Validation of Moderate Resolution Imaging Spectroradiometer leaf area index product in croplands of Alpilles, France

Bin Tan, Jiannan Hu, Ping Zhang, Dong Huang, and Nikolay Shabanov
Department of Geography, Boston University, Boston, Massachusetts, USA

Marie Weiss
Noveltis, Ramonville-Saint-Agne, France

Yuri Knyazikhin and Ranga B. Myneni
Department of Geography, Boston University, Boston, Massachusetts, USA

Received 2 April 2004; revised 19 August 2004; accepted 6 October 2004; published 12 January 2005.

[1] This paper presents results of validating the Collection 4 Moderate Resolution Imaging Spectroradiometer (MODIS) leaf area index (LAI) product using LAI data collected in a 3 × 3 km agricultural (grasses and cereal crops) area near Avignon, France, and 30 m resolution Enhanced Thematic Mapper (ETM+) image. Estimates of the accuracy, precision, and uncertainty with which the ETM+ data convey information about LAI underlie the derivation of a 30 m resolution reference LAI map by accounting for both field measurement and satellite observation errors. The 30 m reference LAI was then extrapolated from sampling points to a 58 km^2 area without loss in the quality and was degraded to a 1 km resolution LAI map. The latter was taken as a reference to assess the quality of the MODIS LAI product. Comparison of the reference and corresponding MODIS retrievals suggests that Collection 4 MODIS LAI is accurate to within an accuracy of 0.3 with a precision and uncertainty of 0.23 and 0.38, respectively. It was found that the Collection 3 MODIS land cover product, input to the Collection 4 operational LAI algorithm, misclassified the 58 km^2 area as broadleaf crops. The use of correct biome type in the operational processing improves the accuracy in LAI by a factor of 2 with an almost unchanged precision and uncertainty. Our results also indicate that the retrieval of LAI from satellite data is an ill-posed problem; that is, small variations in input due to observation errors result in a very low precision of the desired parameter. Any retrieval technique based on a simple model inversion or empirical relationships is unable to generate stable retrievals. The use of information on input errors in the retrieval technique is necessary to generate solutions to the ill-posed problem. The MODIS operational LAI algorithm meets this requirement.

Citation: Tan, B., J. Hu, P. Zhang, D. Huang, N. Shabanov, M. Weiss, Y. Knyazikhin, and R. B. Myneni (2005), Validation of Moderate Resolution Imaging Spectroradiometer leaf area index product in croplands of Alpilles, France, *J. Geophys. Res.*, *110*, D01107, doi:10.1029/2004JD004860.

1. Introduction

[2] Green leaf area index (LAI) and fraction of photosynthetically active radiation (FPAR) (0.4–0.7 μm) absorbed by vegetation characterize vegetation canopy functioning and energy absorption capacity. LAI and FPAR are key parameters in most ecosystem productivity models and global models of climate, hydrology, biogeochemistry, and ecology [*Sellers et al.*, 1997]. For effective use in large-scale models these variables must be collected over a long period of time and should represent every region of the terrestrial surface. The Moderate Resolution Imaging Spectroradiometer (MODIS) is an instrument on board the NASA Terra and Aqua platforms for remote sensing of the Earth's atmosphere, oceans, and land surface. The MODIS standard products include LAI and FPAR. These parameters have been operationally produced from day 1 of science data processing from MODIS. As MODIS LAI and FPAR data become publicly available through the Earth Resources Observation System (EROS) Data Center Distributed Active Archive Center (EDC DAAC), product quality must be ensured through validation.

[3] In general, validation refers to assessing the uncertainty of satellite-derived products by analytical compar-

0148-0227/05/2004JD004860$09.00

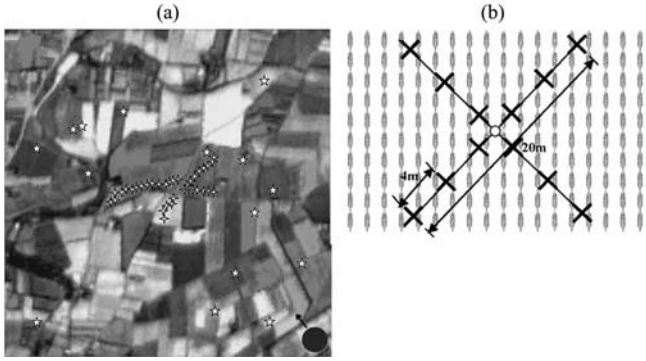

Figure 1. Alpilles field campaign sampling strategy. (a) A 3 × 3 km agricultural area in Alpilles ~25 km southwest of Avignon, France (43.81°N, 4.75°E), from Enhanced Thematic Mapper (ETM+) imagery. The direction from the Sun is depicted as a black circle with an arrow. Pluses depict sampling points on two centrally located transects. Stars depict off-transect sampling points. (b) Data collection at each off-transect point at 4 m intervals on two 20 m lines which form a regularly shaped cross for a total of 12 local points. The average of 12 local measurements was assigned to leaf area index (LAI) at the sampling point.

ison to reference data (e.g., in situ, aircraft, and high-resolution satellite sensor data), which are presumed to represent the target values [*Justice et al.*, 2000]. However, the uncertainty assessment of these products is not straightforward. The 1 km resolution of the MODIS LAI product significantly exceeds the plot size typically used for LAI field measurements. Thus a procedure is needed to correlate the scale of the LAI measurements to the scale of the MODIS pixels. One way of doing this is to employ both field measurements and high-resolution satellite data to produce validated fine-resolution reference LAI maps over a sufficiently extended area, to aggregate these to 1 km resolution, and to use them as benchmarks to validate the MODIS products [*Tian et al.*, 2002a, 2002b]. Disagreement between the reference and MODIS LAI can be expressed in terms of accuracy and precision. The former indicates the deviation (or bias) of the mean LAI from a known value, while the latter characterizes the ability to reproduce this value with a given accuracy. We will follow this approach in our paper.

[4] Several complicated issues arise when one attempts to implement the foregoing strategy. For example, field LAI values are usually obtained by converting downward radiation fluxes measured below the vegetation canopy. Conversion equations are based on certain assumptions which are often not met in reality. This lowers the measurement accuracy. Further, because of spatial heterogeneity the precision of field measurements, especially in the case of sparse vegetation canopies, can be quite low. Also, high-resolution satellite data provide reflective properties of vegetated surface with a certain accuracy and precision, e.g., as a consequence of imperfect atmospheric correction, calibration and geolocation errors, and other reasons. Errors in field and satellite data propagating through extrapolation procedures aggravate these tendencies and, consequently, can result in a reference LAI map of unacceptable quality. The problem then arises as to how the impact of the errors on the reference map can be minimized. This paper has two goals. First, on the basis of information on field measurement and satellite observation errors we derive a reference LAI map of the highest possible quality. Then, this map is used to achieve the second goal: validation of the MODIS LAI product.

[5] The flow of this paper is as follows. Details relevant to collection of LAI data and satellite products used in this study are given in section 2. The aim of section 3 is to derive a set of reference LAI values which account for measurement and observation errors and agree closely with both collected LAIs and those derived from Enhanced Thematic Mapper (ETM+) data. The ability of various retrieval techniques to reproduce the reference values is analyzed in section 4. In section 5 we check if the scale-dependent parameters in the Collection 4 operational LAI and FPAR algorithm are valid. The generation of 30 m resolution maps and their analyses are presented in section 6. Results validating the MODIS LAI product (section 7) and concluding remarks (section 8) complete the paper. Definitions of the accuracy, precision, and uncertainty used in this study are given in Appendix A.

2. Data Used in This Study

2.1. In Situ Data

[6] A 3 × 3 km agricultural area in Alpilles ~25 km southwest of Avignon, France (43.81°N, 4.75°E), was chosen for field data collection to validate the MODIS LAI product. This area is one of the Validation of Land European Remote Sensing Instruments (VALERI) experiment sites (available at http://www.avignon.inra.fr/valeri) (F. Baret et al., VALERI: A network of sites and a methodology for the validation of land satellite products, submitted to *Remote Sensing of the Environment*, 2004). The site is composed of young and fully grown wheat (42.1%), wheat with senescent regrowth (10.6%), bare and partially bare with weeds (30%), orchards (4.1%), grasslands (5.3%), marsh field (5%), and fallow fields (2.9%).

[7] Leaf area index was measured with a LAI-2000 plant canopy analyzer at this site during the period from 26 February to 15 March 2001. The measurements were taken mostly right before sunrise and after sunset. LAI values were calculated according to Miller's derivation [*Miller*, 1967], which is the default method used by LAI-2000. It should be noted that the LAI-2000 converts canopy gap fraction into LAI under the assumption of random spatial distribution of leaves. However, actual foliage distribution is not random which invalidates this assumption. Therefore values from Miller's formulae give an effective leaf area index. The effective LAI underestimates true values for this site [*Jonckheere et al.*, 2004; *Weiss et al.*, 2004].

[8] The 3 × 3 km plot was divided into nine 1 × 1 km grids (Figure 1a). In the middle grid, data were collected on two transects of ~1 km at 50 m intervals from west to east and from southwest to northeast. There were 34 valid readings collected on these transects. In addition, 15 points outside the centrally located grid were selected to represent the distribution of young and fully grown wheat in the whole 3 × 3 km plot. At each sampling point, data were

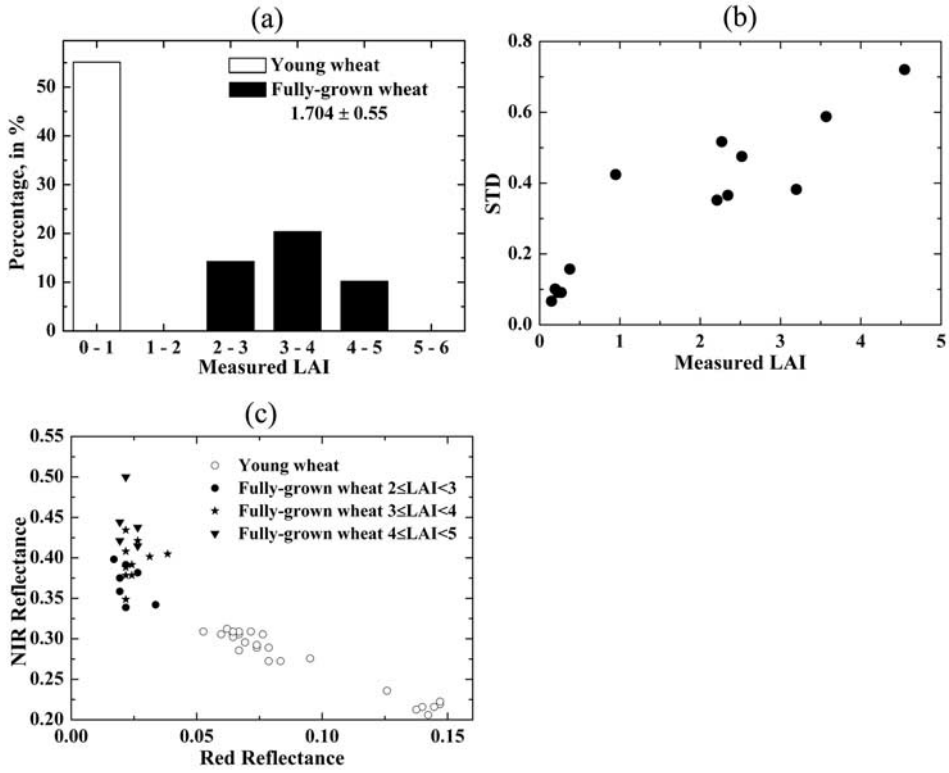

Figure 2. (a) Distribution of 49 values of measured effective LAI. (b) Standard deviation (STD) as a function of the effective LAI for the 15 off-transect points. (c) Variation in atmospherically corrected ETM+ reflectance on the red-near-infrared (NIR) plane.

collected at 4 m intervals on two 20 m lines which form a regularly shaped cross (Figure 1b) for a total of 12 local points. One above-canopy and four below-canopy LAI-2000 measurements were taken at each local point to obtain one local LAI value. The average of 12 local LAI values was assigned to LAI at the sampling point. GPS locations of the center of each cross were made. The location accuracy without differential correction of GPS is about 10–15 m since the deliberate selective availability error was removed from the U.S. GPS system from 2 May 2000 onward. Figure 2a presents histograms of LAI values for fully grown and young wheat canopies. Mean LAI values for fully grown and young wheat canopies are 0.38 (standard deviation is 0.230) and 3.33 (standard deviation is 0.711), respectively. Figure 2b shows values of standard deviation versus LAI for the 15 off-transect sampling points. These values were taken as the precision of field-measured LAI values.

2.2. Satellite Data

2.2.1. Collection 4 MODIS LAI Product

[9] The MODIS LAI and FPAR products are produced at 1 km spatial resolution daily (MOD15A1) and composited over an 8 day period on the basis of the maximum FPAR value. The 8 day product (MOD15A2) is distributed from the EDC DAAC. MODIS product versions are called collections. Collection 1 of LAI/FPAR product runs from February 2000 to February 2001. Collection 2 represents an internal science test of a limited set of data and is not available to the public. Collections 3 and 4 run from November 2000 to December 2002 and from February 2000 to the present time, respectively.

[10] The Collection 4 product is projected on the sinusoidal 10° grid, where the globe is tiled for production and distribution purposes into 36 tiles along the east-west axis and 18 tiles along the north-south axis, each approximately 1200 × 1200 km. The tile coordinate system starts at (0, 0) (horizontal (h) tile number, vertical (v) tile number) in the upper left corner and proceeds rightward (horizontal) and downward (vertical). For each 1 km pixel the product file provides LAI, FPAR, and two quality control variables. A description of data format has been detailed in the FPAR LAI user's guide (available at http://cybele.bu.edu/modis-misr/products/modis/userguide.pdf). The Collection 4 MODIS LAI product from tile h18v04 with the Alpilles site is used in our investigation.

2.2.2. Collection 3 MODIS Land Cover Product

[11] In the MODIS LAI/FPAR algorithm, global vegetation is stratified into six canopy architectural types, or biomes [*Myneni et al.*, 2002]. The six biomes are 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). Collection 3 MODIS land cover product (MOD12Q1) includes a special layer containing the six-biome land cover which is consistent with the LAI/FPAR algorithm and is used as input to the Collection 4 LAI/FPAR operational algorithm. Collection 3 MODIS land cover product is based on 1 year of MODIS observations [*Friedl et al.*, 2002]. Tile h18v04 is used in our investigations.

2.2.3. ETM+ Data

[12] We selected a subset of a Landsat ETM+ image from 15 March 2001 (path 196, row 90) containing the Alpilles site. The ETM+ subset has a spatial resolution of 30 m and covers a 10 × 10 km region. In this study, we use data from band 1 (blue, 450–515 nm), band 2 (green, 525–605 nm), band 3 (red, 630–690 nm), band 4 (near-infrared (NIR), 780–900 nm), band 5 (1550–1750 nm), and band 7 (2090–2350 nm) to generate 30 m resolution LAI maps using different retrieval techniques. The subset is in the universal transverse Mercator (UTM) projection. The image was atmospherically corrected using the 6S radiative transfer code [*Vermote et al.*, 1997]. Aerosol optical depth at 550 nm, input to the 6S code, was obtained by exponential interpolation of data from the Aerosol Robotic Network [*Holben et al.*, 1998]. Figure 2c shows reflectances of pixels where LAI measurements were taken. The fully grown and young wheat canopies are separated sufficiently well in the spectral space. A denser canopy (fully grown wheat) exhibits high NIR and low red reflectance. For the young wheat the opposite effect is seen: high red and low NIR reflectance. An increase in LAI decreases red and increases near-infrared reflectance. Such behavior is seen in canopies with background of intermediate brightness [*Shabanov et al.*, 2002]. On average, the reflectance of fully grown wheat is ∼0.03 and ∼0.40 in the red and NIR bands, respectively. For the young wheat the ETM+ reflectance is 0.1 and 0.25 in the red and NIR spectral bands, respectively.

3. Reference Values

[13] The biggest challenge for validation of moderate- (100–1000 m) and coarse- (>1 km) resolution LAI products is the extrapolation of scarce field data from sampling points to a sufficiently extended area. One way of doing this is to employ both field measurements and high-resolution satellite data (10–30 m) to produce validated fine-resolution LAI maps over a sufficiently extended area, to aggregate these to 1 km resolution, and to use them as benchmarks to validate the MODIS products [*Tian et al.*, 2002a, 2002b]. An important step in the implementation of this approach is the derivation of an accurate site-specific relationship between LAI and surface reflectance using field measurements and satellite data which, in turn, can be taken as a reference. Various retrieval techniques can be tested for their ability to reproduce the relationship and, consequently, for their ability to extrapolate field data.

[14] We have 49 pairs $[L_k, E_k]$ of field-measured LAI, L_k, and observed ETM+ spectral surface reflectance, E_k, of pixels where corresponding LAI measurements were taken (Figure 3). In the case of error-free data, observed relationships $[L_k, E_k]$ can be taken as reference values. A small perturbation, however, is likely to result in a change in a true relationship. Errors in atmospherically corrected ETM+ data and field-measured LAIs are sources of the perturbation. Here we focus on the following problem: Given 49 pairs $[L_k, E_k]$ and their errors, how should values of L_k and E_k be changed within field measurement and satellite observation errors to achieve the most accurate relationship between LAI and ETM+ surface reflectance? The modified pairs will be taken as reference values. They should provide stable relationships in both forward and inverse modes. In the forward mode, i.e., prediction of surface reflectance given LAI, variation in LAI due to field measurement errors should involve corresponding variation in surface reflectance within the satellite observation errors. For the inverse mode, i.e., prediction of LAI given surface reflectance, variation in surface reflectance data due to satellite observation errors should not cause variation in corresponding LAI to fall outside of the measurement errors. The stable relationship converts observed reflectance (measured LAI) to LAI (reflectance) as follows: It replaces the observed reflectance (measured LAI) with the closest reference value and returns the corresponding reference LAI (reflectance). The predicted value will be accurate to within measurement and observation errors. Our goal in this section is to derive 49 pairs of LAIs and spectral surface reflectance satisfying the stability conditions.

[15] Figure 3 illustrates our approach. To examine stability of the forward mode (case A), one selects a LAI value of the kth pixel. Because of measurement errors a true LAI for this pixel belongs to an interval around the selected value. Next, one forms a set of measured LAIs whose values fall in this interval and evaluates coefficients of variation (standard deviation/mean) of LAI and corresponding ETM+ reflectance in the red and NIR spectral bands. By repeating this procedure for each pixel one obtains relationships between variations in LAI and corresponding ETM+ reflectance. We will show that the forward mode satisfies the stability condition. Stability of the inverse mode (case B) is examined in a similar manner (Figure 3). This mode, however, appears to be very sensitive to small variations in surface reflectance, indicating that the prediction of LAI given surface reflectance is an ill-posed problem [*Tikhonov et al.*, 1995]. Finally, by accounting for both measurements and observation errors (case C) we derive 49 pairs of LAIs and spectral surface reflectance satisfying the stability conditions.

3.1. Case A: Stability of the Forward Mode

[16] To assess the impact of measurement errors on the prediction of surface reflectance, we form 49 groups of data, with varying surface reflectance and almost constant values of field LAI, using 49 pairs $[L_k, E_k]$, $k = 1, 2, \ldots, 49$, of field LAIs, L_k, and observed ETM+ reflectance, $E_k = (e_3^k, e_4^k)$, in the red, e_3^k, and NIR, e_4^k, spectral bands as follows (Figure 3). The kth group identified by L_k includes those ETM+ observations for which corresponding field LAIs differ from L_k by an amount equal to or less than the measurement error; that is, $|L_k - L_j| \leq \sigma(L_k)$. The standard deviation, $\sigma(L)$, is approximated by a step function; that is, $\sigma(L) = 0.14$ if LAI < 1, and $\sigma(L) = 0.53$ otherwise. Here 0.14 and 0.53 are averages of standard deviation values shown in Figure 2b over LAI < 1 and LAI ≥ 1, respectively. The correlation (not shown here) between L_k, $k = 1, 2, \ldots, 49$, and mean values over associated groups (i.e., over L_j for which $|L_k - L_j| \leq \sigma(L_k)$) is very strong; $R^2 = 1$, slope is 0.97, and offset is 0.01. If measured values L_k are treated as true values and elements from the associated groups are treated as their estimates, the proximity between them is characterized by a relative accuracy, uncertainty, and average relative precision of 2%, 6%, and 13%, respectively (see Appendix A for definitions of the accuracy, precision, and uncertainty).

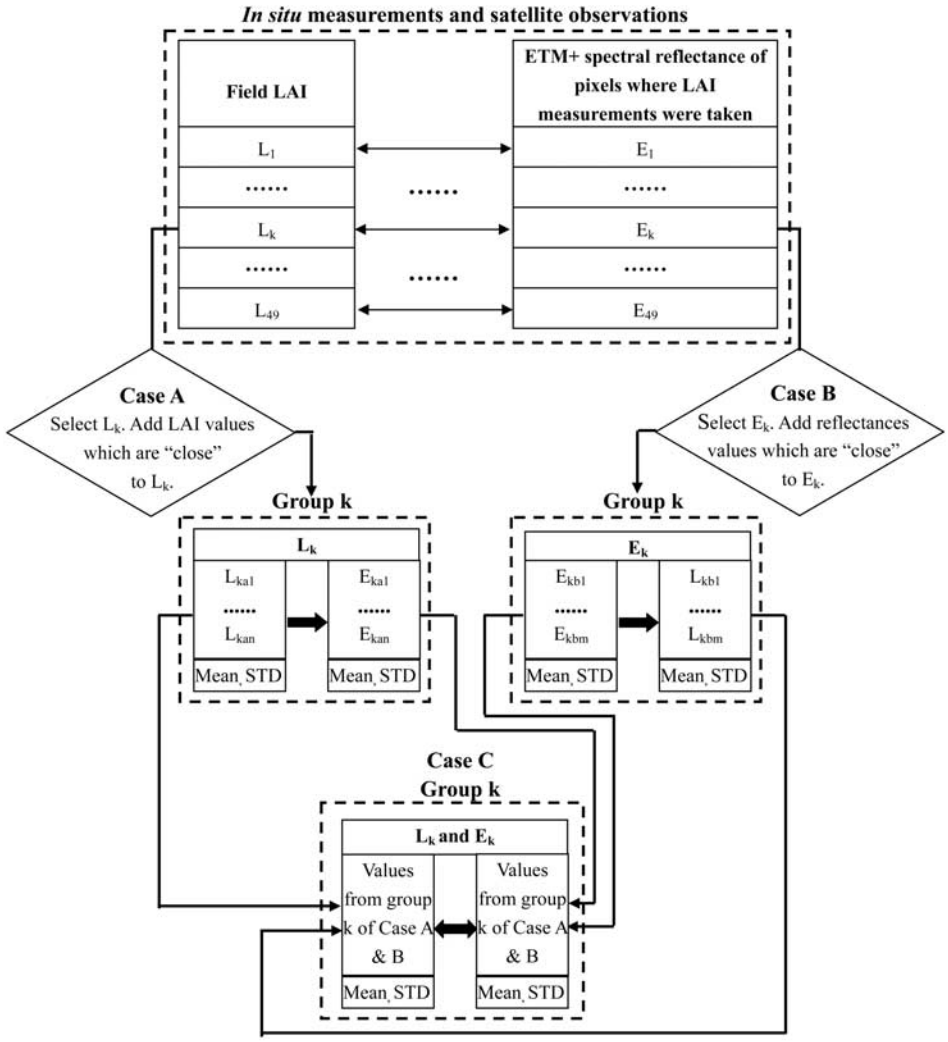

Figure 3. Derivation of reference values. Group k in case A consists of pairs of measured LAI values L_{kal} which are "close" to L_k and corresponding ETM+ surface reflectances E_{kal}. Similarly, group k in case B includes ETM+ surface reflectances E_{kbl} being "close" to E_k and respective measured LAI values L_{kbl}. Group k in case C is obtained by merging groups k in cases A and B.

Thus, to every L_k, there corresponds a set of pairs $[L_j, E_j]$ with varying surface reflectance and "almost" constant value of the leaf area index.

[17] The correlation between observed, E_k, and mean reflectance taken over the corresponding group is shown in Figure 4a. The observed and mean reflectances in red and NIR spectral bands are correlated sufficiently well (Figure 4a) and agree to within a relative accuracy of 2% for red and 1% for NIR spectral bands (Table 1). This suggests that the accuracy in the prediction of surface reflectance given LAI is comparable with the accuracy in LAI. The relative uncertainties (RMS/mean) in the mean red and NIR reflectances with respect to observed values are 27% and 7%, respectively. As one can see from Figure 4b, reflectances from groups can reproduce observed ETM+ reflectance with a relative precision (standard deviation/mean) of 25% at red and 8% at NIR wavelengths. An overall relative precision, (25% + 8%)/2% $\sim$ 16%, is comparable to the relative precision of field measurements (13%). Thus variation in LAI values due to measurement errors results in comparable variation in corresponding surface reflectance.

3.2. Case B: Instability of the Inverse Mode

[18] Let $M = (m_3, m_4)$ be a vector of true values, m_3 and m_4, of the atmospherically corrected ETM+ reflectance in the red (band 3) and NIR (band 4) spectral bands. In-orbit radiances measured by satellite-borne sensors are transformed through an atmospheric correction algorithm into an observation $E = (e_3, e_4)$ of M. Various error sources and factors which are not accounted for by the algorithm cause the observed surface reflectances e_3 and e_4 to deviate from their true values m_3 and m_4. We treat the observations as independent random variables with finite variances σ_k^2, k = 3, 4, and assume that the deviation $\varepsilon_k = (e_k - m_k)/\sigma_k$, k = 3, 4, follows Gaussian distribution. The random variable

$$\chi_\sigma^2[E-M] = \varepsilon_3^2 + \varepsilon_4^2 = \frac{(e_3 - m_3)^2}{\sigma_3^2} + \frac{(e_4 - m_4)^2}{\sigma_4^2} \qquad (1)$$

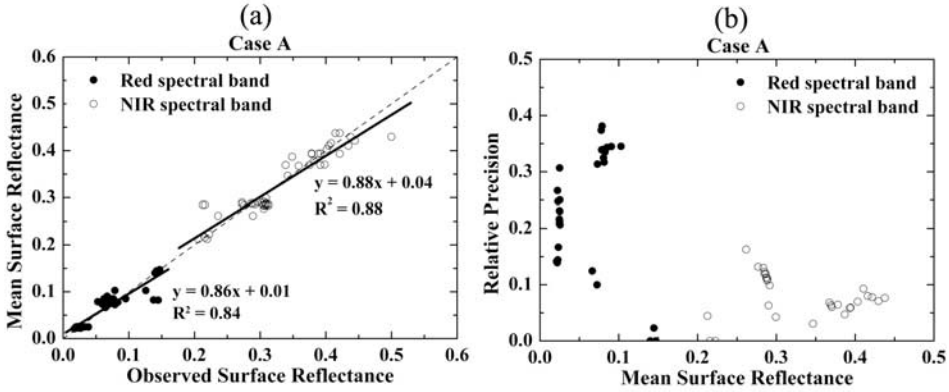

Figure 4. (a) Correlation between observed surface reflectances and mean values over associated groups for red (solid circles) and NIR (open circles) wavelength. (b) Relative precision of the mean surface reflectance in the red (solid circles) and NIR (open circles) spectral bands.

characterizing the proximity of atmospherically corrected data E to true values M has a chi-square distribution with two degrees of freedom. A value of $\chi^2_\sigma \leq 2$ indicates a good observation quality [*Martonchik et al.*, 1998; *Wang et al.*, 2001]. Dispersions $\sigma = (\sigma_3, \sigma_4)$ are precisions in the atmospherically corrected ETM+ surface reflectance in the red and NIR spectral bands, respectively, which, in turn, are parameterized in terms of their relative values, i.e., $\alpha_3 = \sigma_3/e_3$ and $\alpha_4 = \sigma_4/e_4$. Note that the relative precision is used by the MODIS LAI algorithm as input. Since only one observation of ETM+ surface reflectance per pixel was available for this site, we use theoretical estimates to specify the relative precisions, namely, $\alpha_3 = 15\%$ and $\alpha_4 = 6\%$ [*Wang et al.*, 2001; D. Huang et al., Evaluation of Collection 3 MODIS LAI products with respect to input data uncertainties—Case study for grasses, manuscript in preparation, 2004, hereinafter referred to as Huang et al., manuscript in preparation, 2004; E. Vermote and A. Vermeulen, Atmospheric correction algorithm: Spectral reflectances (MOD09), algorithm technical background document, version 4.0, 1999, available at http://eospso.gsfc.nasa.gov/ftp_atbd/review/modis/atbd-mod-08/atbd-mod-08.pdf]. Two observations, E_1 and E_2, are said to be indistinguishable if $\chi^2_\sigma [E_1 - E_2] \leq 2$. In other words, indistinguishable observations are equal within the observation precision.

[19] We form another 49 groups of data using 49 $[L_k, E_k]$, $k = 1, 2, \ldots, 49$, with varying LAIs and almost constant surface reflectances (Figure 3). The kth group identified by E_k includes those measured L_j for which corresponding observed ETM+ reflectance E_j is equal to E_k within the observation precision, i.e., $\chi^2_\sigma [E_k - E_j] \leq 2$. The correlation between E_k, $k = 1, 2, \ldots, 49$, and mean values over corresponding indistinguishable observations (i.e., over E_j for which $\chi^2_\sigma [E_k - E_j] \leq 2$) is very high (Figure 5); they almost lie on the 1:1 line with negligible offsets. The R^2 for red and NIR spectral bands is ~0.99. The relative uncertainty (RMS/mean) is 5% at red and 2% at NIR wavelengths with biases being <0.001. The relative precision (standard deviation/mean) varies with groups between 0% (only one pair in the group) and 15% (mean is 7%) for red and does not exceed 7% (mean is 3%) for NIR spectral band. Thus, if E_k are treated as true values and the indistinguishable observations are treated as their estimates, the proximity between them is characterized by a relative accuracy, uncertainty, and average relative precision of 1%, 5%, and 7% in red and 1%, 2%, and 3% in NIR spectral band, respectively. Thus, to every E_k, there corresponds a set of pairs $[L_j, E_j]$ with varying LAIs and almost constant values of ETM+ reflectance.

[20] Figure 6 demonstrates the response of LAI to indistinguishable variations in surface reflectance. Measured L_k and mean values over corresponding groups are well correlated ($R^2 = 0.95$) and close to the 1:1 line (slope is 0.94; offset is 0.09); the relative uncertainty (RMS/mean) is 20%. If measured L_k are treated as true values and elements from associated groups are treated as their estimates, they agree to within a relative accuracy (bias/mean) of 1% (Figure 6a and Table 1). The relative precision (standard deviation/mean) varies significantly with LAI and can be as high as 130% (Figure 6b). Its mean value is 40%.

[21] For LAI < 1, mean and measured values are poorly correlated; that is, field-measured LAIs can vary significantly with mean values essentially unchanged (Figure 6a). The relative uncertainty and mean relative precision calculated for these LAIs are 46% and 65%, respectively. For

Table 1. Accuracies, Uncertainties, and Precisions of Relationships Between the Effective LAI and ETM+ Surface Reflectance[a]

	Case A			Case B			Case C		
		ETM+ Reflectance		ETM+ Reflectance				ETM+ Reflectance	
	LAI, %	Red, %	NIR, %	Red, %	NIR, %	LAI, %	LAI, %	Red, %	NIR, %
Accuracy	2	2	1	1	1	1	1	1	1
Uncertainty	6	27	7	5	2	20	10	17	5
Precision	13	25	8	7	3	65	33	20	6

[a]ETM, Enhanced Thematic Mapper; LAI, leaf area index; NIR, near-infrared.

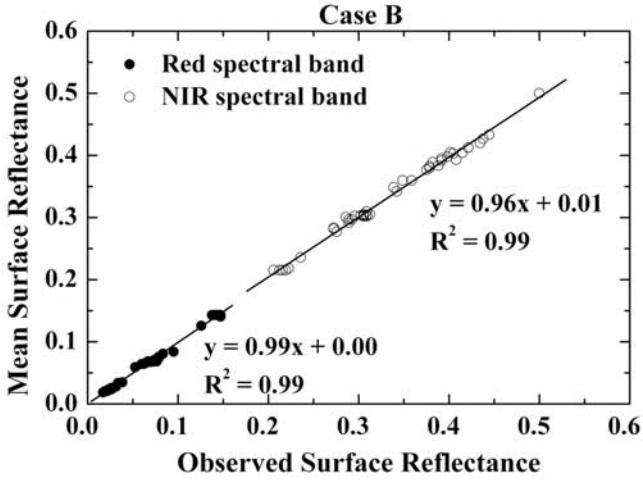

Figure 5. Correlation between observed ETM+ surface reflectance and its mean over ''indistinguishable'' observations for red and NIR spectral bands.

LAI $\geq$ 1 the correlation is stronger (R^2 = 0.60), less uncertain (RMS/mean is 14%), and more precise (standard deviation/mean is 15%). Measured L_k and mean over corresponding groups agree to within a relative accuracy of 1% in both cases.

[22] Thus the inverse mode magnifies the uncertainty and precision in the red surface reflectance by a (sensitivity) factor of about 4 and 9 (Table 1), respectively, making the prediction highly sensitive to input data errors and, consequently, unstable. This result suggests that the retrieval of LAI from satellite data should be treated as an ill-posed problem. The sensitivity factors can be reduced if both measurement and observation errors are used by the retrieval technique [*Tikhonov et al.*, 1995] (see also section 3.3). The MODIS LAI/FPAR algorithm uses the radiative transfer equation and observation precision to establish relationships between LAIs and surface reflectances. This made it possible to reduce the precision sensitivity factor to $\sim$1.2 (Huang et al., manuscript in preparation, 2004). It should be emphasized that the use of error information in the retrieval technique is a must to achieve this level of stability [*Knyazikhin et al.*, 1998a, 1998b]. This is consistent with common knowledge that the estimation of the LAI given reflectances is an ill-posed problem, while prediction of the radiation field given LAIs is a well-posed problem [*Vladimirov*, 1963].

3.3. Case C: Stable Relationship

[23] To assess the overall impact of measurement and observation errors on the relationship between satellite data and field LAI, we form 49 sets of data by joining respective groups introduced in sections 3.1 and 3.2 and in Figure 3. Thus the kth group identified by measured LAI, L_k, and ETM+ reflectance, E_k, is the set of pairs that belongs to at least one of the kth group from cases A and B (Figure 3). For each group, mean values of the ETM+ surface reflectance and field LAI were calculated. The former are taken as reference reflectances which account for both measurement and observation errors. Figure 7a shows correlation between observed and reference reflectances. The relative uncertainties are 17% and 5% at red and NIR spectral bands, respectively (Table 1). They agree to within a relative accuracy of 1%. The reference surface reflectance represents the corresponding observed value with a relative precision of 20% for red and 6% for NIR spectral bands. Both the relative uncertainty and precision, on average, are lower than those obtained in case A.

[24] Mean and field LAI almost lie on the 1:1 line (R^2 = 0.99, slope is 0.96, and offset is 0.07) and agree to within an accuracy of 1% (Figure 7b and Table 1). The relative uncertainty is 28% for LAI < 1, 7% for LAI $\geq$ 1, and 10% for the entire LAI range. The relative precision varies with LAI between 15% and 160% (mean is 51%) for LAI < 1 and between 0% and 22% (mean is 11%) for LAI $\geq$ 1. Its mean value over the entire LAI range is 33%. If both measurement and observation errors are taken into account, the correspondence between LAI and ETM+ reflectance is more precise and less uncertain compared to those obtained in case B. This improvement has resulted from a redistribution of uncertainty and precision values attributed to LAI in case B between LAI and surface reflectance in case C. For example, the 65% precision of LAI in case B has been partitioned between the precision in LAI (33%) and red reflectance (20%) in case C (Table 1). The use of uncertainty information is primarily responsible for a better correspondence between measurements and observations.

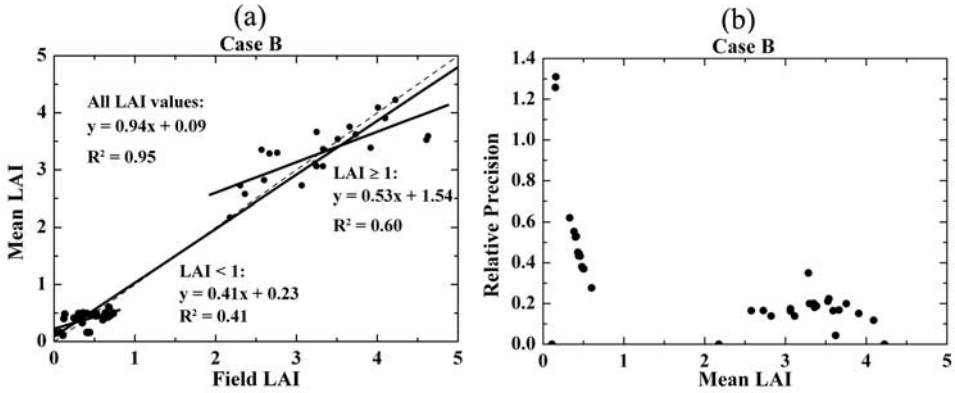

Figure 6. (a) Correlation between field-measured LAIs and their mean values over corresponding groups. (b) Relative precision as a function of the mean LAI.

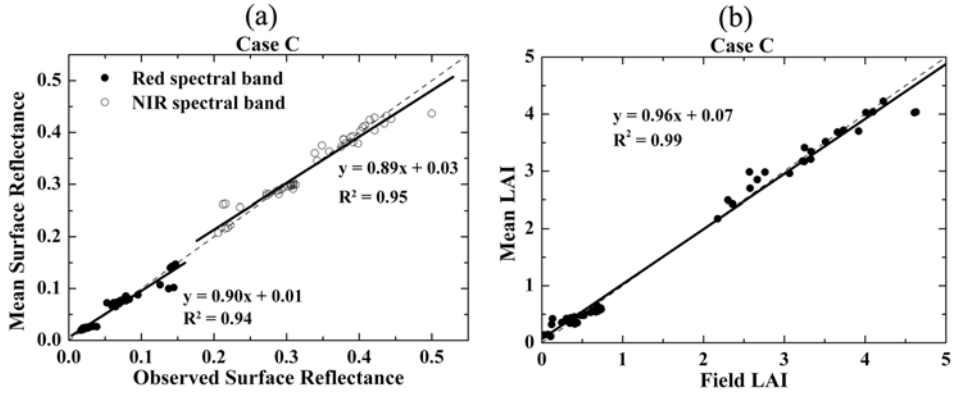

Figure 7. (a) Correlation between observed ETM+ surface reflectances and their mean values over corresponding groups. (b) Correlation between field-measured LAIs and their mean values over corresponding groups.

[25] For each group of ETM+ surface reflectances a mean simple ratio (SR) was calculated. LAI and mean SR appeared to be linearly related,

$$\mathrm{LAI} = 0.22\mathrm{SR} - 0.33. \quad (2)$$

The R^2 is 0.94 with a RMS of 0.36. LAI values obtained by applying this relationship to the 49 reference reflectances are taken as reference values for satellite-derived LAI. Figure 8 demonstrates the proximity of measured and reference LAI values. There is a very strong correspondence between them. The relative uncertainty is 26%. If the measured LAIs are taken as true values, the reference LAIs estimate their mean values over LAI < 1 and LAI ≥ 1 with relative accuracies of 9% and 1%, respectively.

4. Empirically Based Retrieval Techniques and MODIS LAI and FPAR Algorithm

[26] The aim of this section is to derive relationships between LAI and ETM surface reflectance using various techniques and to test for their ability to predict the reference LAI values. Relationships which provide the best agreement will be used to extrapolate field data.

4.1. Simple Ratio–Based Regression Method

[27] Normalized difference vegetation index (NDVI), SR, greenness, and brightness are often used for LAI retrievals [*Jordan*, 1969; *Tucker*, 1979; *Myneni et al.*, 1995; *Knyazikhin et al.*, 1998a; *Elvidge and Chen*, 1995; *Butera*, 1986; *Baret et al.*, 1988; *Nemani et al.*, 1993; *Brown et al.*, 2000; *Wang et al.*, 2004; *Crist and Cicone*, 1984; *Kauth and Thomas*, 1976; *Jackson*, 1983]. The leave-one-out cross-validation technique [*Cawley and Talbot*, 2003; *Gong*, 1986; *Efron and Gong*, 1983] was used to test these indices for their ability to produce a fine-resolution LAI map. An index which minimizes the root mean prediction error sum of squares while keeping the bias within the standard error is taken as the best candidate [*Katila and Tomppo*, 2001]. This technique provides an almost unbiased estimate of the extrapolation ability of the model performance on previously unseen data. We found that only the SR-based relationship satisfies the above criteria in our particular case. This relationship is given by

$$\mathrm{LAI} = 0.197\mathrm{SR} - 0.216. \quad (3)$$

Note that derivation of this equation is based on the ordinary least squares regression method which assumes error-free surface reflectance [*Curran and Hay*, 1986]. Errors in the SR lead to a downward biased estimate of the true regression slope. This effect falls off with increasing R^2 [*Hausman*, 2001]. In our case the R^2 of equation (2) (0.94) is higher than that of equation (3) (0.85). This suggests that equation (2) provides a closer approximation to the relationship between LAI and ETM+ surface reflectance.

[28] Equation (3) takes on negative values for SR < 1.1 (NDVI < 0.05), indicating very low vegetation density. Under the SR-based empirical model the inequality SR < 1.1 is taken as an indicator of nonvegetated pixels. Reflectance of these pixels is linearly related, i.e., NIR = red + 0.02 ($R^2 = 0.98$; RMS error is 0.013). This linear regression describes the distribution of nonvegetated pixel on the red-

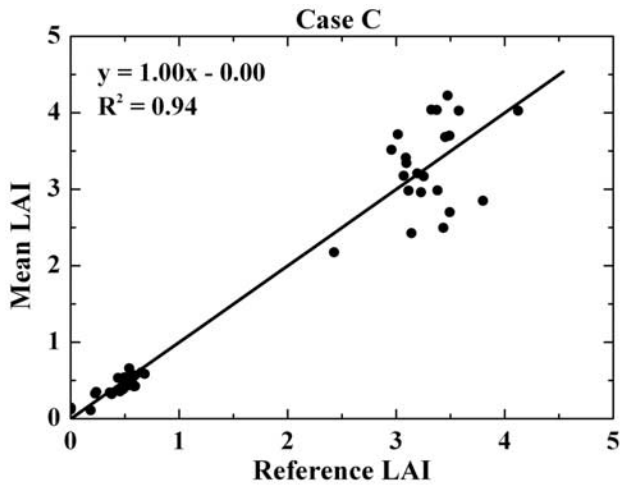

Figure 8. Correlation between field-measured and reference LAI values.

NIR plane, or soil line [*Kauth and Thomas*, 1976]. This regression curve will be used in section 7 to adjust the MODIS algorithm to generate a 30 m fine-resolution LAI map.

4.2. Neural Network Retrieval Technique

[29] We use the neural network technique to derive a mapping function between ETM reflectance at six spectral bands (six input variables) and LAI (one output variable) for given illumination and viewing geometry and site-specific architecture of the vegetation canopy. The back propagation neural network [*Rumelhart et al.*, 1986] available in the MATLAB (version 6) software (The MathWorks, Inc., available at http://www.mathworks.com) was used to derive the map. Two hidden layers with four and two nodes were selected to specify the architecture of the neural network. Two log-sigmoid and linear functions were used to establish the connection between input, two hidden, and output layers. Six parallel networks were run simultaneously. The median value of the six predictions was taken as a final result.

4.3. Fine-Resolution MODIS LAI/FPAR Algorithm

[30] The MODIS LAI/FPAR algorithm uses a biome classification map and atmospherically corrected spectral reflectance as well as model and input precisions to retrieve LAI. It compares measured reflectances with those determined from a suite of canopy models, which depend on biome type, canopy structure, and soil/understory reflectances. The canopy/soil/understory models with which simulated and measured surface reflectances agree within a model and observation precision are used to derive the distribution of all possible solutions, i.e., LAI distribution functions. The mean values of these distribution functions are archived. Thus the operational algorithm was designed to process data in a fashion similar to case C. The three-dimensional transport equation is used to simulate canopy reflectances as a function of biome type, Sun view geometry, and canopy/soil patterns [*Myneni*, 1991]. The precision sensitivity factor is $\sim$1.2 (Huang et al., manuscript in preparation, 2004).

[31] The MODIS LAI/FPAR algorithm adjusted to process ETM+ data [*Tian et al.*, 2002a, 2002b, 2002c] will be used here to generate a 30 m resolution LAI map. There are two parameters in the MODIS LAI/FPAR algorithm which allow accounting for site-specific features. They are the spectral single-scattering albedo [*Tian et al.*, 2002c; *Wang et al.*, 2004] and background reflectance. The leaf spectrum was not available for this site, and thus we use a mean leaf spectrum for grasses (MODIS semiannual report, 1 January 1998 to 30 June 1998). The soil line derived from ETM+ data (section 4.1) was used to specify patterns of site-specific background reflectances. Precisions in atmospherically corrected ETM+ reflectance in the red and NIR spectral bands were set to 20% and 10%, respectively.

4.4. Comparison of Retrieval Techniques

[32] The simple ratio relationship (equation (3)), neural network technique, and fine-resolution MODIS LAI/FPAR algorithm were applied to process 49 ETM+ reflectances of the sample points (Figure 1). Figure 9 demonstrates the proximity of these retrievals to the reference and measured LAI values. We see that the retrievals correlate with the reference values more strongly than with their field-measured counterparts. The R^2 values, slopes, and offsets for the simple ratio–based LAI (SR LAI) and neural network-based LAI (NN LAI) relative to the reference are very close (Figures 9a and 9c). The relative accuracy and uncertainty for the entire LAI range are -1% and 19% for the SR LAI and +5% and 32% for NN LAI, respectively. We exclude the neural network retrievals from further analyses since they rank below the SR-based LAI in quality.

[33] Predicted MODIS LAI and the reference are very well correlated ($R^2 = 0.9$; slope is 1) and agree within an accuracy and uncertainty of 0.3 (relative value is 16%) and 0.58 (relative value is 30%), respectively. To compare, Figure 9d demonstrates correlation between field and MODIS LAIs which appears to be less accurate.

[34] It will be recalled that the reference provides the effective LAI which, in our case, underestimates true values [*Jonckheere et al.*, 2004; *Weiss et al.*, 2004]. The MODIS algorithm, in turn, is designed to retrieve true LAI values. This explains why the MODIS algorithm overestimates the reference values in this particular case. Since the precision sensitivity factor for the MODIS algorithm is $\sim$1.2, the retrieval precision is about $1.2 \times 20\% = 24\%$, where 20% is the precision tolerance level used by the MODIS algorithm. The overall precision of the SRLAI is also 24%. It was calculated by applying equation (3) to the groups introduced in case C and evaluating the mean and standard deviation for each of them. Thus the SRLAI and MODIS LAI are comparable in terms of precision. The difference between true and effective LAI therefore can be obtained by comparing the SR and MODIS algorithm–based retrievals. As one can see from Figure 10, the true LAI, on average, differs from its effective counterpart by 0.28. The difference, $0.28 + 0.07 - 0.30 = 0.05$, between offsets in MODIS LAI–SR LAI plus SR LAI–reference and MODIS LAI–reference relationships can be treated as the accuracy in MODIS LAI and SR LAI retrievals to reproduce reference LAIs.

5. Scaling of the MODIS Algorithm

[35] As demonstrated in section 4.3, the fine-resolution version of the MODIS algorithm can generate a 30 m LAI field with satisfactory accuracy, precision, and uncertainty. However, when the spatial resolution of the imagery becomes significantly coarser than 30 m, then, in general, both the degree of biome mixing within a pixel and the number of mixed pixels in the imagery is increased. If LAI is to be retrieved using data with a pixel size that almost guarantees substantial mixing, the retrieval uncertainty can be quite high if the resolution of the data is not considered in the retrieval technique. Thus the retrieval algorithm must be scale adjustable to allow for spatial-scale effects. Scaling is a process by which it is established that LAI values derived from coarse-resolution sensor data should equal the arithmetic average of values derived independently from fine-resolution sensor data. The MODIS LAI and FPAR algorithm addresses this issue explicitly through two resolution-dependent parameters [*Knyazikhin et al.*, 1998a, 1998b; *Panferov et al.*, 2001; *Wang et al.*, 2003; *Smolander*

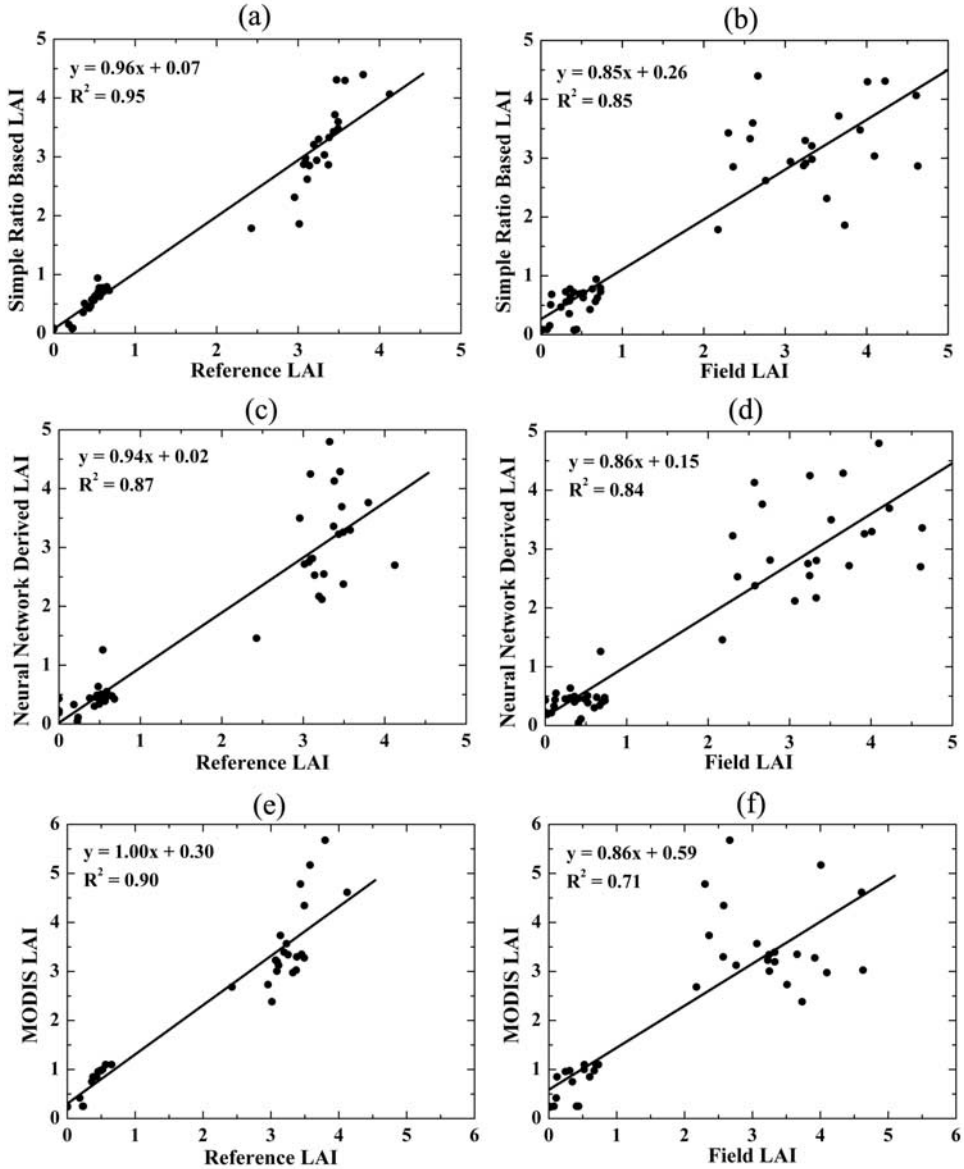

Figure 9. Comparisons of LAI values produced using (a and b) simple ratio LAI (SR LAI) relationship, (c and d) neural network retrieval technique, and (e and f) the fine-resolution Moderate Resolution Imaging Spectroradiometer (MODIS) LAI and fraction of photosynthetically active radiation (FPAR) algorithm against the (left) reference and (right) field LAIs.

and Stenberg, 2003]. A methodology proposed by *Tian et al.* [2002c] was applied to specify the scale-dependent parameters for the Collection 4 MODIS LAI and FPAR algorithm. The aim of this section is to check if the scale-dependent parameters in the Collection 4 operational LAI and FPAR algorithm are valid for the Alpilles site.

[36] Let L_T denote LAI values at a resolution of 1 km obtained by averaging 30 m retrievals produced by the fine-resolution MODIS algorithm. We aggregate the 30 m ETM+ reflectance data to 1 km resolution. Let L denote a LAI value obtained directly from 1 km resolution surface reflectance data using the Collection 4 MODIS operational algorithm. The relative difference between L_T and L, RDL = $|L_T - L|/L_T$, characterizes LAI error incurred by first aggregating reflectances to 1 km resolution and then performing LAI retrievals. The RDL varies with the spatial resolution, and pixel heterogeneity and can be quite high if the resolution of the data is not considered in the retrieval technique [*Tian et al.*, 2002c]. Allowing an overall model and input precision of 20%, the fine-resolution biome 1 look-up table was scaled up in such a way that mean $\overline{\mathrm{RDL}} <$ 20% for biome 1 (grasses and cereal crops) irrespective of within-pixel heterogeneity [see *Tian et al.*, 2002c, Figure 9]. Here $\overline{\mathrm{RDL}}$ is the mean relative difference over grasses and cereal crops (biome 1).

[37] The fine-resolution LAI/FPAR algorithm was applied to produce a 30 m LAI map for a 10 $\times$ 10 km area centered on the Alpilles site. LAI values, L_T, at 1 km resolution were obtained by averaging 30 m LAI retrievals. A 1 km pixel is attributed to biome 1 if the percent of 30 m subpixels with this biome type exceeds 85%. There were 58 pixels satisfying this condition. Figure 11 shows histo-

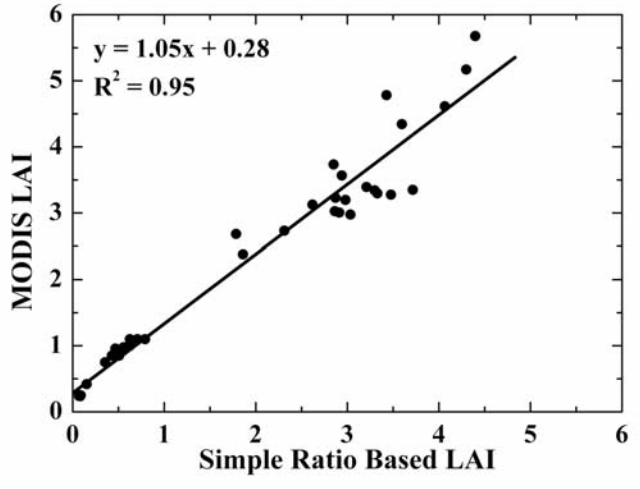

Figure 10. Correlation between LAI values derived from the simple ratio–based technique and the MODIS LAI and FPAR algorithms.

grams of L_T and L for these pixels. Both histograms localize the most probable value of LAI. However, 1 km LAIs derived from 1 km resolution surface reflectance data vary between 0.5 and 1.5, while 87% of values aggregated from 30 m retrievals fall in this interval. In spite of this discrepancy the mean $\overline{\mathrm{RDL}}$ over all 1 km biome 1 pixels does not exceed a threshold value of 20% ($\overline{\mathrm{RDL}}$ = 12%). Mean LAIs obtained by averaging 30 m retrievals and derived directly from 1 km resolution data agree to within an accuracy of 0.15 with a precision and uncertainty of 0.21 and 0.20, respectively. This result suggests that scaling parameters in the operational algorithm are valid for the Alpilles site.

6. Comparison of SR and MODIS Fine-Resolution Maps

[38] The fine-resolution LAI/FPAR algorithm and the SR relationship were applied to produce 30 m LAI maps of a 10 × 10 km area centered on the Alpilles site. The correlation between 30 m MODIS LAI value and 30 m SR LAI is shown in Figure 12a. One can see that the SR LAI can vary significantly with the MODIS LAI essentially unchanged. Because of a high value of the precision sensitivity factor of the empirical relationship (section 3.2), variations within an observation precision typically of about 13–20% (Huang et al., manuscript in preparation, 2004) make the retrievals highly unstable. The MODIS algorithm, in turn, is insensitive to such variations in input since the precision sensitivity factor is ~1.2. Inaccuracies in input therefore do not have such a dramatic impact on the stability of MODIS retrievals. This situation is detailed in case B when LAI values predicted by the empirical relationship exhibit high variations due to observation errors. To account for these factors, the empirically based retrievals should be averaged over LAIs corresponding to "indistinguishable" surface reflectances (section 3.2). Since MODIS algorithm retrievals are stable with respect to the observation precision, the expected values can be obtained by averaging SR LAIs corresponding to a given value of the MODIS LAI. As one can see from Figure 12b, the relationship between averaged SR LAI values and MODIS retrievals coincides with the one derived from the reference (Figure 10), indicating that both MODIS LAIs and averaged SR LAI agree with the reference LAIs. This finding suggests that the empirically based retrievals are of low quality and that they should undergo an additional processing to obtain the reliable product.

[39] The analyses presented here and in section 3.2 indicate that the retrieval of LAI from satellite data is an ill-posed problem; that is, small variations in input due to observation errors result in a very low precision of the desired parameter. Any retrieval technique based on a simple model inversion or empirical relationships is unable to generate stable retrievals [*Tikhonov et al.*, 1995]. The pair, observations and their inaccuracies [*Tikhonov et al.*, 1995, p. 3], "is, in general, the minimal information necessary to construct approximate solution for ill-posed problems." Examples presented here and in section 3 give a good illustration of this fundamental mathematical result. Note that the operational MODIS algorithm admits the minimal information pair and therefore is able to generate stable retrievals (Huang et al., manuscript in preparation, 2004).

7. MODIS Product

[40] To generate a 1 km reference LAI map, the fine-resolution LAI/FPAR algorithm (section 4.3) and the SR relationship were applied to produce 30 m LAI maps of a 20 × 20 km area around the Alpilles site in the UTM WGS84 projection. Pixels with missing MODIS retrievals because of algorithm failure were filled with the SR-based retrievals converted to the MODIS LAI by adding an offset of 0.28 (Figure 10). Thus the 30 m MODIS LAI field includes both MODIS retrievals and adjusted SR LAIs. The reprojection tool provided with the ENVI (version 3.5) software (Research Systems, Inc., available at http://www.rsinc.com/envi) was used to reproject fine-resolution

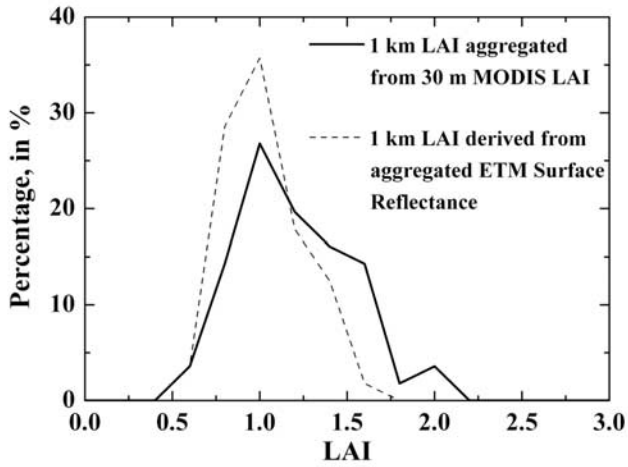

Figure 11. Plot of LAI values at 1 km resolution obtained by (1) averaging 30 m resolution LAI retrievals (solid line) and (2) applying the MODIS operational algorithm to 1 km surface reflectance aggregated from 30 m ETM+ data (dashed line).

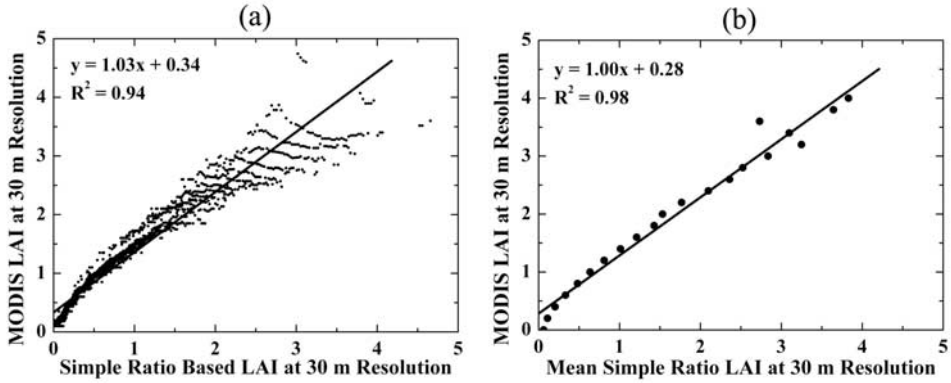

Figure 12. (a) Correlation between LAI values produced by the simple ratio–based technique (SR LAI) and the 30 m resolution MODIS LAI and FPAR algorithm for a 10 × 10 km area centered on the remote sensing data assimilation experiment site. (b) Correlation between the mean SR LAI and MODIS retrievals. The mean values are obtained by averaging SR LAIs corresponding to a given value of the MODIS LAI.

LAI maps into the sinusoidal 10° grid. A 10 × 10 km area centered on the Alpilles site and overlapped with the MODIS standard product was extracted from the reprojected image. Information about coregistration errors was not available to us. Two LAI maps at 1 km resolution were obtained by averaging 30 m SR LAI and 30 m MODIS LAI amplified with the adjusted SR LAIs. The 1 km pixels are attributed to biome 1 if the percent of 30 m subpixels with grasses exceeds 85%. There were 58 pixels satisfying this condition which were selected for our further analysis.

[41] Figure 13 shows the relationship between 1 km LAI fields aggregated from the 30 m SR and MODIS LAIs. Since the 1 km pixels are quite homogeneous with respect to the biome type, the relationship is expected to be close to that derived directly from 30 m LAI fields using only those pixels for which 30 m MODIS retrievals are available. Such a relationship is shown in Figure 12b, which, as expected, almost coincides with one shown in Figure 13. This indicates that the substitution of the missing 30 m MODIS retrievals with the adjusted SR LAIs has not involved the loss in the quality of the 1 km LAI map aggregated from the 30 m MODIS LAI retrievals. This map is taken as a reference 1 km LAI map. Figure 14 shows a fine-resolution LAI map and a contour plot of 1 km resolution LAIs.

[42] The Collection 4 MOD15A2 product for 14 March 2001 is used in our analysis. The Alpilles site is located in tile h18v04, line 744, and sample 411. A 10 × 10 km area centered on the Alpilles site which overlays the reference 1 km LAI map was extracted from the Collection 4 MOD15A2. Fifty-eight pixels specified earlier were selected for further analysis. Note that more than 85% of the selected area is occupied by biome 1 (grasses), while the Collection 3 MODIS land cover product maps it as biome 3 (broadleaf crops). The main algorithm either fails or produces incorrect LAI values in such cases [*Myneni et al.*, 2002; *Tian et al.*, 2000; Huang et al., manuscript in preparation, 2004]. For the selected area, there were no algorithm failures, and thus the retrieval quality is impacted by biome misidentification. To estimate this impact, we reprocessed tile h18v04 using the operational algorithm and the correct biome map. Both the Collection 4 MODIS product and the reprocessed LAI values are compared against the 1 km reference map.

[43] Figure 15 shows distributions of the Collection 4, reprocessed, and reference LAI values. Distributions of the reference and reprocessed values appear to be closely allied. Both plots localize the same most probable value of LAI. Mean LAIs over the 58 km^2 area for these retrievals are coincident. The MODIS operational algorithm, however, overestimates the most probable mean LAI by 0.2. Deviations of the MODIS product and reprocessed LAIs from the reference values are shown in Figure 16. Thus the accuracy, precision, and uncertainty in the Collection 4 MODIS LAI (reprocessed LAI) are 0.2 (0.0), 0.36 (0.25), and 0.48 (0.48), respectively.

[44] It should be noted that the precision of the mean reference LAI over the 58 km^2 area is 0.4; that is, reference values exhibit essential variations within the selected area.

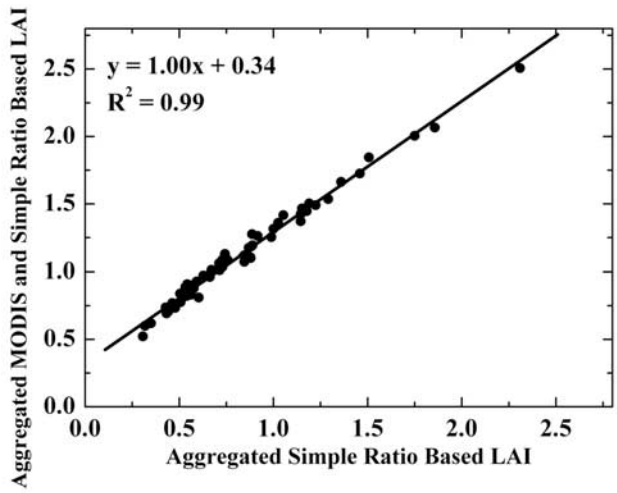

Figure 13. Correlation between 1 km LAI maps obtained by averaging 30 m simple ratio–based retrievals and 30 m MODIS LAI amplified with the adjusted simple ratio–based LAI.

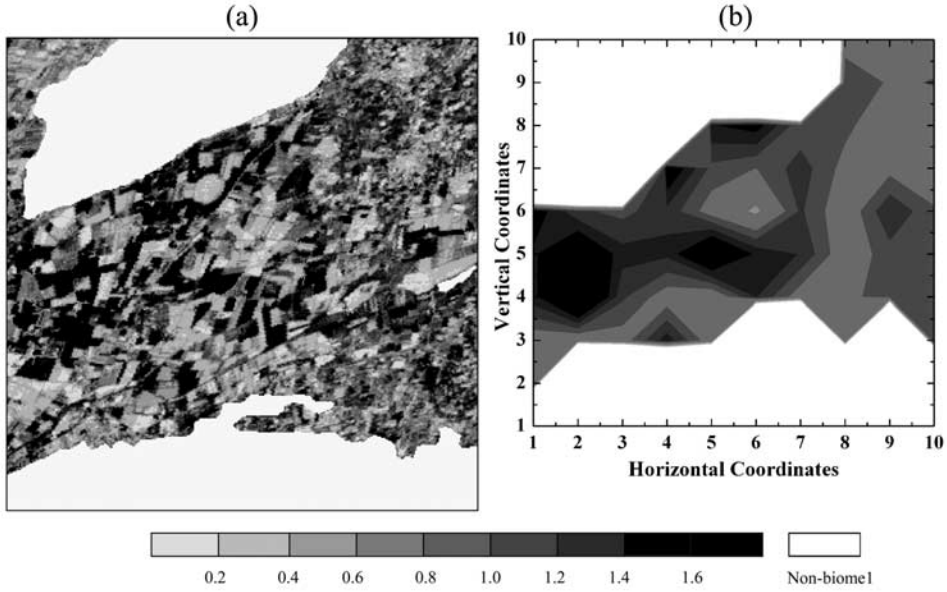

Figure 14. (a) Reference LAI maps at 30 m resolution and (b) contour plot of 1 km resolution LAI aggregated from the 30 m map for a 58 km^2 vegetated area. See color version of this figure at back of this issue.

Therefore we separated two patches on the reference map with LAI values from the intervals [0.8, 1) and [1, 1.2) (Figure 14b), which contain 18 and 12 pixels, respectively. Each patch can be represented by a mean reference LAI value with a precision of 0.06. The MODIS product overestimates the reference mean by 0.3. Its precision and uncertainty are 0.23 and 0.38, respectively. The reprocessed values are accurate to within 0.15 with a precision and uncertainty of about 0.20 and 0.25, respectively. In both cases the retrieval precision does not exceed 0.24, which is the product between the precision sensitivity factor, 1.2, and the model-observation precision, 0.2. These accuracy statements refer to a homogeneous patch of about 12–18 km^2.

8. Conclusions

[45] The biggest challenge for validation of moderate- and coarse-resolution LAI products is the extrapolation of scarce field data from sampling points to a sufficiently extensive area. One way of doing this is to employ both field measurements and high-resolution satellite data to produce validated fine-resolution LAI maps over a sufficiently extended area, to aggregate these to 1 km resolution, and to use them as benchmarks to validate the MODIS products. Retrieving LAI from high-resolution satellite data and field LAI at sampling points, or the extrapolation procedure, is a key element in the implementation of this strategy.

[46] Two factors, errors in field LAI and high-resolution satellite data, have significant impact on the extrapolation of field data. If these factors are ignored, small variations in field and satellite data due to measurements and observation/processing errors result in a low precision of the extrapolated LAI field irrespective of the extrapolation procedure used. In examples presented in this paper the extrapolation precision appeared to be ~1 order of magnitude lower than observation and measurement precisions. This result indicates that the retrieval of LAI from satellite data is an ill-posed problem. As such any retrieval technique based on a simple model inversion or empirical relationships is unable to produce stable retrievals [*Tikhonov et al.*, 1995]. The pair, observations and their inaccuracies [*Tikhonov et al.*, 1995, p. 3], "is, in general, the minimal information necessary to construct approximate solution for ill-posed problems." On the basis of this fundamental mathematical result we derived a 30 m resolution reference LAI of sampling points which accounts for both field measurement and satellite observation errors and agrees with field- and satellite-derived LAI within accuracy and precision of measurements and observations. The 30 m reference was then extrapolated from sampling points to a 58 km^2 area

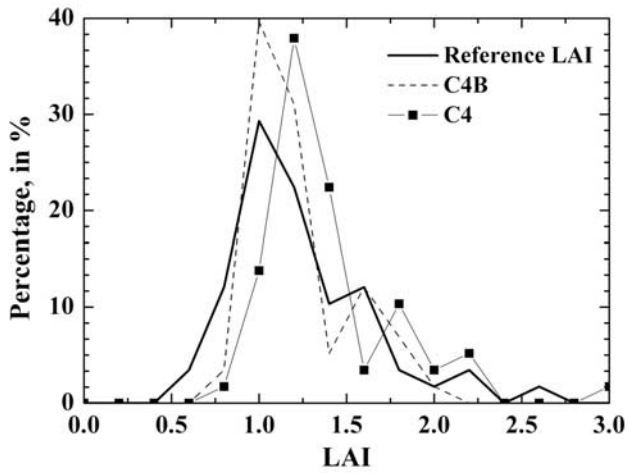

Figure 15. Distribution of the Collection 4 (C4), reprocessed (C4B), and reference LAI values at a resolution of 1 km.

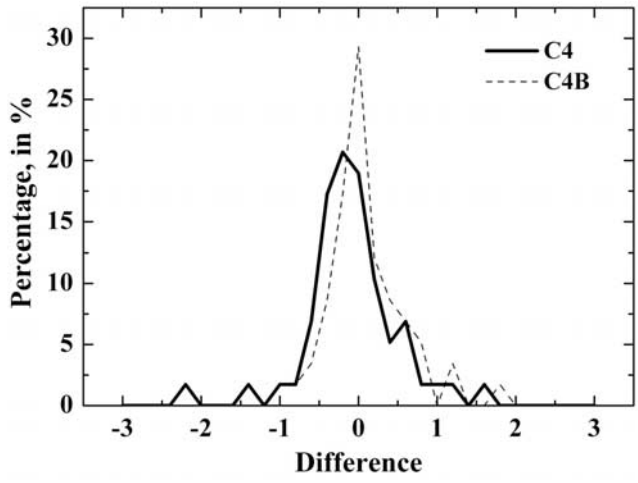

Figure 16. Distribution of the differences between Collection 4 LAI and the reference (C4) and between reprocessed LAI and the reference (C4B).

without loss in the quality and was degraded to a 1 km resolution LAI map. The latter is taken as a reference to specify quality of the MODIS product. Comparison of the reference and corresponding MODIS retrievals suggests satisfactory performance of the MODIS LAI algorithm.

Appendix A: Accuracy, Precision, and Uncertainty

[47] Here we follow S. W. Miller (VIIRS vegetation index (VVI) ATBD (version 5), National Polar-Orbiting Operational Environmental Satellite System, Silver Spring, Maryland, 2002, available at http://npoesslib.ipo.noaa.gov/atbd_viirs_v5.htm). Consider a single true value, T, of a satellite product or field-measured value at the pixel level. Various error sources cause an estimate X_i of T to deviate from T. The index i indicates that any arbitrary number, N, of such estimates can be made. The accuracy, A, is defined as

$$A = |\mu - T|,$$

where μ is the average of all the measured values X_i corresponding to a single true value T. The accuracy is also called a bias, and it is a direct comparison between the measurements X_i and the true value T. The precision, P, is defined as the standard deviation of the measurements:

$$P = \left[\frac{1}{N-1}\sum_{i=1}^{N}(X_i - \mu)^2\right]^{\frac{1}{2}}.$$

Both the accuracy and precision are defined for measurements corresponding to a single value of T.

[48] Consider a number of true values, T_k, and their estimates, Y_k, $k = 1, 2, \ldots, M$. The uncertainty, U, is defined as

$$U = \left[\frac{1}{M}\sum_{k=1}^{M}(Y_k - T_k)^2\right]^{\frac{1}{2}}.$$

The uncertainty is alternatively known as the RMS error between the measurements X_k and the true value T_k.

[49] The relative accuracy (precision) is defined as accuracy (precision) normalized by the average of all the measured values X_i corresponding to a single true value T. The relative uncertainty is the ratio $U/\overline{T}$, where $\overline{T}$ is the average of all T_k.

[50] **Acknowledgments.** This work was funded by the NASA Earth Science Enterprise under the MODIS contract to Boston University. The authors thank the AERONET group for making their field data available. The authors also acknowledge the contributions of several individuals to the MODIS land data processing, notably, N. El Saleous, C. Justice, D. Roy, E. Vermote, and R. Wolfe.

References

Baret, F., G. Guyot, A. Begue, P. Maurel, and A. Podaire (1988), Complementarity of middle infrared with visible and near-infrared reflectance of monitoring wheat canopies, *Remote Sens. Environ.*, *26*, 213–225.

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

Butera, C. (1986), A correlation and regression analysis of percent canopy closure and TM spectral response for selected forest sites in San Juan national forest, Colorado, *IEEE Trans. Geosci. Remote Sens.*, *24*, 122–129.

Cawley, G. C., and N. L. C. Talbot (2003), Efficient leave-one-out cross-validation of kernel Fisher discriminant classifiers, *Pattern Recognition*, *36*, 2585–2592.

Crist, E. P., and R. C. Cicone (1984), A physically-based transformation of thematic mapper data—The TM tasseled cap, *IEEE Trans. Geosci. Remote Sens.*, *22*, 256–263.

Curran, P. J., and A. M. Hay (1986), The importance of measurement error for certain procedures in remote sensing at optical wavelengths, *Photogramm. Eng. Remote Sens.*, *52*, 229–241.

Efron, B., and G. Gong (1983), A leisurely look at the bootstrap, the jackknife, and cross-validation, *Am. Stat.*, *37*, 36–48.

Elvidge, C. D., and Z. Chen (1995), Comparison of broad-band and narrow-band red and near-infrared vegetation indices, *Remote Sens. Environ.*, *54*, 38–48.

Friedl, M. A., D. K. McIver, J. C. F. Hodges, X. Y. Zhang, D. Muchoney, A. H. Strahler, C. E. Woodcock, S. Gopal, A. Schneider, and A. Cooper (2002), Global land cover mapping from MODIS: Algorithms and early results, *Remote Sens. Environ.*, *83*, 287–302.

Gong, G. (1986), Cross-validation, the jackknife, and the bootstrap: Excess error estimation in forward logistic regression, *JASA J. Am. Stat. Assoc.*, *81*, 108–113.

Hausman, J. (2001), Mismeasured variables in econometric analysis: Problems from the right and problems from the left, *J. Econ. Perspect.*, *15*, 57–67.

Holben, B. N., et al. (1998), AERONET—A federated instrument network and data archive for aerosol characterization, *Remote Sens. Environ.*, *66*, 1–16.

Jackson, R. D. (1983), Spectral indices in *n*-space, *Remote Sens. Environ.*, *13*, 409–421.

Jonckheere, I., S. Fleck, K. Nackaerts, B. Muys, P. Coppin, M. Weiss, and F. Baret (2004), Review of methods for in situ leaf area index determination part I. Theories, sensors and hemispherical photography, *Agric. For. Meteorol.*, *121*, 19–35.

Jordan, C. F. (1969), Derivation of leaf-area index from quality of light on the forest floor, *Ecology*, *50*, 663–666.

Justice, C., A. Belward, J. Morisette, P. Lewis, J. Privette, and F. Baret (2000), Developments in the 'validation' of satellite sensor products for the study of the land surface, *Int. J. Remote Sens.*, *21*, 3383–3390.

Katila, M., and E. Tomppo (2001), Selecting estimation parameters for the Finnish multisource National Forest Inventory, *Remote Sens. Environ.*, *76*, 16–32.

Kauth, R. J., and G. S. Thomas (1976), The tasseled cap—A graphic description of the spectral-temporal development of agricultural crops as seen in Landsat, in *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data, June 29–July 1, 1976, the Laboratory for Applications of Remote Sensing, Purdue University, West Lafayette, Indiana*, edited by P. H. Swain, pp. 41–51, Inst. of Electr. and Electron. Eng., New York.

Knyazikhin, Y., J. V. Martonchik, R. B. Myneni, D. J. Diner, and S. W. Running (1998a), Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radia-

tion from MODIS and MISR data, *J. Geophys. Res.*, *103*, 32,257–32,275.

Knyazikhin, Y., J. V. Martonchik, D. J. Diner, R. B. Myneni, M. Verstraete, B. Pinty, and N. Gorbon (1998b), Estimation of vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from atmosphere-corrected MISR data, *J. Geophys. Res.*, *103*, 32,239–32,256.

Martonchik, J. V., D. J. Diner, R. A. Kahn, T. P. Ackerman, M. M. Verstraete, B. Pinty, and H. R. Gordon (1998), Techniques for the retrieval of aerosol properties over land and ocean using multiangle imaging, *IEEE Trans. Geosci. Remote Sens.*, *36*, 1212–1227.

Miller, J. B. (1967), A formula for average foliage density, *Aust. J. Bot.*, *15*, 141–144.

Myneni, R. B. (1991), Modelling radiative transfer and photosynthesis in three dimensional vegetation canopies, *Agric. For. Meteorol.*, *55*, 323–344.

Myneni, R. B., F. G. Hall, P. J. Sellers, and A. L. Marshak (1995), The meaning of spectral vegetation indices, *IEEE Trans. Geosci. Remote Sens.*, *33*, 481–486.

Myneni, R. B., et al. (2002), Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, *Remote Sens. Environ.*, *83*, 214–231.

Nemani, R., L. Pierce, S. Running, and L. Band (1993), Forest ecosystem process at the watershed scale: Sensitivity to remotely-sensed leaf area index estimates, *Int. J. Remote Sens.*, *14*, 2519–2534.

Panferov, O., Y. Knyazikhin, R. B. Myneni, J. Szarzynski, S. Engwald, K. G. Schnitzler, and C. Gravenhorst (2001), The role of canopy structure in the spectral variation of transmission and absorption of solar radiation in vegetation canopies, *IEEE Trans. Geosci. Remote Sens.*, *39*, 241–253.

Rumelhart, D. E., G. E. Hinton, and R. J. Williams (1986), Learning internal representations by error propagation, in *Parallel Distributed Processing*, edited by D. E. Rumelhart and J. McClelland, pp. 318–362, MIT Press, Cambridge, Mass.

Sellers, P. J., et al. (1997), Modeling the exchanges of energy, water, and carbon between continents and the atmosphere, *Science*, *275*, 502–509.

Shabanov, N. V., L. Zhou, Y. Knyazikhin, R. B. Myneni, and C. J. Tucker (2002), Analysis of interannual changes in northern vegetation activity observed in AVHRR data during 1981 to 1994, *IEEE Trans. Geosci. Remote Sens.*, *40*, 115–130.

Smolander, S., and P. Stenberg (2003), A method to account for shoot scale clumping in coniferous canopy reflectance models, *Remote Sens. Environ.*, *88*, 363–373.

Tian, Y., Y. Zhang, Y. Knyazikhin, R. B. Myneni, J. Glassy, G. Dedieu, and S. W. Running (2000), Prototyping of MODIS LAI and FPAR algorithm with LASUR and LANDSAT data, *IEEE Trans. Geosci. Remote Sens.*, *38*, 2387–2401.

Tian, Y., et al. (2002a), Multiscale analysis and validation of the MODIS LAI product. I. Uncertainty assessment, *Remote Sens. Environ.*, *83*, 414–430.

Tian, Y., et al. (2002b), Multiscale analysis and validation of the MODIS LAI product. II. Sampling strategy, *Remote Sens. Environ.*, *83*, 431–441.

Tian, Y., Y. Wang, Y. Zhang, Y. Knyazikhin, J. Bogaert, and R. B. Myneni (2002c), Radiative transfer based scaling of LAI/FPAR retrievals from reflectance data of different resolutions, *Remote Sens. Environ.*, *84*, 143–159.

Tikhonov, A. N., A. V. Goncharsky, V. V. Stepanov, and A. G. Yagola (Eds.) (1995), *Numerical Methods for the Solution of Ill-Posed Problems*, Springer, New York.

Tucker, C. J. (1979), Red and photographic infrared linear combinations for monitoring vegetation, *Remote Sens. Environ.*, *8*, 127–150.

Vermote, E. F., D. Tanre, J. L. Deuze, M. Herman, and J. J. Morcrette (1997), Second simulation of the satellite signal in the solar spectrum, 6S: An overview, *IEEE Trans. Geosci. Remote Sens.*, *35*, 675–685.

Vladimirov, V. S. (1963), Mathematical problems in the one-velocity theory of particle transport, *Tech. Rep. AECL-1661*, At. Energy of Can. Ltd., Chalk River, Ont., Canada.

Wang, Y., Y. Tian, Y. Zhang, N. El-Saleous, Y. Knyazikhin, E. Vermote, and R. B. Myneni (2001), Investigation of product accuracy as a function of input and model uncertainities: Case study with SeaWiFS and MODIS LAI/FPAR algorithm, *Remote Sens. Environ.*, *78*, 296–311.

Wang, Y., W. Buermann, P. Stenberg, H. Smolander, T. Häme, Y. Tian, J. Hu, Y. Knyazikhin, and R. B. Myneni (2003), A new parameterization of canopy spectral response to incident solar radiation: Case study with hyperspectral data from pine dominant forest, *Remote Sens. Environ.*, *85*, 304–315.

Wang, Y., et al. (2004), Evaluation of the MODIS LAI algorithm at a coniferous forest site in Finland, *Remote Sens. Environ.*, *91*, 114–127.

Weiss, M., F. Baret, G. J. Smith, I. Jonckheere, and P. Coppin (2004), Review of methods for in situ leaf area index (LAI) determination part II. Estimation of LAI, errors and sampling, *Agric. For. Meteorol.*, *121*, 37–53.

J. Hu, D. Huang, Y. Knyazikhin, R. B. Myneni, N. Shabanov, B. Tan, and P. Zhang, Department of Geography, Boston University, Boston, MA 02215, USA. (tanbin@crsa.bu.edu)

M. Weiss, Noveltis, Parc technologique du canal, 2 Avenue de l'Europe, Ramonville-Saint-Agne F-31520, France.

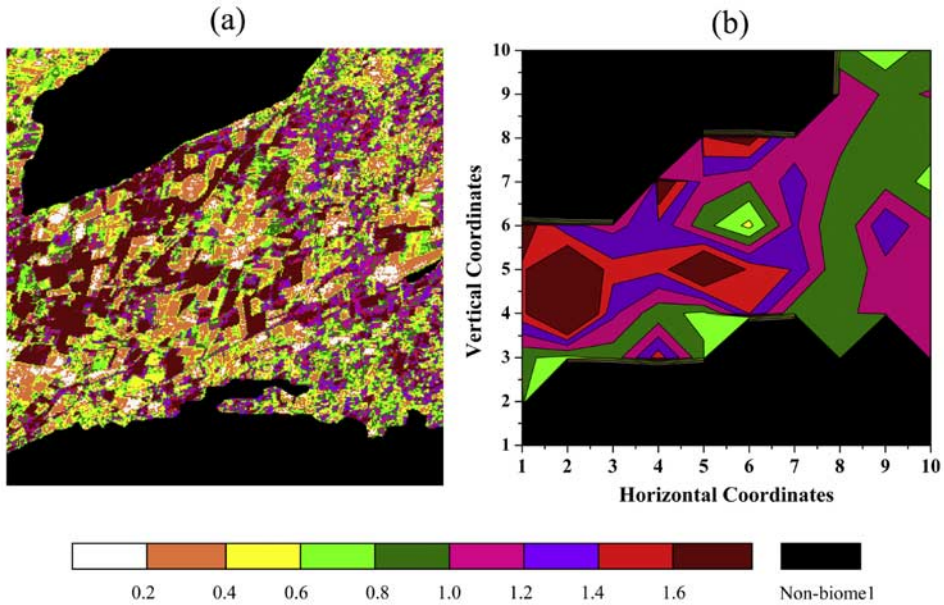

Figure 14. (a) Reference LAI maps at 30 m resolution and (b) contour plot of 1 km resolution LAI aggregated from the 30 m map for a 58 km^2 vegetated area.