Peer Reviewed Paper openaccess Paper Presented at IASIM 2016, July 2016, Chamonix, France

Application of calibrations to hyperspectral images of food grains: example for wheat falling number

Nicola Caporaso,[a,b,*] Martin B. Whitworth[a] and Ian D. Fisk[b]

[a]Campden BRI, Station Road, Chipping Campden, Gloucestershire, GL55 6LD, UK. E-mail: nicola.caporaso@campdenbri.co.uk
[b]Division of Food Sciences, The University of Nottingham, Sutton Bonington Campus, Loughborough LE12 5RD, UK

The presence of a few kernels with sprouting problems in a batch of wheat can result in enzymatic activity sufficient to compromise flour functionality and bread quality. This is commonly assessed using the Hagberg Falling Number (HFN) method, which is a batch analysis. Hyperspectral imaging (HSI) can provide analysis at the single grain level with potential for improved performance. The present paper deals with the development and application of calibrations obtained using an HSI system working in the near infrared (NIR) region (~900–2500 nm) and reference measurements of HFN. A partial least squares regression calibration has been built using 425 wheat samples with a HFN range of 62–318 s, including field and laboratory pre-germinated samples placed under wet conditions. Two different approaches were tested to apply calibrations: i) application of the calibration to each pixel, followed by calculation of the average of the resulting values for each object (kernel); ii) calculation of the average spectrum for each object, followed by application of the calibration to the mean spectrum. The calibration performance achieved for HFN (R^2 = 0.6; *RMSEC* ~ 50 s; *RMSEP* ~ 63 s) compares favourably with other studies using NIR spectroscopy. Linear spectral pre-treatments lead to similar results when applying the two methods, while non-linear treatments such as standard normal variate showed obvious differences between these approaches. A classification model based on linear discriminant analysis (LDA) was also applied to segregate wheat kernels into low (<250 s) and high (>250 s) HFN groups. LDA correctly classified 86.4% of the samples, with a classification accuracy of 97.9% when using an HFN threshold of 150 s. These results are promising in terms of wheat quality assessment using a rapid and non-destructive technique which is able to analyse wheat properties on a single-kernel basis, and to classify samples as acceptable or unacceptable for flour production.

Keywords: hyperspectral imaging, partial least squares calibration, food grains, whole wheat kernels, Hagberg Falling Number (HFN)

Introduction

Hyperspectral imaging

Hyperspectral imaging (HSI) applies spectroscopy coupled with imaging, to give a so-called "hypercube". When used in the near infrared (NIR) wavelength region, this is particularly useful to explore both the chemical composition of a material and the spatial distribution of composition within a sample. However, in comparison to traditional NIR spectroscopy, HSI places additional requirements on data management and processing, which is of paramount importance to extract useful chemical information for samples.[1]

The large amount of recorded data should be processed in order to avoid redundant information but without affecting the useful data. Factors affecting HSI images

Correspondence
N. Caporaso (nicola.caporaso@campdenbri.co.uk)

Received: 9 August 2016
Revised: 3 April 2017
Accepted: 12 April 2017
Publication: 28 April 2017
doi: 10.1255/jsi.2017.a4
ISSN: 2040-4565

Citation
N. Caporaso, M.B. Whitworth and I.D. Fisk, "Application of calibrations to hyperspectral images of food grains: example for wheat falling number", *J. Spectral Imaging* **6,** a4 (2017). https://doi.org/10.1255/jsi.2017.a4

include the presence of bad pixels, noise and spatial and spectral non-uniformity of illumination and detector response for which normalisation procedures are essential.[2] Processing of hypercubes is based on several steps, which have been detailed in previous review papers.[3] For example, a relevant aspect is the pre-processing methods applied to the spectral data to remove noise or to correct for other optical effects before the application of chemometric modelling.[4]

HSI can be used for unsupervised or supervised classification, or for prediction of quantitative properties. Whereas several papers report on the subject of pixel classification based on PCA, sometimes it is difficult to attribute different clusters to a particular compound or cause. On the contrary, supervised methods are based on reference chemical (or physical) measurements and then a statistical model is built as a calibration for the prediction or classification of unknown samples. Applications of HSI in food science have spread widely during the past few years due to its capability for high speed, non-destructive inspection for on-line quality control. The extraction of useful information from hyperspectral images, the application of effective chemometric techniques and the correct data management are crucial steps to build and then apply calibrations for the prediction of food chemical constituents.

Wheat quality criteria

Wheat quality criteria include protein content and Hagberg Falling Number (HFN). The latter parameter is an indirect measurement of the α-amylase activity, which is very relevant for flour functionality in terms of the dough rheology and quality of baked products, especially bread.[5] This enzyme hydrolyses alpha bonds of starch, yielding glucose and maltose. In germinating grain, α-amylase is synthesised by the scutellum and later by the adjacent aleurone,[6] and its activity is affected by germination time and environmental temperature.[7] Especially in wet conditions before the harvesting period or in problematic years, the α-amylase may be activated prematurely and has deleterious effects on the functional properties of the resulting flour. Pre-harvest sprouting has been reported as one of the most important factors for wheat quality, because it can lead to significant economic losses in a short time just before the wheat harvest.[8] The intensity of the damage depends on the climatic conditions and on the susceptibility of particular cultivars to sprout damage. Direct or indirect methods for α-amylase measurement are commonly applied, usually limited to bulk measurements only. HFN is the officially accepted method to assess the quality of wheat and flour related to the α-amylase activity (ICC standard method no. 107/1, 1968).[9]

HFN measurement

The HFN method is based on measurement of the time for a plunger to fall through a wholemeal paste in hot water after mixing.[9] High α-amylase activity causes starch to break down under the conditions of the test, reducing the viscosity of the paste and reducing the time for the plunger to fall. The minimum HFN value is 60 s, this being the initial mixing time, while good bread-making wheats can have values higher than 400 s. HFN values above 250 s, 300 s or in some cases 350 s are required for high-quality grading depending on the receival standards set by the wheat industries in different countries.[10]

It has been highlighted that the use of tools to rapidly analyse sprouted and unsprouted kernels would be beneficial in breeding programmes.[11] Previous researchers compared several methods commonly used to estimate sprout damage in wheat, and a non-linear relationship between HFN and percentage of sprouted kernels was reported, while the Perten liquefaction number, 6000/(HFN-50) results in a linear relationship.[12] Barnard, Van Deventer and Maartens[13] highlighted that HFN does not reflect entirely the activity of α-amylase, as it is influenced by other factors such as starch and fibre. The coefficient of determination between HFN and α-amylase activity was reported to be $R^2 = 0.848$.[13] Moreover, it should be noticed that the presence of a few kernels with very high α-amylase activity, i.e. low HFN, have a large effect on the overall sample.

HFN is relatively time-consuming and requires preparation of a ground sample, an accompanying analysis of moisture content and duplicate analyses. Alternative analytical methods have been investigated, e.g. enzymatic kits for α-amylase activity measurement,[14] antibody tests[15] and spectroscopic methods.[16–18] Despite the development of antibody-based tests for the measurement of α-amylase activity, the traditional HFN method is still regarded as the preferred one because it combines ease of use, reliability of results and wide use throughout the wheat and milling industry, especially for quality control and screening purposes.

HFN requires analysis of a ground sample and cannot give information on single wheat kernels. This might be

useful as a small portion of sprouted grains may be sufficient to drop the HFN below the acceptable quality limits, and the selection and screening of the best lines resistant to specific environmental conditions is one of the most important parameters for breeding programmes, due to the different response of genotypes.[8,19] In fact, it has been reported that screening technologies are needed to identify cases of late maturity α-amylase, which causes low HFN in wheat in the absence of preharvest sprouting, and different genotypes show different degrees of sprout resistance.[10]

Non-destructive analysis using NIR spectroscopy

NIR spectroscopy is rapid, non-destructive and can be applied to whole kernels, and an effective method would therefore be beneficial for analysis of commercial samples to identify problem wheat kernels. Also, breeders are seeking non-destructive methods for screening of kernels to assess sprouting resistance, either at the single kernel level to enable favourable grains to be selected for breeding or because limited sample quantity is available. The application of NIR spectroscopy has been evaluated to estimate sprouting problems in wheat.[16,20] A screening system has been recently described for *in situ* monitoring of wheat quality parameters, and the HFN was predicted with good results, i.e. standard error of cross-validation (*SECV*) of 31 s and $R^2 = 0.70$, using a partial least squares (PLS) calibration model based on bulk NIR reflectance spectroscopy.[21]

HSI technology enables the study of single kernels, instead of just bulk measurements. Many kernels can be imaged in the same hypercube and calibrations applied individually to each kernel. Koç, Smail and Wetzel[22] applied HSI to study sprout damage in wheat, using six wheat varieties with artificially induced sprouting achieved by exposing kernels to high moisture up to 36 h. A classification was tested according to the level of sprouting, and kernels were analysed by principal component analysis (PCA) on a single pixel level within the kernels. Better correlations with germination time were achieved for the embryo region of the kernel than for other regions, but no clear classification of sprouted/unsprouted grains was possible.

HSI in the NIR region has also been applied to predict α-amylase activity in single wheat kernels, using enzymatic kits as the reference method.[20,23] Xing *et al.*[20] explored the possibility of predicting α-amylase activity by HSI in the NIR region for individual Canadian Western Red Spring wheat kernels, in comparison to a Fourier transform (FT)-NIR instrument. The calibration performance was always lower for HSI, being $R^2 = 0.72–0.79$ for the HSI instrument and 0.74–0.82 for FT-NIR. Logarithmic transformation of the reference α-amylase activity value slightly improved the model. However, all the models reported tended to overestimate the enzyme activity at low values and underestimate at high values. In a previous paper, it was reported that the performance of the PLS models to predict the enzymatic activity had R^2 values of 0.54 and 0.73 for Canada Western Amber Durum and Canada Western Red Spring, respectively.[23]

A number of papers have reported on the challenges of developing accurate single kernel calibrations for wheat quality parameters, due to the difficulty of obtaining reference data for single kernels, as many analytical methods require larger amounts of sample. For this reason, a potential approach is to develop calibrations using reference measurements for bulk samples, and to apply them to single kernels.

A further consideration for HSI is the method used for data treatment, in particular whether the hypercube should be treated on a single pixel basis or whether the objects of interest in the hypercubes should be previously selected so that the calibration is applied on the mean spectrum for each object, and whether these different approaches lead to different prediction results.

Therefore, the aims of the present paper were: a) to build a PLS calibration for the prediction and classification of HFN in intact wheat kernels using HSI in the NIR region; b) to compare two approaches for the application of calibrations to HSI images of wheat kernels.

Materials and methods

Wheat samples

Wheat samples cultivated in the United Kingdom were used in the present experiment, including examples of a wide range of varieties (>30) and from multiple growing areas within the UK. The total number of samples was 425, as described in Table 1. These were composed of two sets of samples. The first set comprised 309 samples from a field test involving late harvesting to assess HFN problems, using two bread wheat varieties (Crusoe and Skyfall) grown in six locations across the UK. The second set comprised 116 samples from 30

wheat batches including the most common UK varieties (e.g. Chilton, Claire, Cordial, Crusoe, Einsten, Gallant, Invicta, JB Diego, Mulika, Panorama, Solstice), and examples of a German and a French wheat. These samples were treated in the laboratory to reach a pre-germination step which increases the α-amylase activity, but avoids obvious sprouting. This was achieved by soaking the samples in distilled water for 30 min and subsequently spreading them on plastic trays, covered with wet paper tissue to maintain high relative humidity and stored at 20°C. Sub-samples of *c.* 500 g were taken at discrete intervals (typically 3 h, 9 h, 12 h) during the germination phase up to *c.* 30 h and finally dried at 30°C in an oven (Mitchell dryers, UK), which took approximately 6–8 h. The moisture content of the samples was periodically checked using a SINAR moisture analyser until it was below 12–13%. An aliquot of each sample was analysed for HFN as described below. Previous workers have applied freeze-drying to samples to stop the germination process. However, it was considered that freezing the wet grains may affect the internal microstructure of the kernels and could affect the NIR spectra, and oven drying at mild temperature conditions was therefore used. This is more consistent with the approach used to process commercial wheat samples with excess moisture content.

Falling Number measurement

HFN was used as the reference method, according to the international standardised method.[9] 300 g of each sample was milled using a Laboratory Mill LM 3100 (Perten Instruments, Hägersten, Sweden) with a 0.8 mm sieve to obtain a consistent particle size. The moisture content of ground wheat was checked using a PerCon Inframatic 8611 NIR spectrometer (Perten Instruments) before the HFN measurement to define the amount of water to be added. 7 ± 0.05 g ground wheat were put into a viscometer tube and 25 ± 0.2 mL of distilled water was added, in relation to a previous moisture determination. A Perten FN1700 instrument was used for the analysis, which was done in duplicate.

HSI data acquisition

Hyperspectral images were acquired using a pushbroom approach, using a system described by Millar, Whitworth, Chau and Gilchrist.[24] The instrument was supplied by Gilden Photonics Ltd (Glasgow, UK) and includes a SWIR spectral camera (Specim Ltd, Oulu, Finland) containing a cooled 14-bit 320 × 256 pixel HgCdTe detector and N25E spectrograph, providing 256 spectral bands over a wavelength range of ~990–2500 nm, with a spectral resolution of about 6 nm. Samples were presented on a moveable stage and illuminated with two 500 W incandescent line lamps mounted either side of the lens at an angle of incidence of approximately 45°. The lamps were only turned on for the few seconds required to acquire the image, resulting in minimal heating of the samples. Samples were imaged using a 31 mm focal length lens at a distance of 220 mm, providing a swathe of 35 mm and a pixel size of 0.109 mm for 320 spatial pixels. Images were acquired at a rate of 100 frames s^{-1}, using a stage translation speed of 10.9 mm s^{-1}, providing the same spatial resolution parallel and perpendicular to the scan direction. SpectralCube 3.0041 software (Specim) was used to control the camera and translation stage. The camera shutter was automatically closed for 1 s at the end of each scan and approximately 100 frames were recorded to establish the baseline signal of the detector (black reference) for all pixels. For each batch of samples, approximately 100 frames were also recorded for a white PTFE reference material with approximately 100% reflectance across the entire measured spectral range (white reference). Different exposure times were used for samples and for the white reference to provide optimal dynamic range for each, with black references recorded for each. Thirty wheat kernels were randomly sampled from each batch for the HSI data acquisition. For imaging, the kernels were placed on a black high density polyethylene tray in random orientations, randomly crease-up or crease-down and separated so they did not touch.

HSI data analysis and kernel segmentation

Hyperspectral images were analysed using ENVI-IDL 5.2 software (Harris Geospatial Solutions, Virginia, USA). The first step to extract useful chemical information from the hypercube was to calculate the $\log(1/R)$ values starting from the raw data. Reflectance values, $R(x,y,\lambda)$ were calculated for each pixel (spatial coordinates x,y, wavelength λ) by subtracting the black reference value for the corresponding detector pixel (x,λ) and normalising by the corresponding white reference value and by the ratio of exposure times, assuming a linear response. Absorbance values, $A = \log_{10}(1/R)$, were then calculated for each pixel.

The next step of HSI analysis was removal of spikes. These can be caused by bad pixels in the detector, and can include both high and low signal values. Several algorithms and methods have been described in the litera-

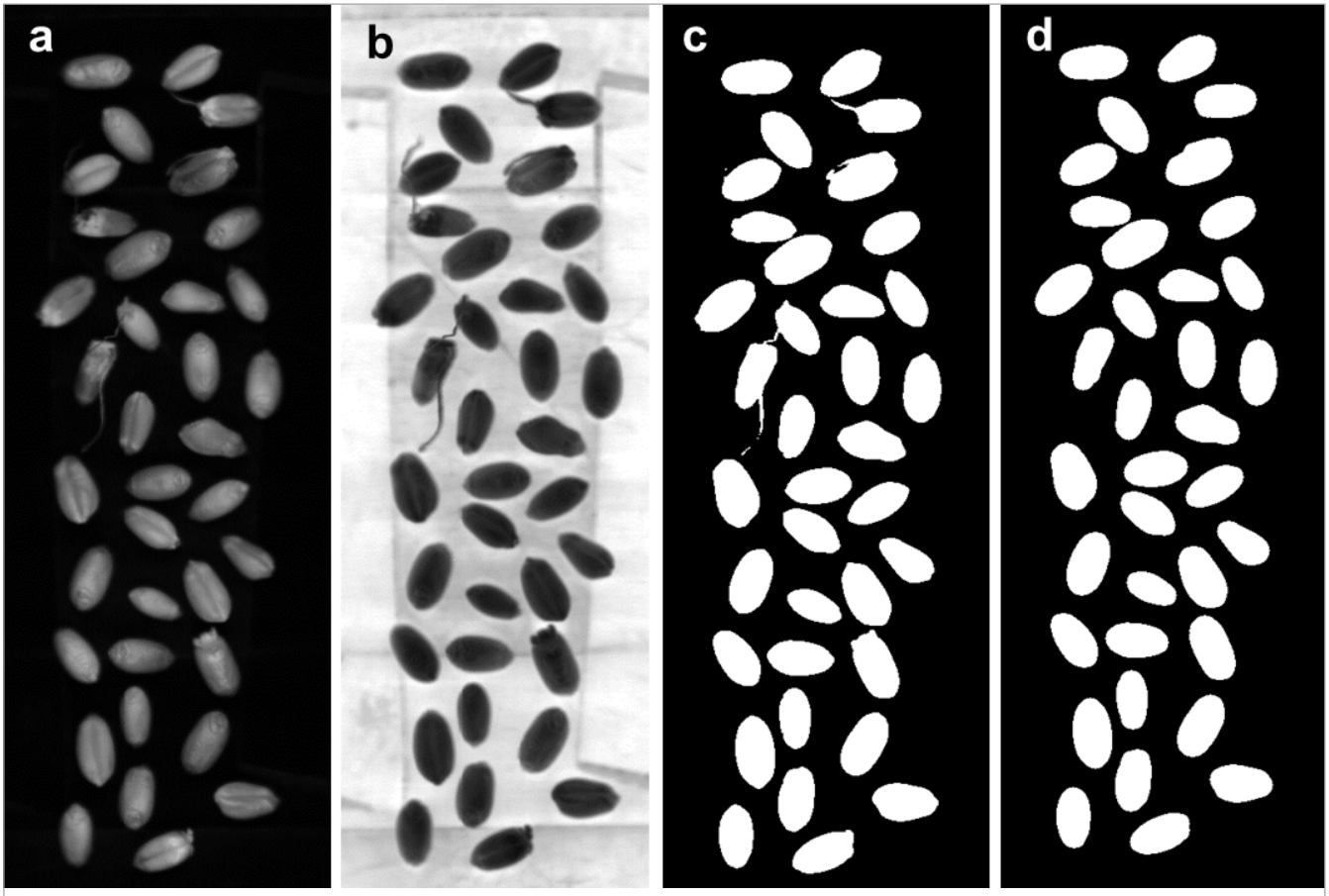

Figure 1. Wheat kernels including sprouted examples. a) reflectance image at 1186 nm; b) log(1/*R*); c) binary thresholding; d) morphological processing for kernel segmentation.

ture to identify and remove spikes from hypercubes.[2,25] Appropriate threshold selection is crucial to effectively identify spikes, sometimes using basic statistics as spikes have usually larger deviation from the normal mean spectrum. The spike removal can be then applied by substituting the specific point by the median of the neighbouring spectra.[2] In this study, an enhancement of this approach was used, in which each image pixel was compared with the median of neighbouring pixels, and a weighted average was calculated, with greater weighting towards the median value if the difference is large.

Following spike removal, images were segmented to identify the grains. Figure 1 shows an example. Figure 1a shows reflectance data for the 1186 nm band after spike removal. Figure 1b shows log(1/*R*) data for the same band. To segment grains, a binary threshold was first applied to this band to select pixels with $\log(1/R_{1186}) < 1$ equivalent to reflectance, $R_{1186} > 0.1$, as shown in Figure 1c. A binary opening operation was then applied with a circular kernel with a radius of 10 pixels (1.09 mm) to remove noise and small projections, including sprouts. The resulting image, shown in Figure 1d, was used to segment kernels for spectral analysis. A flat uniform background, both spatially and spectrally, is generally recommended to aid segmentation of samples from the background.[26] Figure 2 shows example single pixel absorbance spectra for a grain and for the tray. The chosen tray material has high log(1/*R*) values over the entire NIR spectral range, providing good contrast, and a binary threshold is therefore sufficient to achieve good segmentation.[24] 1186 nm was chosen for segmentation, as this wavelength resulted in low log(1/*R*) values for the grain regions, with good contrast and low noise. The effect of the spike removal procedure is also shown. The method is effective in removing spikes due to bad pixels, while preserving spectral bands. A spectrum is also shown for the white reference material. According to the normalisation procedure, this is defined to have a log(1/*R*) of 0 at all wavelengths.

Mean spectra were calculated for each segmented kernel and exported for separate development of

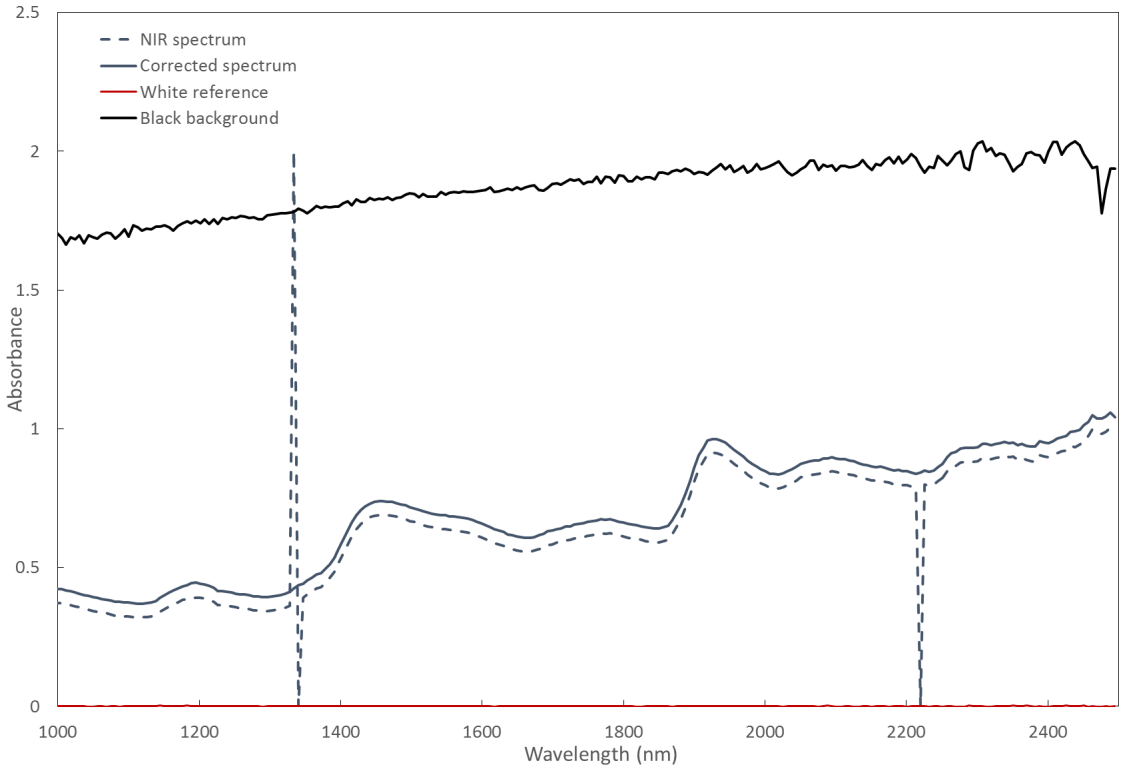

Figure 2. Example log(1/*R*) spectra for pixels corresponding to a wheat kernel before and after spike removal, and the black tray used as a background, showing good contrast at all wavelengths. An additive constant was added to the corrected spectrum for illustration purposes.

calibration models in comparison with reference HFN measurements. For the camera system used, the initial 16 spectral bands did not contain useful data and only bands from ~1100 nm to 2500 nm were used for spectral analysis, due to excessive noise in the detector in these first bands.

Statistical treatment of the data

The exported mean spectra were statistically treated using The Unscrambler X 10.3 software package (CAMO, Sweden). Calibrations were developed using partial least squares (PLS) regression on 240 spectral bands against the reference HFN, to test whether is it possible to quantitatively predict the HFN value at a single wheat kernel level. The mean spectra for each wheat kernel were exported and used as the variables to build the statistical model, while the reference HFN value was used as the response. The algorithm used was the non-linear iterative partial least squares (NIPALS). The PLS regression models were validated using an independent validation dataset and the number of principal components (PC) to use was chosen in order to minimise the root mean square error of prediction (*RMSEP*), leaving the software to choose the best PC. An initial approach was tested using the full set of single kernel spectra and attributing to them the HFN reference value for the corresponding bulk sample, but this approach was then avoided as the calibration performance was inadequate. Instead, the spectral data were averaged for each batch and the mean spectrum was used for the statistical model. Several spectral pre-treatments were tested, including standard normal variate (SNV), second derivative using Savitzky–Golay smoothing (five points smoothing, second order derivative), baseline correction, de-trending and multiplicative scatter correction (MSC).[27] On the whole dataset obtained, multivariate statistical analysis was applied to relate the spectra to the measured HFN of the wheat batches.

Linear discriminant analysis (LDA) is a classification method and was used to discriminate samples into two classes, i.e. high and low HFN value. LDA was calculated using PCA with 20 principal components, and the quadratic method was used. Both a PLS analysis and LDA were tested in order to verify i) the feasibility of HSI for

the prediction of HFN, ii) the possibility of classifying wheat kernels into low-HFN and high-HFN classes based on the reference method. The quality of the models was assessed by evaluating the regression coefficient, R^2, the errors of prediction and by using a separate validation dataset of about 30% of the full set of samples.

Results and discussion

Distribution of HFN in wheat samples

Table 1 shows the statistical description for the reference HFN measurement of wheat samples used in the present experiment. The total sample number (complete dataset) was 425, comprised of field wheat with no further treatment, and a laboratory germination dataset; the complete dataset was randomly split into a calibration dataset and a validation dataset for the PLS regression analysis.

The results of the two groups used indicate that the range of HFN values was similar between the field wheat experiment and the laboratory germination dataset, 62 s being the minimum value and 418 s and 404 s the maximum values for the first and second class, respectively. The average HFN for the dataset was slightly above 250 s, which is considered as the minimum limit for high quality bread-making wheats by the industry.

Figure 3 shows the distribution of reference HFN values for the test samples. A bimodal distribution was obtained, with 53 samples having values below 150 s and 372 samples above. The samples in the lower range are those with high α-amylase activity indicating sprouting problems, and they are desired in the dataset because the aim is to identify problematic samples, especially those with average HFN below 150 s.

PLS calibration for HFN prediction

Table 2 reports the performance of the PLS calibration for HFN prediction in the NIR region (~1100–2500 nm) for several spectral pre-treatments. The spectral pre-treatments provided improved calibration performance. The best performance was achieved for the second derivative treatment, with R^2 of 0.60 and *RMSEC* of 50 s. The R^2 values for the external validation dataset were between 0.41 and 0.47, with little improvement depending on the spectral pre-processing applied. The error was slightly above 60 s, which compares favourably with a repeatability of the reference HFN method of 8.5% for ground wheat (from 5 s to 35 s with our HFN range). The calibration with SNV pre-processing gave slightly poorer calibration performance, but has also been used in the present paper for comparison purposes to evaluate different methods for application of calibrations to HSI images.

The calibration performance compares favourably with the value of 73 s reported by Osborne[16] for an NIR spectroscopy method, with a tested range of 62 s to 377 s, and a R^2 value of 0.56. The author applied NIR spectroscopy on ground wheat and therefore only a bulk measurement was done.

Xing, Van Hung, Symons, Shahin and Hatcher[23] reported the prediction of α-amylase activity in Canadian wheats, a hard red spring and a durum variety,

Table 1. Range of reference HFN in the wheat samples used to develop PLS models.

	Calibration set						Validation set					
	Mean	Median	*SD*	Min	Max	*n*	Mean	Median	*SD*	Min	Max	*n*
Field wheat experiment	265.75	293	84.1	62	418	213	269.72	295	89.4	64	404	96
Laboratory germination	257.74	258	86.1	62	404	87	281.31	279	84.6	62	405	29
Complete dataset	263.43	283	84.6	62	418	300	272.41	295	88.1	62	405	125

Data were obtained from duplicate analysis by the HFN method. *SD*: standard deviation. Mean, median, SD, min and max are expressed in seconds. *n* = number of wheat batches analysed by HFN.

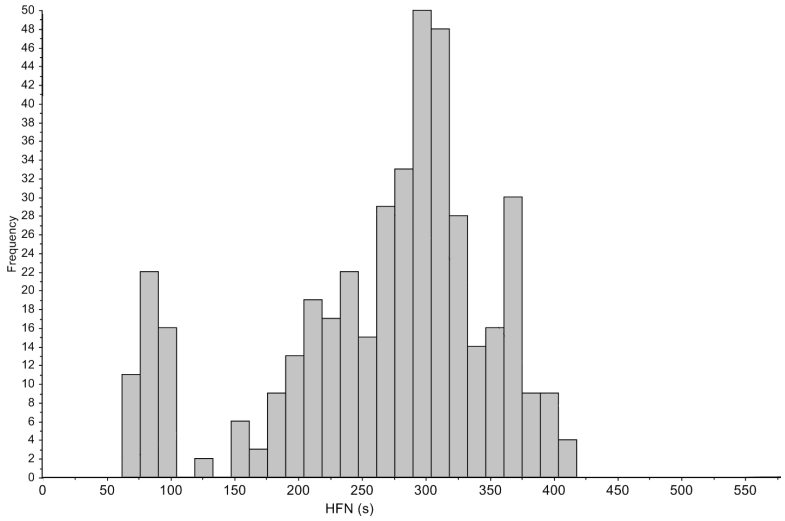

Figure 3. Distribution plot of HFN measured in the whole dataset of wheat samples ($n = 425$).

with R^2 of 0.73 and 0.54, respectively, although creation of a single calibration model for both classes was not successful. This could be sufficient for classification purposes, as the authors reported discriminant analysis accuracy up to 94% for high α-amylase content and 88% for low level samples. More recently, the same authors grouped Canadian West Red Spring wheat samples into three categories, namely low, medium and high activity, according to the measured α-amylase activity. They reported higher correlation coefficients for an FT-NIR spectrometer ($R^2 = 0.82$), while it was 0.79 for the best model using an HSI system. While the log transformation of the reference α-amylase activity helped in improving the model performance, the group with high enzymatic activity had visible sprouting, which could have affected NIR response.

Methods to apply the calibration to hyperspectral images

Several methods were tested to apply calibrations to hyperspectral images of test samples for prediction of HFN. By applying the calibration to each pixel, HSI provides the capability to map the distribution of a chemical compound within a sample, for example within kernels. For assessment of average properties for whole kernels or for bulk samples, potential approaches include application of a calibration to an average spectrum for the region of interest (equivalent to bulk spectroscopic methods), or alternatively to calculate the distribution of the property of interest at the single pixel level and then to calculate the average value for the sample. The effects of these approaches have been compared for the HFN example.

Table 2. Performance of the PLS calibration models for HFN prediction by HSI, by testing several spectral pretreatment methods.

Pre-treatments	Calibration ($n = 300$)				Validation ($n = 125$)				PC factor
	Slope	R^2	*RMSEC*	Offset	Slope	R^2	*RMSEP*	Offset	
Raw	0.397	0.40	62.6	161.7	0.383	0.42	67.0	170.4	8
2nd derivative	0.599	0.60	50.4	110.8	0.508	0.43	62.7	130.2	11
SNV	0.513	0.51	56.7	128.9	0.490	0.47	63.9	139.6	11
Baseline+detrend	0.424	0.42	61.6	153.3	0.409	0.41	67.5	161.5	9
MSC	0.415	0.41	61.9	156.7	0.408	0.42	66.8	159.9	7

R^2: coefficient of determination; *RMSEC*: root mean square error of calibration; *RMSEP*: root mean square error of prediction; PC: principal component; SNV: standard normal variate; MSC: multiplicative scatter correction

In either case, it is advisable to apply any required spike removal procedures prior to further processing to avoid their effect on the spectra.[28] The spike removal processing was done in both cases at the beginning on the hypercube, since any bright bad pixels in the spectral band used for segmentation appear as vertical lines in the segmentation image, making segmentation more complex.

Method 1. Application of the calibrations at single pixel level, followed by averaging the output within each object

Software was written in IDL to apply the relevant spectral pretreatment to each pixel in the hypercube and then to apply the PLS calibration coefficients for each spectral band to assess the spatial distribution of predicted HFN values for each pixel within the kernels. The mean of these values was then calculated to predict the HFN for each kernel.

Method 2. Averaging the spectra within each object followed by application of the calibration on the mean spectra

The alternative approach was to calculate mean spectra for each object (wheat kernel), followed by the application of the calibration, including any spectral pretreatment, to this average spectrum to give the predicted HFN for the kernel. In this case, a lower computational capacity is typically needed, since calculations at the single pixel level involve only calculation of mean spectra rather than application of calibrations, and the latter is a more computationally demanding operation, particularly if spectral pre-treatments are included.

In cases where no spectral pretreatment is applied, the two methods are equivalent, as shown by the following equations. For absorbance values A_{ij} for spectral bands $i = 1$ to n and spatial pixels $j = 1$ to m within a chosen object, the value of a property Y, such as HFN, can be predicted for each pixel using an equation of the form:

$$Y_j = a_0 + \sum_{i=1}^{n} a_i A_{ij} \tag{1}$$

where a_i are the calibration constants derived from the PLS regression.

The mean value of Y for the object (method 1) is given by:

$$\overline{Y} = \frac{1}{m}\sum_{j=1}^{m} Y_j = a_0 + \frac{1}{m}\sum_{j=1}^{m}\sum_{i=1}^{n} a_i A_{ij} \tag{2}$$

Alternatively, the mean spectrum for the object may be calculated as:

$$\overline{A_i} = \frac{1}{m}\sum_{j=1}^{m} A_{ij} \tag{3}$$

Applying the calibration to this mean spectrum yields a predicted measurement (method 2) for the object of:

$$\overline{Y} = a_0 + \sum_{i=1}^{n} a_i \overline{A_i} = a_0 + \frac{1}{m}\sum_{i=1}^{n} a_i \sum_{j=1}^{m} A_{ij} \tag{4}$$

which is equivalent to the result in (2).

For spectral pretreatments, this is not necessarily the case, depending on the pretreatment. For convolution filters, the approaches are also equivalent as shown by Equations 5–8. If a convolution pretreatment is applied to spectra to yield treated spectra:

$$A'_{ij} = \sum_{k=p}^{q} w_k A_{i+k,j} \tag{5}$$

or a treated mean spectrum

$$\overline{A'_i} = \sum_{k=p}^{q} w_k \frac{1}{m}\sum_{j=1}^{m} A_{i+k,j} \tag{6}$$

where w_k are the convolution weights, and for values $1 \le i + k \le n$, then the predicted value for the object, using alternative calibration constants b_i derived for the pretreated spectra, is

$$\overline{Y} = b_0 + \frac{1}{m}\sum_{j=1}^{m}\sum_{i=1}^{n} b_i \sum_{k=p}^{q} w_k A'_{i+k,j} \tag{7}$$

by the first method, and

$$\overline{Y} = b_0 + \frac{1}{m}\sum_{i=1}^{n} b_i \sum_{k=p}^{q} w_k \sum_{j=1}^{m} A'_{i+k,j} \tag{8}$$

by the second method, which are equivalent.

An example is the Savitzky–Golay second derivative pretreatment which gave the best calibration performance for the HFN calibration reported here. Figure 4 shows the effect of the two approaches, using this calibration. Grains for a poor quality sample, i.e. low reference HFN value, are compared to a sample with a high HFN value, and it can be seen that the two approaches are equivalent in predicting the HFN for single kernels. Moreover, the predicted mean HFN values for the batches tested were very close to the reference HFN values, i.e. 117 s vs 103 s for predicted and reference, respectively, in a low-quality wheat batch, and 404 s vs 370 s in the case of good quality wheat. The first method

is useful to visualise differences within a single grain, which is not possible with the second method. In our sample, a distinctive distribution of HFN values is seen within each grain, with lower values at the ends. This is independent of grain orientation, suggesting that it represents a genuine non-uniformity within the grains and not an artefact of the imaging and illumination conditions. Potentially, the distribution relates to the differential production of α-amylase activity in different regions of the grain. It should be noted that lower values are typically seen at both ends of the grain, and are therefore not only associated with the germ and scutellum. These results are in line with previous papers dealing with HSI estimation of germination in wheat, where it was suggested to use the size at the ends of the kernels as a parameter to recognise sprouted kernels, due to the swelling of the embryo during germination.[17] It was also previously reported that other regions within the wheat kernels are more useful for other predictions, e.g. HSI prediction of flour yield, while the germ tip of the kernel was the worst region for softness prediction.[29]

Previous studies applying HSI to HFN observed a non-uniform distribution within the kernel: Smail *et al.*[11] reported that the developed embryos in wheat can be recognised by HSI, as they were observed in the hypercubes of germinated kernels. The authors observed that much of the change in activity during early sprouting occurs in the outer kernel layers, i.e. aleurone and germ. In our case, the grains were placed in random positions in terms of crease up and crease down, and in the same hypercube, kernels in both positions were present at the same time.

Although the two methods for calculating mean HFN values for kernels are equivalent for spectral pretreatments based on convolution, the results may differ for other pretreatment methods. To test the magnitude of this effect, the SNV pretreatment reported in Table 2 was applied and the PLS model using this pretreatment was used to predict unknown samples. As shown in Figure 5, obvious differences in the predicted HFN values were obtained depending on the method used. When the second method was applied, generally the HFN prediction was quite similar between the two equations, i.e. the kernels with very low HFN values were identified. It should be noted that this method corresponds most closely to that used to calculate reference spectra for development of the calibration. However, differences were observed for the first method. The kernels with the lowest HFN in the batch were still predicted successfully, but the absolute value was much higher. The first method led to a systematic overestimation of the HFN in comparison with the second method. In this example, the best prediction of the batch HFN was obtained by using the second method. Also, when the first calibration was applied (Figure 4), the lowest predicted HFN within a single kernel was always localised at the distal portion of the grains, while when the second equation was applied, only the germ portion had obviously lower HFN predicted values (Figure 5).

The application of calibrations for bulk NIR instrumentation follows a different approach. Typically, a single spectrum is measured for a sample presented to a window, corresponding to a mean spectrum for that region of the sample, and is therefore closest to the second method in which calibrations are applied to mean spectra.

Therefore, the sequence of spectral processing should be taken into consideration when applying a PLS calibration on hypercubes, as it can have influence on the final predicted values. This influence depends on the spectral pre-treatments applied when building the calibration: in fact, linear methods such as convolutions do not lead to differences in the prediction, while obvious differences were observed when the SNV pre-treatment was applied. The SNV method involves treating each spectrum by subtracting the mean $\log(1/R)$ over all bands and normalising by their standard deviation, resulting in a pretreated spectrum with mean of 0 and standard deviation of 1:

$$A'_i = \frac{A_i - \overline{A}}{\sigma} \tag{9}$$

where A'_i is the corrected value, $\overline{A}$ is the mean of the band absorbances and σ is their standard deviation. The effect of subtracting the mean is equivalent for both methods, but normalisation of each spectrum by its own standard deviation is not equivalent to normalisation by the overall standard deviation for all kernel pixels and bands. Being sensitive to noise, modifications of SNV have been proposed, e.g. using the median instead of the mean, or using the inner interquartile for mean and standard deviation calculations.[4] This effect might be also tested on other pre-processing techniques, e.g. multiplicative scatter correction (MSC) which is similar to SNV up to a simple rotation and offset correc-

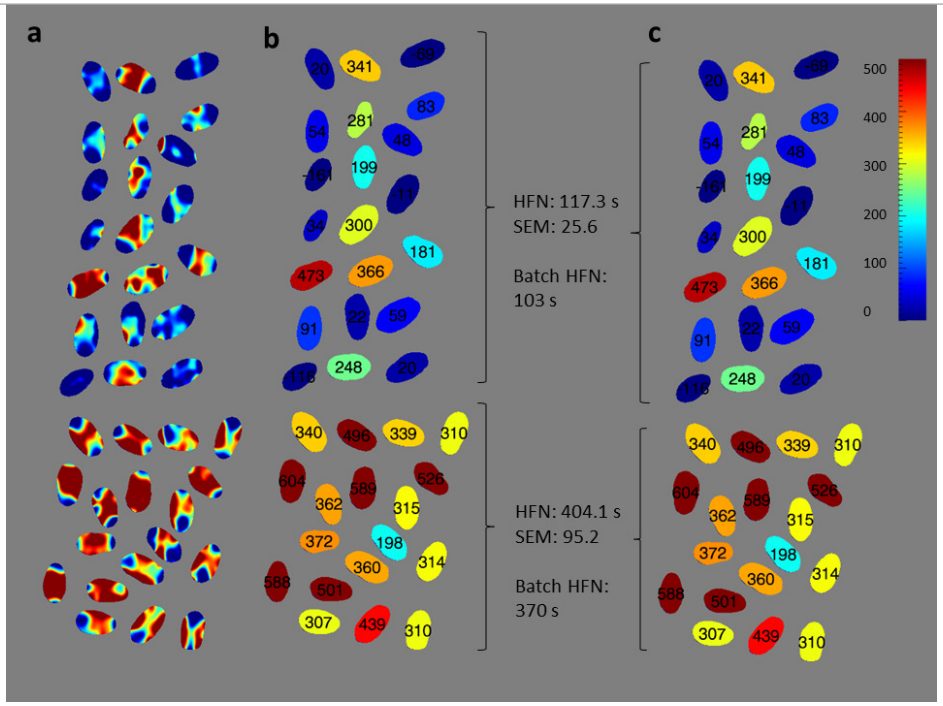

Figure 4. Application of a PLS calibration using second derivative pre-processing for HFN prediction on wheat kernels by HSI, by using two methods: a) pixel-by-pixel application of the calibration and b) average calculation of the predicted HFN; c) application of the calibration on the mean NIR spectrum for each kernel. HFN: predicted value for the batch; SEM: standard error of the mean; Batch HFN: reference measurement value. Numbers on wheat kernels indicate the predicted HFN (s).

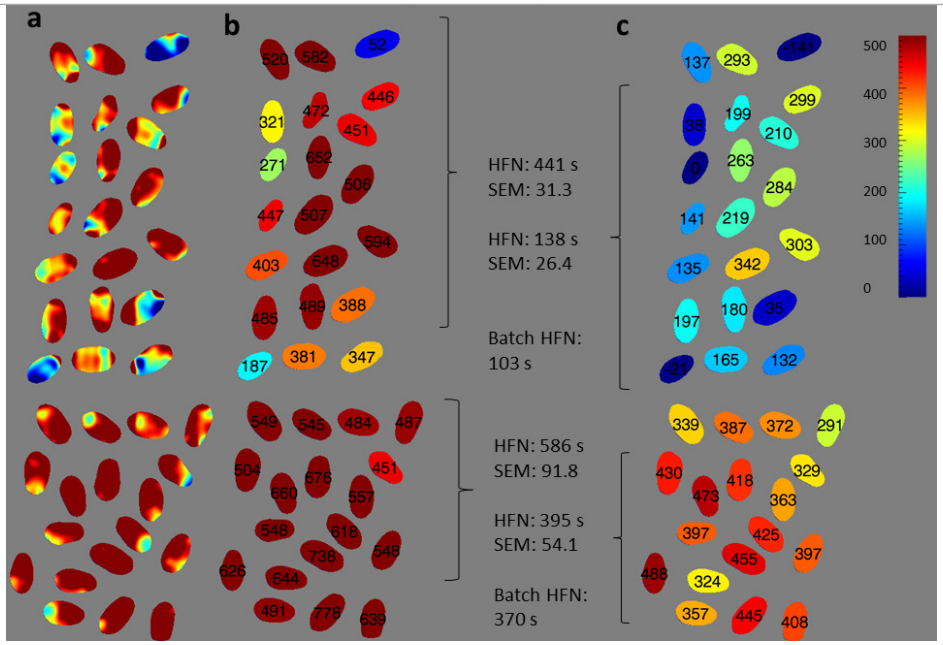

Figure 5. Application of a PLS calibration for HFN prediction on wheat kernels by HSI, using the two methods previously described, and a calibration where the SNV pre-processing was applied to two wheat batches with different HFN values. a) Pixel-by-pixel application of the calibration coefficients; b) second part of the method, average predicted HFN for each kernel (method 1); c) average spectrum calculation followed by application of the calibration coefficients (method 2).

tion. Further work is needed to understand the better approach for application of PLS calibrations to hypercubes for prediction of the chemical composition of granular food commodities, and the effects for different pretreatments.

LDA model for wheat classification related to sprouting problems

Linear discriminant analysis (LDA) was applied as a supervised method to discriminate between samples with high or acceptable HFN and those with excessively low HFN. As the PLS regression model did not result in particularly high performance for prediction of HFN values, this approach can be an alternative way to classify wheat samples for suitability for flour production.

The threshold to classify the dataset into two groups was set at an HFN value of 250 s, corresponding to the typical commercial specification for quality wheat, while also a lower threshold limit was set for grouping into "high" and "low" HFN, i.e. 150 s, which may be more relevant for identification of individual kernels that could contribute to an unacceptably low average value for a bulk sample. As shown in Figure 6, the LDA classification accuracy for the best models was 86.4% when the limit was 250 s, while it was 97.9% for the samples with very low HFN values. The plot shows linear discriminant values for each of the samples, with the axes representing the two categorical variables, i.e. high and low HFN. The classification is done by assigning the higher value to one of the two categories, and when a point is presented by two colours it indicates a misclassified sample.

Assuming that a low average value may be caused by a few much damaged kernels, identification of these at a lower threshold may be useful for practical applications. Interestingly, the better classification performances were obtained in both cases using untreated spectra, and spectral pre-treatments did not result in better scores (data not shown). The performance of this classification might be accurate enough for practical applications in the food industry, and LDA classification might be an alternative to the PLS prediction, when the main aim is not to predict the exact HFN or α-amylase activity but just to segregate according to the quality criteria set in terms of HFN threshold.

Several methods have been reported in the literature with the aim to classify cereal grains from NIR spectroscopy data. An alternative partial least square-discriminant analysis (PLS-DA) model was reported by McGoverin, Engelbrecht, Geladi and Manley[18] to distinguish between viable and non-viable proportions of the image and 61% of the variability in the viable/non-viable classifier was explained using six factors, by applying the classification method on each pixel of the kernels. The authors also calculated a PLS regression as an alternative to PLS-DA, to predict the proportion of viable kernels or incubation time, in which case the obtained model was not sufficiently reliable, consistent with our own results.

Xing *et al.*[17,23] applied alternative methods to predict α-amylase activity in wheat kernels, grouping samples into low, medium and high α-amylase activity. The ratio between log(1/*R*) at 878 nm and 728 nm was used to discriminate between sound and sprouted kernels, but the risk of misclassification is high. The authors applied a PCA on the hypercube to obtain a classification accuracy of 87.9% for severely sprouted kernels, whereas 64.5% of sprouted kernels were misclassified as sound.[17]

In the present experiment, LDA was applied in order to classify the problematic kernels, i.e. those with low HFN values, from those with higher values. Whereas there might be a kernel-to-kernel variation, there is a strong effect of sprouted kernels on the average HFN value, and therefore the application of this classification method might help the food industry and breeders with non-destructive classification of whole wheat kernels.

Conclusions

The present paper reports on the management of hyperspectral data and evaluates several methods for the application of PLS calibrations, with reference to prediction of HFN in single whole wheat kernels. Previous studies have evaluated HSI for classification of sprouted grains and for prediction of enzyme activity measured for single kernel extracts, but this is the first HSI calibration against the industry-standard HFN method. A calibration performance of $R^2 = 0.60$ (error ~50 s) and validation $R^2 = 0.43$ (error ~63 s) was achieved for samples of UK wheat with HFN values in the range 62–318 s, and compares favourably with previous studies using destructive bulk NIR analysis. Even better performance was obtained using classification models based on LDA instead of a PLS calibration. This suggests that an HSI method could be used to provide an effective, rapid, non-destructive assessment of the suitability of wheat sample HFN for flour production using a novel single kernel classification

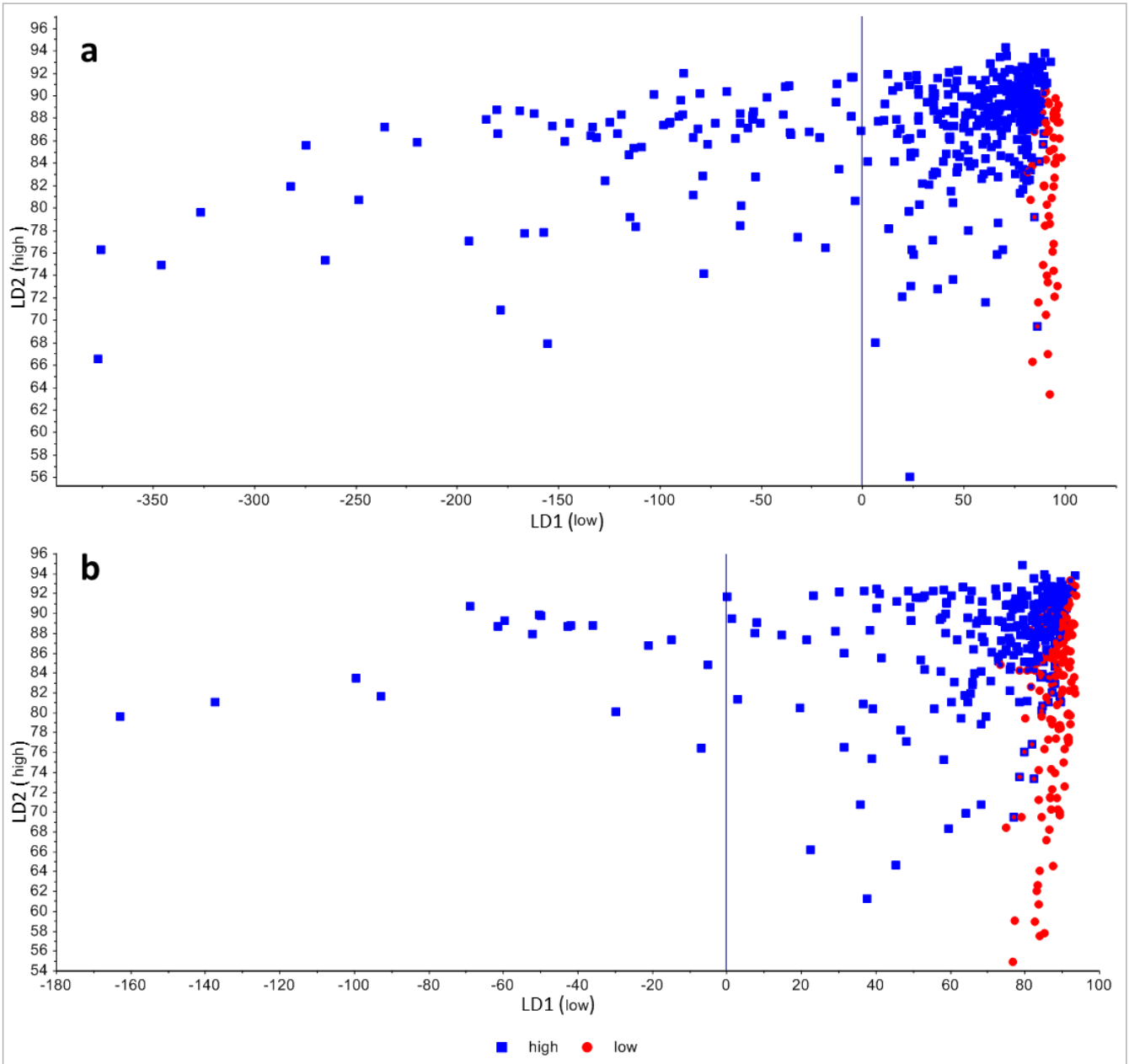

Figure 6. Linear discriminant analysis (LDA) plot for HFN classification of whole wheat kernels by HSI, setting a threshold of 250 s (a) and 150 s (b) to group into low- and high-HFN. Central colour of each point indicates reference classification, position and border colour indicates predicted NIR classification. LDA classification ability: 86.4% (a) and 97.9% (b). Wavelength region: ~1100–2500 nm.

approach. A further application would be for classification of individual kernels for selection of breeding lines for sprout resistance. A similar approach may also be applicable to development of classification methods for other granular commodities based on very low or very high concentrations of a particular compound.

Two approaches for application of calibrations were tested. It was shown that application of calibrations on a single pixel basis followed by calculation of mean values is equivalent to application of calibrations to mean kernel spectra when linear spectral pretreatments are used for calibration development, while non-linear treatments such as SNV lead to obvious differences between these approaches. Further work would be required to assess the effects for other non-linear treatments.

References

1. D. Wu and D.-W. Sun, "Advanced applications of hyperspectral imaging technology for food quality and safety analysis and assessment: A review—Part I: Fundamentals", *Innov. Food Sci. Emerg. Technol.* **19,** 1–14 (2013). https://doi.org/10.1016/j.ifset.2013.04.014
2. M. Vidal and J.M. Amigo, "Pre-processing of hyperspectral images. Essential steps before image analysis", *Chemometr. Intell. Lab. Syst.* **117,** 138–148 (2012). https://doi.org/10.1016/j.chemolab.2012.05.009
3. J. Burger and A. Gowen, "Data handling in hyperspectral image analysis", *Chemometr. Intell. Lab. Syst.* **108,** 13–22 (2011). https://doi.org/10.1016/j.chemolab.2011.04.001
4. Å. Rinnan, F. van den Berg and S.B. Engelsen, "Review of the most common pre-processing techniques for near-infrared spectra", *Trends Anal. Chem.* **28,** 1201–1222 (2009). https://doi.org/10.1016/j.trac.2009.07.007
5. N.L. Kent and A.D. Evers, *Technology of Cereals: An Introduction for Students of Food Science and Agriculture*, *4th Edn.* Pergamon (1994).
6. J.M. Barrero, K. Mrva M.J. Talbot, R.G. White, J. Taylor, F. Gubler and D.J. Mares, "Genetic, hormonal, and physiological analysis of late maturity α-amylase in wheat", *Plant Physiol.* **161,** 1265–1277 (2013). https://doi.org/10.1104/pp.112.209502
7. L. Reddy, T. Ching and R. Metzger, "Alpha-amylase activity in wheat kernels matured and germinated under different temperature conditions", *Cereal Chem.* **61,** 228–231 (1984).
8. D. Mares and K. Mrva, "Wheat grain preharvest sprouting and late maturity alpha-amylase", *Planta* **240,** 1167–1178 (2014). https://doi.org/10.1007/s00425-014-2172-5
9. ICC, International Association for Cereal Chemistry, *Standard Method No 107/1* (1968).
10. D. Mares and K. Mrva, "Late-maturity α-amylase: low falling number in wheat in the absence of preharvest sprouting", *J. Cereal Sci.* **47,** 6–17 (2008). https://doi.org/10.1016/j.jcs.2007.01.005
11. V.W. Smail, A.K. Fritz and D.L. Wetzel, "Chemical imaging of intact seeds with NIR focal plane array assists plant breeding", *Vibr. Spectrosc.* **42,** 215–221 (2006). https://doi.org/10.1016/j.vibspec.2006.02.004
12. H. Perten, "A modified falling-number method suitable for measuring both cereal and fungal alpha-amylase activity [use to improve fermentation of flour]", *Cereal Chem.* **61,** 108–111 (1984).
13. A. Barnard, C. Van Deventer and H. Maartens, "Comparison between methods for estimating sprout damage in wheat", *S. Afr. J. Plant Soil* **22,** 44–48 (2005). https://doi.org/10.1080/02571862.2005.10634679
14. B. McCleary and R. Sturgeon, "Measurement of alpha-amylase in cereal, food, and fermentation products", *Cereal Foods World* **47,** 299 (2002).
15. J.C.K. Verity, L. Hac and J.H. Skerritt, "Development of a field enzyme-linked immunosorbent assay (ELISA) for detection of alpha-amylase in preharvest-sprouted wheat", *Cereal Chem.* **76,** 673–681 (1999). https://doi.org/10.1094/CCHEM.1999.76.5.673
16. B.G. Osborne, "Investigations into the use of near infrared reflectance spectroscopy for the quality assessment of wheat with respect to its potential for bread baking", *J. Sci. Food Agric.* **35,** 106–110 (1984). https://doi.org/10.1002/jsfa.2740350117
17. J. Xing, S. Symons, M. Shahin and D. Hatcher, "Detection of sprout damage in Canada Western Red Spring wheat with multiple wavebands using visible/near-infrared hyperspectral imaging", *Biosyst. Eng.* **106,** 188–194 (2010). https://doi.org/10.1016/j.biosystemseng.2010.03.010
18. C.M. McGoverin, P. Engelbrecht, P. Geladi and M. Manley, "Characterisation of non-viable whole barley, wheat and sorghum grains using near-infrared hyperspectral data and chemometrics", *Anal. Bioanal. Chem.* **401,** 2283–2289 (2011). https://doi.org/10.1007/s00216-011-5291-x
19. S. Dencic, R. DePauw, B. Kobiljski and V. Momcilovic, "Hagberg falling number and rheological properties of wheat cultivars in wet and dry preharvest periods", *Plant Prod. Sci.* **16,** 342–351 (2013). https://doi.org/10.1626/pps.16.342
20. J. Xing, S. Symons, D. Hatcher and M. Shahin, "Comparison of short-wavelength infrared (SWIR) hyperspectral imaging system with an FT-NIR spectrophotometer for predicting alpha-amylase activities in individual Canadian Western Red Spring (CWRS) wheat kernels", *Biosyst. Eng.* **108,** 303–310 (2011). https://doi.org/10.1016/j.biosystemseng.2011.01.002
21. H. Risius, J. Hahn, M. Huth, R. Tölle and H. Korte, "In-line estimation of falling number using near-infrared diffuse reflectance spectroscopy on a combine harvester", *Precis. Agric.* **16,** 261–274 (2015). https://doi.org/10.1007/s11119-014-9374-5

22. H. Koç, V.W. Smail and D.L. Wetzel, "Reliability of InGaAs focal plane array imaging of wheat germination at early stages", *J. Cereal Sci.* **48,** 394–400 (2008). https://doi.org/10.1016/j.jcs.2007.09.015

23. J. Xing, P. Van Hung, S. Symons, M. Shahin and D. Hatcher, "Using a short wavelength infrared (SWIR) hyperspectral imaging system to predict alpha amylase activity in individual Canadian western wheat kernels", *Sens. Instrum. Food Qual. Saf.* **3,** 211–218 (2009). https://doi.org/10.1007/s11694-009-9087-z

24. S. Millar, M. Whitworth, A. Chau and J. Gilchrist, "Mapping food composition using NIR hyperspectral imaging", *New Food* **3,** 34–39 (2008).

25. J. Burger, "Bad pixel detection in hyperspectral staring camera systems", *NIR news* **20(2),** 9–12 (2009). https://doi.org/10.1255/nirn.1117

26. R. Dorrepaal, C. Malegori and A. Gowen, "Tutorial: Time series hyperspectral image analysis", *J. Near Infrared Spectrosc.* **24,** 89–107 (2016). https://doi.org/ 10.1255/jnirs.1208

27. Å. Rinnan, F. van den Berg and S.B. Engelsen, "Review of the most common pre-processing techniques for near-infrared spectra", *Trends Anal. Chem.* **28,** 1201–1222 (2009). https://doi.org/10.1016/j.trac.2009.07.007

28. S. Sasic and Y. Ozaki, *Raman, Infrared, and Near-Infrared Chemical Imaging*. John Wiley & Sons (2011).

29. S.R. Delwiche, E.J. Souza and M.S. Kim, "Limitations of single kernel near-infrared hyperspectral imaging of soft wheat for milling quality", *Biosys. Eng.* **115,** 260–273 (2013). https://doi.org/10.1016/j.biosystemseng.2013.03.015