JOURNAL OF CHEMOMETRICS
J. Chemometrics 2005; **19**: 355–363
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/cem.938

Hyperspectral NIR image regression part I: Calibration and correction

James Burger* and Paul Geladi

Unit of Biomass Technology and Chemistry, Swedish University of Agricultural Sciences, SLU Röbäcksdalen, P.O. Box 4097, SE 90403, Umeå, Sweden

Received 1 January 2005; Revised 27 February 2005; Accepted 24 July 2005

Hyperspectral imaging instruments produce large amounts of raw data. These raw data in A/D converter counts have a number of errors that can be corrected by calibration. The use of multiple Spectralon calibration standards is shown to correct for both spectral and spatial variations. Optimal results are achieved using a two-step calibration and correction process. A series of full field of view or external calibration standards is used to transform raw data counts to reflectance values. A grayscale series of internal standards embedded within each hyperspectral image is used to compensate for instrument instability. Second-order regression models based on these multiple standards provide maximum accuracy. The external standards allow for standardization within a hyperspectral image. The internal standards permit instrument standardization or calibration transfer between hyperspectral images. Copyright © 2006 John Wiley & Sons, Ltd.

KEYWORDS: hyperspectral images; multivariate image analysis; instrument standardization; calibration transfer; internal standards; reflectance calibration; multivariate calibration

1. INTRODUCTION

Laboratory cameras are nowadays capable of producing high-resolution digital images using ultraviolet, visible, near infrared, infrared, and Raman spectral information [1–4]. Individual images measured with red, green, and blue visible light filters are routinely combined to create 'color' digital images. Sets of images measured at different wavelength bands can be stacked or combined to form hyperspectral images or hypercubes. Using more advanced hardware and software wavelength filtering techniques in combination with broadband imaging detectors, it is now possible to measure tens or hundreds of individual grayscale images or channels with much greater spectral resolution [5–7]. The resulting hyperspectral images can be thought of effectively as collections of tens of thousands of spectra with additional two-dimensional spatial information. Such hyperspectral images have a high potential for describing concentration and property distributions in inhomogeneous solids, semisolids, and liquids. Spectroscopic multivariate calibration modeling techniques can be extended to hyperspectral images, permitting qualitative and quantitative analysis. But accurate results require accurate spectra. The objective of this paper is the presentation and discussion of the use of proper external and internal calibration standards in hyperspectral images. Results are based on recalibration of known reference standards. Previous work discussed the use of external standards only [8].

Much work has been published regarding the standardization of spectroscopic instruments and the transfer of multivariate calibration models [9–14]. Three situations can occur which invalidate calibration models: (1) physical or chemical changes in samples such as texture or particle size, (2) instrument changes due to inherent instabilities, aging parts or the use of multiple instruments, or (3) environment changes such as temperature or humidity [13]. The resulting spectral changes may be observed in either the horizontal axis as wavelength shifts or peak broadening, in the vertical axis as baseline offsets or absorbance differences.

Calibration model transfer of hyperspectral images is subject to all the error contributions of conventional one-dimensional spectroscopy (noise, drift, non-linear response of detectors, wavelength-dependent errors) as well as the two-dimensional or spatial error components associated with camera devices and illumination (readout errors, inconsistent detector responses, quantization errors, and non-uniform lighting).

Spatial and spectral standards are very common in photography and television/video work. Gray cards, grayscale wedges, color scales, and other test images can be used to test correctness of the obtained images for different illuminations, film types, or camera settings. Other test standards for resolution, geometry etc. are also available. But similar standard materials are not easily found for the NIR region. One source is a calibration surface made of Spectralon [15]

*Correspondence to: J. Burger, Unit of Biomass Technology and Chemistry, Swedish University of Agricultural Sciences, SLU Röbäcksdalen, P.O. Box 4097, SE 90403, Umeå, Sweden.
E-mail: jim.burger@btk.slu.se

available with different levels of reflectance with calibration values ranging from 250 to 2500 nm. These UV-VIS-NIR reflectance standards are intended for checking and calibrating spectrometers with a typical reflection geometry, but in this paper they are tested as image reflectance standards.

The advent of the increased dimensionality of hyperspectral images has provoked a parallel development in chemometrics tools. PCA has been extended to multivariate image analysis (MIA) [16,17]. Likewise, regression techniques such as OLS, PCR, and PLS have been incorporated into MIR [18–20]. In MIA a common technique is to use PCA of images reorganized to matrices. This allows the use of loadings to interpret the importance of the wavelengths and the use of score images and score plots in conjunction with segmentation to detect concentrations or property regions or gradients. Alternatives to PCA such as factor analysis or curve resolution can also be used. In MIR, the hypercube is regressed against an image of external information, for example concentrations in a calibration step. The goal is to be able to use a hypercube to predict concentrations once the regression model is established.

The accuracy of any modeling or prediction method is limited by the quality of the data used. MIA and MIR can be applied to the raw data given in A/D converter counts, but these counts are not in SI or IUPAC spectroscopic units (reflectance or absorbance), nor are they very reproducible. This results in an under usage of the full potential of MIA and MIR. It is necessary to make the hypercubes consistent both internally (within a hypercube) and also externally (between hypercubes).

Preprocessing steps must be taken to clean or prepare the data for further analysis. Spectral data are routinely transformed to reflectance using a simple two-point transfer function based on dark current and maximum reflectance value measurements. The use of additional calibration standards can improve measurement accuracy and permits linear and non-linear corrections to be made. The current work demonstrates the use of both external and internal calibration standards. External standards are used to correct for pixel-to-pixel variances due to camera inconsistencies and variation in sample illumination. Internal standards are used to compensate for signal drift over time due to changes in power or temperature effects.

2. THEORETICAL BACKGROUND

2.1. Reflectance transforms

Spectral data collected from a CCD device represent detector signal intensity counts not actual reflectance values. Figure 1 shows the typical raw data spectra for a series of NIR materials, which should exhibit flat and smooth uniform spectra. The apparent structure and variation in signal intensity of these spectra is due to the wavelength dependencies of both the sample illumination source and the sensitivity of the detector used. It is generally more useful to correct or transform raw signal data into reflectance or absorbance units by comparing with spectra of standard reference materials. The usual transformation to reflectance values is obtained by correcting sample spectra (C) for detector dark current (Dk) and dividing by a similarly corrected total reflectance spectrum (Ref_{100}) [2,6].

$$R = (C - Dk)(Ref_{100} - Dk)^{-1} \quad (1)$$

Equation (1) transforms instrument measurement values in A/D counts (C, Dk, and Ref_{100}) into unitless reflectance values. This transform is often termed a one-point calibration since the transformation is based on a single reference standard, Ref_{100}. Mathematically, however, this is a two-point

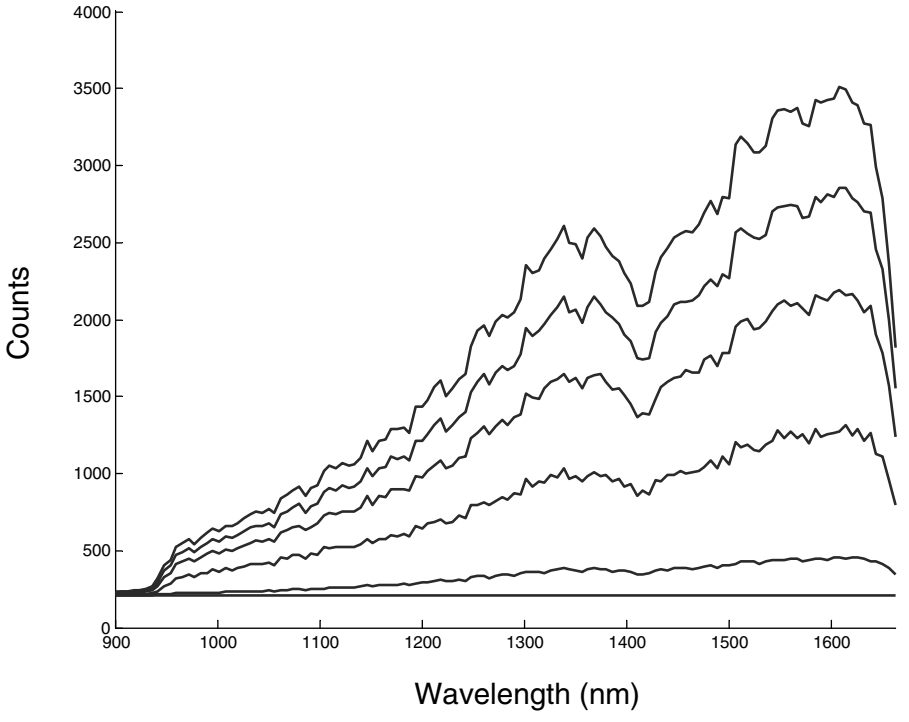

Figure 1. Average spectra (raw counts) of 99 (top curve), 75, 50, 25, 2% Spectralon and dark current (baseline).

linear approximation model with measured endpoint values at 0 and 100% reflectance and for the purpose of this discussion will be termed a *simple* model. Equation (1) can be rearranged and written in a linear bias offset and slope form:

$$R = -\mathrm{Dk}(\mathrm{Ref}_{100} - \mathrm{Dk})^{-1} + (\mathrm{Ref}_{100} - \mathrm{Dk})^{-1}\mathrm{C} \quad (2)$$

Or in general terms:

$$R = b_0 + b_1\mathrm{C} \quad (3)$$

This *simple* transformation model requires only two calibration spectra: one of the total reflectance standard and a dark current spectrum measured by blocking the camera lens. Such a model covers the complete range of reflectance values but gives no indication of system non-linearity.

Optimal transform models are obtained when more than two standards can be measured which span the reflectance or absorbance space of the expected samples. When *true* calibration spectra are available for a collection of standards, an alternative modeling approach is available: fitting least squares regression models to the true versus measured values for the set of all standards. In this paper we present two such regressions: a first order *linear* model with coefficients as in Equation (3) and a second-order quadratic or *quad* model, Equation (4).

$$R = b_0 + b_1\mathrm{C} + b_2\mathrm{C}^2 \quad (4)$$

Although the form of the equations for the *simple* and *linear* transformation models is the same, the $\boldsymbol{b_i}$ coefficients are computed differently. They are based on a single standard and dark measurement for the *simple* model, and more than two standards for the *linear* model. Since the measured response of the system is highly wavelength dependent, individual *simple*, *linear*, and *quad* models must be computed at each wavelength. For the transformation of multiple wavelength spectra, the scalar model coefficients b_i in Equations (3) and (4) are therefore reported as vectors $\boldsymbol{b_i}$ as a function of wavelength. It is also important to note that no dark subtraction is performed for the linear and quad regression models. The dark offset is included as part of the $\boldsymbol{b_0}$ model constant.

2.2. Spatial modeling—external standards

Hyperspectral images add spatial complexity to the spectral transformation problem. Regardless of whether the hypercubes are acquired from a single detector, or a one- or two-dimensional array of detectors, spatial changes in sample illumination or differences in multiple detector sensitivities require that the wavelength-dependent reflectance transform models must be computed for each spatial or pixel location. Two approaches will be considered, which we term *global calibration* and *pixelwise calibration*. Both techniques are spatial extensions to the one-dimensional reflectance transform problem, and necessitate accumulation of complete hypercubes from a series of calibrated reference materials. These hypercubes are measured independently of subsequent sample images and are therefore termed *external standards*.

The *global calibration model* assumes only wavelength dependencies and essentially ignores all spatial dependencies, assuming uniform sensitivities throughout each wavelength channel of the hypercube. This approach is based on the median spectrum of each of the individual calibration standard hypercubes. Figure 2 shows the computation steps necessary to create each model. For each of the calibration standards, a single median spectrum is first computed from the collection of spectra of every pixel in the respective hypercube. This set of median spectra is then used to build the global calibration model at each wavelength. As was described previously, the *simple* transform model is based only on total reflectance and dark hypercubes, while the *linear* and *quad* regression models are based on sets of calibrated reference standards. Future hypercubes are then transformed to reflectance, pixel by pixel, using the wavelength-dependent global model $\boldsymbol{b_i}$ vectors.

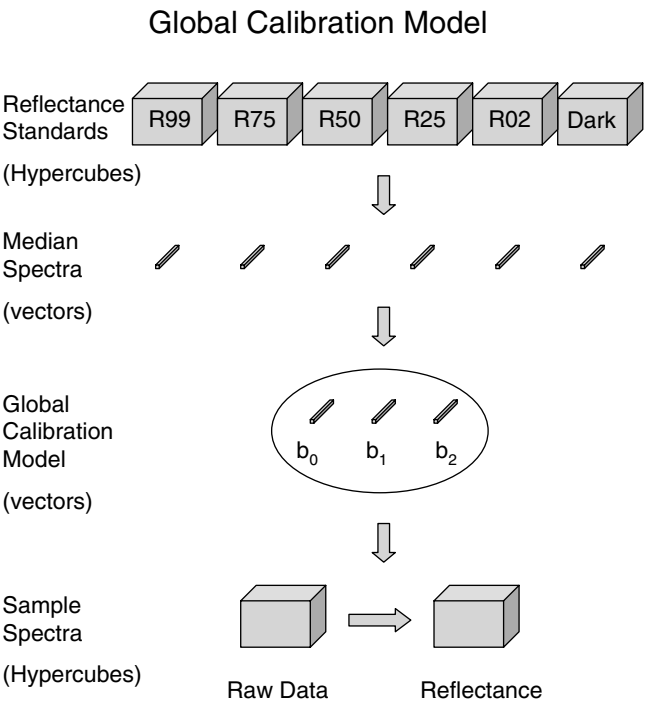

Figure 2. The median spectrum from each of the external standard hypercubes is used to compute a global calibration model one wavelength at a time. The resulting model vectors (individual b_0 and b_1 of Equation (3) and b_0, b_1, and b_2 of Equation (4) at each wavelength) are then used to transform sample hypercubes from instrument counts (raw data) to reflectance at each wavelength.

Independent reflectance transform models can also be created for each individual pixel in the hypercube. This compensates for illumination and detector inhomogeneities. Rather than using the median spectra of the global image, position-specific transform models are computed one pixel at a time from the same pixel location in each of the external standard hypercube images. Figure 3 represents a flowchart for the computation steps of this model process. The complete set of individual regression coefficient vectors computed at each pixel effectively form hypercubes of regression coefficients B_i, in contrast to the global regression vectors $\boldsymbol{b_i}$ in Figure 2. This spatial specific reflectance transform model is termed a *pixelwise calibration model*.

2.3. Internal standards

The use of external calibration standards as discussed above permits the creation of spatial and wavelength-specific

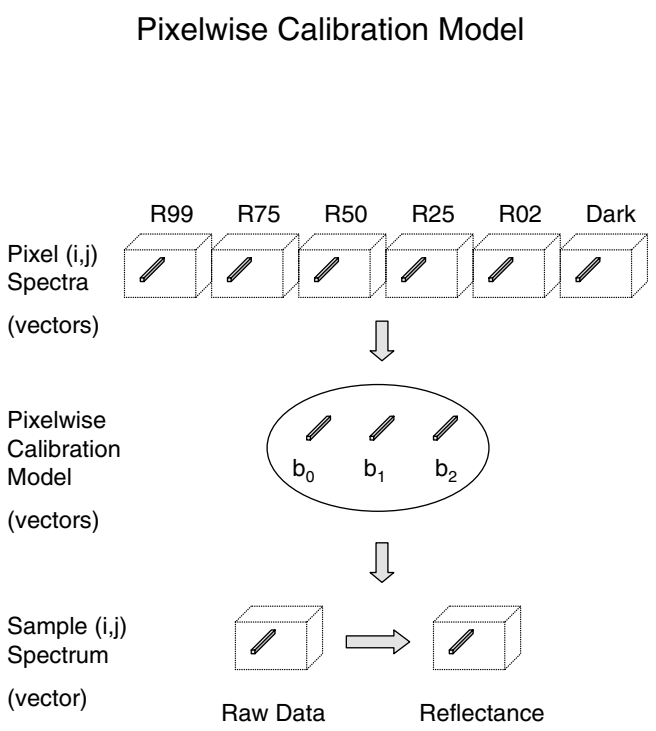

Figure 3. Each of the individual (*i,j*) spectra from the external standard hypercubes are used to compute individual pixel calibration model vectors (b_0 and b_1 of Equation (3) or b_0, b_1, and b_2 of Equation (4) at each wavelength). These pixel (*i,j*) specific model vectors are then used to transform the respective (*i,j*) raw data sample spectra from instrument counts to reflectance.

reflectance transform models. Such models compensate for effects due to differing sensor response sensitivities, sensor non-linearity, or non-uniform lighting. But other factors influence the image quality as well. Thermal drift of the detector and filters, age or temperature effects on the lamps, or power supply voltage fluctuations can all contribute to degradation of image quality. These factors and others have been addressed collectively in spectroscopy as instrument standardization or calibration transfer issues.

One of the advantages of hyperspectral images is the immense amount of data available. This enables large sample statistics to be computed. It also permits the opportunity to target some of the data space specifically for calibration or diagnostic purposes. Adding a small mosaic of internal standard materials, which fills only 5% of the field of view of an 320×256 pixel image, creates over 4000 spectra in every hypercube to perform data validation tests.

If the internal standards are calibrated and have known spectra, then correction models may be computed to realign the measured spectra. Corrections in both wavelength or peak position and reflectance or absorbance intensities can be determined. An alternative approach used in this paper is to simply declare one sample hypercube 'master' and adjust other 'slave' sample hypercubes to match. Normalization of measured intensities is performed by using either a single-point stretch (Equation 5) or a multiple-point linear (Equation 6) or quadratic (Equation 7) regression-based adjustment, depending on the number of internal standards available:

$$R_c = \alpha R \tag{5}$$

$$R_c = \alpha_0 + \alpha_1 R \tag{6}$$

$$R_c = \alpha_0 + \alpha_1 R + \alpha_2 R^2 \tag{7}$$

Here R_c is the corrected reflectance value. The α_i constants are computed from the measured master versus slave values. These internal standard-based corrections are global corrections applied to the entire image. That is, a single correction model based on internal standard ROIs is used to correct the reflectance values of all pixels in the image for a single wavelength channel. These correction models are of course wavelength channel dependent, and therefore α_i vectors must be computed to correct each complete hypercube. This is equivalent to piecewise direct standardization (PDS) with a window width of one [11]. It is assumed that spatial variation dependencies have been previously removed by the external standard-based reflectance transform. Consequently, positional registration of the internal standards from hypercube to hypercube is not necessary.

Figure 4 indicates the two primary steps necessary for these adjustments. A sample raw data hypercube C is first

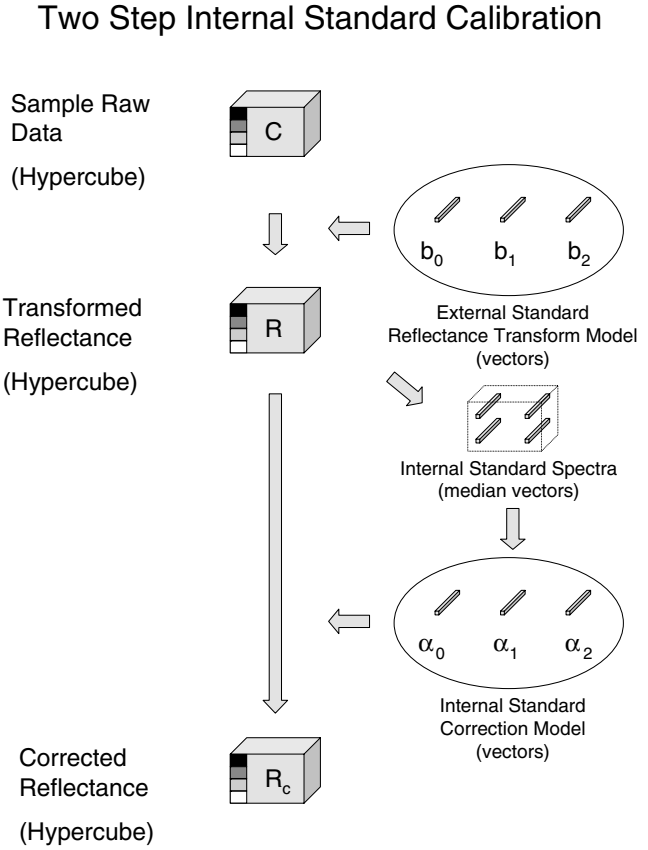

Figure 4. Sample hypercubes C are first transformed to reflectance *R* using the pixelwise wavelength-specific regression coefficient vectors from the external standards calibration model. Median spectra from each of the internal standard regions are then used to compute a second wavelength-specific correction model (α_i of Equations (5–7) at each wavelength), which is then applied to all spectra of the current reflectance hypercube to compute the corrected reflectance hypercube R_c.

transformed to reflectance R using an external standard-based model. Spatial regions of interest are then identified for each internal standard and median spectra computed for each internal standard. These median spectra are then compared with either known calibration spectra or spectra from a master hypercube to create a second correction model, which is then applied globally to each spectrum of the hypercube, resulting in the corrected reflectance hypercube R_c. This correction model is specific to only one hypercube and must be recomputed for each subsequent hypercube. For the purpose of discussion, as with the reflectance transform models, the internal standard-based correction models will be termed *simple*, *linear* or *quad* to represent one-point, or multi-point first- or second-order regression calibration models.

3. EXPERIMENTAL/MATERIALS AND METHODS

3.1. Hyperspectral images

NIR hyperspectral images were collected using a Spectral Dimensions matrix NIR instrument. This instrument uses a liquid crystal tunable filter (LCTF) bandpass filter in combination with an InGaAs diode array detector [21]. Images are digitized with 12 bit resolution. Each image was measured with a 32 ms integration time, repeated 10 times, and averaged. Individual images were acquired between 900 and 1662 nm at 6 nm intervals. This set of individual images when stacked, created a $320 \times 256 \times 128$ hyperspectral image, or hypercube.

Four quartz halogen lamps were positioned to provide as uniform as possible illumination of the image field of view (62×50 mm). The halogen lamps produce a significant amount of heat and their use at maximum power is only possible for inorganic materials. For this study, the lamp power was initially set at 60%, and then sequentially reduced for some images. Individually calibrated tiles of Spectralon standard NIR reflectance material [15] were used for image calibration and test samples. Reference spectra were computed as parabolic interpolations from certified tables provided by Labsphere (Sutton, NH) [8]. The Spectralon material for these samples (99, 75, 50, 25, and 2% reflectance) is created by adding an increasing amount of carbon black to a white Teflon base material, which at high magnification appears inhomogeneous and textured. To avoid imaging these non-uniformities, the external standard samples were positioned on a rotating bearing and manually rotated during image acquisition. Samples were also positioned 5 cm above the normal camera focal plane to purposely blur any surface defects.

3.2. Experiment 1: external standards

To test the effectiveness of the different reflectance transform functions a set of *external standard* hyperspectral images was measured where the large spinning spectralon tile filled the entire camera field of view. (Images R99, R75, R50, R25, and R02) This provided hypercubes of uniform material at every pixel at all wavelengths. An additional uniform dark hypercube was measured by simply blocking the entrance to the camera. Replicate images were measured to allow the construction of different calibration and test datasets.

3.3. Experiment 2: internal standards

Effects of lamp aging and instability were simulated by a drastic reduction of the lamp power. In order to study corrections for this phenomenon a series of *internal standard* images (a mosaic of multiple Spectralon samples) was measured while systematically varying the lamp power level. Four of the Spectralon tiles (25, 50, 75, and 99% reflectance) were arranged to cover the four corners of the field of view. Ten stationary hypercube images of this Spectralon mosaic were then measured while decreasing the lamp power to 60, 59, 58, 57, 56, 55, 50, 45, 40, and 35% power. Spatial surface irregularities in these stationary images were compensated for by computing median spectra of large internal standard region of interest (ROI) areas.

3.4. Data analysis

It is common for InGaAs diode array detectors to have a few elements that respond with unusually high or low values commonly referred to as *dead pixels*. These pixels have high leverage and can lead to erroneous results in calibration models. Rather than compute global average values, median spectra were computed to eliminate the high leverage effects of the outlier dead pixels. Sorting the intensities measured within an ROI by value effectively places the dead pixel measurements at the endpoints of the measurement range, minimizing the impact on median computation. Alternatively, these outlier pixels were detected by using simple univariate thresholding techniques and eliminated from any computational work. All references to average spectra in this paper assume that such outliers have been detected and removed before averaging. No other data pretreatment was performed. All computations were performed using MATLAB version 6.1 [22].

4. RESULTS AND DISCUSSION

4.1. External calibration

Global and pixelwise reflectance transform models were computed from the calibration hypercubes of the Spectralon standard reference materials. The R99 and dark current hypercubes were used to create the simple model. An additional α correction term was added to compensate for the fact that a 99% standard was used to represent total reflectance. This is effectively a combination of Equations (3) and (5):

$$R_c = \alpha(b_0 + b_1 C) \tag{8}$$

The correction term α is wavelength dependent and is equal to the reported reflectance value for the R99 standard reference material and applies to the *simple* model only.

The R99, R75, R50, R25, and R02 hypercubes were used to create linear and quad regression models. Every spectrum in a replicate R50 test hypercube was then individually transformed to reflectance using the global-based models. The median of these reflectance spectra is presented in Figure 5. In contrast to the highly irregular raw counts spectra of Figure 1, fairly flat reflectance spectra are obtained as expected. The deviations in the 900–950 nm range are attributed to the very low signal to noise ratio of this region. It can be clearly seen, however, that the reflectance transform models based on regression models of multiple standards

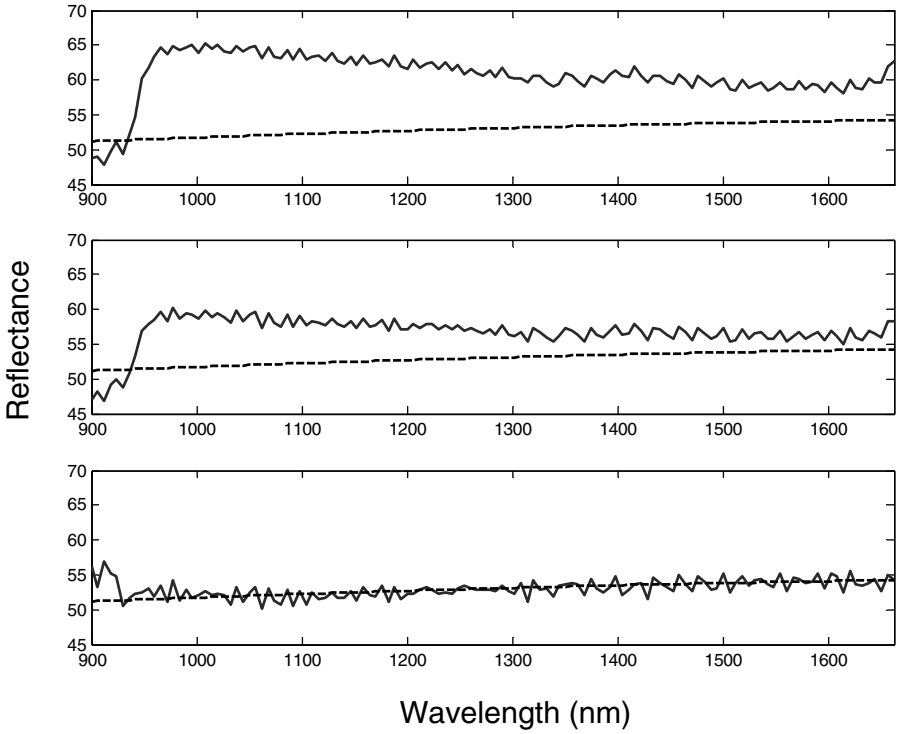

Figure 5. Average predicted reflectance spectra of 50% Spectralon based on global image models of calibration standards. Simple (top), linear (middle), and quad (bottom) models versus true reflectance spectra (- - -).

improve the predicted results. The spectrum based on the second-order quad model was nearly indistinguishable from the expected true spectrum. Similar results were obtained with replicate images of the other calibration standards.

Figure 6 shows a comparison of the results for the global- versus pixelwise-based transforms of a replicate R75 hypercube image. The top plot is based on the quad models of the global image median spectra; the bottom plot is based on the

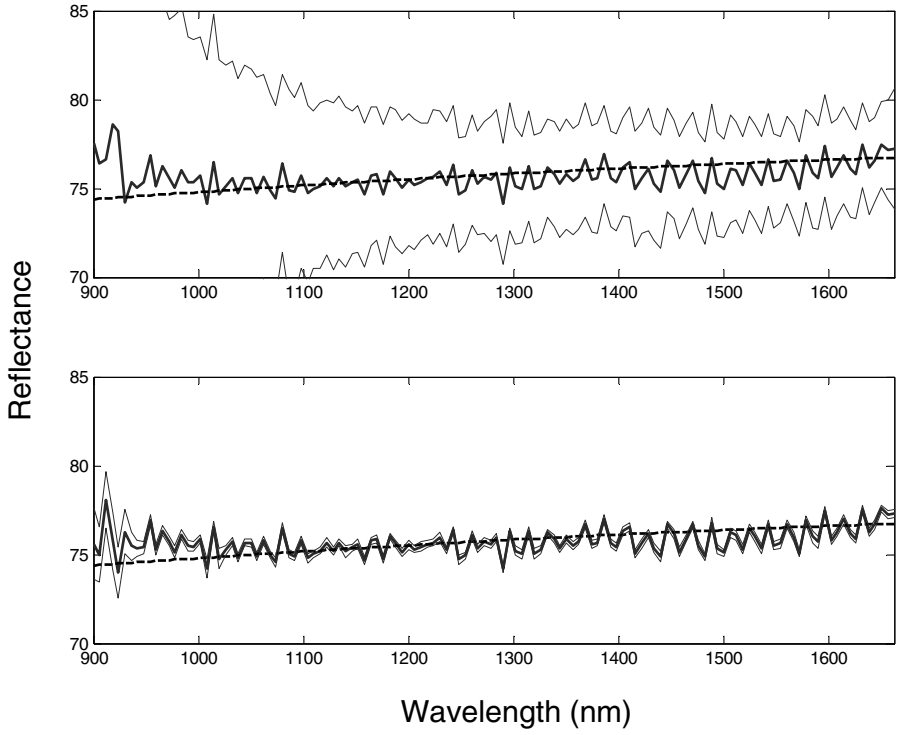

Figure 6. Average predicted spectra of 75% Spectralon (dark line) including standard deviation bands (lighter lines) based on global image quad models (top) versus individual pixelwise quad models (bottom) of calibration standards versus true reflectance spectra (- - -).

Table I. External standard calibration. Average measurement errors (percent reflectance) for the 75% Spectralon standard, 1100–1600 nm

	Model	Bias	Standard deviation
Global	Simple	4.65	2.63
	Linear	0.91	2.78
	Quadratic	−0.34	3.24
Pixelwise	Simple	4.75	0.21
	Linear	1.00	0.19
	Quadratic	−0.26	0.23

pixelwise quad models computed at each individual pixel location. There is a slight improvement in the lower wavelength range, but average spectra results for the two modeling techniques are quite similar. This is because the two single spectra plotted represent the average of thousands of spectra measured from a uniform material, with fairly uniform lighting. The standard deviation spectra show the benefit of the pixelwise modeling: the standard deviations have been significantly reduced. Any deviations between individual sensors of the diode array detector, or differences in lighting effects, have been reduced by the individual pixel modeling. A more typical hyperspectral image containing multiple objects with differing spectral components would clearly benefit from this pixel-specific modeling advantage.

Table I summarizes the average results for predicting the 75% reflectance material between 1100 and 1600 nm. A comparison of the global and pixelwise models indicates nearly identical bias values; however, the standard deviations are reduced by a factor of 10 for the pixelwise models. For non-uniform sample illumination this difference would be even more apparent. Simple calibration models based on only dark current and a single external standard were the least effective. The linear and quad regression model transforms based on multiple external standards performed significantly better. Maximum accuracy for the reflectance transform model was achieved using multiple standards (without a need for a dark current measurement) and computing second-order regression models for each pixel at each wavelength.

4.2. Internal calibration

The inclusion of internal standards was examined using a series of 10 hypercubes where the light illumination intensity was varied by purposely changing the lamp power. The series of reduced power hyperspectral images was first transformed to reflectance using the pixelwise-based quad models based on the previously measured external standard hypercubes at 60% power. Four ROIs were then identified and the median value of each region at each wavelength determined. Again median spectra were computed to minimize leverage contributions of dead pixels and sample inhomogeneity effects of the stationary materials imaged. The data displayed in Figure 7 represent a single-wavelength channel (centered at 1602 nm) but plots at other wavelengths showed similarly shaped response curves. The response signals increase non-linearly with an increase in lamp power (filament temperature) as indicated in the measured reflectance values.

Figure 8 shows the median spectra of the R75 reflectance standard ROI for the eight lamp power settings between 60 and 45%. Note the change in vertical axis scaling for viewing purposes. The top plot shows that decreasing the lamp power reduced the uncorrected reflectance spectrum from

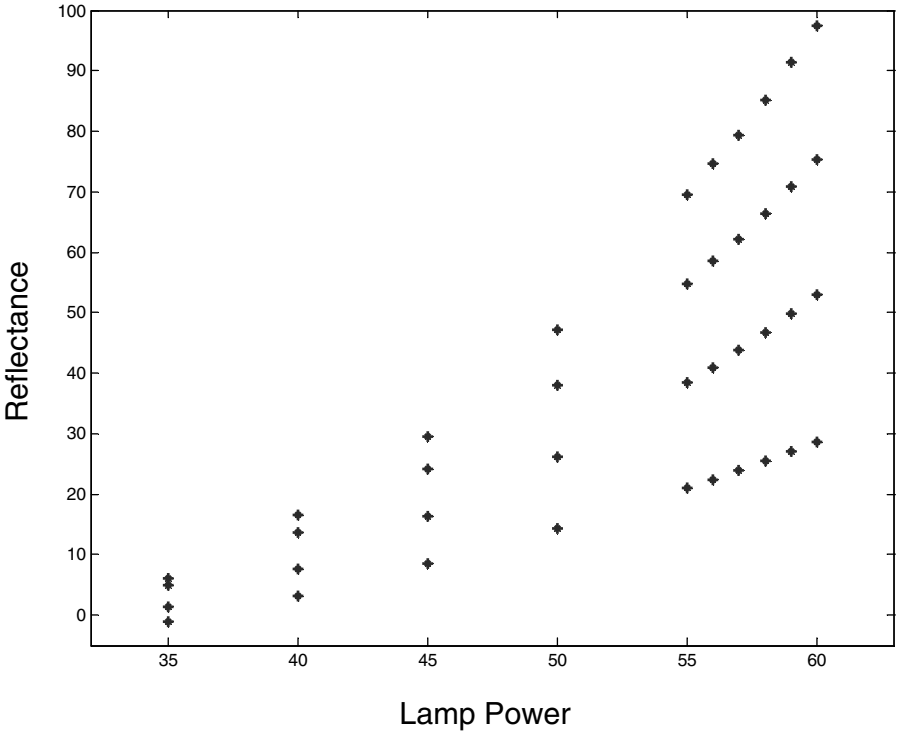

Figure 7. Median reflectance values (1602 nm) versus lamp power for 99 (top), 75, 50, and 25% Spectralon.

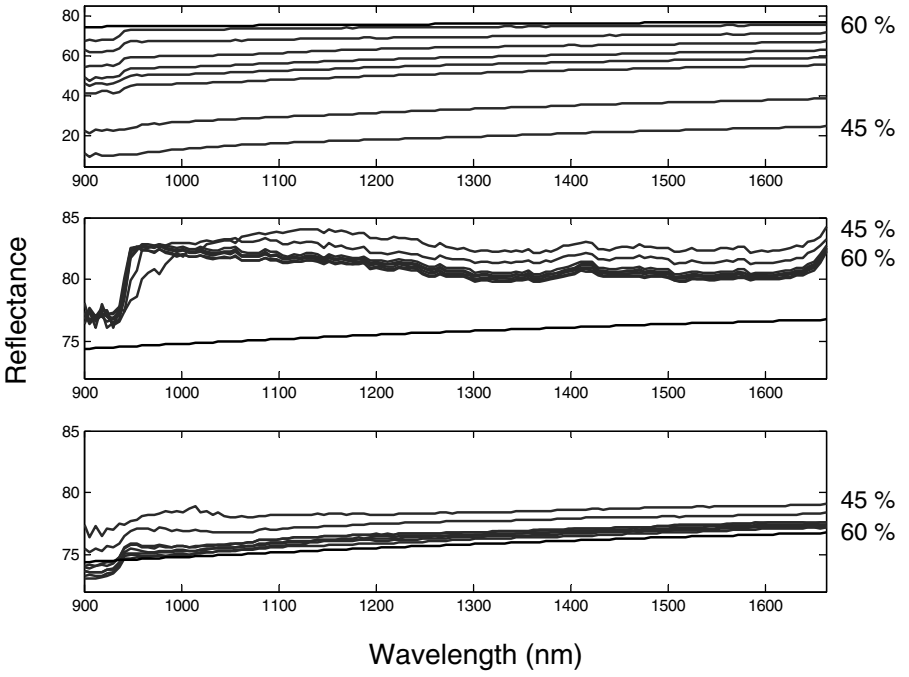

Figure 8. Average reflectance spectra of 75% Spectralon. Uncorrected (top) versus single point (middle), and quad (bottom) corrections based on internal calibration standards. Lamp power sequentially reduced from 60 to 45%. Note change in vertical axis scaling.

the expected 75% to between 10% and 20% reflectance values. The middle plot shows the corrected results using a single internal standard. In this case α scaling factors were computed by scaling the measured median spectra of the R99 internal standard back to the expected value (Equation 5). This simple approach resulted in corrected R75 spectra with only 5%–10% deviations from expected values.

The bottom plot in Figure 8 shows the effect of a quad fit to the median spectra of the R99, R50, and R25 standards (Equation 7). Reducing the lamp power from 60% to 55% reduced the uncorrected spectra from the expected 75% to roughly 40% reflectance (Figure 8 top). But adding the second quad-based internal standard correction resulted in a corrected spectra with only 1–2% error (Figure 8 bottom). Even the 45% lamp power spectra were corrected to within 4% error.

Table II summarizes the results for adjustments to the 75% reflectance standard based on models of the other internal standards. Reducing the lamp power from 60 to 45% is a drastic change, yet even the one-point simple adjustment corrects the bias to within 4–7%. In all cases the bias and standard deviation values increase with a reduction in power, as would be expected from heteroscedastic noise in the system. The best results were obtained with the second-order quadratic correction model. In this example, the correction models were based on only three internal standards, used to predict the values of the 4th standard. Even better results would be expected from creating second-order correction models based on all four or more internal standards.

Table II. Internal standard calibration. Average measurement errors (percent reflectance) for the 75% Spectralon standard, 1100–1600 nm; bias (standard deviation)

Power	Uncorrected	Simple	Linear	Quadratic
60	−1.55(0.17)	4.45(0.50)	0.63(0.07)	0.38(0.04)
59	−6.32(0.69)	4.54(0.51)	0.78(0.09)	0.59(0.06)
58	−11.47(1.26)	4.70(0.53)	0.88(0.10)	0.69(0.08)
57	−16.27(1.79)	4.84(0.54)	0.97(0.11)	0.79(0.09)
56	−19.97(2.20)	4.91(0.55)	1.11(0.13)	0.98(0.11)
55	−24.08(2.65)	5.11(0.57)	1.17(0.13)	0.99(0.11)
50	−41.79(4.59)	5.90(0.65)	1.99(0.22)	1.80(0.20)
45	−55.47(6.09)	6.87(0.76)	2.99(0.33)	2.58(0.28)

5. CONCLUSIONS

Hyperspectral images contain a rich mixture of both spectral and spatial information. For effective use of these images to identify chemical components and to determine their concentration distributions, accurate calibration steps must be performed to transform measured detector responses into reflectance or absorbance values. This requires calibration by the use of standards such as Spectralon NIR standard reflectance material. Problems with surface defects and spatial variations in material composition can be easily corrected. Sample variations can be corrected by either spinning the reference material during image acquisition, or averaging the hundreds or thousands of spectra measured within an internal standard image region. The use of this material as both external and internal standards permits the uniform and accurate transformation of signal counts to reflectance at all pixel locations and at all wavelengths of the hyperspectral image.

Three aspects of hyperspectral image calibration were considered: (1) data-subset selection (global vs. pixelwise modeling), (2) data model type (simple two point linear fit,

or linear or quadratic regression), and (3) type of standards selected (external or internal). Results from these options are summarized in Tables I and II.

The tables show that data subset selection was optimized using the pixelwise models, allowing for corrections of both spatial and wavelength-dependent variations due to non-uniform lighting problems or variations in individual diode sensor sensitivity. Data model type was optimized by the use of multiple external reflectance standards coupled with pixelwise second-order regression models. This combination is recommended for the accurate transformation of hyperspectral images from instrument counts to reflectance values.

Use of internal standards incorporated into the image allowed for detection and correction of time-dependent variations due to temperature- or power-induced changes to the overall image. As with the external standard calibrations, corrections based on only the 99% standard were least effective. Optimum results were obtained using an array of internal standards to compute a first- or second-order regression-based correction model. (Table II) Any measurement variations within the internal standard material were simply removed by computing median spectra of each internal standard ROI. This technique also eliminated any need for accurate registration or placement of the internal standard sample locations from image to image.

The primary goal of this study was the testing of external and internal standards for calibrating and correcting NIR hyperspectral images. The Spectralon standards used in this study all had fairly flat reflectance spectra throughout the NIR region and were selected to span reflectance values between 99 and 2%. The same material was used for both calibration and test samples. But there is a clear need for the development of other secondary NIR standards. These standards should similarly span high and low reflectance values, and be easily calibrated with external reference spectra. Any sharp spectral features would enable testing of wavelength accuracy. This study focused on the use of spectral standards for use in correcting hyperspectral image intensities. The design of image standards for spatial or geometry calibration is also needed. Further development of such NIR hyperspectral image standards is ongoing.

An application of the use of external and internal standard-based corrections to hyperspectral NIR image regression will be presented in Part II.

Acknowledgments

The authors acknowledge financial support from NIRCE, an EU Unizon-Kvarken project. We also acknowledge Torbjörn Lestander, Swedish University of Agricultural Sciences and Robert Cogdill, Iowa State University for related discussions.

Kempestiftelserna grant SMK-2062 provided funds for the equipment.

REFERENCES

1. Wolfe W. *Introduction to Imaging Spectrometers*. SPIE Optical Engineering Press: Bellingham, Washington, USA, 1997.
2. Polder G, vander Heijden G, Keizer L, Young I. Calibration and characterisation of imaging spectrographs. *J. Near Infrared Spectrosc.* 2003; **11**: 193–210.
3. Martinsen P, Schaare P, Andrews M. A versatile near infrared imaging spectrometer. *J. Near Infrared Spectrosc.* 1999; **7**: 17–25.
4. Brianco A, Serafino G, Spoeck G. *An introduction to spectral imaging*. http://www.ctr.at/hypspbvdoc/IntSpImg.pdf. 2005.
5. Hoelter T, Barton J. Extended short wavelength spectral response from InGaAs focal plane arrays. *Proc. SPIE* 2003; **5074**: 481–490.
6. Gat N. Imaging spectroscopy using tunable filters: a review. *Proc. SPIE* 2000; **4056**: 50–64.
7. Tran CD. Infrared multispectral imaging: principles and instrumentation. *Appl. Spectrosc. Rev.* 2003; **38**: 133–153.
8. Geladi P, Burger J, Lestander T. Hyperspectral imaging: calibration problems and solutions. *Chemometr. Intell. Lab. Syst*. 2004; **72**: 209–217.
9. Wang Y, Veltkamp D, Kowalski B. Multivariate instrument standardization. *Anal. Chem.* 1991; **63**: 2750–2756.
10. Wang Y, Kowalski B. Calibration transfer and measurement stability of near-infrared spectrometers. *Appl. Spectrosc.* 1992; **46**: 764–771.
11. Wang Y, Lysaght M, Kowalski B. Improvement of multivariate calibration through instrument standardization. *Anal. Chem.* 1992; **64**: 562–564.
12. Fearn T. Standardisation and calibration transfer for near infrared instruments: a review. *J. Near Infrared Spectrosc.* 2001; **9**: 229–244.
13. Feudale R, Woody N, Tan H, Myles A, Brown S, Ferre J. Transfer of multivariate calibration models: a review. *Chemometr. Intell. Lab. Syst.* 2002; **64**: 181–192.
14. Geladi P. Some recent trends in the calibration literature. *Chemometr. Intell. Lab. Syst*. 2002; **60**: 211–224.
15. Spectralon Diffuse Reflectance Targets, Labsphere, Inc., PO Box 70, 231 Shaker Street North Sutton, NH 03260, USA.
16. Geladi P, Wold S, Esbensen K. Image analysis and chemical information in images. *Anal. Chim. Acta*. 1986; **191**: 473–480.
17. Esbensen K, Geladi P. Strategy of multivariate image analysis. *Chemometr. Intell. Lab. Syst*. 1989; **7**: 67–86.
18. Geladi P, Esbensen K. Regression on multivariate images: principal component regression for modeling, prediction and visual diagnostic tools. *J. Chemometr.* 1991; **5**: 97–111.
19. Esbensen K, Geladi P, Grahn H. Strategies for multivariate image regression. *Chemometr. Intell. Lab. Syst*. 1992; **14**: 357–374.
20. Lied T, Geladi P, Esbensen K. Multivariate image regression (MIR): implementation of image PLSR—first forays. *J. Chemometr.* 2000; **14**: 585–598.
21. Spectral Dimensions, Inc., 3416 Olandwood Court, Suite 210, Olney, MD 20832, USA.
22. The MathWorks, Inc., 3 Apple Hill Drive, Natick, MA 01760-2098, USA.