Research Paper

Detection of sprout damage in wheat kernels using NIR hyperspectral imaging

Jayme G.A. Barbedo [a,*], Eliana M. Guarienti [b], Casiane S. Tibola [b]

[a] Embrapa Agricultural Informatics, Campinas, SP, Brazil
[b] Embrapa Wheat, Passo Fundo, RS, Brazil

ARTICLE INFO

The use of near-infrared (NIR) hyperspectral imaging (HSI) for detecting sprout damage in wheat kernels was investigated. Experiments were carried out to determine which spectral bands had the best potential for discriminating between sound and sprouted kernels. Two wavelengths were selected and combined into an index that was used to indicate the presence or absence of sprouting. Experiments with three wheat cultivars revealed that the proposed method is effective in identifying kernels for which the germination process has initiated, achieving 100% accuracy for the samples used in this study. It was also observed an imperfect correlation with the Falling Number (grain quality), making it challenging to accurately determine the degree of germination, especially if sprouts are not yet clearly visible. These results confirm the usefulness of the near-infrared spectral range for detecting chemical alterations in wheat kernels, as well as the fact that most information is usually contained in a few specific bands within such range.

1. Introduction

Sprouting occurs as a result of germination of wheat kernels following rainfall after maturity, reducing grain quality and value (Biddulph, Plummer, Setter, & Mares, 2008). The chemical properties of the sprouted grain can be significantly changed, causing important alterations on the concentrations of starch, sugar, proteins and dry matter (Lorenz & Valvano, 1981). In particular, the α-amylase enzyme is found in high concentrations in sprouted kernels, which affects baking quality and the premium paid for wheat (Singh, Jayas, Paliwal, & White, 2009). The grinding process is also heavily influenced by sprouting, both in terms of grinding energy requirements and distribution of the particle size (Dziki & Laskowski, 2010).

Sprouted kernels are also more vulnerable to diseases and insect infestations (Singh et al., 2009). Thus, it is very important to accurately measure the damage caused by sprouting so producers are paid fairly and the grains receive a proper destination.

In many cases, grain quality assessment and sprouting detection are performed visually. This visual selection, being a subjective task, is prone to psychological and cognitive phenomena that may lead to bias and optical illusions (Barbedo, 2016). Additionally, some cultivars have visual

characteristics that can lead to misclassifications (Singh et al., 2009). Delwiche, Yang, and Graybosch (2013) used computer vision (black and white images) and machine learning to explore the same visual cues used by human observer while removing the subjectivity of visual inspections. The authors tested several types of damage together, so they did not report how their system performed in the specific case of sprout damage. Another method exploring visual cues was proposed by Ebrahimi, Mollazade, and Babaei (2014), who employed RGB images for detecting several types of damage. Although the results reported for germinated kernels were good, sprouts had to be clearly visible for the system to work properly, preventing early detection.

Another widely employed method to estimate the grain quality loss associated to sprouting is the "Falling Number", which is a measure closely related to the concentration of the α-amylase enzyme in the grains. This number indicates the suitability of the grains for milling (Biddulph et al., 2008): the higher the Falling Number (measured in seconds), the higher is the quality of the grain and, hence, the higher the payment grade. In Brazil, there are four classes defined by the Falling Number: enhancer (FN > 250 s), bread (FN > 220 s), domestic (FN > 220 s) and basic (FN > 200 s) (MAPA, 2010). Despite its advantages, the Falling Number method is destructive and relatively time consuming, making it unsuitable for online inspection (Singh et al., 2009). Additionally, low Falling Numbers are possible in the absence of sprouting, which may lead to misdetection (Mares & Mrva, 2008). Other common methods for detecting sprouted kernels include the measurement of amylograph viscosity and chemical assays (Neethirajan, Jayas, & White, 2007).

One way to overcome some of the limitations of the established approaches is to explore spectral differences between sound and sprouted kernels. Shashikumar, Hazelton, Ryu, and Walker (1993) had relative success applying near-infrared (NIR) spectroscopy to identify sprouted kernels. However, this approach records the spectrum in a specific measurement point rather than the whole seed, thus failing to fully explore the information available (Wu, Zhu, Wang, Ma, & Wang, 2012; Xing, Symons, Hatcher, & Shahin, 2011). Thermal imaging, combined with machine learning, was successfully used by Vadivambal, Chelladurai, Jayas, and White (2010) to differentiate sound and sprouted kernels. Neethirajan et al. (2007) achieved accuracies above 90% using a soft X-ray system (1—100 nm). However, as pointed out by Singh et al. (2009),

this kind of system may pose potential health risks to humans. Some groups are now investigating the possibility of using Terahertz imaging (wavelengths between infrared and microwave bands) as an alternative to X-ray systems (Jiang, Ge, Lian, Zhang, & Xia, 2016), but this is still an incipient technology.

Hyperspectral imaging (HSI) is another recent spectrum-based technique to be explored for analysing wheat kernels. This technique uses the same principles of spectroscopy, but it generates spectra for each pixel in an image, rather than for a small localized area (Barbedo, Tibola, & Lima, 2017). To the authors' knowledge, the first study to use HSI for wheat kernel analysis was carried out by Delwiche and Kim (2000). Since then, this technique has gained momentum and has been applied to several different wheat kernel classification and detection problems. Smail, Fritz, and Wetzel (2006) and Koç, Smail, and Wetzel (2008) were the first to use HSI for sprouting detection. Soon after, Singh et al. (2009) proposed a method to classify wheat kernels into sound, sprouted, and midge-damaged, achieving accuracies close to 100%. Xing, Hung, Symons, and Shahin (2009) used this technique to predict α-amylase activity in wheat kernels, obtaining accuracies above 80%. These authors continued to work on the problem for a few more years, focussing both α-amylase activity (Xing et al., 2011; Xing, Symons, Shahin, & Hatcher, 2010a) and sprouting detection (Xing, Symons, Shahin, & Hatcher, 2010b). More recently, Wu et al. (2012) used hyperspectral images to detect sprouting in whole ears of wheat, coming to the conclusion that under these conditions only severe sprouting is detectable. Although no new investigations on the use of HIS for sprouting detection have been published in the last five years, this kind of technique continues to find several suitable applications for wheat, including deoxy-nivalenol screening (Barbedo et al., 2017), protein content prediction (Caporaso, Whitworth, & Fisk, 2018), classification of contaminants (Ravikanth, Singh, Jayas, & White, 2015), detection of black tip damage (Armstrong, Maghirang, & Pearson, 2015), Fusarium detection (Ropelewska & Zapotoczny, 2018), among others.

Many methods exploring hyperspectral images apply Principal Component Analysis (PCA) to remove redundancy and make the data more treatable. This is the case for most of the references cited above. As powerful as PCA is, its use does not always lead to better results (Barbedo, Tibola, & Fernandes, 2015). More importantly, subtle particularities of specific bands that might be valuable in the classification may be lost in the process. Also, many of those methods extract several features to feed their classification scheme, increasing the chance that at least some features overfit the data and leading to biased results. In this context, the first objective of this study was to investigate how raw reflectance values could be used for sprouting detection, avoiding the use of PCA and other similar techniques. In order to simplify the requirements of a potential kernel screening system, this study also aimed to select a small set of wavelengths representative enough for sprouting detection.

The spectral responses associated to different wheat cultivars may vary considerable. Important differences were observed when hyperspectral imaging was applied to

problems such as *Fusarium* detection (Barbedo et al., 2015), kernel quality (Zhu et al., 2012) and kernel viability (McGoverin, Engelbrecht, Geladi, & Manley, 2011). So far, most studies dealing with sprouting damage have used cultivars grown in Canada. Thus, another important contribution of this work lies in the fact that it used cultivars adapted to the conditions found in Brazil, which are considerably different than those found in Canada. Indeed, significant differences in the spectral responses were observed not only between Brazilian and Canadian cultivars, but also among the cultivars used in this work, as detailed in Section 3.

2. Material and methods

2.1. Wheat material

In order to obtain samples with different levels of germination, wheat kernels from the cultivars BRS Guamirim, BRS Louro and BRS Parrudo were submitted to a sprouting test. The samples were assayed in moistened *Germitest* paper.

Samples were kept inside a seed germination chamber for different periods of time (treatments). Temperature was kept at 20 °C ± 2 °C, and relative air humidity was not controlled. Eight treatments were applied to BRS Guamirim and BRS Louro, and six treatments were applied to BRS Parrudo. Sprouted samples were obtained in quadruplicate at the Laboratory for Seed Analysis of the Embrapa Wheat, Passo Fundo, Brazil.

Samples were dried up for 24 h in a heating chamber with forced air circulation and temperature below 40 °C. This was the time estimated for the water content in the samples to reach 10%, which is considered ideal for interruption of enzymatic activity.

After drying, samples were ground in a Perten mill (Perten Instruments, Hägersten, Sweden). The Falling Number test was carried out using a Perten Falling Number 1900 (Perten Instruments, Hägersten, Sweden), using the 56-81B analysis method (AACC, 2000), with sample weight correction based on moisture content, which was determined using the NIR instrument FOSS XDS (RCA, Hoganas, Sweden).

2.2. Image acquisition

The acquisition of the hyperspectral images was carried out at the Post-harvest Laboratory of Embrapa Wheat, Passo Fundo, Brazil. Kernels were imaged just before grinding for determination of the Falling Number. The spectrometer used in the system (EV/NIR Hyperspec Model 1003B-10151, Headwall Photonics Inc., Fitchburg, MA, USA) employed an InGaAs sensor with a 320×256 pixels focal plane array (FPA) and a XENICS camera (Model XEVA-1246 XC 134, Leuven, Belgium); it measured the reflectance of the kernels in the 528–1785 nm wavelength range (VIS/NIR), with a spectral resolution of 5 nm, resulting in 256 bands. This spectrometer was coupled with a 25-mm C-mount lens (F1.35/25 mm) placed 235 mm above the kernels, resulting in a field of view 50 mm wide. The images were acquired in a line-by-line basis (using push-broom acquisition); a total of 800 lines were scanned for each image, with approximately 500 of

those lines delimiting the region of interest where the kernels were located. The illumination was provided by an external Quartz Tungsten-Halogen lamp, whose light was conveyed via an optical fibre bundle that terminates in a 250-mm long line tilted by approximately 30° with respect to the vertical axis. More details on how the images were captured can be found in Barbedo et al. (2015). The result for each capture was a 3D matrix with dimensions of $320 \times 800 \times 256$, that is, the images have a spatial dimension of 320×800 over 256 different bands. Each side of the pixels has approximately 0.25 mm, and each wheat kernel is represented by 700–1000 pixels.

2.3. Image datasets

The image dataset used in this work has 88 images (32 from BRS Guamirim, 32 from BRS Louro, and 24 from BRS Parrudo cultivars). Two images, one with no sprouting and another with severe sprouting (Falling Number < 70), were selected from each cultivar to test different wavelengths and tune threshold values. Thus, 82 images were used in the tests (Section 3). About 15% of the kernels used in the experiments had some visible sprouts. These were not removed because, in a hypothetical screening system, kernels would be processed as they were, without sprout removal.

2.4. Proposed procedure

First, the tray region was delimited by removing the background, which was done following the guidelines described in Barbedo et al. (2015). In short, tray delimitation is achieved by applying some morphological operations to the 647-nm wavelength, which provides the best contrast between tray and background.

Next, kernels were separated from the tray, effectively isolating the region of interest (ROI) to be processed. The 1017-nm reflectance image was subtracted from the 1115-nm reflectance image. A threshold was then applied to the resulting difference image, where pixels were made equal to one if their values were higher than 0.15, and equal to zero otherwise. The image was then morphologically opened using a 5-pixel diameter disk as structuring element, in order to trim out spurious elements. The resulting binary mask was applied to all 256 bands, effectively isolating the kernels and removing the whole background. All three wavelengths used in the segmentation (647, 1017 and 1115 nm) were selected empirically by visually inspecting their reflectance characteristics, which were explored to promote the best possible segmentation.

In this investigation, the falling numbers used as reference during the experiments were determined collectively for all kernels that appear in each image. As a consequence, all experiments with the proposed methodology also considered the kernels in bulk. However, in many cases it is desirable to detect individual sprouted kernels (Xing, Hung, Symons, Shahin, & Hatcher, 2009). Identifying individual kernels using the binary mask is trivial, but sometimes kernels appear grouped. For that reason, kernel clusters were separated by identifying the concavities generated by touching kernels, and then connecting those by straight separating lines (Barbedo

et al., 2015). The procedure is capable of accurately separating up to three clustered kernels.

As mentioned in the introduction, one of the objectives of this study was to find the simplest setup that could be implemented in a potential screening system for sprouting detection, thus reducing costs and technical requirements. Experiments revealed that there are two bands that provide the best contrast between sound and sprouted kernels (Section 3). Reflectance values associated to each wavelength within those bands were very similar, so only one wavelength was chosen to represent each band, 918-nm and 1411-nm. This confirmed the conclusions of a previous investigation, which indicated that using only a few wavelengths for detection of damaged kernels is more effective than using the entire spectrum (Barbedo et al., 2015). It is worth noting that while the 1411-nm band is associated to water, all kernels used in the experiments had similar moisture contents (around 10%), so spectral differences indicated by this band were likely due to chemical alterations caused by the sprouting.

The expression used to combine the selected wavelengths was:

$$SI = \frac{\sum_{i=1}^{800} \sum_{j=1}^{320} |R_{918}(i,j) - R_{1411}(i,j)|}{\sum_{i=1}^{800} \sum_{j=1}^{320} [R_{1411}(i,j)]}$$

where SI is the "Sprouting Index", R_w is the reflectance of wavelength w, i and j are indices of the pixel, and $|\bullet|$ is the absolute value operator. As it can be seen, SI is given by the sum of all pixels resulting from the absolute difference between the 918-nm and 1411-nm wavelengths, divided by the sum of all pixels of the 1411-nm wavelength. The subtraction operation was chosen because sound and sprouted kernels have opposite behaviours in those bands, thus maximizing their separation. The division by the total reflectance of the 1411-nm was included in order to mitigate the effects of non-uniform pixel sensitivity, which causes the intensities recorded by the detector elements to vary (Jayas, Singh, & Paliwal, 2010). This division normalizes reflectance values, making them more uniform within the image.

The smaller the value of *SI*, the more likely is the presence of sprouting. The ideal threshold that separates sound and sprouted kernels depends on the cultivar, but the value of 0.30 worked relatively well for all cultivars considered in this study (see Table 1).

Table 1 – Accuracy in detecting sound and sprouted kernels arranged randomly and with germ up. Results were obtained using a global threshold (0.30) and cultivar-specific SI thresholds (0.30 for BRS Guamirim and BRS Louro, and 0.35 for BRS Parrudo).

Cultivar	Random arrangement		Germ up	
	Global threshold	Specific threshold	Global threshold	Specific threshold
BRS Guamirim	97%	100%	100%	100%
BRS Louro	100%	100%	100%	100%
BRS Parrudo	36%	68%	88%	100%

2.5. Experimental setup

Five experiments were carried out in the present study.

In the first experiment, wheat samples were divided into three classes according to the cultivar. Each class was further divided into five groups according to the degree of sprouting: sound kernels (FN $\geq$ 350s), beginning of germination (250 s $\leq$ FN < 350 s), low damage (150 s $\leq$ FN < 250 s), moderate damage (70 s $\leq$ FN < 150 s) and severe damage (FN < 70 s). For each cultivar, the mean reflectance spectra obtained for the subgroups were compared in order to identify significant differences, revealing the degree of separability between those groups and how wheat variety affects such separability. Two types of reflectance curves were generated: 1) actual reflectance spectra; 2) the ratios between the reflectance spectra obtained for sprouted kernels and the reference reflectance spectrum obtained for sound kernels. This second set of curves emphasizes differences between both types of kernels, making it easier to identify bands of interest.

In the second experiment, the mean reflectance spectrum for each group described above was collected again, this time placing all kernels with germ up, rather than placing kernels randomly in the tray (standard procedure). The spectra were again compared, as described in experiment 1, in order to determine whether placing the kernels with germ up increased the separability between the groups.

In the third experiment, SI values were used for classifying the samples into sound and sprouted, using a single threshold for all images, and also using individual thresholds for each cultivar. Accuracies were then computed for each case, using as reference the reference labels attached to each image – images with FN > 250 were labelled as sound, and the remaining ones were classified as sprouted.

In the fourth experiment, the correlations between SI and FN were calculated for all kernels, and also for each cultivar individually. The objective was to investigate how well SI can predict FN and, consequently, how suitable is SI for predicting sprouting severity in routine analysis.

In the last experiment, the results obtained using the two selected wavelengths were compared to two other methods. In the first alternative method, the two selected wavelengths were replaced with the mean reflectance values over the entire bands for which the differences between sound and sprouted kernels was larger (844–1140 nm and 1386–1700 nm). In the second alternative method, all wavelengths were combined using PCA. This last method is used in many investigations found in the literature.

3. Results and discussion

The comparison of the reflectance spectra yielded by sound and sprouted kernels showed clear differences in their spectral responses (Fig. 1).

There are two bands for which there is a visible difference between sound and sprouted kernels: 844–1140 nm and 1386–1700 nm. The spectral differences between the two types of kernels can be seen more clearly in Fig. 2, which
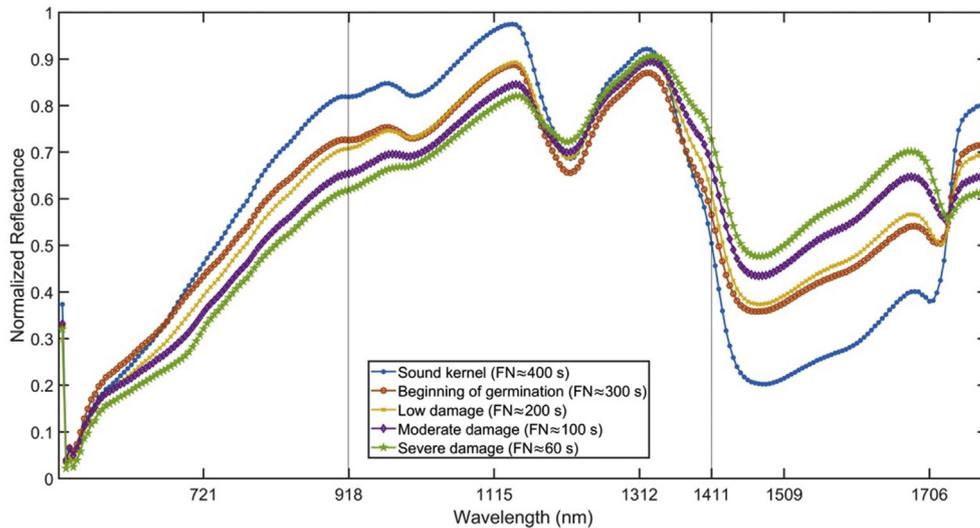
Fig. 1 – Comparison of the spectra yielded by kernels with different degrees of sprouting (BRS Parrudo cultivar).
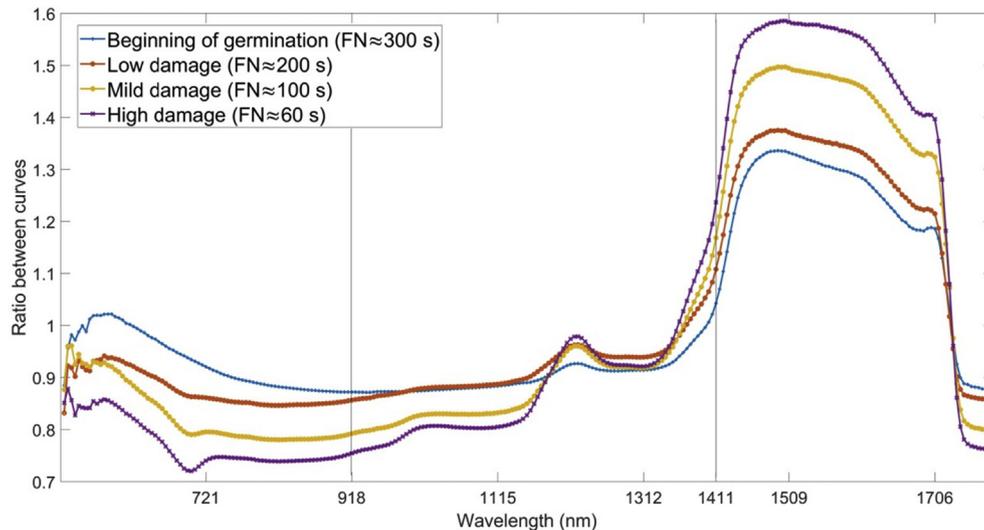

Fig. 2 – Ratios between the spectra of sprouted and sound kernels (BRS Parrudo cultivar).

shows the ratio between the spectra obtained for the sprouted kernels and the one obtained for the sound ones.

In the first band (844–1140 nm), sound kernels have a reflectance 10–20% higher than the sprouted ones. This observation contrasts with the results obtained by Xing et al. (2010b), who stated that sound kernels had a distinctly lower spectral reflectance in the wavelength region between 720 nm and 900 nm. There are two possible explanations for this disparity. First, Fig. 1 was obtained after normalizing each wavelength with respect to the entire spectrum, which may have slightly altered the relationship between the curves – this normalization aimed at compensating for non-uniform pixel sensitivity associated with FPA-based hyperspectral imaging systems, as discussed in Section 2.4 and in Jayas et al. (2010). Second, wheat areas in Brazil and Canada have very different climates and different genealogies. Therefore, the cultivars used in each study probably have significantly

different chemical compositions. It is also important to notice that this band seems to not be strongly associated to any substance present in wheat kernels, rather being the summed response of several types of molecules (Chen et al., 2014; Xing et al., 2010a). Therefore, small alterations in several of those chemical components may result in significantly different spectral responses.

The differences observed in the second band (1386–1700 nm) are more prominent, with the reflectance of sprouted kernels being between 20% and 60% larger than that of sound kernels. This band is strongly associated with water, which has a major absorption wavelength at 1400 nm. In contrast, the spectra obtained by Chen et al. (2014) for kernel embryos alone resulted in lower reflectance for sprouted kernels. This is explained by the fact that just before germination, the endosperm is imbibed with water, reducing the reflectance. As the germination progresses, the embryo,

which occupies only a small part of the kernel, absorbs water. As a result, water content in the endosperm is reduced, increasing the reflectance.

The results reported above indicate that major spectral differences are mostly due to variations in the water content. This is an important observation, because although several physical variables are altered due to sprouting, they seem to have only a mild impact on the reflectance spectrum, mainly in the 844–1140 nm band. There are two possible explanations for this: 1) some substances have their absorption bands located outside the spectral interval considered, e.g. protein (2200 nm); 2) some substances have at least some of their absorption bands located inside the spectral interval considered, but their content variation was small, e.g. fat (1200 nm).

In the specific example shown in Figs. 1 and 2, the more developed the sprouting (more damage), the larger the spectral disparity with respect to sound kernels. However, this was not always the case, as in many cases kernels with more severe sprouting actually generated spectra closer to those of sound kernels. This is exemplified in Fig. 3, where the spectrum with the lowest FN associated (highest damage) was actually the one that departed the least from the spectrum for sound kernels. Thus, although the results indicate that hyperspectral images are suitable for discriminating between sound and sprouted kernels, they do not seem very reliable for estimating the severity of the sprouting. It is important to highlight, however, that significant differences were observed for different cultivars (Figs. 2 and 3), indicating that in some cases it may be possible to successfully estimate sprouting damage.

In all samples analysed in this study, the spectra for sound and sprouted kernels began to diverge just a few hours after the beginning of the sprouting process, even when FN was still above 300 s. This extreme sensitivity to changes caused by sprouting is mostly positive, but it also carries some practical challenges. As discussed before, water content seems to be the main factor shaping the reflectance curves. In contrast, the Falling Number is more related to α-amylase content.

There seems to be a correlation between those quantities, as indicated by experiments presented later in this section, however they are not perfectly aligned. This has to be taken into consideration in the development and use of any HSI-based method expected to replace the FN approach. Another challenge arises from the fact that kernels sprouted in the field, under uncontrolled conditions, may present different damage evolution patterns and, as a consequence, different trends regarding water content variation. Since humidity is such an important variable for shaping the spectral responses, the tendencies observed using artificially sprouted kernels may not hold perfectly under more realistic conditions (Singh et al., 2009).

Any method for differentiating sound and sprouted kernels using hyperspectral images necessarily rely on the differences between the reflectance curves obtained for those kernels, either considering the whole spectrum or specific wavebands. This may be challenging, as typical spectra may vary considerably depending on the cultivar. As a result, methods for detecting sprouted kernels may have to either be trained for each cultivar separately, or include some parameters to be tuned accordingly. To illustrate this issue, Fig. 4 shows the typical spectra for sound kernels of the three cultivars considered in this work. Focussing on the bands of interest (844–1140 nm and 1386–1700 nm), it can be noticed that the spectrum obtained for the BRS Parrudo cultivar has a much higher reflectance than the other two in the first band, but it has almost the same reflectance as BRS Guamirim in the second band. BRS Louro and BRS Guamirim have relatively similar spectra, but they are hardly interchangeable. It is interesting to notice that the spectrum for the BRS Parrudo sprouted kernels remains above the BRS Louro reference spectrum for nearly the entire wavelength range. As a result, the BRS Louro cultivar spectrum for sound kernels should not be used as a reference to detect BRS Parrudo sprouted kernels, as this would most certainly lead to high error rates. This further reinforces that, in general, different cultivars have their own spectral characteristics and should be treated separately.
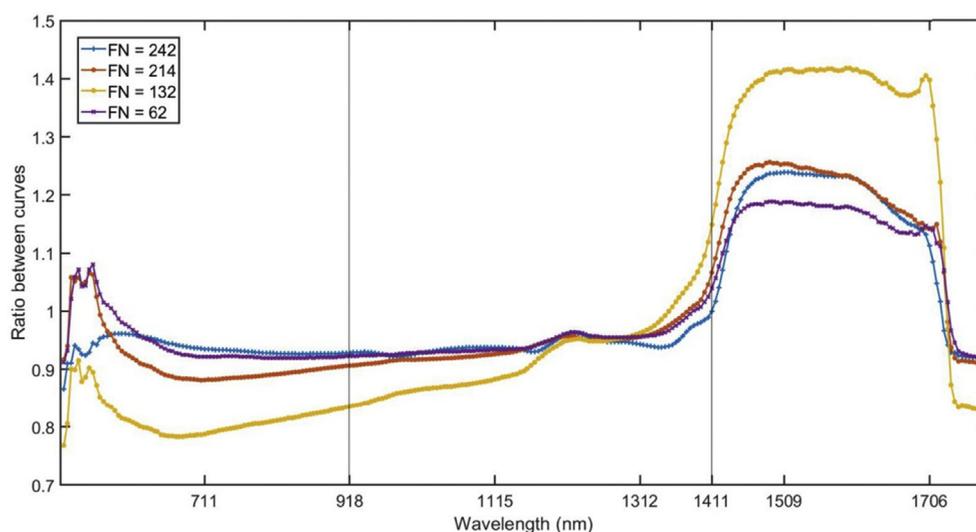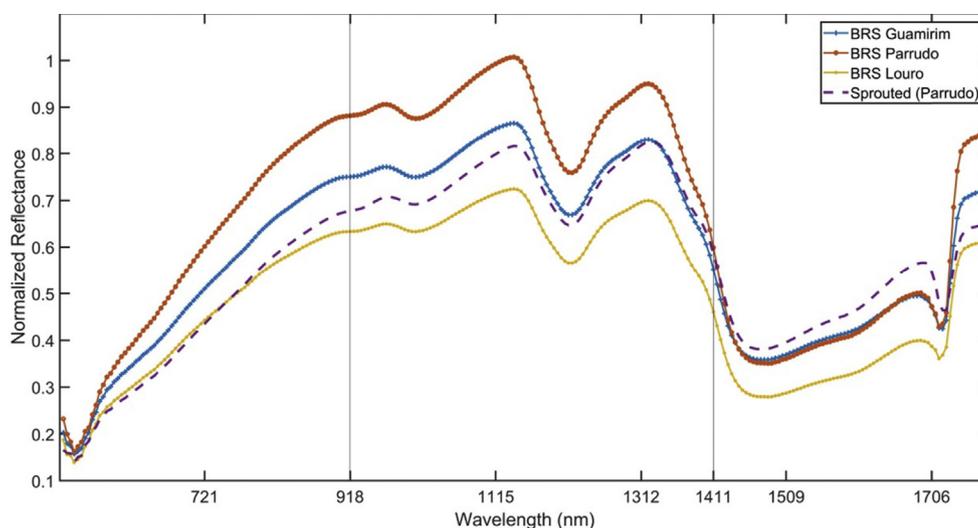

Fig. 3 – Ratios between the spectra of sprouted and sound kernels (BRS Louro cultivar).

Fig. 4 – Reflectance spectra for sound kernels of three different cultivars. The spectrum for sprouted kernels of the BRS Parrudo cultivar was included for comparison (dashed line).

In contrast, the variability found for spectral profiles obtained using different kernels of the same cultivar were negligible for all samples used in this work.

Because of the spectral differences, typical Spectral Index (SI) values, which provide a rough estimate for the sprouting damage (Section 2.4), vary between cultivars. In particular, SI values tend to be 30% higher for BRS Parrudo than for the other cultivars. Thus, finding a single SI threshold value capable of separating sound and sprouted kernels for all cultivars is often not practical. Table 1 shows the results considering global and cultivar-specific SI thresholds. For the samples used in the tests, all misclassifications were due to sprouted kernels being classified as sound. The global threshold was a good match for two cultivars (BRS Guamirim and BRS Louro), but good results for the BRS Parrudo cultivar were only possible when a specific threshold was applied. It is also worth noting that kernel positioning only had a significant impact for the latter. This indicates that physical properties associated to each cultivar determine whether random kernel positioning is acceptable or not.

The results reported in this section indicate that there are noticeable spectral differences between sound (FN > 250) and sprouted kernels (FN ≤ 250). However, the distinctive spectral characteristics of sprouted kernels may be masked depending on their position. Experiments in which all kernels were positioned with the germ up generated more homogeneous and reliable outcomes (Table 1). Most results presented in this section are roughly valid for both cases (e.g. the spectral curves in Figs. 1–4), but when significant differences were observed, the results for both arrangements were presented separately.

The accuracies obtained in the experiments were similar to those reported by Singh et al. (2009) and Xing et al. (2010b), except when a random kernel arrangement was combined with a global threshold, due to poor results for the BRS Parrudo cultivar. It is important to consider, however, that all samples used by Singh et al. (2009) were severely sprouted, while about 25% of the samples used here had FN between 200 s and 300 s,

thus being only slightly sprouted. Thus, as discussed above, sprouting was detectable even at its earliest stages.

The correlations between SI values and FN were also relatively high (Table 2). However, because SI is more related to water content, and FN is more related to α-amylase content, the correspondence between both values is not close enough to allow for a reliable quality classification such as the one described in the Introduction section.

The HSI equipment used in this work captures images in the NIR (Near-Infrared) band. Investigations reported by Xing et al. (2010a) and Xing et al. (2011) indicate that the Short-Wave Infrared (SWIR) band (1000–3000 nm) is more appropriate to directly detect and predict α-amylase content. Using only the spectra of the germ part of the kernels, they were able to detect early sprouting with accuracy around 90%, and predict α-amylase content with a coefficient of determination above 80%. These results seem to indicate that NIR and SWIR bands have similar capabilities regarding sprouting detection, although only the latter can be used to quantify α-amylase content. However, both bands have some limitations. As discussed before, sprouting detection using the NIR-band seems to be mostly based on water content, which can be affected by several aspects other than sprouting, especially in kernels brought directly from the field. On the other hand, the α-amylase detection and quantification requires the germ region of the kernels to be correctly segmented, which is not always possible to be done automatically (Xing et al., 2011). The amount of unknowns that still remain indicates that more studies are necessary for a definitive

Table 2 – Correlation between SI values and Falling Numbers.		
Cultivar	Random arrangement	Germ up
BRS Guamirim	0.78	0.89
BRS Louro	0.83	0.84
BRS Parrudo	0.69	0.80

answer on which band is more appropriate for sprouting detection in wheat.

During the tests, some inconsistencies and unexpected results were observed, such as the slightly erratic behaviour observed for the reflectance curves as FN changed (Fig. 3). Although those irregularities were not too severe, they raised some questions regarding their origin. Two probable causes were identified. First, FPA-based image systems have a number of optical distortions and errors associated, such as chromatic aberrations and non-uniform pixel sensitivity (Jayas et al., 2010). These distortions will inevitably affect the image analysis, and if spectral variations associated to a given phenomenon are slight (which is the case when comparing kernels with close Falling Numbers), detection may become unfeasible. Second, low Falling Numbers may be caused by factors other than sprouting, such as genotype (Mares & Mrva, 2008), cultivar, location, and presence of certain diseases (Wang, Pawelzik, Weinert, Zhao, & Wolf, 2008). Although those are minor factors in comparison with sprouting (Xing et al., 2009), they may have enough influence to cause unexpected results.

Many authors use features to perform image classification (Singh et al., 2009; Xing et al., 2010b). If the number of features is high with respect to the number of images used in the experiments, it is likely that at least some of the features will overfit the data. As a consequence, results may not be valid if larger, more diverse datasets are considered, considerably weakening the conclusions. If only a few samples are available, simpler and more robust approaches should be preferred. In order to avoid overfitting, while being robust to material and condition variations, the proposed method does not employ features, instead including a single variable (SI threshold) that can be easily tuned to meet the characteristics of each cultivar.

Xing et al. (2010b) argued that, in order to improve the grading system and reduce the dependence on starch pasting methods, it is important to create a system that both identifies individual sprouted kernels and determines their degree of damage. All tests in this study were performed on bulk samples, because the reference Falling Numbers were only available for the kernels collectively. However, the segmentation technique proposed in Barbedo et al. (2015) could be applied here, making it feasible to process each kernel individually. Estimating the degree of damage using NIR HSI does not seem to be feasible at the moment, but since imaging technology is continually evolving and new image processing techniques are constantly being proposed, this situation may change in the near future.

The proposed methodology for sprouting detection performed favourably in comparison with other strategies. The accuracy obtained when the reflectance values of the selected wavelengths was replaced with the mean reflectance values fell from 100% to 89% (specific threshold). This is because the specific wavelengths were selected having the maximization of the differences between sound and sprouted kernels as criterion. Calculating the mean over several bands diluted those differences, reducing discrimination power. The best results using PCA were obtained when the first principal was used in isolation, in which case the accuracy was 92%. This drop in accuracy was likely due to the loss of subtle particularities of specific bands that might be valuable in the classification, as mentioned in the Introduction.

Finally, it is worth mentioning that part of the kernels had visible sprouts. The segmentation procedure was successful in removing those and the background altogether. In fact, the few segmentation errors observed consisted of small portions of the background being kept and small regions of the kernels being removed. Sprouting detection accuracies using images manually segmented were statistically identical to those obtained using automatic segmentation, showing that the impact of segmentation flaws was negligible.

4. Conclusions

This paper presented an investigation on the use of NIR hyperspectral images for detection of sprouting in wheat kernels. Two wavelengths, 918-nm and 1411-nm, were found to be the most relevant. They were combined in such a way overfitting is avoided, thus increasing its ability to properly deal with samples from different origins and cultivars. Experiments have shown that NIR hyperspectral images are very sensitive to the physiological changes caused by sprouting, which can be detected at very early stages. On the other hand, it was not possible to reliably estimate the severity of the sprouting, as differences between different degrees of damage were too subtle in comparison with the noise caused by factors intrinsic to the problem. Nevertheless, those limitations may be overcome in the near future as imaging technologies evolve. The complementarity of NIR and SWIR bands for the tasks of sprouting detection and quantification could also be further investigated in order to improve the results obtained so far. The proper exploration of the infrared band may soon allow the conception of equipment tailored for the problem at hand, reducing costs and other practical constraints.

Acknowledgements

REFERENCES

AACC – American Association of Cereal Chemists. (2000). *Approved methods* (10th ed.). MO, USA: Saint Paul.

Armstrong, P. R., Maghirang, E. B., & Pearson, T. C. (2015). Detecting and segregating black tip-damaged wheat kernels using visible and near-infrared spectroscopy. *Cereal Chemistry, 92*, 358–363.

Barbedo, J. G. A. (2016). A review on the main challenges in automatic plant disease identification based on visible range images. *Biosystems Engineering, 144*, 52–60.

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

Barbedo, J. G. A., Tibola, C. S., & Lima, M. I. P. (2017). Deoxynivalenol screening in wheat kernels using hyperspectral imaging. *Biosystems Engineering, 155*, 24–32.

Biddulph, T. B., Plummer, J. A., Setter, T. L., & Mares, D. J. (2008). Seasonal conditions influence dormancy and preharvest sprouting tolerance of wheat (*Triticum aestivum* L.) in the field. *Field Crops Research, 107*, 116–128.

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

Chen, J., Chen, H., Wang, X., Yu, C., Wang, C., & Zhu, D. (2014). The characteristic of hyperspectral image of wheat seeds during sprouting. In *Proceedings of the 7th international conference on computer and computing technologies in agriculture (CCTA)* (pp. 408–421).

Delwiche, S. R., & Kim, M. S. (2000). Hyperspectral imaging for detection of scab in wheat. In *Proceedings of SPIE* (Vol. 4203, pp. 13–20).

Delwiche, S. R., Yang, I.-C., & Graybosch, R. A. (2013). Multiple view image analysis of free falling U.S. wheat grains for damage assessment. *Computers and Electronics in Agriculture, 98*, 62–73.

Dziki, D., & Laskowski, J. (2010). Study to analyze the influence of sprouting of the wheat grain on the grinding process. *Journal of Food Engineering, 96*, 562–567.

Ebrahimi, E., Mollazade, K., & Babaei, S. (2014). Toward an automatic wheat purity measuring device: A machine vision-based neural networks-assisted imperialist competitive algorithm approach. *Measurement, 55*, 196–205.

Jayas, D. S., Singh, C. B., & Paliwal, J. (2010). Classification of wheat kernels using near-infrared reflectance hyperspectral imaging. In D.-W. Sun (Ed.), *Hyperspectral imaging for food quality analysis and control* (pp. 449–470). San Diego: Academic Press.

Jiang, Y., Ge, H., Lian, F., Zhang, Y., & Xia, S. (2016). Early detection of germinated wheat grains using terahertz image and chemometrics. *Scientific Reports, 6*, 21299.

Koç, H., Smail, V. W., & Wetzel, D. L. (2008). Reliability of InGaAs focal plane array imaging of wheat germination at early stages. *Journal of Cereal Science, 48*, 394–400.

Lorenz, K., & Valvano, R. (1981). Functional characteristics of sprout damaged soft white wheat flour. *Journal of Food Science, 46*, 1018–1020.

MAPA. (2010). *Technical regulation of wheat*. Normative Instruction 38/2010.

Mares, D., & Mrva, K. (2008). Late-maturity α-amylase: Low falling number in wheat in the absence of preharvest sprouting. *Journal of Cereal Science, 47*, 6–17.

McGoverin, C. M., Engelbrecht, P., Geladi, P., & Manley, M. (2011). Characterisation of non-viable whole barley, wheat and sorghum grains using near-infrared hyperspectral data and chemometrics. *Analytical and Bioanalytical Chemistry, 401*(7), 2283–2289.

Neethirajan, S., Jayas, D. S., & White, N. D. G. (2007). Detection of sprouted wheat kernels using soft X-ray image analysis. *Journal of Food Engineering, 81*, 509–513.

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

Ropelewska, E., & Zapotoczny, P. (2018). Classification of Fusarium-infected and healthy wheat kernels based on features from hyperspectral images and flatbed scanner images: A comparative analysis. *European Food Research and Technology, 244*(8), 1453–1462. https://doi.org/10.1007/s00217-018-3059-7.

Shashikumar, K., Hazelton, J. L., Ryu, G. H., & Walker, C. E. (1993). Predicting wheat sprout damage by near-infrared reflectance analysis. *Cereal Foods World, 38*, 364–366.

Singh, C. B., Jayas, D. S., Paliwal, J., & White, N. D. G. (2009). Detection of sprouted and midge-damaged wheat kernels using near-infrared hyperspectral imaging. *Cereal Chemistry, 86*, 256–260.

Smail, V. W., Fritz, A. K., & Wetzel, D. L. (2006). Chemical imaging of intact seeds with NIR focal plane array assists plant breeding. *Vibrational Spectroscopy, 42*, 215–221.

Vadivambal, R., Chelladurai, V., Jayas, D. S., & White, N. D. G. (2010). Detection of sprout-damaged wheat using thermal imaging. *Applied Engineering in Agriculture, 26*, 999–1004.

Wang, J., Pawelzik, E., Weinert, J., Zhao, Q., & Wolf, G. A. (2008). Factors influencing falling number in winter wheat. *European Food Research and Technology, 226*, 1365–1371.

Wu, Q., Zhu, D., Wang, C., Ma, Z., & Wang, J. (2012). Recognition of wheat preharvest sprouting based on hyperspectral imaging. *Optical Engineering, 51*(11), 111710.

Xing, J., Hung, P. V., Symons, S., Shahin, M., & Hatcher, D. (2009). Using a short wavelength infrared (SWIR) hyperspectral imaging system to predict alpha amylase activity in individual Canadian western wheat kernels. *Sensing and Instrumentation for Food Quality and Safety, 3*, 211–218.

Xing, J., Symons, S., Hatcher, D., & Shahin, M. (2011). 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. *Biosystems Engineering, 108*, 303–310.

Xing, J., Symons, S., Shahin, M., & Hatcher, D. (2010a). Sprouting detection at early stages in individual CWAD and CWRS wheat kernels using SWIR spectroscopy. *Sensing and Instrumentation for Food Quality and Safety, 4*, 95–100.

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

Zhu, D., Wang, C., Pang, B., Shan, F., Wu, Q., & Zhao, C. (2012). Identification of wheat cultivars based on the hyperspectral image of single seed. *Journal of Nanoelectronics and Optoelectronics, 7*, 167–172.