N95- 33744

CAUSAL CORRELATION OF FOLIAR BIOCHEMICAL CONCENTRATIONS WITH AVIRIS SPECTRA USING FORCED ENTRY LINEAR REGRESSION

Terence P. Dawson and Paul J. Curran
Department of Geography, University of Southampton, Southampton SO17 1BJ, UK.

John A. Kupiec
Scottish Natural Heritage, 2 Anderson Place, Edinburgh EH6 5NP, UK.

"Unguided multiple linear regression methods lack consistency in wavelength selection for a given constituent and can result in over-fitting of the data" (Aber, 1994).

1.0 Introduction

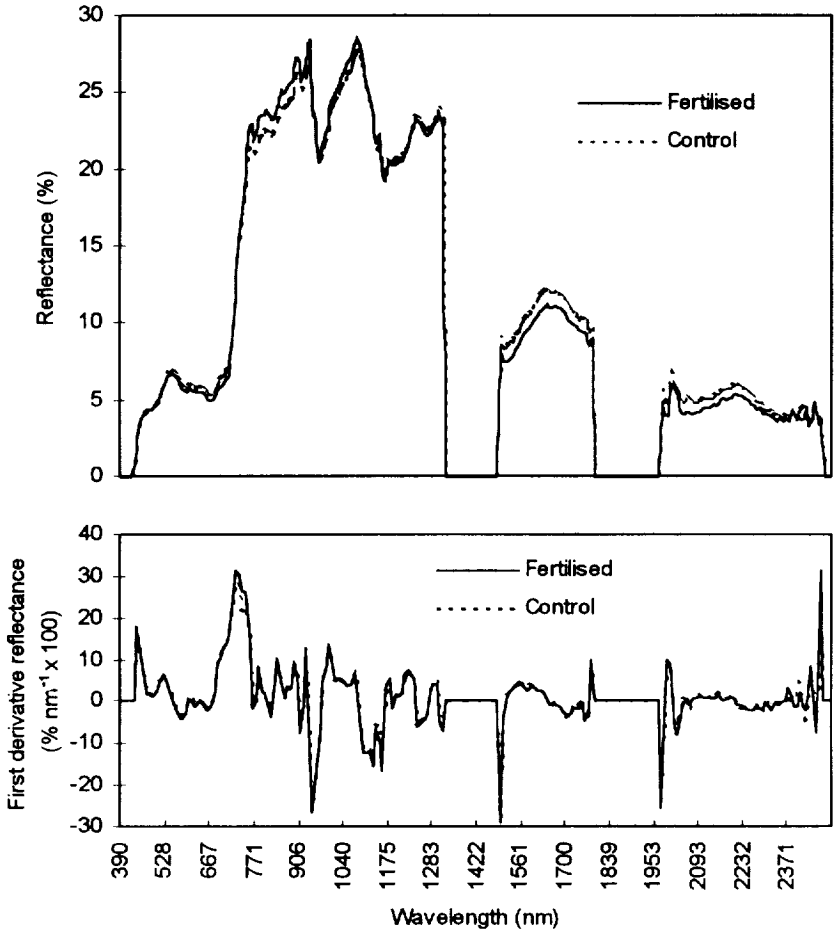
A major goal of airborne imaging spectrometry is to estimate the biochemical composition of vegetation canopies from reflectance spectra (Curran, 1994). Remotely-sensed estimates of foliar biochemical concentrations of forests would provide valuable indicators of ecosystem function at regional and eventually global scales. Empirical research has shown a relationship exisits between the amount of radiation reflected from absorption features and the concentration of given biochemicals in leaves and canopies (Matson et al., 1994, Johnson et al., 1994). A technique commonly used to determine which wavelengths have the strongest correlation with the biochemical of interest is unguided (stepwise) multiple regression. Wavelengths are entered into a multivariate regression equation, in their order of importance, each contributing to the reduction of the variance in the measured biochemical concentration. A significant problem with the use of stepwise regression for determining the correlation between biochemical concentration and spectra is that of 'overfitting' as there are significantly more wavebands than biochemical measurements. This could result in the selection of wavebands which may be more accurately attributable to noise or canopy effects. In addition, there is a real problem of collinearity in that the individual biochemical concentrations may covary. A strong correlation between the reflectance at a given wavelength and the concentration of a biochemical of interest, therefore, may be due to the effect of another biochemical which is closely related. Furthermore, it is not always possible to account for potentially suitable waveband omissions in the stepwise selection procedure.

This concern about the suitability of stepwise regression has been identified and acknowledged in a number of recent studies (Wessman et al., 1988, Curran, 1989, Curran et al., 1992, Peterson and Hubbard, 1992, Martin and Aber, 1994, Kupiec, 1994). These studies have pointed to the lack of a physical link between wavelengths chosen by stepwise regression and the biochemical of interest, and this in turn has cast doubts on the use of imaging spectrometry for the estimation of foliar biochemical concentrations at sites distant from the training sites. To investigate this problem, an analysis was conducted on the variation in canopy biochemical concentrations and reflectance spectra using forced entry linear regression.

2.0 Methods and Results

Our analysis was based upon an AVIRIS data set, together with foliage samples, collected from Gainesville, Florida in 1992. This site, comprising slash pine, provided 14 sample plots: 7 control and 7 fertilised. The biochemicals of interest were

nitrogen, lignin and cellulose and the concentration of these biochemicals, together with others, were measured from the 360 foliar samples using wet chemical assay techniques at the University of New Hampshire (Curran and Kupiec, 1995).



AVIRIS averaged spectra of slash pine.

Selected wavebands identified by Himmelsbach et al. (1988) were entered into a forced entry multivariate linear regression procedure (Norusis, 1992) for analysis of the AVIRIS data set. Himmelsbach's chemometric approach to the interpretation of spectral reflectance and associated biochemical absorption features is highly reliable and consistant having been derived from years of laboratory spectrophotometric experiments by the United States Department of Agriculture (USDA). These wavelengths are the known consequence of bending and stretching vibrations of the chemical bonds, together with their harmonics and overtones, between hydrogen and carbon, nitrogen and oxygen atoms causing spectral absorption in the 400 - 2400 nm spectral region. The five wavebands used were those with both (i) a known causal relationship between biochemical bond and reflectance and (ii) the largest partial correlation coefficient with the biochemical of interest. For the spectral data, a first derivative transformation was made on the selected AVIRIS imagery. This process reduces the effects caused by illumination changes and is widely used although there is an increase in the signal-to-noise ratio. The chosen AVIRIS data channels were within 10 nm of the 'Himmelsbach wavelengths'.

Nitrogen - Correlation coefficient, $R^2 = 0.85$

AVIRIS 1st derivative wavelength (nm)	'Himmelsbach wavelengths' (nm)	Chemical bond
1184.27 ± 8.98	1188	C-H Str. 2nd OT
1222.60 ± 8.98	1230	C-H Str. 2nd OT
1700.28 ± 9.81	1690	C-H Str. 1st OT
2162.56 ± 14.43	2168	2X Amide I + Amide III
2301.45 ± 14.58	2294	N-H Bend NH2 2nd OT

Lignin - Correlation coefficient, $R^2 = 0.83$

AVIRIS 1st derivative wavelength (nm)	'Himmelsbach wavelengths' (nm)	Chemical bond
1193.86 ± 8.98	1192	C-H Str. 2nd OT
1541.59 ± 9.63	1542	C=O Str. 3rd OT Aromatic O-H Str. 1st OT
2043.10 ± 14.19	2048	C=O Str. 2nd OT
2102.88 ± 14.32	2106	2X O-H Def. + 2X
2380.58 ± 14.62	2380	Aliphatic C-H Str. + Def.

Cellulose - Correlation coefficient, $R^2 = 0.48$

AVIRIS 1st derivative wavelength (nm)	'Himmelsbach wavelengths' (nm)	Chemical bond
1193.86 ± 8.98	1194	C-H Str. 2nd OT
1541.59 ± 9.63	1536	O-H Str. 1st OT (Intra H bond)
1739.90 ± 9.84	1736	C-H Str. 1st OT
1768.60 ± 9.86	1772	C-H Str. 1st OT
2350.92 ± 14.61	2356	C-H Dev. 2nd OT

3.0 Discussion

The use of airborne spectrometry for the investigation of canopy biochemical concentrations has raised questions about the statistical techniques used for locating those spectral absorption features that are related to biochemical concentrations. Using forced entry linear regression, based upon known absorption features, it has been possible to account for much of the variation in lignin and nitrogen concentrations and about half of the variation in cellulose concentration can be accounted for using the AVIRIS first derivative spectral reflectance. A problem with the AVIRIS data is that some of the important waveband features identified in the laboratory on dry samples could not be selected due to the lack of data in the major water absorption wavelengths (figure). It is interesting to note that other combinations of wavebands identified by Hemmelsbach can produce similar correlation coefficients for the above biochemicals. Whether this is the result of co-linear harmonics requires further investigation but increasing the number of identified wavebands in the regression equation does increase the R^2 suggesting that there may be a possibility that biochemical concentrations can indeed be reliably detected in airborne spectrometry reflectance. Future research requires the development of radiative tranfer models incorporating foliar biochemistry at the leaf and canopy scales to help our understanding of those effects.

4.0 Acknowledgements

The research was funded by the Natural Environmental Research Council (Research grant GR3/7647 to PJC and studentship GT4/94/407/L to TPD), the University of New Hampshire (Accelerated Canopy Chemistry Program grant to PJC) and the National Aeronautical and Space Administration (AVIRIS flights). The authors are indebted to many individuals who made this work possible, notably Henry Gholz (University of Florida), Geoff Smith (University of Wales, Swansea), John Aber and Mary Martin (University of New Hampshire) and Jennifer Dungan and Dave Peterson (NASA, Ames Research Center).

5.0 References

Aber J. (ed.), 1994, *Accelerated Canopy Chemistry Program Final Report to NASA-EOS-IWG*, NASA - Washington, DC.

Curran P.J. and Kupiec J., 1995, "Imaging spectrometry: a new tool for ecology" In Danson F.M. and Plummer S.E. (eds) *Advances in Environmental Remote Sensing*, Wiley & Sons, Chichester (in press).

Curran P.J., 1994, "Imaging Spectrometry", *Progress in Physical Geography*, 18, 247-266.

Curran P.J., Dungan J.L., Macler B.A., Plummer S.E. and Peterson D.L., 1992, "Reflectrance spectroscopy of fresh whole leaves for the estimation of chemical concentration", *Remote Sensing of Environment*, 39, 153 - 166.

Curran P.J., 1989, "Remote sensing of foliar chemistry", *Remote Sensing of Environment*, 30, 271 - 278.

Himmelsbach D.S., Boer H., Akin D.E. and Barton II F.E., 1988, "Solid-state carbon-13 NMR, FTIR and NIR spectroscopic studies of ruminant silage digestion" in C.S. Erand and A.M.D. Davies (eds.), *Analytical Applications of Spectroscopy*, Royal Society of Chemistry, London, England, 410-413.

Johnson L.F., Hlavka C.A. and Peterson D.L., 1994, "Multivariate analysis of AVIRIS data for canopy biochemical estimation along the Oregon transect", *Remote Sensing of Environment*, 47, 216 - 230.

Kupiec J.A., 1994, *The Remote Sensing of Foliar Chemistry*, Ph.D. Thesis, University College of Wales, Swansea.

Martin M.E. and Aber J.D., 1994, "The Study of Forest Ecosystems Through the Remote Sensing of Canopy Chemistry", *Accelerated Canopy Chemistry Program Final Report to NASA-EOS-IWG*, NASA - Washington, DC.

Matson P., Johnson L.F., Billow C., Miller J. and Pu R., 1994, "Seasonal patterns and remote spectral estimation of canopy chemistry across the Oregon transect", *Ecological Applications*, 4, 280 - 298.

Norusis M.J., 1992, *SPSS for Windows Base Systems User's Guide 5.0*, Prentice Hall, London, 303 - 357.

Peterson D.L. and Hubbard S., 1992, "Scientific issues and potential remote-sensing requirements for plant biochemical content", *Journal of Imaging Science and Technology*, 36, 5, 446 - 456.

Peterson D.L., 1991, *"Report on the Workshop Remote Sensing of Plant Biochemical Content: Theoretical and Empirical Studies"*, NASA-Ames Research Center, Mountain View, CA.

Wessman C.A., Aber J.D., Peterson D.L. and Melillo J.M., 1988, "Remote sensing of canopy chemistry and nitrogen cycling in temperate forest ecosystems", *Nature* 335, 8, 154-156.