Original Article

JNIRS

Journal of Near Infrared Spectroscopy
0(0) 1–10

DOI: 10.1177/0967033518780506
journals.sagepub.com/home/jns
SAGE

Quantification of protein in wheat using near infrared hyperspectral imaging: Performance comparison with conventional near infrared spectroscopy

Ana M. Sillero[1], Juan A. Fernández Pierna[2], George Sinnaeve[2], Pierre Dardenne[2] and Vincent Baeten[2]

Abstract

Hyperspectral imaging is a powerful technique that combines the advantages of near infrared spectroscopy and imaging technologies. Most hyperspectral imaging studies focus on qualitative analysis, but there is growing interest in using such technique for the quantitative analysis of agro-food products in order to use them as universal tools. The overall objective of this study was to compare the performance of a hyperspectral imaging instrument with a classical near infrared instrument for predicting chemical composition. The determination of the protein content of wheat flour was selected as example. Spectra acquisition was made in individual sealed cells using two classical near infrared instruments (NIR-DS and NIR-Perstop) and a near infrared hyperspectral line-scan camera (NIR-HSI). In the latter, they were also acquired in open cells in order to study the possibility of accelerating the measurement process. Calibration models were developed using partial least squares for the full wavelength range of each individual instrument and for the common range between instruments (1120–2424 nm). The partial least squares models were validated using the "leave-one-out" cross-validation procedure and an independent validation set. The results showed that the NIR-HSI system worked as well as the classical near infrared spectrometers when a common wavelength range was used, with an r^2 of 0.99 for all instruments and Root Mean Square Error in Prediction (RMSEP) values of 0.15% for NIR-HSI and NIR-DS and 0.16% for NIR-Perstop. The high residual predictive deviation values obtained (8.08 for NIR-DS, 7.92 for NIR-HSI, and 7.56 for NIR-Perstop) demonstrate the precision of the models built. In addition, the prediction performance with open cells was almost identical to that obtained with sealed cells.

Keywords

Hyperspectral imaging, near infrared spectroscopy, wheat, protein content

Received 5 March 2018; accepted 30 April 2018

Introduction

Since near infrared (NIR) spectroscopy was first applied in agriculture by K.H. Norris in 1964, to measure moisture in grains,[1] this technique has been widely used by the agro-food industry for determining the chemical composition and other quality properties in biological samples. NIR spectroscopy is based on sample absorption at specific wavelengths of incident radiation. This absorption depends on the chemical composition and physical state of the sample. Like other vibrational spectroscopic techniques, NIR is classified as an indirect analytical technique because a calibration step is required in order to subsequently predict a particular property from the spectra of unknown samples. Robust and flexible NIR instruments, as well as easy-to-use software for building calibration models, are now available for routine control. Chemical determination based on NIR spectra is fairly simple, rapid, nondestructive, and cost effective. It requires a low input of reagents and can be implemented at-line, on-line, or in-line at the site of food

[1]Departamento de Ciencias Agroforestales, Universidad de Sevilla, Sevilla, Spain
[2]Valorisation of Agricultural Products Department, Walloon Agricultural Research Centre, Gembloux, Belgium

Corresponding author:
Juan A. Fernández Pierna, Valorisation of Agricultural Products Department, Walloon Agricultural Research Centre, Henseval Building, Chaussée de Namur, 24, Gembloux 5030, Belgium.
Emails: j.fernandez@cra.wallonie.be; foodfeedquality@cra.wallonie.be

or feed production. At the laboratory level, under ISO17025 conditions and taking account of all the sample preparation and analysis steps, up to 100 samples of agro-food products can be analyzed in duplicate in one day and by one analyst.[2–4]

More recently, NIR hyperspectral imaging (HSI) has been developed as a powerful technique that combines the advantages of NIR spectroscopy and imaging technologies. The great advantage of HSI is its ability to acquire simultaneously spectral and spatial information from a sample. Whereas classical NIR instruments collect spectral data from an area of a fraction of the sample, HSI collects spatially distributed NIR spectra responses at many subsample areas or pixels (usually hundreds or thousands) of the sample.[5–7] Each hyperspectral image is represented in a 3D spectral cube, usually called hypercube, with 2D spatial information and 1D spectral information (i.e. absorbance at specific wavelengths). The most important HSI applications in the agro-food industry are in detecting defects,[8,9] discriminating botanical species, cultivars and quality classes,[10–12] determining fruit ripeness[7–13] and chemical composition,[14–16] and detecting and quantifying contaminants.[6,17–24]

Traditionally, hyperspectral instruments are classified into three groups, depending on the way the hypercube is generated: point-scan, line-scan, and plane-scan instruments. Point-scan instruments successively acquire a spectrum at a simple spatial location of the scene of interest. For this, a spectrometer (equipped with a single detector), coupled with a microscope equipped with an automatic sample stage, is used to perform the mapping in order to construct the hypercube. Line-scan instruments, through the use of a 2D focal plane array (FPA) detector, allow the spectra of several pixels from a line to be collected simultaneously. The hypercube is generated by moving the camera or samples horizontally in order to scan the whole scene. Plane-scan instruments allow the absorbance at continuous wavelengths of the scene of interest to be gathered successively. With this equipment, an FPA detector is also used, but a device is needed to select the absorbance at specific frequency, the hypercube being built through a step-by-step acquisition of the absorbance images at the successive wavelengths. In agro-food product analysis, line-scan instruments seem to be the best compromise for speed of analysis, flexibility, and possible use for on-line control.[10,25]

As noted earlier, NIR HSI is used mainly in qualitative studies and its use in quantification is still limited. However, there is an interest for the development of global tools where bulk calibration of classical quantitative constituents for quality control could be determined with the same apparatus used for the determination of properties that are easy or could only be detected with an imaging system mainly related to qualitative aspects as contaminant or fraud detection. In this direction, only in recent years, several studies have demonstrated the feasibility of using an HSI system for the quantitative prediction of the internal composition of agro-food products. Several studies have reported on the use of HSI to predict, inter alia, the content of water, fat, protein, and total saturated and total unsaturated fatty acids in red meat[26–29]; fat and protein content in cheese[30]; moisture and fat content in various species of fish[31]; soluble solids, moisture content, and acidity in fruits and vegetables[32–35]; and moisture, fat, starch, and oleic acid content in cereals.[14,36,37] Studies have also been conducted to predict the content of minor compounds such as anthocyanins and flavonols in grape skin[38] and grape seed,[39] respectively; total pigment in red meat[40]; synthetic astaxanthin coating and total volatile basic nitrogen in fish[41,42]; and total glucosinolate in freeze-dried broccoli.[43] Macronutrients in oilseed rape leaves[44] and wheat leaves[45] have been also determined using HSI. In addition to predicting internal composition, NIR HSI has been used for predicting water holding capacity, color, pH, tenderness, and microbiological attributes, such as total viable counts of bacteria, in red meat[46,47] and in fish,[48–52] as well as firmness in fruits and vegetables.[7,34]

Based on these references, HSI appears to be a powerful technique for the quantitative prediction of the composition of food products. However, it is not known how similar its performance is to any other classical NIR instrument. Knowing more about its comparative performance would enhance the potential of HSI for quantitative aspects until now mainly conducted with classical NIR spectrometers. New HSI instruments are now cheaper and faster, and the development of adequate chemometric tools and computer software will undoubtedly help to optimize the exploration of both spatial and spectral information for qualitative and/or quantitative determination. With the decreasing cost of HSI instruments, there is growing interest in finding out if they can be used in the same way as classical NIR spectrometers for determining quantitative parameters. So far as we know, there have been only a few studies comparing HSI with conventional NIR spectroscopy for quantitative prediction. They have focused on agro-food products usually considered a challenge for classical NIR techniques because of the heterogeneous physical nature of the sample (raw instead of ground material) and the heterogeneous spatial distribution of the parameters of interest. Burger and Geladi[30] compared an HSI instrument based on plane-scan technology with two NIR spectrophotometers (i.e. a scanning grating one with rotating sample holders and an FT-NIR one equipped with a fiber-optic sampling probe) for predicting the protein and fat content of cheese. By selecting a subset region of interest in the hyperspectral images and reducing it to a single average spectrum, the authors obtained almost identical results to those obtained with classical NIR instruments. The prediction error in terms of the average bias of the HSI instrument was 0.6% for protein and 0.2% for fat content compared with those obtained with the classical NIR

instruments. The scanning grating spectrometer showed the best results (0.2% for protein content and 0.5% for fat content) and the FT-NIR the worst results (3.0% for protein content and 0.7% for fat content). Xing et al.[53] compared a line-scan HSI system with an FT-NIR spectrophotometer for predicting alpha-amylase activity in individual wheat kernels. The performance of the HSI system when using the average spectra from the region of interest (i.e. the germ region, where the alpha-amylase is mainly located) was almost identical to the performance achieved with classical FT-NIR instruments. For the HSI system, the best model produced a coefficient of determination r^2 for the validation set of 0.82 and a root mean square error (RMSE) of 0.54. The predictive ability of the FT-NIR was lower; the best model was obtained with the original data (r^2 and RMSE were 0.71 and 11.76, respectively). Mendoza et al.,[54] however, compared the ability of an on-line hyperspectral scattering system and a short NIR spectrometer (USB400, with a range of 460–1100 nm) for predicting the firmness and soluble solids content (SSC) in three apple cultivars over two seasons. Based on SEP and residual predictive deviation (RPD) values, the NIR instrument was always better at predicting SSC, although the RPD values were < 2.5. In contrast, the hyperspectral scattering system was better at predicting firmness. Ignat et al.[55] showed that an HSI system based on electronically tuned band-pass filter (AOTF) performed poorly compared with a short NIR spectrometer (USB2000, with 350–1000 nm of spectral range) and a Liga SWIR spectrophotometer in predicting several quality parameters in intact bell peppers. The best prediction models were found for the SSC and the dry matter content. For SSC, the RPD values were 3.9, 3.3, and 2.6 for the short NIR, Liga SWIR and HSI systems, respectively; for the dry matter, the RPD values were 3.8, 3.0, and 2.4, respectively. Rady et al.[56] compared the potential of an NIR hyperspectral line-scan system with two classical NIR spectrometers (i.e. an NIR transmittance system and a visible/NIR interactance system using a fiber optic), inter alia, for predicting the glucose and sucrose content and the SSC of two potato cultivars. All the instruments performed well for glucose, particularly sliced samples. The visible/NIR interactance system showed the best prediction models in terms of the correlation coefficient r^2 and RPD ($\geq$0.90 and 2.14, respectively) for glucose. For the NIR hyperspectral system, the best model for glucose was an r^2 and RPD of 0.74 (r^2 of 0.55) and 1.49, respectively.

This study sought to assess the potential of an NIR hyperspectral line-scan system for predicting some important quality parameters of agro-food matrices and comparing its performance with that of two classical NIR instruments widely used in the agro-food industry and research laboratories. The overall objective was to determine if an HSI instrument is as efficient for predicting chemical composition as a classical NIR instrument. This is performed because there is an increasing need of developing global instruments, i.e. instruments for the simultaneous determination of different properties. On the basis of this consideration, it could be more worthy to calibrate protein on wheat kernels, for instance, where simultaneously the same analysis could reveal the presence of pests, contaminants, etc. The collection of spectra on a very fine (pixel-based) spatial level, the heterogeneity of the imaged substance as well as size and shape characteristics for intact objects can be obtained for defects/contaminants detection; while with the average of pixels of the scene or fraction of it to obtain a representative spectrum can be of utility to perform quantitative analysis.

In this study, the determination of the protein content of wheat flour, historically one of the most common uses of NIR spectroscopy in the agricultural and food industry,[57] was selected as an example. Wheat flour samples are easy to handle using classical NIR because the grinding process homogenizes the product and protein content is a parameter that is stable over time. Moreover, cereals products are usually studied with a HSI in order to detect their purity, homogeneity, and contamination among others.[58–60]

Materials and methods

Samples

A set of 79 wheat flour samples was used for this study. The samples were collected in 2013, with 57 coming from Belgian mills and 22 from the Bureau Interprofessionnel des Etudes Analytiques (Bipea) in France, an independent association that provides testing programs to laboratories.

Reference analysis

The Dumas combustion method[61,62] was used to determine the total nitrogen, and therefore the crude protein, content. The method involves the total combustion of samples under oxygen. Via subsequent oxidation and reduction tubes, nitrogen is quantitatively converted to N2. Results are given as % or mg nitrogen, which may be converted into protein by using conversion factors.[63] The reference value (% of protein content) for the Bipea samples was 11.06 ± 0.90, with a minimum of 9.50 and a maximum of 12.67. For the rest of the samples, the mean value was 11.10 ± 1.28 and varied between 9.00 and 14.40. The estimated uncertainty of the method for determination of total protein content is of 0.27.

Instrumentation and spectra acquisition

Two classical NIR instruments working in reflection mode, as well as a NIR hyperspectral line-scan camera were used in this study. Table 1 gives the main characteristics of the instruments used.

The first NIR instruments is a NIR System DS2500 (Foss—hereafter referred to as NIR-DS) that operates

Table 1. Characteristics of NIR instruments and NIR-HSI: total wavelength range, common wavelength range, spot size, and analysis time by sample.

Instrument name	Wavelength range (nm)	Common range (nm)	Spot size[a] (cm^2)	Analysis time[b] (s/sample)
NIR-DS	400–2498	1120–2424	1.8	66.0
NIR-Perstop	1100–2498	1120–2424	1.8	57.6
NIR-HSI	1118–2425	1120–2424	3.0	28.8

NIR: near infrared.
[a]For NIR-DS and NIR-Perstop, spot size is fixed at 15 cm in diameter; for NIR HSI, a total of 2601 spectra (51 pixels × 51 pixels) were selected from the center of each image.
[b]Determined from the total time required by instrument for the measurements of all samples (n = 79).

in reflection mode in the 400–2498 nm range, with a spectral resolution of 0.5 nm. It averages a certain number of spectra collected at different locations in a sample cup (rotating sampling device) during analysis. Spectra were collected using the ISIscan Nova software across the original wavelength range. Each spectrum was the average of 32 scans performed on the sample and it was acquired in 66 s. The sample area measured was about 2 cm^2.

The second NIR instruments is a 5000 Autocup DVP6BX instrument (Foss—hereafter referred to as NIR-Perstop) that operates in reflection mode in the 1100–2498 nm range, with a spectral resolution of 0.5 nm. It also has a rotating sampling device.

Spectra were collected using ISIscan software across the original wavelength range and each spectrum was the average of 32 scans performed on the sample. The sample area measured was about 2 cm^2, and 57 s was needed to analyze a cell. The difference of time of analysis (i.e. 9 s) between both NIR instruments (i.e. NIR-DS and NIR-Perstop) was due mainly to the autosampler device of the NIR-Perstop instrument.

The third instrument is a NIR hyperspectral line-scan camera (Burgermetrics—hereafter referred to as NIR-HSI) or push-broom imaging system was used to collect hyperspectral images. The NIR-HSI instrument was a SWIR XEVA CL 2.5 320 TE4 camera using a spectrograph. It has a cooled, temperature-stabilized Mercury–Cadmium–Telluride detector combined with a conveyor belt that presents the sample to the camera. There is more information on this instrument in Vermeulen et al.[23] and Dale et al.[10]

Spectra acquisition was made using HyperPro software (Burger-Metrics SIA, Riga, Latvia). All the images consisted of 420 lines of 320 pixels acquired at 209 wavelength channels (1128–2425 nm) and 32 scans per image. A 15 mm lens was used to analyze the total width of the plate. The lens was set up to cover 10 cm of the width of the conveyor belt. The conveyor belt speed was fixed at 3 mm/s (i.e. 20 lines/s) so as to produce clear images. In order to compensate for offset due to the dark current, the light source temperature drift, and the lack of spatial lighting uniformity, the spectral imaging system was calibrated with a dark image (by blocking the lens entrance) and a white image (background) collected from a standard white reference board (empty Teflon plate). The spectra were then automatically corrected.

After spectral acquisition, a total of 2601 spectra (51 × 51 pixels) corresponding to a sample area of ± 3 cm^2 were selected from the center of each image and averaged in order to obtain the mean spectra of each sample measured. In the configuration selected, the time required to analyze one sample was 28.8 s.

For each instrument, the 79 samples of wheat flour were measured in individual sealed cells with a glass cover, first with the NIR-DS instrument, then with the NIR-HSI instrument, and finally with the NIR-Perstop instrument. Sealed cells were used to prevent small moisture content changes in the samples and to ensure that the same portion of each sample was presented in the same condition to the different instruments. In order to study whether or not the measuring process in the line-scan system could be accelerated, after the NIR-Perstop measurements with sealed cells, spectra were also acquired in open cells and compared with spectra of the sealed cells. In all cases, the samples were measured randomly, although always in the same order for all the instruments. The NIR measurements were done in duplicate (i.e. the batch of samples was measured consecutively two times according to the protocol described before). The analysis of all the samples with the three instruments was simultaneously performed over three consecutive days.

Figure 1 shows the mean spectra of the 79 wheat flour samples obtained with each instrument (i.e. NIR-DS, NIR-Perstop, and NIR-HSI). For all instruments, typical reflectance spectra are obtained with the traditional bands and shape observed for ground wheat.[64] The different bands are related mainly to the different stretching vibrational bands assigned to the O–H, N–H, and C–H groups.[65]

Spectral data analysis

The spectral data were processed using Matlab software, v. 7.0 (The Mathworks, Inc, Natick, MA, USA). Chemometric tools were applied using the PLS Toolbox (v. 4.11, Eigenvector, Inc., Manson, WA, USA).

Principal component analysis (PCA) was used for exploratory analysis in order to extract the maximum amount of information from each set of data. Calibration models were developed using partial least squares (PLS) for predicting the protein content for the three instruments. These models were built using the full wavelength range, which varied from instrument to instrument (see Table 1) and also with the common range between them (reduced wavelength range), which corresponded to 1120–2424 nm. The number of PLS latent variables has been determined through the 'leave-one-out' cross-validation (LOOCV) procedure,

and then an external/independent validation has been applied. For LOOCV, one sample was left out and the multivariate models were constructed with the rest of the samples. The process was repeated until all the samples had been used once in the validation set. For independent validation, the samples were split into two groups using two strategies: a calibration model with 75% of samples randomly selected and validation with the remaining 25% (**Strategy 1**); and a calibration model with the first 60 samples in order of measurement, and validation with the remaining 19 samples (**Strategy 2**). The aim of Strategy 2 was to check stability over time, as mentioned earlier. In both strategies, the samples of the validation set were independent (i.e. they came from different sources, wheat cultivars, growing conditions, harvesting regimes, etc.).

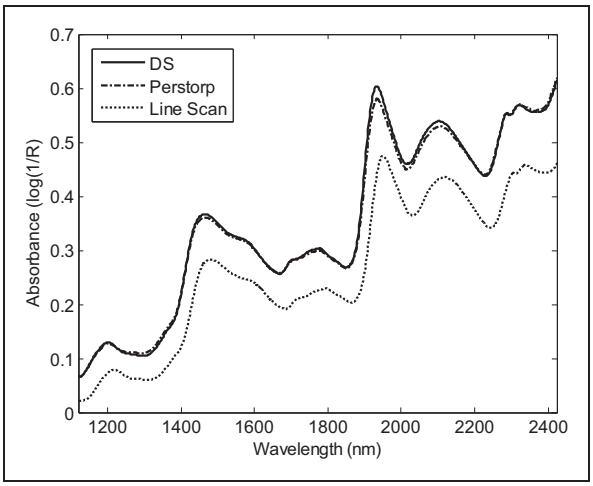

Figure 1. Mean spectrum of wheat flour obtained with the NIRS-DS, NIRS-Perstop and NIR-HSI instruments.

The quality of the constructed models was evaluated using the RMSE for calibration (RMSEC), cross-validation (RMSECV), and external validation (RMSEP). The coefficients for calibration and prediction (r^2) were determined, as well as the RPD for calibration and validation (RDPcal=SD/RMSEC and RDPval=SD/RMSEP, respectively; with SD = standard deviation). The best prediction models were those with higher values of r^2 and RPD and lower values of RMSEC and RMSECV. Values of RPD greater than 3.0 indicate excellent prediction accuracy.[66]

Results and discussion

The PCA score and loading plots of the sealed samples are shown in Figure 2. The first two PCs explained 99.73% of the total variation in the raw NIR spectral matrix. PC1 showed a clear separation between the spectra from all the instruments clearly visible in the loading plot and PC2 allowed the NIR spectra from the NIR-HSI instrument to be discriminated from those coming from the NIR-DS and NIR-Perstop instruments.

The Hotelling's T^2, a measure of the distance from the multivariate mean to the projection of each sample

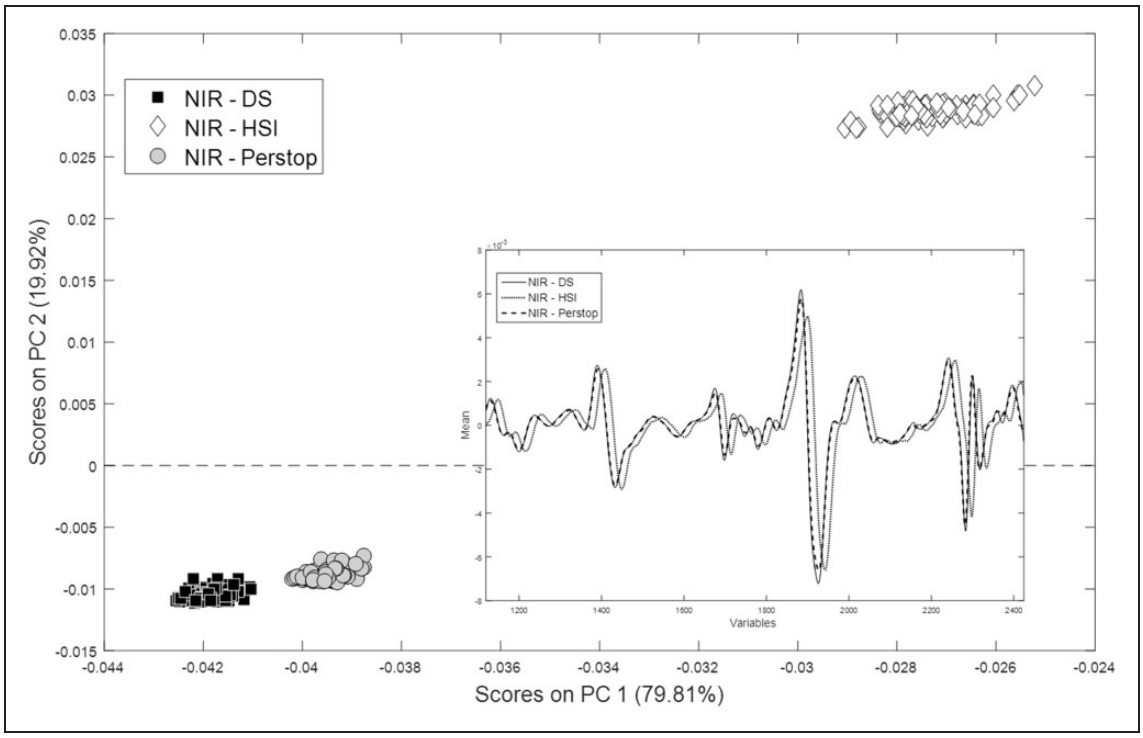

Figure 2. Principal Component Analysis (PC1 and PC2).

onto these two PCs, gave average values of 7.62, 6.87, and 15.38 for NIR-DS, NIR-Perstop, and NIR-HSI, respectively, indicating that the distance between the conventional NIR instruments was smaller, and so they are almost equivalent. These results were confirmed when a three-nearest-neighbors distance calculation was performed, with average values of 0.08, 0.05, and 0.179 for NIR-DS, NIR-Perstop, and NIR-HSI, respectively. This suggests that there was more variability in the raw data coming probably from HSI due to a large measurement surface and non-homogeneous light distribution.

Table 2 shows the results for the PLS models for the determination of protein content for all the instruments tested and using both the whole wavelength ranges and the common range. The most efficient preprocessing tool involved the use of standard normal variate followed by a second derivative Savitzky–Golay (window: 11, polynomial: 2). As shown in the table, good predictions with sealed cells were obtained when considering the total wavelength range with an r^2 of 0.99 for all the instruments. NIR-DS gave the best results in terms of RMSEP (0.14%) compared to NIR-Perstop (0.16%) and NIR-HSI (0.17%). The generally good performance of the NIR spectroscopy instruments (NIR-DS and NIR-Perstop) was not really a surprise, as NIR spectroscopy is used routinely in the cereal industry worldwide for quantifying protein content and other chemical components of wheat flour.[67] However, the similar performance of NIR-HSI demonstrates the potential of this technique for similar quantitative applications. Looking at the RPD values, which were higher than 3.0, the independent validations produced excellent predictions for protein, regardless of instrument.

When the common wavelength range (1120–2424 nm) was considered, Strategy 1 gave the best prediction model, with a similar r^2 (0.99) for all instruments. NIR-HSI showed the same RMSEP values (0.15%) as NIR-DS. NIR-Perstop showed the highest value (0.16%) and an almost identical performance for the total wavelength. The RPD values (8.08, 7.92, and 7.56 for NIR-DS, NIR-HSI, and NIR-Perstop, respectively) showed excellent predictions of protein content. These results demonstrate the potential of NIR-HSI compared with conventional NIR instruments for quantification, agreeing with earlier studies that showed even better prediction of protein and fat content in cheese by hyperspectral plane-scan imaging than FT-NIR,[30] better prediction of alpha-amylase activity in individual wheat kernels by hyperspectral line-scan imaging than FT,[53] and better prediction of firmness in apple fruits using an on-line hyperspectral scattering system than with using a short NIR spectrometer.[54]

The results differed, however, with regard to the outputs of Strategy 2 (i.e. the first measured 60 samples of the wheat flour set were used for calibration and the remaining 19 samples for validation). Here, the performances of the NIR-DS and NIR-Perstop

Table 2. PLS models for the total and common wavelength range.

	Total wavelength range								Common wavelength range Strategy 1[a]								Common wavelength range Strategy 2[a]							
	LV	R^2	RMSEC (%)	RMSECV (%)	RPDcal	r^2	RMSEP (%)	RPDval	LV	R^2	RMSEC (%)	RMSECV (%)	RPDcal	r^2	RMSEP (%)	RPDval	LV	R^2	RMSEC (%)	RMSECV (%)	RPDcal	r^2	RMSEP (%)	RPDval
NIR-DS	8	0.96	0.15	0.22	5.27	0.99	0.14	8.55	6	0.97	0.16	0.20	5.84	0.99	0.15	8.08	6	0.97	0.15	0.20	5.87	0.98	0.16	7.56
NIR-Perstop	6	0.95	0.20	0.25	4.72	0.99	0.16	7.56	6	0.95	0.20	0.24	4.86	0.99	0.16	7.56	5	0.96	0.21	0.25	4.65	0.98	0.14	8.68
NIR-HSI[b]	7	0.97	0.15	0.19	6.18	0.99	0.17	7.07	7	0.96	0.14	0.23	5.09	0.99	0.15	7.92	7	0.98	0.11	0.20	5.96	0.94	0.23	5.15

LV: Latent Variables; PLS: partial least squares; RMSEC: RMSE for calibration; RMSECV: RMSE for cross-validation; RMSEP: Root Mean Square Error in Prediction for an external set; RPD: residual predictive deviation.
[a]Strategy 1: Calibration model with 75% of samples randomly selected and validation with the remaining 25%; Strategy 2: Calibration model with the first 60 samples according to the order of measurement and validation with the remaining 19 samples.
[b]Samples were measured in sealed cells on the NIR-DS instrument, followed by the NIR-HSI instrument, and finally by the NIR-Perstop instrument.

instruments were similar, but worse than in Strategy 1. For both instruments, the r^2 was 0.98, although the RMSEP was higher with NIR-DS (0.16%) than with NIR-Perstop (0.14%). For NIR-HSI, the r^2 was 0.94, the RMSEP 0.23, and the RPD fell by 35% compared with Strategy 1. Although the mean RPD value was higher than 3.0, denoting excellent prediction accuracy, there was an important loss in prediction ability. The results of Strategy 2 therefore indicated a decline in the stability of the NIR-HSI instrument over time, which is undesirable for routine analysis at industrial or laboratory level. This could be easily solved by calibrating the instrument several times a day and before each analysis of a set of samples.[68]

Figure 3 shows the prediction of the calibration and validation sets for the total wavelength range using Strategy 1 and the prediction of the validation set for the reduced/common wavelength range using Strategies 1 and 2.

In addition to the study on sealed cells, a tentative HSI study was done on open cells (not shown in table). The prediction performance for the total wavelength range in open cells (r^2 of 0.98 and RMSEP of 0.20%) was almost identical to that obtained with the HSI instrument in sealed cells (r^2 of 0.99 and RMSEP of 0.17%). The RPD value fell by 17% for the sealed cells, indicating excellent accuracy for protein content prediction. Given that the time needed for analyzing open cells fell by about 50% with the NIR-HSI instrument compared to the other instruments (Table 1), this is an interesting factor to take into account. The use of open cells can therefore increase the speed of measurement with NIR-HSI and could also be accelerated by simultaneous measurement of two cells; hence the time required would fall by up to about 14 s/measurement.

Conclusion

To date, the use of HSI has been based mainly on its capacity for discrimination in qualitative studies, such as those focusing on fraud or contaminant detection. In our study, the comparison with classical NIR instrumentation demonstrated the potential of NIR-HSI for quantifying the chemical composition of the samples.

In particular, the study showed that when a common wavelength range is used for all the instruments and the same sample sets, a hyperspectral line-scan system worked as well as a classical NIR spectrometer and

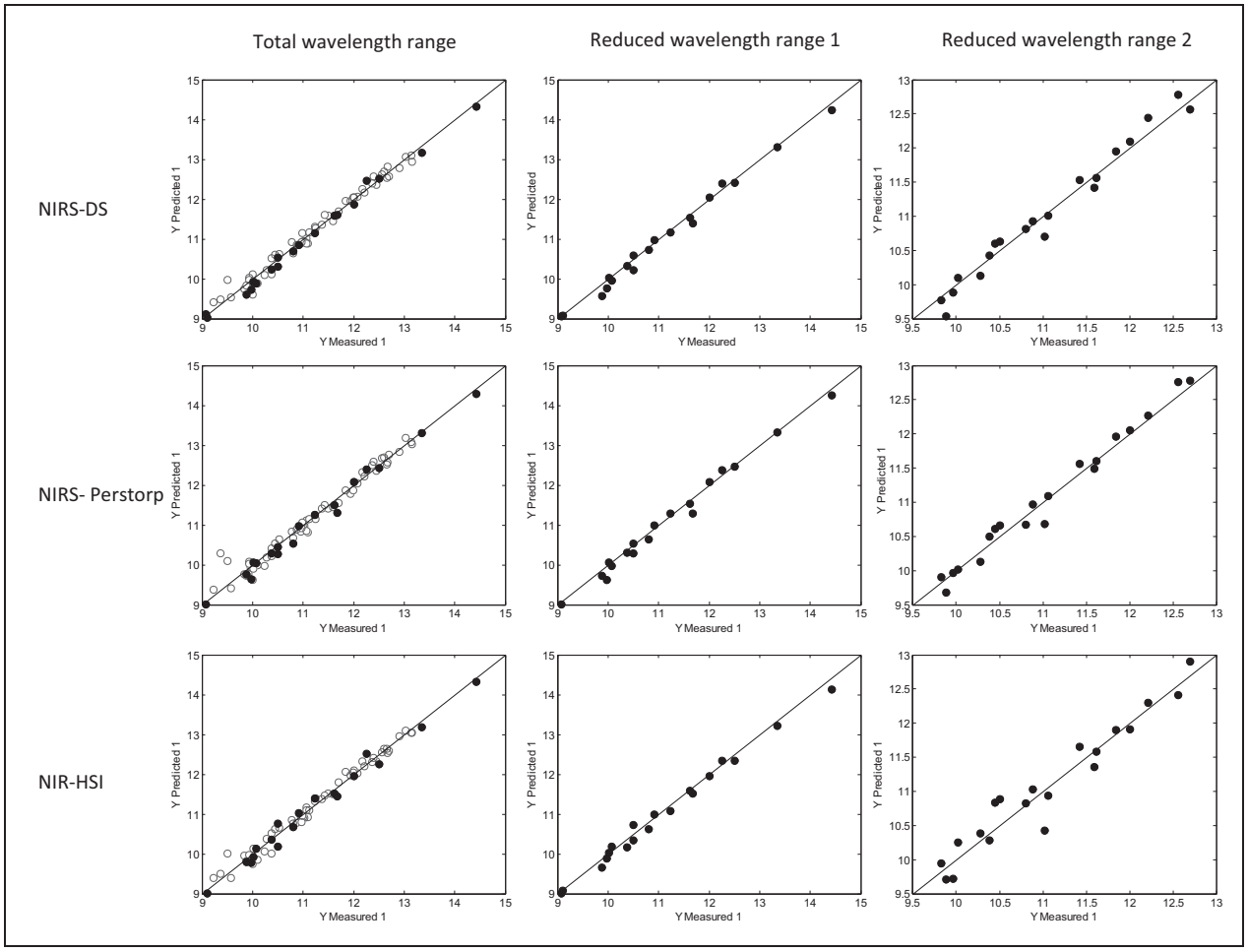

Figure 3. NIR predicted data versus reference data for protein content, for total wavelength range (strategy 1) and common wavelength range (strategy 1 and 2).

the time required for analyzing a sample decrease by at least half. In addition, it showed that using open cells increased the number of samples that could be analyzed over time, paving the way to the development of a complete methodology for on-line analysis.

The study demonstrated, then, the feasibility of using NIR-HSI for predicting the protein content of wheat flour and highlighted the potential of this technology for the cereal industry, given its already demonstrated ability for sorting wheat into different classes according to growing conditions, visible or internal defects, and contaminants.

The main drawback of this type of system is (i) the loss of stability over time, which affects the performance of hyperspectral systems in quantitative studies, but this can be solved by carrying out the appropriate calibration strategy; and (ii) its price, which is probably the main limitation of using this technique for quantification. However, this could be solved as looking at this technique as a multitool instrument, i.e. a system that can be used as a classical NIR spectrometer for quantitative quality estimations, as proved in this study, and at the same time an instrument that gives a fast and reliable solution for the qualitative detection of abnormal ingredients or products, often demonstrated in the scientific literature.

Acknowledgment

We gratefully acknowledge the technical assistance provided by Anne Mouteau, Sandrine Mauro, and Nicaise Kayoka from the CRA-W.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by the European Commission through the 7th Framework Project: Ensuring the Integrity of the European food chain (Food Integrity)—613688: Collaborative Project under the Seventh Framework Programme, KBBE.2013.2.4-01: Assuring quality and authenticity in the food chain.

References

1. Panford JA. Application of near-infrared reflectance spectroscopy in North America. In: Williams P and Norris K (eds) *Near-infrared technology in the agricultural and food industries*. St. Paul, MN: American Association of Cereal Chemists, Inc, 1987, pp.201–211.
2. Mannina L and Di Tullio V. *Food analysis by fingerprint techniques*. University of Molisa, TRACE publications, 2009.
3. Abbas O, Dardenne P and Baeten V. Near-infrared, mid-infrared, and Raman spectroscopy. In: Pico Y (ed) *Chemical analysis of food: techniques and applications*. Burlington, MA: Elsevier Science, 2012, pp.59–91.
4. Baeten V, Rogez H, Fernández Pierna JA, et al. Vibrational spectroscopy methods for the rapid control of agro-food products. In: Toldra F. and Nollet L.M.L. (eds) *Handbook of food analysis*. 3rd ed. 2015, pp.591–614..
5. Baeten V, Fernández Pierna JA and Dardenne P. Hyperspectral imaging techniques: an attractive solution for the analysis of biological and agricultural materials. In: Grahn HF and Geladi P (eds) *Techniques and applications of hyperspectral image analysis*. John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, England, 2007, pp.289–312.
6. Fernández Pierna JA, Vermeulen P, Amand O, et al. NIR hyperspectral imaging spectroscopy and chemometrics for the detection of undesirable substances in food and feed. *Chemometr Intell Lab* 2012; 117: 233–239.
7. Rajkumar P, Wang N, ElMasry G, et al. Studies on banana fruit quality and maturity stages using hyperspectral imaging. *J Food Eng* 2012; 108: 194–200.
8. Ariana DP and Lu R. Evaluation of internal defect and surface color of whole pickles using hyperspectral imaging. *J Food Eng* 2010; 96: 583–590.
9. Nicolai BM, Lotze E, Peirs A, et al. Non-destructive measurement of bitter pit in apple fruit using NIR hyperspectral imaging. *Postharvest Biol Technol* 2006; 40: 1–6.
10. Dale LM, Thewis A, Boudry C, et al. Hyperspectral imaging applications in agriculture and agro-food product quality and safety control: a review. *Appl Spectrosc Rev* 2013; 48: 142–159.
11. Diago MP, Fernandes AM, Millan B, et al. Identification of grapevine varieties using leaf spectroscopy and partial least squares. *Comput Electron Agric* 2013; 99: 7–13.
12. Moscettu R, Saeys W, Keresztes JC, et al. Hazelnut quality sorting using high dynamic range short-wave infrared hyperspectral imaging. *Food Bioprocess Technol* 2015; 8: 1593–1604.
13. Wei X, Liu F, Qiu Z, et al. Ripeness classification of astringent persimmon using hyperspectral imaging technique. *Food Bioprocess Technol* 2014; 7: 1371–1380.
14. Baeten V, Fernández Pierna JA, Vermeulen P, et al. NIR hyperspectral imaging methods for quality and safety control of food and feed products: contributions to 4 European projects. *NIR News* 2010; 21: 10–13.
15. Talens P, Mora L, Morsy N, et al. Prediction of water and protein contents and quality classification of Spanish cooked ham using NIR hyperspectral imaging. *J Food Eng* 2013; 117: 272–280.
16. Whitworth MB, Millar SJ and Chau A. Food quality assessment by NIR hyperspectral imaging. *Proc SPIE* 1987; 7676: 5–12.
17. Baeten V, Vermeulen P, Fernández Pierna JA, et al. From targeted to untargeted detection of contaminants and foreign bodies in food and feed using NIR spectroscopy. *New Food* 2014; 17: 2–9.
18. Fernández Pierna JA, Baeten V, Michotte Renier A, et al. 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.
19. Fernández Pierna JA, Baeten V and Dardenne P. Screening of compound feeds using NIR hyperspectral data. *Chemometr Intell Lab* 2006; 84: 114–118.
20. Fernández Pierna JA, Dardenne P and Baeten V. In-house validation of a near infrared hyperspectral imaging

method for detecting processed animal proteins (PAP) in compound feed. *J Near Infrared Spectrosc* 2010; 18: 121–133.
21. Fernández Pierna JA, Vincke D, Dardenne P, et al. Line scan hyperspectral imaging spectroscopy for the early detection of melamine and cyanuric acid in feed. *J Near Infrared Spectrosc* 2014; 22: 103–112.
22. Ravikanth L, Singh CB, Jayas DS, et al. Classification of contaminants from wheat using near-infrared hyperspectral imaging. *Biosyst Eng* 2015; 135: 73–86.
23. Vermeulen P, Fernández Pierna JA, van Egmond HP, et al. Online detection and quantification of ergot bodies in cereals using near infrared hyperspectral imaging. *Food Addit Contam* 2012; 29: 232–240.
24. Vermeulen P, Fernández Pierna JA, van Egmond HP, et al. Validation and transferability study of a method based on near-infrared hyperspectral imaging for the detection and quantification of ergot bodies in cereals. *Anal Bioanal Chem* 2013; 405: 7765–7772.
25. Fernández Pierna JA, Baeten V, Dubois J, et al. NIR imaging – theory and applications. In: Brown S and Walczak B (eds) *Comprehensive chemometrics*. Vol. 4. Oxford: Elsevier, 2009, pp.173–196.
26. Barbin DF, ElMasry G, Sun DW, et al. Non-destructive determination of chemical composition in intact and minced pork using near-infrared hyperspectral imaging. *Food Chem* 2013; 138: 1162–1171.
27. ElMasry G, Sun DW and Allen P. Chemical-free assessment and mapping of major constituents in beef using hyperspectral imaging. *J Food Eng* 2013; 117: 235–246.
28. Kobayashi KI, Matsui Y, Maebuchi Y, et al. Near infrared spectroscopy and hyperspectral imaging for prediction and visualisation of fat and fatty acid content in intact raw beef cuts. *J Near Infrared Spectrosc* 2010; 18: 301–315.
29. Kamruzzaman M, ElMasry G, Sun DW, et al. 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.
30. Burger J and Geladi P. Hyperspectral NIR imaging for calibration and prediction: a comparison between image and spectrometer data for studying organic and biological samples. *Analyst* 2006; 131: 1152–1160.
31. ElMasry G and World JP. High-speed assessment of fat and water content distribution in fish fillets using online imaging spectroscopy. *J Agric Food Chem* 2008; 56: 7672–7677.
32. Baiano A, Terracone C, Peri G, et al. Application of hyperspectral imaging for prediction of physico-chemical and sensory characteristics of table grapes. *Comput Electron Agric* 2012; 87: 142–151.
33. ElMasry G, Wang N, El Sayed A, et al. Hyperspectral imaging for non-destructive determination of some quality attributes for strawberry. *J Food Eng* 2007; 81: 98–107.
34. Mendoza F, Lu R, Ariana D, et al. Integrated spectral and image analysis of hyperspectral scattering data for prediction of apple fruit firmness and soluble solids content. *Postharvest Biol Technol* 2011; 62: 149–160.
35. Rady AM, Guyer DE and Lu R. Evaluation of sugar content of potatoes using hyperspectral imaging. *Food Bioprocess Technol* 2015; 8: 995–1010.
36. Mahesh S, Jayas DS, Paliwal J, et al. Comparison of partial least squares regression (PLSR) and principal components regression (PCR) methods for protein and hardness predictions using the near-infrared (NIR) hyperspectral images of bulk samples of Canadian wheat. *Food Bioprocess Technol* 2015; 8: 31–40.
37. Weinstock BA, Janni J, Hagen L, et al. 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: 9–16.
38. Fernandes AM, Oliveira P, Moura JP, et al. Determination of anthocyanin concentration in whole grape skins using hyperspectral imaging and adaptive boosting neural networks. *J Food Eng* 2011; 105: 216–226.
39. Rodríguez-Pulido FJ, Hernández-Hierro JM, Nogales-Bueno J, et al. A novel method for evaluating flavanols in grape seeds by near infrared hyperspectral imaging. *Talanta* 2014; 122: 145–150.
40. Xiong Z, Sun DW, Xie A, et al. Quantitative determination of total pigments in red meats using hyperspectral imaging and multivariate analysis. *Food Chem* 2015; 178: 339–345.
41. Cheng J, Sun DW and Pu HB. Non-destructive and rapid determination of TVB-N content for freshness evaluation of grass carp (*Ctenopharyngodon idella*) by hyperspectral imaging. *Innov Food Sci Emerg Technol* 2014; 21: 179–187.
42. Ljungqvist MG, Ersboll BK, Kobayashi KI, et al. Near-infrared hyperspectral image analysis of astaxanthin concentration in fish feed coating. In: *IEEE international conference on Imaging Systems and Techniques* (IST), Manchester, UK, 2012, pp.136–141. Piscataway, USA: IEEE Service Center.
43. Hernández-Hierro JM, Esquerre C, Valverde J, et al. Preliminary study on the use of near infrared hyperspectral imaging for quantitation and localisation of total glucosinolates in freeze-dried broccoli. *J Food Eng* 2014; 126: 107–112.
44. Zhang X, Liu F, He Y, et al. Detecting macronutrients content and distribution in oilseed rape leaves based on hyperspectral imaging. *Biosyst Eng* 2013; 115: 56–65.
45. Vigneau N, Ecarnot M, Rabatel G, et al. Potential of field hyperspectral imaging as a non destructive method to assess leaf nitrogen content in wheat. *Field Crop Res* 2011; 22: 25–31.
46. ElMasry G, Sun DW and Allen P. Near-infrared hyperspectral imaging for predicting colour, pH and tenderness of fresh beef. *J Food Eng* 2012; 110: 127–140.
47. Huang L, Zhao J, Chen O, et al. Rapid detection of total viable count (TVC) in pork meat by hyperspectral imaging. *Food Res Int* 2013; 54: 821–828.
48. Cheng JH and Sun DW. Rapid and non-invasive detection of fish microbial spoilage by visible and near infrared hyperspectral imaging and multivariate analysis. LWT – Food Sci Technol 2015; 62: 1060–1068.
49. He HJ, Wu D and Sun DW. Rapid and non-destructive determination of drip loss and pH distribution in farmed Atlantic Salmon (*Salmo salar*) fillets using visible and near-infrared (Vis-NIR) hyperspectral imaging. *Food Chem* 2014; 156: 394–401.
50. Wu D and Sun DW. Application of visible and near infrared hyperspectral imaging for non-invasively measuring distribution of water-holding capacity in salmon flesh. *Talanta* 2013; 116: 266–276.

51. Wu D and Sun DW. Novel non-invasive distribution measurement of texture profile analysis (TPA) in salmon fillet by using visible and near infrared hyperspectral imaging. *Food Chem* 2014; 145: 417–426.
52. Wu D, Sun DW and He Y. Application of long-wave near infrared hyperspectral imaging for measurement of color distribution in salmon fillet. *Innov Food Sci Emerg Technol* 2012; 16: 361–372.
53. Xing J, Symons S, Hatcher D, et al. Comparison of short-wavelength infrared (SWIR) hyperspectral imaging system with an FT-NIR spectrophotometer for predicting alpha-amylase activities in individual Canadian Western Red Spring (CWRS) wheat kernels. *Biosyst Eng* 2011; 108: 303–310.
54. Mendoza F, Lu R and Cen H. Data fusion of visible/near infrared spectroscopy and spectral scattering for quality assessment. ASABE paper 1111244, 2011.
55. Ignat T, Alchanatis V and Schmilovitch Z. Maturity prediction of intact bell peppers by sensor fusion. *Comput Electron Agric* 2014; 104: 9–17.
56. Rady AM, Guyer DE, Kirk W, et al. The potential use of visible/near infrared spectroscopy and hyperspectral imaging to predict processing-related constituents of potatoes. *J Food Eng* 2014; 135: 11–25.
57. Osborne BG, Fearn T and Hindle PH. *Practical NIR spectroscopy with applications in food and beverage analysis*. 2nd ed. Harlow: Longman Scientific and Technical, 1993.
58. Manley M. Near-infrared spectroscopy and hyperspectral imaging: non-destructive analysis of biological materials. *Chem Soc Rev* 2014; 43: 8200–8214.
59. Mishra P, Herrero-Langreo A, Barreiro P, et al. Detection and quantification of peanut traces in wheat flour by near infrared hyperspectral imaging spectroscopy using principal-component analysis. *J Near Infrared Spectrosc* 2015; 23: 15–22.
60. Scheibelhofer O, Koller DM, Kerschhaggl P, et al. Continuous powder flow monitoring via near-infrared hyperspectral imaging. In: *IEEE instrumentation and measurement technology conference,* May 2012.
61. Buckee GK. Determination of total nitrogen in barley, malt and beer by Kjeldahl procedures and the Dumas combustion method. *Collaborative trial*. J Inst Brew 1994; 100: 57–64.
62. Dumas JBA. Procedes de l'Analyse Organique. *Ann Chim Phys* 1831; 247: 198–213.
63. Müller J. *Dumas or Kjeldahl for reference analysis? A white paper from FOSS*, FOSS, Hilleroed, Denmark, 2014.
64. Williams P and Norris K. Near infrared technology in the agricultural and food industries. In: Williams P (ed) *Near-infrared technology in the agricultural and food industries*. St Paul, MN: American Association of Cereal Chemists, Inc, pp.35–55.
65. Sun DW. *Infrared spectroscopy for food quality analysis and control*. Elsevier's Science & Technology Rights Department, Oxford, UK.
66. Nicolai BM, Beullens K, Bobelyn E, et al. Non-destructive measurement of fruit and vegetable quality by means of NIR spectroscopy: a review. *Postharvest Biol Technol* 2007; 46: 99–118.
67. Fox G and Manley M. Applications of single kernel conventional and hyperspectral imaging near infrared spectroscopy in cereals. *J Sci Food Agric* 2014; 94: 174–179.
68. Geladi P, Burger J and Lestander T. Hyperspectral imaging: calibration problems and solutions. *Chemometr Intell Lab* 2004; 72: 209–217.