Available online at www.sciencedirect.com

SCIENCE DIRECT®

Remote Sensing of Environment 94 (2005) 405–424

Remote Sensing of Environment

www.elsevier.com/locate/rse

A hybrid inversion method for mapping leaf area index from MODIS data: experiments and application to broadleaf and needleleaf canopies

Hongliang Fang*, Shunlin Liang

Department of Geography, University of Maryland, College Park, MD 20742, USA

Received 9 January 2004; received in revised form 27 October 2004; accepted 3 November 2004

Abstract

Leaf area index (LAI) is an important variable needed by various land surface process models. It has been produced operationally from the Moderate Resolution Imaging Spectroradiometer (MODIS) data using a look-up table (LUT) method, but the inversion accuracy still needs significant improvements. We propose an alternative method in this study that integrates both the radiative transfer (RT) simulation and nonparametric regression methods. Two nonparametric regression methods (i.e., the neural network [NN] and the projection pursuit regression [PPR]) were examined. An integrated database was constructed from radiative transfer simulations tuned for two broad biome categories (broadleaf and needleleaf vegetations). A new soil reflectance index (SRI) and analytically simulated leaf optical properties were used in the parameterization process. This algorithm was tested in two sites, one at Maryland, USA, a middle latitude temperate agricultural area, and the other at Canada, a boreal forest site, and LAI was accurately estimated. The derived LAI maps were also compared with those from MODIS science team and ETM+ data. The MODIS standard LAI products were found consistent with our results for broadleaf crops, needleleaf forest, and other cover types, but overestimated broadleaf forest by 2.0–3.0 due to the complex biome types.

Keywords: Leaf area index (LAI); Soil reflectance index (SRI); MODIS; ETM+; Neural network; Projection pursuit regression

1. Introduction

Leaf area index (LAI), one-sided leaf area per unit of ground area, defines an important structural property of terrestrial vegetation canopies. It is a crucial variable in canopy interception, evapotranspiration, and net photosynthesis. Currently, there is considerable interest in developing algorithms for the estimation of LAI to drive the ecosystem productivity models (Running et al., 1989) and some general circulation models (Chase et al., 1996).

Many efforts have been made to estimate LAI from satellite measurements through LAI's statistical relationship with spectral vegetation indices, by physical model inversion or by other nonparametric methods (Liang, 2003; Weiss & Baret, 1999). Using the empirical vegetation indices is simple and easy. Its limits are also obvious, however, such as the limited amount of spectral information, the diversified empirical equations used and their sensitivity to nonvegetation factors. The model inversion method is physically based and independent of vegetation type, but the model inversion process is not always unique and demands time. The nonparametric methods (e.g. neural networks, NN), which provide a direct relationship between the simulated reflectance and the corresponding biophysical variables of interest, are ideal for LAI extraction.

The Moderate Resolution Imaging Spectroradiometer (MODIS) science team in the Earth Observing Program (EOS) is producing an LAI product globally (Justice et al., 1998; Myneni et al., 2002). The MODIS LAI product, a 1-km global data product updated every 8 days, is available for the general user community through the Earth Resources Observation System (EROS) Data Active Archive Center (DAAC). The operational MODIS LAI algorithm uses vegetation maps as a priori information to

* Corresponding author. Tel.: +1 301 405 4050; fax: +1 301 314 9299.
E-mail address: fanghl@geog.umd.edu (H. Fang).

0034-4257/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2004.11.001

constrain the vegetation structural and optical parameter space (Myneni et al., 1997). Six major biomes were used: grasses and cereal crops (biome 1), shrubs (biome 2), broadleaf crops (biome 3), savannas (biome 4), broadleaf forests (biome 5), and needleleaf forests (biome 6). For each land pixel, numerical solutions to a three-dimensional radiative transfer (RT) equation are used to account for the bi-directional reflectance factors (BRF) of the biomes for varying sun-view geometry and canopy/soil patterns (Knyazikhin et al., 1998a, 1998b). A look-up table (LUT) is constructed including a suite of canopy structures and soil characteristics of each biome. The present version of the LUT contains 25 patterns of effective ground reflectances evaluated from the soil reflectance model of Jacquemoud et al. (1992). By comparing the observed and modeled BRFs, LAI is retrieved. The solution is usually not unique; therefore, the mean values of LAI averaged over all acceptable values and their dispersions are taken as the retrievals and their uncertainties (Knyazikhin et al., 1998a, 1998b). Should this main algorithm fail, a back-up algorithm is triggered to estimate LAI using the NDVI (normalized difference vegetation index). The backup algorithm makes use of the pixel NDVI and the straightforward NDVI-LAI relationship for each biome. The LAI product has a value between 0.0 and 8.0 assigned to each 1-km cell of the global grid database.

One important aspect of the MODIS LUT method is that some variables, such as soil reflectance and leaf reflectance and transmittance, need to be fixed with a priori constants. Soil and leaf optical properties are allowed to vary only with biome types by MODIS algorithms in order to facilitate its global application. However, most of these variables vary dramatically. Fixing them with constants carries large uncertainties (Walthall et al., 2000)—we have released the restriction on soil and leaf optical properties in our new approach. In addition, land biome type is indispensable for this algorithm; misclassification will lead to accumulated errors in the final LAI products.

To overcome some of these limitations, we developed a new hybrid LAI estimation approach based on intensive RT model simulation and nonparametric regression methods. Two nonparametric methods, an artificial NN approach and the projection pursuit regression (PPR) approach were explored in creating this hybrid. A similar approach was applied to the fine resolution Landsat ETM+ data (Fang & Liang, 2003) and EO1 ALI data (Liang et al., 2003). However, its applicability to MODIS and over other landscapes is still unknown, requiring further experiments and validation work. In addition, independent methods like ours are crucial for validation of the MODIS standard LAI products (Privette et al., 1998).

This paper starts with a description of the hybrid approach in Section 2. The two test sites and the data used to estimate LAI with the new approach are described in Section 3. Section 4 presents the results. The experimental design flow is illustrated in Fig. 1.

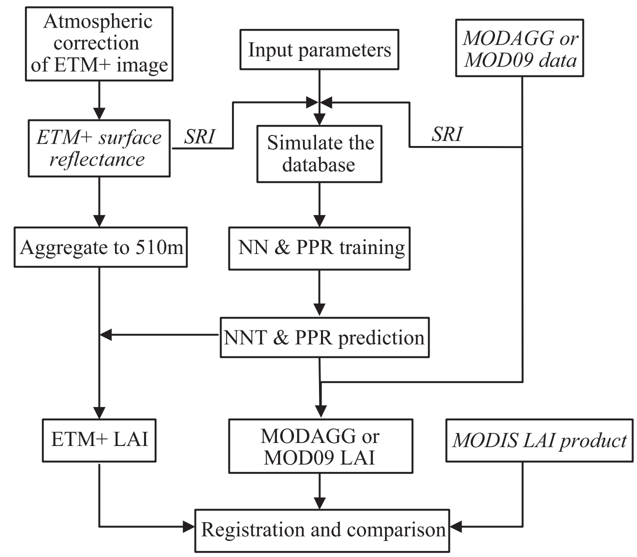

Fig. 1. Work flow of the new hybrid approach to estimate leaf area index with remote sensing imagery.

2. A hybrid LAI estimation approach

The hybrid approach begins with a radiative transfer simulation with various input biophysical parameters (Fig. 1). A new soil reflectance index (SRI) was developed to account for the soil reflectance during the simulation process. Leaf optical properties were simulated with different analytical leaf models for different biome types. Since this new algorithm fixes fewer parameters, the features of different biomes have been implicitly incorporated into the algorithm. The RT simulation processes were designed for two major vegetation types, broadleaf and needleleaf canopies, respectively. Both NN and PPR approaches were integrated with the RT simulation results and used to predict LAI from MODIS data. In this section, the soil reflectance index was first introduced, followed by the RT simulation and LAI estimation with nonparametric methods.

2.1. Determining SRI from MODIS data

The soil reflectance, especially for sparse canopies with small LAIs, is one of the most sensitive parameters in the radiative transfer models. When the LAI increases (>3), the importance of the soil background decreases (Bicheron & Leroy, 1999).

Conventionally, there are several different ways to deal with the soil reflectance in RT model simulations:

(1) using field-measured soil reflectance data directly (Abuelgasim et al., 1998; Qi et al., 2000; Smith, 1993);
(2) using soil reflectances from a soil spectral library (Broge & Leblanc, 2001);

(3) using randomly generated soil reflectance (Kimes et al., 2002) based on the coefficients in green, red and near infrared (NIR) bands; and
(4) using reflectances generated from the soil line (Baret et al., 1995).

Generally, using the field-measured soil reflectance is the most accurate approach, assuming good availability of data. Reflectances from a soil spectral library may not represent the real conditions in the field. Randomly generated or soil line reflectances are appropriate when they are applied to a particular soil background because they are derived from empirical observations.

In this study, the soil reflectance was calculated from the soil reflectance index based on the red–NIR reflectance space. SRI is a simplification of a very complex situation and can be directly derived from the satellite data. The SRI performed very well for the fine resolution ETM+ data (Fang & Liang, 2003), but its characteristics in MODIS imagery still need to be investigated due to different soil line properties. In general, soil pixels are located at the low-right side of the red–NIR scatterplot (Fang & Liang, 2003). Knowing this allows us to develop an automatic method for soil line extraction. Fig. 2 clarifies this process:

(1) Starting from the minimum NIR reflectance ρ_{NIR} determine the pixels (e.g. pixels between points P_1 and P_2) at the higher end of the red reflectance (ρ_{R}). For example, we have $\alpha=(\rho_{\mathrm{R,P1}}-\rho_{\mathrm{R,P2}})/(\rho_{\mathrm{R,Max}}-\rho_{\mathrm{R,Min}})$, and $\rho_{\mathrm{R,P1}}$, $\rho_{\mathrm{R,P2}}$, $\rho_{\mathrm{R,Max}}$, and $\rho_{\mathrm{R,Min}}$ are the red band reflectance at points P_1, P_2, maximum and minimum, respectively. P_1 is equal to the maximum in this case and, typically, an α=10–15% works well for this purpose.

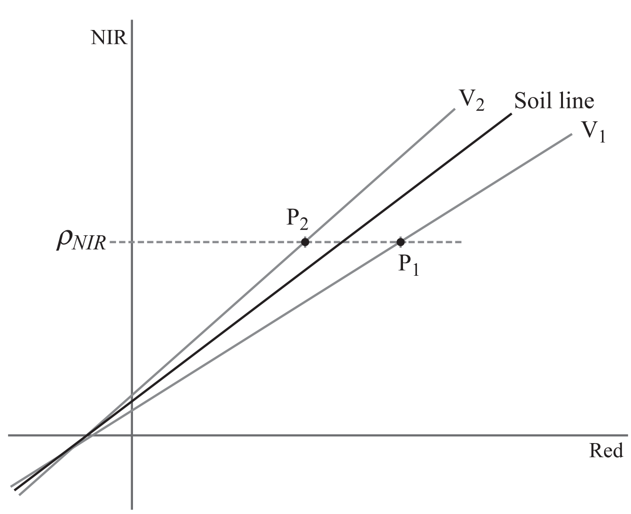

Fig. 2. Automatic extraction of soil line from the red–NIR scatterplot. Points P_1 and P_2 are determined by the upper percentile the red band reflectance at a given NIR reflectance (ρ_{NIR}). The soil line is the regression line of soil pixels within boundary lines V_1 and V_2.

(2) The pixels whose red reflectances are within cross points P_1 and P_2 are treated as soil pixels at the specific ρ_{NIR}.
(3) Repeat (1) for all $\{\rho_{\mathrm{NIR}}\}$ until ρ_{R} is equal to the overall maximum red reflectance. All soil pixels are located within the two boundary lines V_1 and V_2.
(4) Conduct a linear regression analysis to obtain the slope (α) and intercept (β) of the soil line.
(5) SRI can be calculated with

$$\mathrm{SRI} = \frac{\rho_s - \rho_1}{(\rho_2 - \rho_1)} \quad (1)$$

where ρ_1 and ρ_2 are the minimum and maximum band reflectances derived from the soil line (Fang & Liang, 2003).
(6) The soil reflectance (ρ_s) for a pixel is then calculated with

$$\rho_{is} = \rho_{i1} + (\rho_{i2} - \rho_{i1})*\mathrm{SRI} \quad (2)$$

where ρ_{i1} and ρ_{i2} are the minimum and maximum soil reflectance at band i.

The SRI was estimated successfully from MODIS for one of the study areas (BOREAS SSA, see Section 3.2 below) with the method above. This is not unexpected since its red–NIR scatterplot shows a typical vegetated surface. However, calculating SRI from a degraded red–NIR space, e.g. the USDA BARC area (Section 3.1), should be very meticulous. Fig. 3 shows different spatial resolutions and red–NIR spaces upscaling from 30-m (ETM+) to 1-km resolution. The MODAGG red–NIR scatterplot are also shown (Fig. 3e and j). The soil line rotates counter-clockwise when the spatial resolution decreases. In fact, those strips of pixels to the right are not pure soil pixels any more—they simply have more 'soil' than other pixels—which makes soil line identification range from difficult to impossible. Without proper pixel unmixing, it is unrealistic to determine the soil line automatically.

In this circumstance, we explored three methods to get SRI and the soil reflectance (Table 1). The first method uses the coherent high-resolution ETM+ red–NIR space and its soil line parameters. The scaling of surface reflectance could be treated as linear (Liang, 2000). The MODIS soil reflectances are calculated using the spectral transformation formulae between MODIS and ETM+ (Liang et al., 2002a). Soil line parameters obtained this way are very accurate. However, this method is limited since simultaneous ETM+ data are often not obtainable and their processing is time-consuming. The second method uses the soil line obtained from the 250-m MODIS level 2 bands 1 and 2 surface reflectance data. In principle, the SRI obtained in this way is more realistic than from ETM+ and more practical for simulating canopy reflectance. The 250-m MODIS is provided together with the 1-km data. However, the 250-m red–NIR space is also degraded compared with the ETM+ resolution (Table 1). The first two methods assume the soil and vegetation's reflectan-

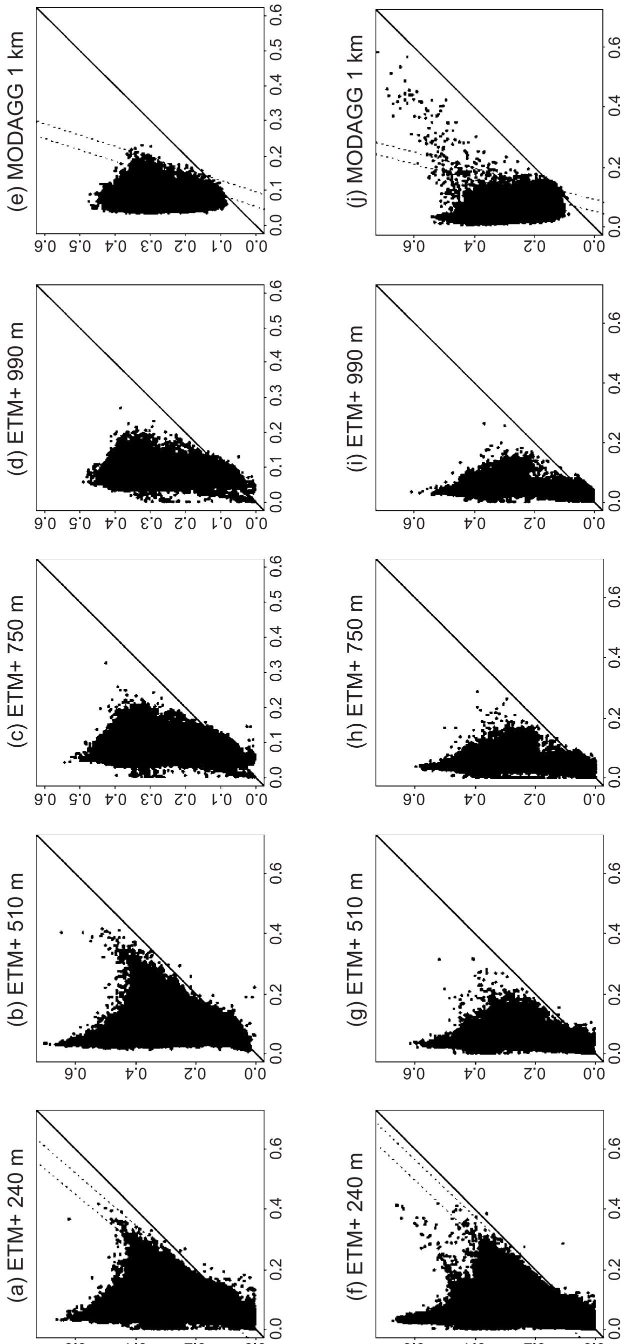

Fig. 3. The scatterplot of the red (abscissa) and near-infrared (ordinate) reflectances for USDA BARC ETM+ and MODAGG imageries on April 28, 2001 (a–e) and August 2, 2001 (f–j). The ETM+ scatterplots are of different resolutions (240, 510, 750, and 990 m). The strip within the dashed lines shows the soil pixels.

ces are constants; their fractions change relative to changing resolutions. Both methods require ancillary reflectance data, sometimes unobtainable, and thus impractical for global application. The third method sets the soil line parameters with the average values computed from literature and soil spectral library. The soil line slope generally varies from 1.0 to 1.3 and intercepts from 0.02 to 0.08 (Table 1). If the automatically calculated parameters are out of the normal range, the empirical values will be used in the simulation. In this study, α=1.2, β=0.04 were used for the USDA BARC area. To test how this selection affects the canopy reflectance and the LAI accuracy, two bias levels ($\pm$10%) were added to the slope. The relative canopy reflectance differences were calculated (see Section 2.2) (Table 2). Since we are using MODIS bands 1 (red) and 2 (NIR), Table 2 lists the average and maximum relative differences for the two bands. The average reflectance varies less than 4% and its maximum variation does not exceed 10%. The reflectance is more stable for higher LAI ($\geq$3). Based on our previous finding that the estimated LAI varies less than 0.5 under 10% reflectance amplitude (Fang & Liang, 2003), we are confident that the soil line parameters meet the accuracy requirement for LAI.

Table 1
Soil line parameters from various sources

	Slope (α)	Intercept (β)	Note
ETM+ image	1.1 (April 28, 2001)	0.04 (April 28, 2001)	
	1.05 (August 2, 2001)	0.04 (August 2, 2001)	
250 m MODIS	1.55 (April 28, 2001)	0.008 (April 28, 2001)	
	1.4 (August 2, 2001)	0.04 (August 2, 2001)	
Literature and soil spectral library	1.2	0.04	(Baret et al., 1993)
	1.253	0.03	(Baret et al., 1995)
	1.2		(Kimes et al., 2002)
	1.159	0.027	(Daughtry, 2001)
	1.00	0.08	JHL[a]

[a] JHL: soil reflectance measurements by J. Salisbury at the John Hopkins Univ.

2.2. Creating databases with canopy radiative transfer models

The MODIS LUT method needs to fix many canopy parameters. To account for the variations, the MODIS LUT algorithm classifies all canopies into six categories. We

Table 2
Relative mean and maximum canopy reflectance variations (%) calculated for $\pm$10% soil line slope biases from the MCRM model

	+10%		−10%	
	Red	NIR	Red	NIR
LAI<3				
Mean	3.68	1.61	−3.68	−1.60
Max	9.68	9.35	−9.67	−9.34
3≤LAI≤8				
Mean	2.66	1.24	−2.66	−1.22
Max	9.26	9.29	−9.65	−9.29

aimed at developing an alternative method adaptable to their variations. No universal canopy radiative transfer model for all canopies exists. Two canopy radiative transfer models were used in this study. The first one is the Markov chain reflectance model (MCRM), which simulates two-layer canopy reflectances for different solar zenith angles (SZA) (Kuusk, 2001). This model has a demonstrable usefulness in most canopies, except needleleaf forests (Fang & Liang, 2003; Kuusk, 1998). The second model is the GeoSAIL model (Huemmrich, 2001), which validation results show is suitable for needleleaf forest and works very well in the boreal area (Huemmrich, 2001).

2.2.1. Simulation with the MCRM model

In MCRM, the canopy is assumed to be horizontally uniform above a horizontal ground surface. This model has been used for the LAI and chlorophyll content retrieval from ETM+ (Kuusk, 1998). It has also been used to retrieve LAI from the POLDER data with a spatial resolution of 7 km (Bicheron & Leroy, 1999).

With the existing LUT method, the ideal way to create large databases is to discreterize more input values and run the RT model for all cases. Since current remote sensing satellites provide the capability of viewing the ground from different directions, it is worthwhile to test the hybrid approach with off-nadir simulations. Multiple viewing angles (MVA) may also improve the retrieval accuracy of land surface parameters.

The MCRM was run with variable solar zenith angles (θ_0: 20°, 30°, 35°, 40°, 45°, 50°, 55°, 60°, 65°, 70°, 75°), view zenith angles θ_v=0–90° by 10°, relative azimuth angles ϕ=0–180° by 15°, LAI (0.1–10 by 0.1), and different SRI values (0.01, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5, 0.6, 0.8, 1.0). Note that the SRI was derived from the empirical soil line. The PROSPECT model (Jacquemoud & Baret, 1990) was used to simulate leaf optical properties. Leaf biophysical parameters were specified by leaf chlorophyll A+B concentration (C_{ab}: 10–90 by 10 ug/cm^2), the Markov parameter (S_z: 0.4–0.9 by 0.1), and the effective number of elementary layers (N: 1.0–3.0 by 0.5). The leaf orientation was assumed spherical. An example of the output MODIS nadir red and NIR reflectances for different LAI and SRI is shown in Fig. 4. This figure is typical of the red–NIR scatterplot of a vegetation canopy. Fig. 4a and b exhibits similar regularities with different input settings. One major difference is the location of the convergence point, which is determined mostly by leaf optical properties and canopy structural features (Shabanov et al., 2002).

An example of the simulation in the principal plane is shown in Fig. 5. When the solar zenith angle is low (30°), the red reflectance is very low and insensitive to LAI change. When LAI increases, the red reflectance decreases very little but the NIR reflectance increases. The red reflectance decreases and NIR reflectance increases related to LAI increase when θ_0=50°. The hotspot effect is obvious for both solar zenith angles.

2.2.2. Simulation with the GeoSAIL model

The GeoSAIL model combined a turbid medium RT model (SAIL) with the Jasinski geometric model to simulate canopy spectral reflectance for discontinuous canopies (Huemmrich, 2001). In the simulation, two canopy components were selected, the leaves (or needles) and twigs. The leaf inclination angle was assumed spherical and twigs planophile. The optical properties of each component, parameterized with the Leaf Incorporating Biochemistry Exhibiting Reflectance (LIBERTY) (Dawson et al., 1998) modeled results, were integrated into the MODIS bands 1 and 2, respectively, with the spectral response function. The twig transmittance was assumed zero. Aspen tree crowns were modeled as cylinders and black spruce crowns use cones. The maximum plant area index was set to 5.0 and 7.0 for aspen and spruce, respectively. Canopy height/width ratios were calculated as 3.5 for aspen and 7.0 for spruce (personal communication with K.F. Huemmrich).

Leaf optical properties for aspen and spruce, simulated from the LIBERTY, were modeled with various leaf cell diameter (m^{-6}) (20–80 by 10), intercellular air space (0.01–0.05 by 0.01), (1, 2, 3, 4, 5, 7, 10), baseline absorptions of (0.0004, 0.0006, 0.0008, 0.001), albino absorptions set at (1, 3, 5, 7, and 9), chlorophyll contents of (100, 200, 300, 400, and 500 mg/m^2), and lignin and four cellulose contents: (10, 30, 50, and 80 g/m^2) (Dawson et al., 1998). Fig. 6 shows an example of leaf reflectance (6a) and transmittance (6b) when the leaf cell diameter is fixed at 20 μm. The view angle is fixed at nadir but GeoSAIL provides a framework for simulating angular reflectances. There are three other free variables: the solar zenith angle, the canopy coverage, and the background reflectance. The soil reflectance index was derived from the MODIS surface reflectance data (MOD09) and was used to calculate the background reflectance for each wavelength. A number of different background fractions (0–1.0 by 0.1) were simulated for spruce and aspen. An example of the simulated red and NIR reflectances for both spruce and aspen is plotted in Fig. 6c and d. Generally, the red reflectance increases with increasing SRI and NIR increases with LAI. The spruce's NIR variation is comparatively smaller than that of aspen.

2.3. Nonparametric training and prediction

Two nonparametric methods were applied in the training process and to predict the LAI. The general idea is that, if we can distinguish needle leaf forest from other canopies, we can apply these two nonparametric regression equations to predict LAI from all canopies. Both the neural network method and the projection pursuit regression method are briefly outlined below.

2.3.1. The neural network method

Basically, the neural network method establishes a mapping function between the simulated reflectance field and the corresponding biophysical variable of interest

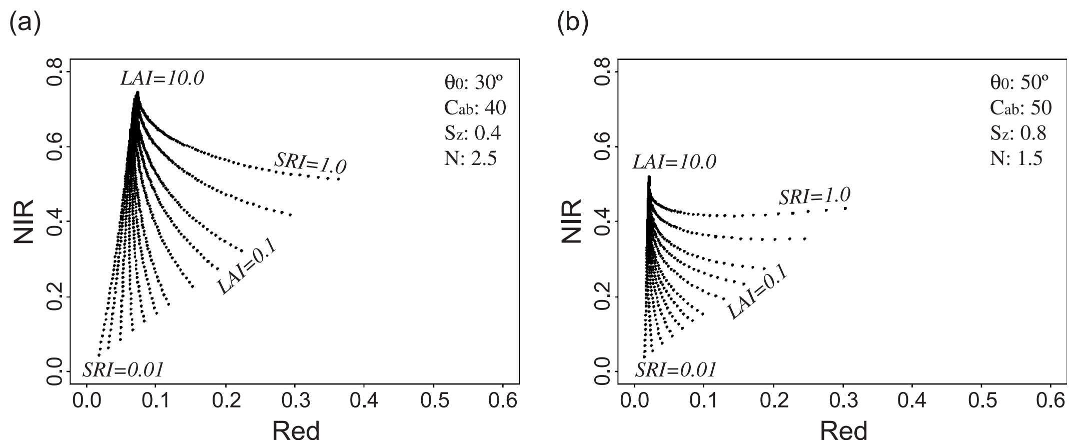

Fig. 4. Examples of the simulated databases (MODIS red and NIR nadir reflectances) for the USDA BARC study site. The two examples are simulated with varying LAI and SRI, fixing (a) θ_0 (30°), C_{ab} (40), S_z (0.4), and N (2.5), and (b) θ_0 (50°), C_{ab} (50), S_z (0.8), and N (1.5). SRI={0.01, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5, 0.6, 0.8, 1}. LAI=0.1–10.0 by 0.1.

(Kimes et al., 2000). Some previous studies have applied neural network methods to invert physically based RT models. Most of these works (Gong et al., 1999; Qi et al., 2000; Smith, 1993) use the simulated database from an RT model in the training and checking processes (part of the simulated data were used for the training and the other testing). The disadvantage of incorporating simulated data is obvious since the simulated databases may not fully represent real environments. It is desirable to apply the training results to reflectance data derived from satellite sensors and calibrate the results with field-measured data.

The training process is usually computationally intensive. Since some of the satellite bands are closely related, only the most information-rich bands are ordinarily applied in the training iteration. Commonly used bands include green, red, NIR and the NDVI, either as single bands or in

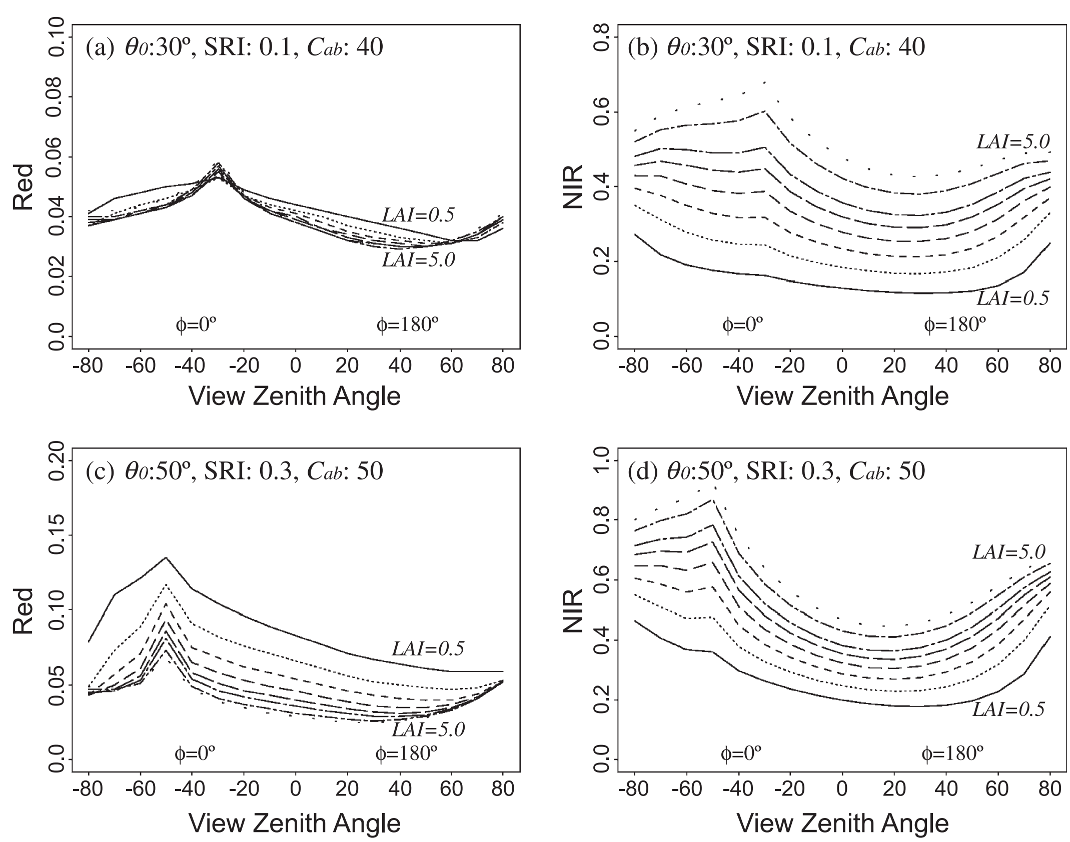

Fig. 5. Simulation of MODIS red and NIR reflectances for various viewing directions. The figure shows reflectances in the principal plane with varying viewing zenith angles and LAI. (a) and (b) are red and NIR reflectances for θ_0 (30°), SRI (0.1), and C_{ab} (40). (c) and (d) for θ_0 (50°), SRI (0.4), and C_{ab} (60). LAI={0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 4.0, and 5.0}.

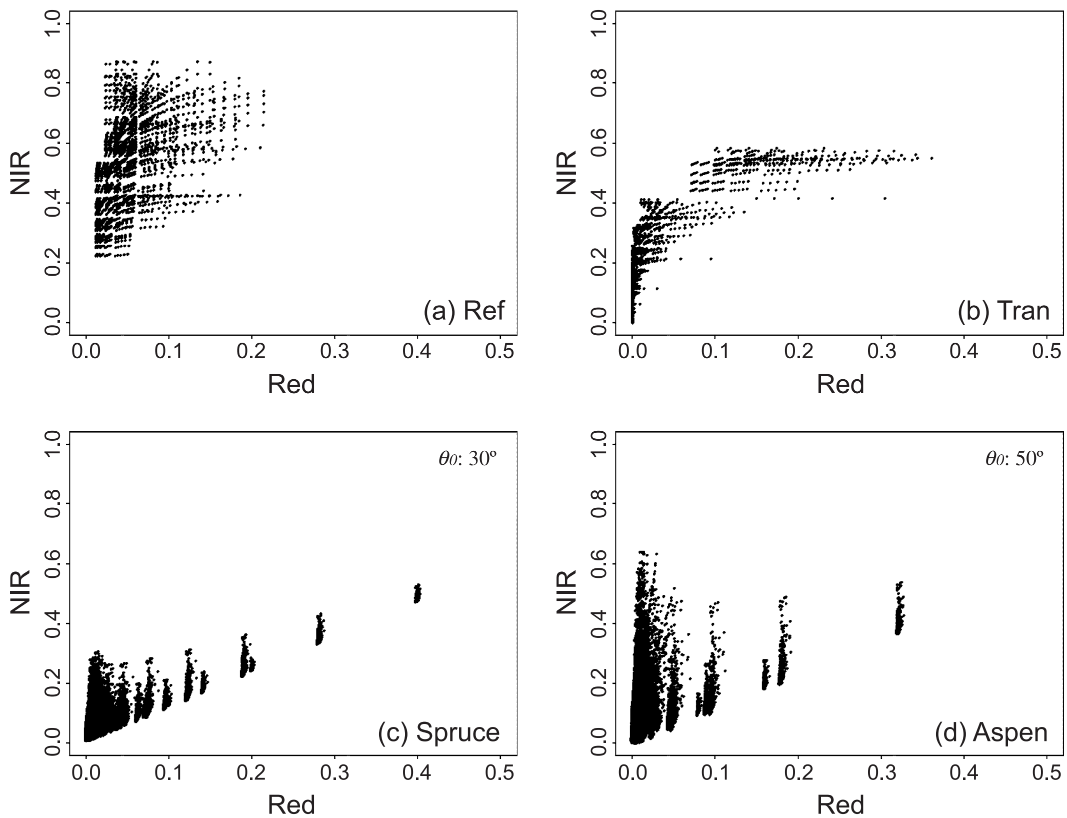

Fig. 6. Leaf optics and canopy reflectances for the BOREAS SSA study site. (a) Leaf reflectance and (b) transmittance (leaf cell diameter=20 m^{-6}). (c) and (d) are the nadir red–NIR reflectances for spruce and aspen, respectively.

combinations (Baret et al., 1995; Kimes et al., 2002; Qi et al., 2000; Smith, 1993). Obviously, different MODIS band combinations can be used to invert LAI from simulated databases. In another similar study, different band combinations were tested to determine their effect on the final LAI accuracy estimated from EO1 ALI data (Liang et al., 2003). In this study, only the MODIS red (band 1) and NIR (band 2) bands were used in the training and prediction process. This selection is similar to the strategy applied in the MODIS LAI algorithm. In addition, our tests have shown that using red and NIR bands can produce accuracy equivalent to using all bands.

2.3.2. PPR training and prediction

Projection pursuit regression is another nonparametric multiple regression method. Its mathematical form is (Friedman & Stuetzle, 1981):

$$Y = a + \sum_{j=1}^{M} W_j \alpha^T X \tag{3}$$

where X and Y are the independent (e.g. LAI) and dependent variables (reflectance) respectively. α is the transformation vector. The dimension (or term) M is to be chosen by the user. W_j is the weights for different terms. Eq. (3) uses an additive model on predictor variables, which are formed by projecting X in M carefully chosen directions.

Many statistical software producers provide the NN and the PPR packages (e.g. Splus, Venables & Ripley, 1994). The PPR process proceeds with the same database used in the NN training process. The database was segregated into different angular bins based on their SZA and VZA (view zenith angle) ranges. For example, the angular bin used for USDA BARC is 20–30° for SZA and 0–30° for VZA. The training procedure was executed for each angular bin to create an input–output relationship between reflectance and LAI for each angular bin. In practice, all angular bins are trained and the results from a specific bin are used to estimate LAI from images with the specific angular setting. In this study, the aggregated ETM+, the MODAGG, and MOD09 red (band 1) and NIR (band 2) reflectances were used to map the LAI based on the trained input–output relationship.

In practice, training with only the simulated databases can lead to aberrant results in the final LAI map because the database may not represent the real environment. To overcome this problem, real image pixels were added to the database before training. The bare ground points are especially necessary because they have zero LAI (or at least very low) and are easily over fitted. Note that we did not include LAI=0.0 in the simulation process (Section 2.2). In this study, the nonvegetated pixels (bare soils, constructed areas, seashores, etc.) were extracted from MODIS imagery using NDVI (<0.3), and bands 6 (<0.1) and 7 (<0.09) as the thresholds. Bare ground reflectances and LAI (=0.0) were added to the database and in the training procedure to represent actual natural conditions.

3. Study area and data preparation

Two study sites were chosen to test the new approach. The first one is centered at the U.S. Department Agriculture (USDA) Beltsville Agricultural Research Center (BARC), Maryland, which is typical middle latitude temperate climate with mixed agricultural land, deciduous broadleaf forest, pasture, and developed areas. The other one is the Boreal Ecosystem–Atmosphere Study (BOREAS) Southern Study Area (SSA), Canada, an area dominated by evergreen needle forests (56.9%) and grasslands (33.6%). The major biome types used by the MODIS LAI algorithms for these two areas appear in Fig. 7.

3.1. USDA BARC

The USDA BARC, listed as one of the initial 24 NASA EOS land validation core sites (Morisette et al., 2002), is adjacent to the NASA Goddard Space Flight Center (GSFC) (39.03°N, 76.85°W). The MODIS image of this area ranges over the Baltimore–Washington metropolitan region and the Chesapeake Bay region. This area is fortunately under the center of a Landsat ETM+ scene (path 15/row 33). In the site, a series of field campaigns were conducted over recent years to measure surface reflectance and LAI data (Fang et al., 2003).

MODIS LAI collection 4 products were downloaded through the EOS data gateway. In the MODIS land data production sequence, the level 2G (MOD09) and level 3 data are accumulated to create the LAI products as well as other land data products (Justice et al., 2002). The MODIS Level 2G (MOD09) data are daily, cloud-cleared, and atmospherically corrected surface reflectances. The level 3 MODIS aggregated (MODAGG) data are 1-km intermediate products aggregated and binned daily from the MOD09 1–7 channels. The MODAGG data are used as the primary input for the MOD43B BRDF/Albedo product (MOD43B), the nadir BRDF adjusted reflectance (NBAR or MOD43B4), and the 16-day enhanced vegetation index product (MOD13A2). MODAGG is directly used to produce the MODIS LAI products because it has the required projections and spatial resolution. These constitute very sound reasons to utilize MODAGG data to test the new approach.

Over the study area, two clear Landsat ETM+ imageries were obtained on April 28 and August 2, 2001, respectively (path 15/row 33). They represent two different vegetation growing states. In late April, the vegetation and crops are in typical early-spring growth. In early August, crops are in the middle-late growing season. ETM+ and MODIS are in the same orbit, about 45 min apart, on Landsat 7 and Terra, respectively. The MODIS imagery over the test site was acquired on the same day as ETM+. The MODIS imagery

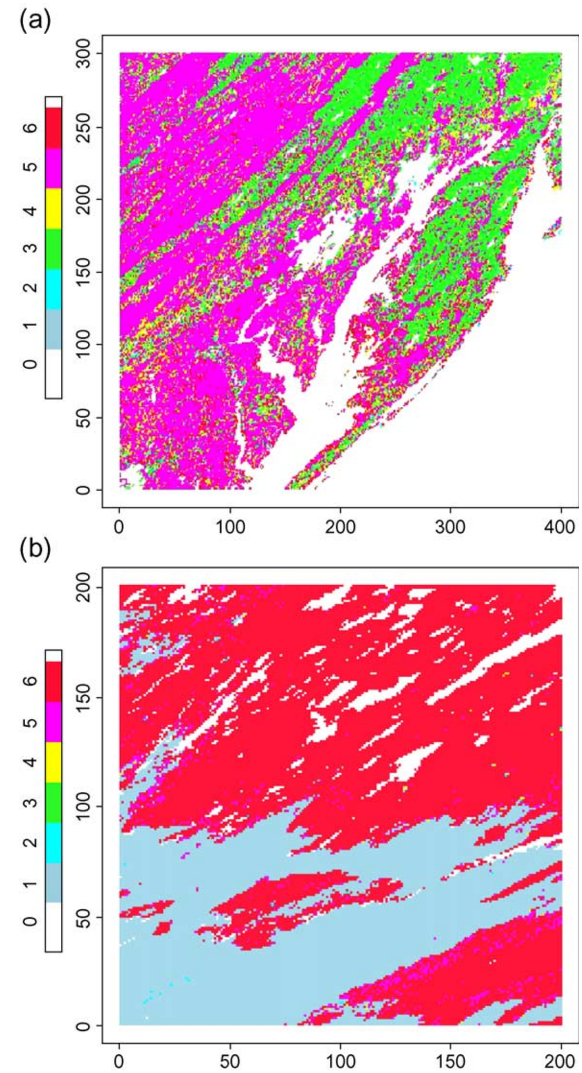

Fig. 7. MODIS LAI biome types at (a) the USDA BARC study site and (b) BOREAS SSA site. Color numbers: 0=water or unclassified, 1=grasses/cereal crops, 2=shrubs, 3=broadleaf crops, 4=savannah, 5=broadleaf forest, 6=needleleaf forest.

had a very small viewing angle (<1° for both images at the 10×10-km core site or <8° over the ETM+ coverage). Hence, both ETM+ and MODAGG reflectance were treated as nadir view. The SZA were 31.41° and 30.19° for the two ETM+ images, respectively, and 27.22° and 25.55° for the MODIS images.

3.1.1. Companion ETM+ data processing

The companion ETM+ imagery in BARC was atmospherically corrected using the measured aerosol optical depths and water vapor content from Sunphotometers (Holben et al., 1998). For heterogeneous haze, a cluster match algorithm was used to estimate the aerosol optical depth and retrieve surface reflectance (Liang et al., 2001). After the ETM+ reflectance data were calibrated with field-measured reflectance, they were aggregated into the MODIS resolutions (Liang et al., 2002b).

The ETM+ data from band 3 and band 4 in the study area were spatially averaged using commercial image software to generate data of 240-, 510-, and 990-m resolutions (Fig. 3), close to the MODIS 250-, 500-, and 1000-m resolutions. This aggregation was accomplished by simply averaging every 8×8, 17×17, and 33×33 ETM+ pixels. The spatial averaging was just a simplification of the complicated spatial convolution and resampling process to aggregate ETM+ imagery to MODIS sensor (Barker et al., 1992). The aggregated 510-m resolution ETM+ imagery was then registered with the 1-km MODIS imagery by manually selecting the common ground control points (GCP), such as rivers, coastal lines, and other distinct features. It is clear that the accuracy of validation depends to a large extent on the accuracy of the spatial registration. In this study, an average registration error of less than one pixel was achieved for both months' images.

To further validate the ETM+ image processing, the 510-m ETM+ reflectances were aggregated to MODIS bands and compared with the actual MODIS data. Fig. 8 compares the red and NIR reflectances for both broadleaf crops and broadleaf forest for April 28 and August 2, 2001. Their mean red/NIR differences are less than 0.005/0.02 for broadleaf crops and less than 0.013/0.015 for broadleaf forests (Fig. 8). Similar accuracy was reported in one of our earlier studies to validate MOD09 surface reflectance with ETM+ data (Liang et al., 2002a). In that paper, the standard deviation of the differences between the retrieved ETM+ surface reflectance and MOD09 products are 0.015 and 0.035 for bands 1 and 2, respectively.

3.2. BOREAS SSA

Selection of the BOREAS SSA (53.656°N, 105.323°W) is particularly for investigating the performance of the hybrid approach for the needleleaf forest. The SSA topography is gentle, with a relief from 550 to 730 m. The western part of SSA is in the Prince Albert National Park (PANP) and the eastern region falls within and around the Narrow Hills Provincial Forest (Newcomer et al., 2000).

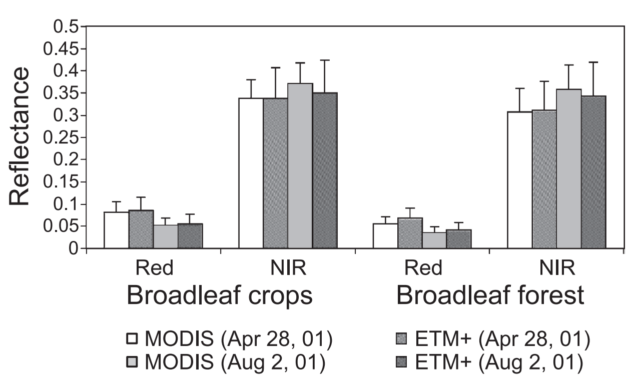

Fig. 8. Comparison of the aggregated ETM+ mean reflectances with the actual MODIS bands for both broadleaf crops and broadleaf forest for April 28, 2001 and August 2, 2001. The error bar shows the standard deviation.

The Southern Study Area has six main sites in and around the PANP and Narrow Hills Provincial Forest. They are fen (FEN), old aspen (OA), old black spruce (OBS), old jack pine (OJP), young aspen (YA), and young jack pine (YJP).

The BOREAS CD-ROM set (Newcomer et al., 2000) is very useful in this study. It includes high spectral resolution reflectance and transmittance factors of individual leaves, needles, twigs (reflectance only), and substrate at the SSA FEN, YJP, YA, and OBS sites (Walter-Shea, 2000). There are also spectral reflectance data for aspen bark and leaves from OA, YA, and ASP sites (Kharouk & Rock, 2000).

MODIS LAI collection 4 products were obtained through the EOS Data Gateway (EDG). For comparison, collection 3 LAI data were also downloaded. MODIS surface reflectance data (MOD09A1, 500 m) on August 5, 2002 were obtained from the MODLAND-BOREAS SSA site (Morisette et al., 2002). Different from the MODAGG used for BARC region, MOD09 data were used directly to estimate LAI for BOREAS SSA.

The BOREAS CD-ROM set also contains ground measurements of LAI made from August 9, 1993 to September 19, 1994 (Chen & Geng, 2000). LAI maps have been calculated from Landsat TM images on August 9, 1991 through the empirical relationship between the field LAI values and the Reduced Simple Ratio (RSR) vegetation index (Brown et al., 2000). Other data sets in the CD-ROM include daily green fraction of absorbed photosynthetically active radiation (FPAR), needle-to-shoot area ratio, clumping index, plant area index, and foliage-to-total area index. These ancillary information sets were used in the database simulation and LAI estimation processes.

4. Results

4.1. USDA BARC area

Fig. 9 shows a color composite image of the LAI derived from the MODIS standard products, the aggregated ETM+ and the MODAGG in the Baltimore–Washington region. This figure also depicts the registration results between the aggregated ETM+ and the MODIS data. The ETM+ image is diamond-shaped and covers the center of the rectangle. MODIS and MODAGG LAI values are displayed in red and blue, respectively, within the rectangle. There are three major vegetation types in the area (Fig. 7): broadleaf forest (50.3%), broadleaf crops (26.9%), and needleleaf forest (10.2%). There are some grasses and cereal crops but they only account for 3.5% of the vegetation. Besides, there are some shrubs, savanna pixels, and unvegetated pixels.

Fig. 10 shows the MODIS standard LAI product from both the main RT method and the empirical backup method on May 1, 2001 and August 5, 2001, respectively. The MODIS LAI quality control (QC) mask was employed-only those pixels whose QC are labeled "main method", "main method with saturation", or "empirical method" were used.

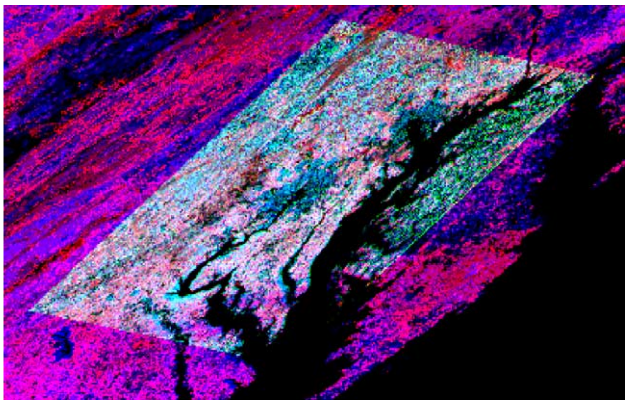

Fig. 9. Presentation of the registration of the MODIS LAI product (red), the aggregated ETM+ LAI (green) and the MODAGG LAI (blue) of the USDA BARC site. The ETM+ (the smaller lozenge region) and MODAGG data are of August 28, 2001. The MODIS LAI product is of May 1, 2001.

In this figure, all water, barren, permanent wetlands/marshes, and built-up areas are filled with zero values. Moreover, all clouds and shadows were excluded in the subsequent comparison work. Table 3 compares the minimum (min), mean, and maximum (max) LAI estimated by different approaches. At the beginning of the growing season (Fig. 10a), the first (0.25) and third (0.75) quantiles of LAI are located at 1.8 and 5.9, respectively. In the middle of the growing season (Fig. 10c), nearly all areas had a larger LAI (more green areas in the figure). The first and third quantiles increased to 3.1 and 6.1, respectively. The mean LAI increased from 3.9 to 4.9 from May 1 to August 5, 2001; maximum LAI reached for both dates was 6.8. Table 4 compares the mean and standard deviation of LAI estimated with different QC masks. LAI values derived from the main method (RT), the main method with saturation (RTs), and the empirical backup method (VI)

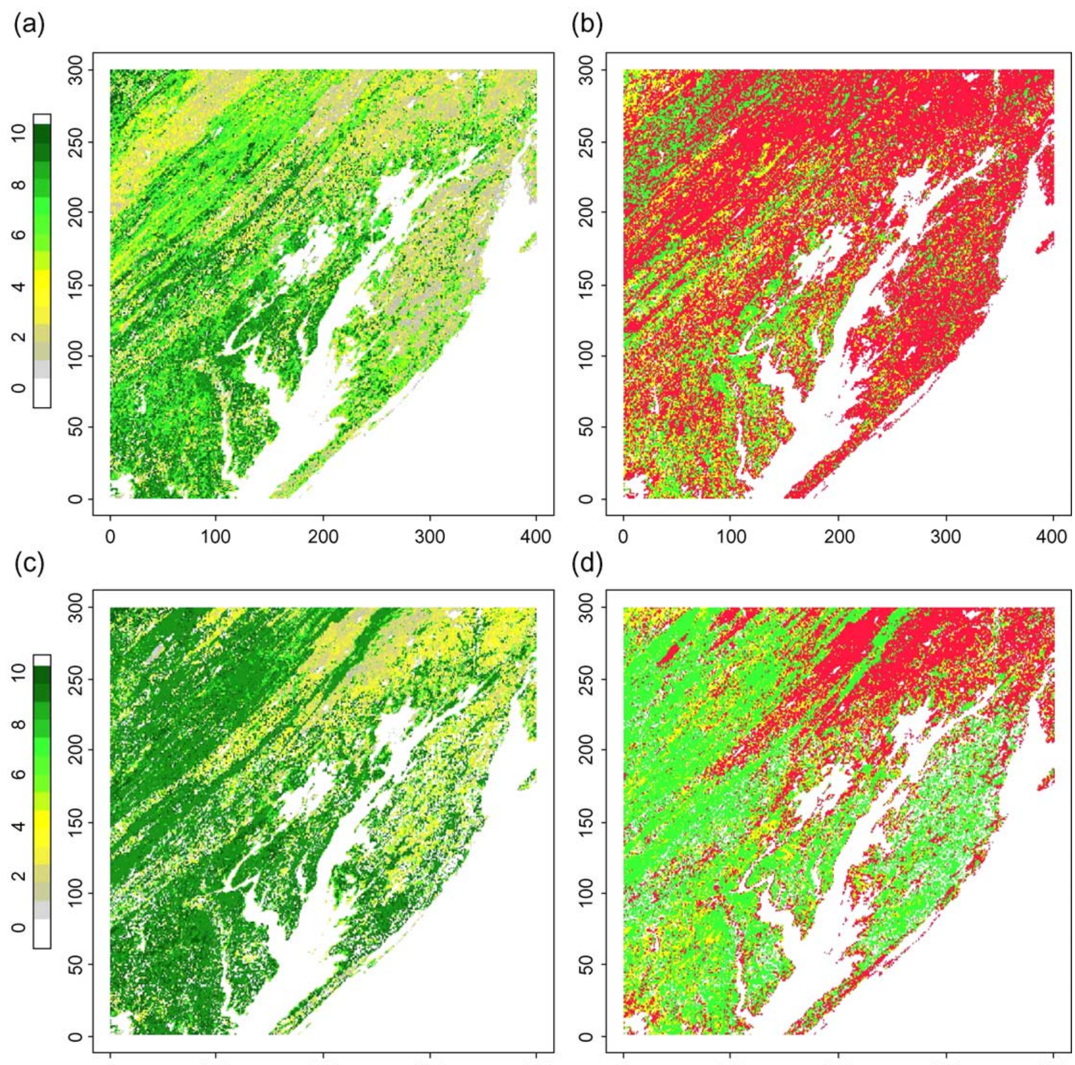

Fig. 10. MODIS LAI products for the USDA BARC study area on (a) May 1, 2001 and (c) August 5, 2001. (b) and (d) depict the MODIS LAI algorithms for the two dates. Red, yellow, and green color represent the main radiative transfer method, the main RT method with saturation, and the empirical backup method, respectively.

Table 3
Statistics of the MODIS LAI products, and LAI estimated from the aggregated ETM+ data and from the MODAGG data in the BARC study site

		April 28, 2001			August 2, 2001		
		Min	Mean	Max	Min	Mean	Max
MODIS LAI[a]		0.2	3.9	6.8	0.1	4.9	6.8
ETM+ LAI	NN	0.0	1.8	7.5	0.0	2.4	9.2
	PPR	0.0	1.6	7.6	0.0	2.3	7.3
MODAGG	NN	0.0	1.9	4.9	0.0	2.6	9.1
LAI	PPR	0.0	1.6	4.5	0.0	2.5	7.3

The minimum (Min), mean and maximum (Max) values of the neural network (NN) and the projection pursuit regression (PPR) methods are shown for two dates.

[a] The MODIS LAI data were of May 1 and August 5, 2001, respectively.

were compared separately. The RT method was mainly used for the broadleaf crops. On April 28, 2001, nearly two thirds of the LAI was produced with the main RT method. The RTs and VI masks produced abnormally higher results (nearly double) than the main RT pixels.

LAI maps for April 28 and August 2, 2001 were generated from MODAGG data using the neural network and the projection pursuit regression methods (Fig. 11). The estimated LAI values ranged from 0.1 to 4.9 on April 28 and from 0.1 to 9.1 on August 2 with the NN method (zero values were masked). The mean LAI values increased from 1.9 to 2.6 from April 28 to August 2, 2001 (Table 3). The PPR method produced very similar results to the NN approach and both results agree very well, spatially. The regional biome map (Fig. 7) shows those high LAI values corresponding to broadleaf forests. Temporal dynamics of the LAI distribution are visually distinct in Fig. 11. On April 28, most of the green patches were located in the central and southwest portions of the image; interestingly, on August 2, the green patches shifted toward the northwest mountain areas.

The LAI maps from the aggregated ETM+ imagery are shown in Fig. 12 with a spatial resolution of 510 m. The neural network-generated spatial pattern of the LAI maps (Fig. 12a and c) is consistent with those generated by the PPR approach (Fig. 12b and d). The color gradient of the ETM+ LAI maps differs from that of the MODAGG results. More regions become greener from April 28 to August 2, which means that the LAI has increased. This trend is also shown statistically. Table 3 shows that the estimated mean LAI increased from 1.8 to 2.4 with the NN and from 1.6 to 2.3 with the PPR. The statistical results of Table 3 represent slightly different areas for ETM+ and MODAGG because of their different resolutions and spatial coverage. For example, the maximum LAI on April 28, 2001, estimated from ETM+ (7.5/7.6 for NN/PPR) were larger than MODAG1G (4.9/4.5). Note that the SZA has a five (5) degree difference between the ETM+ and MODIS images.

LAI was estimated successfully from either the nadir-view or the MVA simulations (Table 4). The absolute LAI difference brought by angular simulation is within −0.5 to +0.5. The maximum difference is for MODAGG LAI with the PPR approach (difference=0.6). Both NN and PPR provided similar LAI. This illustrates that our approach is very useful for estimating LAI from multiple viewing satellites. The simulated database can be applied to MODIS images with large viewing angles. Since the MODIS images here are close to nadir view, more studies are needed when the method is extended to other situations.

4.1.1. Comparison of LAI distribution

To further compare the results from our approach with those from the MODIS standard products, the histograms of the coherent MODIS LAI, ETM+ LAI, and MODAGG LAI are compared in Fig. 13. In Fig. 13, the MODIS QC mask

Table 4
Mean and standard deviation (in brackets) values of the MODIS LAI products, and LAI estimated from the aggregated ETM+ data and from the MODAGG data in the BARC study site for April 28, 2001 and August 2, 2001, respectively

		April 28, 2001			August 2, 2001		
		VI (18.59%)	RTs (16.35%)	RT (65.06%)	VI (47.78%)	RTs (18.84%)	RT (33.38%)
MODIS LAI[a]		5.1 (1.5)	6.4 (0.34)	3.0 (1.62)	5.7 (1.0)	6.3 (0.5)	2.9 (1.4)
Nadir view simulation							
ETM+ LAI	NN	2.2 (0.92)	2.2 (0.87)	1.7 (0.95)	2.7 (0.7)	2.3 (0.7)	2.1 (0.7)
	PPR	1.9 (0.83)	1.9 (0.77)	1.5 (0.83)	2.7 (0.9)	2.2 (0.8)	1.9 (0.8)
MODAGG	NN	2.3 (0.67)	2.2 (0.44)	1.7 (0.63)	2.9 (0.6)	2.4 (0.4)	2.1 (0.6)
LAI	PPR	1.8 (0.53)	1.8 (0.42)	1.5 (0.56)	2.9 (0.7)	2.4 (0.7)	2.1 (0.7)
MVA simulation							
ETM+ LAI	NN	2.1 (0.81)	2.0 (0.74)	1.7 (0.82)	2.9 (0.8)	2.5 (0.7)	2.1 (0.7)
	PPR	2.2 (0.83)	2.1 (0.74)	1.7 (0.83)	3.1 (0.8)	2.6 (0.8)	2.3 (0.8)
MODAGG	NN	2.2 (0.64)	2.1 (0.35)	1.5 (0.59)	3.2 (0.6)	2.6 (0.5)	2.2 (0.7)
LAI	PPR	2.4 (0.62)	2.2 (0.36)	1.7 (0.57)	3.4 (0.6)	2.9 (0.5)	2.4 (0.8)

LAI were derived with the NN and the PPR method from both the nadir view (θ_v=0°) and MVA simulations. RT, RTs, and VI correspond to pixels controlled with collection 4 QC mask (the main RT method, the RT method with saturation, and the backup method, respectively).

[a] The MODIS LAI data were of May 1 and August 5, 2001, respectively.

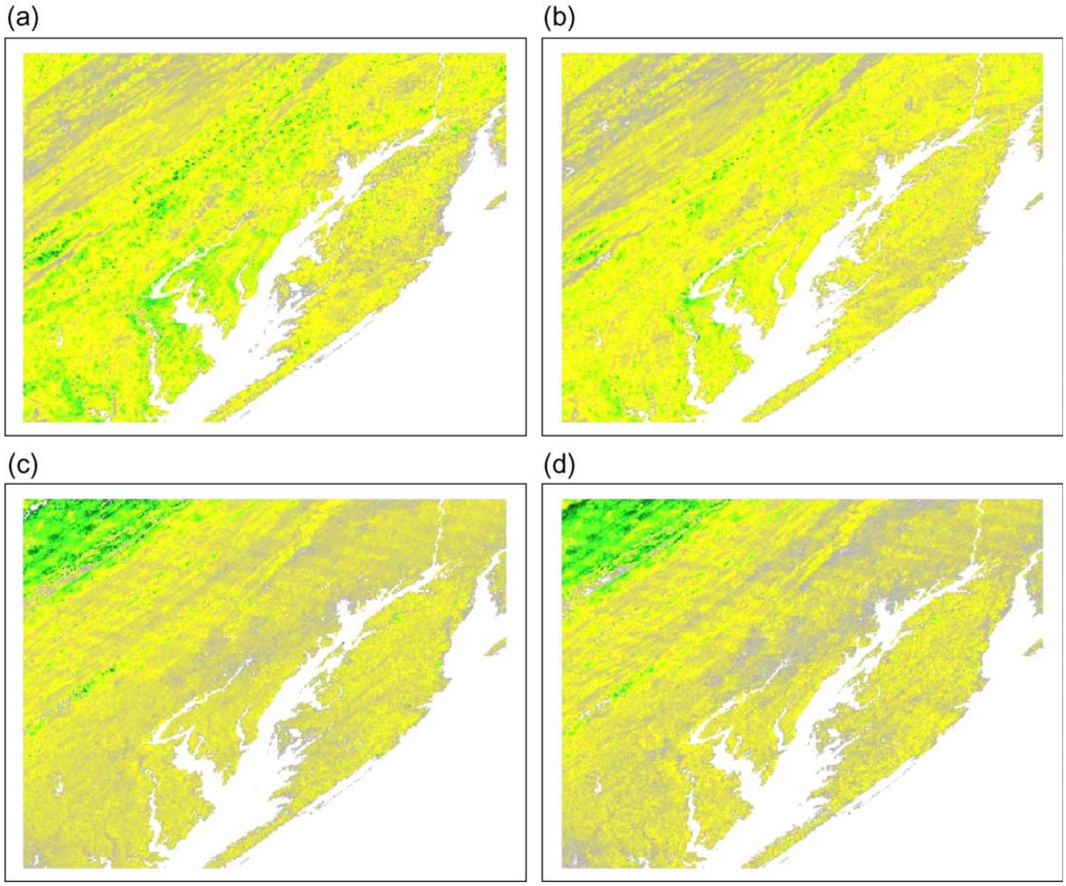

Fig. 11. LAI estimated from the USDA BARC MODAGG data (1 km). The top two are for August 28, 2001 and the bottom two August 2, 2001. (a) and (c) are from the neural network algorithm; (b) and (d) are with the projection pursuit regression method, respectively. LAI legend same as Fig. 10.

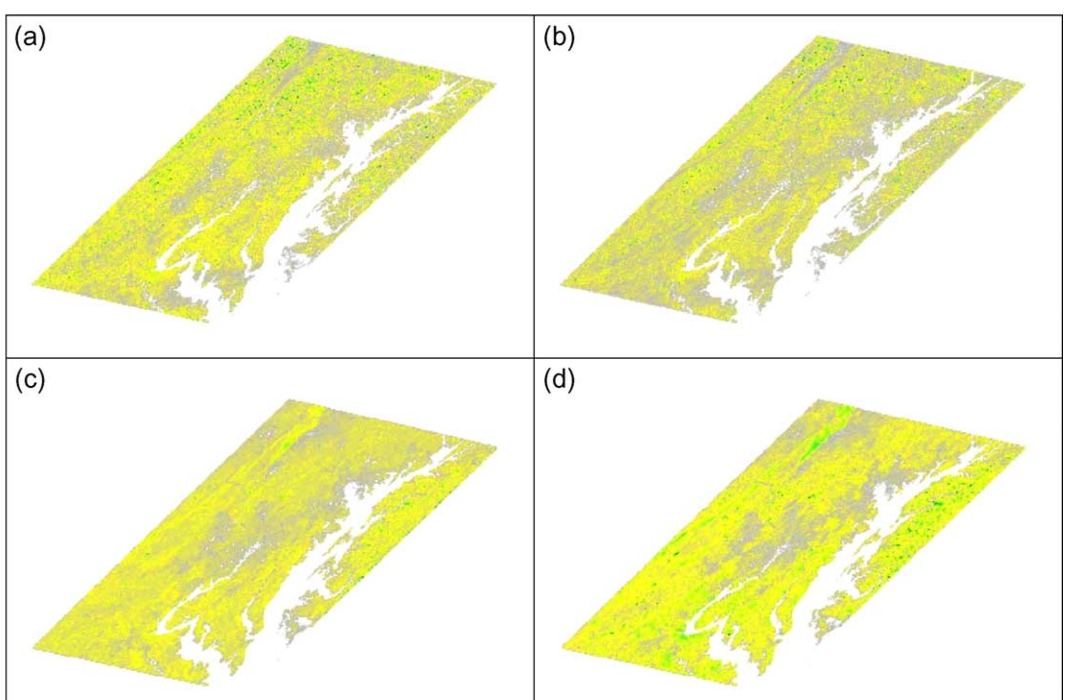

Fig. 12. LAI estimated from aggregated ETM+ reflectance (510 m) for April 28, 2001 (a–b) and August 2, 2001 (c–d). (a) and (c) are from the neural network algorithm; (b) and (d) are with the projection pursuit regression method. LAI legend same as Fig. 10.

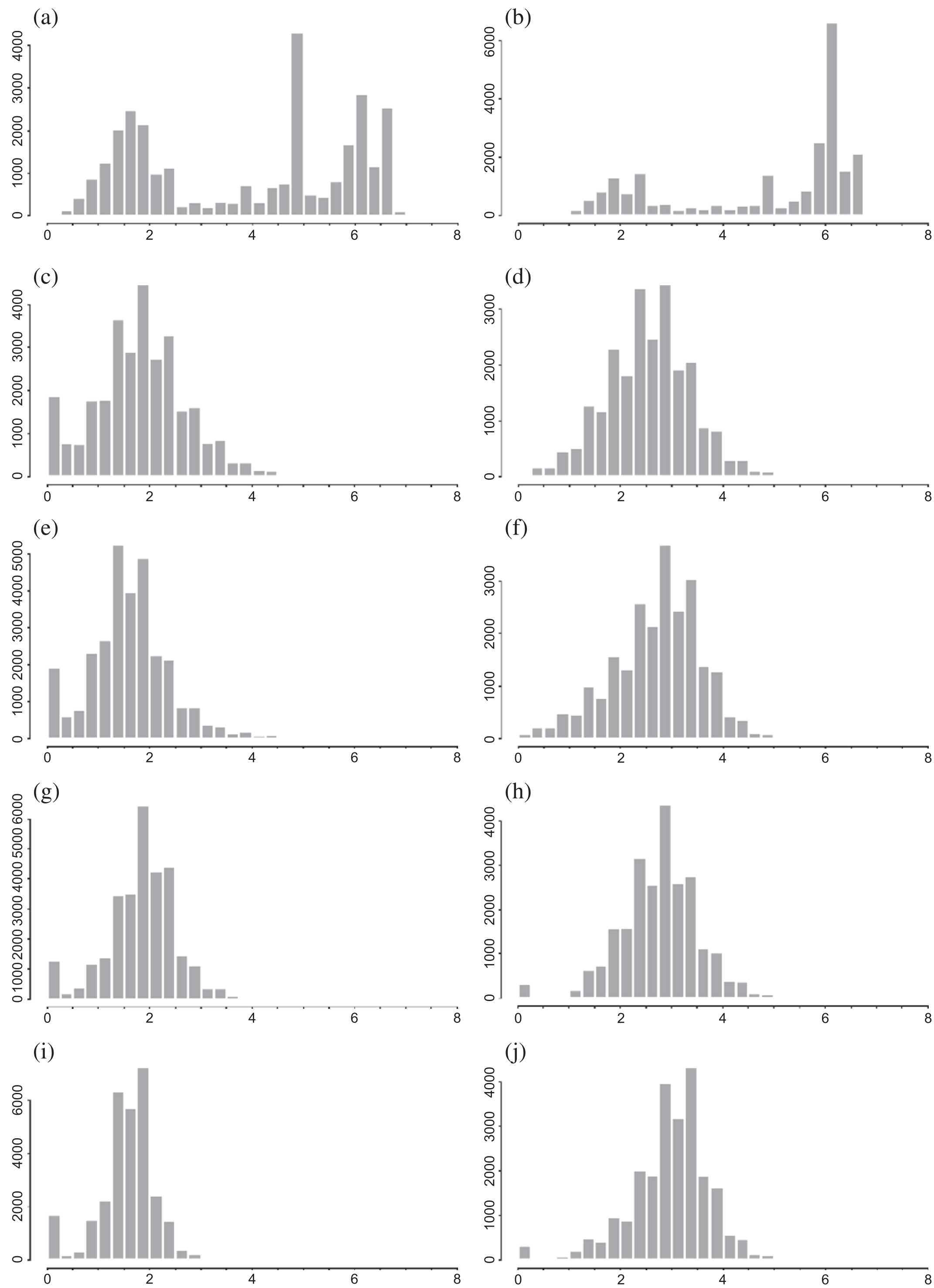

Fig. 13. Histogram comparison of the MODIS LAI products and LAI derived from ETM+ and MODAGG data. The two columns represent the results of two dates (April 28 and August 2, 2001, respectively). (a–b) MODIS LAI products, (c–d) from ETM+ with NN method, (e–f) from ETM+ with PPR method, (g–h) from MODAGG with NN method, and (i–j) from MODAGG with PPR method.

has been applied to exclude the filled values (≥249). The same mask was used to delineate each unique geographic region for all three datasets. The MODAGG LAI (Fig. 13g–j) agreed well with ETM+ LAI (Fig. 13c–f). Their histograms closely resemble one another although the absolute LAI values may differ. The NN and the PPR produce nearly identical results as shown by their LAI distributions. However, MODIS LAI products have significantly more pixels peaked at the high end and less pixels in the middle range of LAI in comparison to the ETM+ and MODAGG LAI outputs. This led to very high MODIS LAI mean values (3.9 and 4.9 in Table 3)—however, their LAI ranges were reasonable in view of the field LAI and literature report (Asner et al., 2003). The percent tree cover of this study area (15–45%) is about half of other typical temperate forests (DeFries et al., 2000) at the 1-km scale. Extensive ground measurements demonstrate that typical LAI of this study area are about 1.2–3.5 for corns, 2.5–5.5 for soybeans, 2.0–5.0 for grasses, and 2.1–3.6 for mixed forests. It is noted that there are several dense wheat fields having high LAI values (above 6.0). The LAI derived from both ETM+ and MODAGG are representative of the mean and seasonal LAI characteristics of this study area. Some of the maximum ETM+ and MODAGG LAI are larger than 8.0, which are treated as outliers in this study. A close examination of these pixels is needed in the future to identify the causes.

To further compare the LAI results from different schemes, the MODIS biome type information was used in the comparison (Fig. 14). This was to demonstrate the discrepancies caused by different classes. On April 28, 2001 (Fig. 14a), the MODIS LAI for broadleaf crops are nearly the same as those estimated from our hybrid approach. In this study area, about one third of the pixels are broadleaf crops. The disagreement for grasses (<0.2) and shrubs (<0.8) is very small. Large discrepancies are observed for savannah (2.5), broadleaf forests (3.7), and needleleaf forests (3.5). On August 2, 2001 (Fig. 14b), very similar

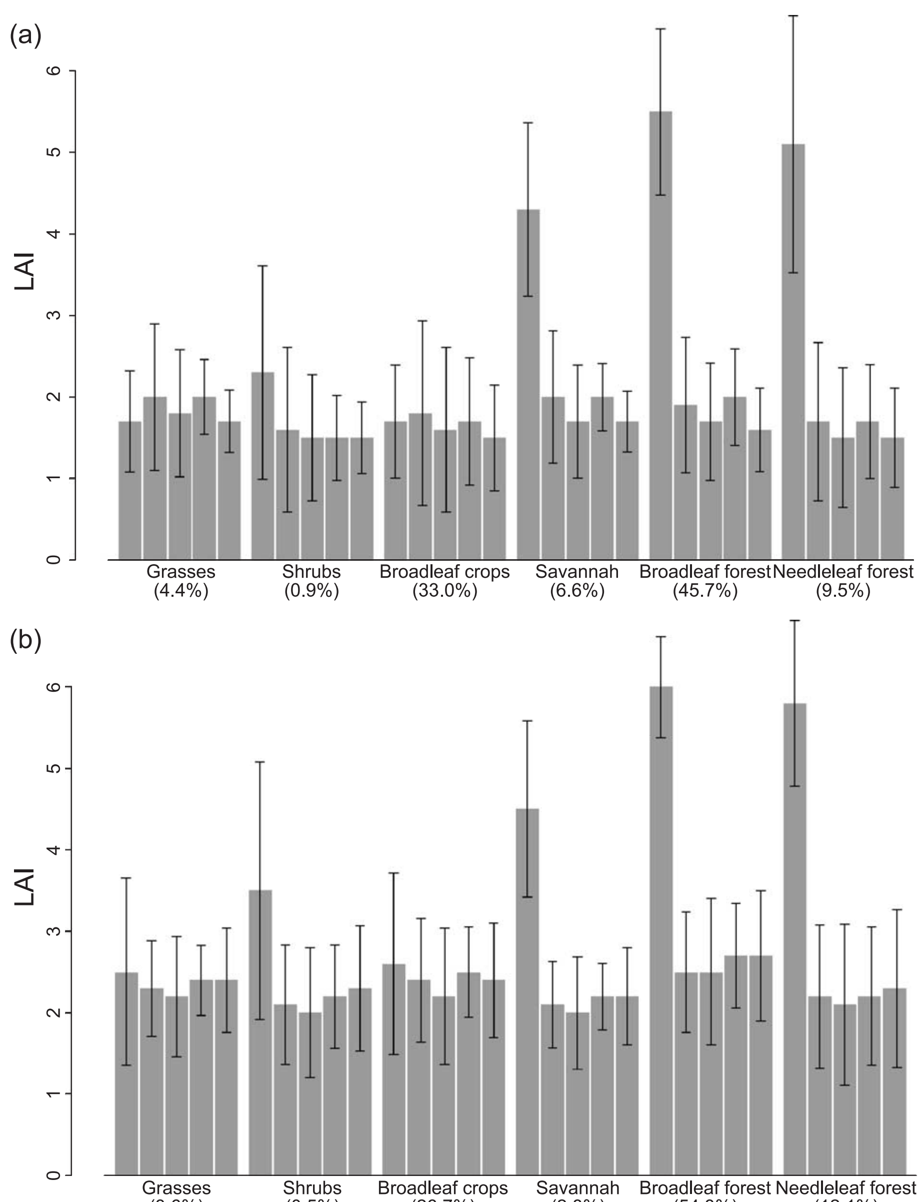

Fig. 14. Mean (bar) and standard deviation (line segment) of the MODIS LAI (first bar), LAI-NN (second), and LAI-PPR (third) estimated from ETM+ and LAI-NN (fourth) and LAI-PPR (fifth) estimated from MODAGG data for different biomes of the USDA BARC region. (a) April 28, 2001 and (b) August 2, 2001. Numbers in brackets are area fractions.

phenomena are seen. The bias for grasses, shrubs and broadleaf crops are very small. There are some biases for shrubs (1.3); however, the fraction of shrubs (0.5%) is negligible. For savannah, broadleaf forests, and needleleaf forests, the biases are as large as 2.3, 3.3, and 3.5, respectively. In general, the MODIS LAI products over-estimated the forests by about 2.0–3.0 in this site, but are consistent with our outputs for grasses and broadleaf crops.

4.2. BOREAS SSA

Two major vegetation zones are distinct in the BOREAS SSA site, mixed grass in the south, and boreal forest in the north (Fig. 7b). Two different simulation databases were applied for the two vegetation canopies. The LIBERTY+-GeoSAIL database was used for the boreal forest and the MCRM for the grassland. MODIS LAI products for BOREAS SSA are shown in Fig. 15, together with the LAI estimated from MOD09 with both NN and PPR approaches. Table 5 compares the LAI for both grasses and needleleaf forests estimated with different approaches. From the MODIS LAI products (collection 4), the LAI for this area ranges from 0.1 to 6.8 and the average LAI for grassland is 1.0 and forest 3.7.

Surface LAI was estimated with the NN and PPR approaches. From the LAI-NN (Fig. 15c and d), the mean LAI for grass is 1.6 and forest 3.4. From the LAI-PPR (Fig. 15e and f), the mean LAI for grass and forest are 1.9 and

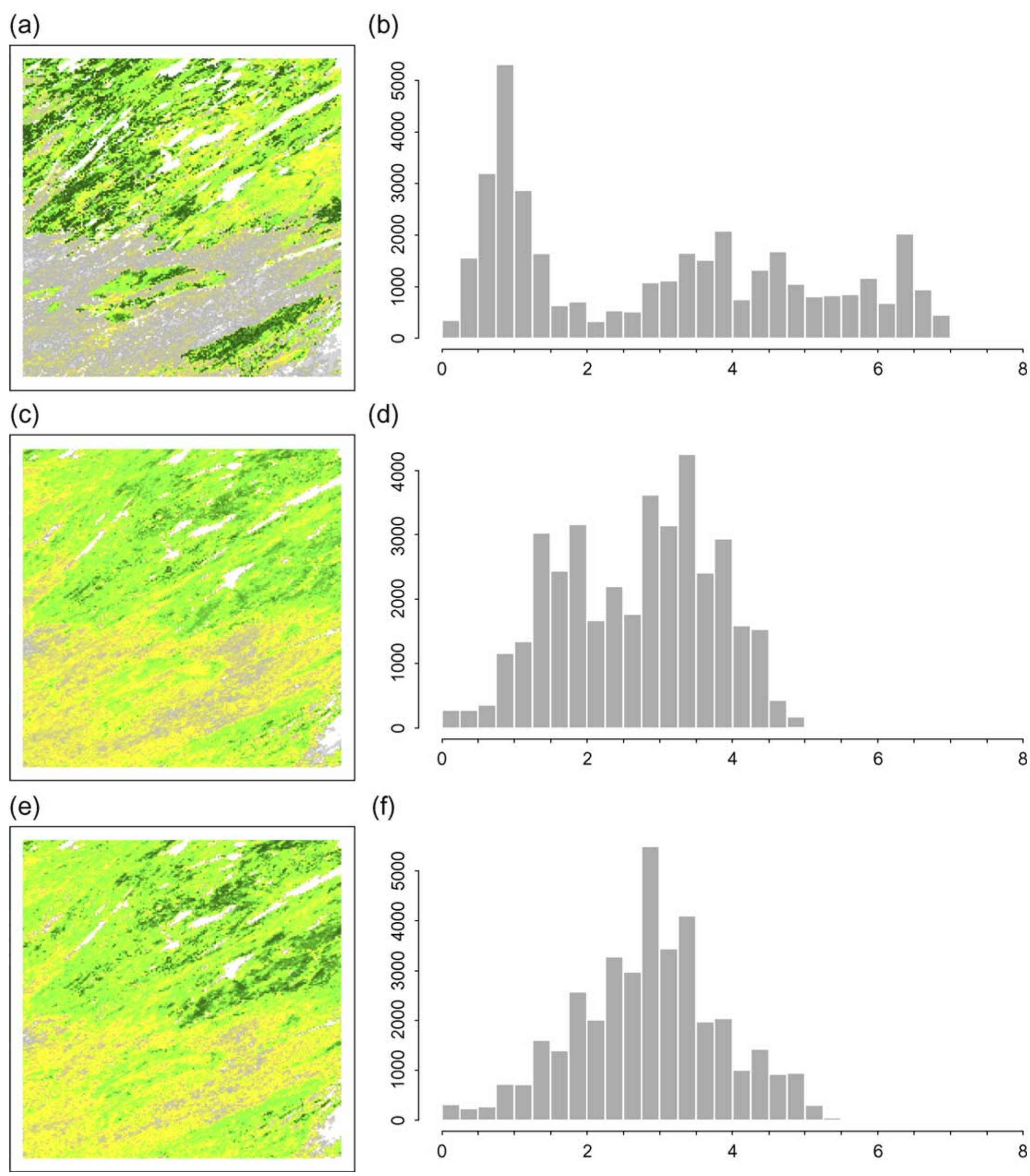

Fig. 15. MODIS LAI products (a) and LAI estimated from MOD09 with the neural network (c) and projection pursuit regression method (e) for the BOREAS SSA site. (b), (d), and (f) are the histograms of the retrieved pixels (for areas, refer to Fig. 16a).

Table 5
Mean and standard deviation (in brackets) of MODIS LAI products, and LAI estimated from MOD09 data with the neural network (LAI-NN) and projection pursuit regression (LAI-PPR) method, respectively, for grasses and needleleaf forests on August 5, 2001 in the BOREAS SSA site

		MODIS LAI	LAI-NN	LAI-PPR
Grasses	(C4:RT)	0.9 (0.32)	1.7 (0.58)	2.0 (0.64)
	(C4:VI)	1.9 (0.70)	2.6 (0.71)	2.9 (0.57)
	(C3:RT)	1.6 (1.04)	1.6 (0.51)	1.9 (0.58)
	(C3:VI)	2.9 (1.49)	2.1 (0.63)	2.5 (0.57)
Forests	(C4:RT)	3.6 (1.09)	3.4 (0.72)	3.5 (0.83)
	(C4:RTs)	6.3 (0.41)	3.1 (0.38)	3.1 (0.33)
	(C4:VI)	4.4 (2.09)	3.1 (1.13)	3.1 (1.09)
	(C3:RT)	4.6 (1.33)	3.3 (0.62)	3.4 (0.71)
	(C3:VI)	5.1 (1.41)	3.6 (0.83)	3.6 (0.88)

LAI from both collection 4 (C4) and collection 3 (C3) are divided into the main (RT) method, the main method with saturation (RTs), and the empirical vegetation index (VI) method based on their QC, respectively.

3.5, respectively. For the forest, the LAI-PPR is nearly identical to LAI-NN, but for grass, the LAI-PPR is a little higher than the LAI-NN. For forest, both LAI-NN and LAI-PPR mean values are similar to the MODIS LAI products, but for grass, they are marginally higher than the latter.

Fig. 15 also displays the histograms of MODIS LAI and LAI derived from MOD09. In the histograms, the filled values and the QC mask in MODIS LAI product have been applied to exclude nonvegetation pixels and a common mask was used to get the same comparison region for all the three datasets. Two LAI peaks, corresponding to grasses and forests, respectively, are obvious in the MODIS LAI histogram. The LAI-NN has two local maximums. LAI-PPR is analogous to a normal distribution and there are no observable grass and forest peaks. More pixels in MODIS LAI are positioned on the high LAI side, a similar phenomenon as the BARC site.

Fig. 16 illustrates the QC mask used for both collection 4 and collection 3 LAI products. Table 5 presents the LAI values from different approaches for both grasses and forests. The main (RT) method was used to produce the MODIS LAI products for most pixels (80.05%). Only a small portion of pixels (16.72%) was processed with the empirical backup (VI) method in collection 4. The VI-processed pixels mainly belong to forest (15.35%). In collection 3, 26.52% of processed pixels used the backup method for both grasses and forests (Table 5). The retrieval index (RI), which indicates the success of the main RT method and the reliability of the retrieval, was calculated (Wang et al., 2001). The RI increased from 0.61 to 0.80 in collection 4, an increase seen also in the number of processed pixels (Fig. 16).

The main RT method and the backup VI method produced different LAI values (Table 5). For example, the mean LAI of grass is 0.9 when obtained with the RT method and 1.9 via the VI method. This phenomenon is also distinct for the forest pixels. The VI method generally produced higher LAI than the RT method. The standard deviation of VI method is also higher and unacceptable in some cases (e.g. forest). Our NN and PPR approach produced similar LAI results with each other. The LAI-NN and LAI-PPR are higher (about 1.0) for grasses but are concomitant with the MODIS LAI products for forests (C4:RT). The LAI difference for grasses may be attributed to the background reflectance and the SRI applicability in this area. Forest LAI estimated with RTs and VI is problematic and needs further refinement.

The collection 3 LAI data were also compared in Table 5. A common mask was applied for both collection 3 and collection 4. Similar to collection 4, in collection 3, the VI method produces higher LAI values than the RT method

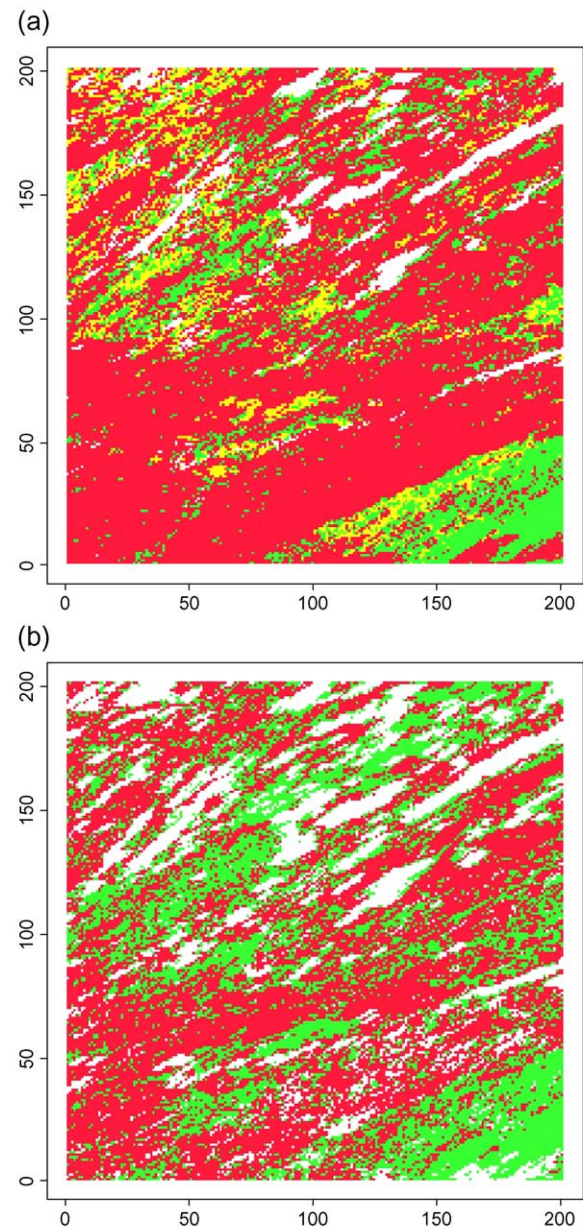

Fig. 16. MODIS LAI algorithms used for the BOREAS SSA site on August 5, 2002. Red, yellow, and green color represents the main RT method, the RT method with saturation, and the empirical backup method, respectively. (a) collection 4 and (b) collection 3. QC control with both collections 3 and 4 was applied.

for either biome type. In collection 3, the main RT approach produced fairly good LAI. Overestimation mainly occurs in the VI approach. For example, for the forest, MODIS LAI overestimated both LAI-NN and LAI-PPR by about 1.5 with the VI approach. The LAI results were significantly improved in the new collection even with the backup algorithm, especially for grasses (Table 5). In general, collection 4 MODIS LAI products have been substantially improved and are in a better consistency with our results.

4.2.1. Forest reflectance and LAI sensitivity test

In this study, the LIBERTY leaf optics model was used to simulate the leaf reflectance and transmittance required in the GeoSAIL canopy reflectance model. To test the sensitivity of the canopy reflectance to leaf optical parameters, two bias levels (±10% and ±20%) were added to the reference leaf reflectance and transmittance values obtained from the field measurements (Walter-Shea, 2000). The spruce forest result is shown in Table 6. In this table, the first line represents canopy reflectances calculated from the reference leaf optics with the GeoSAIL model. Other lines are absolute differences to the first line. For the red band, the maximum relative difference is −13.5% when both ρ and τ increase by 20%. For the NIR band, the maximum relative difference is 37.12% in the same condition. Comparatively, the NIR band is more sensitive to the leaf optics change. For a 10% leaf optics variation, the NIR reflectance could undulate as high as 30%. Higher errors occur when both ρ and τ shift in the same direction. The canopy reflectance errors decrease when ρ and τ change inversely. For example, the relative NIR difference is as low as −0.21% when the leaf reflectance increases 20% but the transmittance decreases 20% ($\rho_{+0.2}$,$\tau_{-0.2}$). However, the relative NIR difference increases to 35.25% when the two leaf optical parameters change the same way ($\rho_{+0.2}$,$\tau_{+0.2}$). These indicate leaf reflectance and transmittance are counter-balancing each other.

LAI values from different noise simulations were compared with those from the reference database (Table 6). For the neural network method, the R^2 of the LAI-NN changed little for all manner of leaf optical uncertainties. Nevertheless, this higher R^2 might be misleading considering the RMSE change. Similar results were observed for the PPR method. If ρ and τ do not change the same way, a good LAI accuracy can be obtained for even 20% leaf optical noise. Otherwise, a 10% error for both ρ and τ or 20% error for one parameter will bring unacceptable errors. Good LAI accuracy can be obtained for a 10% one-parameter bias.

In our method, the requirement for the number of land cover classes is not as strong as other algorithms. Generally, two types of vegetation, broadleaf and needleleaf, are sufficient for this purpose. However, in global land cover mapping contexts, confusion still exists between problematic classes. For example, needleleaf forest pixels can be misclassified as broadleaf forest, or vice versa. To address this issue, some experiments were carried out. For the BOREAS SSA site, a database created from a broadleaf simulation (MCRM) was applied intentionally. The estimated mean LAI for needleleaf forests decreased from 3.29 to 1.64. Similarly, for the USDA BARC site, a database from the needleleaf simulation (GeoSAIL) was used to estimate LAI. The estimated mean LAI increased from 1.31 to 1.94 for April 28, 2001 and from 2.2 to 2.54 for August 2, 2001. This shows how misclassification of needleleaf forest will lead to an underestimation and broadleaf forest

Table 6
Canopy reflectance and LAI variation to different leaf optical biases for the spruce forest in the BOREAS SSA site

Leaf optics	Canopy reflectance				LAI			
	Red		NIR		LAI-NN		LAI-PPR	
	Mean	Relative	Mean	Relative	R^2	RMSE	R^2	RMSE
ρ_0,τ_0	**0.0391**		**0.1541**		**0.9851**	**0.2455**	**0.9889**	**0.212**
$\rho_0,\tau_{-0.1}$	−0.0001	−0.0026	−0.0192	−0.1247	−0.0005	0.8791	−0.041	0.6595
$\rho_0,\tau_{+0.1}$	0.0001	0.0026	0.0296	0.1918	−0.0008	0.8107	−0.0078	0.8599
$\rho_{-0.1},\tau_0$	−0.001	−0.0264	−0.0201	−0.1301	−0.0005	0.8376	−0.0385	0.615
$\rho_{-0.1},\tau_{-0.1}$	−0.0011	−0.0288	−0.0343	−0.2223	−0.0016	1.9404	−0.1142	1.244
$\rho_{-0.1},\tau_{+0.1}$	−0.0009	−0.0238	0.0003	0.0023	0	0.0219	−0.0029	0.0448
$\rho_{+0.1},\tau_0$	0.001	0.0266	0.0273	0.1771	−0.0005	0.7011	−0.0061	0.7564
$\rho_{+0.1},\tau_{-0.1}$	0.0009	0.0238	−0.0002	−0.0015	0	0.0181	−0.0031	0.0361
$\rho_{+0.1},\tau_{+0.1}$	0.0011	0.0293	0.0474	0.3075	−0.0025	1.1643	−0.0074	1.1972
$\rho_0,\tau_{-0.2}$	−0.0002	−0.012	−0.0327	−0.2204	−0.0016	1.9084	−0.1031	1.2638
$\rho_0,\tau_{+0.2}$	0.0002	0.0127	0.0421	0.2811	−0.0012	1.0865	−0.0085	1.1805
$\rho_{-0.2},\tau_0$	−0.0021	−0.1195	−0.0358	−0.2411	−0.0018	1.9956	−0.1051	1.2278
$\rho_{-0.2},\tau_{-0.2}$	−0.0022	−0.1303	−0.0551	−0.3712	−0.01	4.0633	−0.306	2.1
$\rho_{-0.2},\tau_{+0.2}$	−0.0019	−0.1083	0.0008	0.0053	−0.0002	0.0825	−0.0051	0.1002
$\rho_{+0.2},\tau_0$	0.0021	0.1211	0.0527	0.3525	−0.0025	1.2296	−0.0091	1.257
$\rho_{+0.2},\tau_{-0.2}$	0.0019	0.1075	−0.0003	−0.0021	−0.0001	0.0662	−0.0037	0.0653
$\rho_{+0.2},\tau_{+0.2}$	0.0023	0.135	0.0527	0.3525	−0.0013	1.2229	−0.007	1.323

The bold line represents canopy reflectances and LAI from the reference leaf optics. Other lines are absolute differences to the first line. The relative column uses relative difference. The subscripts denote leaf reflectance (ρ) and transmittance (τ) change.

overestimation of LAI. In addition, needleleaf forest misclassification will cause higher LAI errors (about 1.5) than would broadleaf misclassification. The quantitative LAI difference seems very small for broadleaf misclassification, but their spatial distribution is unacceptable (not shown here). This demonstrates the importance of land cover classification. The LAI errors caused by misclassification may vary by time and forest type. Since this is not the focus of this paper, detailed comparisons will be discussed in subsequent papers.

4.3. NN and PPR comparison

Comparing the LAI results derived from both the NN and PPR methods is of great value. Following the simulated databases in Section 2.2, 80% of data were randomly generated for training and the other 20% for testing. The comparison for the two study sites is shown in Fig. 17. The results from the two nonparametric methods are very similar to each other (R^2=0.988 for the BARC site and 0.977 for the SSA site). Their RMSEs are less than 0.28, very small compared with the measurement and model uncertainties. Note that there are some outlying points for the BOREAS SSA site, especially for higher LAI (>3). These differences may come from some different sources, such as the application of the SRI for the area, leaf and canopy reflectance model uncertainties, and algorithm approximations.

5. Summary and discussion

In this paper, a hybrid approach integrating both radiative transfer simulations and nonparametric regression methods was proposed to estimate LAI from MODIS data. Two radiative transfer models, MCRM and GeoSAIL, were used for broadleaf canopy (crop, grass, and broadleaf forest) and needleleaf canopy (needleleaf forest), respectively. Two nonparametric methods, the neural network and the projection pursuit regression methods, were used. To adapt the impacts of underlying soil reflectance, an innovative SRI was devised.

The new approach was tested at two study sites, the USDA BARC site with broadleaf forest and crops, and the BOREAS SSA site with needleleaf forest and grasses. In the USDA BARC, SRI calculated from empirical soil line parameters was used to compute the MODIS soil reflectance. In BOREAS SSA, the background reflectances were calculated via SRI from the MODIS reflectance. MODIS standard LAI products were compared. For the USDA BARC site, the mean LAI values increased from 1.7 to 2.4 from April 28 to August 2, 2001. LAI maps generated with the MODAGG data were very similar to these with the aggregated ETM+ data. The MODIS LAI products agreed excellently with our results for broadleaf crops and grasses but overestimated broadleaf forests by 2.0–3.0. For the BOREAS SSA site, MODIS LAI products were consistent with our results from MOD09 data for both grasses and needleleaf forest. The weak performance of the MODIS algorithm at BARC is attributed to the complicated biome types of this area.

The NN and PPR produced very similar results because of their similar statistical mechanism. Both results agree very well, spatially. It is clear that NN and PPR provide two practical approaches to estimate LAI from MODIS images. The advantage of this hybrid method is that it requires little a priori information as opposed to the MODIS LUT method, which requires many different fixed parameters, although the training process may take a while. The hybrid designates vegetated pixels into two broad categories, much less challenging than the MODIS algorithm that requires six classes, and performs better on mixed landscapes. To make full use of the advantages of different radiative transfer models, different models were used for the two broad biome types, respectively. Further, the hybrid can be used at least as an independent validation method for the MODIS LAI product.

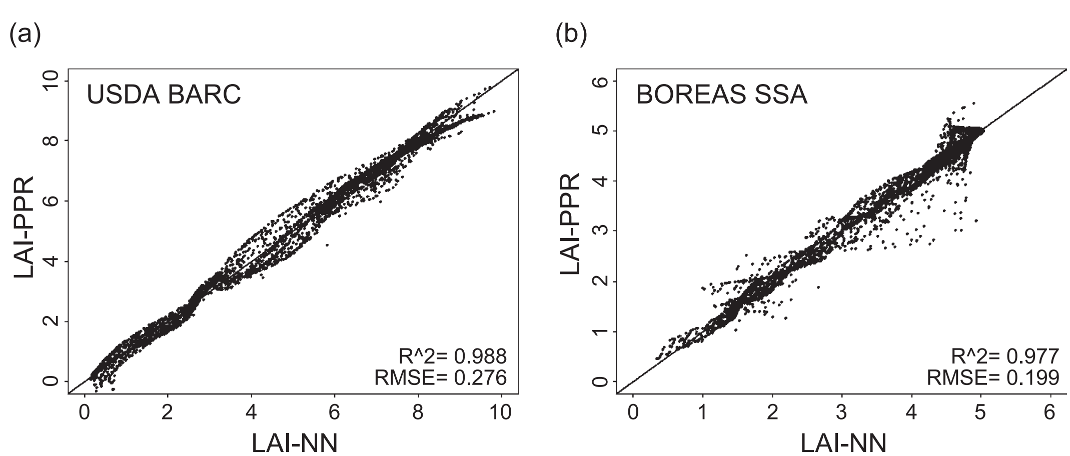

Fig. 17. Comparison of LAI derived from the neural network method (LAI-NN) and the projection pursuit regression (LAI-PPR) method for the USDA BARC (a) and the BOREAS SSA (b) study sites.

More validation studies are needed to evaluate the usefulness and limitations of this hybrid method for other landscapes, especially in sparsely vegetated sites (Fang et al., in press). More tests are also needed to evaluate how this algorithm might be used for other sensors, like the Multi-angle Imaging SpectroRadiometer (MISR). The initial results presented in this paper provide strong evidence supporting the method for MVA application.

To the best of our knowledge, manual determination of the slope and intercept of soil line was used in most studies referenced in the literature. In this study, an automatic soil line identification method was developed based on the pixels' red–NIR reflectance shape. However, automatic soil line identification is complicated and needs more refinements in degraded red–NIR spaces due to low spatial resolution and pixel mixing.

Acknowledgement

The work was partially funded by NASA grants. The authors thank Chad Shuey, Andy Russ, and Wayne Dulaney for their contribution to the field campaigns. They also wish to thank Drs. Andres Kuusk and Karl F. Huemmrich for providing their latest MCRM and GeoSAIL codes. Special thanks go to Drs. Feng Gao and Crystal Schaaf for their providing the MODAGG data.

References

Abuelgasim, A. A., Gopal, S., & Strahler, A. H. (1998). Forward and inverse modelling of canopy directional reflectance using a neural network. *International Journal of Remote Sensing*, *19*(3), 453–471.

Asner, G. P., Scurlock, J. M. O., & Hicke, J. A. (2003). Global synthesis of leaf area index observations: Implications for ecological and remote sensing studies. *Global Ecology and Biogeography*, *12*, 191–205.

Baret, F., Clevers, J., & Steven, M. (1995). The robustness of canopy gap fraction estimates from red and near-infrared reflectances: A comparison of approaches. *Remote Sensing of Environment*, *54*, 141–151.

Baret, F., Jacquemoud, S., & Hanocq, J. F. (1993). The soil line concept in remote sensing. *Remote Sensing Reviews*, *7*, 65–82.

Barker, J. L., Markham, B. L., & Burelbach, J. (1992). *MODIS image simulation from Landsat TM imagery*. In ASPRS/ACSM/RT 92 technical paper, 1 (pp. 156–165).

Bicheron, P., & Leroy, M. (1999). A method of biophysical parameter retrieval at global scale by inversion of a vegetation reflectance model. *Remote Sensing of Environment*, *67*, 251–266.

Broge, N. H., & Leblanc, E. (2001). 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*(2), 156–172.

Brown, L., Chen, J. M., Leblanc, S. G., & Cihlar, J. (2000). A shortwave infrared modification to the simple ratio for LAI retrieval in boreal forests: An image and model analysis. *Remote Sensing of Environment*, *71*(1), 16–25.

Chase, T. N., Pielke, R. A., Kittel, T. G. F., Running, S. R., & Nemani, R. (1996). Sensitivity of a general circulation model to global changes in leaf area index. *Journal of Geophysical Research*, *101*, 7393–7408.

Chen, J., & Geng, X. (2000). Retrieval of boreal forest leaf area index from multiple scale remotely sensed vegetation indices. In J. Newcomer, D. Landis, S. Conrad, et al. (Eds.), *Collected data of the Boreal ecosystem–atmosphere study: BOREAS CD-ROM set*. NASA.

Daughtry, C. S. (2001). Discriminating crop residues from soil by shortware infrared reflectance. *Agronomy Journal*, *93*(1), 125–131.

Dawson, T. P., Curran, P. J., & Plummer, S. E. (1998). LIBERTY-modeling the effects of leaf biochemical concentration of reflectance spectra. *Remote Sensing of Environment*, *65*, 50–60.

DeFries, R., Hansen, M., Townshend, J. R. G., Janetos, A. C., & Loveland, T. R. (2000). A new global 1 km data set of percent ree cover derived from remote sensing. *Global Change Biology*, *6*, 247–254.

Fang, H., & Liang, S. (2003). Retrieve LAI from Landsat 7 ETM+ data with a neural network method: Simulation and validation study. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(6).

Fang, H., Liang, S., & Kuusk, A. (2003). Retrieving leaf area index using a genetic algorithm with a canopy radiative transfer model. *Remote Sensing of Environment*, *85*(3), 257–270.

Fang, H., Liang, S., McClaran, M. P., Leeuwen, W. v., Drake, S., Marsh, S. E., et al. (2004). Biophysical characteristics and management effects on semiarid rangeland observed from Landsat ETM+ data. *IEEE Transactions on Geosciences and Remote Sensing* (in press).

Friedman, J., & Stuetzle, W. (1981). Projection-pursuit regression. *Journal of the American Statistical Association*, *76*, 183–192.

Gong, P., Wang, S. X., & Liang, S. (1999). Inverting a canopy reflectance model using a neural network. *International Journal of Remote Sensing*, *20*(1), 111–122.

Holben, B. N., Eck, T., Slutsker, I., Tanre, T., Buis, J., Setzer, A., et al. (1998). AERONET—A federated instrument network and data archive for aerosol characterization. *Remote Sensing of Environment*, *66*, 1–16.

Huemmrich, K. F. (2001). The GeoSail model: A simple addition to the SAIL model to describe discontinuous canopy reflectance. *Remote Sensing of Environment*, *75*, 423–431.

Jacquemoud, S., & Baret, F. (1990). PROSPECT: A model of leaf optical properties spectra. *Remote Sensing of Environment*, *34*(2), 75–91.

Jacquemoud, S., Baret, F., & Hanocq, J. F. (1992). Modeling spectral and bidirectional soil reflectance. *Remote Sensing of Environment*, *41*(2–3), 123–132.

Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., et al. (2002). An overview of MODIS land data processing and product status. *Remote Sensing of Environment*, *83*(1–2), 3–15.

Justice, C. O., Vermote, E., Townshend, J. R. G., Defries, R., Roy, D. P., Hall, D. K., et al. (1998). The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. *IEEE Transactions on Geoscience and Remote Sensing*, *36*, 1228–1249.

Kharouk, S., & Rock, B. N. (2000). The tree's bark input in tree–atmosphere interactions. In J. Newcomer, D. Landis, S. Conrad, et al. (Eds.), *Collected data of the Boreal ecosystem–atmosphere study: BOREAS CD-ROM set*. NASA.

Kimes, D., Gastellu-Etchegorry, J., & Estève, P. (2002). Recovery of forest canopy characteristics through inversion of a complex 3D model. *Remote Sensing of Environment*, *79*, 320–328.

Kimes, D. S., Knyazikhin, Y., Privette, J. L., Abuelgasim, A. A., & Gao, F. (2000). Inversion methods for physically-based models. *Remote Sensing Review*, *18*, 381–440.

Knyazikhin, Y., Martonchik, J. V., Diner, D. J., Myneni, R. B., Verstraete, M., Pinty, B., et al. (1998a). Estimation of vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MISR data. *Journal of Geophysical Research*, *103*, 32239–32256.

Knyazikhin, Y., Martonchik, J. V., Myneni, R. B., Dine, D. J., & Running, S. W. (1998b). Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *Journal of Geophysical Research*, *103*, 32257–32276.

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

Kuusk, A. (2001). A two-layer canopy reflectance model. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *71*(1), 1–9.

Liang, S. (2000). Numerical experiments on spatial scaling of land surface albedo and leaf area index. *Remote Sensing Reviews*, *19*, 225–242.

Liang, S. (2003). *Quantitative remote sensing of land surfaces* (p. 550). New York: John Wiley and Sons.

Liang, S., Fang, H., & Chen, M. (2001). Atmospheric correction of Landsat ETM+ land surface imagery: I. Methods. *IEEE Transactions on Geosciences and Remote Sensing*, *39*(11), 2490–2498.

Liang, S., Fang, H., Chen, M., Walthall, C., Daughtry, C., Morisette, J., et al. (2002a). Validating MODIS land surface reflectance and albedo products: Methods and preliminary results. *Remote Sensing of Environment*, *83*, 149–162.

Liang, S., Fang, H., Morisette, J., Chen, M., Walthall, C., Daughtry, C., et al. (2002b). Atmospheric correction of Landsat ETM+ land surface imagery: II. Validation and applications. *IEEE Transactions on Geosciences and Remote Sensing*.

Liang, S., Fang, H., Thorp, L., Kaul, M., Niel, T. G. V., McVicar, T. R., et al. (2003). Estimation and validation of land surface broadband albedos and leaf area index from EO-1 ALI data. *IEEE Transactions on Geosciences and Remote Sensing*, *41*(6), 1260–1267.

Morisette, J. T., Privette, J. L., & Justice, C. O. (2002). A framework for the validation of MODIS land products. *Remote Sensing of Environment*, *83*(1–2), 77–96.

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., et al. (2002). Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sensing of Environment*, *83*(1–2), 214–231.

Myneni, R. B., Ramakrishna, R., Nemani, R., & Running, S. W. (1997). Estimation of global leaf area index and absorbed par using radiative transfer models. *IEEE Transactions on Geosciences and Remote Sensing*, *35*(6), 1380–1393.

Newcomer, J., Landis, D., Conrad, S., Curd, S., Huemmrich, K., Knapp, D., et al. (2000). *Collected data of the Boreal ecosystem–atmosphere study: BOREAS CD-ROM set*. NASA.

Privette, J., Myneni, R., Morisette, J., & Justice, C. (1998). Global validation of EOS LAI and FPAR products. *Earth Observer*, *10*(6), 39–42.

Qi, J., Kerr, Y. H., Moran, M. S., Weltz, M., Huete, A. R., Sorooshian, S., et al. (2000). Leaf area index estimates using remotely sensed data and BRDF models in a semiarid region. *Remote Sensing of Environment*, *73*, 18–30.

Running, S. W., Nemani, R. R., Peterson, D. L., Band, L. E., Potts, D. F., Pierce, L. L., et al. (1989). Mapping regional forest evapotranspiration and photosynthesis by coupling satellite data with ecosystem simulation. *Ecology*, *70*, 1090–1101.

Shabanov, N. V., Zhou, L., Knyazikhin, Y., Myneni, R. B., & Tucker, C. J. (2002). Analysis of interannual changes in northern vegetation activity observed in AVHRR data from 1981 to 1994. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(1), 115–130.

Smith, J. A. (1993). LAI inversion using a backpropagation neural network trained with a multiple scattering model. *IEEE Transactions on Geoscience and Remote Sensing*, *31*, 1102–1106.

Venables, W. N., & Ripley, B. D. (1994). *Modern applied statistics with S-PLUS* (p. 462). New York: Springer-Verlag.

Walter-Shea, E. A. (2000). Radiation and gas exchange of canopy elements in a boreal forest. In J. Newcomer, D. Landis, S. Conrad, et al. (Eds.), *Collected data of the Boreal ecosystem–atmosphere study: BOREAS CD-ROM set*. NASA.

Walthall, C., Roujean, J. L., & Morisette, J. (2000). Field and landscape BRDF optical wavelength measurements: Experience, techniques and the future. *Remote Sensing Reviews*, *18*, 503–531.

Wang, Y., Tian, Y., Zhang, Y., El-Saleous, N., Knyazikhin, Y., Vermote, E., et al. (2001). Investigation of product accuracy as a function of input and model uncertainties: Case study with SeaWiFS and MODIS LAI/FPAR algorithm. *Remote Sensing of Environment*, *78*, 296–311.

Weiss, M., & Baret, F. (1999). Evaluation of canopy biophysical variable retrieval performances from the accumulation of large swath satellite data. *Remote Sensing of Environment*, *70*, 293–306.