COMPARISON OF AVIRIS AND LANDSAT ETM+ FOR THE ESTIMATION OF LEAF AREA INDEX

Kyu-Sung Lee[1, 2] and Warren B. Cohen[2]

1. INTRODUCTION

Although there are several parameters of interest in the quantitative aspects of vegetative remote sensing, leaf area index (LAI) has been one of the most useful and important parameters to characterize the vegetation activities from local to global scales. LAI, defined as the sum of the leaf area per unit ground area, can be used to measure the activities (photosynthesis, transpiration, and evapotranspiration) and the production of plant ecosystems (Pierce and Running, 1988; Bonan, 1993). The measurement of LAI on the ground is very difficult and requires a great amount of labor and cost (Gower et al., 1999). Since the plant canopy is composed of leaves, which is a direct source of the energy-matter interactions in most earth-observing remote sensing systems, LAI has been an attractive variable of interest in vegetative remote sensing.

There have been many attempts to estimate LAI using various types of remote sensor data since the early stage of space remote sensing (Badhwar et al., 1986; Peterson et al., 1987; Turner et al., 1999). Remote sensing estimation of LAI has been primarily based on the empirical relationship between the ground measured leaf area index and sensor observed spectral responses (Curran et al., 1992; Peddle et al., 1999). Beside such empirical approaches, there have been a few studies to estimate LAI by the inversion of canopy reflectance models (Smith, 1993; Jacquemoud et al., 1995; Kuusk, 1998). However, these studies were mostly based on broadband multi-spectral data, such as Landsat TM and SPOT images. Recent development of imaging spectrometry technology has brought a new form of remote sensor data, which has shown great potential for many applications. In particular, the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) data, which have been provided since the late 1980s, were used to study specific features of vegetation targets. The 224 continuous spectral band data demonstrated the capability to depict certain characteristics of vegetation canopy, such as biochemical constituent (Wessman et al., 1988; Johnson et al., 1994). There have been a few studies using hyperspectral data in relation to leaf area index. In recent years, hyperspectral data has been used to estimate LAI by the inversion of radiative transfer models, which simulate the mechanics of canopy reflectance (Asner et al., 2000).

The objective of this study is to evaluate the capability of narrowband hyperspectral data for estimating leaf area index by comparing them with those of broadband multispectral instruments. Although there were many studies dealing with the remote sensing estimation of LAI by various types of image data, it has been rare to find the advantage of using the narrowband hyperspectral data over the broadband multispectral data. The lack of comparative studies might be explained by the shortage of both hyperspectral and broadband data sets available at the same location. Even though it was a simulated form of hyperspectral data, the hyperspectral-based vegetation indices did not provide better results than broadband data for estimating LAI (Jacquemoud et al., 1995; Broge and Leblanc, 2000). Lefsky et al. (2001) have compared several data types, including Landsat TM, high spatial data, AVIRIS, and scanning lidar data, for estimating forest stand structures and found that AVIRIS data were not better than other data types for relating stem diameter, biomass, and basal area. Some studies have compared vegetation indices derived from both narrowband and broadband data to relate them for discriminating different vegetation groups or cover types (Teillet et al., 1997; Galvao et al., 2000). Hyperspectral data may or may not be better than broadband data to estimate LAI. The empirical approach is often very useful for analyzing and comparing different types of remote sensing data by relating them to the ground-measured biophysical variables. This study attempts to compare AVIRIS and Landsat ETM+ data by relating them to LAI over the two study sites of tall-grass prairie and mixed forest, which have well-established ground measurement plots.

[1] Inha University, Department of Geoinformatic Engineering, Inchon, Korea (ksung@inha.ac.kr)
[2] USDA Pacific Northwest Research Station, Forestry Sciences Laboratory, Corvallis, Oregon

2. DATA AND METHODS

2.1 Study Sites and Data

This study was conducted in collaboration with the BigFoot project, which was designed to provide local validation of global estimates of biophysical variables and processes using MODIS data (Cohen and Justice, 1999). Study areas include two sites of mixed forest and grass prairie. These areas are also the sites of the long-term ecological research plots and a rich source of ground information. The Harvard Forest (HARV) site has a 5 x 5 km^2 area located about 50 km northwest of Worcester, Massachusetts. The forest has mixed species of eastern hemlock, red pine, and eastern hardwood. The Konza Prairie (KONZ) site is located approximately 10 km southwest of Manhattan, Kansas, and mostly covered by grass and shrub.

At each site there were one hundred, 25 x 25 m^2 plots where land cover, LAI, absorbed radiation, and net primary production were measured/observed. Each plot has four to nine subplots. Subplot measurements were averaged to provide a single value for each measured variable at each plot. Plot locations were determined using a real-time differential GPS. The accuracy of the system was <0.3 m in both the *x* and *y* dimensions. At the HARV site, LAI was measured at two time periods (June and August) during the growing season in 2000. Due to unexpected problems during the field measurements, only 73 of the 100 plots were measured. At the KONZ site, LAI measurements were made in all the 100 plots in June 2000.

Considering the phenological variation of vegetation target, both AVIRIS and Landsat ETM+ data were obtained as close to the ground data collection as possible (Table 1). For the KONZ study site, the data acquisition date was close in time to the ground measurements. The HARV study site, however, shows some discrepancy between image acquisition and ground measurement. We used two sets of LAI measurements separately for analyzing AVIRIS and ETM+ data. AVIRIS data were analyzed with the field LAI data of June while ETM+ data were analyzed with the LAI measurement of June.

Table 1. Date of data acquisition for AVIRIS and TM image data and ground measurements of LAI.

Study site	AVIRIS	Landsat ETM+	Field LAI measurement
Harvard Forest (HARV), Massachusetts	May 16, 2000		June 18, 2000
		August 31, 1999	August 4, 2000
Konza Prairie (KONZ), Kansas	June 22, 2000	June 7, 2000	June 6, 2000

2.2 Preprocessing of Image Data

Both AVIRIS and ETM+ images were georeferenced, radiometrically calibrated, and converted to surface reflectance value. Initially, Landsat ETM+ data were georeferenced to UTM coordinates by using USGS digital orthophoto quadrangles. The AVIRIS data were then georectified into the ETM+ image by image-to-image registration. The average positional accuracy of the geometric correction was less than 0.5 pixels.

In quantitative analysis of remote sensing images, atmospheric correction is a critical step to convert the digital number (DN) value to a more absolute magnitude of surface reflectance. Due to the carefully designed experimental sensor system, AVIRIS data have relatively well-developed atmospheric correction algorithms. Atmospheric water-vapor absorption is the most influential factor in the atmospheric correction of remote sensing data. The atmospheric correction of hyperspectral data has a clear advantage over multispectral data since the magnitude of water-vapor effects can be directly assessed from certain spectral channels of their own data sets. AVIRIS data used in this study were atmospherically corrected by using the ATREM program (Gao et al., 1993), which is based on a radiative transfer model. ATREM uses two water absorption spectra (around 940 nm and 1140 nm) channels in the AVIRIS channels to estimate the amount of water vapor at the time of data acquisition.

Although there have been several studies dealing with the atmospheric correction of Landsat TM and ETM+ data, all of them have certain limitations to calculate the absolute magnitude of waver vapor absorption. For our study, we used the COST radiometric correction model of Chavez (1996), which is an empirical model based upon 'dark' objects. Radiometrically 'dark' objects were assumed to have 2 % reflectance across all bands. Excluding the thermal infrared band, the six reflective bands of ETM+ data were converted to surface reflectance.

2.3 AVIRIS Data Manipulation

From the aspect of the number of spectral bands and the range of spectrum involved, the relative amount of information content of AVIRIS may be far more than that of Landsat ETM+. Therefore, any absolute comparisons between these two data types may not be reasonable. Also, the comparison between ETM+ and AVIRIS can be subjective, depending on the method applied. In this study, we generated four sets of AVIRIS data to be compared with the six reflective bands of Landsat ETM+. These AVIRIS data sets were mostly out of 89 channels covering the wavelength range of six ETM+ bands, with the exception of the data set obtained by principal component analysis, which used all AVIRIS channels excluding the water-absorption and high-noise channels. Table 2 shows the list of AVIRIS data sets generated for the comparison with ETM+ data.

Table 2. Sets of AVIRIS data to be compared with Landsat ETM+ data.

AVIRIS data set	Number of AVIRIS channels used	Method
Simulated ETM+ (sim_TM)	89 channels covering the spectrum of six ETM+ bands	• Weighted average of TM corresponding channels • Weighting factors were determined to represent spectral responsivities of each ETM+ band
Center wavelength of TM (cen_TM)	6 channels matching with the center wavelength of each ETM+ band	• ch 12 (481nm) for ETM+ band 1 • ch 21 (568nm) for ETM+ band 2 • ch 31 (665nm) for ETM+ band 3 • ch 52 (837nm) for ETM+ band 4 • ch 137 (1653nm) for ETM+ band 5 • ch 195 (2220nm) for ETM+ band 7
Selected combination (sel_AVIRIS)	7 channels (HARV) 14 channels (KONZ)	Selection by multiple stepwise regression using 89 channels covering ETM+ spectrum
Principal component (pc_AVIRIS)	187 channels	Principal component transformation of all 187 channels (excluding water-absorption and noise channels from original 224 channels)

2.4 Multiple Regression and Canonical Correlations Analysis

At first, we extracted surface reflectance values from the all sets of AVIRIS and Landsat ETM+ data for the location of every ground plot for each study site. Field-measured LAI was then related to spectral reflectance values extracted from the image data by using the multiple regression method. The coefficient of determination (R^2) of each regression equation is used as a simple measure of predicting LAI by the image data. The 89 AVIRIS channels would have too many independent variables for relating LAI by the multiple regression method. We used the stepwise selection method to select the best sets of AVIRIS channels.

Another way of looking at the closeness between LAI and image-extracted spectral reflectance was canonical correlation analysis (CCA). Canonical correlation analysis is a multivariate statistical method to look at the relationships between two sets of variables (multiple *X*s and multiple *Y*s) (Tabachnick and Fidell, 1989). CCA maximizes the correlation between a composite of variables from one set with a composite of variables from another set. When there is only one *X* (such as the LAI in this study), CCA provides a set of coefficients to transform the *Y*s (selected image bands) such that they have the maximum correlation with *X*. The canonical correlation coefficients were used as a single measure of relationship between LAI and a set of multiple image bands selected for each data

set. Since CCA is a generalized form of multiple regression, canonical correlation coefficient (r) is basically the same as the square root of the coefficient of determination (R^2) of the multiple regression model.

3. RESULTS

As an initial effort to examine the relationship between the field-measured LAI and AVIRIS spectral reflectance, correlation coefficients were calculated. As can be observed in Figure 1, two study sites show a rather different pattern of correlation along the wavelength. In the HARV study site, where the variation of LAI and species composition is relatively large, relatively high correlations were found in visible and middle infrared spectra. The statistical significance level is proportional to the absolute value of correlation coefficients (r), and the confidence limit is less than 95% when $|r|$ is below 0.15. It turned out that the near infrared spectrum (730–1300 nm) showed no significant correlation with LAI. The weak correlations in the near infrared spectrum might be explained by the saturated spectral reflectance at the high LAI (around 5 on average) of the HARV study site. The strongest correlations were found at the red-edge position (around 720 nm) and the middle infrared spectrum (2070–2400 nm). In the KONZ study site, where the range of LAI is rather narrow and cover type is relatively homogeneous grass prairie, relatively high correlation coefficients with the field-measured LAI were found at the near infrared spectrum (730−1300 nm). The field-measured LAI in this study area were relatively small (less than 2 on average), and they were relatively sensitive in the near infrared spectrum. Unlike the HARV study site, no significant correlations were found at the visible spectrum. In the middle infrared spectrum, significant correlations were only found in the wavelengths longer than 2100 nm. Such high correlations in the middle infrared spectrum have not been well explained and probably need further analysis. Again, the strongest correlation with the field-measured LAI was found at the red-edge position around 730 nm, which has been reported by several previous studies (Danson, 1996).

For the comparison, the similar correlation coefficients were also obtained by using the ETM+ reflectance for each study site (Figure 2). Unlike with the AVIRIS reflectance, no significant correlations were found at the visible and near infrared spectra. The only statistically significant correlations were found in the second middle infrared band (ETM+ band 7), which has a spectrum range between 2090 and 2350 nm.

Comparing the two correlation plots in Figures 1 and 2, we might anticipate the outcome of the multiple regression models using AVIRIS and ETM+ data. Due to the relatively weak correlations along the entire range of wavelength spectra in both AVIRIS and ETM+, the coefficients of determination (R^2) of the regression models were relative low. Overall, most AVIRIS data sets did not show any striking improvements of R^2 values over the ETM+ data although certain combinations of AVIRIS channels showed some improvement to explain the LAI (Table 3). The simulated ETM+ data or the six AVIRIS channels matching the center wavelength of ETM+ bands was not much different from the ETM+ data in predicting the LAI for both sites. From the viewpoint of the sensor dynamic range, pixel gray level, and atmospheric correction procedure, AVIRIS can be said to have better radiometric resolution than ETM+ data. The simulated ETM+ data have exactly the same spectral range as the ETM+ and probably have improved radiometric quality than ETM+. However, the high radiometric resolution of AVIRIS does not help to explain the LAI as compared to the ETM+.

Table 3. Coefficients of determination (R^2) of multiple regression models to relate the field measured LAI.

Data set used	Harvard Forest	Konza Prairie
ETM+ 6 bands	0.1729	0.2815
AVIRIS (simulated ETM+)	0.3348	0.1986
AVIRIS (6 channels of ETM+ center wavelength)	0.3336	0.2557
AVIRIS (best subset channels)	0.5138	0.5571
AVIRIS (first 15 principal components)	0.4376	0.3486

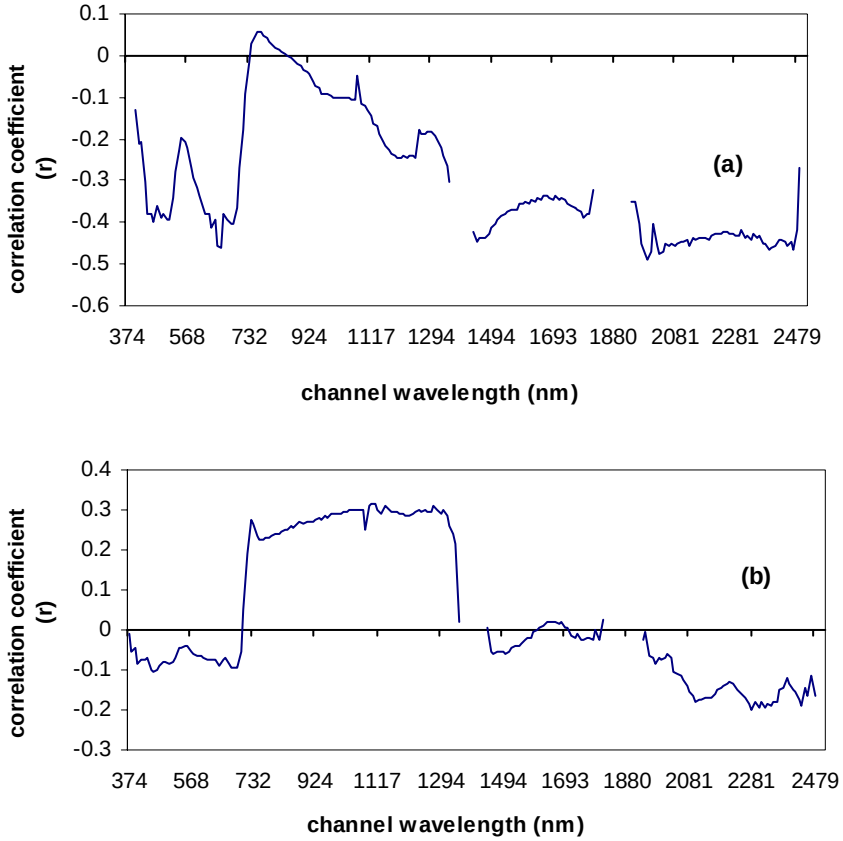

Figure 1. Correlation coefficients between the field-measured leaf area index (LAI) and reflectance along the AVIRIS spectral channels for Harvard Forest (a) and Konza Prairie (b). Statistical significance level is proportional to the absolute value of r (less than 95% confidence limits when $|r| < 0.15$).

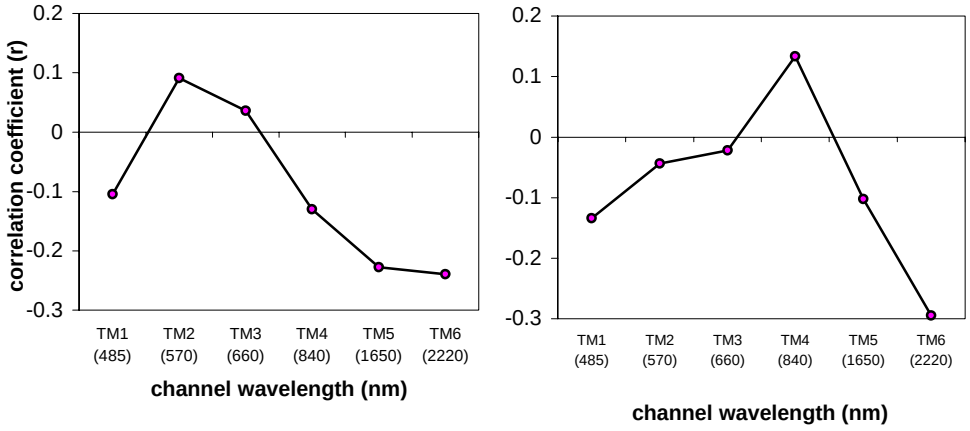

Figure 2. Correlation coefficients between the field-measured leaf area index (LAI) and reflectance along the ETM six bands for the Harvard Forest (left) and Konza Prairie (right). Statistically significant correlations were not found in visible and near infrared spectra.

Principal component analysis (PCA) has been frequently used in dealing with multi-dimensional remote sensing data. In particular, with the capacity of assembling total variation of whole data into a reduced number of bands, the principal component transformation is often used in analyzing hyperspectral data. In our study, we applied PCA to the 187 AVIRIS channels after excluding those water-absorption and high-noise channels from the original 224 channels. The first 15 principal components include 99.9% of total variations of the 187 channels AVIRIS data used. Even though the 15 bands of PCA transformed data have the most information content, they were not very helpful for explaining the LAI. The R^2 value of the regression model using PCA transformed data shows only a slight improvement.

In both study sites, the best-fit regression model was achieved when we used only a few selected channels of AVIRIS data. The coefficients of determination (R^2) were higher than the ETM and other subsets of AVIRIS data. The best channel selection was performed by the stepwise multiple regression method. Table 4 shows the final results of the stepwise multiple regression to select the best AVIRIS channels to explain LAI. Although we could protect against including any channels that do not contribute to the predictive power of the regression model by specifying small significance level (i.e., 5 percent), we tried to include more channels that could contribute to predicting LAI by increasing the significance level up to 15 percent (SAS, 1988). Only 7 AVIRIS channels were selected in the HARV study site while 14 channels were selected for the KONZ study site. These selected AVIRIS channels do not correspond exactly for the two study sites except for the one middle infrared channel (2310 nm).

Table 4. The result of stepwise multiple regression analysis to select the best AVIRIS channels to explain the field measured LAI for each study site.

step	Harvard Forest		Konza Prairie	
	channel entered (nm)	R^2	channel entered (nm)	R^2
1	Ch 11 (471)	0.2914	Ch 56 (876)	0.0731
2	Ch 53 (847)	0.3225	Ch 54 (856)	0.2141
3	Ch 204 (2310)	0.3628	Ch 201 (2281)	0.2781
4	Ch 24 (597)	0.4059	Ch 186 (2131)	0.3358
5	Ch 49 (808)	0.4523	Ch 59 (904)	0.3603
6	Ch 15 (510)	0.4869	Ch 208 (2350)	0.3809
7	Ch 58 (894)	0.5138	Ch 31 (665)	0.4025
8			Ch 185 (2121)	0.4423
9			Ch 204 (2310)	0.4663
10			Ch 205 (2320)	0.4867
11			Ch 192 (2191)	0.5100
12			Ch 196 (2231)	0.5274
13			Ch 199 (2261)	0.5447
14			Ch 206 (2330)	0.5571

Although the magnitude of R^2 for each regression model is the same as the squared value of the canonical correlation coefficient, the scatter plot between the canonical variable and LAI shows effectively the interrelationships (Figures 3 and 4). In both study sites, there was not much differences among the three scatter plots (on the left side) for the canonical variable of ETM+, simulate ETM and 6 ETM center channels although the canonical correlation coefficients were slightly different. As indicated by the R^2 value of the regression model, the highest canonical correlation was found for the canonical variable of the best subset of AVIRIS channels.

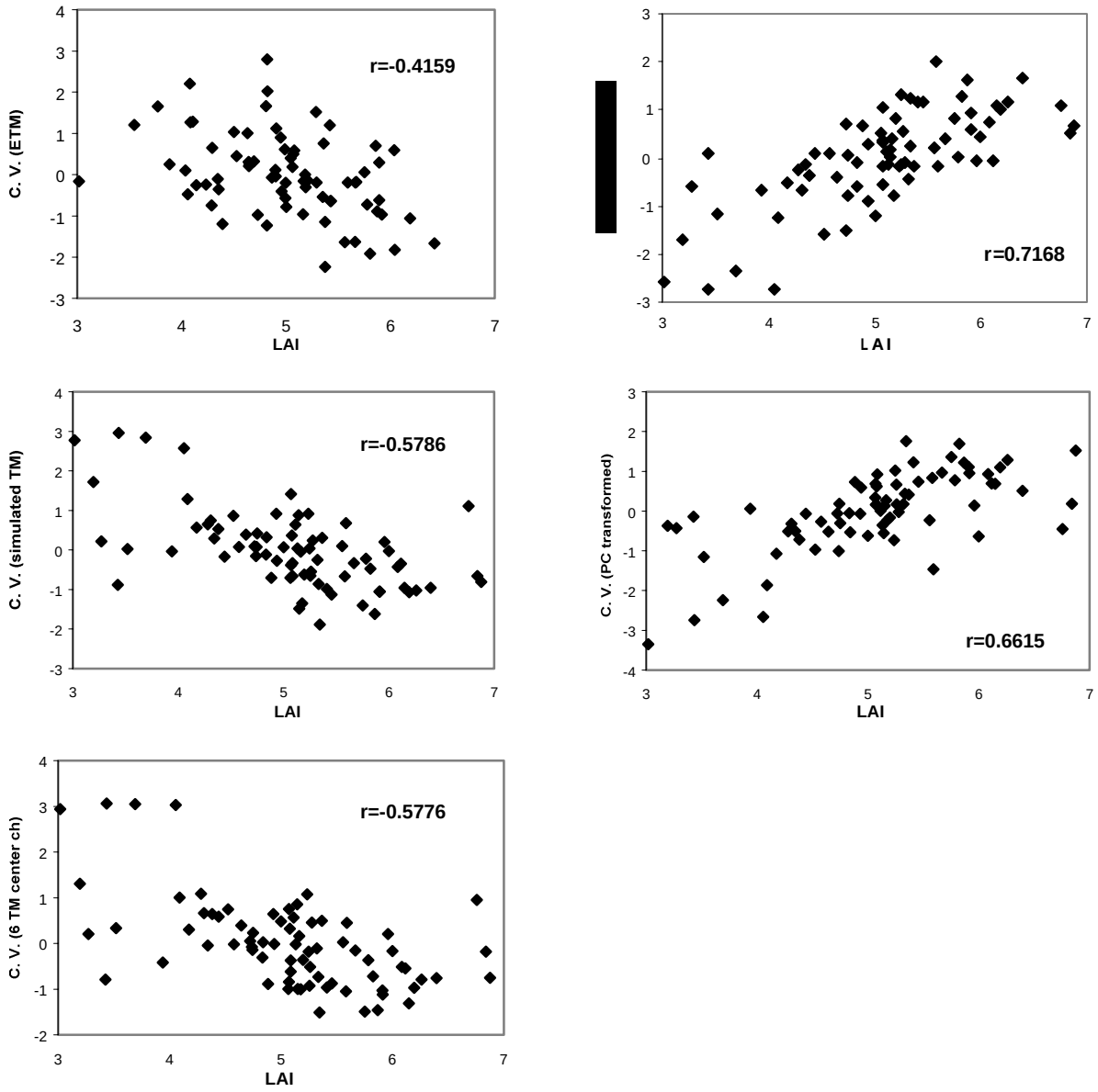

Figure 3. Canonical correlations of ETM and selected sets of AVIRIS data aligned with the leaf area index for the Harvard Forest study site.

4. CONCLUSIONS

It has been more than ten years since hyperspectral images were introduced. There are several airborne imaging spectrometer systems around the world. Further, we have already seen the spaceborne imaging spectrometers and are expecting more hyperspectral sensor systems in the near future. Although hyperspectral image data are known to have great potential in many applications of vegetative remote sensing, there have been very few studies that actually show the advantages of narrowband hyperspectral data over the traditional broadband multispectral data.

In an attempt to compare the predictive power of AVIRIS and ETM+ data, several subsets of AVIRIS data were analyzed in relation to the field-measured LAI for the two study sites of mixed forest and grassland. Among the four subsets of AVIRIS data, the best combination of selected AVIRIS channels was better than ETM+ or other

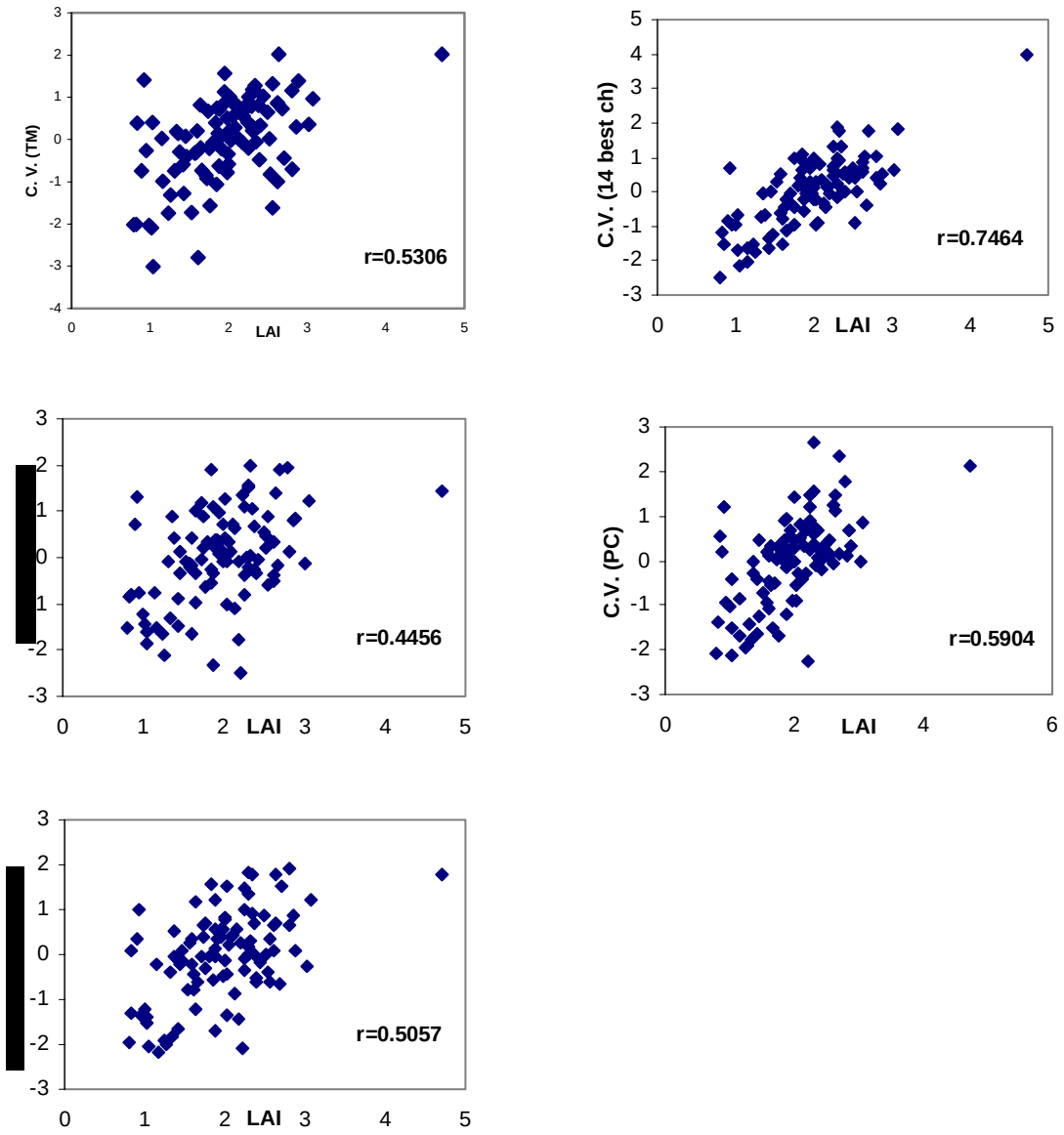

Figure 4. Canonical correlations of ETM and selected sets of AVIRIS data aligned with the leaf area index for the Konza Prairie study site.

subsets of AVIRIS in predicting LAI. There was no difference between the ETM-like subsets (simulated ETM+ and the six channels corresponding to the center wavelength ETM+ bands) and the actual ETM+. Although principal component analysis has been frequently used in hyperspectral data analysis, the transformed principal components were not necessarily aligned to predict the LAI.

In many applications of hyperspectral data, it may be rare to use all the spectral channels of the original hyperspectral image. As seen in this study, the 7 to 14 selected channels worked better to predict LAI than the principal component transformed data set. This empirical study demonstrated that narrowband hyperspectral data could be better than the broadband ETM+ data in estimating LAI. Our next question would be how we can select those best channels to use for certain applications. We believe that channel selection should be based on the integrated approach of theories, effective data processing, and empirical validation.

5. REFERENCES

Asner, G.P., R.N. Treuhaft, B. E. Law, 2000. A hyperspectral photon transport system to simulate imaging spectrometer observations of terrestrial ecosystems, Proc. of 2000 JPL AVIRIS Workshop.

Badhwar, G.D., R.B. MacDonald, and N.C. Mehta, 1986. Satellite-derived leaf area index and vegetation maps as input to global carbon cycle models—a hierarchical approach, *International Journal of Remote Sensing*, 7(2):265–281.

Bonan, G., 1993. Importance of leaf area index and forest type when estimating photosynthesis in boreal forests, *Remote Sensing of Environment*, 43:303–314.

Broge, N.H. and E. Leblanc, 2000. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density, *Remote Sensing of Environment*, 76:156–172.

Chavez Jr., P., 1996. Image-based atmospheric corrections—revised and improved, *Photogrammetric Engineering & Remote Sensing*, 62:1025–1036.

Cohen, W. and C. Justice, 1999. Validating MODIS terrestrial ecology products: linking in situ and satellite measurements, *Remote Sensing of Environment*, 70:1–3.

Curran, P.J., J. Dungan, and H.L. Gholz, 1992. Seasonal LAI measurements in slash pine using Landsat TM. *Remote Sensing of Environment*, 39:3–13.

Danson, F.M., 1996. Developments in the Remote Sensing of Forest Canopy Structure, in Advances in Environmental Remote Sensing, edt. by Danson and Plummer, John Wiley & Sons, New York, USA, 53–69.

Galvao, L.S., I. Vitorello, and M.A. Pizarro, 2000. An adequate band positioning to enhance NDVI contrasts among green vegetation, senescent biomass, and tropical soils, *International Journal of Remote Sensing*, 21(9):1953–1960.

Gao, B.C., Hedebrecht, K., and Goetz, A.F.H., 1993. Derivation of scaled surface reflectances from AVIRIS data, *Remote Sens. Environ.*, 44:165–178.

Gower, S., C. Kucharik, and J. Norman, 1999. Direct and indirect estimation of leaf area index, fAPAR, and net primary production of terrestrial ecosystems, *Remote Sensing of Environment*, 70:29–51.

Jacquemond, S., F. Baret, B. Andrieu, F. M. Danson, and K. Jaggard, 1995. Extraction of vegetation biophysical parameters by inversion of the PROSPECT + SAIL models on sugar beet canopy reflectance data. Application to the TM and AVIRIS sensors, *Remote Sensing of Environment*, 52:163–172.

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

Kuusk, A., 1998. Monitoring of vegetation parameters on large areas by the inversion of a canopy reflectance model, *Inter. J. Remote Sensing.*, 19(15):2893−2905.

Lefsky, M., W. Cohen, and T. Spies, 2001. An evaluation of alternative remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon, *Canadian Journal of Forest Research*, 31:78–87.

Peddle, D.R., F.R. Hall, and E.F. LeDrew, 1999. Spectral mixture analysis and geometric-optical reflectance modeling of boreal forest biophysical structure, *Remote Sensing of Environment*, 67:288–297.

Peterson, D.L., M.A. Spanner, S.W. Running, and K. Teuber, 1987. Relationship of Thematic Mapper data to leaf area index of temperate coniferous forests, *Remote Sensing of Environment*, 22:323–341.

Pierce, L.L and S.W. Running, 1988. Rapid estimation of coniferous forest leaf area index using a portable integrating radiometer, *Ecology*, 69:1762–1767.

SAS, 1988. SAS/STAT User's Guide, Release 6.03 Edition, SAS Institute Inc., Cary, NC, USA

Smith, J.A., 1993. LAI Inversion using a back-propagation neural network trained with a multiple scattering model, *IEEE Trans. Geoscience and Remote Sensing*, 31(5):1102–1106.

Tabachnick, B., and L. Fidell, 1989. *Using Multivariate Statistics, 2nd Edition*. 1989. Harper Collins Publishers, Inc., United Kingdom.

Teillet, P.M., K. Staenz, and D.J. Williams, 1997. Effects of spectral, spatial, and radiometric characteristics on remote sensing vegetation indices of forested regions, *Remote Sensing of Environment*, 61:139–149.

Turner, D., W. Cohen, R. Kennedy, K. Fassnacht, and J. Briggs, 1999. Relationships between leaf area index and Landsat TM spectral vegetation indices across three temperate zone sites, *Remote Sensing of Environment*, 70:52–68.

Wessman, C.A., J.D. Aber, D.L. Peterson, and J.M. Melilo, 1988. Remote sensing of canopy chemistry and nitrogen cycling in temperature forest ecosystems, *Nature*, 335:154–156.