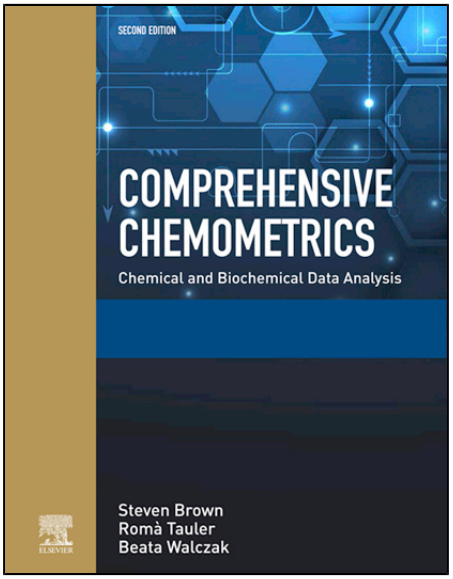

From Fernández Pierna, J. A.; Vermeulen, P.; Eylenbosch, D.; Burger, J.; Bodson, B.; Dardenne, P.; Baeten, V. Chemometrics in NIR Hyperspectral Imaging: Theory and Applications in the Agricultural Crops and Products Sector. In Comprehensive Chemometrics: Chemical and Biochemical Data Analysis; Brown, S., Tauler, R., Walczak, B., Eds., Elsevier, 2020; pp 361–379.
ISBN: 9780444641656

4.14 Chemometrics in NIR Hyperspectral Imaging: Theory and Applications in the Agricultural Crops and Products Sector☆

Juan Antonio Fernández Pierna and Philippe Vermeulen, Walloon Agricultural Research Centre (CRA-W), Quality of Agricultural Products Department, Gembloux, Belgium
Damien Eylenbosch, Walloon Agricultural Research Centre (CRA-W), Production and Sectors Department, Gembloux, Belgium
James Burger†
Bernard Bodson, Gembloux Agro-Bio Tech (GxABT) Faculty, University of Liege, AgroBioChem Department, Gembloux, Belgium
Pierre Dardenne and Vincent Baeten, Walloon Agricultural Research Centre (CRA-W), Quality of Agricultural Products Department, Gembloux, Belgium

This is an update of J.A. Fernandez Pierna, V. Baeten, P. Dardenne, J. Dubois, E.N. Lewis, J. Burger, 4.06 - Spectroscopic Imaging, in Comprehensive Chemometrics, edited by Steven D. Brown, Romá Tauler, Beata Walczak, Elsevier, 2009, https://doi.org/10.1016/B978-044452701-1.00004-1.

Abbreviations

ANN Artificial neural networks
DBSCAN Density-based spatial clustering of applications with noise
FOV Field of view
FPA Focal plane array
***k*-NN** *k*-nearest neighbors
LCTF Liquid crystal tuneable filter
LDA Linear discriminant analysis
MCR Multivariate curve resolution
MLP Multi-layer perceptron
MSC Multiplicative scatter correction
NIR Near infrared
NIR-HSI Near infrared hyperspectral imaging
PC Principal component
PCA Principal component analysis
PCR Principal component regression
PLS Partial least squares
PLS-DA Partial least squares discriminant analysis
PMSC Piecewise multiplicative scatter correction
RGB Red green blue
RMSEP Root mean square error of prediction
ROI Region of interest

☆*Change History*: November 2019. JA Fernández Pierna, P Vermeulen, D Eylenbosch, B Bodson, P Dardenne V Baeten updated text and references.
†Deceased.

Comprehensive Chemometrics, Second Edition, 2020, 361–379

RPD Ratio between the standard error on reference values of the validation set and the RMSEP
SNV Standard normal variate
SVM Support vector machines

4.14.1 Introduction

In the last two decades, multispectral and hyperspectral chemical imaging have become powerful analytical approaches in several areas including remote sensing to tackle environment, agricultural or mineralogy challenges, assessment of painting features, and for a variety of applications for troubleshooting and quality assurance of pharmaceutical,[1,2] chemical[3,4] and biological[5,6] products, medical and forensic analysis[7,8] or for archeological issues,[9,10] among other.

In the agrofood sector, these optical imaging techniques are increasingly considered the preferred tools in the design of nondestructive food/feed inspection instruments with applications in sample characterization, the measurement of chemical species distribution, the safety evaluation and quality control and the detection of contamination and defects in agrofood products.[11–13]

Of particular interest from a public health standpoint are instruments designed for multispectral or hyperspectral near-infrared (NIR) imaging analysis, which already play and will increasingly hold a key role for automatic food and feed inspection.

This technological advances introduced with the imaging analysis are such that much data is currently easily collected on individual samples. These data are therefore heterogeneous due to their origin. They can be quantitative, qualitative/categorical, repeated in time and space, etc. They may also be highly dependent on each other. This complex set of data, acquired quickly and stored in large volume, is commonly referred as massive data (also defined usually as big data).

A few years ago, the factor limiting the use of this massive data was the computer power. But improvements in storage and calculation make this constraint obsolete. For analysts and chemometricians, this is something of a revolution with hundreds or thousands of spectra (including tens or hundreds of variables) being collected for each sample, instead of the unique average spectrum typically collected with classical spectroscopic instrumentation. The challenge is to handle, extract, and exploit the relevant information contained in this large amount of data now available.[14–17]

The aim of this article is to discuss the integration of chemometric processing tools for studying NIR imaging data. The first part of this article briefly introduces NIR hyperspectral imaging (NIR-HSI) including the factors of success of this technology. The second part is related to the integration of chemometric processing tools for the study of NIR imaging data. Some of the most popular supervised and unsupervised chemometrics tools are reviewed. The third part of this article explains how images and spectra are acquired and finally processed. Finally, a number of applications of chemometrics in the development of NIR imaging analytical methods are explored, with a focus on agricultural crops and products.

4.14.2 Introduction to NIR Imaging

Near infrared imaging instruments probe the spatially resolved chemical composition of samples. The spectral range of the instrument depends on its illumination source, wavelength selection mechanism and detector. Long wavelength infrared imaging systems probe the combination, first and second overtone bands of NH, CH, and OH bonds; intermediate-wavelength instruments measure the first, second and third overtones; short-wavelength systems access only the third overtone. The chemical specificity of the spectral information, greater in the combination bands and gradually decreasing through the first, second and third overtone bands, is the key to the value of this family of techniques aptly named chemical imaging. The success of NIR chemical imaging can be attributed to a combination of different factors: (1) nondestructive method, (2) high performance and availability of uncooled NIR-sensitive two-dimensional array detectors, (3) digitally tuneable infrared optical filters and (4) drastic increase in both speed and capacity of laboratory computing platforms. The integration of these elements has already shown promising results in the determination of quality parameters for complex matrices such as pharmaceutical blends (e.g., Lyon et al.,[18] Reich[19]), food products such as the detection of apple surface defects (e.g., Lu[20]) and contaminations (e.g., Mehl et al.[21]) or for the feed industry (e.g., Fernández Pierna et al.,[22] Fernández Pierna et al.[23]). Indeed, it differs from classical NIR spectroscopy in that the quality attributes can be assessed from both spatial and spectral (and therefore chemical) information. This makes the hyperspectral imaging technique better than classical NIR spectroscopy in extracting details on a smaller scale with the objective of obtaining a better quality evaluation.

Hyperspectral images or hypercubes are three-dimensional data sets containing light intensity measurements where two dimensions (x and y) represent spatial pixel coordinates, while a third dimension (λ) represents spectral variation such as wavelength (Fig. 1). They can be interpreted as stacks of typically hundreds of two-dimensional spatial images at different wavelengths, or tens of thousands of absorbance spectra, aligned in rows and columns. Three instrumentation approaches are used to acquire hyperspectral images. These different approaches can be termed 'point', 'line' or 'plane' scan, based on the orientation of the scanning dimension relative to the two-dimensional spatial sample axes. These three acquisition modalities are described in detail later.

Regardless of the scanning acquisition technique, instruments may be multispectral and hyperspectral in nature. With multispectral instruments, only a limited number of wavelengths (below 10) are collected. Typically, these make use of a set of interference

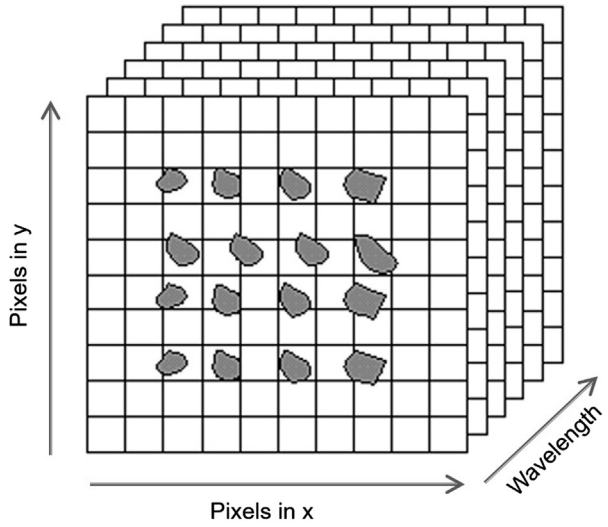

Fig. 1 The three-dimensional image hypercube.

filters to measure the radiation of a fixed or predefined number of spectral bands. The band pass of the filters used may vary from a few tenths up to hundreds of nanometres. On the contrary, hyperspectral instruments allow collection of continuous spectral information from many contiguous or discrete wavelengths (typically >100 wavelength channels) within a specified spectral range.

4.14.3 Chemometrics in NIR Imaging

Traditionally, chemometrics are used with spectroscopic data to yield a measure of composition. The advent of imaging spectrometers adds a new dimension to the data set and it is imperative that the chemometric tools be designed to use this information as well. Structurally intricate samples are well suited for chemical imaging because of spatial and chemical complexity. Pharmaceutical products are a good example of chemically complex architectures and they have become one of the most prolific fields of use of NIR chemical imaging (e.g., Veronin et al.,[24] Dubois et al.,[25] Lee et al.,[26] Lewis et al.[27]). Combining the spatial and chemical information brings about a need for different data-processing modalities in order to exploit the information that is present in the chemical imaging data set. In the following sections, the multistep chemometric approaches that attempt to include the spatial information present in the image as well as the chemistry to better describe and segment complex mixtures are presented and discussed.

The fundamental requirement for NIR chemical imaging experiments is that sample building blocks possess different chemical and physical properties. A chemometric model is used to segment the image based on chemical parameters that are specific to that system, just like it would be used on individual spectra from regular NIR spectroscopy. The contrast contained in the results of the chemometric approaches allows a segmentation of the sample into smaller parts that are characterized by their chemistry and physical parameters such as texture and particle size (like the example of sugar and salt later in this article); these segments may further be measured and described with size and shape parameters such as their elongation, circularity, and convexity. In this article we explore how this processing applied to individual samples can provide valuable phenotypic information or insight into plants and roots.

The strength of chemical imaging resides in the availability of a massive amount of information in the data set; it can never be overemphasized that carefully designed data processing is required for accessing this information. Image-processing techniques can be applied to individual images measured at single-wavelength channels (or combinations of channels) to obtain spatial information content such as feature recognition. Chemometric techniques can also be applied to the spectra from complete hyperspectral data sets as well as subsets of spectra measured within spatial regions of interest (ROI's).

Overall, the purpose of using Chemometrics in NIR Imaging can be summarized into three categories:

- Clustering aiming to create subgroups of samples based on their similarities;
- Classification aiming to explain a qualitative variable based on explanatory variables and lists samples into different classes;
- Regression aiming to explain a quantitative variable from data and provides a prediction;

Mentioned in this way, clustering and classification seem to be very similar but are nevertheless distinguished by a major fact. The first is an unsupervised method and the second is supervised. In other words, regression and classification define their data model/structure based on a target to be reached that corresponds to the variable to be explained. For clustering, there is no target. This method will highlight the links between the different samples.

4.14.3.1 Unsupervised Data Analysis

Unsupervised methods refer to methods not needing any reference information for modeling. In this article, the applications of unsupervised methods are focused on exploratory or clustering purposes.

In pattern recognition problems (e.g., discrimination of incoming batches, authentication of food/feed products, or detection of contaminants), all the variables used to describe the data may not be equally distinctive and informative. In terms of hyperspectral imaging data, it means that it is not difficult to get overwhelmed by the high dimensionality of the data set, which leads to the so-called *curse of dimensionality*. For this reason, optimal variable selection and variable combination methods are important topics in these fields.

In chemometrics, the most widely used method for exploratory analysis is the PCA.[28,29] PCA is simple, easy to use and amply discussed, especially for its use as a pattern recognition and data compression method for signal processing. In the case of image processing, volume reduction is a common objective of the PCA. PCA is a way of identifying patterns in data by reducing the number of dimensions without much loss of information. In other words, PCA is a linear transformation that tries to reduce the dimensionality. This is performed in such a way that the reduced number of dimensions captures most of the informative variance of the data. The new variables are called principal components (PCs) and correspond to the largest eigenvalues of the covariance matrix that account for the largest possible variance in the data. In other words, PCA is based on a decomposition of the data matrix X into two orthogonal matrices: the scores (T) matrix and the loadings (P) matrix. The loadings can be understood as the weights for each original variable when calculating the principal component. The scores matrix contains the original X data in a rotated coordinate system.

Applied on hyperspectral imaging, scores can be displayed as score maps or as density scatter plots (as options to represent individual pixel information). Fig. 2 shows a PCA scheme using a NIR hyperspectral image of a mushroom represented by the hypercube in 3D (X). This hypercube is first unfolded into a 2D matrix before performing PCA and the determination of scores and loadings. As observed, the overlapped PC1-PC2 score map allows putting into evidence the differences between the mushroom itself and the presence of soil. The loading plot of the associated score map describes the most relevant spectral features. The 1450 and 1950 nm wavelengths are the absorption bands with larger peaks, which are mainly related to water content.

Applied on digital photographs, PCA performs a coordinate transformation of the color image represented in the space of fundamental colors (i.e. red, green, and blue, known as RGB). After this transformation, the new axes are the largest eigenvalues of the covariance matrix of the three input images. Then three new images are obtained by projecting the RGB axes onto the three resulting axes. The first axis corresponds to the largest eigenvalue. The two other axes, which must be orthogonal to the previous axis, are linear combinations of the input images that lead to the remaining variability or information that is not correlated to the first axis or to any other PC of higher order.

Fig. 3A represents a digital photograph of size *MxNx3* of a worldwide known Belgian scene. The picture includes the information about intensity of color components stored in three planes. These R, G, and B components are represented in Fig. 3B. Fig. 3C represents the pseudocolor maps of the three reconstructed components for each eigenvalue and of the PCA residuals. The first PC contains most of the intensity information. The other two PCs have a lower signal-to-noise ratio and they would require a filter in order to improve the visualization; this is clear when looking at PC 2.

Owing to the recent availability of imaging spectrometers with high spectral resolution, hyperspectral image compression has become increasingly important.[30] In these cases, the hyperspectral image cube can be considered as a high- dimensional feature

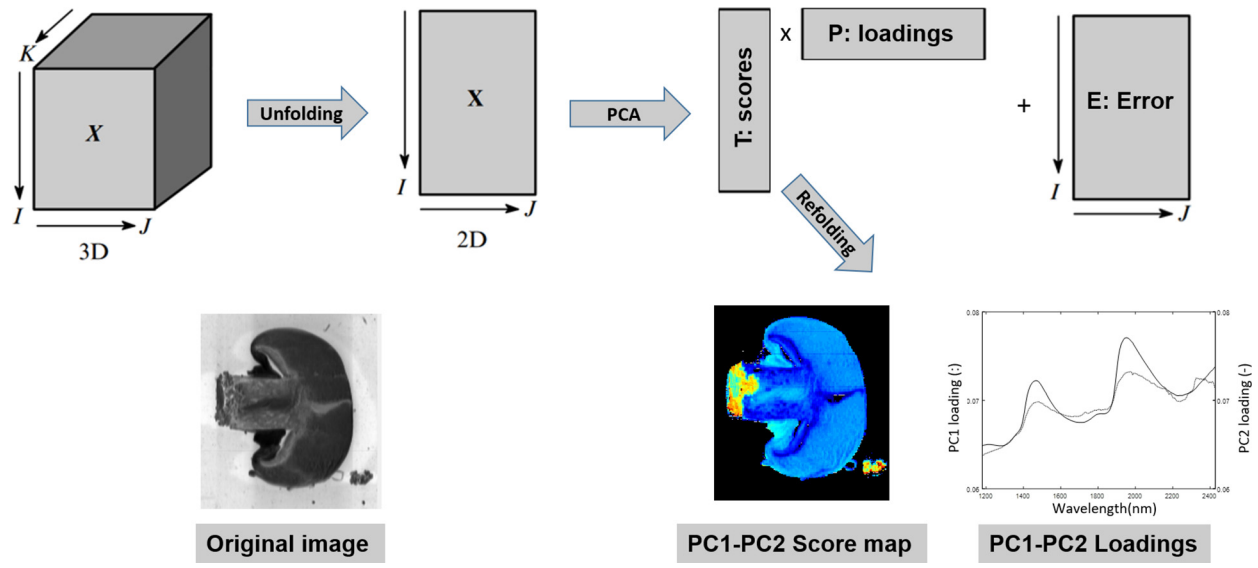

Fig. 2 PCA representation using a NIR hyperspectral image of a mushroom.

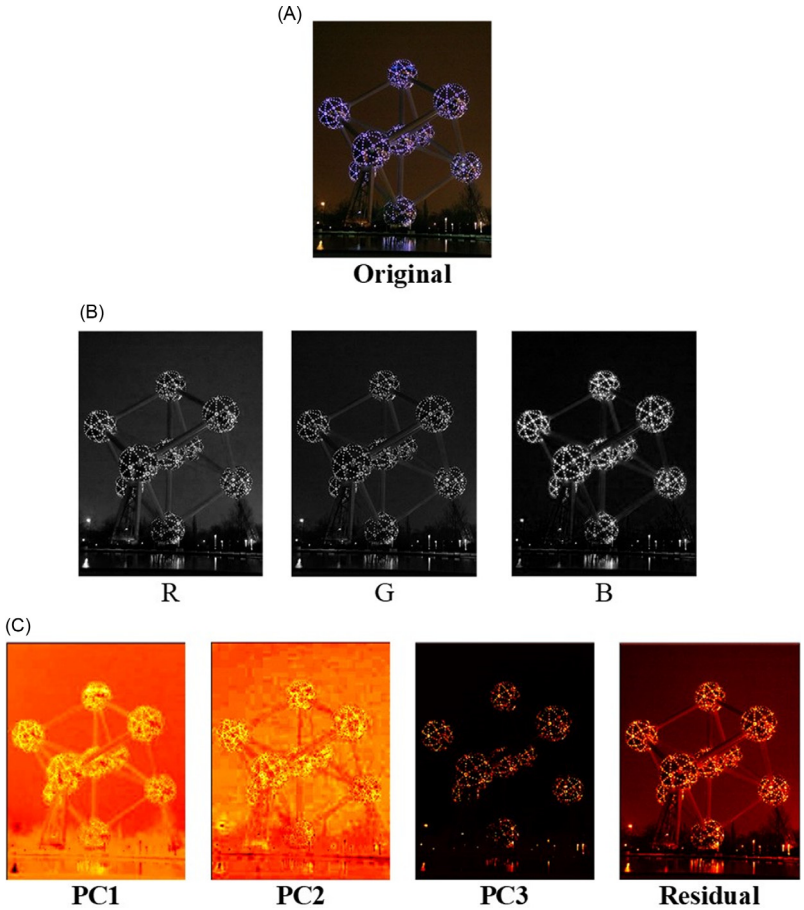

Fig. 3 A worldwide known Belgian scene: (A) digital image; (B) R, G, and B components; and (C) pseudocolor score maps of the three reconstructed components for each eigenvalue and of the PCA residuals.

space where each feature is represented as a spectral image. The (reduced) dimensions obtained after PCA are defined as the ones that preserve the most information among the hyperspectral data cube, that is, the ones that present the best representation of the original data.[31] Examples of the use of PCA in chemical image analysis of agrofood products are presented later in the article.

Even if the applications of unsupervised methods presented in this article are specially focused on exploratory purposes, other techniques must be mentioned as MCR, which is also an unsupervised method with the aim of unmixing pure compound signal and distribution maps out from the sole hyperspectral image information.[32,33]

4.14.3.2 Supervised Data Analysis

Supervised methods refer to methods needing a reference information for modeling, e.g., classification methods, using qualitative class information, and calibration methods, using quantitative reference information. NIR-HSI is used mainly in qualitative/classification studies.

NIR image supervised methods can be performed on the minimum image unit, i.e., the pixel and thus are called per-pixel supervised models. Classical chemometric methods, such as Partial Least Squares (PLS)[34] or Artificial Neural Networks (ANN)[35] are well-known, proven techniques for both classification and regression analysis of multivariate data, such as NIR spectra.[36] More recently, Support Vector Machines (SVM) has been introduced in the hyperspectral imaging treatment. The use of SVM as a chemometric tool applied to spectroscopic data has been proposed by Cogdill and Dardenne[37] among others.[38,39] The use of SVM in the analysis of hyperspectral imaging data is still sparse and the focus is on discrimination problems.[22,23,40–43] Because the theory about all those supervised methods has been explained elsewhere in this book, only examples of their applications are presented in this article.

In contrast to per-pixel methods, object-based image classification methods first segment image pixels into homogeneous objects and classify them into classes.[44] A popular method for this is the density-based spatial clustering of applications with noise method (DBSCAN).[45,46] This technique groups together pixels/spectra that are closely packed together in the original or in the score space, i.e., with many nearby neighbors, and marking as outliers all those points that lie alone in low density regions, i.e., whose nearest neighbors are too far away. In other words, DBSCAN locates regions of high density that are separated from each other by regions of low density. The density at a point P is defined as the number of points within a circle of a certain radius from point P; and

a dense region is considered when, for each point in the cluster, the circle with the defined radius contains at least a minimum number of points previously fixed. After segmentation, features as mean and standard deviation of spectral values can be extracted and used as input for the supervised modeling.

Independently of the method used, the results of the classification/discrimination models are expressed in a confusion matrix in terms of sensitivity and specificity. Sensitivity refers to the percentage of samples (pixels or objects in the NIR-HIS context) from the class studied that have been correctly classified by the model. Specificity refers to the percentage of samples (pixels or objects) not from the class studied that have been correctly classified by the model.[47]

In the case of quantitative analysis, the performance of the models must be assessed in terms of the root mean square error of prediction (RMSEP) and the ratio between the standard error on reference values of the validation set and the RMSEP (RPD).

4.14.4 Image and Spectra Acquisition

Spectroscopic data typically comprise variations in sample illumination and systematic throughput dependencies on wavelengths that must be accounted and corrected for; this is especially true when working with spectroscopic imaging systems. Equally important are any variations due to spatial differences not linked to the chemical composition of the sample, for example, variations of physical origin such as density or glare, and they must be corrected for to focus the chemometric analysis on the chemical information. Moreover and because of the nature of the detectors in the imaging systems, there is a need to use one-dimensional, two-dimensional or three-dimensional references for dark and white in order to perform instrumental calibration prior to further analysis. For this, a dark image is collected by blocking the access of light to the camera and a white image using a standard white reference board. The reflectance is calculated for each pixel at each wavelength using the following equation:

$$R = \frac{\text{Sample} - \text{Dark}}{\text{White} - \text{Dark}}$$

where Sample = image of the sample, Dark = dark current image, and White = white current image.

This is performed to remove noise and to compensate for offset (b_j) due to dark current, light source temperature drift, and spatial lighting non uniformity across the scene.

As introduced earlier, there are different approaches to collect hyperspectral image data (Fig. 4). The process of converting instrument measurement signals to units of reflectance or absorbance depends on the type of instrumentation used for image acquisition.

A point scan instrument (Fig. 4A) acquires a spectrum at a single spatial location using a Fourier transform or grating-type spectrometer. Hyperspectral images are obtained by successively measuring spectra while the sample is repositioned in the X and Y spatial dimensions. Assuming the lighting source remains fixed relative to the spectrometer, a one-dimensional (spectral) instrumental calibration reference for dark and white is performed. This instrument configuration is often used in microscopy utilizing a high-precision X-Y motion stage.

For line-scan imaging systems (Fig. 4B), which project a line of light onto a two-dimensional focal plane array (FPA), a two-dimensional (spatial-spectral) instrumental calibration reference for dark and white is needed to account for variation in sample illumination and instrument throughput. This instrumentation is best suited for remote sensing by aircraft or online process measurement since the Y spatial axis may be arbitrarily long.

Finally, plane-scan imaging systems (Fig. 4C) position the measurement camera parallel to the sample surface, obtaining X-Y spatial images with fixed sizes limited by the dimensions (pixels) of the camera detector. Hyperspectral images are obtained by modulating the radiation reaching the camera using band pass or tuneable filters positioned in front of the camera. In this configuration, a complete three-dimensional (spatial-spatial-spectral) instrumental calibration reference for dark and white must be determined. This instrumental calibration typically uses only two points (dark and bright images), but calibrations with intermediate values have also been designed to account for system nonlinearity.[48]

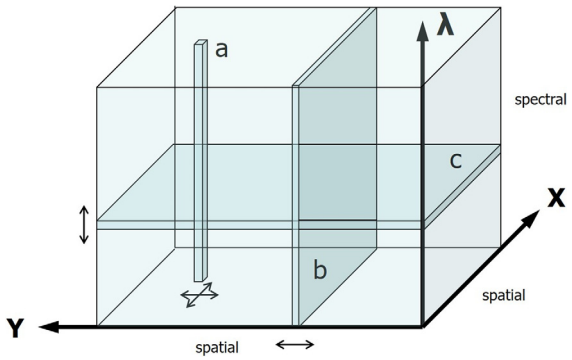

Fig. 4 Modes of acquisition used for hyperspectral image data collection: (A) single-point mapping; (B) line-scan imaging and (C) plane-scan imaging. Vermeulen, P. (2018). Analytical Tools For The Monitoring Of Food Fraud CRA-W INFO 58, 2.

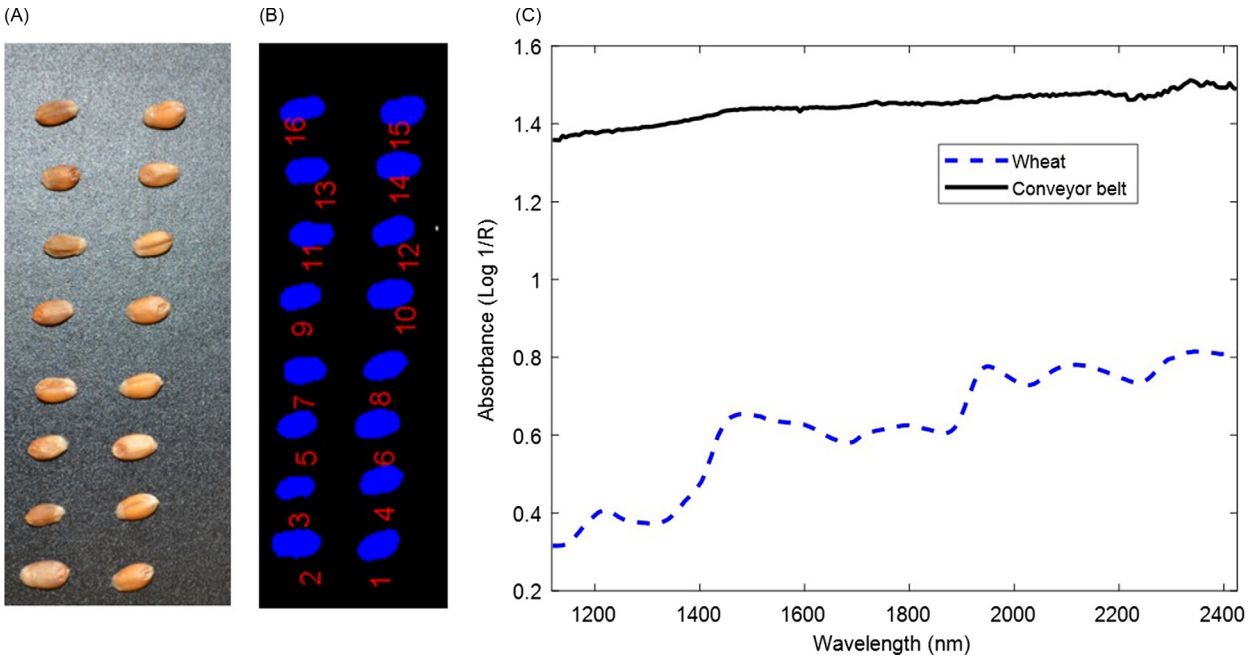

Fig. 5 Wheat kernel: (A) RGB image, (B) NIR image after applying a mask based on the saturation of the absorbance (C) of the conveyor belt.

After instrumental calibration and NIR image acquisition, the next step consists to extract, from the image, the different ROI related to the problem to be studied. For instance, in Fig. 5A, a RGB picture displaying 16 wheat kernels on a conveyor belt is shown. A first mask can be constructed based on the detection of pixels/spectra in the image showing a saturation of the absorbance corresponding to the conveyor belt (Fig. 5B and C).

Moreover, the unsupervised technique PCA can be used to get information on the different spectra present in the image allowing, then, to extract the ROI and therefore build specific spectral libraries. Fig. 6A shows the RGB picture of a young sugar beet plant spread between a tile and a glass, all laid on a conveyor belt in front of a line-scan NIR hyperspectral system. Fig. 6B shows the first PC score, which explains 96.8% of the variance, and is suitable for identifying the background (conveyor belt, tile, glass), the soil and the plant (root, stem, leaves). Spectral libraries can be, then, built by selecting spectra on each ROI (Fig. 6C). Fig. 6D shows the specific mean NIR spectral profiles of each part of the image.

To improve spectral identification, DBSCAN could be applied to study the neighborhood of the pixels detected as ROI. As previously explained, this technique is one of the most common clustering algorithms. It allows grouping together pixels/spectra in clusters and, at the same time, removing the isolated pixels out of the ROI, grain or plant for example.

One final consideration in data conditioning is that an FPA detector may contain a small number of bad or "dead" pixels originating from sensor elements that either fail to respond or respond erroneously. The type of imaging instrumentation setup implies a different impact of such bad sensor pixels on the data set and dictates different approaches to their detection and removal.[49]

A bad sensor element in a line scan configuration affects only the part of the spectrum imaged onto that element. Removing this data value would require eliminating either a complete wavelength channel or the affected spatial channel from the entire hypercube. Such "bad" data elements are routinely simply replaced instead, using median filters that examine the surrounding neighborhood of data values.

Bad pixels in a plane scan image system affect all values within a single spectrum. Correction steps depend on the type of analysis. In the case of chemometric processing of sets of spectra, bad spectra can be detected as outliers and removed. In the case of spatial image processing, bad pixels can be replaced with the median value of neighboring pixels. The impact of the number of bad pixels must be considered both as a function of their spatial distribution and a proportion to the total number of detectors. Clusters of tens of bad pixels greatly impact the image by removing all spatial resolution for the entire area of the cluster, while tens of isolated bad pixels in an array containing tens of thousands of pixels have relatively no impact.

How are bad data values detected? Simple thresholding will identify sensors stuck "on" or "off." For example, with 12-bit digital data values ranging between 0 and 4095, data values <100 or >4000 may be suspect. PCA score plots of image spectra will often reveal significant outliers where an individual sensor may give an erroneous result that otherwise lies within the acceptable data value range. For example, Fig. 7 shows PC1-PC2 score plots (Fig. 7A) before and (Fig. 7B) after replacement of bad line scan data with neighborhood values. The data are taken from 160,000 reflectance spectra of a single cheese sample. Each line scan image (320×256 pixels) contained 298 bad pixels. The strange clustering of points observed is due primarily to sets of spectra containing individual values near 0 or 4095. These extreme outliers have very significant leverage and must be detected and removed to obtain accurate PCA or PLS calibration models. Because hyperspectral images often contain tens or even hundreds of thousands of spectra, outlier spectra can be liberally removed while retaining quality spectra representative of the true sample nature.

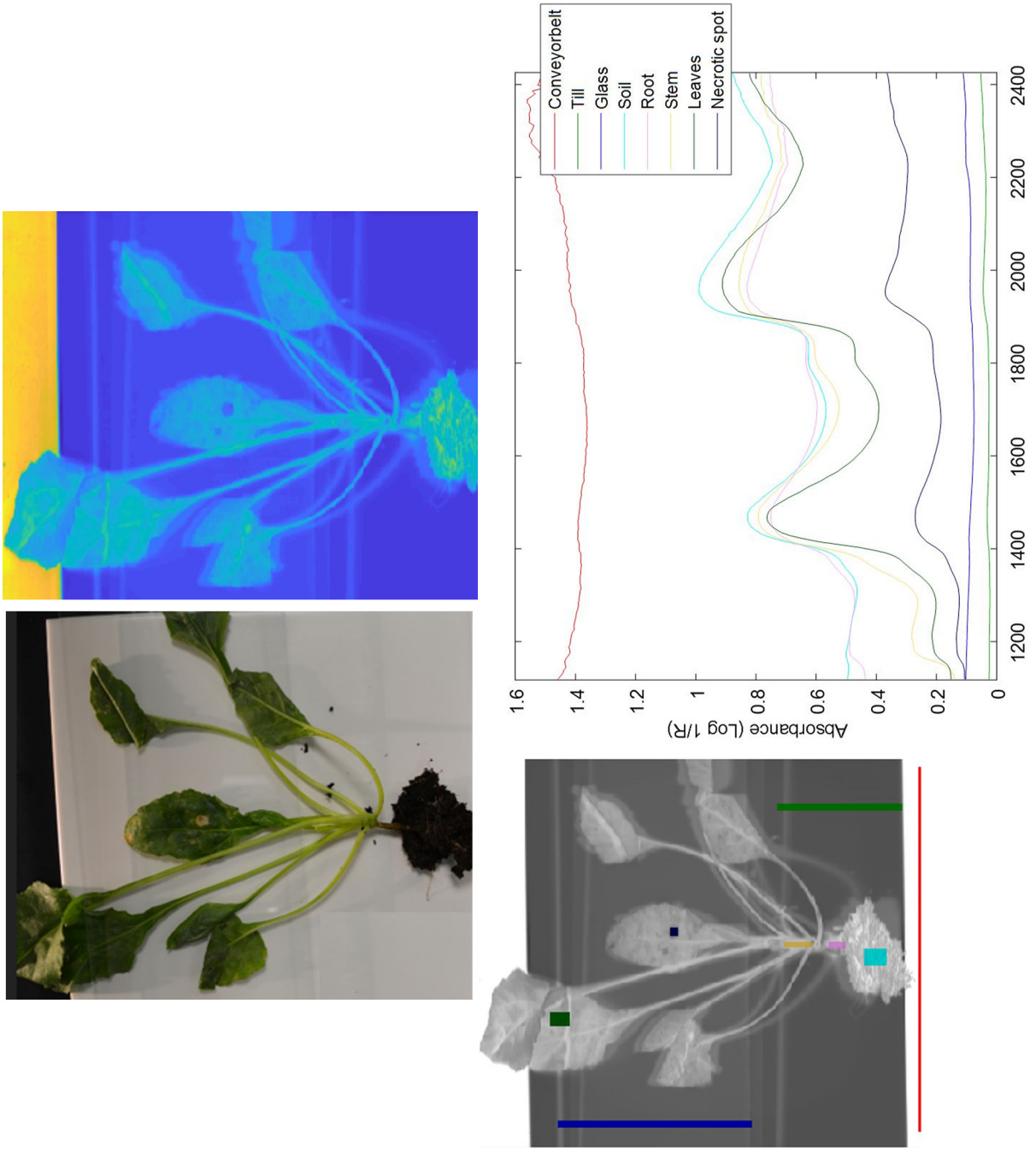

Fig. 6 Sugar beet plant: (A) RGB image, (B) score map of PC1 applied on the NIR image, (C) selection of ROI, and (D) spectral profiles of ROI.

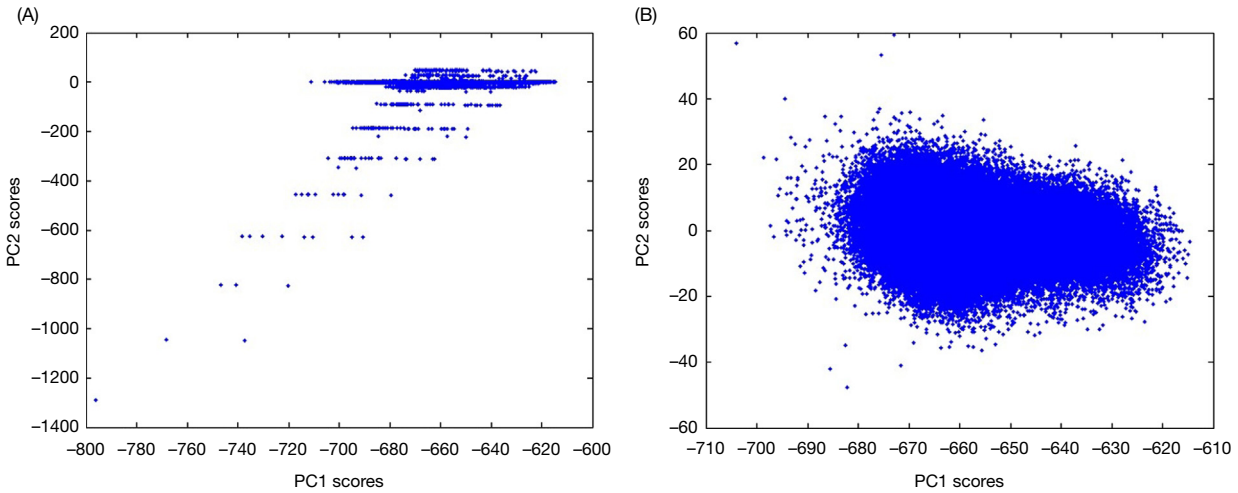

Fig. 7 PC1-PC2 score plots (A) before and (B) after replacement of bad line scan data with neighborhood values.

4.14.5 Spectral Processing

As with conventional spectroscopy, hyperspectral image reflectance spectra can be further transformed or filtered to remove unwanted nonlinear, additive, and multiplicative effects. Transformation to absorbance units estimated as the base 10 logarithm of reflectance, or application of the Kubelka-Munk transform[50] may help to "linearize" diffuse reflectance spectra making them proportional to chemical constituent concentrations, which is frequently the objective of spectroscopic analysis.

Changes in sample orientation, particle size distributions, packing, instrumentation hardware, and/or analytical environment such as lamp intensity, temperature, or detector response may result in background signal that is added throughout the spectrum. The application of first and second derivative Savitzky-Golay transforms[51–54] can compensate for constant additive effects.

Multiplicative Scatter Correction (MSC)[55,56] can be applied to correct for particle light scatter effects when sample preparation or data acquisition tools cannot be adapted to limit this effect. Since the scatter effect may not be exactly the same for all wavelength ranges, Isaksson and Kowalski[57] proposed correcting the spectral value at each wavelength with independent offset and slope correction terms. This technique is called Piecewise Multiplicative Scatter Correction (PMSC). Alternatively, the Standard Normal Variate (SNV) transform was proposed by Barnes et al.[58] and has been shown to be equivalent to the MSC transform, differing only in scaling factors.[59]

Other preprocessing methods (smoothing, detrend, etc.) could be used but they are not the subject of this article.[60]

Hyperspectral imaging offers unique opportunities to closely examine the interaction between preprocessing treatments, instrumentation, and sample problems. Fundamental particle scattering effects can be explored, for instance, by examining the large populations of spectra contained within individual hyperspectral images. Large population statistics and data visualization tools can be employed to compare the effectiveness of different preprocessing treatments.

A series of hyperspectral NIR images (960–1662 nm) were acquired of ordinary table salt and sugar, separated into eight different particle size fractions ranging between >400 and <63 μm. From each image 250 spectra were randomly selected from the two sets of images (salt and sugar), thereby providing a combined spectral data set containing 4000 spectra. PCA was applied to mean-centered data, following different spectral preprocessing treatments. The results are shown in Fig. 8A, where the plus symbols represent sugar and the circles represent salt. The arrow indicates the direction from small to large particle size fractions. Each size fraction has been color coded as well.

Sugar has a much stronger NIR absorbance than salt; consequently, the scattering effects are greater. This is indicated in the first plot for absorbance spectra (Fig. 8A). The variation in score values for each particle size fraction is much greater for the sugar samples, and increases with particle size. The first derivative transform (Fig. 8B) reduces the variance in sugar relative to that in salt; however, the trend in particle size group clusters is still clearly evident. The second derivative transformation (Fig. 8C) results in a similar reduction in relative variance in the sugar samples.

One approach to MSC corrections was based on first computing individual mean spectra for each of the eight size fractions of both sugar and salt samples. These 16 spectra then became the target spectra for the respective size fraction MSC corrections. As indicated in Fig. 8D, this technique greatly reduces the within-fraction variance—each cluster is significantly tighter when compared to the untreated absorbance (Fig. 8A). The general trend between different size fractions is the same. Fig. 8E indicates the results when a single target spectrum is used for each salt and sugar (the mean of the eight respective size fraction spectra) and piecewise MSC is applied. Here, the variation in sugar spectra has been dramatically reduced. Compared with sugar, the salt has very little absorbance—consequently, the piecewise MSC models a much higher noise contribution indicated by the larger relative cluster sizes. The overall trend in particle size dependency has been reduced for both salt and sugar.

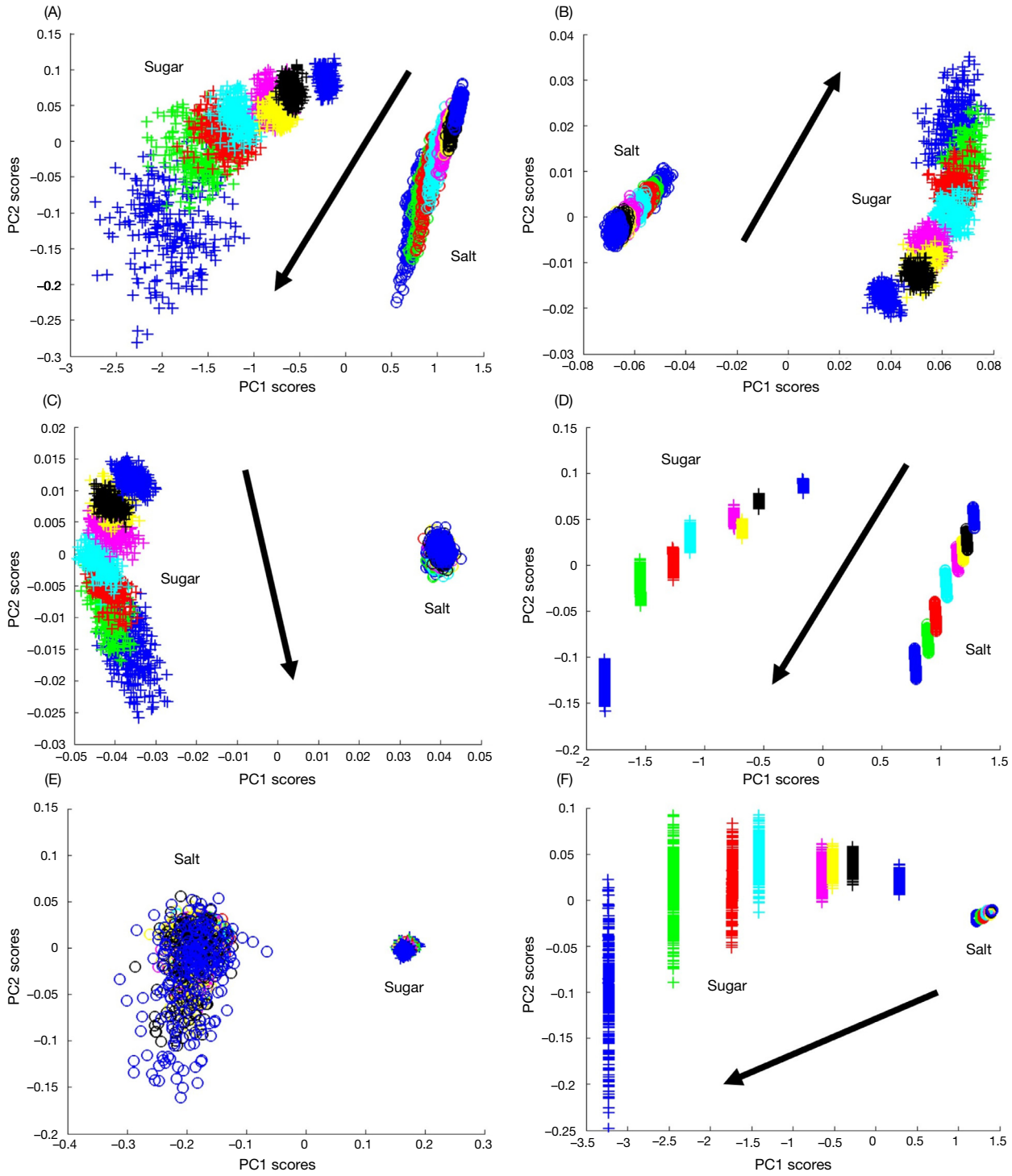

Fig. 8 PCA applied to salt and sugar mixture. The plus symbol represents sugar and the circle represents salt. An *arrow* has been added to indicate the direction from small to large particle size fractions. (A) Absorbance, (B) first derivative, (C) second derivative, (D) Absorbance + image MSC, (E) Absorbance + global PMSC, and (F) Kubelka-Munk + image MSC.

To indicate the exploratory nature of this process, Fig. 8F shows the results of applying MSC with targets based on particle size fractions as in Fig. 8D, but to spectra first transformed using the Kubelka-Munk transformation on reflectance spectra. As with the second derivative transform (Fig. 8C), the size dependency of the salt fractions has been nearly completely removed. These figures indicate some of the possible ways to combine exploratory analysis with preprocessing and PCA to examine the group populations

of spectra obtained with hyperspectral imaging. Such exploratory analysis of class populations is realistically possible only with the immense number of spectra available from hyperspectral images.

After pre-processing methods, classification models can be developed based on the spectral libraries as reference, defined by previous knowledge or by PCA for instance, in order to discriminate plant from soil, classes of grains, … Depending on the study, additional libraries can be built to identify germ, starch and hull in a kernel, to identify impurities in grain, to identify stresses on plants. Models can be developed and applied in a dichotomist way.[61]

4.14.6 Applications on Agricultural Products and Crops

4.14.6.1 Cereal Grains

The following examples of NIR images aim to demonstrate the potential of this technique to assess damaged kernels and identify kernel types. Fig. 9 shows a NIR chemical image at 1430 nm following processing with a second derivative using the Savitzky-Golay algorithm (third polynomial degree and a gap of 15 points). The image shows three maize grains with different characteristics: The grain at the left part of the image corresponds to a damaged grain containing an insect (weevil larvae), the grain in the middle is a damaged empty grain, and the grain at the right is a healthy grain. An enlargement of the image at the left that contains the grain with the insect allows to study the spectral difference between the "healthy" part and the damaged part (containing the insect) of the grain.

NIR imaging and PCA can also be used for the detection of the presence of rot in cereals. Fig. 10 shows an example of this problematic with two grains of maize, one of them being rotted (the left part of the image) and the other being healthy. This figure shows the images taken at two different wavelengths: At 1160 nm, the difference between the two grains is enhanced as it is also shown in the spectra, and at 1560 nm, the rotten part of the grain becomes clearer. PCA, and mainly the seventh PC, is also useful in order to put in evidence the rotten part of a grain mainly because of the large loading at 1560 nm as indicated in Fig. 11.

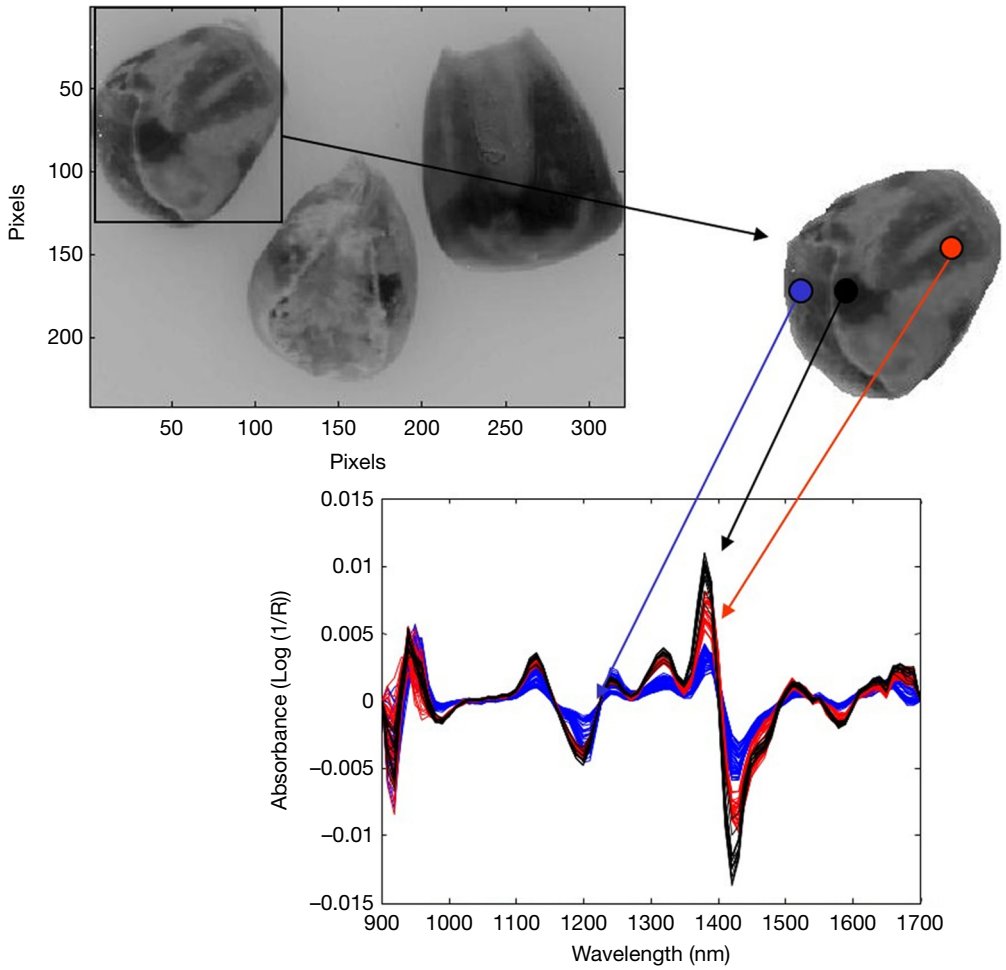

Fig. 9 NIR chemical image at 1430 nm of three maize grains with different characteristics and spectra of different parts for one grain.

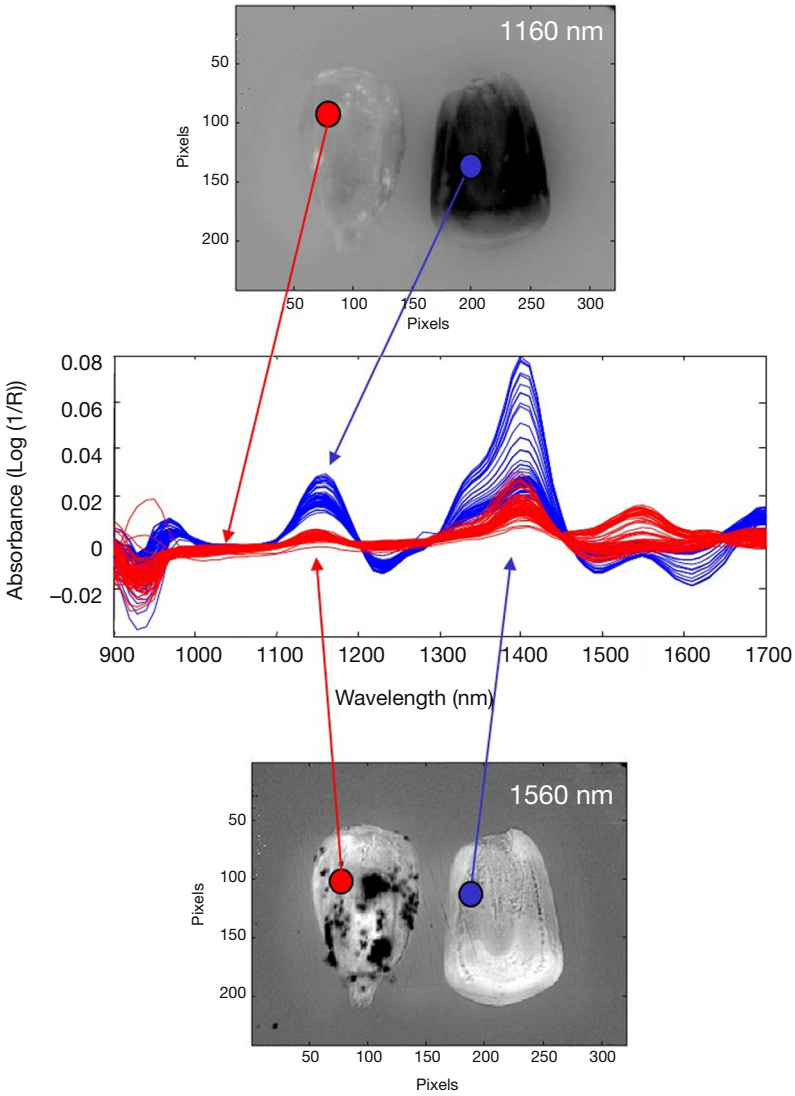

Fig. 10 Example of two grains of maize: one of them being rotted (at the left part of the image) and the other being healthy; and their corresponding spectra.

NIR chemical imaging has shown the potential to provide both speed for high-throughput analysis of large numbers of kernels for quality assessment and a high content of chemical information, two important criteria for plant breeding studies (e.g., Smail et al.[62]; Weinstock et al.[63]). The selection of a larger field of view to image multiple kernels to extract average measurements per unit described previously highlights the value of the spectroscopic image for high throughput analysis. If kernels are spectroscopically imaged using a smaller field of view, i.e., at greater magnification, the data acquired may contain a breath of additional valuable information, which may help establish phenotypic differences. The previous examples made this point clear.

In the next example, a barley kernel was sequentially sectioned and imaged slice by slice in the NIR spectral range 1100–2400 nm for 29 times. Each image is 50 μm deeper into the barley kernel than the previous one. It is noteworthy that a single image acquired from a cross-section of the core of the kernel could be enough to access a large amount of phenotypic information. The interest of the multiple slices is to follow slice after slice, the chemical composition of the kernel in relation to the known structures of the kernel. Fig. 12A shows the RGB picture of a barley kernel sliced at 1 mm displaying the germ, the endosperm, and the hull. Fig. 12B shows the corresponding predicted image by partial least squares discriminant analysis (PLS-DA) displaying the three main fractions: in red the pixels predicted as lipid, in blue the pixels predicted as starch, and in yellow the pixels predicted as cellulose. Fig. 12C shows the percentage of each component for each slice from top to inside the kernel by thickness of 50 μm.

Another application of supervised approaches is the use of PLS regression to perform quantitative analysis of chemical components on the ROI identified previously. To illustrate this, the next application is dedicated to define a new cereal sorting scale based on the quality at the kernel level, as requested by the food/feed sector. For this, several sets of 96 single kernels were analyzed individually by reference chemical analysis for protein content and images were acquired using a NIR hyperspectral system. After applying a mask on the images (Fig. 13A), the mean spectra by kernel was calculated and a PLS model was developed. Fig. 13B

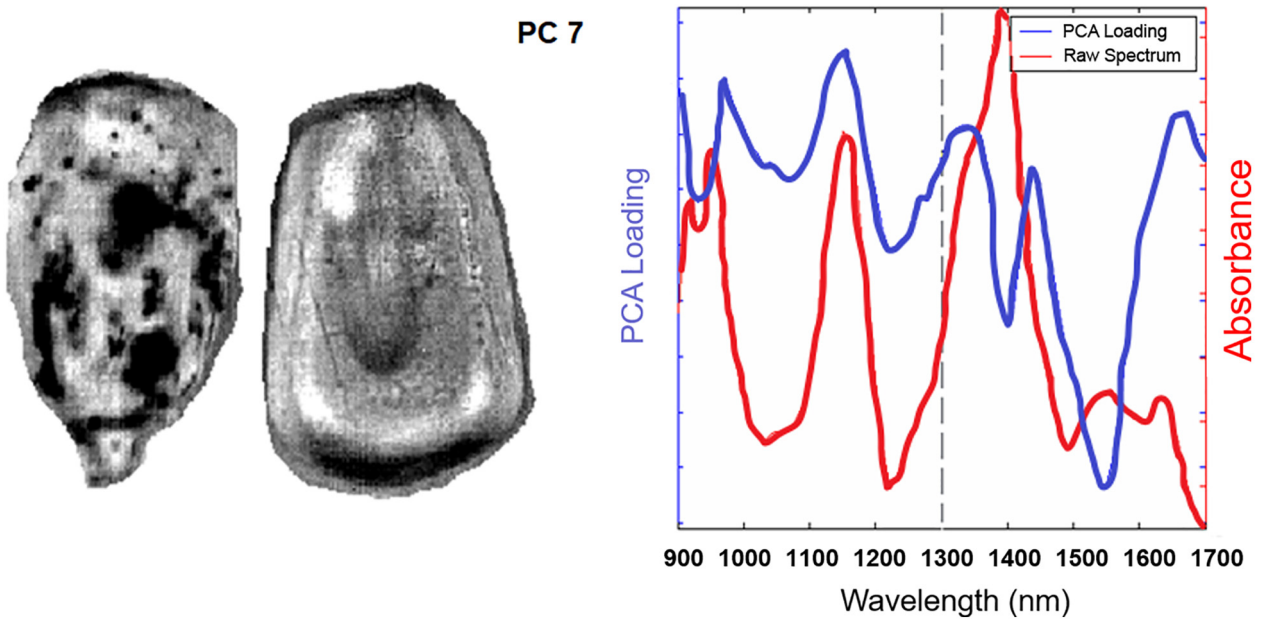

Fig. 11 *Left*: The seventh principal component of **Fig. 9**. *Right*: The *blue line* represents the loading of the 7th PC and the red line is a raw spectrum of the healthy grain.

shows the performance of the PLS model on a validation set: RMSEP = 0.35% and RPD = 3.3. This application has proved that NIR-HSI can do what it is usually performed by classical NIR spectroscopy but at kernel level.[64]

Another application aimed to show the potential of NIR-HSI to assess the quality of pesticide coating treatment on cereals seeds. The application of the correct chemometric tools offers new future prospects for the quality control of the coating efficiency of pesticides on seeds. Vermeulen et al.[65] have shown that PLS-DA models allow sorting the seeds based on the species, on the treated/untreated status and in some cases on the type of pesticide. On another hand, PLS regression models allow classifying the treated seeds in two groups according to the average dose of pesticides: underdosing and overdose against the target dose. This methodology allows also to assess the pesticide coating homogeneity between seeds inside one batch but also to assess the pesticide coating homogeneity at the seed level. **Fig. 14** shows a RGB picture and a predictive NIR hyperspectral image of a wheat seed displaying the treated area in white and the non-treated area in gray.

Other studies performed on that topic concern mainly quality parameter determination in the case of single kernel analysis. Cogdill et al.[66] have used hyperspectral imaging spectroscopy for the study of single maize kernel analysis. After comparison of different spectral preprocessing methods, they have developed predictive calibrations for moisture and oil content using PLS and principal component regression (PCR). The most accurate results were obtained for moisture calibration using PLS on the raw data. The results for the oil content calibration were worse not as accurate as the moisture calibration mainly due to the reference method rather than the spectrometer. In their paper, Stevermer et al.[67] have designed an automated single kernel analysis and sorting system that allows them to construct models for protein determination. The performance of this system has been tested by predicting constituents using PLS and comparing the predicted values with the reference values. Gorretta et al.[68] proposed a combination of a hyperspectral system with PLS-DA for the classification of durum wheat kernels according to their vitreousness in order to create an automatic method to replace the visual method stipulated by the European Union regulations. They obtained a classification rate of up to 94% when separating two nonvitreous classes, and a 100% separation when separating vitreous and nonvitreous kernels. More recently, Vermeulen et al.,[69] were able to discriminate with high level of confidence between durum and common wheat analyzed using hyperspectral imaging kernel by kernel. For this, different criteria were combined in a data fusion based methodology including visual and morphological information, the protein content calculated with a reference method as well as the NIR spectral profile.

In the cereal domain, recent studies include the characterization of specific cereals as fonio,[70] the online detection and quantification of fusarium, as dangerous impurity, in cereals,[71–73] the assessment of salt stress tolerance,[74] cereal classification according to hardness[75] or according to grading regulations[76] or the quantification of protein in wheat using hyperspectral imaging,[77] among others.

4.14.6.2 Plants

Plant phenotyping refers to a quantitative description of the plant's anatomical, physiological and biochemical properties. Today, rapid developments are taking place in the field of non-destructive, image-analysis-based phenotyping that allow for a characterization of plant traits in high-throughput.[78] Discrimination of plants according to botanical families,[79] tracing of plant disease symptoms,[80] estimation of quality control parameters[81] or root detection in soil samples[82] are some of the most recent applications

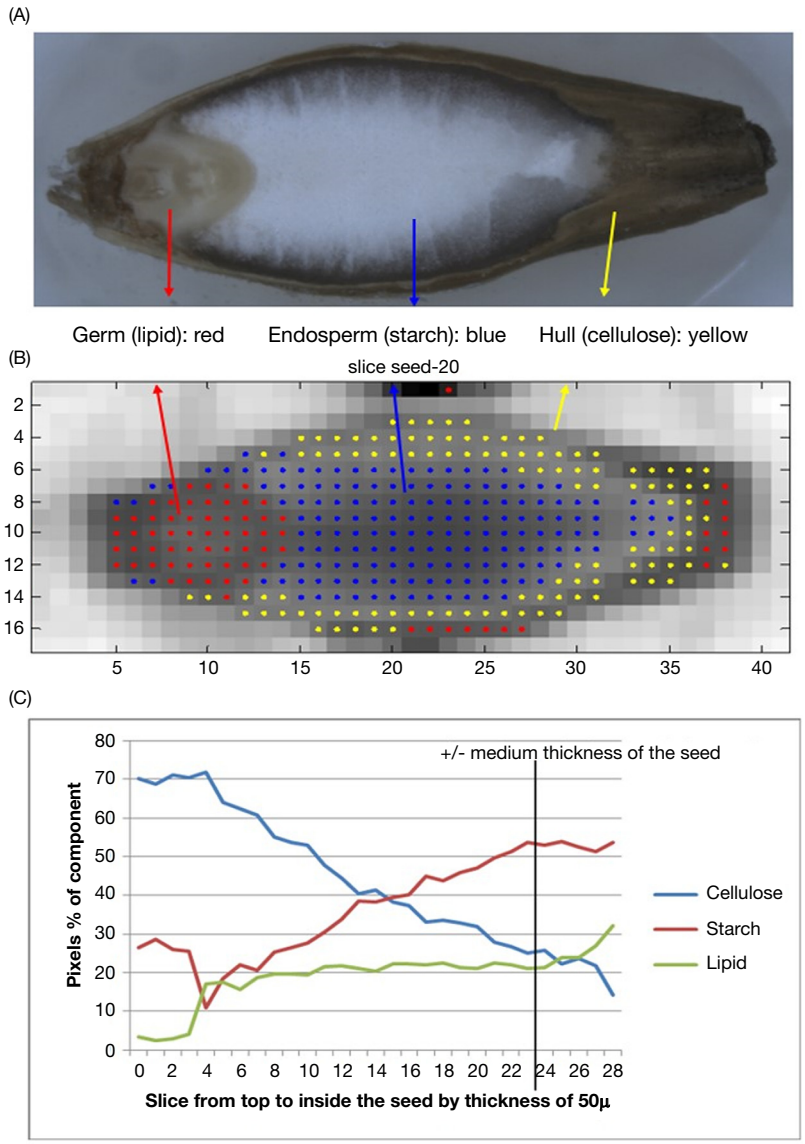

Fig. 12 Barley kernel slice: (A) RGB image, (B) corresponding NIR predicted image by a PLS-DA classification model, and (C) composition change based on the PLS-DA model through the consecutive slices.

where hyperspectral imaging has been performed proving the performance of this technique in combination with chemometrics for the plant characterization and phenotyping.

As example for plant characterization, a sugar beet plant represented in Fig. 15A, was analyzed using a line scan hyperspectral camera. The spectral libraries were selected and used to develop PLS-DA models for identifying leaves in the images in order to assess biotic/abiotic stresses. Fig. 15B shows the plant after applying a mask defined using a PLS-DA model "background (conveyor belt, tile and glass) vs. soil and plant" and after applying DBSCAN. Two clusters were identified: A small one corresponding to a fragment of soil (cluster 2), and a big one corresponding to the object of interest (cluster 1). Successive PLS-DA models were applied in a dichotomist way to discriminate underground from aerial parts of the plant (Fig. 15C), roots from soil (Fig. 15D) and finally, stems from leaves (Fig. 15E).

The following study case is an example of supervised data analysis applied in hyperspectral images for plant characterization. In this case, NIR-HSI and chemometrics were used in order to reduce the time needed to study root systems of legume plant, in particular the composition of nodules present on the roots of peas (*Pisum sativum* L.).[83]

Legume plants can fix atmospheric nitrogen thanks to a symbiosis with bacteria enclosed in nodules, special organs located on their roots. This nitrogen fixation is possible thanks to leghaemoglobin, a protein synthetized in nodules and keeping O_2 concentration in an optimal range in order to supply bacterial respiration and protect nitrogenase enzyme from oxidation.

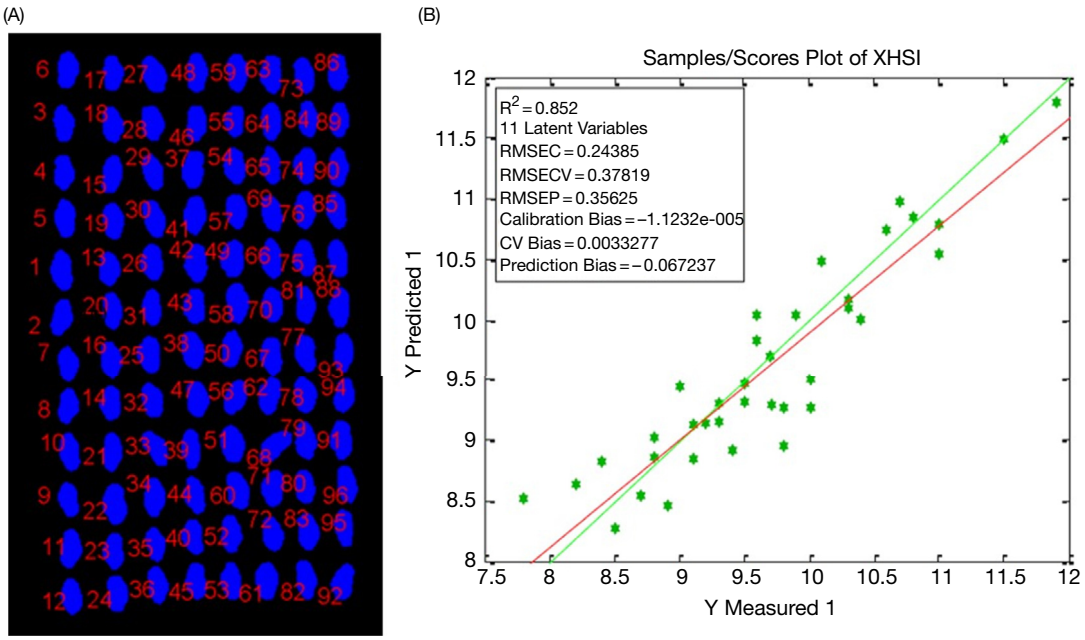

Fig. 13 Set of 96 wheat kernels: (A) mask applied on NIR image, (B) PLS model performance in validation, used for protein prediction at a kernel level. Vermeulen, P. , Flemal Pauline , Pigeon, O. , Dardenne, P. , Fernández Pierna, J.A. & Baeten, V. (2017). Assessment of pesticide coating on cereal seeds by near infrared hyperspectral imaging. J. Spectral Imaging, 6: (a1), 1-7.

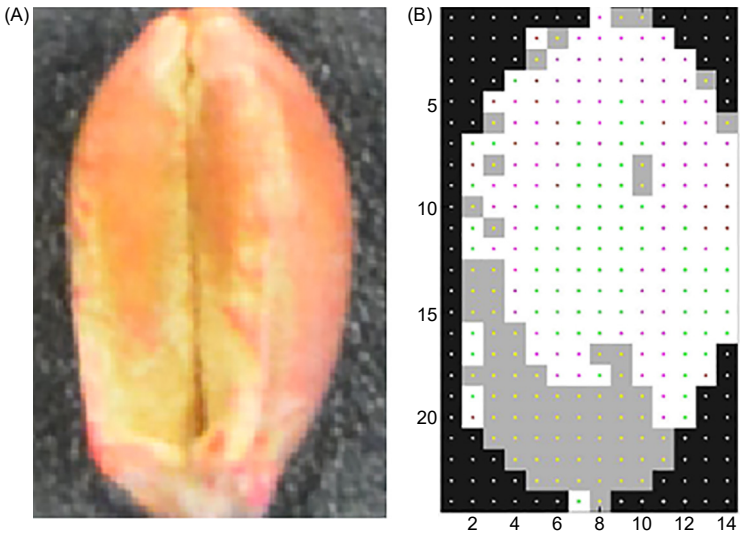

Fig. 14 Coated wheat seed: (A) RGB image, and (B) predictive NIR image displaying the treated area in white and the untreated area in gray.

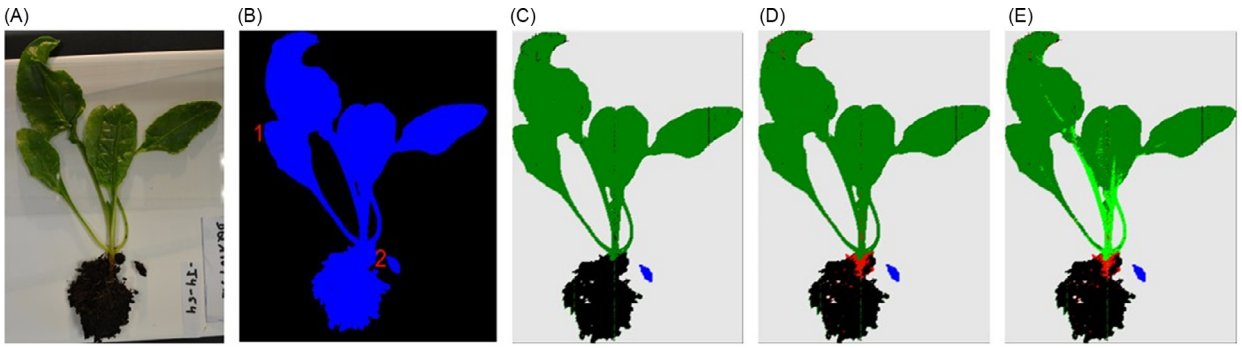

Fig. 15 Sugar beet plant: (A) RGB image, (B) mask applied on NIR image to separate plant from background, (C) predicted image after applying PLS-DA "soil-root vs. stem-leaf," (D) predicted image after applying PLS-DA "soil vs. root," and (E) predicted image after applying PLS-DA "stem vs. leaf."

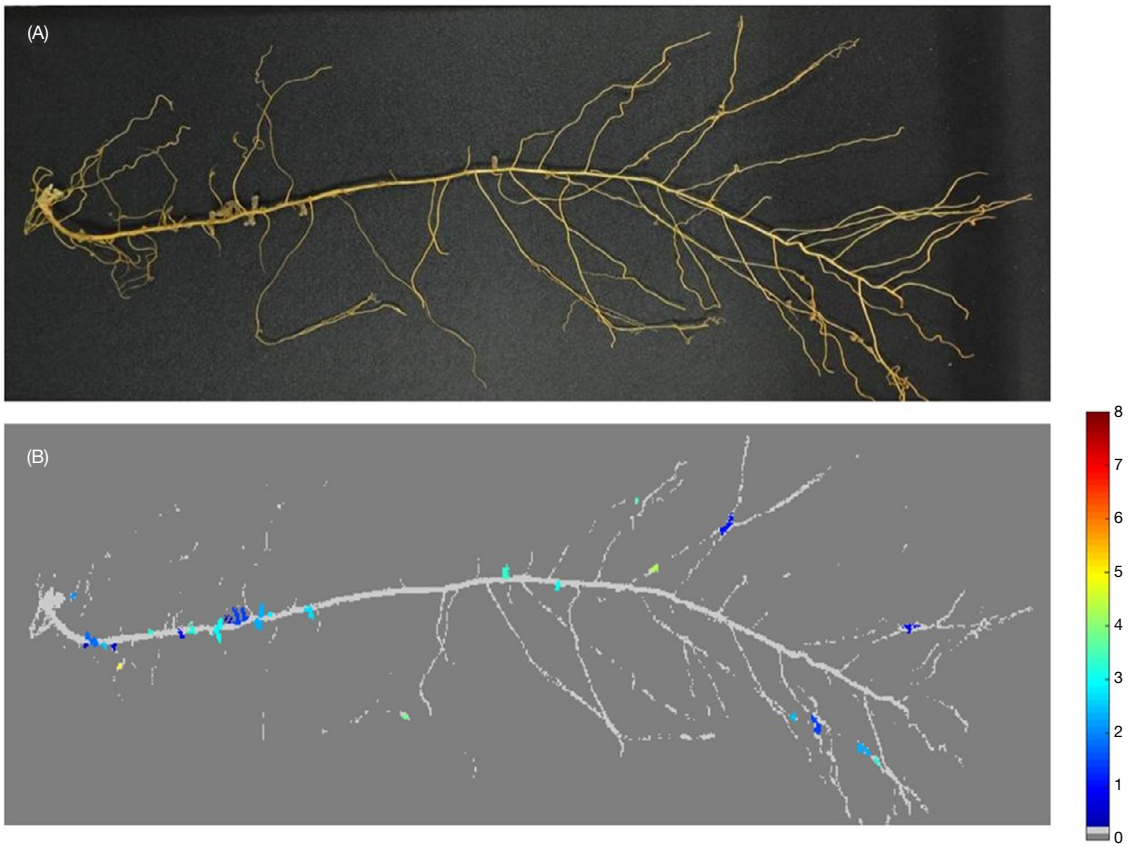

Fig. 16 (A) Pea roots with nodules and (B) image of the same root sample constructed in false color with the predicted leghaemoglobin content of nodules (expressed in mg leghaemoglobin* g^{-1} fresh nodule) after analysis of the hyperspectral image with discriminant and regression models.

Leghaemoglobin content being linked to nitrogen fixing activity of nodules, its quantification gives information on the ability of nodules to fix atmospheric nitrogen.

SVM was used in order to discriminate roots and nodules and PLS was used to quantify leghaemoglobin content in nodules using NIR hyperspectral images acquired using a line-scan instrument. Fig. 16A shows a RGB image of a root system to be analyzed by NIR imaging. With PLS, it was possible to predict the leghaemoglobin content of a single nodule. This single nodule prediction allows then the study of leghaemoglobin repartition among nodules present on a same root system. It can be visualized by constructing a false color image with a color scale linked to leghaemoglobin content (Fig. 16B).

With this method, nodules are evaluated based on NIR hyperspectral images and do not need to be previously separated manually from roots before leghaemoglobin quantification. Leghaemoglobin content is evaluated rapidly and do not need laborious chemical extraction with toxic reagents and the repartition of leghaemoglobin content can be visualized on the false color image of the root system.

4.14.6.3 Other

Another important area of work in the use of hyperspectral imaging and chemometrics in the agronomical domain is the food sector, in particular for the quality control of fruits and vegetables as well as for meat and fish products.[84–90] For instance, in a work aiming to study the ripeness of tomatoes, Polder et al.[91–93] used PCA and Fisher's linear discriminant analysis (LDA; a classical classification method) in order to visualize the data and calibrate the instrument, respectively. The technique, based on NIR-HSI to determine the maturity stage of preclimacteric apples,[94] applied PCA in order to distinguish the starch concentrations within one apple and among several apples during maturation. In the same direction, Menesatti et al.[95] have also applied multivariate analysis to determined starch index in apples using hyperspectral NIR images. Mehl et al.[96] demonstrate that hyperspectral imaging system allowed them to determine scabs, fungal and soil contamination and bruising using either PCA or the absorption intensities at a specific frequency.

Lawrence et al.[97] demonstrated the usefulness of PCA for distinguishing the contaminants inside poultry carcasses. In their study to detect poultry skin tumors, Chao et al.[98] used PCA to select useful bands for detecting tumorous regions. Different studies have been performed for the authentication as well as tenderness estimation in meat and fish.[99–103]

Hyperspectral imaging combined with SVM were also used in the development of an automatic system for pollen identification based on its texture classification.[40] In this study, a texture feature extractor computes image properties on selected ROIs; these texture features are then used for pollen load classification. In a comparison of the classification ability of *k*-NN (*k*-nearest neighbors), MLP (multilayer perceptron), and SVM using optical microscopy, SVM showed a 76% classification rate for the discrimination of the different geographical origin compared to 69% for MLP and 67% for *k*-NN.

Other study tried to automatically sort the different parts of onions, dry peel outer skin and onion bulbs, produced during specific industrial processes based on PLS-DA.[104] Also several studies have been published where hyperspectral imaging is applied for the detection of melamine and cyanuric acid in feed[105] or in milk powder,[106] for the authentication of spices,[107] and cacao,[108] or for the origin of rice,[109] among others.

4.14.7 Conclusion

Spectroscopic imaging, in particular chemical imaging, has rapidly developed over the last two decades into a technology that is deployed in numerous fields of endeavor. In this particular section, we have described a few applications in the area of agricultural science; which has proven to be a fruitful area for the exploitation of this particular technology. One of the reasons for this is the significant spatial/chemical heterogeneity that exists in this area, where structure-function relationships drive performance. On the contrary, because of the complexity/diversity of the data, this presents challenges for the traditional data mining methodologies. It must be solved by adding a completely new multivariate descriptor of the data, which relies, for instance, on the shape and size of intrinsic "objects" within a complex chemical system as an additional discriminator.

References

1. Sacre, P. Y.; De Bleye, C.; Chavez, P. F.; Netchacovitch, L.; Hubert, P.; Ziemons, E. Data Processing of Vibrational Chemical Imaging for Pharmaceutical Applications. *J. Pharm. Biomed. Anal.* **2014,** *101,* 123–140.
2. Amigo, J. M. Practical Issues of Hyperspectral Imaging Analysis of Solid Dosage Forms. *Anal. Bioanal. Chem.* **2010,** *398,* 93–109.
3. Caballero, D.; Bevilacqua, M.; Amigo, J. M. Application of Hyperspectral Imaging and Chemometrics for Classifying Plastics With Brominated Flame Retardants. *J. Spectral Imaging* **2019,** *8,* 1–16.
4. Vidal, M.; Gowen, A.; Amigo, J. M. NIR Hyperspectral Imaging for Plastics Classification. *NIR News* **2012,** *23* (1), 1–13.
5. Levenson, R.; Mansfield, J. R. Multispectral Imaging in Biology and Medicine: Slices of Life. *Cytometry A* **2006,** *69* (8), 748–758.
6. Pisani, M.; Zucco, M.; Caricato, V.; Egidi, A. Hyperspectral Imaging: A Tool for Biological Measurements. Proceedings of the 16th International Congress of Metrology; 2013; p 14007.
7. Lu, G.; Fei, B. Medical Hyperspectral Imaging: A Review. *J. Biomed. Opt.* **2014,** *19* (1). https://doi.org/10.1117/1.JBO.19.1.010901.
8. Edelman, G. J.; Gaston, E.; van Leeuwen, T. G.; Cullen, P. J.; Aalders, M. C. Hyperspectral Imaging for Non-Contact Analysis of Forensic Traces. *Forensic Sci. Int.* **2012,** *223* (1–3), 28–39.
9. Linderholm, J.; Fernández Pierna, J. A.; Baeten, V. NIR Hyperspectral Image Analysis in Archaeology—Analysing Bone Matter in Soils and Sediments. *J. Near Infrared Spectrosc.* **2013,** *21,* 459–466.
10. Vincke, D.; Miller, R.; Stassart, E.; Otte, M.; Dardenne, P.; Collins, M.; Wilkinson, K.; Stewart, J.; Baeten, V.; Fernández Pierna, J. A. Analysis of Collagen Preservation in Bones Recovered in Archaeological Contexts Using NIR Hyperspectral Imaging. *Talanta* **2014,** *125,* 181–188.
11. Baeten, V.; Dardenne, P. Applications of Near-Infrared Imaging for Monitoring Agricultural Food and Feed Products. In *Spectrochemical Analysis Using Infrared Multichannel Detectors;* Bhargava, R., Levin, I. W., Eds., Blackwell Publishing Ltd.: Oxford, UK, 2005.
12. Baeten, V.; Fernández Pierna, J. A.; Dardenne, P. Hyperspectral Imaging Techniques: An Attractive Solution for the Analysis of Biological and Agricultural Materials. In *Techniques and Applications of Hyperspectral Image Analysis;* Grahn, H. F., Geladi, P., Eds., John Wiley & Sons Ltd.: Chichester, UK, 2005.
13. Mendez, J.; Mendoza, L.; Cruz-Tirado, J. P.; Quevedo, R.; Siche, R. Trends in Application of NIR and Hyperspectral Imaging for Food Authentication. *Sci. Agropecu.* **2019,** *10* (1), 143–161.
14. Russ, J. C. *The Image Processing Handbook,* 3rd ed.; CRC Press LLC: Boca Raton, Florida, 1999, 771 p.
15. Amigo, J. M.; Babamoradi, H.; Elcoroaristizabal, S. Hyperspectral Image Analysis. A tutorial. *Anal. Chim. Acta* **2015,** *896,* 34–51.
16. Liu, D.; Sun, D. W.; Zeng, X. A. Recent Advances in Wavelength Selection Techniques for Hyperspectral Image Processing in the Food Industry. *Food Bioproc. Tech.* **2013,** *7* (2), 307–323.
17. Boldrini, B.; Kessler, W.; Rebner, K.; Kessler, R. Hyperspectral Imaging: A Review of Best Practice, Performance and Pitfalls for In-Line and On-Line Applications. *J. Near Infrared Spectrosc.* **2012,** *20* (5), 483–508.
18. Lyon, R. C.; Lester, D. S.; Lewis, E. N.; Lee, E.; Yu, L. X.; Jefferson, E. H.; Hussain, A. S. Near-Infrared Spectral Imaging for Quality Assurance of Pharmaceutical Products: Analysis of Tablets to Assess Powder Blend Homogeneity. *AAPS PharmSciTech* **2002,** *3* (3). article 17.
19. Reich, G. Near-Infrared Spectroscopy and Imaging: Basic Principles and Pharmaceutical Applications. *Adv. Drug Deliv. Rev.* **2005,** *57,* 1109–1143.
20. Lu, R. Detection of Bruises on Apples Using Near-Infrared Hyperspectral Imaging. *Trans. ASAE* **2003,** *46* (2), 523–530.
21. Mehl, P. M.; Chao, K.; Kim, M.; Chen, Y. R. Detection of Defects on Selected Apple Cultivars Using Hyperspectral and Multispectral Image Analysis. *Appl. Eng. Agric.* **2002,** *18* (2), 219–226.
22. Fernández Pierna, J. A.; Baeten, V.; Michotte Renier, A.; Cogdill, R. P.; Dardenne, P. Combination of Support Vector Machines (SVM) and Near Infrared (NIR) Imaging Spectroscopy for the Detection of Meat and Bone Meat (MBM) in Compound Feeds. *J. Chemometr.* **2004,** *18,* 341–349.
23. Fernández Pierna, J. A.; Baeten, V.; Dardenne, P. Screening of Compound Feeds Using NIR Hyperspectral Data. *Chemom. Intel. Lab. Syst.* **2006,** *84,* 114–118.
24. Veronin, M. A.; Youan, B. C. Magic Bullet Gone Astray: Medications and the Internet. *Science* **2004,** *305,* 481, 23.
25. Dubois, J.; Wolff, J. C.; Warrack, J. K.; Schoppelrei, J.; Lewis, E. N. NIR Chemical Imaging for Counterfeit Pharmaceutical Products Analysis. *Spectroscopy* **2007,** *22* (2), 40–50.
26. Lee, E.; Huang, W. X.; Chen, P.; Lewis, E. N.; Vivilecchia, R. V. High-Throughput Analysis of Pharmaceutical Tablet Content Uniformity by Near-Infrared Chemical Imaging. *Spectroscopy* **2005,** *21* (11), 25–32.
27. Lewis, E. N.; Schoppelrei, J.; Lee, E. Near-Infrared Chemical Imaging and the PAT Initiative. *Spectroscopy* **2004,** *19* (4), 22–31.

28. Massart, D. L. M.; Vandeginste, B. G. M.; Buydens, L. M. C.; De Jong, S.; Lewi, J. P.; Smeyers-Verbeke, J. *Chemometrics: A Textbook*; Vol. 2; Elsevier: Amsterdam, The Netherlands, 1988.
29. Vandeginste, B. G. M.; Massart, D. L.; Buydens, L. M. C.; De Jong, S.; Lewi, P. J.; Smeyers-Verbeke, J. In *Handbook of Chemometrics and Qualimetrics: Part B;* Vandeginste, B. G. M., Rutan, S. C., Eds., Elsevier: Amsterdam, The Netherlands, 1998; pp 87–160. Chapter 31.
30. Burger, J.; Geladi, P. Hyperspectral NIR Image Regression Part I: Calibration and Correction. *J. Chemometr.* **2005,** *19,* 355–363.
31. Cheng, X.; Chen, Y. R.; Tao, Y.; Wang, C. Y.; Kim, M. S.; Lefcourt, A. M. A Novel Integrated PCA and FLD Method on Hyperspectral Image Feature Extraction for Cucumber Chilling Damage Inspection. *ASAE Trans.* **2004,** *47* (4), 1313–1320.
32. Forchetti, D. A.; Poppi, R. J. Detection and Quantification of Adulterants in Roasted and Ground Coffee by NIR Hyperspectral Imaging and Multivariate Curve Resolution. *Food Anal. Methods* **2019,** 1–6. https://doi.org/10.1007/s12161-019-01502-x.
33. Zhang, X.; de Juan, A.; Tauler, R. Multivariate Curve Resolution Applied to Hyperspectral Imaging Analysis of Chocolate Samples. *Appl. Spectrosc.* **2015,** *69* (8), 993–1003.
34. Martens, H.; Naes, T. *Multivariate Calibration;,* 2nd ed.; Vol. 1; Wiley: Chichester, UK, 1989.
35. Despagne, F.; Massart, D. L. Neural Networks in Multivariate Calibration. *Analyst* **1998,** *123,* 157R–178R.
36. Naes, T.; Isaksson, T.; Fearn, T.; Davies, T. *A User Friendly Guide to Multivariate Calibration and Classification*, NIR Publications: Chichester, UK, 2002.
37. Codgill, R. P.; Dardenne, P. Least-Squares Support Vector Machines for Chemometrics: An Introduction and Evaluation. *J. Near Infrared Spectrosc.* **2004,** *12* (1), 93–100.
38. Dardenne, P.; Fernández Pierna, J. A. Soil Parameter Quantification by NIRS as a Chemometric Challenge at 'Chimiométrie 2006'. *Chemom. Intel. Lab. Syst.* **2008,** *91,* 94–98.
39. De la Haba, M. J.; Fernández Pierna, J. A.; Fumière, O.; Garrido-Varo, A.; Guerrero, J. E.; Pérez-Marín, D. C.; Dardenne, P.; Baeten, V. Discrimination of the Class Origin of Bones Present in the Sediment Fraction of Animal By-Products Using Near Infrared Microscopy (NIRM). In *Proceedings of the NIR2005; NIR Publications: New Zealand, 2005. Near Infrared Spectroscopy: Proceedings of the 12th International Conference;* Burling-Claridge, G. R., Holroyd, S. E., Sumner, R. M. W., Eds., Near Infrared Spectroscopy Society Incorporated: New Zealand, 2006.
40. Sá Otero, P. Improved Classification of Pollen Texture Images Using SVM and MLP. In 3rd IASTED Int. Conf. on Visualization, Imaging and Image Processing (VIIP2003); Fernández Delgado, M., Carrion Pardo, P., Cernadas García, E., Gálvez Gálvez, J. F., Hamza, M. H., Eds.; Benidorm, Spain, 2003; Vol. 2, pp 686–691.
41. Zhong, S.; Chang, C.; Zhang, Y. Iterative Support Vector Machine for Hyperspectral Image Classification. In *25th IEEE International Conference on Image Processing (ICIP), Athens;* 2018.
42. Moughal, T. A. Hyperspectral Image Classification Using Support Vector Machine. *J. Phys.: Conf. Ser.* **2013,** *439,* 1–9.
43. Marcelo, M. C. A.; Soares, F. L. F.; Ardila, J. A.; Dias, J. C.; Pedó, R.; Kaiser, S.; Pontes, O. F. S.; Pulcinellia, C. E.; Sabina, G. P. Fast Inline Tobacco Classification by near-Infrared Hyperspectral Imaging and Support Vector Machine-Discriminant Analysis. *Anal. Methods* **2019,** *11,* 1966–1975.
44. Jiang, Z.; Shekhar, S. *Spatial Big Dat Science—Classification Techniques for Earth Observation Imagery*, Springer International Publishing: Cham, Switzerland, 2017.
45. Ester, M.; Kriegel, H. P.; Sander, J.; Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases With Noise. In *KDD-96 Proceedings;* 1996; pp 226–231.
46. Daszykowski, M.; Walczak, B.; Massart, D. L. Looking for Natural Patterns in Data: Part 1. Density-Based Approach. *Chemom. Intell. Lab. Syst.* **2001,** *56,* 83–92.
47. Salzer, R.; Siesler, H. W. Infrared and Raman Spectroscopic Imaging. *Anal. Bioanal. Chem.* **2015,** *407* (19), 5551–5552.
48. Tsai, F.; Lin, E. K.; Yoshino, K. Spectrally Segmented Principal Component Analysis of Hyperspectral Imagery for Mapping Invasive Plant Species. *Int. J. Remote Sens.* **2007,** *28* (5), 1023–1039.
49. Burger, J. Bad Pixel Detection in Hyperspectral Staring Camera Systems. *NIR News* **2009,** *20,* 9–12.
50. Kubelka, P.; Munk, F. Ein Beitrag zur Optik der Far-banstriche. *Z. Tech. Phys.* **1931,** *12,* 593–604.
51. Savitzky, A.; Golay, M. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Anal. Chem.* **1964,** *36,* 1627–1639.
52. Hopkins, D. Derivatives in Spectroscopy. *Near Infrared Anal.* **2001,** *2,* 1–13.
53. Giesbrecht, F.; McClure, W.; Hamid, A. The Use of Trigonometric Polynomials to Approximate Visible and Near Infrared Spectra of Agricultural Products. *Appl. Spectrosc.* **1981,** *35,* 210–214.
54. Norris, K.; Williams, P. Optimization of Mathematical Treatments of Raw Near-Infrared Signal in the Measurement of Protein in Hard Red Spring Wheat. Influence of Particle Size. *Cereal Chem.* **1984,** *61,* 158–165.
55. Geladi, P. Multivariate Linearity Transformations for Near-Infrared Spectrometry. In *Proceedings of the Nordic Symposium on Applied Statistics;* Martens, H., Jensen, S., Christie, O., Eds., Stokkand Forlag: Stavanger, 1983; pp 205–233.
56. Geladi, P.; McDougall, D.; Martens, H. Linearization and Scatter-Correction for Near-Infrared Reflectance Spectra of Meat. *Appl. Spectrosc.* **1985,** *39,* 491–500.
57. Isaksson, T.; Kowalski, B. Piece-Wise Multiplicative Scatter Correction Applied to Near-Infrared Diffuse Transmittance Data From Meat Products. *Appl. Spectrosc.* **1993,** *47,* 702–709.
58. Barnes, R.; Dhanoa, M.; Lister, S. Standard Normal Variate Transformation and Detrending of Near Infrared Diffuse Reflectance. *Appl. Spectrosc.* **1989,** *43,* 772–777.
59. Dhanoa, M.; Lister, S.; Sanderson, R.; Barnes, R. The Link Between Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV) Transformations of NIR Spectra. *J. Near Infrared Spectrosc.* **1994,** *2,* 43–47.
60. Rinnan, A.; van den Berg, F.; Balling Engelsen, S. Review of the Most Common Pre-Processing Techniques for Near-Infrared Spectra. *TrAC Trends Anal. Chem.* **2009,** *28* (10), 1201–1222.
61. Fernández Pierna, J. A.; Vermeulen, P.; Amand, O.; Tossens, A.; Dardenne, P.; Baeten, V. NIR Hyperspectral Imaging Spectroscopy and Chemometrics for the Detection of Undesirable Substances in Food and Feed. *Chemom. Intel. Lab. Syst.* **2012,** *117,* 233–239.
62. Smail, V. W.; Fritz, A. K.; Wetzel, D. L. Chemical Imaging of Intact Seeds With NIR Focal Plane Array Assists Plant Breeding. *Vib. Spectrosc.* **2006,** *42,* 215–221.
63. Weinstock, B. A.; Janni, J.; Hagen, L.; Wright, S. Prediction of Oil and Oleic Acid Concentrations in Individual Corn (Zea mays L.) Kernels Using Near-Infrared Reflectance Hyperspectral Imaging and Multivariate Analysis. *Appl. Spectrosc.* **2006,** *60* (1), 9–16.
64. Dijon, A. *Évaluation de différents instruments de spectroscopie proche infrarouge pour la prédiction de la teneur en protéines de céréales analysées en graine à graine.* PhD Thesis, Haute Ecole Charlemagne—ISIa: Huy, Belgium, 2017, 108p.
65. Vermeulen, P.; Flemal, P.; Pigeon, O.; Dardenne, P.; Fernández Pierna, J. A.; Baeten, V. Assessment of Pesticide Coating on Cereal Seeds by Near Infrared Hyperspectral Imaging. *J. Spectral Imaging* **2017,** *6* (a1), 1–7.
66. Cogdill, R. P.; Hurburgh, C. R.; Rippke, G. R. Single Kernel Maize Analysis by Near-Infrared Hyperspectral Imaging. *Trans. ASABE* **2004,** *47* (1), 311–320.
67. Stevermer, S. W.; Steward, B. L.; Codgill, R. P.; Hurburgh, C. R. Automated Sorting and Single Kernel Analysis by Near-Infrared Hyperspectral Imaging. In *Presented at the 2003 ASAE International Meeting, Paper Number: 036159, Las Vegas, Nevada, USA, July 27–30*, American Society of Agricultural and Biological Engineers: St. Joseph, MI, USA, 2003.
68. Gorretta, N.; Roger, J. M.; Aubert, M.; Bellon-Maurel, V.; Campan, F.; Roumet, P. Determining Vitreousness of Durum Wheat Kernels Using Near Infrared Hyperspectral Imaging. *J. Near Infrared Spectrosc.* **2006,** *14,* 231–239.
69. Vermeulen, P.; Suman, M.; Fernández Pierna, J. A.; Baeten, V. Discrimination Between Durum and Common Wheat Kernels Using Near Infrared Hyperspectral Imaging. *J. Cereal Sci.* **2018,** *84,* 74–82.
70. Fernández Pierna, J. A.; Vermeulen, P.; Stilmant, D.; Dupuis, B.; Dardenne, P.; Baeten, V. Characterisation of fonio millet by near infrared hyperspectral imaging. In *Proceedings of the 14th International Conference on NIR Spectroscopy, Bangkok, Thailand;* Saranwong, S., Ed.; 2010. Chichester, UK.
71. Vermeulen, P.; Fernández Pierna, J. A.; van Egmond, H. P.; Dardenne, P.; Baeten, V. Online Detection and Quantification of Ergot Bodies in Cereals Using near Infrared Hyperspectral Imaging. *Food Addit. Contam.* **2012,** *29* (2), 232–240.

72. Vermeulen, P.; Ebene, M. B.; Orlando, B.; Fernández Pierna, J. A.; Baeten, V. Online Detection and Quantification of Particles of Ergot Bodies in Cereal Flour Using Near-Infrared Hyperspectral Imaging. *Food Addit. Contam., Part A* **2017,** *34* (8), 232–240.
73. Williams, P. J.; Geladi, P.; Britz, T. J.; Manley, M. Near-Infrared (NIR) Hyperspectral Imaging and Multivariate Image Analysis to Study Growth Characteristics and Differences Between Species and Strains of Members of the Genus *Fusarium*. *Anal. Bioanal. Chem.* **2012,** *404* (6–7), 1759–1769.
74. Moghimi, A.; Yang, C.; Miller, M. E.; Kianian, S. F.; Marchetto, P. M. A Novel Approach to Assess Salt Stress Tolerance in Wheat Using Hyperspectral Imaging. *Front. Plant Sci.* **2018,** *9,* 1182–1193.
75. Williams, P.; Geladi, P.; Fox, G.; Manley, M. Maize Kernel Hardness Classification by Near Infrared (NIR) Hyperspectral Imaging and Multivariate Data Analysis. *Anal. Chim. Acta* **2009,** *653,* 121–130.
76. Sendin, K.; Manley, M.; Baeten, V.; Fernández Pierna, J. A.; Williams, P. J. Near Infrared Hyperspectral Imaging for White Maize Classification According to Grading Regulations. *Food Anal. Methods* **2019,** *12* (7), 1612–1624.
77. Morales Sillero, A.; Fernández Pierna, J. A.; Sinnaeve, G.; Dardenne, P.; Baeten, V. Quantification of Protein in Wheat Using NIR Hyperspectral Imaging: Performance Comparison With Conventional NIR Spectroscopy. *J. Near Infrared Spectrosc.* **2018,** *26* (3), 186–195.
78. Walter, A.; Liebisch, F.; Hund, A. Plant Phenotyping: From Bean Weighing to Image Analysis. *Plant Methods* **2015,** *11,* 14.
79. Dale, L. M.; Thewis, A.; Boudry, C.; Rotar, I.; Păcurar, F. S.; Abbas, O.; Dardenne, P.; Baeten, V.; Pfister, J.; Fernández Pierna, J. A. Discrimination of Grassland Species and their Classification in Botanical Families by Laboratory Scale NIR Hyperspectral Imaging: Preliminary Results. *Talanta* **2013,** *116,* 149–154.
80. Behmann, J.; Bohnenkamp, D.; Paulus, S.; Mahlein, A. K. Spatial Referencing of Hyperspectral Images for Tracing of Plant Disease Symptoms. *J. Imaging* **2018,** *4,* 143.
81. Dale, L. M.; Fernández Pierna, J. A.; Vermeulen, P.; Lecler, B.; Bogdan, A. D.; Păcurar, F. S.; Rotar, I.; Thewis, A.; Baeten, V. Research on Crude Protein and Digestibility of Arnica montana L. Using Conventional NIR Spectrometry and Hyperspectral Imaging NIR. *J. Food Agric. Environ.* **2012,** *10* (1), 391–396.
82. Eylenbosch, E.; Bodson, B.; Baeten, V.; Fernández Pierna, J. A. NIR Hyperspectral Imaging Spectroscopy and Chemometrics for the Discrimination of Roots and Crop Residues Extracted From Soil Samples: Discrimination of Roots and Crop Residues on Hyperspectral Images. *J. Chemometr.* **2017,** *32* (1). https://doi.org/10.1002/cem.2982. e2982.
83. Eylenbosch, D.; Dumont, B.; Baeten, V.; Bodson, B.; Delaplace, P.; Fernández Pierna, J. A. Quantification of Leghaemoglobin Content in Pea Nodules Based on Near Infrared Hyperspectral Imaging Spectroscopy and Chemometrics. *J. Spectral Imaging* **2018,** *7,* a9.
84. Pu, Y. Y.; Feng, Y. Z.; Sun, D. W. Recent Progress of Hyperspectral Imaging on Quality and Safety Inspection of Fruits and Vegetables: A Review. *Compr. Rev. Food Sci. Food Saf.* **2015,** *14* (2), 176–188.
85. Manley, M. Near-Infrared Spectroscopy and Hyperspectral Imaging: Non-Destructive Analysis of Biological Materials. *Chem. Soc. Rev.* **2014,** *43,* 8200–8214.
86. Mahesh, S.; Jayas, D. S.; Paliwal, J.; White, N. D. G. Hyperspectral Imaging to Classify and Monitor Quality of Agricultural Materials. *J. Stored Prod. Res.* **2015,** *61,* 17–26.
87. Dale, L. M.; Thewis, A.; Boudry, C.; Rotar, I.; Dardenne, P.; Baeten, V.; Fernández Pierna, J. A. Hyperspectral Imaging Applications in Agriculture and Agro-Food Product Quality and Safety Control: A Review. *Appl. Spectrosc. Rev.* **2013,** *48,* 142–159.
88. Baeten, V.; Fernández Pierna, J. A.; Vermeulen, P.; Dardenne, P. NIR Hyperspectral Imaging Methods for Quality and Safety Control of Food and Feed Products: Contributions to Four European Projects. *NIR News* **2010,** *21* (6), 10–13.
89. Huang, H.; Liu, L.; Ngadi, M. Recent Developments in Hyperspectral Imaging for Assessment of Food Quality and Safety. *Sensors* **2014,** *14* (4), 7248–7276.
90. Wu, D.; Sun, D. W. Advanced Applications of Hyperspectral Imaging Technology for Food Quality and Safety Analysis and Assessment: A Review—Part I: Fundamentals. *Innov. Food Sci. Emerg. Technol.* **2013,** *19,* 1–14.
91. Polder, G.; van der Heijden, G. W. A. M.; Young, I. T. *Hyperspectral Image Analysis for Measuring Ripeness of Tomatoes. Presented at the 2000 ASAE International Meeting, Paper number 003089, Milwaukee, WI, USA,* 2000.
92. Polder, G.; van der Heijden, G. W. A. M.; Young, I. T. Spectral Image Analysis for Measuring Ripeness of Tomatoes. *Trans. ASAE* **2002,** *45* (4), 1155–1161.
93. Polder, G.; van der Heijden, G. W. A. M.; Young, I. T. Tomato Sorting Using Independent Component Analysis on Spectral Images. *Real-Time Imaging* **2003,** *9,* 253–259.
94. Peirs, A.; Scheerlinck, N.; De Baerdemaeker, J.; Nicolaï, B. M. Starch Index Determination of Apple Fruit by Means of a Hyperspectral Near Infrared Reflectance Imaging System. *J. Near Infrared Spectrosc.* **2003,** *11,* 379–389.
95. Menesatti, P.; Zanella, A.; D'Andrea, S.; Costa, C. Supervised Multivariate Analysis of Hyper-Spectral NIR Images to Evaluate the Starch Index of Apples. *Food Bioproc. Tech.* **2009,** *2,* 308–314.
96. Mehl, P. M.; Chen, Y. R.; Kim, M. S.; Chan, D. E. Development of Hyperspectral Imaging Technique for the Detection of Apple Surface Defects and Contaminations. *J. Food Eng.* **2004,** *61* (1), 67–81.
97. Lawrence, K. C.; Windham, W. R.; Park, B.; Buhr, R. J. A Hyperspectral Imaging System for Identification of Faecal and Ingesta Contamination on Poultry Carcasses. *J. Near Infrared Spectrosc.* **2003,** *11,* 269–281.
98. Chao, K.; Mehl, P. M.; Chen, Y. R. Use of Hyper- and Multi-Spectral Imaging for Detection of Chicken Skin Tumors. *Appl. Eng. Agric.* **2002,** *18* (1), 113–119.
99. Naganathan, G. K.; Grimes, L. M.; Subbiah, J.; Calkins, C. R.; Samal, A.; Meyer, G. E. Partial Least Squares Analysis of Near-Infrared Hyperspectral Images for Beef Tenderness Prediction. *Sens. & Instrumen. Food Qual.* **2008,** *2* (3), 178–188.
100. ElMasry, G.; Sun, D.-W.; Allen, P. Near-Infrared Hyperspectral Imaging for Predicting Colour, pH and Tenderness of Fresh Beef. *J. Food Eng.* **2012,** *110* (1), 127–140.
101. Kamruzzaman, M.; ElMasry, G.; Sun, D. W.; Allen, P. Non-Destructive Prediction and Visualization of Chemical Composition in Lamb Meat Using NIR Hyperspectral Imaging and Multivariate Regression. *Innov. Food Sci. Emerg. Technol.* **2012,** *16,* 218–226.
102. Kamruzzaman, M.; Barbin, D.; ElMasry, G.; Sun, D. W.; Allen, P. Potential of Hyperspectral Imaging and Pattern Recognition for Categorization and Authentication of Red Meat. *Innovative Food Sci. Emerg. Technol.* **2012,** *16,* 316–325.
103. He, H.-J.; Wu, D.; Sun, D. W. Potential of Hyperspectral Imaging Combined With Chemometric Analysis for Assessing and Visualising Tenderness Distribution in Raw Farmed Salmon Fillets. *J. Food Eng.* **2014,** *126,* 156–164.
104. Vincke, D.; Baeten, V.; Sinnaeve, G.; Dardenne, P.; Fernández Pierna, J. A. Determination of Outer Skin in Dry Onions by Hyperspectral Imaging Spectroscopy and Chemometrics. *NIR News* **2014,** *25* (2), 9–12.
105. Fernández Pierna, J. A.; Vincke, D.; Dardenne, P.; Yang, Z.; Han, L.; Baeten, V. Line Scan Hyperspectral Imaging Spectroscopy for the Early Detection of Melamine and Cyanuric Acid in Feed. *Journal of NIRS* **2014,** *22* (2), 103–112.
106. Lim, J.; Kim, G.; Mo, C.; Kim, M. S.; Chao, K.; Qin, J.; Fu, X.; Baek, I.; Cho, B. K. Detection of Melamine in Milk Powders Using Near-Infrared Hyperspectral Imaging Combined With Regression Coefficient of Partial Least Square Regression Model. *Talanta* **2016,** *151,* 183–191.
107. Kiani, S.; van Ruth, S. M.; van Raamsdonk, L. W. D.; Minaei, S. Hyperspectral Imaging as a Novel System for the Authentication of Spices: A Nutmeg Case Study. *LWT- Food Sci. Technol.* **2019,** *104,* 61–69.
108. Rogez, H.; Fernández Pierna, J. A.; Souza, J.; Baeten, V. Application of NIR Hyperspectral Spectroscopy for the Analysis of Cocoa Beans. 17th International Conference on Near Infrared Spectroscopy 2015, Campinas, Brazil.
109. Mo, C.; Lim, J.; Kwon, S. W.; Lim, D. K.; Kim, M. S.; Kim, G.; Kang, J.; Kwon, K. D.; Cho, B. K. Hyperspectral Imaging and Partial Least Square Discriminant Analysis for Geographical Origin Discrimination of White Rice. *J. Biosyst. Eng.* **2017,** *42* (4), 293–300.