

Article

The Impact of Potential Land Cover Misclassification on MODIS Leaf Area Index (LAI) Estimation: A Statistical Perspective

Hongliang Fang ^{1,*}, Wenjuan Li ¹ and Ranga B. Myneni ²

¹ LREIS, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; E-Mail: liwj@lreis.ac.cn

² Department of Geography and Environment, Boston University, Boston, MA 02215, USA; E-Mail: ranga.myneni@gmail.com

* Author to whom correspondence should be addressed; E-Mail: fanghl@lreis.ac.cn; Tel.: +86-10-6488-8005; Fax: +86-10-6488-9630.

Received: 18 December 2012; in revised form: 4 February 2013 / Accepted: 4 February 2013 /

Published: 15 February 2013

Abstract: Understanding the impact of vegetation mixture and misclassification on leaf area index (LAI) estimation is crucial for algorithm development and the application community. Using the MODIS standard land cover and LAI products, global LAI climatologies and statistics were obtained for both pure and mixed pixels to evaluate the effects of biome mixture on LAI estimation. Misclassification between crops and shrubs does not generally translate into large LAI errors (<0.37 or 27.0%), partly due to their relatively lower LAI values. Biome misclassification generally leads to an LAI overestimation for savanna, but an underestimation for forests. The largest errors caused by misclassification are also found for savanna (0.51), followed by evergreen needleleaf forests (0.44) and broadleaf forests (~0.31). Comparison with MODIS uncertainty indicators show that biome misclassification is a major factor contributing to LAI uncertainties for savanna, while for forests, the main uncertainties may be introduced by algorithm deficits, especially in summer. The LAI climatologies for pure pixels are recommended for land surface modeling studies. Future studies should focus on improving the biome classification for savanna systems and refinement of the retrieval algorithms for forest biomes.

Keywords: leaf area index (LAI); uncertainty; land cover; biome type; subpixel mixture; biome misclassification; MODIS

1. Introduction

Leaf area index (LAI) quantifies the amount of live green leaves in the canopy per unit of ground surface. It is an important parameter in various vegetation ecosystem and land surface process models [1]. Global LAI products have been operationally provided through several satellite remote sensing projects, such as MODIS [2,3], CYCLOPES [4] and GLOBCARBON [5,6]. Understanding the uncertainties of these products is critical in order to assimilate LAI into the ecosystem and land surface models [1,7]. Product uncertainty information can be categorized into two types, theoretical and physical [8]. Theoretical uncertainties are caused by uncertainties in the input data and model imperfections and are usually estimated and reported in the quantitative quality indicators (QQIs) [4,9,10]. Physical uncertainties are derived through comparison with values representing the ground truth—field measurements or estimations from higher-resolution imagery. In practice, both uncertainties are used as complementary indicators of product quality [1].

Previous studies have identified three key sources of LAI uncertainties [11–13]: (1) uncertainties in the input data; (2) model uncertainties and problems of ill-posed retrieval; and (3) errors in the ancillary information (e.g., land cover type). Many studies have discussed LAI uncertainties introduced by the input reflectance, e.g., [12,14] and imperfect modeling, e.g., [2,3,5,9]. However, few studies have addressed the question of uncertainties in land cover products and their impact on LAI uncertainties [15,16].

Land cover information is commonly used to parameterize canopy radiative transfer models or models that require land cover stratification, such as MODIS [3], GLOBCARBON [5] and ECOCLIMAP [17]. Errors in classifying land cover type may thus propagate into LAI uncertainties during the parameter retrieval process. Given the modest accuracy of current global land cover products (overall accuracy 60%–80%) [18,19], understanding the impact of land cover misclassification on LAI estimation is fundamental to improving LAI retrieval from remote sensing imagery.

Previous studies of the effect of land cover misclassification on LAI estimation have adopted either a deterministic or a statistical approach. The deterministic approach simulates the physical radiative transfer processes within vegetation canopies that depend on land cover type. The relationship between misclassification and LAI uncertainty can be explored using numerical experiments and trial-and-error methods [20]. Depending on the LAI estimation methods used, biome misclassification may lead to incorrect selection of look-up tables, inappropriate radiative transfer models and/or estimation algorithms [16,20,21]. The statistical approach directly explores the relationship between the input biome classification map and the resulting LAI uncertainty, which avoids the usually complex radiative transfer simulation and parameter retrieval processes. The effects of biome misclassification can be evaluated through the retrieval index, mean LAI and the histogram of the retrieved LAI distribution [9,22]. Despite its simplicity, few studies have been conducted systematically using this approach on a global scale.

In this context, the focus of this paper is to explore the effect of biome misclassification on MODIS LAI estimation using a novel statistical approach. The MODIS products contain both collocated land cover and LAI products and were thus explored in this study. A confusion matrix was constructed to investigate the correspondence between biome misclassification and LAI uncertainties. Global LAI climatologies were compared for pure and mixed pixels to explore the impact of different biome misclassification on LAI errors.

2. Data and Methods

2.1. MODIS LAI and Land Cover Products

The 8-day synergistic LAI (MCD15 C5, 1 km) and land cover (MCD12Q1 C5, 500 m) combined products from the TERRA and AQUA platforms for 2003–2009 were available from the NASA WIST (Warehouse Inventory Search Tool) website (WIST, Available online: <http://wist.echo.nasa.gov> (accessed on 1 March 2012)). The MODIS LAI retrieval algorithm is based on the inversion of a 3D radiative transfer model that simulates radiation absorption and scattering in vegetation canopies [2,23,24]. The main algorithm employs a look-up table (LUT) method that searches for LAIs for specific solar and view zenith angles, observed bidirectional reflectance factors (BRFs) at certain spectral bands and biome type. The output is the LAI mean value averaged over all acceptable solutions. The standard deviation (LaiStdDev) serves as a measure of accuracy and is provided as a product QQI. The current Collection 5 LAI products use eight biome types as *a priori* information to constrain the vegetation structural and optical parameter space: (1) grasses and cereal crops, (2) shrubs, (3) broadleaf crops, (4) savanna, (5) deciduous broadleaf forest (DBF), (6) evergreen broadleaf forest (EBF), (7) deciduous needleleaf forest (DNF) and (8) evergreen needleleaf forest (ENF) [13].

MODIS global land cover products are generated utilizing a supervised classification method that exploits a global training database obtained from high-resolution imagery in association with ancillary data [18]. The classification algorithm requires spectral and temporal information from MODIS nadir BRDF-adjusted reflectance (NBAR) bands 1–7, supplemented by the enhanced vegetation index (EVI), and the MODIS land surface temperature (LST) is required to obtain the land cover types and phenology information. Several classification schemes are adopted, principally, the International Geosphere Biosphere Program (IGBP) (Type 1), the University of Maryland (UMD) (Type 2) and LAI/FAPR biomes (Type 3). In the IGBP classification scheme, in addition to the primary types, an alternative secondary type was assigned for each pixel when the confidence in the primary type is not high [25]. Pixels with no secondary label were considered unequivocal with high confidence [25]. The paired primary and secondary land cover types make it possible to study the potential LAI uncertainty caused by biome misclassification. In this study, ‘misclassification’ represents both subpixel mixture and biome misclassification.

2.2. Data Analysis

The uncertainty of LAI retrieval resulting from biome misclassification was analyzed on the basis of the primary and secondary land cover types. For ease of analysis, the Type 3 MODIS LAI/Fraction of Absorbed Photosynthetically Active Radiation (FPAR) eight biome types were used in the confusion matrix and subsequent analysis. Since the secondary types are only provided in the Type 1 IGBP scheme, they were converted to the Type 3 MODIS LAI/FPAR biome types in the confusion matrix and subsequent analysis [18]. Figures 1 and 2 show the global primary and secondary land cover types, respectively, in the Type 3 classification system in 2003. The secondary biome type indicates a potential vegetation mixture or a misclassification contributing to LAI uncertainties. For the secondary biome types, land cover types included in the Type 1, but with no equivalent Type 3 class, were not included in the analysis in order to minimize the discrepancies between different

classification schemes. These types include the mixed forest, permanent wetlands and cropland/natural vegetation mosaic (white in Figure 2). Pixels with high confidence (pink in Figure 2) and of identical primary and secondary type were regarded as representing ‘pure’ biome types with minimal biome mixing or misclassification. Otherwise, the pixels were regarded as ‘mixed’ or ‘misclassified’—both terms are used interchangeably in the text. Assuming that the secondary type mainly indicated biome misclassification, a confusion matrix was constructed to investigate the impact of such misclassification on LAI uncertainties. We consider the primary biome type as the actual class and the secondary type as the predicted class in the confusion matrix. LAI uncertainty was calculated as the differences between the predicted and actual vegetation LAIs (Equation (1)):

$$\text{LAI Uncertainty} = \text{LAI}_{\text{mixed}} - \text{LAI}_{\text{pure}} \quad (1)$$

where $\text{LAI}_{\text{mixed}}$ and LAI_{pure} are the average LAI values for the mixed (predicted) and pure (actual) pixels, respectively. LAI uncertainties were then examined by comparing their climatologies for all biomes. The misclassification induced errors (MIEs) were further compared with the theoretical uncertainties reported in MODIS LAI QQIs. Our goals to calculate the global LAI climatology were to obtain an overview of the LAI uncertainties and their performances against the MODIS LAI QQIs. Only LAI values retrieved from the main algorithm and ‘good quality’ land cover data based on the quality assessment layer were analyzed in this study. Considering the yearly land cover variation, the collocated LAI and land cover data from 2003 to 2009 were used for the calculation.

Figure 1. Global distribution of the primary biome types based on the MODIS Leaf Area Index (LAI)/Fraction of Absorbed Photosynthetically Active Radiation (FPAR) (Type 3) classification system. Data from the MODIS (MCD12Q1 C5) land cover product in 2003 (1 km).

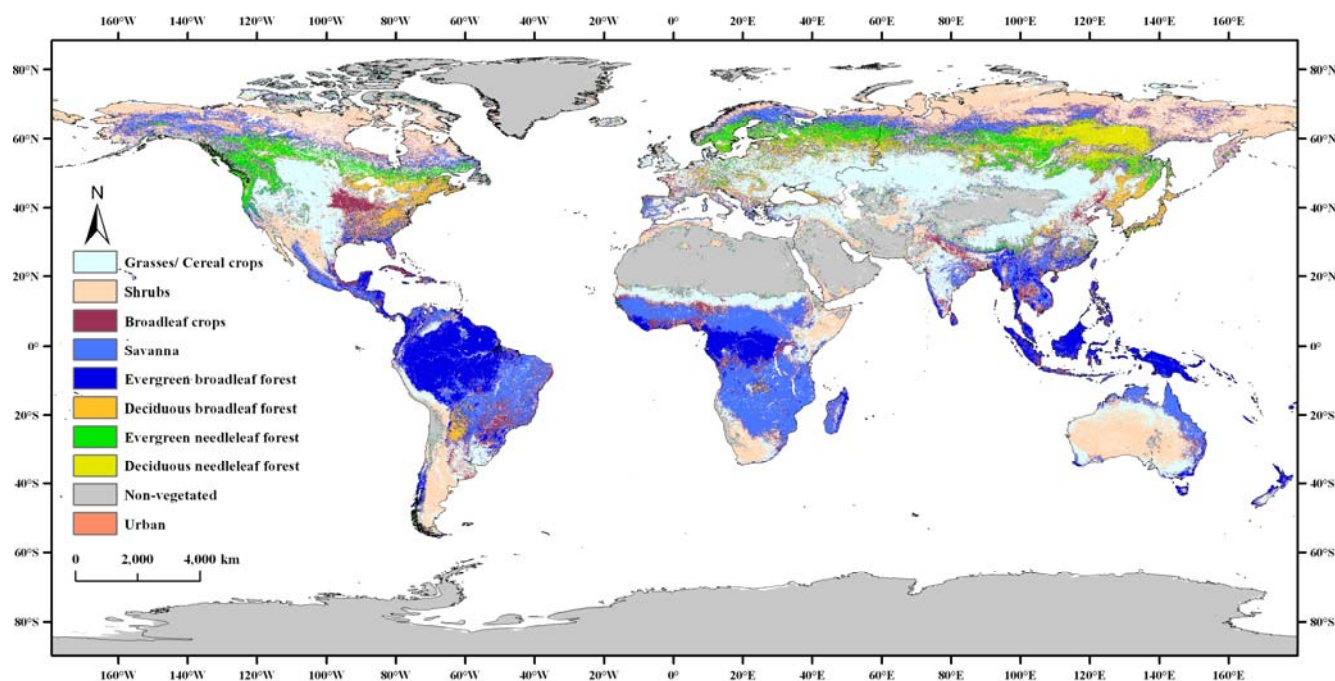
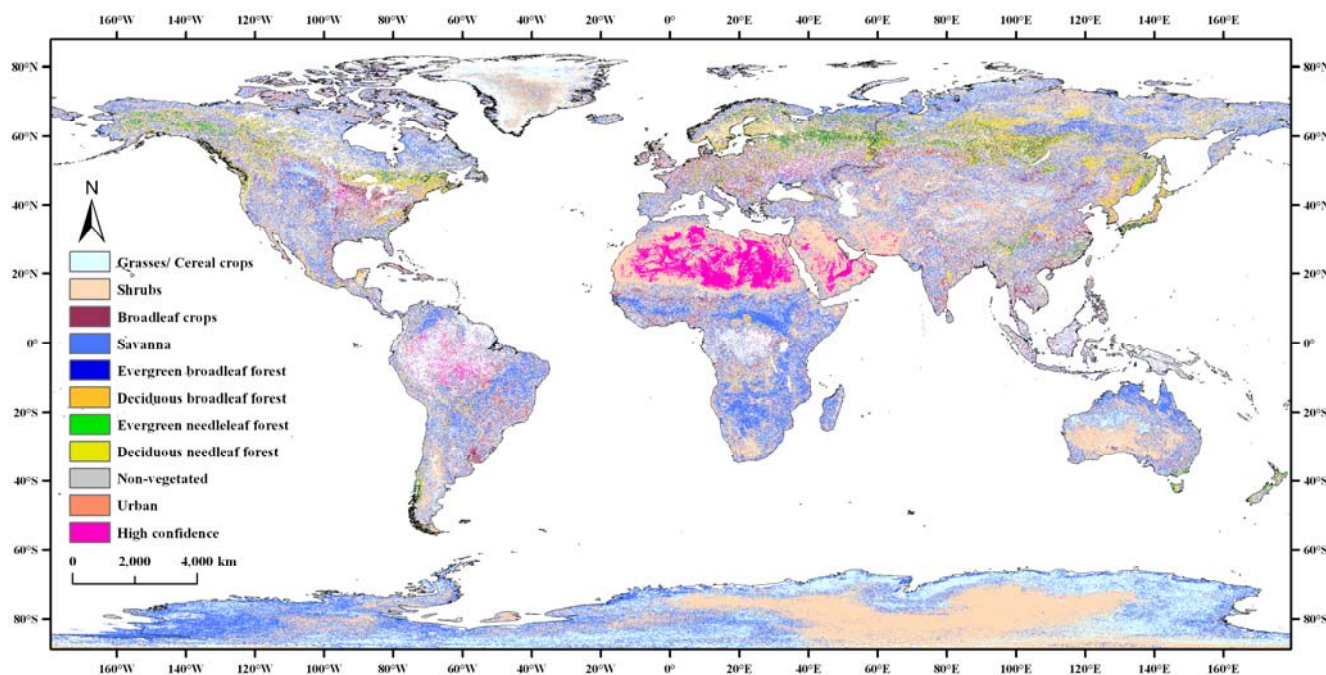


Figure 2. Global distribution of the secondary biome types based on the MODIS LAI/FPAR (Type 3) classification system. Data converted from the MODIS (MCD12Q1 C5) International Geosphere Biosphere Program (IGBP) (Type 1) classification system (1 km, 2003). Pink pixels show the primary biome types with high confidence. White areas are IGBP classes (e.g., mixed forest) with no equivalent MODIS LAI/FPAR classes.



3. Results

3.1. Misclassification between Different Biome Types

Table 1 shows the statistics of the primary and secondary biome types. Pure pixels are located on the diagonal cells, and all others are considered mixed pixels. The last row shows the percentage of pure pixels over the globe. The table reveals that secondary biome types are common and correspond reasonably to the primary types. The total percentage of high confidence (HC) pixels is only 0.92% over the globe. Overall, only 28.74% are pure pixels, leaving most (71.26%) as mixed pixels. Grasses, crops, shrubs and savannas are easily confused with each other, but unlikely to be misclassified as forests. For example, grasses/cereal crops (35.69%) and broadleaf crops (36.94%) are likely to be confused with savannas, but only 6.15% of savannas are mistaken for deciduous broadleaf forest. Because of the mixed composition of trees and grasses, 49.37% of savannas, higher than all other biome types, are classified as pure pixels.

Of the forest biomes in Table 1, deciduous broadleaf forest contains the highest percentage of pure pixels (26.72%). However, all forest types are often confused with one another. For example, evergreen broadleaf forest is easily mistaken for deciduous broadleaf forest (10.64%) and *vice versa* (8.96%), and evergreen needleleaf forest is often mistaken for deciduous needleleaf forest (16.21%). Deciduous broadleaf forests are very likely to be misclassified as deciduous needleleaf forests (16.69%), and deciduous needleleaf forests are also likely to be misclassified as deciduous broadleaf forests (12.21%). Forest biomes are likely to be misclassified as shrubs (>21%) or savanna (>16%)

(Table 1). It is rarely possible to obtain pure pixels for evergreen broadleaf forest (3.87%), because of their confusion with shrubs (25.09%) and savanna (31.72%). Deciduous needleleaf forests are most likely to be misclassified as shrubs (33.84%) and savanna (30.96%).

Table 1. MODIS primary biome types and the percentage of each collocated secondary biome types from 2003–2009. The last row shows the percentage of pure pixels (diagonal) over the globe. The HC (high confidence) column shows the percentage of high confidence pixels for each primary biome type. The last column shows the percentage of each primary biome type over the globe. EBF, DBF, ENF and DNF stand for evergreen broadleaf forest, deciduous broadleaf forest, evergreen needleleaf forest and deciduous needleleaf forest, respectively.

↓Primary Biome Type	Secondary Biome Type									%
	Grasses/ Cereal Crops	Shrubs	Broadleaf Crops	Savanna	EBF	DBF	ENF	DNF	HC (%)	
Grasses/cereal crops	16.74	31.44	14.96	35.69	0.19	0.53	0.16	0.28	0.02	22.90
Shrubs	20.16	40.54	4.94	30.79	0.22	1.73	0.15	1.42	0.05	21.7
Broadleaf crops	19.35	19.74	19.04	36.94	0.55	2.52	0.30	1.08	0.48	9.42
Savanna	12.22	19.24	7.04	49.37	1.65	6.15	1.06	3.27	0.00	21.17
EBF	8.75	25.09	5.69	31.72	3.87	10.64	2.15	4.16	7.93	10.61
DBF	2.93	22.58	1.03	16.90	8.96	26.72	4.14	16.69	0.05	5.83
ENF	3.00	21.56	1.01	19.24	10.88	15.94	11.98	16.21	0.19	6.43
DNF	2.42	33.84	0.82	30.96	5.29	12.21	3.52	10.75	0.19	1.94
Pure pixels (%)	3.83	8.79	1.79	10.45	0.41	1.56	0.77	0.21	0.92	28.74

3.2. Misclassification Induced LAI Errors (MIEs)

Table 2 shows the LAI mean values for both the pure and mixed pixels and the variability in LAI values induced by potential vegetation mixture and misclassification. The table provides an insight into the extent to which the confusion between two biome types affects LAI retrievals. The impact of biome misclassification varies between biome types. The bias due to the misclassification of grasses/cereal crops ranges from -0.29 (-51.8%) for shrubs to 0.14 (25.0%) for broadleaf crops. The misclassification of shrubs as savanna leads to an overestimation of up to 60% . By contrast, the misclassification of broadleaf crops as shrubs underestimates the LAI by 0.37 (-27.0%). Misclassification of savannas as any of the herbaceous types overestimates the LAI (>0.57), with the greatest bias being 0.84 (100.0%) for broadleaf crops.

Forests, especially evergreen forests, are easily misclassified as shrubs or savannas. Significant underestimation of LAI is observed when evergreen broadleaf forest (0.67 , -16.1%) or evergreen needleleaf forest (0.62 , -31.2%) is mixed with shrubs. The mixture with savanna leads to the largest underestimation (0.87 , -43.7%) for evergreen needleleaf forest. The large errors related to shrubs and savanna for the evergreen forests could be due to the selection of wrong LUTs [26]. Evergreen broadleaf forest can be confused with deciduous broadleaf forest, which causes an underestimation of LAI up to 0.41 (-9.8%). However, misidentifying deciduous broadleaf forest as evergreen broadleaf forest has nearly no effect on LAI values (0.03 , -1.4%). Overall, confusion between broadleaf forest and needleleaf

forest generally leads to an LAI underestimation. Confusion between evergreen and deciduous needleleaf forests causes an underestimation of about 0.24 (−12.1%). When evergreen needleleaf forest is misclassified as broadleaf forest, the underestimation is up to 0.23 (−11.6%). Misclassification of deciduous needleleaf forest as broadleaf forest slightly affects LAI values (0.07, −4.2%).

Table 2. (a) Confusion matrix for LAI mean values for pure (bold) and mixed pixels and (b) the relative LAI errors induced by biome misclassification. Statistics based on data from 2003–2009. Mean values for mixed pixels are significantly different from those of the pure pixels (*t*-test, $p < 0.001$). Relative errors are calculated relative to the diagonal pure values. The cells in brackets indicate a smaller number of pixels (<5%) from Table 1.

(a) Mean LAI values								
↓Primary Biome Type	Secondary Biome Type							
	Grasses/ Cereal Crops	Shrubs	Broadleaf Crops	Savanna	EBF	DBF	ENF	DNF
Grasses/cereal crops	0.56	0.27	0.70	0.39	(1.29)	(0.94)	(0.93)	(0.91)
Shrubs	0.41	0.29	(0.78)	0.46	(1.03)	(0.83)	(1.04)	(0.81)
Broadleaf crops	1.29	1.00	1.37	1.10	(1.34)	(1.21)	(1.23)	(1.18)
Savanna	1.41	1.63	1.68	0.84	(1.91)	2.18	(1.25)	(1.99)
EBF	4.02	3.50	4.48	3.97	(4.17)	3.76	(3.09)	(3.08)
DBF	(1.91)	2.07	(1.97)	2.07	2.14	2.17	(2.18)	2.04
ENF	(0.72)	1.37	(0.72)	1.12	1.86	1.76	1.99	1.75
DNF	(1.33)	1.52	(1.12)	1.51	1.60	1.64	(1.54)	1.67

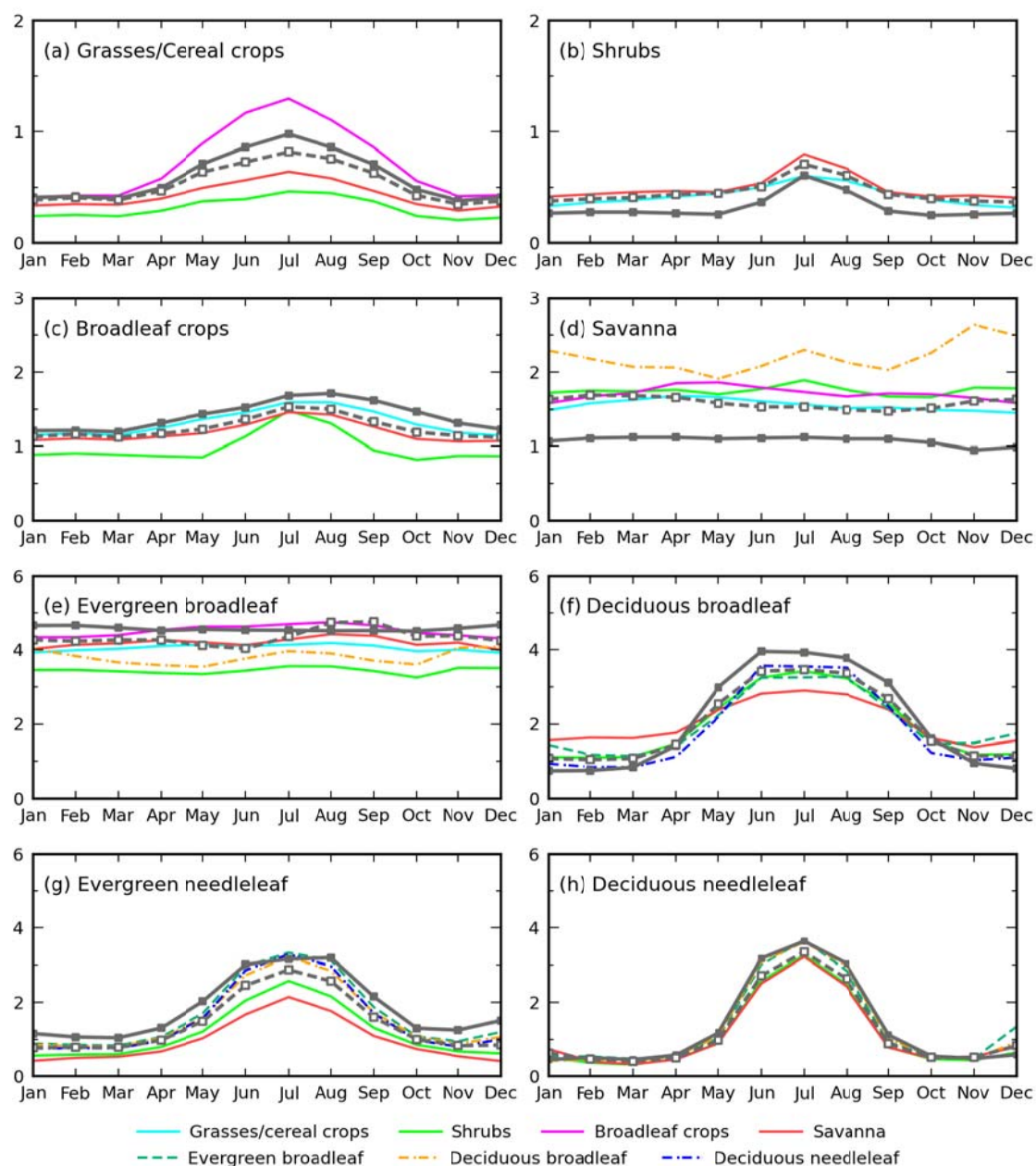
(b) Relative errors								
Grasses/cereal crops	0.0	−51.8	25.0	−30.4	(130.4)	(67.9)	(66.1)	(62.5)
Shrubs	41.4	0.0	(169.0)	58.6	(255.2)	(186.2)	(258.6)	(179.3)
Broadleaf crops	−5.8	−27.0	0.0	−19.7	(−2.2)	(−11.7)	(−10.2)	(−13.9)
Savanna	67.9	94.1	100.0	0.00	(127.4)	159.5	(48.8)	(136.9)
EBF	−3.6	−16.1	7.4	−4.8	(0.0)	−9.8	(−25.9)	(−26.1)
DBF	(−12.0)	−4.6	(−9.2)	−4.6	−1.4	0.0	(0.5)	−6.0
ENF	(−63.8)	−31.2	(−63.8)	−43.7	−6.5	−11.6	0.0	−12.1
DNF	(−20.4)	−9.0	(−32.9)	−9.6	−4.2	−1.8	(−7.8)	0.0

3.3. LAI Climatologies for Pure and Mixed Pixels

Figure 3 illustrates the global LAI climatologies for both pure and mixed pixels in different months. For each primary biome type, the climatologies show similar temporal profiles and seasonal variations. Seasonally, except for savanna, the deviations between pure and mixed LAI values are generally higher in summer than in winter. For grasses/cereal crops and broadleaf crops, misclassification of these two biomes as any other types produces LAI deviations less than 0.70. Similarly, the overestimation of LAI for shrubs is also small (~0.19). The greatest effect of subpixel mixture on LAI estimation is for savanna, for which the mixed LAI values are consistently higher than those of the pure pixels, ranging from about 0.41 (37.3%) when misclassified as grasses/cereal crops in August to 1.69 (178.0%) as deciduous broadleaf forest in November.

LAI values for evergreen broadleaf forest are greatly underestimated when misclassified as any other type, reaching by about 1.25 (−27.8%) in October when misclassified as shrubs. Misclassification of deciduous broadleaf forest as savanna lowers the estimated LAI in summer by about 1.14 (−28.9%). Misclassification of deciduous broadleaf forest as evergreen broadleaf forest increases the LAI by about 0.94 (116.0%) in winter. The underestimation of LAI for evergreen needleleaf forest is up to 1.43 (−44.7%) in August. For deciduous needleleaf forest, misclassification generally leads to an underestimation in summer (0.68 or −21.4%), but an overestimation in winter (0.77 or 132.8%).

Figure 3. LAI climatologies for both pure and mixed pixels from 2003 to 2009. Each panel is for a primary biome type, and the temporal profiles indicate the secondary biome types in the legend. The solid dark lines are for pure pixels (same primary and secondary types), while the dashed dark lines are the average of all mixed pixels.



3.4. Comparison of MIEs and MODIS LAI Uncertainty Indicators

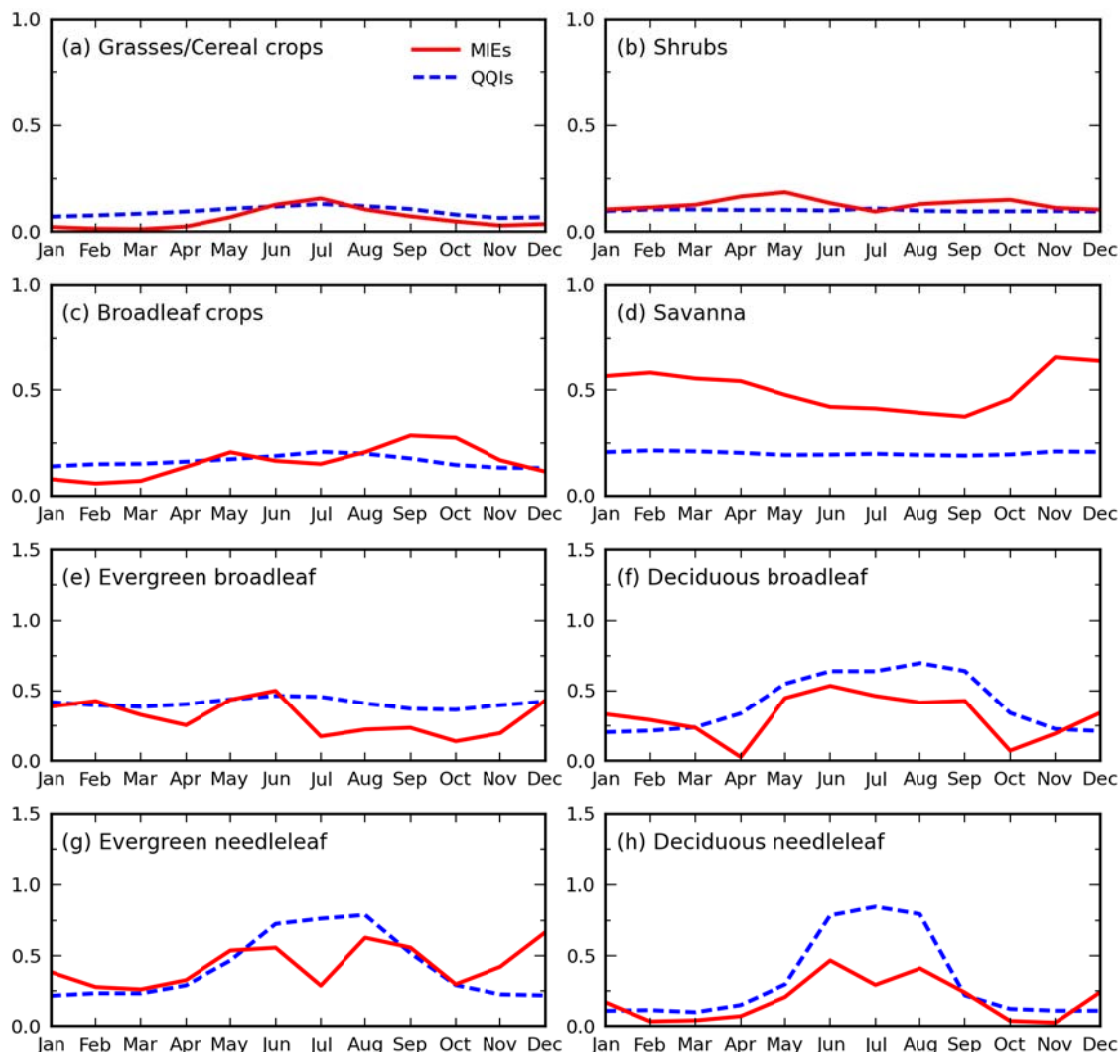
Table 3 illustrates the monthly average of MIEs calculated from the pure and mixed profiles in Figure 3 with Equation (1). The last column shows the average of the absolute errors from Equation (1). Biome misclassification generally leads to an LAI overestimation for savanna (0.51), but an underestimation for forests (0.08–0.44). Misclassification of woody biomes (≥ 0.19) generally causes higher errors than those for herbaceous types (≤ 0.16). The largest errors caused by misclassification are found for savanna (0.51), followed by evergreen needleleaf forest (0.44) and broadleaf forests (~ 0.31).

Table 3. Monthly average of the misclassification induced LAI errors (MIEs) for the eight primary biome types (2003–2009). The mean absolute errors (MAEs) are calculated from the average of the absolute monthly errors.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Average	MAEs
Grasses/ cereal crops	-0.02	-0.01	-0.01	-0.02	-0.07	-0.13	-0.16	-0.11	-0.07	-0.05	-0.03	-0.04	-0.06	0.06
Shrubs	0.11	0.12	0.13	0.17	0.19	0.14	0.10	0.13	0.14	0.15	0.11	0.10	0.13	0.13
Broadleaf crops	-0.08	-0.06	-0.07	-0.14	-0.21	-0.17	-0.15	-0.21	-0.29	-0.28	-0.17	-0.11	-0.16	0.16
Savanna	0.57	0.58	0.55	0.54	0.48	0.42	0.41	0.39	0.37	0.46	0.66	0.64	0.51	0.51
EBF	-0.39	-0.43	-0.33	-0.25	-0.44	-0.50	-0.17	0.22	0.24	-0.14	-0.20	-0.43	-0.24	0.31
DBF	0.33	0.29	0.24	0.03	-0.45	-0.54	-0.46	-0.42	-0.43	-0.07	0.19	0.34	-0.08	0.32
ENF	-0.38	-0.28	-0.26	-0.33	-0.54	-0.56	-0.29	-0.63	-0.56	-0.30	-0.42	-0.67	-0.44	0.44
DNF	0.17	-0.03	-0.04	-0.07	-0.21	-0.47	-0.29	-0.41	-0.24	-0.04	0.03	0.25	-0.11	0.19

To acquire a better understanding of both quality indicators, the uncertainties induced by misclassification are compared with those reported in QQIs (Figure 4). QQIs mainly reflect uncertainties induced by the retrieval algorithms and the input reflectance, because the retrieval algorithm is applied for each biome type separately. In this figure, the misclassification induced errors (MIEs) are represented with the absolute differences between the pure and mixed profiles in Figure 3. MIEs and QQIs are generally small (< 0.20) and consistent for grasses/cereal crops, shrubs and broadleaf crops. This signifies that biome misclassification and algorithm limitations contribute evenly to LAI uncertainties for these biome types. For savanna, the MIEs (0.51) are significantly higher than those of QQIs (0.20). This reveals that biome misclassification is a major factor contributing to LAI uncertainties for savanna. This contrasts with earlier analyses with a deterministic approach claiming that MIEs do not exceed uncertainties in the model, e.g., [21]. The discrepancies are partly due to the different sensor (MISR) and study area (Africa) explored in [21]. For forest biomes, MIEs are generally smaller than QQIs for most of the season, indicating the robustness of the MODIS LAI algorithm to forest misclassification. Slightly higher MIEs can only be observed in winter when the LAI values are low, and hence, QQIs are usually lower. For needleleaf forests, the lower MIEs (< 0.30) in comparison to QQIs (0.75–0.85) in July indicate that, other than biome misclassification, the uncertainties may be largely introduced by algorithm deficits in summer.

Figure 4. Comparison of misclassification induced LAI errors (misclassification induced errors (MIEs) in absolute values, solid red) with theoretical uncertainties reported in quantitative quality indicators (QQIs) (dashed blue). Data calculated globally from 2003 to 2009.



4. Discussions

4.1. Misclassification and LAI Estimation Errors

Various estimates of the effects of incorrect classification have previously been reported for different data sets and estimation methods. For example, deterministic studies have indicated that misclassification can lead to 15%–50% differences in LAI estimation [3,16,20]. In general, the present results using a statistical approach (Table 2(b)) are comparable to findings from similar studies by others. Previous studies have also indicated that confusion between herbaceous and woody biomes fatally impacts on LAI retrievals [22]. On a global scale, confusion between these two biome types rarely occurs because of their distinctive features [18]. In contrast, misclassification can easily take place between grasses/cereal crops and shrubs because of their spectral and structural similarities [27–29]. However, the average LAI values for herbaceous pixels are generally low; therefore, confusion

between these two biomes does not cause large errors (<0.30), which is consistent with deterministic findings from earlier LAI collections [21,30,31].

The primary classification confusion is related to savanna and forest. The ambiguity between savanna and forest properties has been reported in other studies [18,32]. The main source of classification error has been attributed to the continuum of fractional cover and canopy structure [30,32]. These issues, as shown in the present study, result in an overestimation of LAI values consistently exceeding 0.37 for savanna. Misclassification of forest as any other biome type, however, generates an underestimation in LAI retrievals, especially for evergreen needleleaf forest (Table 3). Better characterization of the savanna biome [33] and refinements of the LAI algorithms for woody vegetation [34] may help address some of these concerns.

While other forest biomes have shown to be relatively robust to land cover uncertainties, evergreen needleleaf forest LAI exhibits the highest sensitivity to biome misclassification. This is related to the similarity of the top-of-canopy reflectances of needleleaf forests with a low ground cover and a bright (green) understory [21]. Our earlier local studies have shown that misclassification of needleleaf forest leads to an underestimation of LAI, whereas misclassification of broadleaf forest leads to an overestimation [20]. These local trends were confirmed for needleleaf forests globally in the present study. The discrepancies for broadleaf forests are attributable to differences in the Type 1 and Type 3 MODIS classification systems and the small study area in [20].

4.2. Future Prospects

LAI has long been used by the modeling community in land surface modeling studies [15,35,36]. Land surface process models use climatology-based LAI values and uncertainties as input. It should be noted that, in most land surfaces models, LAI is defined for pure pixels or a mosaic of pure vegetations [37,38]. Direct calculation of remote sensing LAI climatology based on the MODIS LAI biome types obtains the climatology of mixed biome types. The difference between pure and mixed climatologies may cause land surface models to produce incorrect simulations, especially for savanna and forest biomes (Figure 3). Improved model parameterization will need to consider the biome mixture and assign more realistic climatology values. Table 4 shows the global LAI climatologies derived from the pure pixels (2003–2009). The climatologies for pure pixels are useful for land surface models that treat pixels as a mosaic of pure vegetation types, e.g., [37]. The prescribed LAI values in such land surface models may be updated with the LAI climatologies derived here.

In this study, pixels with the same primary and secondary biome types were treated as pure pixels; others were considered to be a mixture of primary and secondary biome types. The results of this study thus depend on the proportion or percentage of the two types in the mix. Confidence in our results will be enhanced when data for the proportion of the secondary biome becomes available. Information about the fractional vegetation cover (FVC) can be explored to study the impact on LAI retrievals [31]. The effects of FVC will be quantified when the full MODIS vegetation continuous field (VCF) products (MOD44B), including the percentage of trees, grasses and bare ground, are refined (MODIS Land. Available online: <http://modis-land.gsfc.nasa.gov/vcc.html> (accessed on 1 February 2013)). It is expected that the increasing quality of the MODIS land cover products will lead to improved LAI retrieval in the future. It is important to note that the sources of uncertainties, such as biome

misclassification and retrieval algorithms, are interrelated. Rigorous quantification of the physical relationship between these factors and the influence of misclassification on LAI estimation can be made with a deterministic approach through the retrieval algorithms.

Table 4. Global LAI climatologies for the eight biome types derived from pure pixels (2003–2009).

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Overall
Grasses/cereal crops	0.41	0.42	0.40	0.50	0.71	0.86	0.98	0.86	0.71	0.48	0.38	0.41	0.59
Shrubs	0.27	0.28	0.28	0.27	0.26	0.37	0.61	0.48	0.29	0.25	0.26	0.27	0.32
Broadleaf crops	1.21	1.21	1.19	1.31	1.43	1.52	1.68	1.71	1.62	1.47	1.31	1.23	1.41
Savanna	1.07	1.11	1.12	1.12	1.10	1.11	1.12	1.10	1.10	1.05	0.95	0.99	1.08
EBF	4.65	4.66	4.59	4.52	4.55	4.53	4.52	4.51	4.51	4.50	4.57	4.67	4.57
DBF	0.75	0.76	0.84	1.42	2.99	3.95	3.93	3.78	3.11	1.61	0.95	0.81	2.08
ENF	1.16	1.06	1.04	1.31	2.02	3.01	3.16	3.20	2.15	1.30	1.25	1.51	1.85
DNF	0.45	0.50	0.45	0.57	1.18	3.18	3.64	3.04	1.12	0.55	0.48	0.58	1.31

5. Conclusions

Through an analysis of global MODIS products, this study has quantified the LAI discrepancies induced by potential subpixel mixture and biome misclassification. The statistics show that 28.74% of LAI products are for pure pixels and the other 71.26% are retrieved as mixed biome types. When misclassification between distinct biome types occurs, it does not generally translate into strong disagreement in LAI retrievals. Misclassification between herbaceous types has minimal impact on LAI retrievals (<0.37 or 27.0%). Biome misclassification generally leads to an LAI overestimation for savanna, but an underestimation for forests. The largest errors caused by misclassification are found for savanna (0.51), followed by evergreen needleleaf forest (0.44) and broadleaf forests (~0.31). Biome misclassification is a major factor contributing to LAI uncertainties for savanna, while for forests, the main source of uncertainties may be due to algorithm deficits, especially in summer. To reduce the LAI uncertainties, further efforts should therefore be focused on improving the biome classification for the structurally complex savanna systems and refinement of the retrieval algorithms for forest biomes.

Acknowledgments

This work was partly supported by the Hundred Talent Program of the Chinese Academy of Sciences and the National Natural Science Foundation of China (41171333). MODIS LAI and land cover products were obtained through NASA's Warehouse Inventory Search Tool (WIST) (WIST. Available online: <http://wist.echo.nasa.gov> (accessed on 1 March 2012)).

References

1. Global Climate Observing System (GCOS). Systematic Observation Requirements for Satellite-Based Products for Climate, 2011 Update, Supplemental Details to the Satellite-Based Component of the Implementation Plan for the Global Observing System for Climate in Support of the UNFCCC (2010 Update); 2011; p. 138. Available online: <http://www.wmo.int/pages/prog/gcos/Publications/gcos-154.pdf> (accessed on 13 November 2012).
2. Huang, D.; Knyazikhin, Y.; Wang, W.; Deering, D.W.; Stenberg, P.; Shabanov, N.; Tan, B.; Myneni, R.B. Stochastic transport theory for investigating the three-dimensional canopy structure from space measurements. *Remote Sens. Environ.* **2008**, *112*, 35–50.
3. Myneni, R.B.; Hoffman, S.; Knyazikhin, Y.; Privette, J.L.; Glassy, J.; Tian, Y.; Wang, Y.; Song, X.; Zhang, Y.; Smith, G.R.; *et al.* Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sens. Environ.* **2002**, *83*, 214–231.
4. Baret, F.; Hagolle, O.; Geiger, B.; Bicheron, P.; Miras, B.; Huc, M.; Berthelot, B.; Niño, F.; Weiss, M.; Samain, O.; *et al.* LAI, fPAR, and fCover CYCLOPES global products derived from VEGETATION Part 1: Principles of the algorithm. *Remote Sens. Environ.* **2007**, *110*, 275–286.
5. Deng, F.; Chen, J.M.; Plummer, S.; Chen, M.; Pisek, J. Algorithm for global leaf area index retrieval using satellite imagery. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2219–2229.
6. Plummer, S.; Arino, O.; Simon, M.; Steffen, W. Establishing a earth observation product service for the terrestrial carbon community: The globcarbon initiative. *Mitig. Adapt. Strat. Gl.* **2006**, *11*, 97–111.
7. Morisette, J.T.; Baret, F.; Privette, J.L.; Myneni, R.B.; Nickeson, J.E.; Garrigues, S.; Shabanov, N.V.; Weiss, M.; Fernandes, R.A.; Leblanc, S.G.; *et al.* Validation of global moderate-resolution LAI products: a framework proposed within the CEOS land product validation subgroup. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1804–1817.
8. Fang, H.; Wei, S.; Jiang, C.; Scipal, K. Theoretical uncertainty analysis of global MODIS, CYCLOPES and GLOBCARBON LAI products using a triple collocation method. *Remote Sens. Environ.* **2012**, *124*, 610–621.
9. Knyazikhin, Y.; Glassy, J.; Privette, J.L.; Tian, Y.; Lotsch, A.; Zhang, Y.; Wang, Y.; Morisette, J.T.; Votava, P.; Myneni, R.B.; *et al.* MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm Theoretical Basis Document; 1999; p. 126. Available online: http://modis.gsfc.nasa.gov/data/atbd/land_atbd.php (accessed on 13 November 2012).
10. Pinty, B.; Andreadakis, I.; Clerici, M.; Kaminski, T.; Taberner, M.; Verstraete, M.M.; Gobron, N.; Plummer, S.; Widlowski, J.L. Exploiting the MODIS albedos with the Two-stream Inversion Package (JRC-TIP): 1. Effective leaf area index, vegetation, and soil properties. *J. Geophys. Res.* **2011**, *116*, D09105.
11. Gemmell, F.; Varjo, J.; Strandstrom, M.; Kuusk, A. Comparison of measured boreal forest characteristics with estimates from TM data and limited ancillary information using reflectance model inversion. *Remote Sens. Environ.* **2002**, *81*, 365–377.

12. Wang, Y.; Tian, Y.; Zhang, Y.; El-Saleous, N.; Knyazikhin, Y.; Vermote, E.; Myneni, R.B. Investigation of product accuracy as a function of input and model uncertainties: case study with SeaWiFS and MODIS LAI/FPAR algorithm. *Remote Sens. Environ.* **2001**, *78*, 296–311.
13. Yang, W.; Tan, B.; Huang, D.; Rautiainen, M.; Shabanov, N.V.; Wang, Y.; Privette, J.L.; Huemmrich, K.F.; Fensholt, R.; Sandholt, I.; *et al.* MODIS leaf area index products: from validation to algorithm improvement. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1885–1898.
14. Vermote, E.F.; El Saleous, N.Z.; Justice, C.O. Atmospheric correction of MODIS data in the visible to middle infrared: First results. *Remote Sens. Environ.* **2002**, *83*, 97–111.
15. DeFries, R.S.; Los, S.O. Implications of land-cover misclassification for parameter estimates in global land-surface models: An example from the simple biosphere model (SiB2). *Photogramm. Eng. Remote Sensing* **1999**, *65*, 1083–1088.
16. Gonsamo, A.; Chen, J.M. Evaluation of the GLC2000 and NALC2005 land cover products for LAI retrieval over Canada. *Can. J. Remote Sens.* **2011**, *37*, 302–313.
17. Masson, V.; Champeaux, J.L.; Chauvin, F.; Meriguet, C.; Lacaze, R. A global database of land surface parameters at 1-km resolution in meteorological and climate models. *J. Clim.* **2003**, *16*, 1261–1282.
18. Friedl, M.A.; Sulla-Menashe, D.; Tan, B.; Schneider, A.; Ramankutty, N.; Sibley, A.; Huang, X. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **2010**, *114*, 168–182.
19. Herold, M.; Mayaux, P.; Woodcock, C.E.; Baccini, A.; Schmullius, C. Some challenges in global land cover mapping: An assessment of agreement and accuracy in existing 1 km datasets. *Remote Sens. Environ.* **2008**, *112*, 2538–2556.
20. Fang, H.; Liang, S. A hybrid inversion method for mapping leaf area index from MODIS data: Experiments and application to broadleaf and needleleaf canopies. *Remote Sens. Environ.* **2005**, *94*, 405–424.
21. Hu, J.; Martonchik, J.V.; Diner, D.J.; Knyazikhin, Y.; Myneni, R.B.; Tan, B.; Shabanov, N.; Crean, K.A. Performance of the MISR LAI and FPAR algorithm: A case study in Africa. *Remote Sens. Environ.* **2003**, *88*, 324–340.
22. Tian, Y.; Zhang, Y.; Knyazikhin, Y.; Myneni, R.B.; Glassy, J.M.; Dedieu, G.; Running, S.W. Prototyping of MODIS LAI and FPAR algorithm with LASUR and Landsat data. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 2387–2401.
23. Knyazikhin, Y.; Martonchik, J.V.; Diner, D.J.; Myneni, R.B.; Verstraete, M.; Pinty, B.; Gobron, N. Estimation of vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MISR data. *J. Geophys. Res.* **1998**, *103*, 32239–32256.
24. Knyazikhin, Y.; Martonchik, J.V.; Myneni, R.B.; Dine, D.J.; Running, S.W. Synergistic algorithm for estimating vegetation canopy leaf area index and fraction of absorbed photosynthetically active radiation from MODIS and MISR data. *J. Geophys. Res.* **1998**, *103*, 32257–32276.
25. Friedl, M.A.; McIver, D.K.; Hodges, J.C.F.; Zhang, X.Y.; Muchoney, D.; Strahler, A.H.; Woodcock, C.E.; Gopal, S.; Schneider, A.; Cooper, A. Global land cover mapping from MODIS: Algorithms and early results. *Remote Sens. Environ.* **2002**, *83*, 287–302.

26. Kaptue Tchente, A.T.; Roujean, J.-L.; Faroux, S. ECOCLIMAP-II: An ecosystem classification and land surface parameters database of Western Africa at 1 km resolution for the African Monsoon Multidisciplinary Analysis (AMMA) project. *Remote Sens. Environ.* **2010**, *114*, 961–976.
27. Pandya, M.R.; Singh, R.P.; Chaudhari, K.N.; Bairagi, G.D.; Sharma, R.; Dadhwal, V.K.; Parihar, J.S. Leaf area index retrieval using IRS LISS-III sensor data and validation of the MODIS LAI product over central India. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 1858–1865.
28. Tan, B.; Hu, J.; Zhang, P.; Huang, D.; Shabanov, N.; Weiss, M.; Knyazikhin, Y.; Myneni, R.B. Validation of MODerate Resolution Imaging Spectroradiometer leaf area index product in croplands of Alpillis, France. *J. Geophys. Res.* **2005**, *110*, doi: 10.1029/2004JD004860.
29. Yang, P.; Shibasaki, R.; Wu, W.; Zhou, Q.; Chen, Z.; Zha, Y.; Shi, Y.; Tang, H. Evaluation of MODIS land cover and LAI products in cropland of North China Plain using *in situ* measurements and Landsat TM images. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3087–3097.
30. Lotsch, A.; Tian, Y.; Friedl, M.A.; Myneni, R.B. Land cover mapping in support of LAI and FPAR retrievals from EOS-MODIS and MISR: classification methods and sensitivities to errors. *Int. J. Remote Sens.* **2003**, *24*, 1997–2016.
31. Tian, Y.; Woodcock, C.E.; Wang, Y.; Privette, J.L.; Shabanov, N.V.; Zhou, L.; Zhang, Y.; Buermann, W.; Dong, J.; Veikkanen, B.; *et al.* Multiscale analysis and validation of the MODIS LAI product I. Uncertainty assessment. *Remote Sens. Environ.* **2002**, *83*, 414–430.
32. Hill, M.J.; Senarath, U.; Lee, A.; Zeppel, M.; Nightingale, J.M.; Williams, R.D.J.; McVicar, T.R. Assessment of the MODIS LAI product for Australian ecosystems. *Remote Sens. Environ.* **2006**, *101*, 495–518.
33. Hanan, N.P.; Hill, M.J. *Savannas in a Changing Earth System: The NASA Terrestrial Ecology Tree-Grass Project*; 2011; p. 53. Available online: http://cce.nasa.gov/terrestrial_ecology/scoping.html (accessed on 1 March 2012).
34. Myneni, R.; Knyazikhin, Y.; Shabanov, N. Leaf Area Index and Fraction of Absorbed PAR Products from Terra and Aqua MODIS Sensors: Analysis, Validation, and Refinement. In *Land Remote Sensing and Global Environmental Change: NASA's Earth Observing System and the Science of ASTER and MODIS*, 1st ed.; Ramachandran, B., Justice, C.O., Abrams, M.J., Eds.; Springer: New York, NY, USA, 2011; pp. 603–633.
35. Bach, H.; Mauser, W. Methods and examples for remote sensing data assimilation in land surface process modeling. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1629–1637.
36. Sellers, P.J.; Mintz, Y.; Sud, Y.C.; Dalcher, A. A simple biosphere model (SiB) for use within general circulation models. *J. Atmos. Sci.* **1986**, *43*, 505–531.
37. Bonan, G.B.; Levis, S.; Kergoat, L.; Oleson, K.W. Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. *Glob. Biogeochem. Cy.* **2002**, *16*, 1–25.
38. El Maayar, M.; Chen, J.M. Spatial scaling of evapotranspiration as affected by heterogeneities in vegetation, topography, and soil texture. *Remote Sens. Environ.* **2006**, *102*, 33–51.