Aceh International Journal of Science and Technology, 1(1): 30-39
April 2012
ISSN: 2088-9860

Functional Data Analysis of Multi-Angular Hyperspectral Data on Vegetation

[1]Sugianto, and [2]Shawn Laffan
[1]Remote Sensing and GIS Laboratory, Syiah Kuala University, Banda Aceh 23111, Indonesia,
[2]School of Biological Earth and Environmental Science, The University of New South Wales, Australia.

Abstract - The surface reflectance anisotropy can be estimated by directional reflectance analysis through the collection of multi-angular spectral data. Proper characterization of the surface anisotropy is an important element in the successful interpretation of remotely sensed signals. A signal received by a sensor from a vegetation canopy is affected by several factors. One of them is the sensor zenith angle. Functional data analysis can be used to assess the distribution and variation of spectral reflectance due to sensor zenith angle. This paper examines the effect of sensor zenith angles on the spectral reflectance of vegetation, example on cotton leaves. The spectra were acquired in a green house trial in order to address the question 'how much information can be obtained from multi-angular hyperspectral remote sensing of vegetation?' The goals of the functional data analysis applied in this paper is to examine the Functional Data Analysis approach was applied to analysis multi-angular hyperspectral data on cotton, highlighting various characteristics of cotton spectra due to sensor view angles, and to infer directional variation in an outcome or dependent variable with different zenith angles.

Keywords: Functional data analysis, multi angular hyperspectral, zenith angle.

Introduction

The physical interpretation of hyperspectral remote sensing for vegetation has been hampered by anisotropic surface reflectance (Bruegge *et al.*, 2000). This is because vegetation properties are not generally perfectly diffuse reflectors, and sunlight reflected from vegetation exhibits a significant degree of anisotropy (Settle, 2004). The surface reflectance anisotropy can be estimated by directional reflectance analysis through the collection of multi-angular spectral data (Chooping *et al.*, 2003; D'Urso *et al.*, 2003; Widen, 2004). Proper characterization of the surface anisotropy is an important element in the successful interpretation of remotely sensed signals. However, proper characterization of surface anisotropy needs perfect measurement of illumination and modelling. This is not an easy task in remote sensing. It involves consideration of factors such as land surface properties, sensor view angles, reflectance and scattering, and sun incident angles.

This problem has received attention in remote sensing through research on multi-angular measurement of hyperspectral data. The effect of sensor viewing geometry on vegetation spectra can be assessed by measuring the spectral reflectance of the target at different sensor positions. A handheld spectroradiometer can be mounted on a goniometer, a hemispherical structure, to simulate multiple viewing angles and sensor positions. The importance of directional spectral reflectance is that different zenith angles can create variation in spectral reflectance, inhibiting the analysis of the spectral characteristics.

Recent developments in multi-angular space borne hyperspectral remote sensing provide directional spectral reflectance of the vegetation canopies. The first step, however, is to measure the multi-angular spectral reflectance at ground level. Crucial factors that need to be considered when collecting multi-angular hyperspectral data are how the data is distributed along the wavelength, the variance of the data, and what additional information can be generated by collecting spectra from multiple angles.

Data obtained with a spectroradiometer can be used to model the spectral reflectance. However, there are still assumptions that need to be considered in dealing with spectral data collected from the field such as variation and distribution of the data due to viewing sensor geometry. Sensor viewing geometry is related to the sun illumination. Handling this factor can be achieved by measuring multi-angular spectral reflectance and analysing them in functional relationship between factors affecting them and the spectral reflectance characteristics of the targets. The functional data analysis implemented in this research includes an assessment of the effect of smoothing basis functions, functional principal component analysis, functional linear modelling and functional analysis of variance (Ramsay and Silverman 1997, 2002) There are two types of analysis presented in this paper, wavelength basis and zenith- angle basis.

Materials and Methods

Goniometer setting and cotton spectral collection

A field goniometer was used with mounted of spectradiometer. A goniometer is a mechanised device that characterises radiance as a function of angle and allows the operator to acquire spectra at different zenith angles. The spectroradiometer was put one meter above the canopy attached to goniometer with a 10° instantaneous field of view (IFOV). The spectroradiometer used is an Analytical Spectral Devices (ASD) Fieldspec UV-NIR CCD spectroradiometer (ASD, 1999). The ASD spectroradiometer records a spectrum from 325 - 1075 nm (Visible to Near Infra Red wavelenght). This device is post-dispersive, meaning it can be exposed much higher levels of ambient light outside laboratory controlled conditions. The average band interval is 1.5 nm. For the purposes of this research, a subset of spectra is used, selecting the spectral wavelength range 400 to 950 nm (412 bands).

The cotton plant used in the trial is Siokra V-16 cultivar. This cultivar is a conventional and non-genetically modified variety (CSD, 2003), The cultivar were grown in a greenhouse experiment at the University of New South Wales. The temperature and humidity were maintained at 18-30°C with 55-70% humidity during the experimental growing season. The 25 cm diameter-pots were used to grow the cultivar. Watering was monitored to replicate water consumption during the growing period of the cotton cultivar at the field at Iffley Farm, between 410 mm and 610 mm during a seven-month growing period. Water consumption of cotton increases significantly through the plants' development, during vegetative and generative period, and then decreases near harvest time. The cotton spectra were collected using the ASD spectroradiometer attached to the field goniometer (Figure 1). Calibration was performed for each measurement using a spectralon® standard reflectance (ASD,1999).

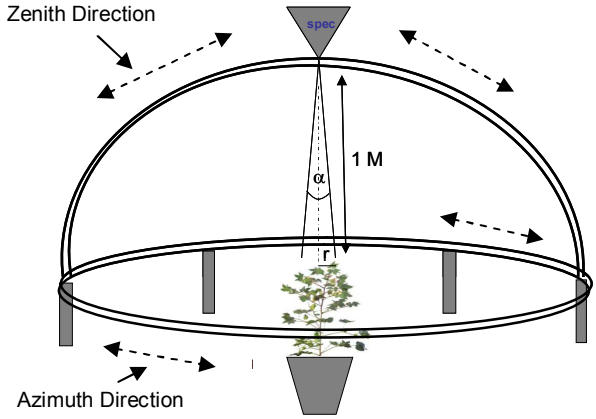
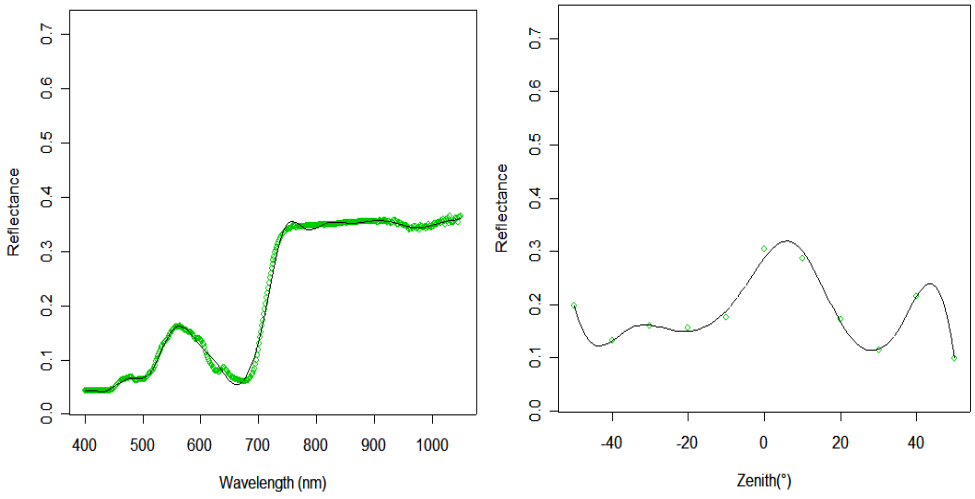


Figure 1. Illustration of the goniometer set up for multi-angular spectral data collection

To prevent unwanted background scattering and absorption from the surrounding area, a black plastic cover was placed underneath the canopy and over the surrounding area to minimise reflectance from the surrounding area. The spectroradiometer was located one metre above the foliage. A 10° spectroradiometer foreoptic was used to create a FOV area of 240.46 cm². Spectral reflectance readings were taken as the average of 25 readings and repeated five times for each view angle position relative to the target. The azimuth and zenith angles position of the sensor were determined using a compass and clinometers attached to the spectroradiometer.

Spectral measurement was conducted under natural light conditions between 10 AM and 2 PM during clear sky conditions. The hot spot effect was avoided by not measuring the spectra where sensor zenith angle and solar zenith angle were at the same position. The solar azimuth angle was assumed to have small variation during spectral data collection in this time. The direction of zenith angles to collect multi-angular reflectance was set facing northeast (45° azimuth) to southwest (225° azimuth). The angle between northeast to nadir is denoted using a positive sign and nadir to southwest direction is denoted using a negative sign. The zenith angles are 50°, 40°, 30°, 20°, 10°, 0°, -10°, -20°, -30°, -40°, and -50°, where 0° is the nadir position.

Two factors need to be considered with the analysis of spectral reflectance data during functional data conversion, wavelength and zenith angle basis function analyses. The wavelength basis function treats the reflectance value as a function of bands (or channels). For the wavelength basis, the *x* value is the wavelength and *y* value is the reflectance. For the zenith angle function, the *x* axis is the angle and the *y* axis is the reflectance. Figure 2 illustrates the difference between the analysis of wavelength and zenith angle basis functions. The wavelength functions treat the wavelength as the major component that affects the reflectance for individual zenith angles. Thus, every band can be retrieved in the plot at any individual zenith angle, such as at -50°, 50°, 0°, 40° and so on. Zenith basis functions assume that the zenith angle is the main factor that affects the reflectance for individual bands. We used a 20-basis function to represent the functional spectra for wavelength basis analysis, and 5-basis function to represent the functional spectra for zenith function for the January and February data.



(a) (b)

Figure 2. Illustration of functional curve representation of the hyperspectral data, wavelength (a), zenith basis (b) basis analysis

Functional data analysis routines

The descriptive statistic functions consist of the mean and standard deviation functions of the cotton spectra was calculated. The mean and standard deviation of the two data sets were plotted to show the variance of the curve functions of the cotton spectra. Variance-covariance functions and the functional derivatives of the cotton spectra was generated in FDA (Ramsay and Ramsay, 2002; Ramsay and Silverman, 1997). The purpose of the descriptive statistics is to assess the nature of the data by observing the mean value and standard deviation function plots, variance-covariance function, and first and second derivative of the cotton spectra acquired with different zenith angles. The derivative of cotton spectra was set for wavelength basis analysis. The variance-covariance function was calculated for sensor zenith basis analysis.

The basis function for spectral smoothing was applied using B-spline smoothing in FDA to the wavelength basis analysis (413 values), but not for zenith basis analysis due to the limited number of data points (11 angles). In order to examine the smallest factors of smoothing for spectral curves of vegetation, different numbers of basis function (*K)* were assessed namely 5, 10, 15, 20, 30, 50 and 100 for the wavelength analysis. Then, the effect of the smoothing was examined by visual comparison of the smoothed spectra to the original spectrum, and by comparing the residual noise spectrum with each smoothing method to find the best fit to the data.

The four principal components along with the residual and harmonic data were calculated. The main objective is to demonstrate the features of the functional principal component analysis results by recounting the variance value for each principal component of the cotton spectra, which, when plotted, will expose the dynamics of multi-angular spectra. A functional linear model can be used to make predictions, and so it was applied to the cotton spectral data to assess the effect of azimuth and zenith angles on spectral reflectance by comparing the functional linear curve, and followed by functional Analysis of Variance (fANOVA) to test the significance of the model. Direction of scanning (e.g. forward (negative zenith) and backward) from the nadir positions was used to display the directional spectra in a functional linear model. It is assumed that the different sensor positions at each zenith angle will relate to different spectral reflectance at nadir. This method was applied for wavelength basis analysis.

Results and Discussion

The effect of sensor zenith angle on spectral plots

The spectral reflectance signature of the cotton spectra acquired at different sensor zenith angles shows a vertical 'shift' (increase or decrease) of spectral reflectance values in the visible to NIR. Some wavelength regions above 900 nm have excessive noise. This spectral signature consists of 413 bands and 11 different zenith angles. It is clear that forward direction (positive angles) tend to have higher spectral reflectance value in the *Red-NIR* regions. The functional spectral reflectance for all eleven-zenith angles with a total of 412 bands for both January and February data, fitted with 20-basis functions, are shown in Figure 4. Each curve represents a functional data object for each zenith angle (different color and line types) of cotton spectra for the spectrum range and gives a trend of spectral reflectance over different angles.

There is variance across the zenith angles, with lower variance in the January data when compared to February data. For the January data, at zenith angles 0° (nadir), -10° and 10°, the first three functional spectra have high reflectance in the red to near infrared wavelengths (the red-edge) (see Figure 4). For the data from February, 0°, -10° and -50°, all show a shift in spectral reflectance in the *Red –NIR*. When compared, the 50° azimuth of the January and February data is similar, less than 0.4 or 40% reflectance for the cotton spectra. Noise is visible at above 900 nm, particularly for the February data. As is common for green vegetation, reflectance in the visible range was quite low, with peak response in the green. In the case of data collected from different angles using the goniometer, the range in reflectance for the visible region is important, showing directional variation in the visible to near infrared region.

The functional mean and functional standard deviation for the 11-zenith angles is shown in Figure 5. The functional mean and functional standard deviation for both datasets are similar except for a slight difference at 900 nm in the January data. Both datasets have the same tendency and distribution, which is not surprising because the spectral results come from the same plant at different dates of acquisition. The functional derivative analysis of cotton plant spectra indicates spectral variation due to sensor zenith angles. The reflectance at +10° and 0° (nadir) are higher for both dates of spectral collection (Figure 6). The first derivative has shown some critical absorption of the wavelength for cotton for vegetation discrimination such as green, red and near infrared regions.

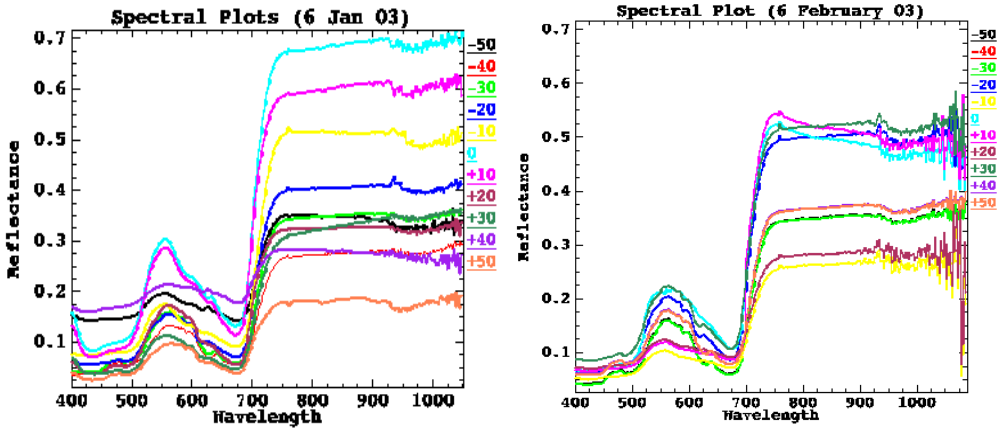


Figure 3. Typical spectral signature of cotton plant acquired at difference zenith angles before the FDA process.

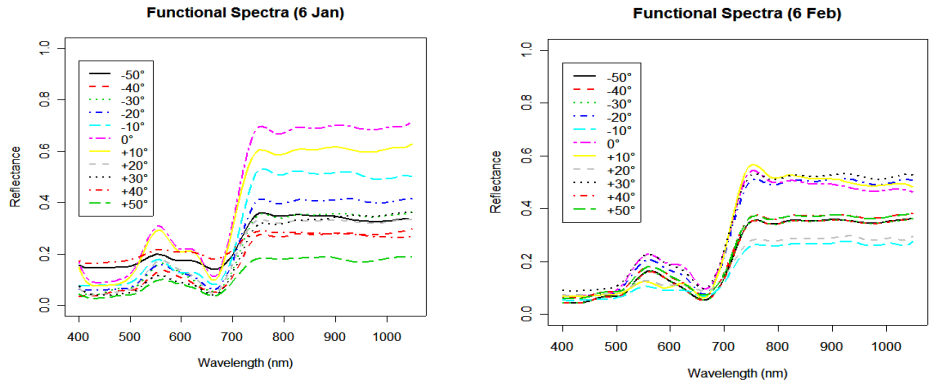


Figure 4. Spectral functional at different zenith angles using 20-basis for January and February data

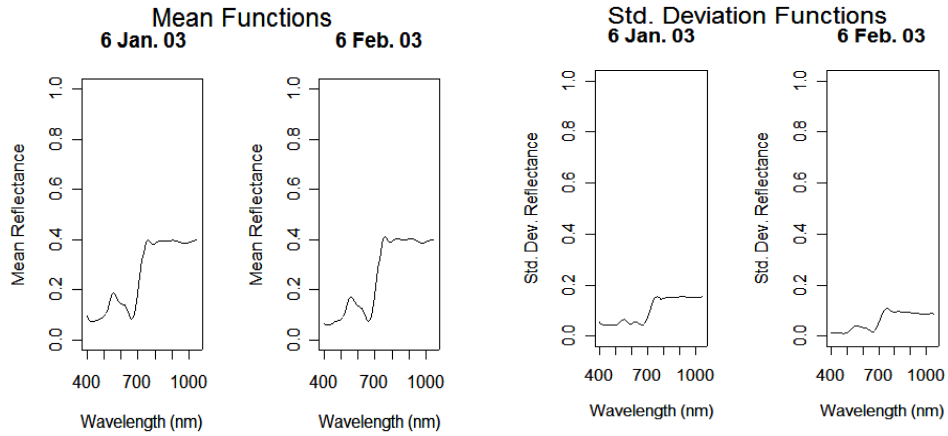


Figure 5. Functional mean and standard deviation of cotton spectra across the eleven zenith angles.

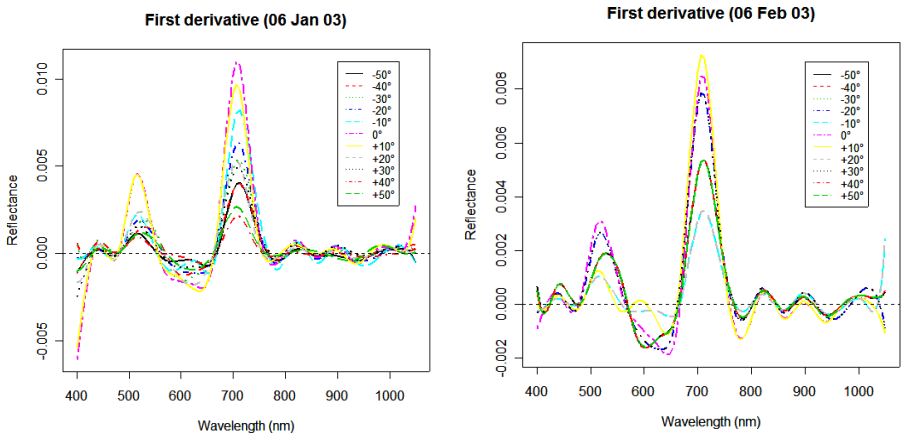


Figure 6. First derivative curve showing spectral variables for different zenith angles of cotton spectra acquired on 06 January and 06 February 2003

Basis function for spectra smoothing

Figures 7 and 8 show the smoothing results of multi-angle cotton spectra using different basis functions. The lower limit function determines the lowest function that can be applied to the data sets and still affect the smoothing basis, but not perfectly fit the dataset. The upper limit determines the number of basis functions that can show a perfect fit to the dataset, but not exceed the number of data points that need to be smoothed. In this experiment, the lowest basis function that can be applied to the data is five. Applying the 5-basis function show very smooth curve along the spectrum range, and the spectral data was shifted to a model far from the real data value, resulting in a poor representation of the absorption features of the vegetation spectra. Using 10-basis functions, the curve starts to follow the real data values. The function that is smoothed has limited curvature, depending on the number of basis functions used. The greater the basis used up to 20-basis, the closer the curve line follows the original data values. The 50 basis curves provide no more variation than the 20-basis curves. In this example, the 50-basis function can be considered as the upper limit for this data set. It is still possible to apply higher bases, but a basis of 412 must not be exceeded as this is the maximum number of range data value for the dataset. Since the plots (Figures 7 and 8) are on the same scale, it can be seen that the range of variation in the polynomial closely matches the range of variation in the data. With the use of a 20-basis function, the line model fits approximately well to the real data, but there are still some points where the line model strays from the real data values. The use of a 30-basis function, the line model fits well to the real data. In addition, a 30-basis function is not much different to 50- and 100- basis functions. The 20-basis is sufficient to represent smoothed spectra.

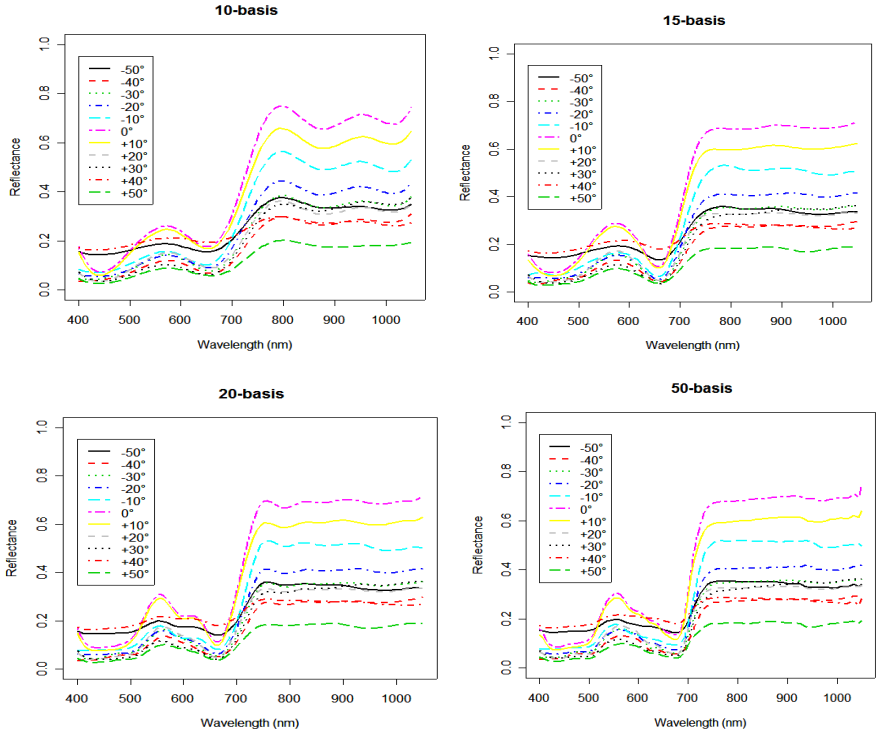


Figure 7. Spectral curves with different basis function models (January data), the 20 basis function curves give a good approximation of the data.

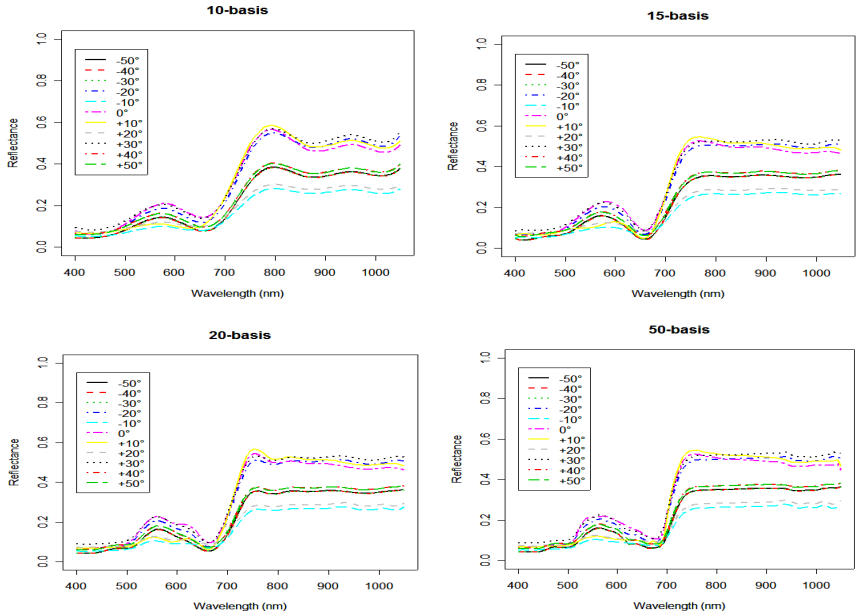


Figure 8. Spectral curves with different basis function models (February data), the 20 basis function curves give a good approximation of the data.

The fitting basis function of spectral data (Figures 9 and 10) shows the construction of a smoothing curve of B-splines with given coefficients. The use of the same basis functions (5-basis) on different zenith angles shows a variation in the residual fit (Figure 9) from 0.05 at Znt 8 (-40°) to 0.09 at Znt 6. The results show that different residual numbers could determine the level of fit to the particular zenith angle. These findings indicate that the specific basis function needed for smoothing purposes depends on two factors: (1) the level of noise or variations of spectral data determines what basis function is required. More variation in the spectrum requires greater basis functions in order to obtain a better fit for the smoothed model to the original data, (2) the selection of basis for individual spectra can determine how well the smoothing line fits the data. The greater the number of basis functions used for a particular spectrum, the better the approximate models fits the data.

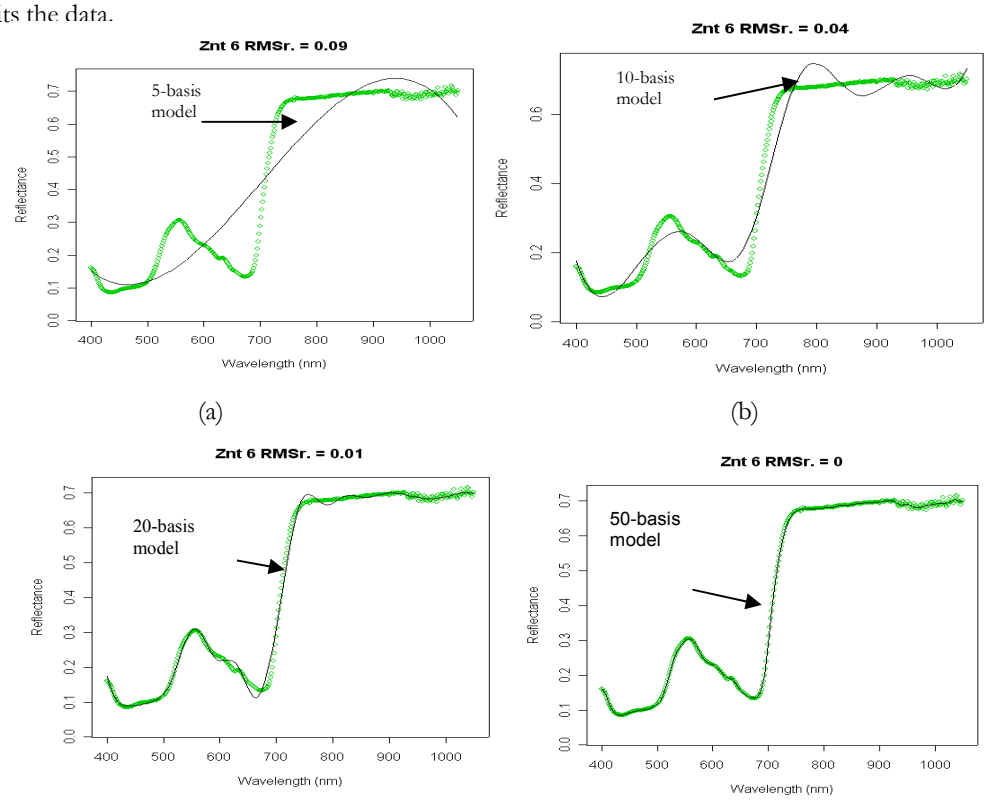
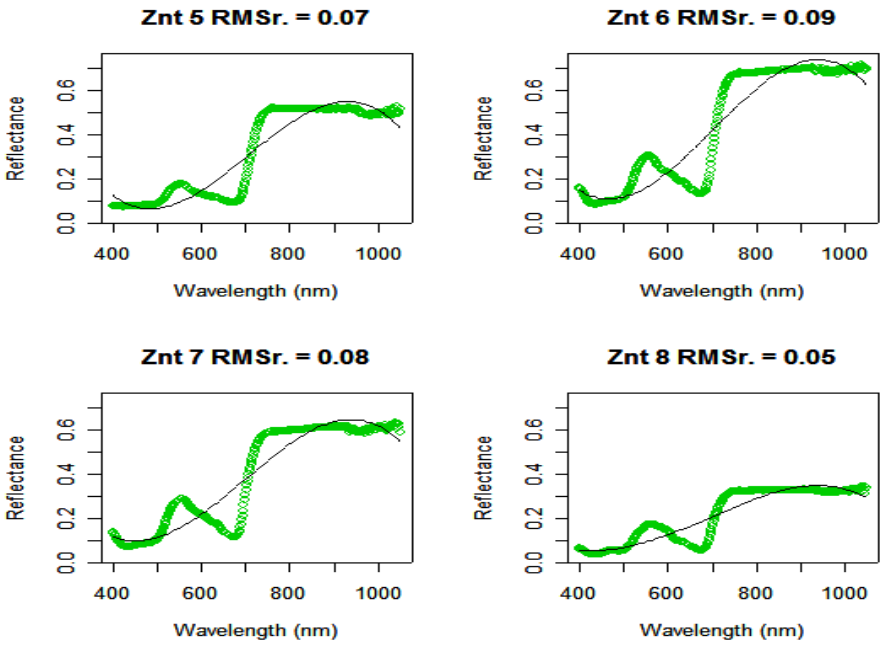


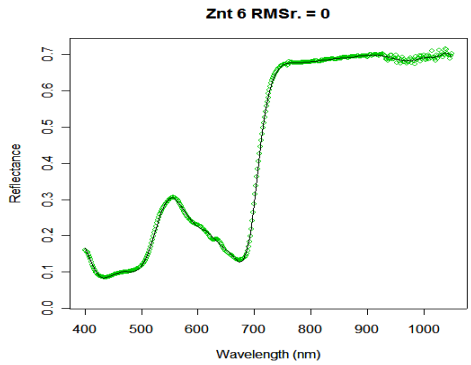
Figure 9. Fitting the curve at 0° (nadir) view angle (Znt 6) for different basis; 5-basis (a), 10-basis (b), 20-basis (c), and 50-basis (d) (January data. RMSr=Root mean square residual.

Figures 10 shows the comparison of fitting spectral curves using 5-basis function for the four zenith angles(*Znt 5, Znt 6, Znt 7, Znt 8*, representing -10°, 0°, +10° and +20° respectively) and 30-basis function for smoothing at *Znt 6* (0° nadir). The use 30-basis function for all zenith angles has zero residual for the January spectra. This result shows that level of noise due to the sensor view angles could be identified by examining the residual. The noise not only comes from sensor, but also from environmental conditions that may affect the spectral reflectance. In this example, the February spectra are noisier than the January spectra. The use of 30-basis functions on the January data displays no residual, but the February data still has residual value with 30 basis functions. The 30-basis function was used for the remaining analyses in this chapter.

Through functional basis analysis, a 30-basis function was found to be the best for smoothing the data with minimal residual. Using lower a basis than 30 for the cotton spectra provided a poorer fit of the functional model to the original data. Values above the 30-basis function resulted in no further change to the approximate functional model. Thus, more than 30-basis is considered overfit to the curve. Various basis function test worked for wavelength basis analysis. For zenith basis analysis, 5-basis function still works well in representing an approximation of the basis function model. However, fitting the model to the original data was not properly fitted. The main important point for basis function for smoothing of cotton spectral data is that the smoothing using basis function can be estimated by inverting mathematical models of the Bidirectional Reflectance Different Function against directional reflectance data sampled at several different sensor view angles as with the soils.



(a)



(b)

Figure 10. Fitting the curve at different view angle (Znt 5-to 8; 50°, 0°, -50° and -40°) for the same basis (5-basis) (a). The 30-basis functions for Znt 6 (0°) (b) show zero residual; a perfect fitting for January data.

Functional principal component analyses (FPCA)

Figure 11 shows the bivariate plot of the principal component score for the first four principal components of January and February data; the second principal component was plotted against first, third against first, fourth against first and second against third. The harmonic values, a summary statistic used in analyses of frequency data display a clearly unordered pattern. Four principal component directions (eigenvectors) and the percent of total variation of each direction are shown. In each case, the solid curve (black) is the overall mean spectra for the given wavelength range and the circle (green) and dashed (red) curves shows the effect of adding and subtracting a multiple of each principal component to clarify the effect of the components.

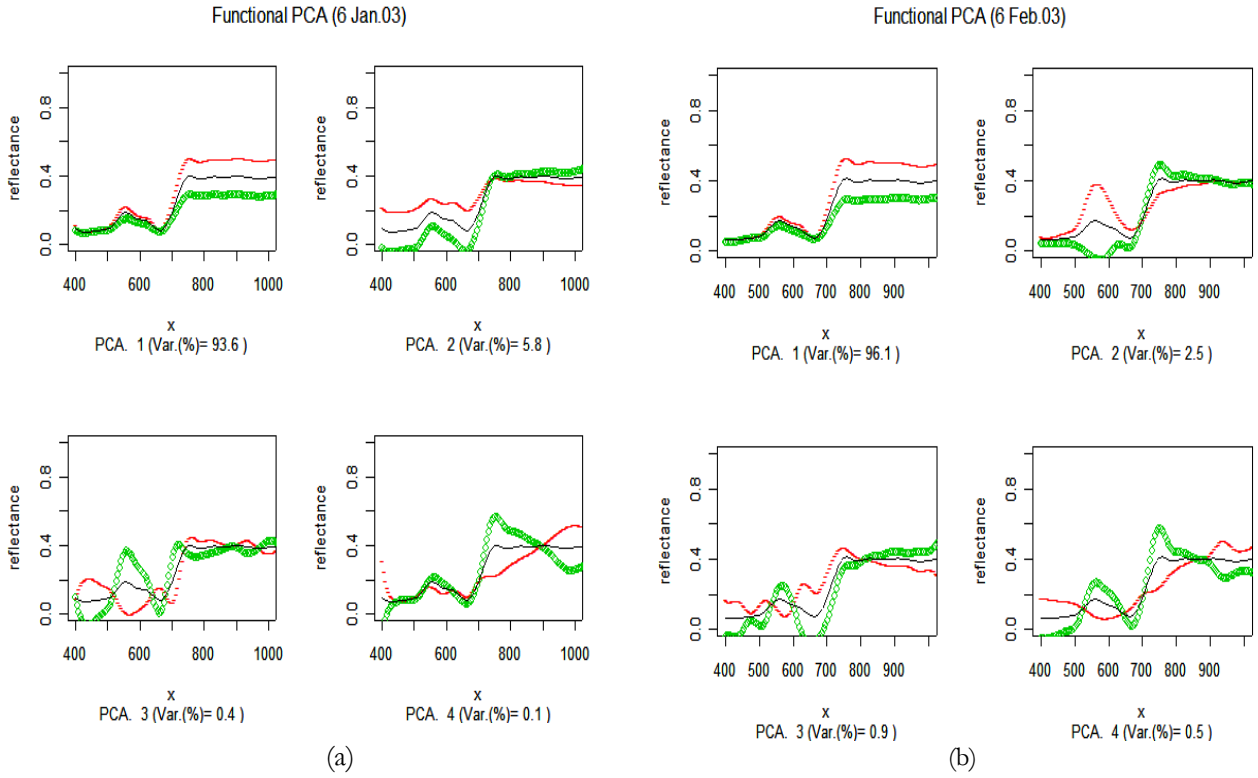


Figure 11. Results of the fPCA (PCA 1 to PCA 4) on the multi-angular spectral data on cotton. Note the first principal component direction of first fPCA explains most of the variation 93.6% (January) (a), and 96.1% (February) (b) in the data. The shape of first fPCA indicates that most of the variation of the data occurs at the two angle measurements.

The first principal component direction, fPCA, explains most of the variation, about 93.6% and 96.4% respectively for the January and February data. The variation occurs for all wavelengths, especially in the 700 to 900 nm. The fPCA value is essentially the common spectral variation over different zenith angles, and is due to the specific reflectance or absorption along the wavelength range for vegetation. The second principal component function, which accounts for 5.8% and 2.6% respectively for January and February, picks up the differences between high and low reflectance values regarding the NIR regions. The third and the fourth components account for less than one percent of the variation. These could be spurious and therefore do not provide useful information, contributing to error effects in the spectral analysis. An important aspect of PCA is the examination of the score function of each curve on each component. Each angle is identified by the level given to the angles, as described earlier. Some positions of the score have been adjusted to the same degree of score component-scale to improve legibility. Variation of spectral data was almost indistinguishable from the harmonic score values. The -10°, 10° and 0° appear in the same order for the first three harmonic comparisons. The rest of data points are clustered in the adjacent values, except for harmonic 2 versus harmonic 3 (Figure 12). The -10°, 0° are in the upper right corner for January data because they have a higher reflectance value in the *Red-NIR* than most of the other zenith angles (on PCA 1). Similarly, the -50°, 20°, 0° and -30° February data are higher in spectral reflectance in this region than most of the other zenith angles.

The results of functional principle component analysis of the wavelength basis analysis vary for each date of collection and sensor zenith angles. The first functional principle component of zenith angles accounts for 93.6% and 96.4% respectively for January and February data. The variation of the first fPCA was high in the NIR reflectance. This region basically is sensitive reflectance location for most green vegetation. The second fPCA accounts for 5.8% and 2.6% respectively for January and February, picking up the differences between high and low reflectance values with respect to the NIR regions.

This second fPCA still provide the information of spectral reflectance, especially below 700 nm wavelengths. Vegetation pigment and chloroplast play important role in this region (Lichtenthaler *et al*., 1996; Yoder and Pettigrew-Crosby, 1995). Thus, the second fPCA provides important information on plant physiology. The third and the fourth fPCA account for less than one percent of the variation. These results could be spurious and so do not provide useful information which is contributed to a large error effect in the spectral analysis. The key point of using fPCA for January and February data of cotton spectra is that the variation for both data is not much different in the distribution of variance explained for the first four fPCA. February data has a slightly higher variance for the first fPCA.

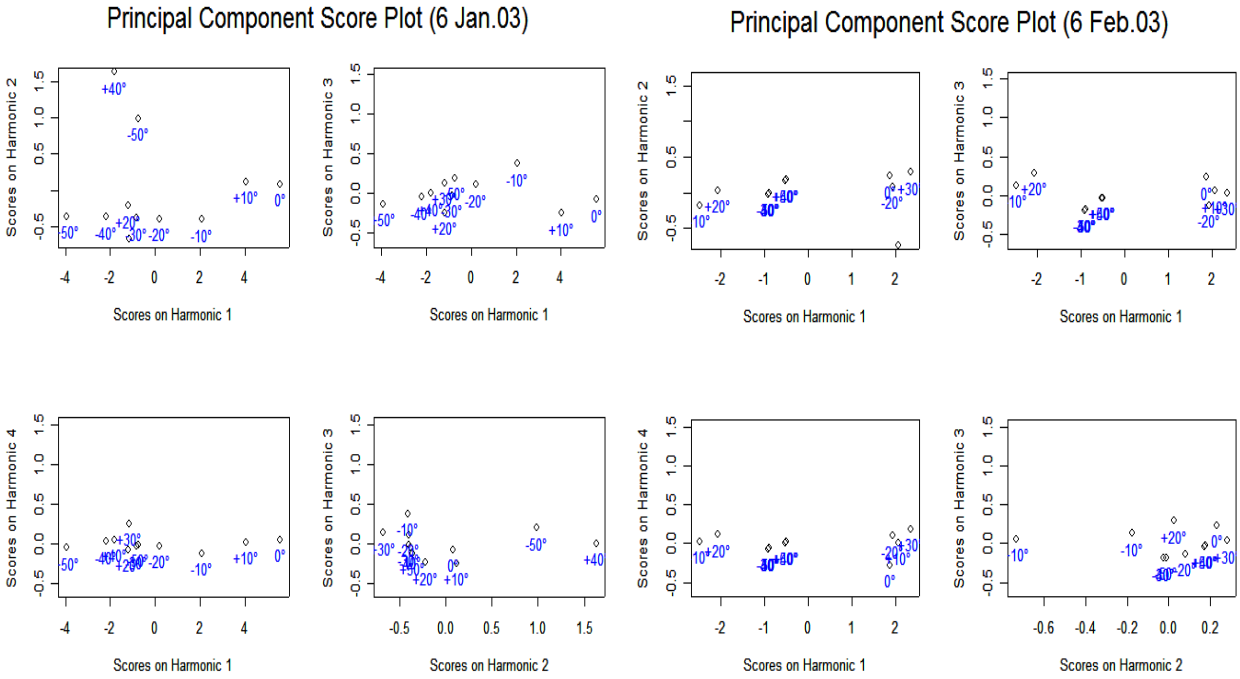


Figure 12. Principal component scores of second fPCA versus first fPCA, third fPCA versus first fPCA, fourth fPCA versus first fPCA and second fPCA versus third fPCA for January and February 2003 data.

Conclusions

Various characteristics of cotton spectra due to sensor view angles have been considered in the results, which indicate directional variation as an outcome of the dependent variable with different zenith angles. Two principal approaches have been used to represent cotton plant spectra in the FDA process: wavelength and zenith angles. These result in different curve function output, the basis function, functional analysis of variance and the linear model of multi-angular cotton spectra. The spectral noise produced by the sensor presented in a smoothed curve approximation has different RMS residual, depending on the basis function and zenith angles. The use of different numbers of basis functions clearly affects the fit of the spectral curves obtained from eleven-zenith angles. One implication of the findings from this directional analysis of multi-angular spectra of cotton is that the spectral reflectance data can be considered as functions. By using a functional approach for modelling, the smoothing, predicting and analysis of the cotton spectra can be achieved. However, the main implication is that there is only a very small amount of additional information provided by multi-angular measurements within the range of zenith angles assessed.

Acknowledgements

We thank to Geoff Mcdonald for helping me during this research. Yanni Zakaria and John Owen for their assistance in computer matters at the University of New South Wales.

References

ASD (1999). Analytical spectral devices. Inc *(ASD),* Technical Guide, Boulder ,Colorado

Bruegge, C.J., Helmlinger, M. C., Conel, J.E., Gaitley, B.J., and Abdou, W.A. (2000). PARABOLA III: a sphere-scanning radiometer for field determination of surface anisotropic reflectance functions. Remote Sensing Reviews, 19:75-94.

Chooping, M. J., Rango, A., Havstad, K. M., Schiebe, F. R., Ritchie, J. C., Schmugge, T. J., French, A. N., Su, L. H., McKee, L., and Davis, M. R. (2003). Canopy attributes of desert grassland and transition communities derived from multi-angular airborne imagery. Remote Sensing of Environment, 85(3): 339-354.

CSD (2003). CSIRO cotton breeding program. Cotton Seed Distribution Ltd., Narrabri New South Wales, Australia

D'Urso, G., Dini, L., Vuolo, F., Alonso, L., and Guanter, L. (2003). Retrieval of leaf area index by inverting hyperspectral, multi-angular CHRIS/PROBA data from SPARC 2003. In ESA/ESRIN (Ed.), Proceedings of 2nd CHRIS/Proba Workshop, (pp. CD-room). Frascati, Italy, ESA.

Lichtenthaler, H. K., Gitelson, A. A., and Lang, M. (1996). Non-destructive determination of chlorophyll content of leaves of a green and an aurea mutant tobacco by reflectance measurements. Journal of Plant Physiology, 148: 241-252.

Milton, E. J. (1987). Principles of field spectroscopy. International Journal of Remote Sensing, 8(3): 1807-1827.

Ramsay, J. B., and Dalzell, C. (1991). Some tools for functional data analysis. Journal of the Royal Statistical Society, 53: 539-572.

Ramsay, J. O. (1996). Principal differential analysis: data reduction by differential operators. Journal of the Royal Statistical Society Series B, 58: 495-508.

Ramsay, J. O., and Ramsay, J. B. (2002). Functional data analysis of the dynamics of the monthly index of nondurable goods production. Journal of Econometrics,.107: .327-344.

Ramsay, J. O., and Silverman, B. W. (1997). Functional data analysis. Springer, New York.

Ramsay, J. O., and Silverman, B. W. (2002). Applied functional data analysis. Springer, New York.

Settle, J. (2004). On the dimensionality of multi-view hyperspectral measurements of vegetation. Remote Sensing of Environment, 90(1): 235-242.

Walthall, C. L., Norman, J. M., Welles, J. M., Campbell, G., & Blad, B. (1985). Simple equation to approximate the bidirectional reflectance from vegetated canopies and bare soil surfaces. Applied Optics, 24: 383-387.

Widen, N. (2004). Assessing the accuracy of land surface characteristics estimation from multi-angular remote sensed data. International Journal of Remote Sensing, 25(3): 1105-1117.

Yoder, B. J., and Pettigrew-Crosby, R.E. (1995). Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400–2500 nm) at leaf and canopy scales. Remote Sensing of Environment, 53(3): 199–211.