review & Editing; software

Paul J Williams – conceptualization; Supervision; validation; Writing - Review & Editing; Project administration; funding acquisition

Supplementary Information:

Hierarchical classification pathway for white maize, defect and foreign material classification using spectral imaging

Kate Sendin[1], Marena Manley[1], Federico Marini[1,2] & Paul J. Williams

Architecture of the hierarchical classification pathway

The structure of the hierarchical model is given in Table S1. An object starts at Level 1 and is classified at each level and sub-level as one of the two classes. According to this classification, it follows the relevant 'Proceed to' instruction to subsequent levels until reaching a 'Final classification' instruction.

The classification pathway begins at Level 1, where all maize classes and all foreign materials were separated based on a single PLS-DA latent variable (LV) ($Q^2 = 0.86$). If an object was classified as a maize kernel, it proceeded to classification in Level 2. If it was a foreign material, it proceeded to Level 3.

Level 2 was the most challenging section of the hierarchical model. Twelve closely related classes had to be separated sequentially, which included the main categories sound maize, yellow maize, pinked maize and defective maize (9 sub-categories). By calculating models of one class vs. grouped class of all remaining classes, the order was determined as follows: 2a – Screenings ($Q^2 = 0.74$); 2b – heat damage ($Q^2 = 0.55$); 2c – *Fusarium* fungal damage ($Q^2 = 0.68$); 2d – immature kernels ($Q^2 = 0.75$); 2e – water damage ($Q^2 = 0.50$); 2f – *Diplodia* fungal damage ($Q^2 = 0.61$); 2g – yellow maize ($Q^2 = 0.68$); 2h – sound white maize ($Q^2 = 0.81$); 2i – sprouted kernels ($Q^2 = 0.76$); 2j; frost damage ($Q^2 = 0.76$); 2k – pinked white maize and pest damage ($Q^2 = 0.75$).

The classification of foreign materials in Level 3 was less challenging, as this was separating different commodities, not classes of a single commodity. The previous findings of Sendin et al. (2019) revealed that the spectral signature of plant materials and sunflower seeds lacked absorbance by starch, leading to easy differentiation from wheat, soy and sorghum [1]. Instead, plant materials and sunflower seeds were characterised by a cellulose-rich surface chemistry. Thus, foreign materials were first separated as cellulose-rich vs. starchy. A two-way PLS-DA model was calculated for plant material vs. sunflower seeds ($Q^2 = 0.89$). Only one LV was required, as the model error increased with the addition of LVs. Due to sufficient differences between wheat, soy and sorghum, a step-wise approach was not necessary, and a three-way PLS-DA was calculated ($Q^2 = 0.95$).

The secondary classification step was introduced to minimise misclassification between closely related classes. Due to the large number of classes used, a clear separation between each class was not expected. During hierarchical model development, the classification results of the one class vs. all remaining classes were examined. If a large number of the misclassifications (*ca.* 5+) were due to confusion with a specific class, a secondary step was included. If only one or two misclassifications were due to a specific class, the step was not included in order to avoid overfitting. Using similar classes screenings (broken kernels) and rodent damage (bitten kernels) as an illustration, the classification of screenings occurs early in the hierarchical model (Level 2a) when many classes remain. As rodent damage had not yet been classified (Level 2k), all rodent damage kernels should be classified in the group class and continue to subsequent classification steps. However, many were misclassified as screenings, and thus did not continue to the following steps. As a corrective measure, all objects classified as screenings were predicted by a second two-way PLS-DA model of screenings vs. rodent damage, where the result of this secondary step is taken as the final classification result. The secondary step classification models had higher Q^2 values and excellent cross-validated classification accuracies (often 98 – 100%), and single classes were well-defined and easily separated. The number of errors was greatly reduced by including this step.

To illustrate how a kernel ideally flows through the hierarchical model decision pathway from beginning to final classification, a heat damaged kernel is used as an example (see Table S1):

1. Level 1: Classified as the class 'Group: sound, all defects, pinked & yellow', where the instruction 'Proceed to Level 2' is given.
2. Level 2a: Model of screenings vs. grouped class (heat, *Fusarium*, immature, water, *Diplodia*, yellow, sound, sprouted, frost, pinked & pest) classified the kernel as the grouped class, where the instruction 'Proceed to Level 2b' is given.
3. Level 2b: Model of heat vs. grouped class (*Fusarium*, immature, water, *Diplodia*, yellow, sound, sprouted, frost, pinked & pest) classified the kernel as heat damaged, where the instruction 'Proceed to 2nd classification' is given
4. Second classification step: Model of heat damage vs. yellow maize classified the kernel as heat damage, giving a final classification of 'heat damage'.

References

[1] K. Sendin, M. Manley, V. Baeten, J.A. Fernández Pierna, P.J. Williams, Near Infrared Hyperspectral Imaging for White Maize Classification According to Grading Regulations, Food Anal. Meth., 12 (2019) 1612-1624.

Table S1 Full spectrum hierarchical model structure, consisting of 3 main levels and 15 sub-levels, with a total of 25 PLS-DA classification models. Each object enters the decision pathway at Level 1 and follows the relevant instructions according to classification result by the PLS-DA model.

	CLASS ONE	CLASS TWO	CLASS THREE	2nd CLASSIFICATION	LVs; Q^2
LEVEL 1: MAIZE vs. FOREIGN MATERIALS					
1	Group: sound, all defects, pinked & yellow *Proceed to LEVEL 2*	Group: soy, sorghum, sunflower, wheat & plant *Proceed to LEVEL 3*	-	-	1: 7; 0.861
LEVEL 2: MAIZE CATEGORIES & SUBCATEGORIES					
2a	Screenings *Proceed to 2nd classification step*	Group: heat, *Fusarium*, immature, water, *Diplodia*, yellow, sound, sprouted, frost, pinked & pest *Proceed to 2b*	-	Screenings vs. pest (rodent) *Final classification = predicted class*	2a: 12; 0.744 2nd: 4; 0.810
2b	Heat *Proceed to 2nd classification step*	Group: *Fusarium*, immature, water, *Diplodia*, yellow, sound, sprouted, frost, pinked & pest *Proceed to 2c*	-	Heat vs. yellow *Final classification = predicted class*	2b: 16; 0.554 2nd: 7; 0.847
2c	*Fusarium* *Final classification =* 'Fusarium'	Group: immature, water, *Diplodia*, yellow, sound, sprouted, frost, pinked & pest *Proceed to 2d*	-	-	2c: 11; 0.678
2d	Immature *Proceed to 2nd classification*	Group: water, *Diplodia*, yellow, sound, sprouted, frost, pinked & pest *Proceed to 2e*	-	Immature vs. water *Final classification = predicted class*	2d: 10; 0.754 2nd: 8; 0.704
2e	Water *Proceed to 2nd classification*	Group: *Diplodia*, yellow, sound, sprouted, frost, pinked & pest *Proceed to 2f*	-	Water vs. yellow *Final classification = predicted class*	2e: 7; 0.501 2nd: 7; 0.938
2f	Diplodia *Final classification=* 'Diplodia'	Group: yellow, sound, sprouted, frost, pinked & pest *Proceed to 2g*	-	-	2f: 12; 0.612

2g	Yellow *Proceed to 2nd classification*	Group: sound, sprouted, frost , pinked & pest *Proceed to 2h*	-	Yellow vs. heat *Final classification = predicted class*	2g: 14; 0.678 2nd: 7; 0.847
2h	Sound *Final classification = 'Sound'*	Group: sprouted, frost, pinked & pest *Proceed to 2i*	-	-	2h: 14; 0.809
2i	Sprouted *Proceed to 2nd classification*	Group: frost, pinked & pest *Proceed to 2j*	-	Sprouted vs. water *Final classification = predicted class*	2i: 15; 0.757 2nd: 5; 0.867
2j	Frost *Proceed to 2nd classification*	Group: pinked & pest *Proceed to 2k*	-	Frost vs. water *Final classification = predicted class*	2j: 11; 0.758 2nd: 6; 0.835
2k	Pinked *Proceed to 2nd classification (1)*	Pest *Proceed to 2nd classification (2)*	-	(1)Pinked vs. sound (2) Pest vs. *Diplodia* *Final classification = predicted class*	2k: 4; 0.751 2nd (1): 11; 0.948 2nd (2): 16; 0.856
LEVEL 3: STARCHY vs. CELLULOSE-RICH FOREIGN MATERIALS					
3a	Soy, sorghum & wheat *Proceed to 3b*	Sunflower & plant *Proceed to 3c*	-	-	3a: 7; 0.968
3b	Soy *Final classification = 'Soy'*	Sorghum *Final classification = 'Sorghum'*	Wheat *Final classification = 'Wheat'*	-	3b: 6; 0.954
3c	Sunflower *Final classification = 'Sunflower'*	Plant *Proceed to 2nd classification*	-	Plant vs. screenings *Final classification = predicted class*	3c: 1; 0.893 2nd: 8; 0.913