Standardization of near infrared hyperspectral imaging for wheat single kernel sorting according to deoxynivalenol level

Antoni Femenias\textsuperscript{a}, Maria Belén Bainotti\textsuperscript{a}, Ferran Gatius\textsuperscript{b}, Antonio J. Ramos\textsuperscript{a}, Sonia Marín\textsuperscript{*}\textsuperscript{a}

\textsuperscript{a} Applied Mycology Unit, Food Technology Department, University of Lleida, UTPV-XaRTA, Agrotecnio, Av. Rovira Roure 191, 25198, Lleida, Spain

\textsuperscript{b} Departament de Química, Universitat de Lleida (UdL), Av. Rovira Roure, 191, Lleida, 25198, Spain

Corresponding author:

Sonia Marín Sillue

Food Technology Department, University of Lleida

Av. Rovira Roure 191, 25198, Lleida, Spain

Tel: (+34) 973 70 25 42

Fax: (+34) 973 70 25 96

Email: sonia.marin@udl.cat
Abstract

The spatial recognition feature of near infrared hyperspectral imaging (HSI-NIR) makes it potentially suitable for Fusarium and deoxynivalenol (DON) management in single kernels to break with heterogeneity of contamination in wheat batches to move towards individual kernel sorting and provide more quick, environmental-friendly and non-destructive analysis than wet-chemistry techniques, and to replace commonly used time-consuming and destructive techniques. The aim of this study was to standardize HSI-NIR for individual kernel analysis of Fusarium damage and DON presence, to predict the level of contamination and classify grains according to the EU maximum limit (1250 µg/kg).

Visual inspection on Fusarium infection symptoms and HPLC analysis for DON determination were used as reference methods. The kernels were scanned in both crease-up and crease-down position and for different image captures. The spectra were pretreated by Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV), 1st and 2nd derivatives and normalisation, and they were evaluated also by removing spectral tails.

The best fitted predictive model was on SNV pretreated data ($R^2 0.88$ and RMSECV 4.8 mg/kg) in which 7 characteristic wavelengths were used. Linear Discriminant Analysis (LDA), Naïve Bayes and K-nearest Neighbours models classified with 100% of accuracy 1st derivative and SNV pretreated spectra according to symptomatology and with 98.9 and 98.4% of correctness 1st derivative and SNV spectra, respectively. The starting point results are encouraging for future investigations on HSI-NIR technique application to Fusarium and DON management in single wheat kernels to overcome their contamination heterogeneity.
Keywords: Hyperspectral imaging; Deoxynivalenol; Single Kernel; Near infrared;

Cereal sorting.

Nomenclature

DON Deoxynivalenol
HPLC High Performance Liquid Chromatography
UHPLC Ultra High Performance Liquid Chromatography
ELISA Enzyme-Linked Immuno-Sorbent Assays
NIR Near Infrared
HSI-NIR Hyperspectral Imaging Near Infrared
SK-NIR Single Kernel Near Infrared
FDK Fusarium-Damaged Kernels
LDA Linear Discriminant Analysis
PCA Principal Component Analysis
FHB Fusarium Head Blight
LC-MS Liquid Chromatography – Mass spectrometry
PLS Partial Least Squares
PLS-DA Partial Least Square Discriminant Analysis
WL Wavelength
PC Principal Component
IAC Immunoaffinity Chromatography
Deoxynivalenol (DON), also known as vomitoxin, is a *Fusarium*-produced mycotoxin that presents high incidence in wheat. Its ingestion results in a wide range of harmful effects, in which gastroenteritis and potential chronic diseases are included (Eriksen & Pettersson, 2004; Pestka & Smolinski, 2005). A significant cereal consumption in occidental countries associated to an increased incidence of mycotoxins due to global climatic change, in which raised temperatures promote their production, is an issue of special concern for human and livestock health (Nesic, Milicevic, Nesic, & Ivanovic, 2015; Uhlig et al., 2013). Due to this exposure, the European Commission established DON maximum limits for unprocessed durum wheat (1750 μg/kg), unprocessed wheat other than durum wheat (1250 μg/kg) for human foodstuffs (European Comission, 2006b) and recommended limits for animal feed (5 mg/kg) (European Comission, 2006a). As its presence seems to be insufficiently reduced in the pre-harvest stages, post-harvest strategies are highly demanded for mycotoxin reduction in food industry. It has been
demonstrated that DON is not fully destroyed during food processing (Vidal, Sanchis, Ramos, & Marín, 2016), consequently its reduction needs to be implemented before entering in food industry.

To date, several analytical methods are accessible for DON detection in wheat batches, including HPLC, ELISA and mass spectrometry. These techniques are time-consuming, expensive and, results depend on sample collection in food industry entrance. They can present false negatives or positives depending on the sample collection, increase batch rejections. Those refusals or acceptances are caused due to the DON heterogeneity in wheat lots. The methods of cereal sampling established by the European Comission (2006) lay down that an aggregate sample weight of 10 kg obtained from 100 incremental samples is analysed for batches larger than 50 tonnes. Consequently, some highly contaminated kernels can disrupt the whole batch admission in food industry according to UE legal limit, with high impact in food chain sustainability. This problem could be avoided by a rapid analysis detection of these over-contaminated grains which could be discarded from the whole batch. As a result, a subbatch in which all the kernels exceed the DON established limits could be obtained and discarded.

HSI-NIR has been proposed as a rapid, low cost and eco-friendly technique which permits the analysis of chemicals in a specific sample position. This method combines the spectral features of each pixel location in the image acquired. Thus, the spatial skill of HSI allows the discernment of an individual kernel light spectrum (Fox & Manley, 2014). The use of these information could be used to calibrate models for DON quantification and to discard over-contaminated kernels from whole batches.

Revision articles compiled several studies referring to similar purposes attempting effective systems for cereal sorting according to fungal and DON contamination
The application of SK-NIR has been used by authors for assessing FDK and DON contamination. Some studies with the same aim used near infrared technology recording single grain reflectance (Jin et al., 2014; Peiris, Bockus, & Dowell, 2016; Peiris et al., 2010).

However, the technology developments introduced spatial features to spectroscopic measurements which permitted single kernel recognition in a whole image. Several authors have used HSI-NIR for *Fusarium* damage detection and classification in wheat single grain. Singh, Jayas, Paliwal, & White (2009) and Delwiche, Kim, & Dong (2010, 2011) built discriminant analysis to detect FHB and sprout in samples with high DON contaminations, by selected wavelengths in NIR and Vis/NIR region spectra, respectively. All these studies used visual inspection as reference method to discern as sprout, midge or *Fusarium* damaged and healthy kernels, thus DON detection quantification was not carried out. Moreover, the same method was used by Shahin & Symons (2011), which classified FDK and healthy kernels by LDA based on PCA scores. The FDK reference characterisation was according to visual classification of abundant or low mycelial growth, thus they used mouldy kernels with advanced contamination instead of natural infected ones found in field. In addition, whole Vis/NIR spectrum (450-950 nm) and six selected wavebands were used to define the most accurate model, also based on visual inspection. Both studies from Barbedo, Tibola, & Fernandes (2015) and Ropelewska & Zapotoczny (2018) used HSI-Vis/NIR technology for FHB and *Fusarium graminearum* infection in individual wheat kernels, respectively. The first study used a complex system based on *Fusarium* index, which reveals the likelihood of the kernel to be infected, to build a probability distribution function discerning between FHB and sound kernels and correlated it with DON. DON levels were
analysed as the mean of the contamination in each image, so the kernels were not analysed individually. Based on these results, Barbedo, Tibola, & Lima (2017) built a confusion matrix a classification algorithm to separate kernels based on k values to classify kernels in two and three categories (<500; 500-1250 and >1250 μg DON/kg), below 500 μg DON/kg, between 500 and 1250 μg DON/kg and above 1250 μg DON/kg (legal EU limit). In both cases, they used direct competitive DC-ELISA and LC-MS to analyse reference DON concentrations of 30-50 kernel lots instead of single kernels. Conversely, the second research selected three characteristic wavelengths to compare LDA, decision tree (LMT), partial decision tree (PART), Naïve Bayes and K star classifications from artificially infected and visually examined kernels. Delwiche, Rodriguez, Rausch, & Graybosch (2019) used the potential of HSI-NIR to classify sound and FDK by PLS-DA and LDA. The first classifier obtained the best results for two wavelengths selection and the second model obtained the best accuracies selecting four wavelengths between 1000-1400 nm. Single kernels arising from a severe Fusarium infected cultivar were also visually checked, for which not common field contamination was assessed were examined and they introduced the inspector subjectivity to the analysis. Furthermore, Alisaac et al. (2019) investigated FHB in artificially inoculated wheat kernels and flour by HSI-Vis/NIR by the comparison of the spectral signature reflectance. Spectral profiles of FDK presented higher reflectance’s than non-infected ones. HSI-NIR could also be used to early recognize FHB in field by spikelet’s analysis. Alternatively, Zhang, Pan, Feng, & Zhao (2019) targeted Fusarium damage in artificially inoculated wheat by FHB classification index. They worked directly on spikelets instead of single wheat kernels using was calculated by a first an extraction of four sensitive wavelengths and then a PLS to determine the best difference spectral index (based on
two bands divergences (two bands extraction) to calculate both best combinations to detect
and classify healthy and damaged wheat kernels according to the index.

Despite of the studies applying HSI-NIR on single wheat kernels for detection of DON
and classification, batches heterogeneity still remains a trouble in the so far published
studies as they use DON contaminations of a kernel set or whole batches to as reference
values and sometimes artificial contaminations. The aim of the present work is to
overcome the kernel diversity by a standardization of HSI-NIR to screen kernels in
accordance to *Fusarium* symptoms and DON levels. To reach this purpose, this study
works on naturally-contaminated samples and the same kernels scanned by HSI-NIR are
individually analysed by the reference method. In addition, spectral pretreatments,
wavelength selection and model calibrations were performed, mainly using PLS for
single kernel DON quantification and LDA for individual kernel classification according
to DON and symptomatology.

2. Materials and methods

2.1. Wheat samples

Wheat samples were supplied by a feed producing agricultural cooperative during 2018-
2019. Their origin was the plain area of Lleida province. They were taken within its
quality control programme from each incoming truck. From the whole homogenized
sample, a subsample (200–500 g) was sent to our laboratory. A total of 18 kernels from
a non-contaminated sample were used to validate the UHPLC analytical method and 50
kernels from a highly-contaminated one were used for the experimental analysis. A
flowchart representation of the work conducted in this study is provided in the Fig. 1.
2.2. Visual symptoms assessment

The 50 kernels used for the experimental study were manually selected and categorized into three levels (symptomatic, mildly-symptomatic and asymptomatic) according to visual symptoms of fungal infection. Discoloured, shrivelled and wrinkled kernels were considered as symptomatic (S). Kernels with part of this symptoms were categorised as mildly-symptomatic (M) and kernels with no visible signs as asymptomatic. The kernels were selected trying to cover, as wide as possible, all the visual kernels characteristics. Consequently, the percentage of kernels with visual symptoms in our working sample set was higher than in the original sample.

2.3. HSI-NIR experimental work

2.3.1. Instrumentation and data acquisition by HSI-NIR

A Pika NIR-320 camera assembled in RESONON Inc. (Boezman, MA, USA) based on a push-broom hyperspectral imaging system was used. The device consists in an InGaAS sensor line scan camera with 320 × 256-pixel resolution, a 30 × 30 μm pixel size and a 14-bit resolution A/D spectrograph (Goldeye G-008 SWIR TEC1, Allied Vision Technologies GmbH, Germany). The framerate was 520 fps in combination with a spectral resolution of 4.9 mm (164 spectral bands from 900 to 1700 nm) and a spatial resolution of 320 pixels. The objective lens was positioned 220 mm above the scanning surface and it had 24 mm of focal length (F/1.4 SWIR, 0.9–1.7 μm, 21 mm image format, c-mount). The illumination device consisted in four halogen lamps lighting system with Lambertian filters fixed on an adjustable tower. The lights were turned on at least 20 minutes before taking the images to ensure the stability of light beams. The illumination system was powered by Samplexpower® power converter (SEC-1223CE, Burnaby, BC, V5A 0C6, Canada) which provides a highly regulated output DC voltage of 13.8 V at 23
Amps with an AC input of 230 V, 50 Hz. Lastly, the fixed optical devices could register full samples due to a motor-powered linear translation stage of 600 mm.

The Spectronon PRO software permitted the Resonon’s benchtop control before and after the image acquisition. The captured intensities of each sample data array were automatically transformed to reflectance by dividing the dark current-subtracted intensity by the dark current-subtracted white standard intensity at each of the corresponding wavelengths (1). A dark current intensity image was collected before samples' scanning to remove dark current noise by covering the camera lens. Likewise, intensity from a 99% reflectance standard, made of polytetrafluoroethylene (PTFE) (Spectralon™, SRT-99-120, Labsphere, North Sutton, NH, USA) to correct lighting effects, was registered immediately after the dark current image. These two images were applied to the subsequent sample intensities as follows:

\[
I = \frac{I_0 - I_b}{I_w - I_b}
\]  

(1)

where \(I_0\) is the raw hyperspectral image obtained, \(I_w\) is the white reference and \(I_b\) is the dark current reference. In addition to dark and absolute reflectance response, the pixel illumination saturation was also adjusted by the camera controls. Framerate and integration time were established so that no pixel on the image was saturated.

The 50 wheat kernels were first scanned in triplicate individually for both positions (crease-up and crease-down). The kernel location was adjusted, so all kernels were scanned at the same image position. Mean raw spectrum and mean first derivative spectrum for each kernel was recorded by the Spectronon PRO software. Subsequently, the same 50 kernels were placed on the same scanning tray but shifting positions and the process was repeated in triplicate for both kernel faces. Images were adjusted to 350 bands for horizontal size and approximately 90 mm of vertical size. The pixel selection
was done by the collection of the mean and mean first derivative reflectance's of similar
spectrum pixels by Euclidian distance that are best adjusted to the ROI to remove the
background signal. Both spectra for each kernel were recorded as text file for their
subsequent exporting to the spectral analysis software.

2.4. Determination of DON concentration in wheat kernels by UHPLC

2.4.1. Reagents and chemicals

Water was obtained from a Milli-Q® SP Reagent water system from Millipore Corp.
(Brussels, Belgium). Methanol (HPLC grade) was purchased from Scharlab (Sentmenat,
Spain). Mycotoxin standards of DON were purchased from Romer Labs (Tulln, Austria).
Imunoaffinity columns for DON (DONPREP®) were acquired from R-Biopharm
(Rhone LTD Glasgow, UK).

2.4.2. Preparation of DON solutions

DON concentration in the stock solution was checked by UV spectroscopy according to
AOAC Official Methods of Analysis, Chapter 49 (AOAC, 2005), obtaining a
concentration of the stock solution of 962 μg/mL. Standard solutions of DON were
prepared in methanol at a concentration of 10 μg/mL and stored at 4 °C. Calibration
curves were prepared by appropriate dilution of known volumes of the stock solution with
the mobile phase.

2.4.3. Immunoaffinity column extraction of DON from wheat kernels for
analytical method validation

A total of 9 kernels from a non-contaminated sample, previously analysed twice by
UHPLC (< LOD), were used. Individual kernels were spiked at different contamination
levels prior to DON extraction with specific immunoaffinity columns (DONPREP®)
following the manufacturer's instructions. A single wheat kernel previously crushed and
pulverized with a small mortar and pestle was mixed with 1.5 mL of MiliQ water in a 1.5 mL Eppendorf tube, followed by 10 min stirring. Then, samples were centrifuged for 10 min at 1780×g. This process was repeated three times, obtaining a supernatant extracted final volume of 4.5 mL. The filtrate was passed through the immunoaffinity column and it was subsequently washed with 10 mL of bi-distilled water and the toxins were eluted with 3 mL of methanol HPLC-grade (the first 1.5 mL performing back-flushing). Samples were evaporated under a low nitrogen stream at 40 °C and resuspended in 0.3 mL of mobile phase (methanol:water, 10:90, v/v). Every resuspended extract was filtered through a nylon filter (0.4 μm) before being injected into the UHPLC-DAD system.

2.4.4. Direct extraction of DON from wheat kernels for analytical method validation

Other 9 kernels from the same sample used in the previous section were also spiked at two different levels before their direct extraction. Concisely, individual wheat kernels, previously ground with a small laboratory mortar and pestle, were mixed with 0.3 mL of MiliQ water in 1.5 mL Eppendorf tubes, followed by 10 min vortex. Then, samples were centrifuged for 10 min at 1780×g. Supernatant was filtered through a nylon filter (0.4 μm) before being injected into the UHPLC-DAD system.

2.4.5. Performance of the UHPLC analytical method

Selectivity was checked by injecting 50 μL of standard solution at least three times (150 μg/L) and comparing retention time and peak resolution between injections. For linearity check, a calibration curve of eight concentration levels for DON solutions (20, 30, 50, 100, 250, 500, 1000, 3000 μg/L solutions) was prepared and injected into the system, generating a linear regression plotting solutions' concentration versus peak area according to the methodology used by (Wall-Martínez et al., 2019). Method performance was assessed according to Commission Regulation (EC) 401/2006 (European Comission,
2006c). Precision was evaluated preparing blank kernels and kernels spiked with DON at several concentration levels and recovery percentages were determined: 64–105% (direct extraction), 68–125% (IAC extraction) (Table 1).

The recovery rates obtained for DON in single kernels for direct and IAC extraction were similar. Still, the direct extraction seems to be more suitable for small samples in order to avoid steps in which the toxin could be lost. Furthermore, the direct extraction was selected, apart from being effective according to the European Commission (2006c), for its quickness and lower costs.

2.4.6. Direct DON extraction from wheat kernels for experimental work

A highly-contaminated sample was selected, previously analysed twice by UHPLC (2682.8 and 2403.5 µg/kg). A total of 50 wheat kernels from this sample were selected with a mean weight of 33 mg in a range between 11.3 to 53.4 mg per kernel. DON was extracted from each single kernel with Mili-Q water following the same methodology used in the direct DON extraction for the validation of the analytical method.

2.4.7. UHPLC system

The determination of DON was performed using an Agilent Technologies 1260 Infinity UHPLC system (California, USA) coupled with an Agilent 1260 Infinity II Diode Array Detector (DAD). A Gemini® C18 column from Phenomenex 150 × 4.6 mm (California, USA) with a particle size of 5 µm and a pore size of 110 Å was used. Absorption wavelength was set at 220 nm. The mobile phase was composed of methanol:water (10:90, v/v) and set at a flow rate of 1 mL/min. The column temperature was 40 °C, the injection volume was 50 µL and total run time was 15 min for mycotoxin analyses. The performance of the method for the quantification of DON in wheat was tested, in which the limit of detection (LOD) was considered to be three times the signal of the noise (100
μg/kg). Once the concentrations were achieved, the kernels were categorised as contaminated above the EU legal limit (C) and below the limit (B) for subsequent classifications.

### 2.5. Hyperspectral data modelling for quantification and classification of DON contaminated kernels

Spectral data were processed using The Unscrambler software (version 7.6 SR1, CAMO, Oslo, Norway, 2001). First, to compare the spectral differences between high-contaminated (>1250 µg/kg) and low-contaminated kernels (<1250 µg/kg), a linear representation of the mean spectral profile was performed. Once the reflectance differences were established, the spectral information of single kernels was used to calibrate a DON prediction model. The calibration was performed by recording mean reflectances of the characteristic pixels from the kernel as the explanatory variables (X) and the DON concentrations (from <LOD to 79.7 mg/kg) obtained by UHPLC as the dependent variables (Y). A total of 150 images (arising from 50 kernels) were obtained to develop leave-one-out cross-validated prediction and classification models, in which a single sample is left out of the training set for each iteration in order to use it for the validation of the corresponding model. This validation is considered to be the most realistic validation in order to obtain the closest parameters of prediction to the independent set validation. However, it is used when it is not possible to obtain large sample size to separate them between calibration and validation sets that can guarantee the representativeness of the population of samples. In our case, our sample size is based on 150 images (from 50 kernels), thus we used leave-one-out cross-validation, in which a single sample is left out of the training set for each iteration. The recorded spectra from Spectronon PRO were obtained as the mean of the characteristic pixel’s spectra and the mean spectral curve within the ROI as a function of wavelength in the spectral plotter of
crease-up and crease-down positioned kernels for single kernel image and fifty kernel image.

First, the spectral data were preprocessed by the application of different spectral modifications apart from the raw and first derivative ones obtained directly from the Spectronon PRO software. The spectral array was imported to The Unscrambler software and scatter correction pretreatments were applied to minimise the non-linear effect of light scatter due to particle size differences among samples, as Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV). First and second derivatives were used to avoid noise, multiplicative and additive effects, by applying a 5-point Savitzky-Golay Smoothing. There is no standard optimum pretreatment for a specific type of data, as it depends on the signal (reflectance, absorbance or transmittance), sample features, instrument configuration or trial (prediction or discrimination). Consequently, its selection needs a trial and error process followed by experience.

Possible nonlinearities between reflectances in the extremes of the spectra and the compound for single kernel seeds were reported by Delwiche (1998), thus we aimed to check if there was a substantial enhancement in the model when the extremes were omitted. Moreover, Esteve-Agelet, Armstrong, Romagosa-Clariana, & Hurburgh (2012) also improved their model performance by removing the noisy extremes of the NIR range in single kernel analysis. Consequently, we also subtracted the spectral tails, leaving a remaining short wavelength range of 1000-1650 nm.

Prediction models were developed for short and large spectral range and for each pretreatment by PLS regression. The models were refined to obtain the lower Root Mean Square Error of Prediction (RMSEP). The RMSEP is defined as:

\[
RMSEP = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{n}}
\]
In which $\hat{y}_i = \text{i}^{th}$ validation sample predicted value; $y_i = \text{i}^{th}$ validation measured values; $\bar{y}$ = mean of the n values (measured or predicted), n = number of samples.

For their depuration, data points which differed significantly from the other samples were considered outliers. The criteria followed for outlier exclusion was to represent the residual influence plot, which represented the residual Y variance versus the leverage. Those samples that presented high values of both parameters were removed one by one and the models were recalculated for each rejection. A maximum percentage of 10% of the original images spectra were removed. In addition, the overriding criterion for the optimal Principal Components (PC) of the model selection was the first PC which presented a minimum on the RMSEP curve. Finally, the characteristic wavelength (the spectral band which had the highest weight on the model) was selected by the regression coefficient representation to build a PCA able to separate high (> 1250 µg/kg) and low (< 1250 µg/kg) contaminated DON kernel groups, marked as (B) and (C) in the previous section. The performance parameters determined the fitness of the models, which consisted in the slope and offset, correlation (R), coefficient of determination ($R^2$), Root Mean Square Error of Calibration (RMSEC), RMSEP, Square Error of Prediction (SEP) and bias. Finally, the ratio of percent deviation (RPD), which is the ratio between the laboratory measured data standard deviation to the RMSE of cross-validation, was estimated (2) according to Rossel, Mcglynn, & Mcbratney (2006):

$$\text{RPD} = \frac{\text{Sdev}}{\text{RMSE}}$$

(2)

The RPD was interpreted by distinguishing between five levels: excellent predictions (RPD > 2.5); good (RPD of 2.0 to 2.5); approximate quantitative predictions (RPD of 1.8 to 2.0); possibility to distinguish high and low values (RPD of 1.4 to 1.8); and unsuccessful (RPD < 1.40).
The best fitted models were selected following the combination of the closest to 1 for slope and $r^2$, the highest RPD and the lowest RMSECV and number of PC. Most of the models presented excellent predictions as they achieved RPD > 2.5, due to low errors of predictions in models and a high standard deviation in reference values. For the best models, characteristic wavelengths were distinguished for their high positive and negative regression coefficients. They were remodelled by using only the spectral bands which contributed substantially to explain DON variance within the model. The same 50 kernels were used to classify them according to visual symptoms and DON. Statistical analysis was performed with JMP PRO 14.1.0 (SAS Institute Inc., 2018) software using Linear Discriminant Analysis (LDA), Naïve Bayes Confusion, K-Nearest Neighbours and Neural Networks. Classification models were also developed by leave-one out cross-validation. The concentration threshold was established was 1250 µg/kg. As stated before, the spectra from kernels above this limit formed the group (C) and spectra from kernels below the limit formed the group (B). The models performances were evaluated according to the classification accuracy and the false negative percentages. Neural networks tended to over-fit data for such a reduced data set, thus it was discarded from the results. However, future testing on a wider set of samples will be essential to develop a robust neural network model.

3. Results

3.1. Comparison of spectral profiles

The spectral mean profiles of crease-down and crease-up positioned kernels for single kernel images and crease-down located for one image containing the fifty kernels are represented in Fig. 24. The blue-coloured spectral profiles that belong to the high-contaminated kernels (> 1250 µg/kg) and the red-coloured profiles that refers to the low
contaminated grains (< 1250 µg/kg) are represented also in the Fig. 2. Contaminated kernels presented higher reflectances than low-contaminated ones, although kernels spectral shapes remained visually similar in all the cases. Consequently, as we cannot attribute any change in a specific spectral band associated to DON concentration, multivariate analysis methods are needed to explain the spectral variability in association to DON contamination.

3.2. Single kernel DON variability

Table 2 includes the descriptive statistics of the kernels used in the experimental work. As indicated before, the results correspond to the direct DON extraction methodology. The parameters reported are the DON content and single kernel weight. The minimum DON content of a single kernel was below the LOD, while the maximum was 78.7 mg/kg. The mean concentration of the dataset was 12.9 mg/kg. Interestingly, the differences between this mean and the original sample concentration, analysed twice (2682.8 and 2403.5 µg/kg), must be explained by the manual selection of the wider symptomatology range. Furthermore, the single kernel weight range includes loads from 11.3 mg to 53.4 mg. Fig. 3 shows the frequency and distribution of single kernel DON content measured by the UHPLC system used in this study. Fig. 3a describes the frequency and distribution of samples in regard to DON contents. A 68 % of the kernels had a DON content covering a range from below the LOD to 10 mg/kg, which included the maximum UE legal limit (1.25 mg/kg). The rest of the kernels presented lower frequencies for higher intervals, as higher concentrations of DON are more rarely found in naturally contaminated wheat. The distribution of the reference DON content in the single kernels is also represented. It is asymmetric under the mean as a 40% of the total sample set is presented DON concentrations below the LOD. Thus, the distribution does not follow a Gaussian distribution due to many environmental and harvest causes. Fig. 3b shows the highest
frequency in kernels between 20-25 mg, which corresponds to the 20% of the entire sample set. Kernel weight could vary due to fungal infection, as its weight loss is a characteristic symptom of fungal growth.

3.3. Quantification of DON concentration in single wheat kernels based on spectral data

The quantifications were performed considering the effect of several conditions (pretreatment, spectral region, kernel position and number of images). All the regression performance parameters were collected in Table S1 for cross-validated models. No notable differences were detected for crease-up or crease-down position of kernels, either for individual kernel image or 50 kernel image. Also the effect of the kernel position and removing spectral bands from the extremes were estimated, leaving a wavelengths range from 1000-1650 nm.

Different spectral pretreatments were tested for DON quantification prior to PLS regression and they showed a substantial effect on the models performances. However, four models were selected for their high-performance and good adjustment for which their regression plots were represented in Fig. 43. The two best pretreatments were SNV and 1st derivative, or the combination of both.

The bands and the performance parameters of the models are collected in Table 3. The model A presented enhanced results, as it presents higher $R^2$ (0.88) and lower RMSECV (4.8 m/kg) with only 6 PC and 7 bands. In addition, model C presented slightly under-fitted results in comparison to the whole spectra model, thus a well-fitted model is obtained with a reduced number of wavelengths (4). As the models complexities are reduced to multispectral dimension by optimal wavelengths selection, they are more close to on-line DON detection.
3.4. Classification of single kernels according to visual symptoms by LDA

In order to discern between symptomatology, a first screening was performed by PCA to estimate the ability of HSI-NIR to discriminate between levels of symptomatology (asymptomatic, mildly-symptomatic and severe-symptomatic). Most of the models showed a tendency of the method to differentiate between those groups, as it was shown in Fig. 5. However, 1st derivative seemed to be the pretreatment which presented the best separation into three groups. Thus, classification models as LDA, Naïve Bayes and K-nearest neighbours were performed to assess the accuracy to classify samples grains according to typical visual symptoms (shrivelling, discolouration and wrinkling) in fungal-contaminated wheat. A classification model was built for each spectral pretreatment used in the DON quantification section. The classifications were accomplished contemplating the effect the pretreatments used in the DON quantification section.

The results showed that no differences were observed for those conditions, in exception of the spectral preprocessing. Thus, the mean accuracies for each analysis type and pretreatment were collected in Table 4. Naïve Bayes algorithms presented inefficient classifications because the conditional probabilities of class membership are assumed independent. Correct classifications above the 88.8 % to 97.6% were accomplished for K-nearest neighbour’s predictions, although they correspond only to its calibration, thus those results would be overestimated in comparison with a validated model. The highest accuracies were achieved for LDA cross-validated models, for 1st derivative spectra and for the combination of 1st derivative and SNV pretreatments in which the correctness was up to the 100%. The classification results including all the conditions independently are represented in Table 4.
3.5. Classification of single kernels according to DON levels by LDA

The 68.3% of the kernels agreed that they were either, on the one hand, symptomatic and contaminated over the EU limit or, on the other hand, asymptomatic and contaminated below limit. Thus, the remaining 31.7% were the grains that, although they had no symptoms, were heavily contaminated or that even if they had symptoms did not have DON. In addition, the DON content regarding to other components of the sample, e.g. starch, protein and moisture, makes the ability to discern between levels challenging.

A prior PCA analysis was performed to assess the ability of the NIR spectra to distinguish between sample grains contaminated above and below the EU maximum limit of DON (1250 µg/kg) (Fig. 65). The classifications of kernels according to 1250 µg/kg threshold presented results similar to symptomatology discrimination, although the match between symptomatology and DON levels was low. The highest discrimination accuracies were obtained for the cross-validated LDA models with a mean of 98.9 % for the 1st derivative spectra and 98.4 % for the combination of 1st derivative and SNV, represented in Table 5. For Naïve Bayes and K-nearest neighbour’s algorithms the classifications were from 73.4 to 81.9 % and 85.2 to 96.4 %, respectively. The compilation of the results for each condition is presented in Table 5.

4. Discussion

DON heterogeneous distribution from a single batch has been evidenced demonstrated in this irregular distribution, as its kernels can be infected from extremely high concentrations up to 79.7 mg/kg to undetectable levels. In addition, kernel weight could also affect NIR images acquisition, consequently its management is important.

Differences on reflectance intensities were observed both for Fusarium damage and DON in which high-DON level and symptomatic kernels seemed to present higher intensities.

Briefly, to summarize the results, the standardization of the methodology to quantify
DON and classify kernels according to visual symptoms and EU maximum limit presented positive results for a reduced sample set. It was determined that the spectra pretreatment with 1st derivative, SNV or the combination of both presented improved results for PLS regressions and LDA classifications. Moreover, the outcomes confirmed that the kernel position and the image acquisition (one image for each kernel or one image for all kernels) did not have effects on the results, thus the single image for dorsal positioned kernels was selected for its low complexity in operation. In addition, the recalibration of models using only the optimal wavelengths reduced the complexity and maintained their adjustment. Consequently, this methodology will offer a rapid scanning of a large amount of kernels for its possible adaptation to industry, as most of the applied on-line detection models in industry are based on multispectral analysis.

The present study achieved lower errors of prediction for DON than one of the first studies trying to quantify DON concentrations in single wheat kernels (Dowell, Ram, & Seitz, 1999) for which standard error was 44 mg/kg and R² of 0.64 for its calibration. For a validation with the 20% of kernels from the original 88 kernel set an error of 52 mg/kg and an R² of 0.66 were obtained. Even though our research was cross-validated and it is not entirely comparable with their validation results, the calibration resulted also weaker than the R² of 0.88 and 4.8 mg/kg RMSECV obtained in our work. Jin et al. (2014) and Peiris et al. (2010) evaluated the SK-NIR reflectance for FDK and DON estimation. The results obtained were different from ours since they worked with DON contamination ranges from 0.49 to 29.5 mg/kg and 0.2 to 1008.4 mg/kg, respectively. Nevertheless, the correlations obtained for Visual-FDK/SK-NIR FDK and GC-MS-DON/SK-NIR-DON were weak (0.72 and 0.49). Peiris et al. (2010) obtained an R² of 0.72 and a SECV of 154.2 mg/kg for DON quantification and a 98.8 and 99.9% of correct classification of
sound and damaged kernels and a 95.7 and 96.7% of correctness of DON contamination above and below 60 mg/kg, respectively. Thus, our results presented higher DON prediction and classification accuracies, despite of the different range of contamination and classification threshold. The inspection not only of the surface but also of the inner part of the kernel to detect *Fusarium* was performed by Polder, Van Der Heijden, Waalwijk, & Young (2005) measuring transmittance instead of reflectance's.

The spatial ability of hyperspectromics permitted to analyse a single kernel as the ROI from a whole image. As in our previous study (Femenias, Gatius, Ramos, Sanchis, & Marín, 2020a), PCA was evaluated previous to LDA classification in order to show differences in kernels symptomatology and DON contamination. The results confirmed that exists a tendency of discrimination of classes by extracting the most relevant information compressing it into new orthogonal variables, however, a LDA model is needed to overcome the covariance in data matrix. The use of a discriminant analysis by selecting specific wavelengths of the NIR and the Vis/NIR spectra was established by Singh et al. (2009), Delwiche et al. (2011) and Shahin & Symons (2011). The first research accomplished a LDA classification of 100% between healthy and damaged kernels with only three wavelengths (1101.7, 1132.2 and 1305.1 nm). Our studies also achieved the same accuracy using the whole spectra but using the 1st derivative or combined 1st derivative + SNV spectra pretreatment and our characteristic bands only can be comparable in the 1100-1200 nm region. Furthermore, Delwiche et al. (2011) LDA results achieved a mean accuracy of 95% for which the 1200 nm region (related with ergosterol absorption) was a key wavelength for damage determination. Finally, Shahin & Symons (2011) classified FDK kernels by a combined LDA-PCA method for which they classified correctly up to the 92% the FDK kernels for the validation set. Even though
we cross-validated the LDA, our accuracies are higher than the calibration results obtained by these authors (93%).

A method based on *Fusarium* index, which is the likelihood of a kernel of being infected by *Fusarium*, was used by Barbedo, Tibola, & Fernandes (2015) to assess FHB in 50 individual grains by HSI-NIR screening. It is essential to clearly state that visual inspection is subjective in terms of the fungal contamination perception. In some cases, damage is not perceived appreciable when fungal infection is found in early stages, thus its detection by human and devices remains indiscernible. The direct detection of FHB was 91%, however, the index correlation with DON contamination was (0.84).

Moreover, Ropelewska & Zapotoczny (2018) also classified *Fusarium* infected kernels using HSI-Vis/NIR and some of the discrimination models that we used (LDA and Naïve Bayes), even though they used a single selected wavelength. Their higher accuracies were obtained for ventral side for 550 and 710 nm bands, which are not included in our spectral range, for which the correctness were from 90-100% for LDA and Naïve Bayes. In our work, the classifications for LDA were more correct than for the Naïve Bayes models, which are in discordance with their results for which both models were similar. This can be explained because of our data scarcity and continuous variables in Naïve models can result in numerical instabilities and higher misclassification rates in comparison with a single band used in the discussed study.

Delwiche et al. (2019) and Zhang et al. (2019) evaluated FHB by HSI in individual kernels. The first study used LDA to evaluate the health status of kernels for which they granted a value of 0 and 1 for sound and damaged, respectively, in which 200 randomly orientated kernels were scanned. Only mean centering (except for one test in which SNV was used) was applied as spectral data pretreatment and a cross-validated model with selected wavelengths (1100, 1197, 1308 and 1394 nm) was developed. It was able to
classify with a 97.1% of correctness *Fusarium* damaged kernels and 96.4% of healthy ones. Their results were close to the obtained in this study, however, they achieved slightly lower accuracies probably due to the higher number of kernels (556).

**Moreover,** as in our study, determined that not important differences were observed between crease-up and crease-down kernel position, thus it would be interesting to test randomly individual kernels as in their study that is being discussed. On the other hand, Zhang et al. (2019) build a FDK *Fusarium Head Blight* classification index (FCI) to determine its damage on spikelets using hyperspectral microscopy imaging. They extracted four characteristic wavelengths for which two of them (668 and 417 nm) were used to calculate the FCI. Even though the technique was quite different than HSI-NIR, they obtained a FDK overall classification accuracy of 89.9%. This was quite similar to Alisaac et al. (2019) findings, for which HSI-Vis/NIR spectral signatures were used to determine correlations between fungal DNA and DON in wheat kernels.

Some studies tried to classify wheat by HSI-NIR according to DON contamination levels. The first study which reached this purpose was that of Barbedo et al. (2017), in which a two and three categories confusion matrix was developed (< 1250 µg/kg; > 1250 µg/kg and < 500; 500-1250; > 1250 µg/kg) for which they classified wheat batches with a 81 and 72 % of correctness, respectively. Even though they classified naturally-contaminated wheat according to DON, important differences with sampling conditions are noticeable, as they scanned 30-50 kernels instead of single kernel analysis and they analysed all the kernels together by ELISA as the reference method. Recently, Liang et al. (2020, 2018) also investigated on DON detection, but they used 70 wheat kernel samples instead of single kernels. Consequently, this research had a different aim from ours, as they assessed bulk wheat samples. Nevertheless, in their last publication, the SNV was assigned as the best spectral pretreatment for HSI-NIR (1000-2500 nm).
classifications which, in accordance with the present paper, it is appropriate for wheat
classification according to DON levels.

5. Conclusion

Preliminary analysis demonstrated that 16 from the 50 kernels analysed were *Fusarium*
damaged kernels not contaminated by DON or vice versa. We accomplished that the measured reflectances of DON contaminated kernels were higher than those of kernels with DON levels above below the UE limit. In addition, this study concluded, using the PCA analysis, that for all the spectral pretreatments showed a tendency pattern in the separation according to FDK and DON for all the spectral pretreatments used. This research dealt proposed both with quantification and sorting according to DON levels of single grains, for which it has been demonstrated that, despite of the RMSECV (4.8 mg/kg) is higher than the EU maximum limit, the high contaminations could be identified with a simplified model (7 WL and 6 PC). Moreover, the LDA cross-validated classifications presented promising results, achieving a 100% of accuracy for symptomatology prediction and 98.9% for DON (according to EU maximum level). Thus, based on these findings, HSI-NIR could be applied as an accurate kernel sorting technique. Once standardized the single kernel analysis standardized, further research will be required in order to build better fitted prediction models and increasing sample size to build robust classifications.

6. Conflicts of interest

The authors declare that they have no conflict of interest.
7. Acknowledgements

The authors are grateful to the University of Lleida (predoctoral grant), and to the Spanish Ministry of Science, Innovation and Universities (Project AGL2017-87755-R) for funding this work.
8. Bibliography


Dowell, F. E., Ram, M. S., & Seitz, L. M. (1999). Predicting scab, vomitoxin, and


Zhang, N., Pan, Y., Feng, H., & Zhao, X. (2019). Development of Fusarium head blight classification index using hyperspectral microscopy images of winter wheat
Figure Captions

Fig. 1. Flowchart of the experiments executed in this research.

Fig. 2. Symptomatology and DON effect on spectral profiles depending on kernel position and image acquisition (raw reflectance’s). A) Mean spectra of crease-down positioned kernels for each symptomatology category (one image per kernel). B) Mean spectra of crease-up positioned kernels for each symptomatology category (one image per kernel). C) Mean spectra of crease-down positioned kernels for each symptomatology category (one image for the 50 kernels). Categories: A = Asymptomatic (red); M = Mildly-symptomatic (blue); S = Symptomatic (green). D) Mean spectra of crease-down positioned kernels DON level (one image per kernel). E) Mean spectra of crease-up positioned kernels DON level (one image per kernel). F) Mean spectra of crease-down positioned kernels DON level (one image for 50 kernels). DON levels: > 1.25 mg/kg (blue); < 1.25 mg/kg (red).

Fig. 23. Distribution of DON content and kernel weight in single wheat kernels on the full dataset. A) Total kernel DON frequency and distribution. B) Total kernel weight frequency and distribution.

Fig. 34. Predicted vs measured plots of cross-validated models. Spectral range: 1000-1650 nm. A) SNV pretreated model for crease-up kernel position; n =135; optimum PCs = 10. B) 1st derivative pretreated model for crease-up kernel position; n =137; optimum PCs = 13. C) 1st derivative + SNV pretreated model for crease-up kernel position; n =135; optimum PCs = 8.
Fig. 4. PCA scores for visual symptoms screening. A = asymptomatic; M = mildly-symptomatic; S = symptomatic. A) Raw spectra, X-exp: 96%, 4%. B) Multiplicative Scatter Corrected spectra, X-exp: 83%, 7%. C) Standard Normal Variate corrected spectra, X-exp: 83%, 7%. D) 1st Derivative spectra, X-exp: 71%, 12%. E) 1st Derivative + Standard Normal Variate corrected spectra, X-exp: 62%, 17%. F) 2nd Derivative spectra, X-exp: 50%, 28%. G) Normalised spectra, X-exp: 99%, 1%.

Fig. 5. PCA scores for visual DON screening. B = < 1250 µg/kg kernels; C = > 1250 µg/kg kernels. A) Raw spectra, X-exp: 96%, 4%. B) Multiplicative Scatter Corrected spectra, X-exp: 83%, 7%. C) Standard Normal Variate corrected spectra, X-exp: 83%, 7%. D) 1st Derivative spectra, X-exp: 77%, 14%. E) 1st Derivative + Standard Normal Variate corrected spectra, X-exp: 62%, 17%. F) 2nd Derivative spectra, X-exp: 50%, 28%. G) Normalised spectra, X-exp: 99%, 1%.
Table 1. Performance of methods for the determination of DON from wheat kernels

<table>
<thead>
<tr>
<th>Mycotoxin</th>
<th>Extraction</th>
<th>LOD (µg/kg)</th>
<th>LOQ (µg/kg)</th>
<th>n</th>
<th>Spiking level (µg/kg)</th>
<th>Recovery (%)</th>
<th>RSDr (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DON</td>
<td>Direct</td>
<td>100</td>
<td>300</td>
<td>5</td>
<td>1000</td>
<td>87 ± 10</td>
<td>12</td>
</tr>
<tr>
<td>DON</td>
<td>IAC</td>
<td>100</td>
<td>300</td>
<td>4</td>
<td>1500</td>
<td>93 ± 17</td>
<td>19</td>
</tr>
<tr>
<td>DON</td>
<td>IAC</td>
<td>100</td>
<td>300</td>
<td>4</td>
<td>1250</td>
<td>80 ± 17</td>
<td>21</td>
</tr>
</tbody>
</table>

*a* LOD = limit of detection. *b* LOQ = limit of quantification. *c* Mean value ± standard deviation. *d* RSDr = relative standard deviation.
Table 2. Descriptive statistics for the variables used in the present work.

<table>
<thead>
<tr>
<th>Data set</th>
<th>Mean</th>
<th>Range</th>
<th>SDev</th>
<th>CV (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DON content (mg/kg)</td>
<td>12.9</td>
<td>&lt; 0.1 – 79.7</td>
<td>21.3</td>
<td>163.2</td>
</tr>
<tr>
<td>Kernel weight (mg)</td>
<td>33.0</td>
<td>11.3 – 53.4</td>
<td>11.4</td>
<td>34.6</td>
</tr>
</tbody>
</table>

SDev = standard deviation. CV = coefficient of variation.

Table 3. Performance parameters of PLS regressions from selected optimal wavelengths.

<table>
<thead>
<tr>
<th>Model</th>
<th>Optimal wavelengths (nm)</th>
<th>Slope</th>
<th>RMSECV</th>
<th>R²</th>
<th>SEP</th>
<th>PC</th>
<th>RPD</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. SNV</td>
<td>1198, 1322, 1353, 1428, 1445, 1497, 1549</td>
<td>0.88</td>
<td>4.8</td>
<td>0.88</td>
<td>4.8</td>
<td>6</td>
<td>4.4</td>
</tr>
<tr>
<td>B. 1st Derivative</td>
<td>1112, 1205, 1345, 1401, 1452, 1499, 1525, 1541</td>
<td>0.79</td>
<td>8.1</td>
<td>0.78</td>
<td>8.1</td>
<td>6</td>
<td>2.6</td>
</tr>
<tr>
<td>C. 1st D + SNV</td>
<td>1325, 1396, 1406, 1421</td>
<td>0.81</td>
<td>6.1</td>
<td>0.81</td>
<td>6.1</td>
<td>3</td>
<td>3.4</td>
</tr>
</tbody>
</table>
Table 4. Classification accuracies for single kernel classification according to symptomatology. A = asymptomatic; M = mildly-symptomatic; S = symptomatic. Correctly-classified kernels correspond to grey cells numbers.

<table>
<thead>
<tr>
<th>Symptomatology</th>
<th>Pretreatment</th>
<th>Raw spectra</th>
<th>MSC</th>
<th>SNV</th>
<th>1st Derivative</th>
<th>2nd Derivative</th>
<th>1stD + SNV</th>
<th>Normalisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>56.2 4.7 0.3</td>
<td>63.2 3.5 0.3</td>
<td>62.5 4.3 0</td>
<td>69 0 0</td>
<td>58.3 3 5.8</td>
<td>69 0 0</td>
<td>59.7 4 0</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>9.3 17.8 10</td>
<td>3.3 22 3.7</td>
<td>3.5 20.5 5.5</td>
<td>0 27 0</td>
<td>4 21 2.2</td>
<td>0 27 0</td>
<td>7.7 19.8 6.3</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>3.5 4.5 43.7</td>
<td>2.5 1.5 50</td>
<td>3 2.2 48.5</td>
<td>0 0 54</td>
<td>6.7 3 46</td>
<td>0 0 54</td>
<td>1.7 3.2 47.7</td>
<td></td>
</tr>
<tr>
<td>Naïve Bayes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>55.5 15.3 10.2</td>
<td>54.5 11.8 8</td>
<td>31.7 6.8 5.2</td>
<td>46.5 3.5 0.5</td>
<td>53.2 5 1.3</td>
<td>55 7.5 2.8</td>
<td>50.5 3.2 0</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3.5 3 12</td>
<td>7.5 10 10.2</td>
<td>29.8 14.8 15.7</td>
<td>19.5 20 5.5</td>
<td>12.2 18.5 6.5</td>
<td>6.2 13.2 7</td>
<td>14.5 18.8 11.7</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>10 8.7 31.8</td>
<td>7 5.2 35.8</td>
<td>7.5 5.3 33.2</td>
<td>3 3.5 48</td>
<td>3.7 3.5 46.2</td>
<td>7.8 6.3 44.2</td>
<td>4 5.0 42.3</td>
<td></td>
</tr>
<tr>
<td>K-nearest Neighbours</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>64 4.3 0.7</td>
<td>62.7 4.8 2.2</td>
<td>62.7 4.7 2.8</td>
<td>66.3 3.8 1.8</td>
<td>65 3 3.2</td>
<td>66.2 3.3 3.3</td>
<td>68 1.5 0.3</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>3 20.2 3.2</td>
<td>4.5 20.2 1.5</td>
<td>4 20.5 1.3</td>
<td>1.7 21.3 1.3</td>
<td>3.3 21.7 1.7</td>
<td>1.8 22.2 1.2</td>
<td>1 24.8 0.2</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>2 2.5 50.2</td>
<td>1.8 2 50.3</td>
<td>2.3 1.8 49.8</td>
<td>1 1.8 50.8</td>
<td>0.7 2.3 49.2</td>
<td>1 1.5 49.5</td>
<td>0 0.7 53.5</td>
<td></td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td></td>
<td>78.4 90.1</td>
<td>87.7</td>
<td>100</td>
<td>83.6</td>
<td>100</td>
<td>84.8</td>
<td></td>
</tr>
</tbody>
</table>

36
Table 5. Classification accuracies for single kernel classification according to DON levels. B = < 1250 µg/kg; C = >1250 µg/kg. Correctly-classified kernels correspond to grey cells numbers.

<table>
<thead>
<tr>
<th>Pretreatment</th>
<th>Raw spectra</th>
<th>MSC</th>
<th>SNV</th>
<th>1st Derivative</th>
<th>2nd Derivative</th>
<th>1st D + SNV</th>
<th>Normalisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>LDA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DON levels</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B C</td>
<td>65.5</td>
<td>24.3</td>
<td>74.2</td>
<td>10.7</td>
<td>72.8</td>
<td>12.7</td>
<td>74.5</td>
</tr>
<tr>
<td>Accuracy (%)</td>
<td>9.5</td>
<td>50.7</td>
<td>0.8</td>
<td>64.3</td>
<td>2.2</td>
<td>62.3</td>
<td>0.5</td>
</tr>
</tbody>
</table>

| Naïve Bayes    |             |     |     |                |                |             |              |
| DON levels     |             |     |     |                |                |             |              |
| B C            | 63.7        | 28.5| 68.3| 20.5           | 66.2           | 28.3        | 61           | 18.2         | 63.5         | 21.7         | 62           | 21.5         | 58.7         | 23.5         |
| Accuracy (%)   | 11.3        | 46.5| 6.7 | 54.5           | 8.8            | 46.7        | 14           | 56.8         | 11.5         | 53.3         | 13           | 53.5         | 16.3         | 51.5         |

| K-nearest Neighbours |             |     |     |                |                |             |              |
| DON levels         |             |     |     |                |                |             |              |
| B C                | 64.2        | 11.3| 70.8| 7              | 69.5           | 7.7         | 69.5         | 6.7          | 67.7         | 7.3          | 70.2         | 6.3          | 72.3         | 2.7          |
| Accuracy (%)       | 10.8        | 63.7| 4.2 | 68             | 5.5            | 67.3        | 5.5          | 68.3         | 7.3          | 67.7         | 4.8          | 68.7         | 2.7          | 72.3         |

|             |             |     |     |                |                |             |              |
| B C         | 85.2        | 92.6| 91.2| 91.9           | 90.2           | 92.6        | 96.4         |              |              |              |              |              |              |              |
Figure 1

1. HSI-NIR single kernel scanning
   - ROI selection
   - Spectral extraction
   - Spectral preprocessing
   - Model calibration
   - Model validation
   - Characteristic wavelength extraction
   - Visual symptoms assessment

2. DON direct extraction
   - UHPLC analysis

3. IAC extraction
   - UHPLC performance determination

4. Kernel classification according to Fusarium damage
   - DON quantification
   - Kernel classification according to DON
Graphical Abstract
CRediT author statement

Antoni Femenias: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization.

Maria Belén Bainotti: Methodology, Software, Validation, Formal analysis, Investigation, Data Curation.

Ferran Gatius: Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Review & Editing.

Antonio J. Ramos: Resources, Writing - Original Draft, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition.

Sonia Marín: Conceptualization, Methodology, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Supervision, Project administration, Funding acquisition.
Declaration of interests

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

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