Use of hyperspectral imaging as a tool for *Fusarium* and deoxynivalenol risk management in cereals: A review

Antoni Femenias[1], Ferran Gatius[2], Antonio J. Ramos[1], Vicente Sanchis[1] and Sonia Marín[1*].

[1]Applied Mycology Unit, Food Technology Department, University of Lleida, UTPV-XaRTA, Agrotecnio, Av. Rovira Roure 191, 25198 Lleida, Spain

[2]Departament de Química, Universitat de Lleida (UdL), Av. Rovira Roure, 191, Lleida 25198, Spain

Abstract

Hyperspectral imaging (HSI) is an emergent, rapid, cost-effective and non-destructive technique in which the spectral data are obtained for each pixel location in a sample's image. The application of this technique to assess mycotoxins and mycotoxigenic fungi in cereals is considered promising to replace time-consuming wet-chemistry methods and for its potential grain sorting ability, in order to reduce food and feed contamination and the associated toxic effects in human and animals. *Fusarium* is a plant pathogen and deoxynivalenol (DON)-producer which presents high incidence in cereals such as wheat, maize and barley. The following review encompasses detailed information about the HSI principle and an updated outlook of its applications in the detection and quantification of *Fusarium* and DON in cereals. Moreover, HSI prediction algorithms for DON quantification are novel approaches which present high complexity owed to the asymptomatic nature of the grain despite of high mycotoxin concentrations. The spatial faculty of this system may be able to overcome the contamination heterogeneity of the grain for its elimination, enhancing risk management and rising the economic performance. Additionally, HSI is also proposed as a powerful grain sorting instrument due to high accuracies obtained in classification of single kernels according to *Fusarium* and DON infection. Therefore, an overview of the HSI applications for on-line and massive cereal sorting in grain industry is also presented.

Keywords: Hyperspectral imaging; near infrared; *Fusarium*; deoxynivalenol; cereal sorting

1. Introduction

Deoxynivalenol (DON) is a known mycotoxin produced mainly by the plant pathogens *Fusarium graminearum* and *F. culmorum*. Several environmental factors affect fungal growth and therefore cereal damage and toxin accumulation, such as temperature, water activity, pH and nutrient composition (Kokkonen, Ojala, Parikka, & Jestoi, 2010; Wegulo, 2012). Small grains *Fusarium* contamination is an existing problem around the world, especially in temperate regions in which these DON-producing species are commonly found (Sudakin, 2003). DON is

one of the major trichothecenes found in corn and small-grain cereals, and it is responsible of substantial economic impact, mainly in the major affected products, such as wheat, barley, oat and rye (Zain, 2011). It is known that mould growth and potential DON production before harvest and during storage can occur, favoured by cold and damp environment in field, or due to high moisture level by insufficient drying of cereals during trading (Magan, Hope, Cairns, & Aldred, 2003; Reddy et al., 2010). In addition to DON contamination, Fusarium Head Blight (FHB) is one of the most severe plant diseases, that causes loss of quality and reduction of the grain yield (Krnjaja, Levic, Stankovic, & Stepanic, 2011).

Numerous harmful effects have been reported owing to acute or prolonged DON exposure. Elevated dose ingestion induces headache, throat irritation, diarrhoea, nausea, vomiting and gastrointestinal haemorrhage. Severe consequences, as death, can be achieved at excessive exposures in a short period of time (European Food Safety Authority, 2007; Pestka & Smolinski, 2005). Otherwise, chronic administration of DON in animals is associated to nutritional disorders such as weight loss and anorexia and to immune problems, depending on the dose and exposure ratio, including immunosuppression and immunostimulation (Eriksen & Pettersson, 2004). Moreover, growth reduction related to emetic syndrome was detected in farm animals and, consequently, diminution in economic performance, especially in pigs (Placinta, D'Mello, & Macdonald, 1999). International Agency on Research on Cancer (IARC) evaluation of carcinogenicity in humans classified DON in Group 3 (not classifiable as to its carcinogenicity to humans) (IARC, 2012). However, considering the abovementioned adverse effects, the EU Scientific Committee on Food established a Tolerable Daily Intake (TDI) of 1 μg/kg bw/d (Sobrova et al., 2010).

Wheat-containing products such as flour, bread and other derived products are considered common daily consumption foodstuffs in European countries. Due to DON high stability, it is difficult to eliminate during cereal processing and in some cases, as bread-making, its concentration can also be increased (Vidal, Sanchis, Ramos, & Marín, 2016) In light of the exposure to DON through cereal intake, the European Commission (EC) set maximum levels for unprocessed durum wheat (1750 μg/kg), unprocessed cereals other than durum wheat, oats and maize (1250 μg/kg), flour and pasta (750 μg/kg) and bread (500 μg/kg) (EC, 2006). Consequently, adequate instrumentation is needed for DON detection in cereals.

Various techniques have been used for the detection and quantification of *Fusarium* and DON contamination in cereals. Typical chemical methods including Enzyme-linked Immunosorbent Assay (ELISA) and High Performance Liquid Chromatography (HPLC) are being used to determine DON concentration in wheat, but they have shown some drawbacks (de Girolamo, Lippolis, Nordkvist, & Visconti, 2009; Dowell, Ram, & Seitz, 1999). The most frequent disadvantages present in ELISA are false positives due to antibodies' cross-reactivity and

matrix dependence and false negatives caused by low sensitivity. Nevertheless, chromatography is characterized by a high sensitivity and selectivity, achieving precise mycotoxin quantification. Despite of the accuracy shown by these methods, they are time-consuming, expensive and sample-destroying. Consequently, farmers and food industry require alternative detection and quantification techniques able to substitute expensive laboratory facilities and qualified staff. Hence, the present article aims to review the latest improvements in HSI technology for fungal and mycotoxin assessment focused on *Fusarium* and DON detection.

2. Basis of Hyperspectral Imaging (HSI)

Near Infrared Spectroscopy (NIR) employs the spectral range from 780 to 2500 nm, offering information about the overtones and the combination of the molecular vibrations of the hydrogen molecular bonds (O-H, C-H, N-H and S-H) from the tested object (Cen & He, 2007). When the organic molecules are exposed to NIR frequencies, the vibration of the abovementioned bonds absorbs the spectral energy. Then, the rest of the chemical bounds reflect or transmit the other beams at different infrared wavelengths, which are dispersed and measured by the detector. In solid samples, the spectral bands in the NIR regions are wide and overlapped, so the spectra obtained is characterized by a smooth shape.

Preliminary studies of Dowell et al. (1999) demonstrated the ability of NIR for DON classification. They observed absorption patterns of DON which are determined by its functional groups (O-H, C-H and N-H), comprised between the NIR spectral region. The characteristic overtones of each group were (750, 950 and 1400 nm), (1200, 1400 and 1650 nm) and (1050 and 1500 nm) respectively, and the absorption wavelengths were used for DON detection.

NIR spectroscopy is a spatial limited technique, especially for heterogeneous samples measurement (Manley, 2014). Consequently, efficient technologies in the spatial characterization of the samples are required.

HSI is a potent technique based on the electromagnetic spectrum collection through the spatial positions of the measured object. Figure 1 shows a schematic representation of the three dimensions captured by the image, two of them corresponding to the spatial location of the pixel and the third one equivalent to the spectral data acquisition through NIR wavelength range. Additionally, an illustration of the spectral response for both individual pixel and the whole image is also presented. The multiple combinations between the imaging vectors (X, Y), which determine the pixel location in the image, and the spectral information vector for each pixel at different wavelengths create a three-dimensional hypercube containing an elevated amount of information (Dale et al., 2013). The HSI technique generates a spectral variation map showing many advantages such as a minimum and non-destructive sample manipulation, environmental-

friendly, extremely rapid measurements once validated, low cost analysis and detection of different chemical compounds at specific sample locations (Sendin, Williams, & Manley, 2018).

2.1. Equipment

The most common hyperspectral image acquisition method is the push-broom due to the capability of online scanning line by line while the sample is moving. A classic push-broom HSI system incorporates optical systems, illumination devices, a moving unit and a data acquisition instrument. The optical systems include the following three components: a charged coupled device (CCD) camera characterized by semiconductor electronic properties for spectral and spatial detection; an spectrograph, which is considered the key of the optical system, disperses the light into different wavelengths to generate an spectrum for each pixel of the image (Elmasry, Kamruzzaman, Sun, & Allen, 2012); and an objective lens which is coupled to the spectrograph to focus the light beam from the scanned object to the detector (camera). The illumination unit is characterized by the production of an homogeneous focus of light on the sample which may be free of radiation that could alter the sample (ElMasry & Sun, 2010). The sample is scanned line by line due to the incorporation of a moving system which permits the measurement of the full target having the optical systems fixed. The translation stage and the motor are the two main components of the moving device and they control the sample movement speed (Delwiche et al., 2017). Finally, the data acquisition instrument is based on a computer processing software which converts the raw data obtained from the measurements to band image data. Furthermore, it may also permit dark and light correction, image improvement, simple mathematical operations and high volume of data storage as spectral band images (Kim & Chen, 1998). Figure 2 is a schematic representation of a tower push-broom HSI system.

Recent studies have used HSI devices to screen *Fusarium* and DON contamination of cereal kernels. Tekle, Måge, Segtnan, & Bjørnstad (2015) used a SWIR (Short Wavelength Infrared) camera coupled to a Mercury Telluride (HgCdTe) detector with a spectral range of 1000-2500 nm. The data obtained was processed by a SpectralDAQ software. The imaging system used by Barbedo, Tibola, & Lima, (2017) was a XENICS camera combined with a VIS/NIR spectrometer working at wavelengths of 528-1785 nm. The illumination system consisted of a Quartz Tungsten-Halogen Lamp. In recent publications, Ropelewska & Zapotoczny (2018) acquired images with a charge-coupled device (CCD) camera, a VIS/NIR (400-1100 nm) spectrometer and a Fiber Optic Illuminator associated to a supplementary infrared lamp (600-1100 nm). Liang et al. (2018) also used a CCD camera but a different spectrometer (ImSpector V10) detecting 400-1000 nm wavelengths. The illumination source was a 150-W halogen lamp light. Finally, Alisaac, Behmann, Kuska, Dehne, & Mahlein (2018) used two different cameras

(Hyperspectral Camera ImSpector V10 and SWIR-camera) with wavelengths ranging from 400 to 1000 and 1000 to 2500 nm, respectively. The samples were illuminated with Analytical Spectral Devices PRO-Lamps and the spectral measurements of both cameras were controlled by a Spectral Cube Software.

2.2. Measurements

The three foremost methods for hyperspectral images production are push-broom imaging (two-dimension spectral information across the spatial axis line by line), whisk-broom scan (spectrum generation for a single pixel at a time) and staring imaging (generation of a spectral plan across the wavelength axis).

2.2.1. Push-broom imaging

A push-broom imaging scanner is a system which obtains spectral measurements for each pixel in a line. The basis of a push-broom imager is based on the sample movement in the Y axis direction, in which the spectrometer records the spectra for each pixel in the X-axis line. This technique displays good relationships between spatial and spectral resolution, hence it is widely used for on-line and in-line measurements (Boldrini, Kessler, Rebnera, & Kessler, 2012).

2.2.2. Whisk-broom imaging

A point-scanning method (or whisk-broom method) is an imaging system obtaining spectral measurement pixel by pixel. The hyperspectral image is acquired while the sample or the detector moves in the X and Y axis for a single position spectral acquisition. In this case, an exhaustive pixel by pixel data collection is needed to obtain high spectral resolutions of the physical and chemical information of the sample (Qin, 2010). Nevertheless, the method shows low spatial resolutions due to the long times used to measure the sample point by point.

2.2.3. Staring Imaging

The staring method is an area scanning technique characterized by the 2-D spatial (x, y) plane acquisition at a single spectral wavelength. The hypercube is built up with the collection of the spatial areas through the spectral domain for a determined number of wavelengths. Its selection is achieved by the illumination of the sample at a specific wavelength or by the monochromatic reflection analysis in the detector. In this case, the advantage is that the images obtained have high spatial resolution although lower spectral resolutions are obtained (ElMasry & Sun, 2010; Gupta, 2011).

The comparison between the studies using HSI shows that the predominant method used for HSI-NIR is a line-by-line scanning (push-broom imaging) at different pixel resolution (x, y, λ). Barbedo, Tibola, & Fernandes, (2015) and Barbedo, Guarienti, & Tibola, (2018) used this method to capture a 3D hypercube with a 320 x 800 x 256 dimension. Besides, Delwiche, Kim,

& Dong (2011) also acquired a push-broom imagine but with a different pixel resolution of 320 x 320 x 288. The hypercube dimension depends on the wavelength range (spectral resolution) of the camera and the field of view (spatial resolution) selected for the image capture.

2.3. Object measurement modes

Hyperspectral measurement modes are generally based on reflectance, transmittance or absorbance. The difference between the modes depends on how the light beam reaches the detector after its interaction with the sample. The light variances after the interaction with the cereal samples have a large impact in the spectral data acquisition and should be interpreted to evaluate sample features and correlate them with changes in its contamination. The light radiation of the illumination unit to the sample can be reflected, transmitted or absorbed, so that one of them is measured. Thus, the relationship between incident illumination on a sample with characteristic optical properties and the released radiation can be used to obtain information about the sample properties (Pasquini, 2003).

For cereal grain evaluation, diffuse reflectance is the most suitable mode to distinguish features from the NIR spectra, due to the its ability to partially penetrate into the sample and to show its physical and chemical properties. Otherwise, in specular reflectance the sample works as a mirror, so, as the light beam do not penetrate inside the grain samples, only the information of the surface is detected, and both physical and chemical information is overlapped in the NIR spectra. In cereal grain studies, the worst performances were obtained by transmittance mode, commonly used for liquid and gas sample analysis (Caporaso, Whitworth, & Fisk, 2018). As the light radiation goes through the sample, it is detected in the opposite site of the illumination unit. In pasta studies, the predominant measurement mode is the transflectance, which consists in amplifying the light beam by a mirror in order to make it pass twice through the sample. Additionally, the interactance mode, usually used in meat analysis, gives higher incidences of light on the sample due to the location of the detector distant from the emission position. Consequently, the light beam can penetrate deeply inside the sample and it contains more useful information for heterogeneous products where the surface does not represent the whole sample (Gou et al., 2013). No studies have been published on the interactance application in cereals. Finally, light absorption is also related to food chemical and biological properties and it can be estimated from the logarithm of the inverse measured reflectance.

2.4. Reflectance Calibration

When the hyperspectral image is captured in the reflectance mode, the raw data obtained from the measurements is expressed in absolute reflectance. These values are required to be compensated by removing the dark current noise and correcting the white response (Ngadi &

Liu, 2010). Thus, a reflectance correction is performed on the original measurements obtaining the relative reflectance, according to the equation 1:

$$I = \frac{I_0 - I_b}{I_w - I_b} \quad (1)$$

where I_0 is the raw hyperspectral image obtained, I_w is the white reference and I_b is the dark current reference. In practice, the compensation of dark and white references is performed by covering the lens with a zero reflectance cap and using a white fluoropolymer with the highest known reflectance (99%), respectively (Huang, Liu, & Ngadi, 2014).

3. Hyperspectral data pre-processing

Pre-processing techniques are useful tools to reduce the data variability and to highlight desired spectral characteristics prior to modelling. Frequently used pre-processing techniques in HSI-NIR spectroscopy can be classified in two main groups, such as scatter-correction methods and spectral derivatives. The main objectives of these approaches are to improve subsequent exploratory analysis, to ameliorate linear calibration modes or to improve classification or prediction models (Martens, Nielsen, & Engelsen, 2003). The success of the models depends on the selection of the suitable pre-processing technique in order to reduce or maintain model complexity. On the other hand, too strict pre-processing steps should be avoided because they can hide valuable information from the spectral data.

Scatter correction approaches are designed to solve multiple light scatter or additive effects which produce non-linearities for the calibration model. The extensively used technique for NIR pre-processing is the Multiplicative Scatter Correction (MSC). When samples are solid or emulsions, multiplicative scattering effects occur due to the deviations in the optical path length. This is caused by the influence of the vibrational effect of the neighbour chemical bounds of the sample. Coefficients of regression which describe the dispersion are obtained from the original spectrum to build a scatter corrected spectrum (Wu et al., 2019). The first step involves the estimation of the correction coefficients that are obtained from a least square linear regression (LS) against the average optical spectrum from the calibration samples, by:

$$y_i = a + b \cdot y_{ref} + e \quad (2)$$

Where y_i is the previously measured spectrum from the sample; a is the specular effect of the reflection of the sample; $\frac{1}{b}$ are the estimated scatter interferences in the sample; y_{ref} is a reference spectrum used for the pre-processing of the data set, which usually corresponds to the average spectrum of the calibration set; and *e* is the error which contains the chemical information not explained by physical variations.

The second step is the correction of the recorded spectrum using the coefficients obtained previously, following the equation:

$$y_{corr} = \frac{y_{ref} - a}{b} = y_{ref} + \frac{e}{b} \quad (3)$$

Where y_{corr} is the corrected spectrum similar to y_{ref} in terms of linear regression. By this transformation, the spectra would have a consistent baseline (Rinnan, Berg, & Engelsen, 2009).

The second most applied technique used for scatter correction of NIR is the Standard Normal Variate (SNV) (4). This pre-processing method has similar benefits to MSC, as it is suitable for removing multiplicative and additive interferences of scatter in light distance and particle size (Barnes, Dhanoa, & Lister, 1989; Cen & He, 2007). The mean spectrum and the standard deviations are calculated and each point in the spectrum is recalculated subtracting the mean and dividing by the standard deviation.

$$SNV_i = \frac{(A_i - \bar{A})}{\sqrt{\frac{(A_i - \bar{A})^2}{n-1}}} \quad (4)$$

Where, SNV_i are individual standard normal variations for i wavelengths (corrected value); A_i is reflectance value at i wavelength; $\bar{A}$ is the mean of the A_i reflectance's for all the wavelengths and n is the number of wavelengths from the used range (Caporaso, Whitworth, & Fisk, 2017).

First and second derivatives are widely applied in analytical spectroscopy in data in which noise is a problem. Such transformations remove also both multiplicative and additive effects. The differences between applying first or second derivative is that first derivative is based on the difference between two subsequent spectral points and baseline is removed. Second derivative is based on the successive spectral points of the first derivative so both baseline and derivative trends are removed (Tsai & Philpot, 1998).

Moreover, Savitzky-Golay smoothing and differentiation is a derivative technique based on the computation of a local polynomial regression to obtain a similar but smoothed function. The advantages of this approach are noise reduction, smoothing of spectra and a single-step derivative computation with the application of a filter. The difference with other pre-processing techniques is that the initial distribution, relative maxima and minima and peak width are preserved (Savitzky & Golay, 1964).

Two more pre-processing techniques, such as averaging and normalization, are also used to transform data sets into a suitable matrix for subsequent modelling. Averaging characteristics are based on the reduction of variables or objects in the data set, noise and uncertainty measurements. On the other hand, normalization is based on the vector scaling (to 1.00 or 100) of the sample set to obtain variables with the same size. This pre-processing is useful to

compensate variances in analytical measurements and it is similar to SNV (Esbensen, Guyot, Westad, & Houmoller, 2002).

4. Hyperspectral data analysis: Chemometric tools

4.1. Calibration models

Hyperspectral measurements are commonly characterized by a high complexity, so they should be interpreted to obtain efficient quantitative and qualitative information. Once the data have been pre-processed, multivariate statistical tools are applied to find relationships between the samples and the numerous variables obtained. Figure 3 is a graphical representation of multivariate qualification (projection and classification models) and quantification (regression models) methods that have been used as chemometric tools to deal with *Fusarium* and DON contaminated samples using HSI-NIR (Dale et al., 2012; Kumar, Bansal, Sarma, & Rawal, 2014).

Multivariate qualification methods are applied to reduce large amount of data to a limited number of variables, called principal components (PC). Principal component analysis (PCA) is based on the extraction of the most relevant information from the raw data and a reduction of its dimension by compressing the spectral data into a new group of orthogonal variables. It primarily permits the detection of sample groups, which can be used to classify them (Gatius, Lloveras, Ferran, & Puy, 2004). In order to build the best classification method, the selection of the number of PC should be considered. The optimum number is achieved when introducing one more PC, the performance does not improve, so we are increasing the computation time and the complexity of the model. The use of a cross-validation method should be considered to determine the number of PC to be selected for the best performance of the analysis (Jiang, Zhu, & Tao, 2010). A supervised classification model can be constructed making class models independently, each one described by a PCA. Unknown samples can be assigned to the known classes to classify them. Several studies of fungal contamination in cereals have used this technique to explain their results (Barbedo et al., 2015; Serranti, Cesare, & Bonifazi, 2013; Shahin & Symons, 2011). The structure of the PCA model is characterized by the equation 5:

$$X = t_1 p_1^T + t_2 p_2^T + \cdots + t_A p_A^T + E_A \qquad (5)$$

Where X are the coordinates expressed over the original X variables; t corresponds to the coordinates of the objects over the PCs; p corresponds to the vectors of the new subspace where the original variables X are projected; E corresponds to the noise characterized by the residuals; and A corresponds to the number of PC used.

Polder, Heijden, Pioneer, Waalwijk, & Young, (2005) used Fuzzy c-means clustering as statistical method for a prediction model development. Cluster analysis is a method which

consists in the categorization of the observations into groups (clusters), in which each observation of the group has more similar features than other groups. More specifically, the function used for the characterization of each observation inside the clusters by ranging the observation between 0 and 1 (fuzzy partition) is called Fuzzy C-means clustering. By minimizing iteratively, the function to obtain optimal fuzzy partition, an algorithm is built to discriminate the relationship between the observation and the other groups. Values close to 1 indicate the similarity of the observation with its cluster and values close to 0 indicate less similarity (Bezdek, Ehrlich, & Full, 1984).

An additional extensively used classification method in HSI is discriminant analysis (DA), which purpose is to find recognition patterns that permits the separation of the observations into classes. The rules obtained should also allow the association of new data into one of these groups, so it is categorized as a supervised method (Hubert & Van Driessen, 2004). If a common covariance matrix is obtained from the data estimation, a Linear Discriminant Analysis (LDA) is used. LDA uses a linear combination of features in the multivariate data which are able to divide observations into two or more groups by the maximization of the ratio between variances of the compared observations in relation with the group variance (Esteki, Shahsavari, & Simal-Gandara, 2018). This latest method is more widely applied because of the reduced number of parameters involved in contrast with the unequal covariance matrix structures. Consequently, both procedures have been applied for hyperspectral data analysis in *Fusarium* and DON contamination in cereals (Delwiche & Kim, 2000; Delwiche et al., 2011; Delwiche, Kim, & Dong, 2010; Rinnan et al., 2009; Ropelewska & Zapotoczny, 2018; Serranti et al., 2013; Shahin & Symons, 2012, 2011; Singh et al., 2012; Tekle et al., 2015).

Linear and non-linear regression techniques have been used for the prediction of unknown sample concentrations from the spectral data (Westad, Bevilacqua, & Marini, 2013). Regression methods require a calibration set, in which the coefficients of the relation between concentrations and spectra are calculated, and a validation set, in which these coefficients obtained from the calibration set are checked with a new set of samples in order to calculate the prediction error. (Boldrini et al., 2012). The most commonly used methods in hyperspectral image data analysis are Principal Component Regression (PCR) and Partial Least Squares (PLS) (Caporaso et al., 2018; Viscarra Rossel, 2008) .

MLR is the most basic regression method which can provide successful models for data matrix with few variables X. Thus, for simple components prediction does not present substantial differences with other regression procedures. However, its application in complex systems (more X variables than Y samples) gives ineffectiveness because it is not able to deal with interferences, noise, errors and collinearity between X variables. For that reason, MLR is not

commonly used in HSI-NIR calibration (Balabin, Safieva, & Lomakina, 2007; Fox, Onley-Watson & Osman, 2002).

PCR is a two-step model which involves PCA on the spectral data and a subsequent MLR of the prediction parameters on the scores obtained from the PCA. The regression is performed on the scores with the optimum number of PC, which is much lower than the amount of original X variables, so MLR weaknesses are avoided. The scores are orthogonal, so collinearity problems disappear. In addition, the reduction of the number of variables permits to solve the model with a lower number of equations (Næs & Martens, 2005).

PCR is used when the X information corresponding to the spectral data is desired because it works with the direction of maximum variance of X. However, the HSI-NIR calibration purpose is focused on the Y information about the analyte. PLS releases instrumental information to focus on the Y information to obtain a well fitted model and lower errors of prediction.

PLS is, as PCR, a two-step model, that, instead of projecting on the directions of maximum variance of X, it does on the maximum covariance of X-Y, so it improves the subsequent regression of the dependent variable on these directions. For reaching this cooperation, latent variables are computed to model the covariance structure for matrix and dependent variables (Mehmood, Liland, Snipen, & Sæbø, 2012; Westad et al., 2013). PLS algorithms can be applied to single Y variable (PLS1) or multiple Y variables (PLS2). Both PCR and PLS methods are useful for wavelength selection because it permits the analysis of the original variables, so the significant ones, which introduce much information, can be discerned from the irrelevant ones.

Finally, Support Vector Machines (SVM) is a method that has been applied to kernel classification due to its applications in recognition and detection. Theoretically, SVM purpose is to find an optimal hyperplane surface that divides the maximum of the data points into classes by representing the sample points in the space. The space is divided by this hyperplane, which consists in a vector between two points, of the two classes. When a new sample is introduced in the model, depending on the position in the space previously separated by the hyperplane, will be classified into one of the groups (X. Yang, Hong, You, & Cheng, 2015).

5. Validation and performance of the calibration model

The common procedures for calculating a regression model comprises the acquirement of reference values and spectral data (usually pre-treated) for the calibration set. For validation, a cross-validation procedure or and independent sample set may be introduced, in order to obtain a realistic error of prediction to achieve concentration values as close as possible to the real concentration of the analytes of interest in unknown samples from its spectral data. The previous calibration step involves the reference values and spectra to be considered for the training set, both related with a regression model that has a basic form as the equation 6:

$$Y = b_0 + b_1X_1 + b_2X_2 + \cdots + b_kX_k \quad (6)$$

Where Y is the unknown variable to be measured; X_i are characteristic wavelengths used for the regression model; b_i are the regression coefficients, which estimates the unknown parameters; and b_0 is the offset (Givens, De Boever, & Deaville, 2005). In the calibration step, spectral values are used to calculate the regression coefficients. Subsequently, a set of new spectral data and the previously obtained regression coefficients are used to predict and measure the unknown variables (Y) in the validation set. The information from the reference method and the spectral data used for calibration and validation should be representative of the population that you work with.

To fit adequately the model to the data, the number of PCs to be used in the model should be adjusted to explain efficiently the variability of the calibration and validation sets. The criteria used to select the number of PCs to optimize the model is to detect the PC where a break on the curve of the validation residual variance occurs or a minimum in the prediction error is observed. Other more specific considerations about the analyte should also be reviewed to obtain a feasible method (Viscarra Rossel, 2008).

Calibration fitness can be determined by its statistical performance described by the following parameters: Coefficient of Determination of calibration (R^2_{Cal}), Standard Error of Calibration (SEC), the Root Mean Square Error of Calibration (RMSEC) (Chavez et al., 2013). The closest R^2_{cal} value to 1 and the closest SEC and RMSEC values to 0 correspond to the best calibration model.

In order to achieve the internal test and know which model is the best fitted to the data, a cross-validation procedure is commonly used when it is not possible to have a large number of autonomous samples from the calibration set. When a single sample is left out of the training set for each iteration, a leave-one-out cross-validation is performed. Otherwise, a k-fold cross-validation is carried out when a group of samples are left out. They are separated into different groups (*k*), *k-1* used for the calibration training and one remaining group for the test set. The test set group is changed until all samples have been tested (Ramírez-Morales, Rivero, Fernández-Blanco, & Pazos, 2016). Thus, the number of latent variables and the parameters (such as Root Mean Square Error of Cross-Validation - RMSECV) of the model are evaluated by its internal implementation.

Full Cross-Validation is considered the most realistic estimation when only one sample set is available, although it will always be more optimistic than the prediction error calculated from two sample sets. Two sample sets are not always accessible because it requires large number of samples and, if they are not representative of the population you are working with, the model will not be so realistic. Thus, in order to build an ideal model, large and representative sets of

samples from the calibration and external validation of the model are needed. This will prove the linearity, specificity and accuracy of the model to predict future sample concentrations (Levasseur-Garcia, 2018).

The performance of the model is assessed by statistic parameters which have to be considered for the model selection. Table 1 presents the most important statistic parameters to estimate the statistic performance of the models. The Coefficient of Determination of prediction (R^2_p) estimates the variance between reference and predicted values and the reference values versus the total variance. The Standard Error of Prediction (SEP) determines how precise is the model. However, it should be corrected by the bias, which considers the difference between the expected SEP value and its true value. The Root Mean Square Error of Prediction (RMSEP) determines how much accurate is the calibration and it is closely related with SEP and bias. Finally, the Ratio of Performance to Deviation (RPD) is the ratio between the SEP and the standard deviation of the samples and it is under discussion for its advantages in NIR calibration. Thus, some publications consider RPD a redundant parameter (Bellon-Maurel, Fernandez-Ahumada, Palagos, Roger, & McBratney, 2010). The model to be used should have the R^2_p closer to 1 and the SEPc (corrected by bias) and the RMSEP closer to 0.

6. HSI and *Fusarium* damage detection in cereals

Optical detection of fungal contamination, specially from *Fusarium* species, using hyperspectral technologies has been achieved recently by some authors, not only to obtain spectral features, but also spatial characteristics of numerous cereals (Xing et al., 2019). The research of a precise and accurate model to distinguish and classify kernels as damaged or healthy has motivated researchers to publish many works. In general, it has been concluded that NIR spectra are more suitable than VIS spectra for *Fusarium* damage detection (Polder et al. 2005; Delwiche et al. 2011). This section has focused on the studies based on detection of *Fusarium* contamination with HSI-NIR method, with special interest in wheat samples. Table 2 is a summary of the most relevant features of studies with this purpose.

Many works were concerned about *Fusarium* damage characterization in cereals by HSI using visual inspection as reference method for the classification of *Fusarium* Damaged Kernels (FDK) from colour and textural features (shrivelled, pink discoloration, white chalky, weight loss, etc.). Delwiche & Kim (2000) used a spectral range of 430-860 nm obtaining classifications accuracies of 86.8 and 98.4% in two different wheat types. Moreover, Delwiche et al. (2010) combined two pairs of wavelengths between the spectral range of 1000-1700 nm (1199, 1474 nm) and (1315, 1474 nm) obtaining a LDA classification accuracy of 82.5% in high visual contrast kernels. A year later, Delwiche et al. (2011) repeated the experiment at similar conditions but adding a VIS camera (400-1000 nm) and different wavelength pairs (502, 678 nm) and (1198, 1498 nm) achieving a better LDA correct classification accuracy (95%).

These authors placed the kernels in a template, each kernel crease-down located in a well, and averaged the spectra of the pixels belonging to each kernel. Moreover, they showed that the pixels located in the endosperm gave more useful information than those located in the germ. A LDA was also built by Shahin & Symons, (2011), using a wavelength range of 400-1000 nm, collected images of 800 kernels placed crease-down. The best LDA model was performed on PCA score features and selecting six characteristic wavelengths (484, 567, 684, 817, 900, 950 nm). The classification accuracy was 92.25% for the validation set and it was similar compared to previous works. The study of Singh et al. (2012) used different conditions, as *Penicillium* and *Aspergillus* artificial inoculation instead of *Fusarium* and only an individual wavelength (870 nm) selection (highest loading factor for the first PC). Classification by statistical LDA reached a range from 88.7 to 98.0% of accuracy. The last study reported using a LDA method for *Fusarium*-damage assessment was that of Ropelewska & Zapotoczny (2018) which used three characteristic wavelengths (550, 710 and 850 nm) to classify two wheat varieties. The classification accuracy obtained based on textural parameters of ventral and dorsal sides was 85-98%, so the model presented similar precisions to the previous studies published. Even though the reports used different spectral ranges or specific wavelengths, their results were similar.

Alternatively, PLS-DA models were also used in numerous studies for FDK classification. Williams, Manley, Fox, & Geladi (2010) proved the discrimination power of HSI-NIR of *Fusarium*-infected maize samples by different pre-processing techniques (MSC, SNV and non-processed). Selecting the wavelengths 1960 nm and 2100 nm for the variation of the infected, associated to carbohydrates and protein from fungal presence, and 1450 nm, 2300 nm and 2350 nm for the variance of the non-infected kernels a classification accuracy between 94.0-97.7%, a coefficient of determination of 0.73 and 0.86 for each camera and a RMSECV of 0.23 were reached. The best conditions to classify were applying an MSC step in infected kernels spectra. A slightly poorer accuracy (90.5%) was achieved by Shahin & Symons, (2012) using 4 wavelength bands (494, 578, 639, 678 nm) and mean-normalized spectra. Moreover, the R^2 was also lower (0.62) leading to a RMSECV of 0.31%. Serranti et al. (2013) used a general least square weighting algorithm (GLSW) as pre-processing method and selected 12 effective wavelength bands construct a PLS-DA model based on characteristic features. They obtained good results to correctly classify FDK with 91% (RMSECV 3.8%) and 92% (RMSECV 2.4%) for all wavelengths and 12 effective variables selected, respectively. Tekle et al. (2015) analysed oat kernels by HSI-NIR obtaining reflectance spectrums, which were transformed to absorbances and pre-processed by applying SNV. The characteristic wavelength bands for high infection spectra were 1925, 2070 and 2140 nm, whereas 1400, 1626 and 1850 were related to negative infections, obtained from the PCA loading weights. The average percentages of damaged pixels in each group were 73.3% for severely damaged kernels, 46.9% for mildly

damaged kernels, 29.3% for asymptomatic and 26.5% for uninoculated, so it reflects the spatial ability, in comparison to whole kernel classification, of the HSI system to detect levels of *Fusarium* damage. Finally, the recent study of Delwiche, Rodriguez, Rausch, & Graybosch (2019) displayed percentages of correct classification of cross-validation higher than 92% for sound and *Fusarium*-damaged kernels statistically processed by both LDA and PLS-DA. In this case, only mean-centring was applied except in the case of one assay in which SNV was used. LDA models were performed with different wavelength properties (1000, 1197, 1308, 1394 nm) and PLS-DA with different number of latent variables (2, 4 and 7). An independent validation set of external samples was used to establish the SEP which were ranged between 4.9 and 6.6%, so the results for the percentage of FDK determination were accurate. Finally, Barbedo et al. (2015) proposed an algorithm for image pre-processing (region of interest (ROI) delimitation, kernel and background segmentation and cluster splitting) previous to *Fusarium* Index (FI) calculation. The higher the proportion of pixels with high reflection in the 1411 nm band was, the larger value of FI was obtained. Moreover, for large FI values, FHB presence in the kernel was more likely. According to them, the classification accuracy of the algorithm was approximately 91%. The results obtained are similar to those in the abovementioned studies based on the same purpose, but they are not comparable because they were obtained under different parameters. In summary, all authors tried to identify the most relevant wavelengths in order to build a multispectral equipment, affordable to be implemented in the food industry, however, the identified wavelengths were not the same in the different studies.

The study of Polder et al. (2005) also used HSI-NIR technique for wheat analysis, but they employed PLS regression for *Fusarium* quantitative analysis. The reference values were fungal DNA concentration obtained with Taqman Real-Time PCR. Normalization and second polynomial order SG-smoothing were applied for spectral pre-processing. The regression model permitted a well-defined identification for *Fusarium* DNA in concentrations above 6000 pg and the prediction of concentrations higher than 100 pg with an R^2 of 0.8. Although the correlation between the *Fusarium* DNA and the spectra was acceptable, further investigations are needed to improve fungal and DON concentration predictions.

7. HSI and DON estimation

HSI-NIR is not only applied to improve the detection of features that can be seen by eye so it is also used for detecting components in a sample that cannot be appreciated visually. DON is commonly found in asymptomatic grains and its detection is more complex than other visible traits. Still, the detection of *Fusarium* damaged kernels, can be a rough indirect measurement of DON presence. The application of the HSI-NIR technique to DON detection is a novel approach which would be able to solve the problem of heterogeneity of the toxin through the sample, due to the capacity of spatial examination (Fox & Manley, 2014), but still few studies exist. Recent

studies based on visualization, screening and quantification of DON in cereals have been summarised in Table 3, with the purpose of presenting the last progresses about the topic.

Tekle et al. (2015) aimed to measure DON content from oat average NIR spectra by a PLS regression model. The reference contamination of the samples was obtained with gas chromatography coupled to mass spectrometry (GC-MS) and the images were acquired by HSI-NIR at a wavelength range of 1000-2500 nm. Reflectance spectra were transformed into absorbances and SNV was applied to remove scattering impact. PLS was optimized by cross-validation. For 112 validation set of images and using five PCs for the model optimization, an R_{val}^2 of 0.81 was obtained. An alternative PLS-LDA using the ratio of damaged pixels in the kernel was also performed. The correlation between predicted and measured DON was 0.79, so it was proved that both were valid for DON prediction.

Barbedo et al. (2015) algorithm also aimed the estimation of DON contamination of wheat samples. Full-image DON estimation was performed (25-50 kernels per image), so individual kernel analysis was not feasible. They studied the correlation between FI and DON concentrations and they obtained a strong correlation of 0.84. However, at low DON levels, the correlations with FI seems to be weaker because visual discrimination is more complex. A later study of Barbedo et al. (2017) focused on DON screening of wheat samples and developed a new algorithm. DON reference measurements were performed by direct competitive enzyme-linked immunosorbent assay (DC-ELISA) and liquid chromatography coupled to mass spectrometry (LC-MS). Two wavelengths (627 and 1411 nm) were selected because they seemed to converge with DON presence. The algorithm was based on a DON preliminary index (DPI) calculation, which is calculated from the kernel brightness. It was able to classify the samples into the three classes proposed (<500 µg/kg, 500-1250 µg/kg, and >1250 µg/kg) with an accuracy of 72%. Furthermore, the accuracy increased to 81% when the classes were reduced into two, separated by the UE legal limit (1.25 mg/kg). Liang et al. (2018) also used algorithms for DON detection in bulk wheat samples containing 250 – 5000 µg/kg in the Vis/NIR of 400-1000 nm. First, samples were analysed by (LC-MS) and they were divided into three groups. Images were acquired for each level of contamination (70 wheat kernels each) and the spectra obtained were pre-processed by SNV and MSC, in order to reduce the RMSECV. Optimal wavelengths were inspected for both pre-processing methods for Successive Projection Algorithm (SPA) and Random Frog (RF) to reduce dimensionality. The best combination between pre-processing and the selected algorithm was the MSC-SPA-SVM, obtaining an accuracy of 100% and 97.72% in the training and the testing set, respectively. A visual representation of the DON contamination was also achieved using the same model, which provided information about the levels of the toxin within the sample.

8. HSI as a cereal sorting tool

Generally, mono or dichromatic cameras have been used for fungal inspection purposes. High-power LED pulses have been applied as inspection systems, in which the reflectance of the samples was measured. Although a kernel classification can be reached, spatial characterization of the sample is not possible, so the sample is processed as a whole, not being able to select its regions of interest. Consequently, one measurement analysis for the whole sample or each ROI is needed, so the selection of a the spectrum of an specific area (as a kernel) cannot be achieved (Delwiche, 2008, 2009). Thus, massive and rapid classification systems are needed to achieve discriminations of tonnes of grain in few hours (Fox & Manley, 2014).

The background of optical cereal sorting started in the Single Kernel-NIR (SK-NIR) spectroscopy, in which the automatic classification aptitudes were demonstrated. This automatization was achieved in numerous cases, providing wheat kernels discernment according to DON contamination at a limit of 60 mg/kg in 96% of the cases (Peiris et al., 2010). Moreover, positive and negative fungal infection discrimination in corn kernels was reached with an efficacy of 96% and 98%, respectively (Pearson & Wicklow, 2006). This system automatically feeds single wheat kernels to a spectrometer viewing area, which slows the sorting process.

Other rapid cereal sorting methods have also been tested for the reduction of mould-damage and DON contamination. Rejection of contaminated wheat kernels during freefalling have been evaluated for its rapid discrimination (Delwiche, 2006, 2007). Although the studies reached good results, samples were tested in a laboratory-controlled conditions, so further reports are needed to improve fungi and DON sorting in wheat. However, recent studies also reached good in-line results for AFB1 with the 99% of the peanut samples accepted below 10 μg/kg with a laser-based system (Liu et al., 2019).

HSI calibration is proposed for routine inspection of grain in food industries. Its ability to see spatially all over the kernels and across them in the cases the grain presents fissures or insect damages, have driven this instrument to be promising for real time *Fusarium* damage and mycotoxin assessment. The presence of asymptomatic kernels hinders the visual evaluation, so fungal detection is not as simple as for symptomatic ones. HSI-NIR is capable to evaluate fungal growth and spatially ubicate its presence on a single kernel. Moreover, albeit DON is produced by *Fusarium*, the contamination is not proportional. Asymptomatic samples with high DON concentrations can be fungal-free. For that reason, a precise technique with the ability of chemically inspect and sort contaminated cereals is needed (Tatzer, Wolf, & Panner, 2005).

9. Conclusions

Based on the last findings, the combination of spectroscopy tools and imaging opens a new field in the detection of toxigenic fungi and their metabolites in cereals. HSI has been considered as a rapid, non-destructive and low cost (after bearing the initial cost of the device) method for reaching this aim. The review of the studies showed that this technique is able to classify kernels between fungal-damaged or healthy kernels with high accuracies by simple analysis. However, the visual inspection as a reference method used to build the model should be changed for more robust techniques, in order to obtain wide representations of the sample. Furthermore, the classification of grain in specific DON levels has been reached in some studies, trying to approximate the threshold to legal limits in order to be used in real applications. The advances described in this review are the preliminary models for the future classification systems applied to the industry. Besides, DON quantification results showed important improvements in algorithm construction and optimization of the models. Nevertheless, lower prediction errors and reduced limits of detection are needed, to introduce HSI-NIR as a routine inspection tool. Despite the improvements presented by numerous works, enhancements in sample representativeness, in pre-processing techniques to highlight spectral information and in model performances should be reached to construct routine massive sample analysis.

HIS-NIR implementation may have an impact in food safety management of cereals, at two main levels. First, industrial reception of grain is a critical step in which HSI-NIR applications would be interesting for the substitution of classical analysis methods. Still, the heterogeneity of the batches often presents troubles in the representativeness of the samples to be analysed. Secondly, HSI-NIR would be useful to overcome this heterogeneity by rapid kernel or pixel spectra evaluation (Gruna, Vieth, Michelsburg, & León, 2010). Consequently, industry-applied instruments for on-line sorting of grain would be promising for fungal and mycotoxins contamination management. In short, HSI is an encouraging analysis technique for fungal damage and mycotoxin analysis due to the spatial dimension introduced and its fast scanning capacity.

Conflicts of interest

The authors declare that they have no conflict of interest.

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.

References

Agelet, L. E., & Hurburgh, C. R. (2010). A tutorial on near infrared spectroscopy and its calibration. *Critical Reviews in Analytical Chemistry*, *40*(4), 246–260. https://doi.org/10.1080/10408347.2010.515468

Alisaac, E., Behmann, J., Kuska, M. T., Dehne, H.-W., & Mahlein, A.-K. (2018). Hyperspectral quantification of wheat resistance to Fusarium head blight: comparison of two *Fusarium* species. *European Journal of Plant Pathology*, *152*(4), 869–884. https://doi.org/10.1007/s10658-018-1505-9

Balabin, R. M., Safieva, R. Z., & Lomakina, E. I. (2007). Comparison of linear and nonlinear calibration models based on near infrared (NIR) spectroscopy data for gasoline properties prediction. *Chemometrics and Intelligent Laboratory Systems*, *88*(2), 183–188. https://doi.org/10.1016/j.chemolab.2007.04.006

Barbedo, J. G. A., Guarienti, E. M., & Tibola, C. S. (2018). Detection of sprout damage in wheat kernels using NIR hyperspectral imaging. *Biosystems Engineering*, *175*, 124–132. https://doi.org/10.1016/j.biosystemseng.2018.09.012

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. https://doi.org/10.1016/j.biosystemseng.2015.01.003

Barbedo, Jayme Garcia Arnal, Tibola, C. S., & Lima, M. I. P. (2017). Deoxynivalenol screening in wheat kernels using hyperspectral imaging. *Biosystems Engineering*, *155*, 24–32. https://doi.org/10.1016/j.biosystemseng.2016.12.004

Barnes, R. J., Dhanoa, M. S., & Lister, S. J. (1989). Standard normal variate transformation and de-trending of near-infrared diffuse reflectance spectra. *Applied Spectroscopy*, *43*(5), 772–777. https://doi.org/10.1366/0003702894202201

Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J. M., & McBratney, A. (2010). Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *TrAC - Trends in Analytical ChemistryTrends in Analytical Chemistry*, *29*(9), 1073–1081. https://doi.org/10.1016/j.trac.2010.05.006

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, *10*(2–3), 191–203. https://doi.org/10.1016/0098-3004(84)90020-7

Boldrini, B., Kessler, W., Rebnera, K., & Kessler, R. W. (2012). Hyperspectral imaging: A review of best practice, performance and pitfalls for in-line and on-line applications. *Journal of Near Infrared Spectroscopy*, *20*(5), 483–508. https://doi.org/10.1255/1003

Caporaso, N., Whitworth, M. B., & Fisk, I. D. (2017). Application of calibrations to hyperspectral images of food grains: example for wheat falling number. *Journal of Spectral Imaging*, *6*, 1–15. https://doi.org/10.1255/jsi.2017.a4

Caporaso, N., Whitworth, M. B., & Fisk, I. D. (2018). Near-Infrared spectroscopy and hyperspectral imaging for non-destructive quality assessment of cereal grains. *Applied Spectroscopy Reviews*, *53*(8), 667–687. https://doi.org/10.1080/05704928.2018.1425214

Cen, H., & He, Y. (2007). Theory and application of near infrared reflectance spectroscopy in determination of food quality. *Trends in Food Science and Technology*, *18*(2), 72–83. https://doi.org/10.1016/j.tifs.2006.09.003

Chavez, P.-F., De Bleye, C., Sacré, P.-Y., Rozet, E., Hubert, P., & Ziemons, E. (2013). Validation methodologies of

near infrared spectroscopy methods in pharmaceutical applications. *European Pharmaceutical Review*, *18*(1), 3–6.

Dale, L. M., Thewis, A., Boudry, C., Rotar, I., Dardenne, P., Baeten, V., & Pierna, J. A. F. (2013). Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: A review. *Applied Spectroscopy Reviews*, *48*(2), 142–159. https://doi.org/10.1080/05704928.2012.705800

Dale, L. M., Thewis, A., Rotar, I., Pierna, J. A. F., Boudry, C., Vidican, R. M., & Baeten, V. (2012). Chemometric Tools for NIRS and NIR Hyperspectral Imaging. *Bulletin UASVM Agriculture*, *69*(1), 70–76. https://doi.org/10.15835/buasvmcn-agr:8658

de Girolamo, A., Lippolis, V., Nordkvist, E., & Visconti, A. (2009). Rapid and non-invasive analysis of deoxynivalenol in durum and common wheat by Fourier-Transform Near Infrared (FT-NIR) spectroscopy. *Food Additives and Contaminants - Part A Chemistry, Analysis, Control, Exposure and Risk Assessment*, *26*(6), 907–917. https://doi.org/10.1080/02652030902788946

Delwiche, S. R. (2006). Improvements in optical sorting of mold-damaged wheat. In Y.-R. Chen, G. E. Meyer, & S.-I. Tu (Eds.), *Optics for Natural Resources, Agriculture, and Foods* (Vol. 6381, pp. 1–8). https://doi.org/10.1117/12.686027

Delwiche, S. R. (2007). Optical characterization of free-falling mold-damaged wheat kernels. *Optics for Natural Resources, Agriculture, and Foods II*, *6761*, 67610G. https://doi.org/10.1117/12.735662

Delwiche, S. R. (2008). High-speed bichromatic inspection of wheat kernels for mold and color class using high-power pulsed LEDs. *Sensing and Instrumentation for Food Quality and Safety*, *2*(2), 103–110. https://doi.org/10.1007/s11694-008-9037-1

Delwiche, S. R., & Kim, M. S. (2000). Hyperspectral imaging for detection of scab in wheat. *Biological Quality and Precision Agriculture II*, *4203*, 13–20. https://doi.org/10.1117/12.411752

Delwiche, S. R., Kim, M. S., & Dong, Y. (2010). Damage and quality assessment in wheat by NIR hyperspectral imaging. In M. S. Kim, S.-I. Tu, & K. Chao (Eds.), *Sensing for Agriculture and Food Quality and Safety II* (Vol. 7676, pp. 1–8). https://doi.org/10.1117/12.851150

Delwiche, S. R., Kim, M. S., & Dong, Y. (2011). *Fusarium* damage assessment in wheat kernels by Vis/NIR hyperspectral imaging. *Sensing and Instrumentation for Food Quality and Safety*, *5*(2), 63–71. https://doi.org/10.1007/s11694-011-9112-x

Delwiche, S. R., Qin, J., Chao, K., Chan, D., Cho, B., & Kim, M. (2017). Line-scan hyperspectral imaging techniques for food safety and quality applications. *Applied Sciences*, *7*(2), 125. https://doi.org/10.3390/app7020125

Delwiche, S. R., Rodriguez, I. T., Rausch, S. R., & Graybosch, R. A. (2019). Estimating percentages of *Fusarium*-damaged kernels in hard wheat by near-infrared hyperspectral imaging. *Journal of Cereal Science*, *87*, 18–24. https://doi.org/10.1016/j.jcs.2019.02.008

Dowell, F. E., Ram, M. S., & Seitz, L. M. (1999). Predicting scab, vomitoxin, and ergosterol in single wheat kernels using near-infrared spectroscopy. *Cereal Chemistry*, *76*(4), 573–576. https://doi.org/10.1094/CCHEM.1999.76.4.573

EC. (2006). 1881/2006. Setting maximum levels for certain contaminants in foodstuffs. *Official Journal of the European Communities*, *364*(1881), 5–24.

Elmasry, G., Kamruzzaman, M., Sun, D., & Allen, P. (2012). Principles and applications of hyperspectral imaging in quality evaluation of agro-food products: a review. *Critical Reviews in Food Science and Nutrition*, *52*(11),

999–1023. https://doi.org/10.1080/10408398.2010.543495

ElMasry, G., & Sun, D. W. (2010). Principles of hyperspectral imaging technology. In *Hyperspectral Imaging for Food Quality Analysis and Control* (pp. 3–43). Elsevier. https://doi.org/10.1016/B978-0-12-374753-2.10001-2

Eriksen, G. S., & Pettersson, H. (2004). Toxicological evaluation of trichothecenes in animal feed. *Animal Feed Science and Technology*, *114*(1–4), 205–239. https://doi.org/10.1016/j.anifeedsci.2003.08.008

Esbensen, K. H., Guyot, D., Westad, F., & Houmoller, L. P. (2002). *Multivariate data analysis-in practice: an introduction to multivariate data analysis and experimental design*. CAMO. Retrieved from https://books.google.es/books?id=Qsn6yjRXOaMC

Esteki, M., Shahsavari, Z., & Simal-Gandara, J. (2018). Use of spectroscopic methods in combination with linear discriminant analysis for authentication of food products. *Food Control*, *91*, 100–112. https://doi.org/10.1016/j.foodcont.2018.03.031

European Food Safety Authority. (2007). Opinion of the scientific panel on contaminants in the food chain on a request from the commission related to deoxynivalenol (DON) as undesirable substance in animal feed (Question N° EFSA-Q-2003-036). *EFSA Journal*, *73*, 1–42. https://doi.org/10.2903/j.efsa.2005.235

Fox, G., & Manley, M. (2014). Applications of single kernel conventional and hyperspectral imaging near infrared spectroscopy in cereals. *Journal of the Science of Food and Agriculture*, *94*(2), 174–179. https://doi.org/10.1002/jsfa.6367

Fox, G. P., Onley-Watson, K., & Osman, A. (2002). Multiple linear regression calibrations for barley and malt protein based on the spectra of hordein. *Journal of the Institute of Brewing*, *108*(2), 155–159. https://doi.org/10.1002/j.2050-0416.2002.tb00534.x

Gatius, F., Lloveras, J., Ferran, J., & Puy, J. (2004). Prediction of crude protein and classification of the growth stage of wheat plant samples from NIR spectra. *Journal of Agricultural Science*, *142*(5), 517–524. https://doi.org/10.1017/S0021859604004691

Givens, D. I., De Boever, J. L., & Deaville, E. R. (2005). The principles, practices and some future applications of near infrared spectroscopy for predicting the nutritive value of foods for animals and humans. *Nutrition Research Reviews*, *10*(01), 83. https://doi.org/10.1079/nrr19970006

Gou, P., Santos-Garcés, E., Høy, M., Wold, J. P., Liland, K. H., & Fulladosa, E. (2013). Feasibility of NIR interactance hyperspectral imaging for on-line measurement of crude composition in vacuum packed dry-cured ham slices. *Meat Science*, *95*(2), 250–255. https://doi.org/10.1016/j.meatsci.2013.05.013

Gruna, R., Vieth, K.-U., Michelsburg, M., & León, F. P. (2010). Hyperspectral imaging - from laboratory to in-line food sorting. In *Proceedings of the 2nd International Workshop on Image Analysis in Agriculture (CIGR'10), Budapest, Hungary, August 26-27, 2010; 1 CD-Rom* (pp. 79–90). BCE, Budapest.

Gupta, N. (2011). Development of staring hyperspectral imagers. *Proceedings - Applied Imagery Pattern Recognition Workshop*, (i), 1–8. https://doi.org/10.1109/AIPR.2011.6176379

Huang, H., Liu, L., & Ngadi, M. O. (2014). Recent developments in hyperspectral imaging for assessment of food quality and safety. *Sensors*, *14*(4), 7248–7276. https://doi.org/10.3390/s140407248

Hubert, M., & Van Driessen, K. (2004). Fast and robust discriminant analysis. *Computational Statistics and Data Analysis*, *45*(2), 301–320. https://doi.org/10.1016/S0167-9473(02)00299-2

International Agency for Research on Cancer (IARC). (2012). Mycotoxins and human health. *IARC Scientific*

Publications, (158), 87–104. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed10b&NEWS=N&AN=23477198

Jiang, L., Zhu, B., & Tao, Y. (2010). Hyperspectral image classification methods. In *Hyperspectral Imaging for Food Quality Analysis and Control*, pp. 79–98. Elsevier. https://doi.org/10.1016/B978-0-12-374753-2.10003-6

Kokkonen, M., Ojala, L., Parikka, P., & Jestoi, M. (2010). Mycotoxin production of selected *Fusarium* species at different culture conditions. *International Journal of Food Microbiology*, *143*(1–2), 17–25. https://doi.org/10.1016/j.ijfoodmicro.2010.07.015

Krnjaja, V., Levic, J., Stankovic, S., & Stepanic, A. (2011). *Fusarium* species and their mycotoxins in wheat grain. *Zbornik Matice Srpske Za Prirodne Nauke*, (120), 41–48. https://doi.org/10.2298/ZMSPN1120041K

Kumar, N., Bansal, A., Sarma, G. S., & Rawal, R. K. (2014). Chemometrics tools used in analytical chemistry: an overview. *Talanta*, *123*, 186–199. https://doi.org/10.1016/j.talanta.2014.02.003

Levasseur-Garcia, C. (2018). Updated overview of infrared spectroscopy methods for detecting mycotoxins on cereals (corn, wheat, and barley). *Toxins*, *10*(1). https://doi.org/10.3390/toxins10010038

Li, H., Wu, G., Fang, Q., Zhang, M., Hui, X., Sheng, B., … Jia, W. (2018). Fibroblast growth factor 21 increases insulin sensitivity through specific expansion of subcutaneous fat. *Nature Communications*, *9*(1), 272. https://doi.org/10.1038/s41467-017-02677-9

Liang, K., Liu, Q. X., Xu, J. H., Wang, Y. Q., Okinda, C. S., & Shena, M. X. (2018). Determination and visualization of different levels of deoxynivalenol in bulk wheat kernels by hyperspectral imaging. *Journal of Applied Spectroscopy*, *85*(5), 953–961. https://doi.org/10.1007/s10812-018-0745-y

Liu, J. J., Cai, Z., Liao, Y., Zhao, L., Moulin, J., & Hartmann, C. (2019). Validation of a laser based in-line aflatoxin sorting technology in Spanish type raw peanut in factory-scale production conditions. *Journal of Food Safety*, *39*(2), 1–11. https://doi.org/10.1111/jfs.12611

Magan, N., Hope, R., Cairns, V., & Aldred, D. (2003). Post-harvest fungal ecology: impact of fungal growth and mycotoxin accumulation in stored grain. *European Journal of Plant Pathology*, *109*(7), 723–730. https://doi.org/10.1023/A:1026082425177

Manley, M. (2014). Near-infrared spectroscopy and hyperspectral imaging: Non-destructive analysis of biological materials. *Chemical Society Reviews*, *43*(24), 8200–8214. https://doi.org/10.1039/c4cs00062e

Marin, S., Ramos, A. J., Cano-Sancho, G., & Sanchis, V. (2013). Mycotoxins: Occurrence, toxicology, and exposure assessment. *Food and Chemical Toxicology*. Elsevier Ltd. https://doi.org/10.1016/j.fct.2013.07.047

Martens, H., Nielsen, J. P., & Engelsen, S. B. (2003). Light scattering and light absorbance separated by extended multiplicative signal correction. Application to near-infrared transmission analysis of powder mixtures. *Analytical Chemistry*, *75*(3), 394–404. https://doi.org/10.1021/ac020194w

Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods in Partial Least Squares Regression. *Chemometrics and Intelligent Laboratory Systems*, *118*, 62–69. https://doi.org/10.1016/j.chemolab.2012.07.010

Næs, T., & Martens, H. (2005). Principal component regression in NIR analysis: viewpoints, background details and selection of components. *Journal of Chemometrics*, *2*(2), 155–167. https://doi.org/10.1002/cem.1180020207

Ngadi, M. O., & Liu, L. (2010). Hyperspectral Image Processing Techniques. In *Hyperspectral Imaging for Food Quality Analysis and Control* (First Edit, pp. 99–127). Elsevier Inc. https://doi.org/10.1016/B978-0-12-

374753-2.10004-8

Pasquini, C. (2003). Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, *14*(2), 198–219. https://doi.org/10.1590/S0103-50532003000200006

Pearson, T. C., & Wicklow, D. T. (2006). Detection of corn kernels infected by fungi. *Transactions of the ASABE*, *49*(4), 1235–1245. https://doi.org/10.13031/2013.21723

Peiris, K. H. S., Pumphrey, M. O., Dong, Y., Maghirang, E. B., Berzonsky, W., & Dowell, F. E. (2010). Near-infrared spectroscopic method for identification of Fusarium head blight damage and prediction of deoxynivalenol in single wheat kernels. *Cereal Chemistry*, *87*(6), 511–517. https://doi.org/10.1094/CCHEM-01-10-0006

Pestka, J. J., & Smolinski, A. T. (2005). Deoxynivalenol: Toxicology and potential effects on humans. *Journal of Toxicology and Environmental Health - Part B: Critical Reviews*, *8*(1), 39–69. https://doi.org/10.1080/10937400590889458

Placinta, C. ., D'Mello, J. P. ., & Macdonald, A. M. . (1999). A review of worldwide contamination of cereal grains and animal feed with *Fusarium* mycotoxins. *Animal Feed Science and Technology*, *78*(1–2), 21–37. https://doi.org/10.1016/S0377-8401(98)00278-8

Polder, G., Van Der Heijden, G. W. A. M., Waalwijk, C., & Young, I. T. (2005). Detection of *Fusarium* in single wheat kernels using spectral imaging. *Seed Science and Technology*, *33*(3), 655–668. https://doi.org/10.15258/sst.2005.33.3.13

Qiao, T., Ren, J., Craigie, C., Zabalza, J., Maltin, C., & Marshall, S. (2014). Comparison between Near Infrared Spectroscopy and hyperspectral imaging in predicting beef eating quality.

Qin, J. (2010). Hyperspectral Imaging Instruments. In *Hyperspectral Imaging for Food Quality Analysis and Control* (First Edit, pp. 129–172). Elsevier. https://doi.org/10.1016/B978-0-12-374753-2.10005-X

Ramírez-Morales, I., Rivero, D., Fernández-Blanco, E., & Pazos, A. (2016). Optimization of NIR calibration models for multiple processes in the sugar industry. *Chemometrics and Intelligent Laboratory Systems*, *159*, 45–57. https://doi.org/10.1016/j.chemolab.2016.10.003

Reddy, K. R. N., Salleh, B., Saad, B., Abbas, H. K., Abel, C. A., & Shier, W. T. (2010). An overview of mycotoxin contamination in foods and its implications for human health. *Toxin Reviews*, *29*(1), 3–26. https://doi.org/10.3109/15569541003598553

Rinnan, Å., Berg, F. van den, & Engelsen, S. B. (2009). Review of the most common pre-processing techniques for near-infrared spectra. *TrAC - Trends in Analytical Chemistry*, *28*(10), 1201–1222. https://doi.org/10.1016/j.trac.2009.07.007

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

Savitzky, A., & Golay, M. J. E. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*(8), 1627–1639. https://doi.org/10.1021/ac60214a047

Sendin, K., Williams, P. J., & Manley, M. (2018). Near infrared hyperspectral imaging in quality and safety evaluation of cereals. *Critical Reviews in Food Science and Nutrition*, *58*(4), 575–590. https://doi.org/10.1080/10408398.2016.1205548

Serranti, S., Cesare, D., & Bonifazi, G. (2013). The development of a hyperspectral imaging method for the detection of *Fusarium*-damaged, yellow berry and vitreous Italian durum wheat kernels. *Biosystems Engineering*, *115*(1), 20–30. https://doi.org/10.1016/j.biosystemseng.2013.01.011

Shahin, M. A., & Symons, S. J. (2011). Detection of *Fusarium* damaged kernels in Canada Western Red Spring wheat using visible/near-infrared hyperspectral imaging and principal component analysis. *Computers and Electronics in Agriculture*, *75*(1), 107–112. https://doi.org/10.1016/j.compag.2010.10.004

Shahin, M. A., & Symons, S. J. (2012). Detection of *Fusarium* damage in Canadian wheat using visible/near-infrared hyperspectral imaging. *Journal of Food Measurement & Characterization*, *6*(1–4), 3–11. https://doi.org/10.1007/s11694-012-9126-z

Singh, C. B., Jayas, D. S., Paliwal, J., & White, N. D. G. (2012). Fungal damage detection in wheat using short-wave near-infrared hyperspectral and digital colour imaging. *International Journal of Food Properties*, *15*(1), 11–24. https://doi.org/10.1080/10942911003687223

Sobrova, P., Adam, V., Vasatkova, A., Beklova, M., Zeman, L., & Kizek, R. (2010). Deoxynivalenol and its toxicity. *Interdisciplinary Toxicology*, *3*(3), 94–99. https://doi.org/10.2478/v10102-010-0019-x

Sudakin, D. L. (2003). Trichothecenes in the environment: Relevance to human health. *Toxicology Letters*, *143*(2), 97–107. https://doi.org/10.1016/S0378-4274(03)00116-4

Tatzer, P., Wolf, M., & Panner, T. (2005). Industrial application for inline material sorting using hyperspectral imaging in the NIR range. *Real-Time Imaging*, *11*(2), 99–107. https://doi.org/10.1016/j.rti.2005.04.003

Tekle, S., Mage, I., Segtnan, V. H., & Bjornstad, A. (2015). Near-infrared hyperspectral imaging of *Fusarium*-damaged oats (Avena sativa L.). *Cereal Chemistry*, *92*(1), 73–80. https://doi.org/10.1094/CCHEM-04-14-0074-R

Tsai, F., & Philpot, W. (1998). Derivative analysis of hyperspectral data. *Remote Sensing of Environment*, *66*(1), 41–51. https://doi.org/10.1016/S0034-4257(98)00032-7

Vidal, A., Sanchis, V., Ramos, A. J., & Mar??n, S. (2016). The fate of deoxynivalenol through wheat processing to food products. *Current Opinion in Food Science*, *11*, 34–39. https://doi.org/10.1016/j.cofs.2016.09.001

Viscarra Rossel, R. A. (2008). ParLeS: Software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems*, *90*(1), 72–83. https://doi.org/10.1016/j.chemolab.2007.06.006

Wegulo, S. N. (2012). Factors influencing deoxynivalenol accumulation in small grain cereals. *Toxins*, *4*(11), 1157–1180. https://doi.org/10.3390/toxins4111157

Westad, F., Bevilacqua, M., & Marini, F. (2013). Regression. In *Diagnostic Histopathology* (Vol. 22, pp. 127–170). https://doi.org/10.1016/B978-0-444-59528-7.00004-1

Williams, P., Manley, M., Fox, G., & Geladi, P. (2010). Indirect Detection of *Fusarium verticillioides* in Maize (*Zea mays L.*) Kernels by near Infrared Hyperspectral Imaging. *Journal of Near Infrared Spectroscopy*, *18*(1), 49–58. https://doi.org/10.1255/jnirs.858

Wu, Y., Peng, S., Xie, Q., Han, Q., Zhang, G., & Sun, H. (2019). An improved weighted multiplicative scatter correction algorithm with the use of variable selection: application to near-infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, *185*, 114–121. https://doi.org/10.1016/j.chemolab.2019.01.005

Xing, F., Yao, H., Liu, Y., Dai, X., Brown, R. L., & Bhatnagar, D. (2019). Recent developments and applications of hyperspectral imaging for rapid detection of mycotoxins and mycotoxigenic fungi in food products. *Critical*

Reviews in Food Science and Nutrition, *59*(1), 173–180. https://doi.org/10.1080/10408398.2017.1363709

Yang, I.-C., Delwiche, S. R., Chen, S., & Lo, Y.-M. M. (2009). Enhancement of Fusarium head blight detection in free-falling wheat kernels using a bichromatic pulsed LED design. *Optical Engineering*, *48*(2), 023602. https://doi.org/10.1117/1.3081092

Yang, X., Hong, H., You, Z., & Cheng, F. (2015). Spectral and image integrated analysis of hyperspectral data for waxy corn seed variety classification. *Sensors (Switzerland)*, *15*(7), 15578–15594. https://doi.org/10.3390/s150715578

Zain, M. E. (2011). Impact of mycotoxins on humans and animals. *Journal of Saudi Chemical Society*, *15*(2), 129–144. https://doi.org/10.1016/j.jscs.2010.06.006

Figure captions

Figure 1. Hyperspectral image diagram for a wheat sample and the relationship between spatial (x, y) and spectral axis (λ).

Figure 2. Push-broom HSI equipment.

Figure 3. Diagram of the most common used multivariate data analysis systems in HSI-NIR. Principal component analysis (PCA); component analysis (CA); DA (discriminant analysis); linear discriminant analysis (LDA); partial least squares (PLS) regression; principal component regression (PCR); multiple linear regression (MLR); supported vector machines (SVM).

Table 1. Performance statistic parameters of the validation set (adapted from Agelet & Hurburgh, 2010; Levasseur-Garcia, 2018).

Validation set parameters		
$\mathbf{R^2_p}$	Coefficient of determination of prediction	$R_p^2 = \frac{\Sigma\,(\hat{y}_i - \bar{y})^2}{\Sigma\,(y_i - \bar{y})^2}$
$\mathbf{d_i}$	Residual	$d_i = \hat{y}_i - y_i$
Bias	Bias	$Bias = \frac{\Sigma\, d_i}{n}$
SEPc	Standard Error of Prediction (corrected by the bias)	$SEPc = \sqrt{\frac{(d_i - bias)}{n-1}}$
RMSEP	Root Mean Square Error of Prediction	$RMSEP = \sqrt{\frac{\Sigma\,(d_i)^2}{n}}$
RPD	Ratio of Performance to Deviation	$RPD = \frac{Sdev_{ref}}{SEPc}$

$\hat{y}_i$ = ith validation sample predicted value, y_i = ith validation measured values, $\bar{y}$ = mean of the n values (measured or predicted), n = number of samples; $Sdev_{ref}$ = Standard deviation of reference.

Table 2. Hyperspectral imaging studies for *Fusarium* management in cereals

Fungi or Disease	Crop	Number of samples	Wavelength range	Reference of contamination	Model	Performance and characteristic wavelength	Reference
Fusarium head blight	Wheat single kernels	32 scabby kernels, 32 control kernels	HSI-Vis/NIR 430-860 nm	Visual inspection	LDA	Increasing 5 bands from 458-844 nm Cross-validation classification accuracy: 83-98%	(Delwiche & Kim, 2000)
***Fusarium culmorum* infection**	Wheat single kernels	96 kernels 2 images (Vis and NIR)/kernel	HSI-Vis 430-900 nm HIS-NIR 900-1750 nm	Visual inspection and TaqMan RT-PCR	PLS, Fuzzy c-means clustering	Identification of >6000 pg *Fusarium* DNA Prediction of >100 pg DNA with R^2 of 0.8	(Polder et al., 2005)
Fusarium verticillioides	Maize single kernels	15 kernels (5 infected, 5 asymptomatic, 5 control)	HSI-NIR 960–1662 nm and 1000–2498 nm (2 camera)	Fungal artificial inoculation	PCA, PLS-DA	Positive infection 1960, 2100 nm Negative infection: 1450, 2300, 2350 nm Camera 1 R^2 = 0.73 Camera 2 R^2 = 0.86 RMSEP = 0.23	(Williams et al., 2010)
Fusarium head blight	Wheat single kernels	8 samples of 2 g of intact seed (~60 kernels each sample)	HSI-NIR1000-1700 nm	Visual inspection	LDA	1001.7, 1126.9, 1199.2, 1314.8, 1473.8 nm Accuracy ND	(Delwiche et al., 2010)
Fusarium head blight	Wheat single kernels	8 samples of 2 g of intact seed (~60 kernels each sample)	HSI-Vis/NIR 400-1000 nm and 1000-1700 nm	Visual inspection	LDA	502, 678, 1198, 1496 nm Classification accuracy: 95%	(Delwiche et al., 2011)
***Fusarium*-damaged kernels**	Wheat single kernels	Calibration set: 200 SND Validation set: 200 FDK each 24-36 kernels/image	HSI-Vis/NIR 400-1000 nm	Visual inspection	PCA, LDA	484, 567, 684, 817, 900, 950 nm Classification accuracy: 92% Prediction accuracy: 86%	(Shahin & Symons, 2011)
Penicillium* spp., *Aspergillus glaucus, A. niger	Wheat single kernels	Calibration set: 240 kernels Validation set: 60 kernels 5 kernels/image	HSI-NIR 700-1100 nm	Fungal artificial inoculation	LDA, QDA, Mahalanobis	870 nm FD classification accuracy: 94-98.3% False positives: 6.3-11.3%	(Singh et al., 2012)
***Fusarium*-damaged kernels**	Wheat single kernels	Calibration set: 1073 SND and 1538 FDK Validation set: 1074 SND and 1536 FDK Prediction set: 799 kernels 24-36 kernels/image	HSI-Vis/NIR 400-1000 nm	Visual inspection	PLS-DA	494, 578, 639, 678 nm R_C^2 = 0.60; R_V^2 = 0.62 RMSEC = 0.31; RMSEV = 0.31 Classification accuracy: 90% False positives: 9%	(Shahin & Symons, 2012)
***Fusarium*-damaged kernels**	Wheat kernels	Calibration set: 5 g of bulk kernels Validation set: 20-30 single kernels	HSI-NIR 1000-1700 nm	Visual inspection	PCA, PLS-DA, iPLS-DA	1209-1230, 1489-1510, 1601-1622 nm PLS-DA classification accuracy: 96% iPLS-DA: 94%	(Serranti et al., 2013)
Fusarium head blight	Wheat single kernels	25-50 kernels/image 27 hyperspectral images 830 total kernels	HSI-Vis/NIR 528-1785 nm	Visual inspection	*Fusarium* index, PCA	ROI: 647 nm; PCA: 1411 nm Classification accuracy: 91%	(Barbedo et al., 2015)
***Fusarium*-damaged kernels**	Oat single kernels	Calibration set: 30 kernels Validation set: 14 kernels (4 categories each)	HSI-NIR 1000-2500 nm	Fungal artificial inoculation and SEM	PLS-LDA	Positive infection: 1925, 2070, 2140 nm Negative infection: 1400, 1626, 1850 nm Accuracy ND	(Tekle et al., 2015)
***Fusarium graminearum* infection**	Wheat single kernels	1 kernel/image Ventral/dorsal scan 240 hyperspectral images	HSI-Vis/NIR 600-1100 nm	Visual inspection	LDA, K Star, PART, LMT	550, 710, 850 nm Classification accuracy: Ventral 78-100%; dorsal: 78-98%; both: 76-98%	(Ropelewska & Zapotoczny, 2018)
Fusarium head blight	Wheat single kernels	5 calibration samples 82 validation samples	HSI-NIR 938-1654 nm	Visual inspection	PLS-DA, LDA	1000, 1197, 1394, 1308 nm (LDA) LDA R = 0.772 and 0.811 PLS-DA classification accuracy: >92% PLS-DA/LDA SEP: 4.9-6.6%	(Delwiche et al., 2019)

FDK = *Fusarium* damaged kernel, iPLS-DA = interval partial least squares – discriminant analysis, LDA = Linear discriminant analysis, LMT = decision tree classifier, ND = not defined, PART = rules classifier, PCA = Principal component analysis, QDA = quadratic discriminant analysis, SEM = scanning electron microscope, SND = sound.

Table 3. Hyperspectral imaging studies for DON management in cereals.

Crop	Number of samples	Wavelength range	Reference of contamination	Spectral pre-processing and characteristic wavelength	Model	Performance	Reference
Oat single kernels	Calibration set: 248 kernels; 31 kernels/category; 4 dorsal, 4 ventral Validation set: 112 kernels; 14 kernels/category; 4 dorsal, 4 ventral	HSI-NIR 1000-2500 nm	GC-MS	Positive infection: 1925, 2070, 2140 nm Negative infection: 1400, 1626, 1850 nm SNV	PLSR, PLS-LDA	Calibration R^2 = 0.75 Cross-validation R^2 = 0.71 Correlation PLSR = 0.81 Correlation PLS-LDA = 0.79	(Tekle et al., 2015)
Bulk wheat kernels	25-50 kernels/image 6 hyperspectral images	HSI-Vis/NIR 528-1785 nm	LC-MS	ROI: 647 nm; PCA: 1411 nm	*Fusarium* index	Correlation FI/DON = 0.84	(Barbedo et al., 2015)
Bulk wheat kernels	30-50 kernels/image 3 levels of contamination 251 total images Calibration set: 33 images Validation set: 218 images	HSI-Vis/NIR 528-1785 nm	DC-ELISA, LC-MS	Wavelength observation 623, 1411 nm	Confusion matrix, k values	Classification accuracy: 72% in three classes, 81% in two classes (EU limit)	(Barbedo et al., 2017)
Bulk wheat kernels	70 kernels/image 60 images/level of contamination Calibration set: 44 images Validation set: 16 images	HSI-Vis/NIR 400-1000 nm	LC-MS/MS	SNV, MSC 14, 12, 7, 14 wavelengths for each pre-processing method	SVM, PLS-DA	Classification accuracy: 100% in training set; 97.92% in testing set	(Liang et al., 2018)

FI = *Fusarium* index, PLS-DA = partial least squares regression – discriminant analysis, PLSR = partial least squares regression, SVM = support vector machine.

Figure 1

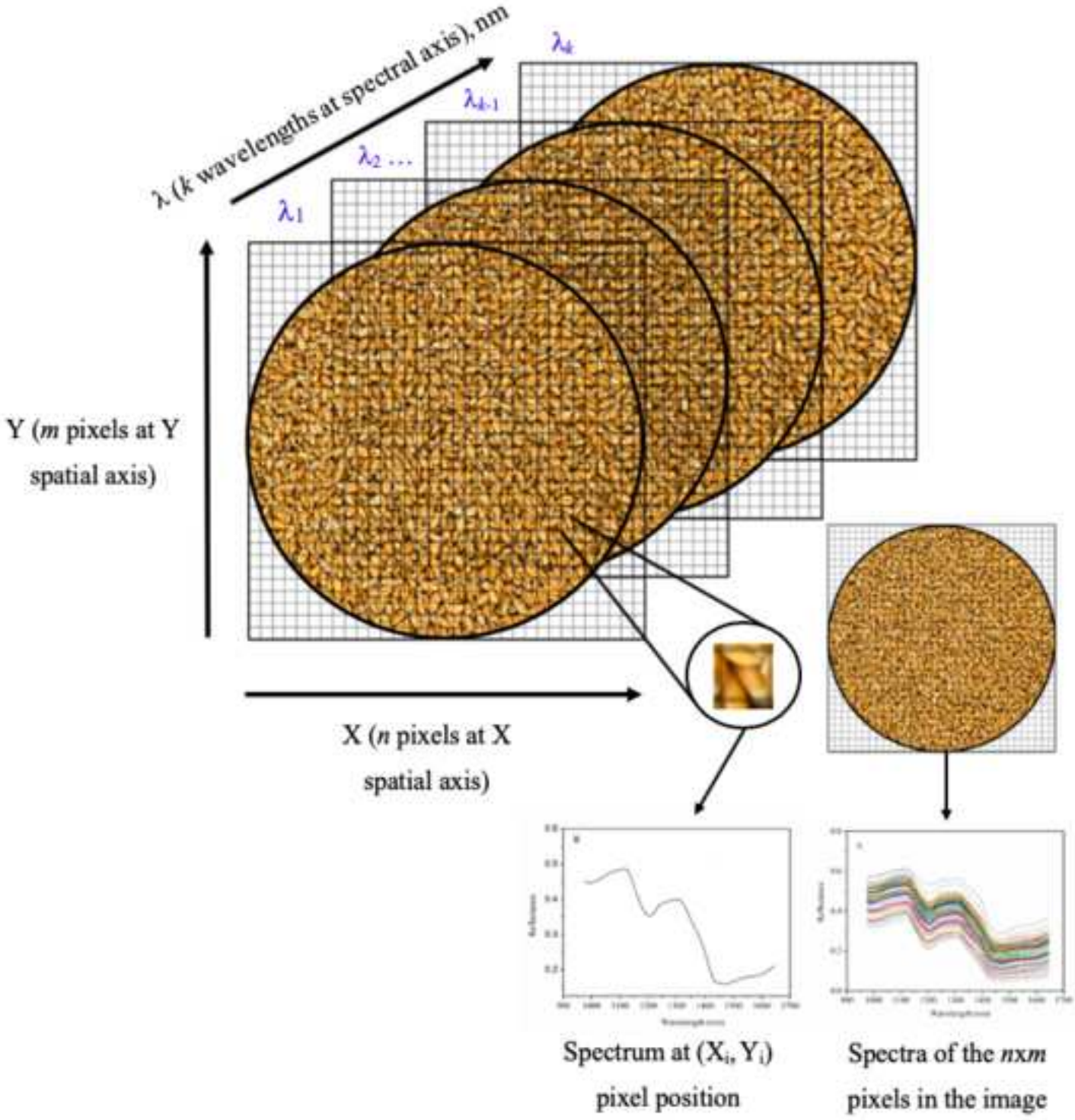

λ (k wavelengths at spectral axis), nm
λ1
λ2 …
λk-1
λk
Y (m pixels at Y spatial axis)
X (n pixels at X spatial axis)
Spectrum at (Xi, Yi) pixel position
Spectra of the nxm pixels in the image

Figure 2

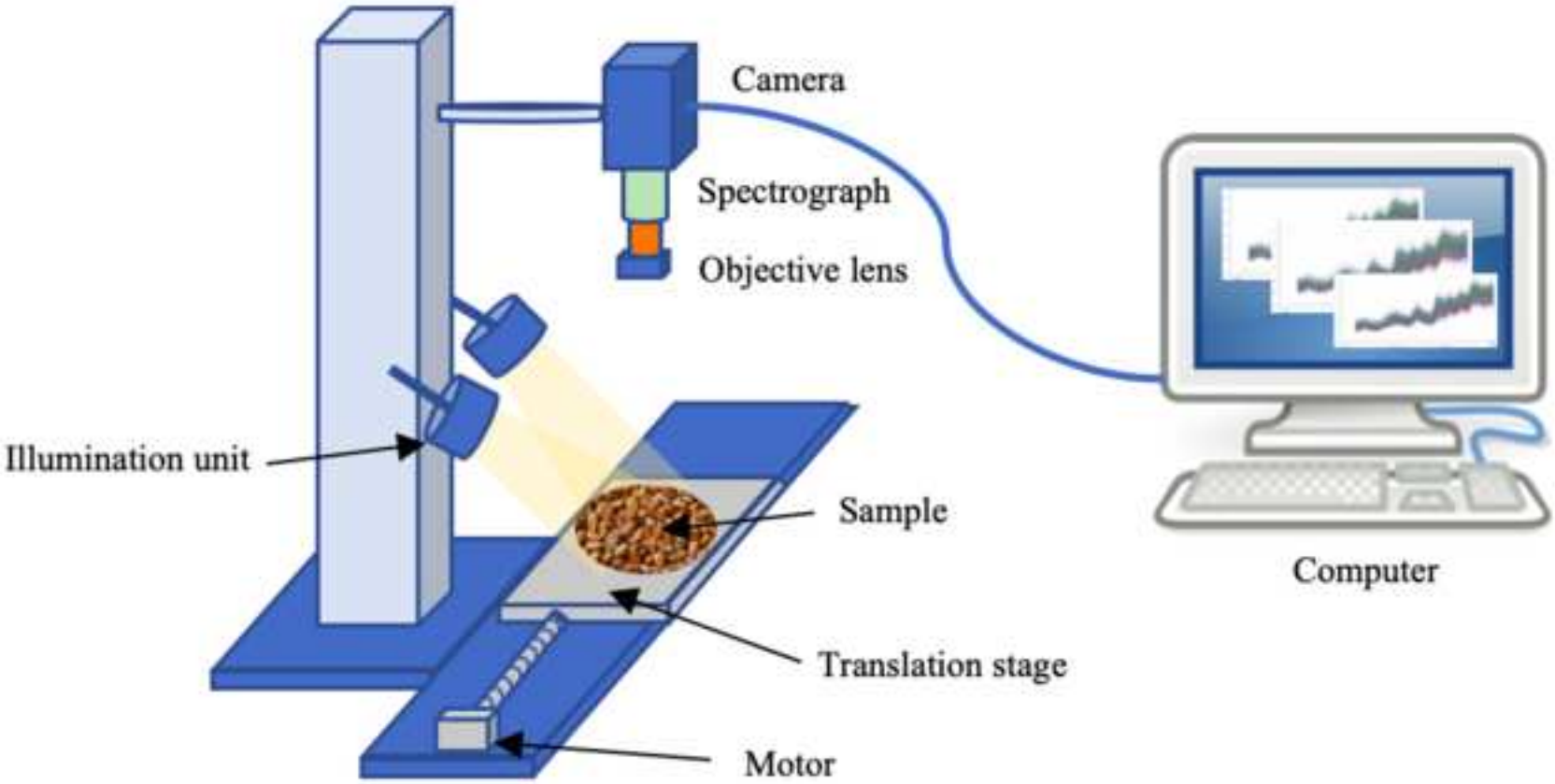
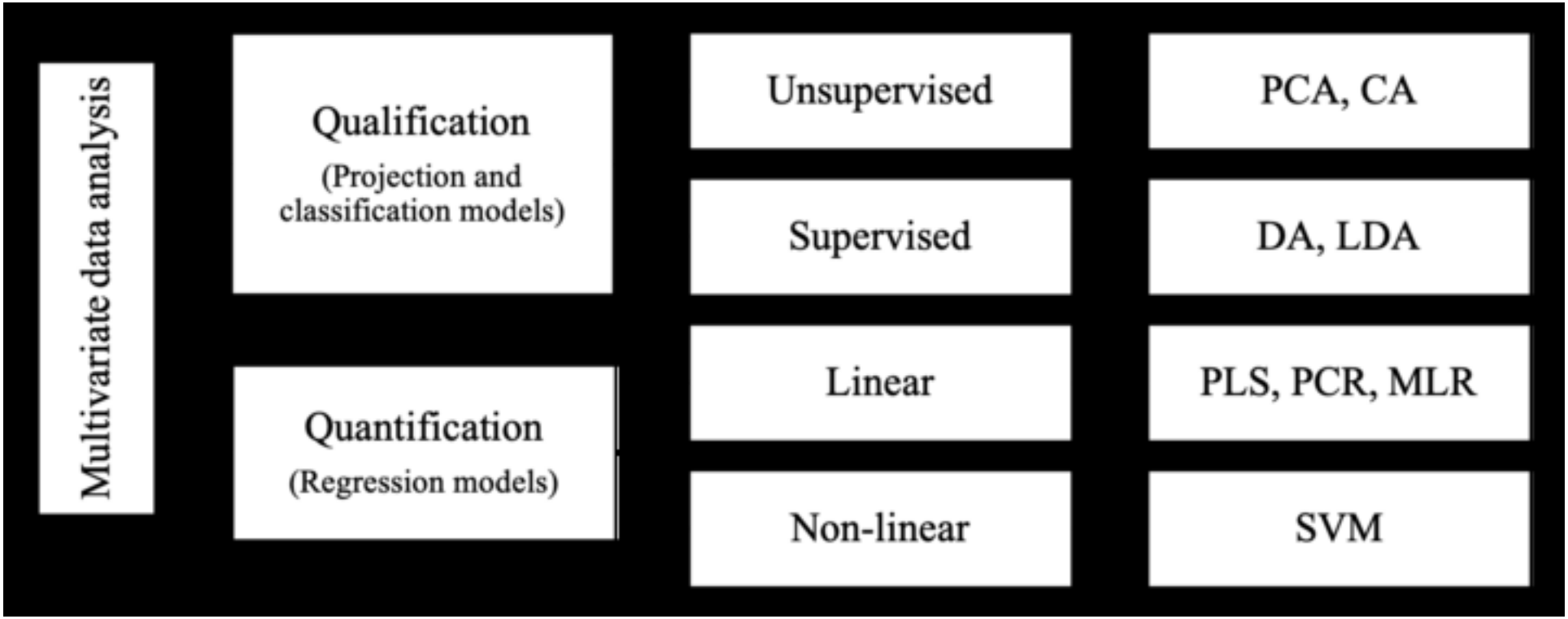

Camera
Spectrograph
Objective lens
Illumination unit
Sample
Translation stage
Motor
Computer

Figure 3


Multivariate data analysis
Qualification
(Projection and classification models)
Quantification
(Regression models)
Unsupervised
Supervised
Linear
Non-linear
PCA, CA
DA, LDA
PLS, PCR, MLR
SVM