ELSEVIER

Derivative Analysis of Hyperspectral Data

Fuan Tsai and William Philpot**

With the goal of applying derivative spectral analysis to analyze high-resolution, spectrally continuous remote sensing data, several smoothing and derivative computation algorithms have been reviewed and modified to develop a set of cross-platform spectral analysis tools. Emphasis was placed on exploring different smoothing and derivative algorithms to extract spectral details from spectral data sets. A modular program was created to perform interactive derivative analysis. This module calculated derivatives using either a convolution (Savitzky–Golay) or finite divided difference approximation algorithm. Spectra were smoothed using one of the three built-in smoothing algorithms (Savitzky–Golay smoothing, Kawata–Minami smoothing, and mean-filter smoothing) prior to the derivative computation procedures. Laboratory spectral data were used to test the performance of the implemented derivative analysis module. An algorithm for detecting the absorption band positions was executed on synthetic spectra and a soybean fluorescence spectrum to demonstrate the usage of the implemented modules in extracting spectral features. Issues related to smoothing and spectral deviation caused by the smoothing or derivative computation algorithms were also observed and are discussed. A scaling effect, resulting from the migration of band separations when using the finite divided difference approximation derivative algorithm, can be used to enhance spectral features at the scale of specified sampling interval and remove noise or features smaller than the sampling interval. ©*Elsevier Science Inc., 1998*

INTRODUCTION

Until relatively recently, imaging remote sensing systems have been limited to multispectral devices, that is, devices using only a few, carefully selected and usually discontinuous wavebands. The four-band MSS, the seven-band TM, and the three-band SPOT are typical of this type of imaging system. Data generated by multispectral sensors have been valuable sources of remote sensing data, and numerous effective methods have been developed and applied successfully for spatial or spectral analysis of these data, usually derived from established methods in multivariate statistics (Duda and Hart, 1973; Richards, 1993).

When higher resolution, spectrally continuous remote sensing data (e.g., AVIRIS—Airborne Visible/Infrared Imaging Spectrometer) became available, researchers in remote sensing tended to select suitable bands to optimize the existing algorithms of multispectral data analysis (Hoffbeck and Landgrebe, 1993) or to generate new algorithms based on traditional multispectral concepts (Chappelle et al., 1992; Curran et al., 1992; Martin and Aber, 1993; Peñuelas et al., 1994). However, typical multispectral analysis methods treat each spectral band as an independent variable—a reasonable assumption for multispectral data, but not really appropriate for hyperspectral data. Only a few researchers have tried to employ approaches commonly used in spectroscopy (Talsky, 1994) or have manipulated data as truly spectrally continuous data (Demetriades-Shah et al., 1990).

Not all methods used in spectroscopy can be directly adapted to remote sensing analysis because there are significant differences between these two types of data. Data used in spectroscopy are usually collected under controlled laboratory conditions with full control of the intensity and spectral distribution of the illumination as well as viewing geometry. The sample has usually been prepared to enhance detection of the target substance. A single sample is assumed to be homogenous and uniformly illuminated over the viewing area. Finally laboratory procedures usually involve the use of a standard. This may be a reflection, absorption, or fluorescence standard with well-defined properties, or a "known" preparation of the target substance.

In contrast, hyperspectral remote sensing data are

*CIRIS, Cornell University, Ithaca

Address correspondence to Fuan Tsai, 453 Hollister Hall, CIRIS/Cornell Univ., Ithaca, NY 14853. E-mail: ft11@cornell.edu

Received 31 July 1997; revised 13 April 1998.

REMOTE SENS. ENVIRON. 66:41–51 (1998)

655 Avenue of the Americas, New York, NY 10010

0034-4257/98/$19.00
PII S0034-4257(98)00032-7

collected using natural illumination. In most cases, the spatial resolution for a single pixel is greater than several meters in diameter (often several tens of meters), and it is rare for a single object or target feature to fill any one pixel. Even the topography may vary over a single pixel. Thus, the characteristics of any pixel can rarely be considered truly homogenous. The spectral resolution of hyperspectral systems is usually lower (0.005–0.01 μm) than that of laboratory spectrophotometers and is usually not constant over the spectrum. Furthermore, references with known properties for hyperspectral remote sensing data are difficult to obtain. Laboratory references may sometimes be helpful (Gamon et al., 1992), but they are not always effective in reducing noise because of illumination differences and atmospheric effects. Because of these differences, it is first necessary to examine thoroughly some particular algorithms spectroscopists have been using in order to adapt them to remote sensing spectral analysis successfully.

Among the techniques that have been developed in spectroscopy, derivative analysis is particularly promising for use with remote sensing data. Derivatives of second order or higher should be relatively insensitive to variations in illumination intensity whether caused by changes in sun angle, cloud cover, or topography. At the spectral sampling interval typical of hyperspectral systems, derivatives should also be relatively less sensitive to the spectral variations of sunlight and skylight. Nonetheless, only a few researchers so far have addressed applications of spectral derivatives in remote sensing (Demetriades-Shah et al., 1990; Peñuelas et al., 1994; Philpot, 1991). Although some of these studies have used high-order derivatives (Butler and Hopkins, 1970), first- and second-order derivatives have been the most common. In addition, whether in analytical spectroscopy or in remote sensing, most of the studies using spectral derivative analysis were directed toward particular applications, such as research on water constituent concentration of freshwater lakes (Dick and Miller, 1991), correction of atmospheric effects (Philpot, 1991), band decomposition (Huguenin and Jones, 1986), and estimation of suspended sediment concentration (Chen and Curran, 1992).

Because of the inherent differences between laboratory spectroscopic analysis and remote sensing analysis, several problems arise when applying derivative analysis to remote sensing data. Derivatives are notoriously sensitive to noise. Thus, smoothing or otherwise minimizing random noise is a major issue. Among the various methods for smoothing spectral data, algorithms based on least-square fits are most commonly selected. Perhaps the most common method is that of Savitzky and Golay (1964) which provides a simplified least square procedure for simultaneously smoothing and differentiating of data.

The Savitzky–Golay least-square-fitting procedure implicitly assumes that random noise has similar characteristics throughout the spectrum and that it can be handled by an invariant procedure over the spectrum. Kawata and Minami (1984), argued that, since random noise usually varies over the spectrum, polynomial curve-fitting might alter the signal waveform instead of eliminating the noise. They presented an alternative based on minimizing the mean-squared error. The process is essentially a locally variable filter which is adjusted for the local signal-to-noise ratio in order to smooth the target spectrum point by point.

Another commonly used smoothing algorithm is the mean-filter algorithm, which also locally smooths data within a predetermined smoothing window but without any polynomial curve-fitting or least-mean-square procedure in the smoothing computation. Instead, a mean-filter uses the mean value of samples within the local smoothing window as the new value of the middle sampling point in the smoothing window.

Neither Kawata and Minami's algorithm nor the mean-filter algorithm compute derivatives. Therefore, a separate derivative computation module is necessary for performing derivative spectral analysis (Demetriades-Shah et al., 1990). One alternative is a simple "finite divided difference approximation" of the derivative (Chapra and Canale, 1988), referred to hereafter simply as the finite approximation. The result of the approximation is dependent on the resolution (band separation) and the filter size or bandwidth (Philpot, 1991). Another benefit of using the numerical, finite approximation is that it is straightforward in computation and is less likely to introduce artifacts—a possible side effect of applying polynomial curve-fitting procedures.

The purpose of this study was to 1) treat remotely sensed hyperspectral data as truly spectrally continuous data and 2) develop a set of spectral analysis tools based on the use of spectral derivatives. These tools were designed specifically for dealing with fine resolution, spectrally continuous data. Emphasis was placed on exploring appropriate algorithms for smoothing spectra (noise reduction) and calculating derivatives with the goal of extracting subtle spectral features. The scaling effects that occur in derivatives as the sampling interval changes were also examined. Spectra of targets with strong spectral variability in remote sensing data were the focus of this study.

It is hoped that with these exploratory tools one would be able to smooth hyperspectral data, obtain derivatives of any order, and extract useful spectral features for further analysis. More importantly, these operations would be performed with no need to assume that the data consisted of homogeneous pixels or were generated in highly controlled environments. The eventual goal is to apply these algorithms and tools to the analysis of image data. However, for the sake of reducing the computation time, high resolution and spectrally continuous spectra of vegetation collected in laboratories and several synthetic spectra were used as test data.

METHODOLOGY

Several computation modules were developed in MATLAB (Mathworks, 1993) script for analyzing hyperspectral data in this study. With these modules, known collectively as HyperSpec,[1] one can smooth high resolution, spectrally continuous spectra, compute derivatives of different orders and at different sampling intervals. Four options were developed for smoothing and derivative calculation: 1) Savitzky–Golay smoothing and derivative computation, 2) Kawata–Minami smoothing and finite approximation derivative computation, 3) mean-filter smoothing and finite approximation derivative computation, and 4) Savitzky–Golay smoothing and finite approximation derivative computation.

Savitzky–Golay Smoothing and Derivative Computation

Savitzky and Golay (1964) used simplified least-square-fit convolution for smoothing and computing derivatives of a set of consecutive values (a spectrum). The general equation of the simplified least square convolution can be represented as follows:

$$Y_j^* = \frac{\sum_{i=-m}^{m} C_i Y_{j+i}}{N}, \tag{1}$$

where Y is the original spectrum, Y^* is the resultant (smoothed) spectrum, C_i is the coefficient for the ith spectral value of the filter (smoothing window), and N is the number of convoluting integers. The index j is the running index of the original ordinate data table. The smoothing array (filter size) consists of $2m+1$ points, where m is the half-width of the smoothing window. Savitzky and Golay (1964) provided several tables of coefficient values. Corrections to errors in the original published tables were provided by Steinier et al. (1972). From the point of view of computerized hyperspectral analysis, these tables have two disadvantages. First, for applications using limited combinations of derivative and polynomial orders, the table lookup approach might be efficient, provided all lookup tables have been constructed and computerized in advance. However, for applications which require exploring the spectra in various orders of derivatives and polynomials, a general analytical form for calculation would be preferable. Second, the tables provided by Savitzky and Golay have a limit of up to 25 points for a smoothing array ($m=12$).

This study does not employ a "universal" equation that is available for unlimited points in any order of derivative, nor does it use Savitzky and Golay's (Savitzky and Golay, 1964) tables directly. Instead, the computation is based on equations presented by Madden (1978). Madden's equations overcome the limit of 25 points for a smoothing array and use one equation for each order of derivatives. Following Madden (1978) Eq. (1) can be rewritten as Eq. (2):

$$Y_j^* = \sum_{i=-m}^{m} P_i^{(0)} Y_{j+i}, \tag{2}$$

where $P_i^{(0)}$ is the coefficient for the ith point of the filter in the zeroth order of derivative computation. Y_j is the midpoint of the smoothing array (window). Accordingly, the smoothed qth-order derivative of the midpoint is then represented as Eq. (3):

$$\frac{d^q \overline{Y_j}}{dx^q} = \sum_{i=-m}^{m} P_i^{(q)} Y_{j+i}. \tag{3}$$

Madden's equations allow calculation of the coefficients of least-square-fit convolution from zeroth order to the sixth order of smoothed derivatives.

A derivative spectrum can be generated by applying the procedures described above for calculating the smoothed derivative of the midpoint of the smoothing array over the spectrum. However, since the procedures cannot be applied correctly at the ends of the spectrum, the resultant spectrum is shorter than the original by the width of the filter.

Kawata and Minami's Linear Least Mean-Square Smoothing

Kawata and Minami (1984) presented an alternative algorithm for smoothing spectral data. Their algorithm is also based on the least mean-square process but attempts to account explicitly for random noise that may vary with wavelength. The observed spectrum $r(\lambda)$ can be described as the sum of the true signal of the spectrum, $s(\lambda)$, plus the noise $n(\lambda)$ in Eq. (4):

$$r(\lambda) = s(\lambda) + n(\lambda), \tag{4}$$

where λ denotes a wavelength or a wavenumber. The purpose of the smoothing process is merely to obtain an accurate $s(\lambda)$ from the observed spectrum $r(\lambda)$. Ideally the noise spectrum would be determined experimentally or derived for a particular instrument and experimental situation. For the purpose of this study, noise is assumed to be stationary random with zero mean and constant variance. The estimated noise-free spectrum $\hat{s}(\lambda)$ can then be expressed in terms of the current value of the observed data in Eq. (5):

$$\hat{s}(\lambda) = f(\lambda) r(\lambda) + g(\lambda). \tag{5}$$

Mean Filter Smoothing

A mean filter simply takes the mean spectral value of all points within the specified window as the new value of the middle point of the window in Eq. (6)

$$\hat{s}(\lambda_j) = \frac{\sum s(\lambda_i)}{n}, \tag{6}$$

[1] HyperSpec was developed as an exploratory tool for use in this study and was not prepared as a public release. However, readers who are interested in obtaining a copy of the module can contact the authors directly.

where n (number of sampling points) is the filter size and j is the index of the middle point of the filter. If the user specifies an even number of points as the filter size, the mean is assigned as the new value of the nearest point right of the center (longer wavelength).

Finite Approximation

Finite approximation can be used to estimate derivatives by suitable difference schemes in accordance with a finite band resolution, $\Delta\lambda$. In this study, the first derivative is estimated by Eq. (7):

$$\left.\frac{ds}{d\lambda}\right|_i \approx \frac{s(\lambda_i)-s(\lambda_j)}{\Delta\lambda}, \tag{7}$$

where $\Delta\lambda$ is the separation between adjacent bands, $\Delta\lambda=\lambda_j-\lambda_i$ and $\lambda_j>\lambda_i$, and the interval between bands is assumed to be constant. The second derivative can be derived from the first derivative as in Eq. (8):

$$\left.\frac{d^2s}{d\lambda^2}\right|_j = \frac{d}{d\lambda}\left(\frac{ds}{d\lambda}\right)\Bigg|_j \approx \frac{s(\lambda_i)-2s(\lambda_j)+s(\lambda_k)}{(\Delta\lambda)^2} \tag{8}$$

where $\Delta\lambda=\lambda_k-\lambda_j=\lambda_j-\lambda_i$, $\lambda_k>\lambda_j>\lambda_i$.

Accordingly, higher-order derivatives are computed iteratively and any order of derivative is accessible using the finite approximation. In general, the nth derivative can be represented as

$$\left.\frac{d^ns}{d\lambda^n}\right|_j = \frac{d}{d\lambda}\left(\frac{d^{(n-1)}s}{d\lambda^{(n-1)}}\right) = \cdots \approx \frac{s(\lambda_i)-\cdots+s(\lambda_{i+n})}{(\Delta\lambda)^n} = \frac{\sum_i^{i+n} C_k s(\lambda_k)}{(\Delta\lambda)^n} \tag{9}$$

where $j=(2i+n)/2$, if $(2i+n)$ is even, or $j=(2i+n+1)/2$, if $(2i+n)$ is odd; this means that if the position of the resultant derivative falls between sampling points, it is assigned to the sampling point at the next larger wavelength or wavenumber. The coefficients C_k are calculated using an iteration scheme.

RESULTS

Smoothing

All smoothing approaches remove random noise from spectra regardless of whether or not derivative computation has been specified. Savitzky and Golay (1964) and Kawata and Minami (1984) both smooth spectra by convolving a spectrum with a shaped filter in order to minimize the effect on spectral details. In contrast, the mean-filter removes noise using a simple average of points within the filter window. The primary factor controlling the extent of smoothing is the size (bandwidth) of the filter window used for convolution or averaging. In general, the greater the size of the filter window, the smoother the result. In this study, no quantitative measure of noise was implemented. Instead, the filter size was increased gradually until relatively smooth spectra or derivatives were obtained. As the filter size increases, however, spectral details may also be suppressed.

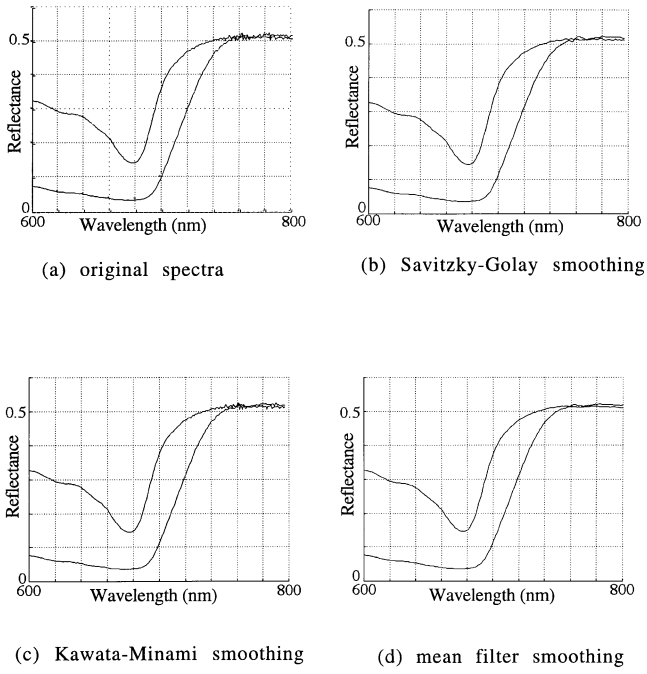


Figure 1. Results of smoothing in HyperSpec: filter size=7 points. a) Original spectra; b) Savitzky–Golay smoothing; c) Kawata–Minami smoothing; d) mean filter smoothing.

Figure 1 shows the results of the smoothing of rice reflectance spectra collected using a MSR-7000 spectrometer (Opt Research Inc.), with a wavelength range from 280 nm to 2500 nm by 1 nm, at the same filter size for each of the three methods. The mean filter algorithm achieved the strongest smoothing effect but probably also suppressed more details than the other algorithms. In contrast, the Kawata–Minami algorithm did not remove much noise. Noise suppression in the Kawata–Minami procedure is based on an estimation of machine (random) noise. The greater the noise, the more the smoothing. Realistically, machine noise will be a function of wavelength and sometimes of amplitude; however, noise was assumed to have constant variance with wavelength in the example in Figure 1. While the selected noise value appears to have been appropriate for wavelengths less than 750 nm, it may have been an underestimate for wavelengths greater than 750 nm, with the result that the spectrum remains noisy in this range.

There is always a trade-off between noise removal and the ability to resolve fine spectral detail. Ideal smoothing will remove noise without altering real spectral features, and the optimal filter size for approaching this ideal depends both on the particular spectra and on the smoothing algorithm. For the rice spectra in Figure 1, a minimum of five points as the filter size was required to produce usable derivative spectra.

The mean-filter algorithm will not perform as well

as Savitzky–Golay or Kawata–Minami filtering when the spectral features are smaller than or on the same order as the filter size. Indeed, the Savitzky–Golay and Kawata–Minami procedures were developed precisely because a simple mean (boxcar, bandpass) filter would be unable to discriminate between signal and noise in those cases. When the spectral features are broad and the noise is relatively high frequency compared to the real spectral features, there should be little significant difference among the three methods. The other smoothing methods were explored because it was hoped that there will be more useful information in some of the higher spectral resolution remote sensing data. However, for the data used in the examples presented here, one should not expect much difference among the filtering methods. In the following examples, the mean-filter smoothing algorithm is used most often, because it is straightforward and requires the least time for computation.

Derivative Analysis

Interactive derivative analysis was the primary focus of this study; an example is displayed in Figure 2. Figure 2a shows a set of reflectance spectra of soybean plants over the wavelength range of 400–800 nm (Adams et al., 1995), while Figures 2b and 2c are second-order derivatives of the same soybean spectra at different band separations. These spectra are from soybean plants provided with different levels of manganese ranging from sufficient to deficient. Plants with sufficient manganese were healthy and green. Plants deficient in manganese grew slowly and were chlorotic (Adams et al., 1995). The derivative spectra were generated using mean-filter smoothing in conjunction with the finite approximation algorithm. As shown in Figure 2a and 2b, although the original spectra differ in magnitude between 500 nm and 650 nm, they have almost identical values of the second order derivative over the same wavelength range. Several interesting spectral features are apparent in the derivative spectra that were obscure in the original spectra. For example, there are subtle changes in curvature between 550 nm and 650 nm that are consistent over the entire set of spectra. These changes appear as a shift from positive (at 575 nm) to negative (at 600 nm) to positive (at 625 nm) values in the second derivative indicating changes in curvature from concave up to concave down and back to concave up, indicating subtle, consistent variations in reflectance.

With the finite approximation, the spectral details revealed in the derivative spectra are a function of the band separation $\Delta\lambda$. In Figure 2b, the band separation is fixed at one sampling interval. (The actual sampling interval varies with wavelength—a characteristic of the radiometer—and is between 2.5 nm and 3.1 nm for the plotted wavelength range.) Figure 2c displays the same second derivative spectra except that the band separation

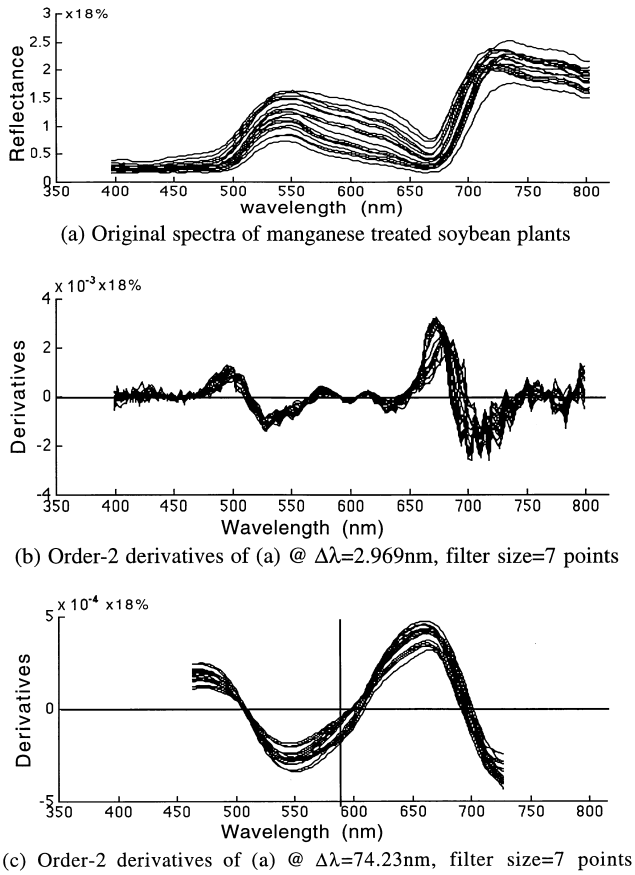
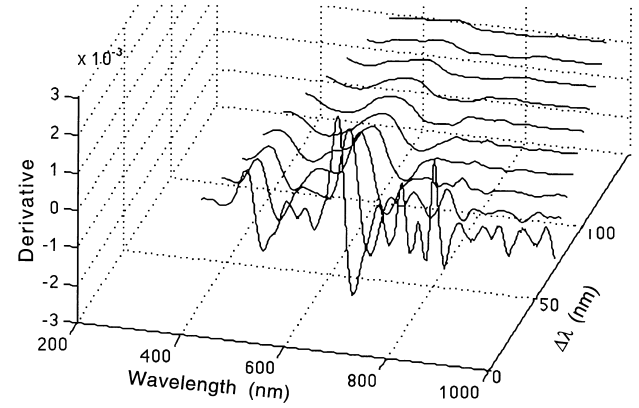
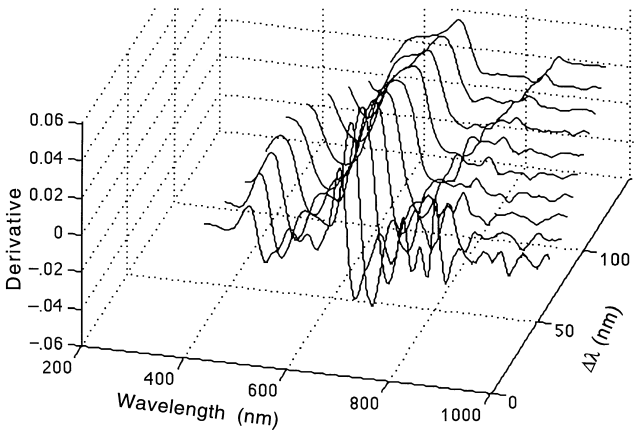


Figure 2. Derivative spectra of manganese treated soybean plants using mean-filter smoothing and finite approximation at b) 1 and c) 25 sampling intervals. The reflectance values in a) were measured relative to a Kodak garycard (18% reflectance).

was increased to 25 sampling intervals ($\Delta\lambda \approx 75$ nm). As might be expected, the small fluctuations between 550 nm and 650 nm which were apparent at the smaller band separations have disappeared. On the other hand, the spectra now fall into several groups, indicating that there are distinguishing features in the shapes of these spectra at this sampling interval. For example, at about 580 nm in Figure 2c, the second derivatives divide into two groups. This indicates that the original spectra at this wavelength have two ranges of curvature. One group of spectra has second derivatives that are close to zero, corresponding to spectra that are nearly flat at these wavelengths; the other group has second derivatives with more negative values, indicating that the original spectra have a concave down shape. The implication is that distinguishing spectral features occur at different wavelength scales (as determined by the band separation, $\Delta\lambda$) and that the derivative will extract different information dependent on the choice of $\Delta\lambda$. For the manganese treated soybean spectra, since chlorophyll concentrations are directly related to manganese treatment, the differences in chlorophyll concentrations of the spectra are

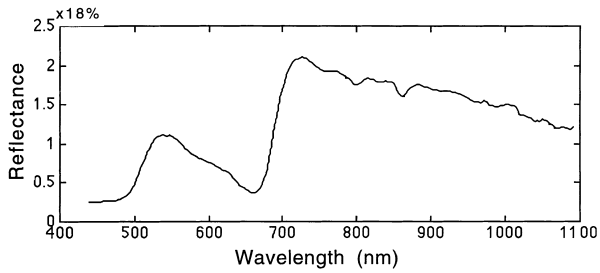


(a) Ordinary (unenhanced) second derivative spectra

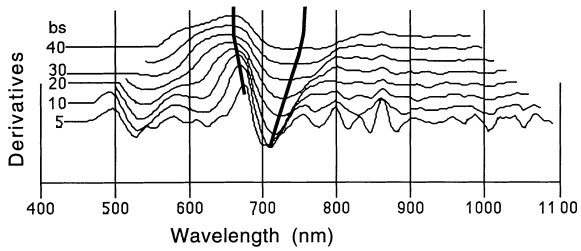


(b) Enhanced second derivative spectra

Figure 3. Enhancement effects in finite approximation algorithm.



(a) A spectrum of manganese treated soybean



(b) Order-2 Digital Derivatives after Savitzky-Golay smoothing

Figure 4. The shift effect of spectral features in a series of band separations. Spectra are displaced vertically in order to demostrate the spectral shift introduced by the changing band separations: a) a spectrum of manganese treated soybean measured relative to a Kodak graycard (18% reflectance); b) second derivatives calculated with Sovitzky–Golay smoothing (polynomial order 2, filter size 5) and finite approximation with band separations from 5 to 40 sampling intervals by 5 intervals.

probably the cause of these spectral differences (Adams et al., 1995).

As the band separation $\Delta\lambda$ increases, the magnitude of the spectral derivatives will be depressed because the derivatives are normalized by a power of $\Delta\lambda$ in the finite approximation [Eq. (9)]. This is illustrated in Figure 3a, which shows the second derivative for a single spectrum computed for several values of $\Delta\lambda$. Clearly, the magnitude of the derivative is damped as the band separation increases. In order to facilitate visual comparison of the derivatives as $\Delta\lambda$ increases, the finite approximation was "enhanced" by replacing the denominator of Eq. (9) with the band separation, $\Delta\lambda$, regardless the order of the derivatives. Figure 3b shows the same sequence of second derivatives as Figure 3a, but with the enhancement. The spectra are now more comparable, and the spectral detail at larger band separations is more apparent. This enhancement is an empirical choice based on experience with computing and comparing many different orders of derivative spectra and was used only to simplify visual interpretations when comparing spectra analyzed at different band separations.

Another shortcoming of using wide band separations is that, because the finite approximation calculates the derivative value of the middle point of a wavelength range, it is not possible to compute derivatives near the beginning or ending of the original spectra. If the derivative calculation requires samples over a bandwidth $\Delta\lambda$, then derivatives cannot be calculated at wavelengths closer than $\Delta\lambda$ to the end points of the original spectra.

Band separation also causes other deviations of spectral features. One of them is the displacement of spectral attributes in the wavelength domain if the original spectrum is not symmetric. Figure 4 presents one of the reflectance spectra used in Figure 2 and its enhanced second order derivatives at a series of band separations from five sampling intervals to 40 sampling intervals at a five-interval increment. Derivative spectra in this figure are displaced vertically in order to demonstrate the spectral shift introduced by changing band separation. The positions of the significant valley point vary from around 700 nm toward 750 nm and the peak at about 680 nm shifts toward shorter wavelengths. Other shifts toward longer or shorter wavelengths also occur over the spectra to-

Table 1. Parameters of the Six Constituents of the Synthetic Spectrum Displayed in Figure 5

Band	*Peak Position*	*Peak Value (%)*	*Half-Width*
1	9500	30	2355
2	11,500	42	3040
3	14,500	30	1990
4	16,000	34	2150
5	18,500	60	2033
6	20,500	80	2150

gether with a loss of spectral detail. This generally represents a loss of high resolution information. Alternatively, it also represents the spectral structure at a different scale. This effect is illustrated in the following example application.

An Example of Derivative Applications (Band Decomposition)

Huguenin and Jones (1986) introduced an algorithm for detecting the absorption band positions from reflectance spectra—a form of spectral decomposition—using derivative analysis. The criteria for determining the occurrence of an absorption band position is that the fifth derivative of the spectrum equals zero, the fourth derivative is positive, and the second derivative is negative. These criteria and a synthetic spectrum used by Huguenin and Jones (1986) are used here to test the performance of HyperSpec procedures in detecting the original peak positions of the subspectra that compose the synthetic spectrum. The spectrum was composed of six overlapping Gaussian curves, each with a different mean, peak strength, and half-width (full width at half maximum, FWHM) values. Each Gaussian band can be represented as Eq. (10):

$$f(k)=h*\exp[-0.5*(k-k_0)^2/\sigma^2], \qquad (10)$$

where h is the peak strength, k_0 is the position of mean center, and σ is the standard deviation of the distribution function, whose value, in this case, was generated from the predetermined half-width in Eq. (11):

$$W=2*(2*\ln 2)^{1/2}\sigma. \qquad (11)$$

Wavenumber (cm^{-1}) is used here to be consistent with Huguenin and Jones (1986). The band resolution for generating the synthetic spectrum is one wavenumber. Table 1 lists the values of band center, maximum band strength, and half-width of the six elementary Gaussian bands. Figure 5 shows the six elementary spectra and the composite spectrum.

The mean-filter smoothed, finite approximation algorithm was selected to examine the various orders of derivatives. Figure 6 combines the results of the second, fourth, and fifth derivatives with a filter size of 11 sampling points and band separation of 50 cm^{-1}. It illustrates that HyperSpec successfully identified the absorption

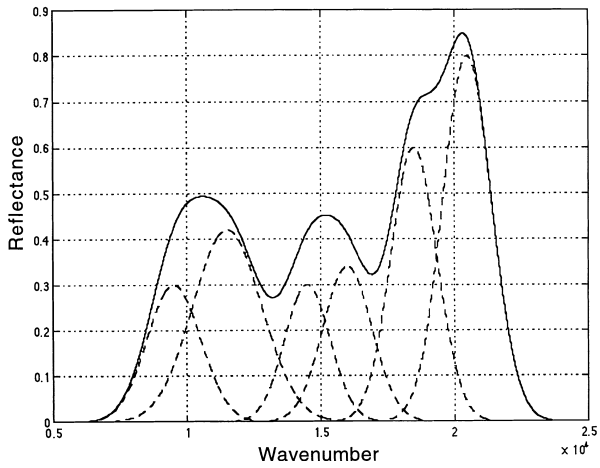


Figure 5. Synthetic spectrum (solid line) composed of six Gaussian constituents [(---) see Table 1].

band positions at 9500 cm^{-1}, 14,500 cm^{-1}, 18,500 cm^{-1} and 20,500 cm^{-1}, and located the fourth peak (16,000 cm^{-1}) slightly lower than it was supposed to be, but misallocated the position of 11,500 cm^{-1}. In fact, this result is very close to and even slightly improves upon the test result of Huguenin and Jones (1986), which located the absorption band positions at 9500 cm^{-1}, 14,500 cm^{-1},

Figure 6. The second (top), fourth (middle), and fifth (bottom) derivatives of the synthetic spectrum (Fig. 5). The six solid vertical lines indicate the peak positions of the six constituents.

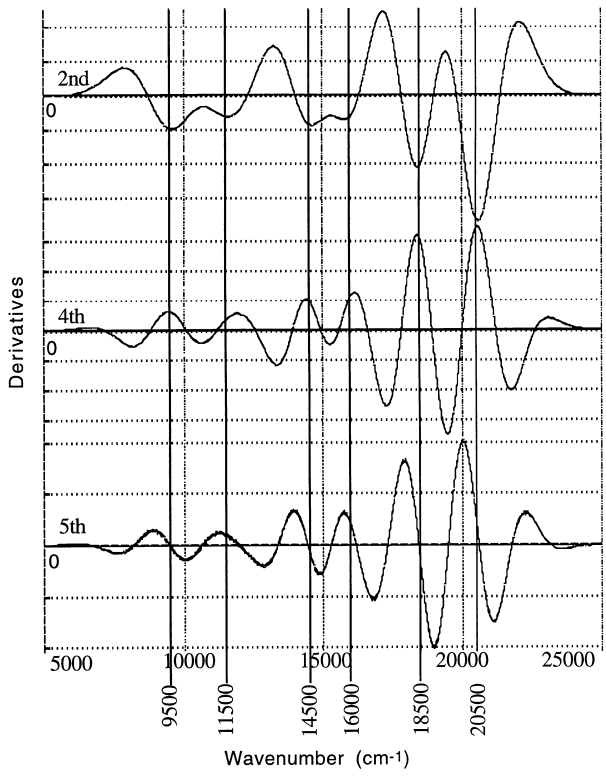


Table 2. Parameters of the Five Constituents of the Synthetic Spectrum Displayed in Figure 7

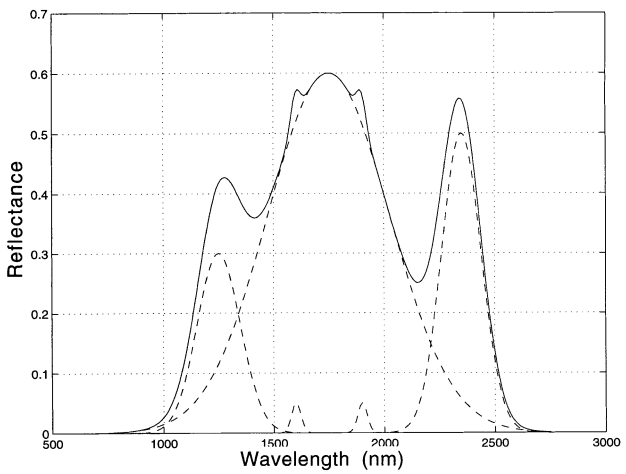
Band	*Peak Position*	*Peak Value (%)*	*Half-Width*
1	1250	30	230
2	1600	5	50
3	1750	60	650
4	1900	5	50
5	2350	50	210

and 20,500 cm^{-1}, but lost the precision at 11,500 cm^{-1} and 20,500 cm^{-1} in the fifth derivative and 16,000 cm^{-1} in the second derivative.

The finite band separation used in the finite approximation algorithm is a potential source of error in the derivative computation. Spectral features that are narrower than the band separation will be lost. In this case, smoothing will further decrease the sensitivity of the finite approximation. On the other hand, as long as the spectral feature of interest is larger than the band separation, it should be detected by the finite approximation derivative, and finer scale noise will be reduced. To test this hypothesis, another synthetic spectrum was constructed as described in Table 2 and Figure 7. This time the computation was performed in terms of wavelength, and the sampling resolution was taken to be 10 nm. It is a spectrum composed of five constituent Gaussian features, two of which (the second and the fourth features) are much smaller in amplitude and width than the other three features.

Assume that the two smallest features are artifacts and that the central peak is the primary feature of interest. Further assume that the critical factor for detecting the feature positions is the zero-crossing of the fifth derivative. Figure 8 shows the results of the fifth derivatives of the synthetic spectrum at band separations of 1, 5, 10, 15, 20, and 30 sampling intervals ($\Delta\lambda$=10 nm, 50 nm, 100 nm, 150 nm, 200 nm, and 300 nm, respectively). From Table 2, the peak positions of two artifacts are 150 nm less and greater than the feature centered at 1750 nm. When the band separation is smaller than the distance between the noise peak and the center of the third band ($\Delta\lambda$=10 nm, 50 nm, and 100 nm), the zero-crossings in the fifth derivative occur almost exactly at the three central peak positions. When the band separation is equal to 150 nm, zero-crossings still occur close to the peak of the artifacts. (Note, as discussed before, that if a derivative occurs at the middle of two adjacent sampling points it is actually assigned to the closest right neighboring point. Therefore, the derivative curves at $\Delta\lambda$=1, 5 and 15 sampling intervals in Fig. 8 are right-shifted.) As $\Delta\lambda$ approaches and then exceeds the distance between the artifact center and the neighboring spectral feature, the effect of the artifact on the derivative first diminishes and then disappears. This appears as a change of the derivative to a simpler, smoother curve as $\Delta\lambda$ increases. This quasi-smoothing effect is a direct result of the wide band separation. As described in Eq. (9), the estimated derivative is proportional to $1/(\Delta\lambda)^n$, where n is the order of derivative. As $\Delta\lambda$ increases, the artifact has less influence on the derivative. If the band separation $\Delta\lambda$ is greater than the distance between a "noise" center and the adjacent feature, the "noise" becomes un-

Figure 7. Synthetic spectrum (solid line) composed of five Gaussian constituents [(---) see Table 2].



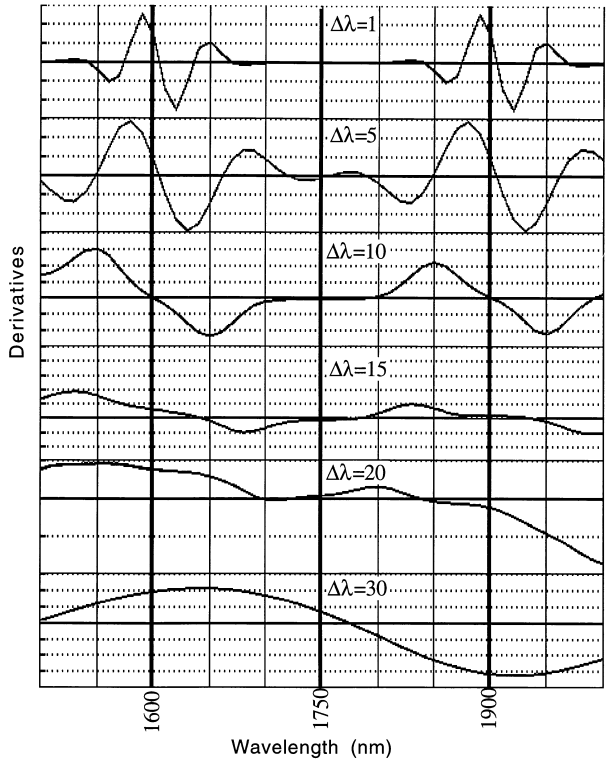


Figure 8. Zero-crossing positions in the fifth derivatives of synthetic spectrum (Fig. 7) at band separation 1, 5, 10, 15, 20, and 30 sampling intervals (10 nm, 50 nm, 100 nm, 150 nm, 200 nm, and 300 nm).

detectable. As a result, in Figure 8, the derivative at 20 sampling intervals (200 nm) loses both "noise" positions.

This scaling effect of $\Delta\lambda$ provides an additional advantage for using the finite approximation algorithm. That is, by selecting a suitable value of $\Delta\lambda$, features smaller than $\Delta\lambda$, whether noise or otherwise, will be lost while features at the scale of $\Delta\lambda$ will be enhanced. The best band separation depends on the scale of features to be detected, the orders of derivatives and the purposes of the analysis. The effect of using $\Delta\lambda$ to match the scale of the desired spectral feature was already illustrated in Figures 2b and 2c.

In real applications, the Huguenin and Jones (1986) criteria for locating absorption band positions can be very helpful in identifying spectral features in complex laboratory or remotely sensed data. Figure 9 shows a portion of a soybean leaf fluorescence spectrum at excitation wavelength of 337 nm (Philpot et al., 1996) and its second, fourth, and fifth derivatives. The plant had been subject to diesel treatment for simulating the effect of diesel exhaust pollution. The spectrum was smoothed using the mean filter algorithm with a filter size of 11 sampling points. Derivatives were then computed using the finite approximation at a band separation of 16 nm (eight sampling intervals for this data set). From this figure, the peak at 600 nm was identified with no difficulty. A second feature, one which was inconspicuous in the original spectrum, was located around 530 nm. This matches the observation of Lang et al. (1992) who suggest that a shoulder of blue-green fluorescence emission occurs at wavelength about 525 nm due to the fluorescence emission from the inner leaf material.

There is ambiguity in detecting the band position at 450 nm. In the fifth derivative, there are four zero-crossings near 450 nm. All four positions fulfill the criteria of the occurrence of an absorption band. However, if the criteria are modified such that the fifth derivative equals zero, the fourth derivative is positive and at or near a local maximum (i.e., within the range defined by FWHM), and the second derivative is negative and at or near a local minimum, the interpretation changes. Applying these expanded criteria to the data in Figure 9, the possibility of a band peak at 490 nm is now excluded because the second derivative is beyond the FWHM of the local minimum, and the occurrence of band position at 450 nm is no longer sustained because the fourth derivative at 450 nm is in a local minimum, although the value is positive. Instead, both 440 nm and 460 nm would be considered as potential absorption band positions, in spite of the fact that there is no evidence in the fluorescence literature supporting the presence of more than one peak around 450 nm in the soybean spectra. It is possible that the ambiguity results from an improper selection of band separation or filter size. According to Figure 9, if the band separation is increased, the local minimum at 450 nm in the fourth derivative is expected

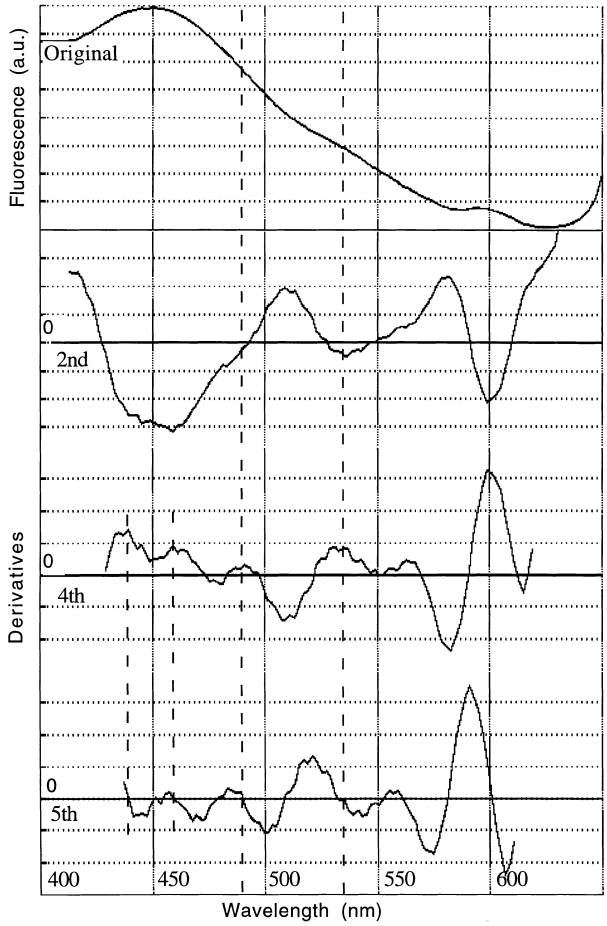


Figure 9. Fluorescence spectrum of a soybean leaf with diesel treatment at excitation wavelength of 337 nm and its derivatives.

to disappear and the two local peaks are expected to incorporate into one peak. Thus, 450 nm would be identified as an absorption band position even according to the expanded criteria.

The test of absorption band position detection in the soybean fluorescence spectrum demonstrates the proposed use of the tools developed in this research: exploring the spectra carefully to discover subtle spectral features and then using other analysis tools to verify the detected potential features. As discussed before, the performance of the developed tools depends on iterative examination of the spectra to select suitable parameters (method for smoothing, filter size, band separation, etc.) for each particular set of spectral data and for each specific purpose of the analysis.

SUMMARY

Hyperspectral derivative analysis on several test spectra for exploring subtle spectral features has been described. The tools developed in this study adopt several smoothing and derivative algorithms commonly used in the field

of spectroscopy. These algorithms are modified for use with remote sensing data rather than typical spectroscopic data. A key issue is to allow users to adapt the filter size to optimize noise reduction and to adjust the effective sampling interval (band separation) to better match the scale of the spectral features of interest.

Filter size is the principal factor affecting the results in the three smoothing methods used in this research. Selection of a suitable filter size for smoothing is spectrum-dependent. In this study, the noise variance used in the Kawata–Minami smoothing algorithm is assumed to be constant. However, the true noise variance is a function of wavelength. Consequently, for spectra with a wide range of wavelengths, the assumption of constant noise variance may be an underestimation or overestimation at certain ranges of wavelength and affects the performance of smoothing.

The finite approximation allows the derivatives to be computed at various band separations. Band separation not only determines the resolution of the results, but also causes a scaling effect unless using the enhanced finite approximation algorithm. The scaling effect is not always a disadvantage. One benefit of the scaling effect is the quasismoothing effect. However, there is always a tradeoff between noise removal and feature extraction. Selecting a sampling interval suitable to the scale of the spectral features of interest will usually enhance the desired spectral features.

The test of absorption band position detection on two synthetic spectra and a spectrum of a soybean leaf exposed to diesel exhaust also suggests the advantage of using this tool to identify spectral features from raw spectral data as well as to demonstrate the interaction between the noise suppression and the change of sampling interval when using finite approximation to calculate the derivative.

Derivative analysis can be an effective tool to analyze hyperspectral data with a different emphasis than traditional remote sensing algorithms. Treating hyperspectral data as truly continuous allows access to information that is often suppressed by the standard analysis methods.

Several improvements and extensions to the tool developed in this research would be worthwhile:

1. Improve the Kawata–Minami smoothing algorithm to incorporate wavelength-dependent noise variances.
2. Modify the Savitzky–Golay smoothing and derivative computation algorithm to allow wider band separations.
3. Develop procedures for selecting appropriate filter size and band separations.
4. Formalize and extend the procedure of spectral decomposition for more general application.
5. Improve the program algorithms and data structure to speed up the process for handling large hyperspectral images.
6. Establish the connections between the analysis modules and an appropriate visualization utility as well as the linkages with other relevant analysis applications.

REFERENCES

Adams, M., Norvell, W. A., and Philpot, W. D. (1995), Fluorescence and reflectance of manganese deficient soybean leaves: effects of leaf selection. In *The Annual Conference of ASA-CSSA-SSSA-CMS*, No. S4-105P.

Butler, W. L., and Hopkins, D. W. (1970), Higher derivative analysis of complex absorption spectra. *Photochem. Photobiol.* 12:439–450.

Chappelle, E. W., Kim, M. S., and McMurtrey, J. E. I. (1992), Ratio analysis of reflectance spectra (RARS): an algorithm for the remote estimation of the concentrations of chlorophyll A, chlorophyll B, and carotenoids in soybean leaves. *Remote Sens. Environ.* 39:239–247.

Chapra, S. C., and Canale, R. P. (1988), *Numerical Methods of Engineers*, 2nd. ed., McGraw-Hill, New York, pp. 93–97.

Chen, Z., and Curran, P. J. (1992), Derivative reflectance spectroscopy to estimate suspended sediment concentration. *Remote Sens. Environ.* 40:67–77.

Curran, P. J., Dungan, J. L., Macler, B. A., Plummer, S. E., and Peterson, D. L. (1992), Reflectance spectroscopy of fresh whole leaves for the estimation of chemical concentration. *Remote Sens. Environ.* 39:153–166.

Demetriades-Shah, T. H., Steven, M. D., and Clark, J. A. (1990), High resolution derivatives spectra in remote sensing. *Remote Sens. Environ.* 33:55–64.

Dick, K., and Miller, J. R. (1991), Derivative analysis applied to high resolution optical spectra of freshwater lakes. In *The 14th Canadian Symposium on Remote Sensing*, Calgary, Alberta, Canada.

Duda, R. O., and Hart, P. E. (1973), *Pattern Classification and Scene Analysis*, Wiley, New York.

Gamon, J. A., Peñuelas, J., and Field, C. B. (1992), A narrow-waveband spectral index that tracks diurnal changes in photosynthetic efficiency. *Remote Sens. Environ.* 41:35–44.

Hoffbeck, J. P., and Landgrebe, D. A. (1993), Classification of high dimensional multispectral image data. In *The Fourth Annual JPL Airborne Geoscience Workshop*, Washington D.C.

Huguenin, R. L., and Jones, J. L. (1986), Intelligent Information extraction from reflectance spectra: absorption band position. *J. Geophys. Res.* 91(B9):9585–9598.

Kawata, S., and Minami, S. (1984), Adaptive smoothing of spectroscopic data by a linear mean-square estimation. *Appl. Spectrosc.* 38(1):49–58.

Lang, M., Siffel, P., Braunová, Z., and Lichtenthaler, H. K. (1992), Investigations of the blue-green fluorescence emission of plant leaves. *Bot. Acta* 105:435–440.

Madden, H. (1978), Comments on the Savitzky–Golay convolution method for least-squares fit smoothing and differentiation of digital data. *Anal. Chem.* 50(9):1383–1386.

Martin, M. E., and Aber, J. D. (1993), Measurements of can-

opy chemistry with 1992 AVIRIS data at Blackhawk Island and Harvard Forest. In *The Fourth Annual JPL Airborne Geoscience Workshop*: AVIRIS (R. O. Green, Ed.). Jet Propulsion Laboratory, Pasadena, CA, Vol. 1, pp. 113–116.

Mathworks (1993) *MATLAB: High Performance Computation and Visualization Software*, The Mathworks, Natick, MA.

Peñuelas, J., Gamon, J. A., Fredeen, A. L., Merino, J., and Field, C. B. (1994), Reflectance indices associated with physiological changes in nitrogen- and water-limited sunflower leaves. *Remote Sens. Environ.* 48:135–146.

Philpot, W. D. (1991), The derivative ratio algorithm: avoiding atmospheric effects in remote sensing. *IEEE Trans. Geosci. Remote Sens.* 29(3):350–357.

Philpot, W., Duggin, M., Raba, R., and Tsai, F. (1996), Analysis of reflectance and fluorescence spectra for atypical features: fluorescence in yellow-green. *J. Plant Physiol.* 148:567–573.

Richards, J. A. (1993), *Remote Sensing Digital Image Analysis*, 2nd ed., Springer-Verlag, New York.

Savitzky, A., and Golay, M. J. E. (1964), Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36:1627–1639.

Steinier, J., Termonia, Y., and Deltour, J. (1972), Comments on smoothing and differentiation of data by simplified least squares procedure. *Anal. Chem.* 44(11):1906–1909.

Talsky, G. (1994), *Derivative Spectrophotometry*, Weihein, Verlagsgesellschaft, Weihein, Germany.