Hindawi Publishing Corporation
Dataset Papers in Science
Volume 2015, Article ID 564279, 11 pages
http://dx.doi.org/10.1155/2015/564279

Dataset Paper

Mapping Biophysical Parameters for Land Surface Modeling over the Continental US Using MODIS and Landsat

Lahouari Bounoua,[1] Ping Zhang,[1,2] Kurtis Thome,[1] Jeffrey Masek,[1] Abdelmounaime Safia,[3] Marc L. Imhoff,[4] and Robert E. Wolfe[5]

[1] *Biospheric Sciences Laboratory, NASA's Goddard Space Flight Center, Greenbelt, MD 20771, USA*
[2] *Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD 20742, USA*
[3] *Center for Research and Applications in Remote Sensing (CARTEL), Sherbrooke University, Sherbrooke, QC, Canada J1K 2R1*
[4] *Joint Global Change Research Institute, Pacific Northwest National Laboratory and the University of Maryland, College Park, MD 20740, USA*
[5] *Terrestrial Information Systems Laboratory, NASA's Goddard Space Flight Center, Greenbelt, MD 20771, USA*

Correspondence should be addressed to Lahouari Bounoua; lahouari.bounoua-1@nasa.gov

Received 8 October 2014; Revised 12 January 2015; Accepted 12 January 2015

Academic Editor: Jai Vaze

In terms of the space cities occupy, urbanization appears as a minor land transformation. However, it permanently modifies land's ecological functions, altering its carbon, energy, and water fluxes. It is therefore necessary to develop a land cover characterization at fine spatial and temporal scales to capture urbanization's effects on surface fluxes. We develop a series of biophysical vegetation parameters such as the fraction of photosynthetically active radiation, leaf area index, vegetation greenness fraction, and roughness length over the continental US using MODIS and Landsat products for 2001. A 13-class land cover map was developed at a climate modeling grid (CMG) merging the 500 m MODIS land cover and the 30 m impervious surface area from the National Land Cover Database. The landscape subgrid heterogeneity was preserved using fractions of each class from the 500 m and 30 m into the CMG. Biophysical parameters were computed using the 8-day composite Normalized Difference Vegetation Index produced by the North American Carbon Program. In addition to urban impact assessments, this dataset is useful for the computation of surface fluxes in land, vegetation, and urban models and is expected to be widely used in different land cover and land use change applications.

1. Introduction

Rapid urbanization is replacing large areas of vegetated land, which by some estimates represent about 3% of the Earth surface and host more than 50% of its population [1, 2]. In terms of ecological impact, urbanization is one of the most significant and long lasting forms of land transformation and its extent of increase is at least proportional to population growth and economic development. Viewed from the perspective of the amount of space it currently occupies, urbanization appears to be a minor form of land transformation. However, it occupies the most fertile and productive lands [3]. The cumulative signature of this anthropogenic land cover disturbance is reaching higher proportions in some regions; for example, about 15% of the best agricultural soils in California are urbanized [4], which may already be altering the surface fluxes of carbon, water, and energy with implications for local to regional biological, hydrological, and energy cycles.

Land surface models are useful tools to understand and simulate the exchange of carbon, energy, water, and momentum between the soil, vegetation, and the atmosphere. These land surface models are generally coupled to atmospheric general circulation models (GCMs) and often operate at the same spatial resolution as the host GCM. Despite the progress in land surface models (LSMs) and better estimates of surface fluxes, previous studies describing biophysical data [5], analyzing the sensitivity of climate [6], or simulating carbon fluxes [7] were performed at spatial resolution of

$1^\circ \times 1^\circ$ (about 100 km × 100 km) or coarser and monthly time scales. It is acknowledged here that even though previous studies were constrained by the GCM's coarse resolution, enormous effort has been put into incorporating subgrid variability of biophysical parameters into LSMs (e.g., [5]); and the need for a better description of the landscape heterogeneity was the motivating driver for more accurate and finer land cover and biophysical data not previously available. These LSMs require a detailed characterization of land cover elements and a specification of a set of physical and biological parameters for each of these elements. LSMs are highly sensitive to these parameters [8]. It has been shown, for example, that land cover misclassification may result in large uncertainties in the computation of leaf area index and roughness length, both of which have significant effects on the fluxes of carbon, water, and energy at the land-atmosphere interface [9].

Most previous land surface datasets were generated using the Advanced Very High Resolution Radiometer (AVHRR) in the late 1990s and early 2000s [5, 10–12] at spatial resolution of ($1^\circ \times 1^\circ$) and temporal resolution of 1 month. Furthermore, these previous datasets used the dominant land cover type in a grid to characterize the entire grid and as such inherently carried an unquantifiable error in the specification of the land cover type and the description of its seasonal phenology which percolates into the land surface model to affect the calculations of the surface fluxes [9]. MODIS and Landsat provide a useful combination and more capabilities for more accurate and consistent estimates of these canopy parameters for a finer-scale ($0.05^\circ \times 0.05^\circ$) continental land surface modeling. In this paper we describe production details and structure of a continental scale gridded land cover dataset including fractions from 13 land cover classes obtained from higher resolution MODIS and Landsat products. For each of these classes, we also provide time series of commonly used biophysical parameters at 8-day time interval over the continental US for the year 2001 obtained using algorithms developed for the Simple Biosphere model (SiB2) [13]. Although the grid is an equal angle grid of $0.05^\circ \times 0.05^\circ$ latitude/longitude, for simplicity, in the remainder of this paper, we will refer to the grid resolution as 5 km × 5 km.

The dataset described in this paper has been produced for simulation of surface state variables with the Simple Biosphere model (SiB2) land-vegetation model and will be used to explore the impact of urbanization dynamics on surface climate. The dataset can also be used in a myriad of other land surface models of comparable complexity, such as the Community Land Model [14], the Biosphere-Atmosphere-Transfer Scheme (BATS) [15], Mosaic [16], Noah [17], the Variable infiltration Capacity (ViC) [18], and the Catchment model [19] and many other land surface vegetation models requiring characterization of terrestrial vegetation to represent not only the exchanges of energy, water, and momentum across the land-atmosphere interface but also how ecosystems and water resources respond to climate and atmospheric environment [20]. The dataset can also be used for ecological modeling and other land cover and land use change diagnostic studies.

2. Methodology

We combined the 30 m impervious surface area (ISA) from the National Land Cover Database (NLCD) and the 500 m MODIS collection 5 land cover (MCD12Q1) for the year 2001 and aggregated them to a common Climate Modeling Grid (CMG) at 5 km × 5 km over the contiguous US. The fractional ISA of the NLCD were retrieved using Landsat 7 Enhanced Thematic Mapper Plus and IKONOS discriminating man-made surfaces from natural or vegetated lands. The ISA product was generated by the United States Geological Survey (USGS) [21–23] and was not altered in the processing. The ISA and its distribution obtained from Landsat compare well with independently derived census data estimating urban populations in the USA [24], represent phenologically different environments [25], and are positively linked to urban warming effects in the US long-term climate record [26]. The MODIS MCD12Q1-Type 1 (IGBP, International Geosphere Biosphere Program) land cover classification uses a dominant type to produce 17 distinct cover types including build-up, ice and water classes [27]. The data fusion has been done considering Landsat data as the ground truth and was implemented as follows. (1) We aggregated the different cover types from the 500 m MCD12Q1 and obtained fractions in the 5 km × 5 km CMG. (2) We aggregated the Landsat ISA from 30 m × 30 m to 5 km × 5 km and coregistered them to the CMG. (3) We then imposed the aggregated Landsat ISA fractions as ground truth into the CMG replacing the build-up fraction of MODIS. When imposing the Landsat ISA into the CMG, differences between the ISA fraction obtained from Landsat and the build-up fraction obtained from MODIS were proportionally distributed over other nonurban land cover types coexisting in the CMG. The distributions were weighted by the fractions of the existing vegetation classes. In cases where the build-up class from MODIS was 100%, the difference was distributed over nonurban cover types imported from surrounding grids in the immediate vicinity.

The reconstructed product provides a gridded dataset in a CMG at 5 km × 5 km spatial resolution over the continental US. Each CMG can have up to 13 land cover classes (Table 1) along with fractions of each class within the CMG obtained from higher resolution MODIS and Landsat data. For each of the land cover classes in the CMG a set of 7 biophysical parameters is mapped for 46 time periods representing an 8-day composite annual cycle for the year 2001 (Figure 1).

We aggregated the 500 m gap-filled North American Carbon Program (NACP) MODIS Normalized Difference Vegetation Index (NDVI) to each CMG by averaging the 500 m NDVI values over each land cover class. In doing so, we obtain an average NDVI value for each land cover class coexisting in the CMG. The NACP MODIS NDVI is an 8-day product over North America obtained using an advanced interpolation algorithm [28]. We used Landsat images to generate NDVI over a range of urban areas having different ISA values. We find that maximum NDVI values for densely impervious pixels (>25% ISA) are less than or equal to 0.1 with NDVI values of approximately 0.05 for pixels with maximum ISA of 73% [29]. Since Landsat NDVI is not continuous in space and time, we use the sample of Landsat-derived urban

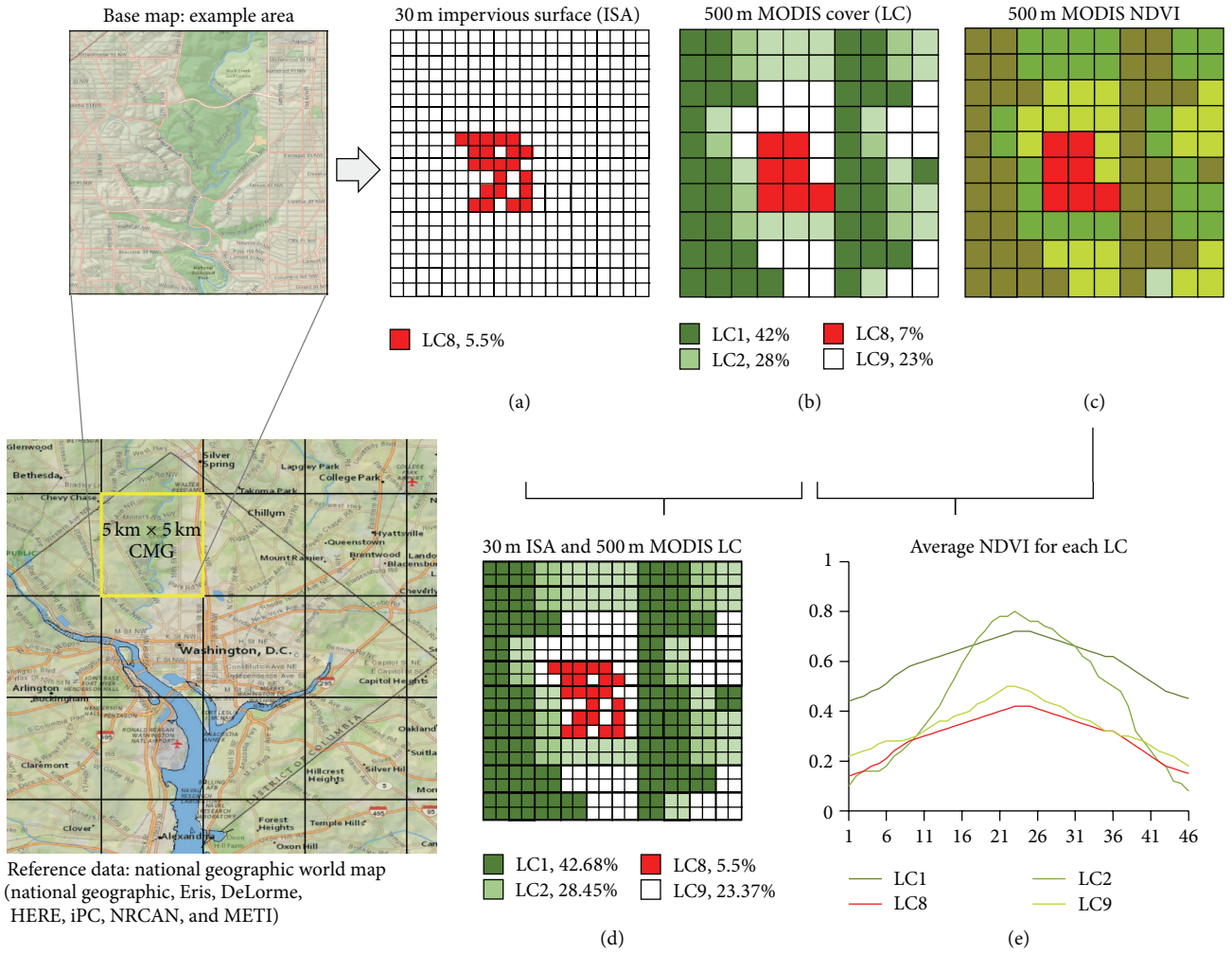

FIGURE 1: Flow chart of data fusion. First, the 30 m impervious surface area (ISA) is aggregated to the CMG (a) and then superimposed on the 500 m MODIS land cover (b) to produce a hybrid CMG containing fractions of vegetation cover from MODIS and fractions of ISA from the Landsat 30 m (d). Note the difference in the fraction of LC8 between (a) and (b) and its proportional redistribution over other classes in (d) (see text for details). Second, we combine the 500 m MODIS NDVI (c) with the 500 m MODIS land cover (b) to produce the average NDVI for each land cover class in the CMG (e). NDVI over the urban areas was scaled using Landsat retrieved NDVI values (see text for details).

NDVI values as a factor to scale down the MODIS NDVI over urban areas based on the ISA fractions. We found the NDVI values over urban classes to be less than 0.1 at the CMG scale but maintained them nevertheless to represent the phenology of the small fractions of vegetation in urban setting.

Using the aggregated average NDVI values, a set of biophysical and physical parameters were then computed for each cover type coexisting in the CMG for 2001 (Table 2). This set includes the following variables: the fraction of photosynthetically active radiation (FPAR) absorbed by the canopy, the leaf area index (LAI), the canopy greenness fraction (G), the canopy roughness length (Z_0), the canopy zero-plane displacement (D), a bulk boundary-layer resistance coefficient ($C1$), and the ground to canopy air space resistance coefficient ($C2$) for different land cover types. It is to be noted that biophysical parameters do not exist over the inland water class, and because the tundra class does not exist in the selected domain it also does not have biophysical parameters associated with it.

These biophysical parameters are computed directly from the NDVI and the land cover classes derived from MODIS using algorithms derived and explained in [30]. The computation of FPAR follows the procedure in which linearity is assumed between the NDVI and FPAR. Because the relationships between NDVI, FPAR, and the nonstressed canopy assimilation and conductance rates are quasi-linear, they should be largely scale invariant. Spatial averages of NDVI can therefore be used to compute spatial integrals of canopy assimilation and conductance rates and this holds true whether the canopy varies in depth, in cover fraction, or in both. While these relationships are fairly conservative, the relationship between NDVI and LAI may vary widely between vegetation classes [30]. LAI is nonlinearly related to FPAR. Its computation differs for broadleaf and needleleaf vegetations and

Table 1: Land cover codes and nomenclature (adapted from [11]). Fill value for all classes is 14.

Land cover	Code	Fraction	Name
LC0	0	1.31	Inland water
LC1	1	0.26	Evergreen broadleaf
LC2	2	6.67	Deciduous broadleaf
LC3	3	6.43	Mixed forest (broadleaf and needleleaf)
LC4	4	5.10	Evergreen needleleaf
LC5	5	0.03	Deciduous needleleaf
LC6	6	7.64	Savannas
LC7	7	27.91	Grassland
LC8	8	1.03	Urban and buildup
LC9	9	10.07	Shrubs with bare soil
LC10	10	0	Tundra
LC11	11	1.42	Barren
LC12	12	32.12	Croplands

uses a linear mixing model when both types are present [30]. FPAR is used directly in the computation of photosynthesis which affects the stomatal conductance and the fluxes of carbon and water. LAI is used in many components of the energy and water balances such as albedo, transpiration, and conductance. The greenness fraction (G) modulates the LAI between time periods and alters the fraction of green versus dead vegetation. Z_0, D, $C1$, and $C2$ are used in the aerodynamic computation of turbulent fluxes in LSMs and are estimated from functional relationships based on satellite-derived parameters, land cover dependent properties, and standard bulk aerodynamic theory [13, 29]. We used the SiB2 preprocessor code (MAPPER) to map and georeference the biophysical parameters onto each land cover type.

It is a challenge to assess or validate the data produced in this paper as large scale observations of biophysical parameters such as LAI, FPAR, greenness fraction, and roughness length are uncommon and nearly absent. LAI is the major biophysical parameter determining the amount of light intercepted by the canopy for photosynthesis. Under optimum growth conditions, its maximum value for a closed canopy is related to the ability of the lower leaves in the canopy to intercept enough light to sustain a positive carbon balance. LAI is the primary biophysical parameter from which FPAR, G, Z_0, and other parameters can be derived. It is greatly variable with species and canopy structure, and values estimated for large sample areas such as model grids include the average of a range of point values often comprising different species, canopy structure, and bare ground. As such, large areas LAI values may display lower maxima and lower variance than point measurements. For example, the distribution of approximately 1000 historical estimates of LAI, summarized by biome type, indicates that, for the temperate deciduous broadleaf forest (LC2), the average LAI peak over 130 records is about $5.0\,m^2{\cdot}m^{-2}$ with a standard deviation sdev = $1.80\,m^2{\cdot}m^{-2}$, whereas for croplands the historical average LAI estimated over more than 100 records is about $4.25\,m^2{\cdot}m^{-2}$ with sdev = $3.1\,m^2{\cdot}m^{-2}$ [31].

To put our estimated LAI in perspective and given the limitations, we compare it with that obtained directly from the MODIS algorithm over the $5 \times 5\,km^2$ CMG. These two model derived LAI estimates are then compared to historical LAI measurements developed by [31] and compiled for the eastern United States by [32]. It is important to recognize the difference in approach used to estimate our LAI and MODIS LAI. While MODIS LAI considers the heterogeneity of canopy structure and NDVI in the CMG, our method averages the NDVI for each land cover class within the CMG and computes the LAI for each class assumed as standing alone or "pure." This difference in approach makes our LAI estimates appear slightly higher than those obtained from MODIS, especially outside the growing season where bare ground can be predominant in the sample grid.

Figure 2 shows the annual cycle of SiB2 LAI, averaged over the US, with that obtained from MODIS for the major forest classes, the urban (LC8) and cropland (LC12) classes. For the evergreen broadleaf class (LC1), the SiB2 estimated LAI shows little variation over the year with a peak value, of $5\,m^2{\cdot}m^{-2}$, within the observed historical average range of 4.8 to $5.8\,m^2{\cdot}m^{-2}$. For this class, MODIS values approach $4\,m^2{\cdot}m^{-2}$ during the growing season but drop to values around $2\,m^2{\cdot}m^{-2}$ during winter. For LC2, while the peak LAI for both estimates is within the standard deviation obtained from the historical average (3.2 to $6.8\,m^2{\cdot}m^{-2}$) of [31], the SiB2 LAI estimate is closer to the observed historical average of $5.0\,m^2{\cdot}m^{-2}$. Outside the growing season, the SiB2 estimated LAI value is slightly higher than that obtained using the MODIS algorithm. This was expected as MODIS estimate represents the LAI over the entire grid whereas SiB2 LAI represents the LAI of the specific land cover type within the grid. Similar results are obtained for the evergreen needleleaf forest (LC4) with an average peak LAI estimated at $5\,m^2{\cdot}m^{-2}$ for SiB2 versus a peak LAI from MODIS slightly less than $3\,m^2{\cdot}m^{-2}$, with winter differences approaching $1.5\,m^2{\cdot}m^{-2}$. For the mixed forest (LC3), containing broadleaf and needleleaf trees, SiB2 estimated peak LAI is around $6\,m^2{\cdot}m^{-2}$ while MODIS estimate peak is around $4\,m^2{\cdot}m^{-2}$ with even larger difference outside of the growing season when bare soil is more apparent under deciduous trees and occupies a larger fraction of the CMG. This effect is more evident in the average LAI annual cycle of the urban class (LC8) where the SiB2 estimate is based on "pure" impervious surface whereas MODIS estimate includes a mosaic of urban lands and vegetated areas contained within the coarser 500 m × 500 m grid. In general, the difference between our estimates and those from MODIS is smaller for short vegetations. To further narrow down the assessment, we compare both LAI estimates to LAI measurements reported by biome type over the eastern US [32]. Table 3 shows the peak LAI values from SiB2, MODIS, and estimates from historical measurements averaged over the eastern US, as well as the LAI range defined by the difference between the average peak and the average value outside the growing season. Except for shrubs with bare soil (LC9), the difference between the average LAI peaks estimated by SiB2 and the observed LAI is much smaller than $1\,m^2{\cdot}m^{-2}$. For land cover type 9, the SiB2 approach

Table 2: Biophysical parameters in the data file for each land cover class.

Order	Name	Definition	Unit	Fill value
1	FPAR	Fraction of photosynthetically active radiation	Unitless	−999
2	LAI	Total leaf area index	$m^2 \cdot m^{-2}$	−999
3	G	Canopy greenness fraction	Unitless	−999
4	Z_0	Canopy roughness length	Meter (m)	−999
5	D	Canopy zero-plane displacement	Meter (m)	−999
6	$C1$	Bulk boundary-layer resistance	$(m\,s^{-1})^{1/2}$	−999
7	$C2$	Ground to canopy air-space resistance	Unitless	−999

Table 3: Average peak LAI and LAI range for the Eastern US estimated from MODIS algorithm, SiB2, and observations from Steyaert and Knox (S&K) (2008) [32]. Roughness length estimates are from SiB2, S&K (2008) [32], and Borak et al. (2005) [34]. See text for details.

Land cover	Peak LAI			LAI range			Roughness		
	MODIS	SiB2	S&K 2008 [32]	MODIS	SiB2	S&K 2008 [32]	SiB2	S&K 2008 [32]	Borak et al. 2005 [34]*
LC1	3.28	5.04	5.00	1.92	0.20	1.00	2.80	2.30	NA
LC2	3.52	5.05	5.00	3.14	3.95	4.00	0.90	1.50	1.60
LC3	3.50	6.20	5.50	2.83	3.49	2.50	1.20	1.20	1.60
LC4	2.09	5.17	5.50	1.55	2.70	1.00	1.20	1.35	1.36
LC5	1.29	5.40	NA	1.23	4.25	NA	1.20	NA	NA
LC6	2.46	3.46	3.13	1.88	1.96	2.13	0.10	0.60	0.09
LC7	1.36	2.56	2.00	1.07	1.89	1.50	0.10	0.02	0.06
LC8	0.71	0.07	0.00	0.63	0.02	0.00	2.50	2.00	NA
LC9	1.10	2.66	0.70	0.88	1.91	0.60	0.07	0.05	0.08
LC11	0.35	0.38	0.70	0.27	0.25	0.60	0.08	NA	NA
LC12	2.66	4.07	5.00	2.33	3.24	4.50	0.10	0.06	0.08

*Values estimated using the normalized mean Z_0/h developed by Borak et al. (2005) [32] and the height allocated to similar biomes in SiB2.

overestimates the peak LAI by about $2\,m^2 \cdot m^{-2}$ compared to observations, while difference between MODIS estimated LAI and observation is only $0.4\,m^2 \cdot m^{-2}$. Overall peak LAI values obtained in this study are closer to observations than those estimated using the MODIS algorithm, especially for forests, although both estimates remain within the observed variance. Differences in range of up to $1.7\,m^2 \cdot m^{-2}$ are noted between LAI estimated using the SiB2 algorithm and observations for evergreen needleleaf forest (LC4). For cropland (LC12), this difference is about $1.3\,m^2 \cdot m^{-2}$ for SiB2 and about $2.2\,m^2 \cdot m^{-2}$ for MODIS (Table 3).

Overall there is a fair agreement between the LAI estimates from SiB2 and those obtained from the MODIS algorithm, and both estimates are within the variance of historical values obtained as average over large number of site observations. It is well recognized that misclassification of the land cover, which is an input to the LAI algorithm in both SiB2 and MODIS, is one of the major sources of errors in estimating LAI [33].

As stated earlier, the dataset presented here has the advantage of being continuous in time and space with no missing data and is therefore suitable for land surface and ecological modeling. The discrepancy in LAI outside the growing season has minimal effect in modeling the carbon cycle as the latter is additionally constrained by cold temperatures in most LSMs.

Virtually all recent LSMs require knowledge of the aerodynamic roughness length to compute the turbulent exchanges of momentum, energy, water, and carbon. The turbulent exchanges between the surface and the atmosphere determine not only the local scale momentum transfer but also the transport of sensible and latent heat away from the surface. Two fundamental momentum aerodynamic variables are of importance for the turbulent exchanges: the roughness length, Z_0, and the zero-plane displacement, D. Z_0 is defined as the height above the surface at which the mean logarithmic wind profile theoretically reaches zero and D is the height above the land surface to which roughness elements (e.g., vegetation) have effectively displaced the momentum-absorbing properties of the surface [34]. The turbulent submodel of SiB2 calculates Z_0 and D based on static parameters specific to land cover types such as the canopy top and base heights and other biophysical parameters defining the seasonality such as LAI, cover fraction, and leaf dimension. Over the normal range of LAI values, SiB2 estimates of Z_0 show a weak exponential dependence on LAI [30].

Most published values of Z_0 are model derived and actual observations at the scale of the CMG are virtually inexistent. Again to put the values of Z_0 estimated using the SiB2 aerodynamic submodel into context, we compare them to those of [32] averaged over the Eastern US for canopies with comparable heights and morphological properties. For example, for the deciduous broadleaf forest (LC2), the SiB2 estimated Z_0 is 0.90 m compared to the reported average value of 1.50 m. On the other hand, the SiB2 evergreen

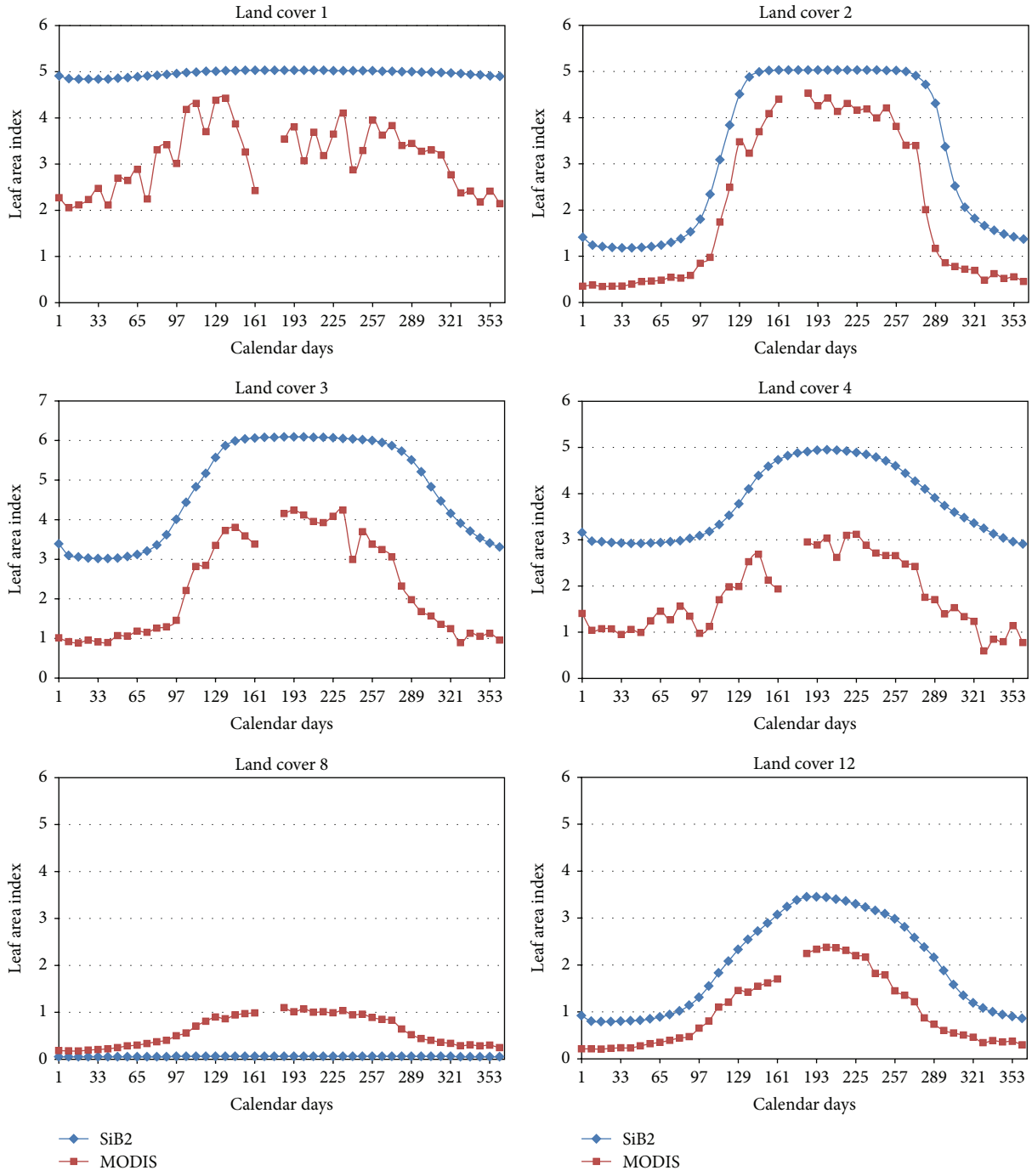

FIGURE 2: Leaf area index estimates from SiB2 and MODIS averaged over the continental US for major forests (LC1–LC4), the urban class (LC8), and cropland class (LC12). Note that the MODIS estimates are obtained from LAI values representing an average over an entire 500 m × 500 m pixel, whereas SiB2 estimates are obtained for a specific class (see text for details). Different land cover classes (LCs) are described in Table 1.

needleleaf forest (LC4) Z_0 estimate of 1.20 m is remarkably close to the average 1.35 m value reported in [32]. In general, roughness length values obtained using the SiB2 algorithm are comparable to those reported in the literature for the eastern US and the differences do not exceed 0.5 m for any land cover type. These estimates are also comparable to those obtained by another approach that used wind-tunnel and field data to fit the mathematical formulation of the normalized roughness length Z_0/h, with h being the height of the roughness element [34]. Canopy top heights of SiB2 land cover classes [30] are used in the model of [34] to provide a second approach to roughness estimates (Table 3). This assessment indicates that roughness estimates from SiB2 and those estimated by [32, 34] are in close agreement.

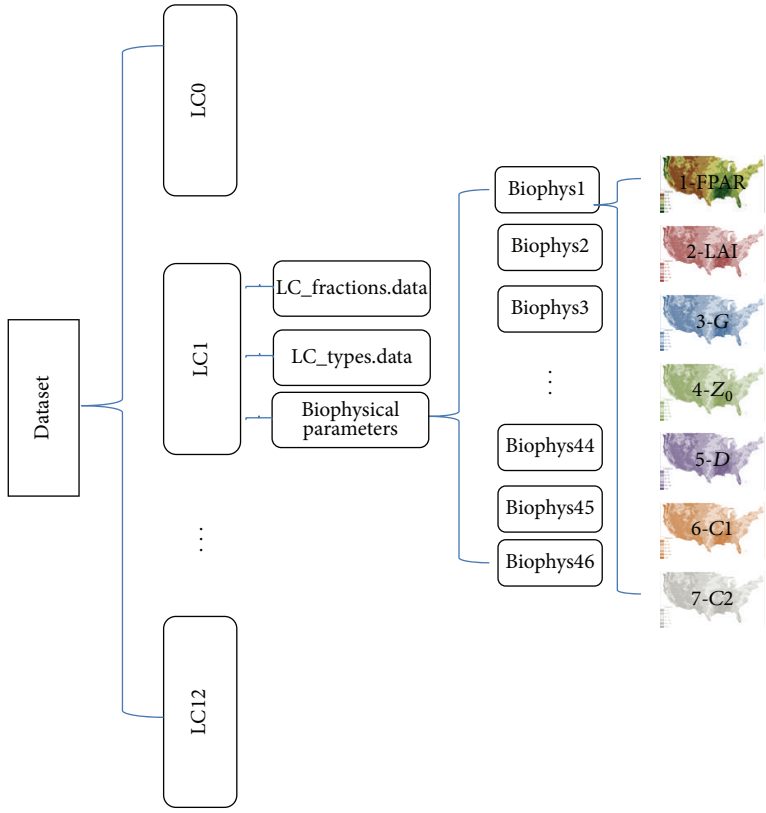

Figure 3: Dataset structure.

3. Dataset Description

The dataset associated with this Dataset Paper consists of 4 items which are described as follows.

Dataset Item 1 (Binary Data). It shows 13 different land cover classes, classes 0 to 12 (Table 1 and Figure 3). Each class contains two files representing the land cover type (LC<type>_types.data) and the land cover type fraction (LC<type>_fractions.data), where "<type>" is the land cover type (0 to 12), and 46 files, for each of the time periods, labeled "biophys<period>," namely, biophys1 to biophys46, where <period> is the 8-day time period varying from 1 to 46, each containing 7 biophysical parameters for that land cover class. For example, to obtain the biophysical parameters for the Deciduous Broadleaf class (LC2) for the 18th time period in 2001, access the LC2_biophys18 file. As an example, Table 4 illustrates the land cover class, the fraction (%), and the LAI ($m^2 \cdot m^{-2}$) of all 10 classes coexisting in the CMG defined by (row = 208, column = 992) for period 18. To determine the latitude and longitude from the row/column of a point, one can use the following relationship: Latitude = NL − (row-1) ∗ gridsize and Longitude = WL + (column-1) ∗ gridsize, where the NL is the northernmost latitude (49.475°), WL is the westernmost longitude (125.025°), and the gridsize is equal to 0.05°. Using these calculations, the point (row = 208, column = 992) should be at latitude 39.125° north and longitude 75.475° west. For each class, the LC<type>_types.data file shows the code of that particular cover type in the CMG (Table 1) wherever the class exists with a fraction greater than 1% and fill value (−999.0) elsewhere. The threshold fraction of 1% is chosen to build the land cover types and is supposed to represent the error in classification emanating from the MCD12Q1. Users are encouraged to choose their own threshold fraction to build land cover type masks based on the land cover fractions. The LC<type>_fractions.data file contains the coregistered fraction of the particular land cover type for all pixels (Figure 4). Each data file contains 1160 columns and 490 rows. The land cover types and fractions are floating point values from 0 to 12 for the types and from 0 to 100% for the fractions. Though we have included "tundra" as a land cover class, there is no tundra recorded in the contiguous US and its land cover fraction is not provided. For portability and general access, the format storing data using the American Standard Code for Information Interchange (ASCII) is provided for scientists who prefer to write their own code to manipulate the information. Data files are 2-dimensional array with the longitude varying from west (125.025 W) to east (67.075 W) and then the latitude varying from north (49.475 N) to south (25.025 N) with an equal grid spacing of 0.05 degrees. The biophysical parameters are written in 46 data files corresponding to 46 8-day periods over which MODIS data have been composited with period 1 corresponding to day 1 to day 8 of the calendar year and period 2 corresponding to day 9 to day 16, and so on. Each file contains seven biophysical parameters for that 8-day period written as seven appended layers in the order shown in Table 2. Similar to the land cover types and fractions, the

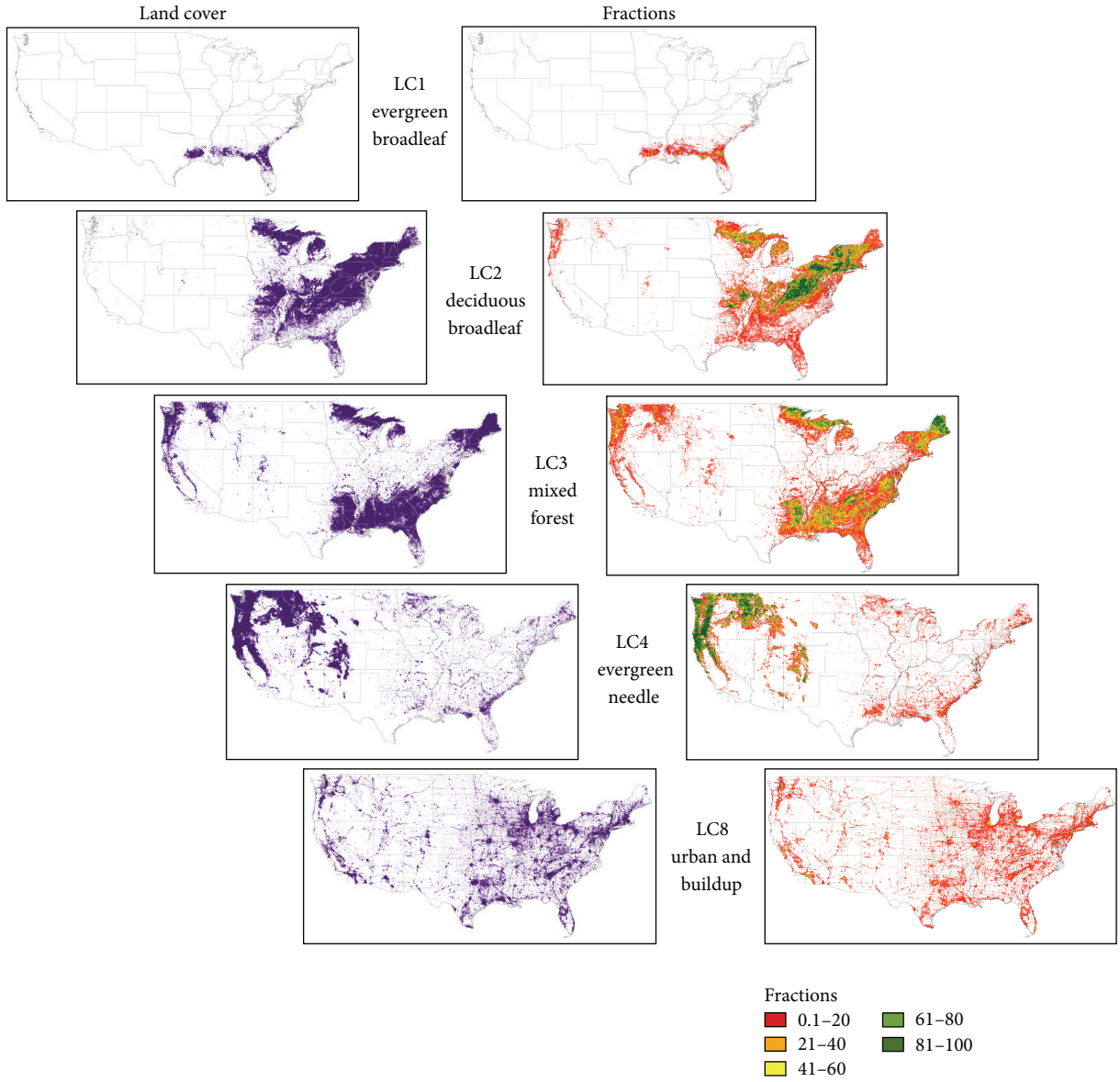

Figure 4: Geographical distribution of selected land cover classes (left) and corresponding fractions in percent (right) over the contiguous US.

biophysical parameters are also written in ASCII format. The biophysical data files are presented as a 3-dimensional array with the longitude varying first, then the latitude, and then the biophysical parameters changing in the sequence shown in Table 2. Note that biophysical parameters are not defined over inland water (class 0) and tundra (class 10). Figure 5 shows an example of the temporal profile of the seven biophysical parameters associated with cropland in one grid cell.

Dataset Item 2 (Source Code). A simple code (test_read_biophysical_ascii.f) written in standard FORTRAN is also included with the dataset to help read the provided ASCII data files.

Dataset Item 3 (Binary Data). For portability and general access, we also provided the same dataset in binary form which is readable by some remote sensing and geographical information systems (GIS) software such as ENVI and Arcgis. For example, users can use the parameters listed in Table 5 to read the data within this item in ENVI.

Dataset Item 4 (Source Code). We also provide a simple code (test_read_biophysical_binary.f) written in standard FORTRAN to help read the provided binary data files.

4. Concluding Remarks

This dataset was developed in response to a demand by different land surface and ecological modeling groups to provide a phase coherent, internally consistent, space and time continuous description of land surface cover elements and their biophysical characterization. The dataset is at continental

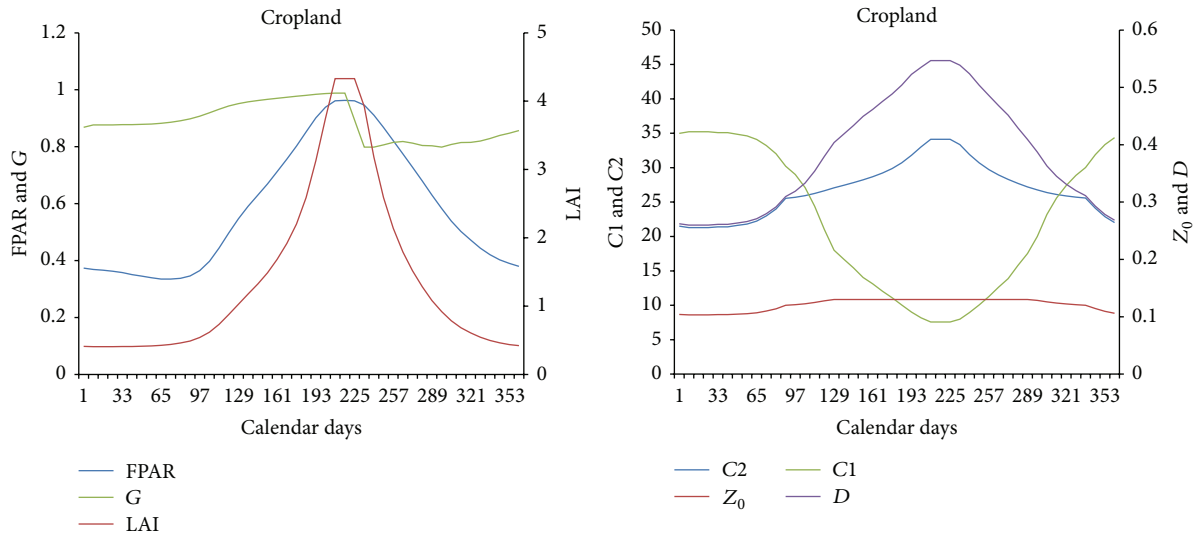

Figure 5: Example of the biophysical parameter annual cycle for a cropland.

Table 4: Land cover class, fraction (%), and LAI ($m^2 \cdot m^{-2}$) of all classes coexisting in CMG defined by row = 283, column = 992 for time period 18. The −999.0 means that the class does not exist in the CMG.

Class	Fraction	LAI
1	−999.0	−999.0
2	5.43	2.71
3	1.36	6.15
4	2.71	0.99
5	2.71	0.85
6	10.18	3.62
7	1.36	1.40
8	2.27	0.06
9	5.43	0.55
11	3.39	0.50
12	45.47	1.47

Table 5: Parameters to use when using ENVI to open Dataset item 3 (Binary data). *For all land cover type and fraction files, the band (layer) is 1. For the biophysical parameters, the band (layer) is 7.

Samples	1160
Lines	490
Bands (Layers)	1/7*
Data type	Floating point
Interleave	bsq

scale with a spatial and temporal resolution capable of capturing the physics of changes in canopy-scale physiology and urban metabolism. Here, we provide the dataset for the year 2001 and for the contiguous US at a spatial resolution of 0.05° and temporal resolution of 8 days, but we recognize the need to develop it at global scale.

In addition to providing a common platform for land surface and ecological modelers, the dataset is the first of its kind to gather all parameters required to characterize land surface cover from plot to continental-scale in one single location. The dataset is critical for the assessment of surface fluxes over land and is expected to be widely used in different applications and research related to the modeling and diagnostic studies of carbon, water, and energy fluxes, including interactions with the atmospheric boundary layer. This dataset, though with slightly different format, will be also built into the NASA Land Information System [35], which is designed to perform customized land surface data assimilation at different spatial resolutions. Not only do the LAI values presented here compare well with those obtained from the MODIS algorithm but also peak values are closer to observations. Both the SiB2 and MODIS LAI values are within the range of observed variance. Similarly, the estimated roughness length is in close agreement with observed historical averages obtained over the eastern US and with estimates obtained from a different approach found in the literature.

Dataset Availability

The dataset associated with this Dataset Paper is dedicated to the public domain using the CC0 waiver and is available at http://dx.doi.org/10.1155/2015/564279/dataset. In addition, it is available for download from the Terrestrial Information Systems Laboratory at Goddard Space Flight Center (GSFC). The data can be accessed using the file transfer protocol (ftp) at the following location: ftp://landsc1.nascom.nasa.gov/outgoing/BIOPARAM_LSM (for Firefox) or ftp://landsc1.nascom.nasa.gov/pub/outgoing/BIOPARAM_LSM (for Internet Explorer).

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was funded by NASA Interdisciplinary Research in Earth Science NNH09ZDA001N-IDS (2012). Thanks are due to Garik Gutman, program manager.

References

[1] CIESIN News Archives and The Growing Urbanization of the World, "The Growing Urbanization of the World: GRUMP mapping project finds urban areas increasing in surprising ways," Report, 2005, http://www.earth.columbia.edu/news/2005/story03-07-05.html.

[2] UNFPA (United Nations Population Fund), *State of World Population. Unleashing the Potential of Urban Growth*, UNFPA, 2007, http://www.unfpa.org/swp/2007/english/introduction.html.

[3] E. L. Nizeyimana, G. W. Petersen, M. L. Imhoff et al., "Assessing the impact of land conversion to urban use on soils with different productivity levels in the USA," *Soil Science Society of America Journal*, vol. 65, no. 2, pp. 391–402, 2001.

[4] M. L. Imhoff, W. T. Lawrence, D. C. Stutzer, and C. D. Elvidge, "A technique for using composite DMSP/OLS 'city lights' satellite data to map urban area," *Remote Sensing of Environment*, vol. 61, no. 3, pp. 361–370, 1997.

[5] S. O. Los, N. H. Pollack, M. T. Parris et al., "A global 9-yr biophysical land surface dataset from NOAA AVHRR data," *Journal of Hydrometeorology*, vol. 1, no. 2, pp. 183–199, 2000.

[6] L. Bounoua, G. J. Collatz, S. O. Los et al., "Sensitivity of climate to changes in NDVI," *Journal of Climate*, vol. 13, no. 13, pp. 2277–2292, 2000.

[7] J. G. Masek and G. J. Collatz, "Estimating forest carbon fluxes in a disturbed southeastern landscape: Integration of remote sensing, forest inventory, and biogeochemical modeling," *Journal of Geophysical Research G: Biogeosciences*, vol. 111, no. 1, Article ID G01006, 2006.

[8] G. B. Bonan, D. Pollard, and S. L. Thompson, "Influence of subgrid-scale heterogeneity in leaf area index, stomatal resistance, and soil moisture on grid-scale land-atmosphere interactions," *Journal of Climate*, vol. 6, no. 10, pp. 1882–1897, 1993.

[9] L. Bounoua, J. Masek, and Y. M. Tourre, "Sensitivity of surface climate to land surface parameters: a case study using the simple biosphere model SiB2," *Journal of Geophysical Research D: Atmospheres*, vol. 111, no. 22, Article ID D22S09, 2006.

[10] S. O. Los, C. O. Justice, and C. J. Tucker, "A global 1° by 1° NDVI data set for climate studies derived from the GIMMS continental NDVI data," *International Journal of Remote Sensing*, vol. 15, no. 17, pp. 3493–3518, 1994.

[11] P. J. Sellers, "A global 1◦ by 1◦ NDVI data set for climate studies. Part 2: the generation of global fields of terrestrial biophysical parameters from the NDVI," *International Journal of Remote Sensing*, vol. 15, no. 17, pp. 3519–3545, 1994.

[12] C. J. Tucker, J. E. Pinzon, M. E. Brown et al., "An extended AVHRR 8–km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data," *International Journal of Remote Sensing*, vol. 26, no. 20, pp. 4485–4498, 2005.

[13] P. J. Sellers, D. A. Randall, G. J. Collatz et al., "A revised land surface parameterization (SiB2) for atmospheric GCMs. Part I: model formulation," *Journal of Climate*, vol. 9, no. 4, pp. 676–705, 1996.

[14] Y. Dai, X. Zeng, R. E. Dickinson et al., "The common land model," *Bulletin of the American Meteorological Society*, vol. 84, no. 8, pp. 1013–1023, 2003.

[15] R. E. Dickinson, A. Henderson-Sellers, and P. J. Kennedy, "Biosphere atmosphere transfer scheme (BATS) version 1e as coupled to the NCAR community climate model," NCAR Technical Notes TN-387+STR, 1993.

[16] R. D. Koster and M. J. Suarez, "Modeling the land surface boundary in climate models as a composite of independent vegetation stands," *Journal of Geophysical Research*, vol. 97, no. 3, pp. 2697–2715, 1992.

[17] F. Chen, K. Mitchell, J. Schaake et al., "Modeling of land surface evaporation by four schemes and comparison with FIFE observations," *Journal of Geophysical Research D: Atmospheres*, vol. 101, no. 3, pp. 7251–7268, 1996.

[18] X. Liang, D. P. Lettenmaier, and E. F. Wood, "One-dimensional statistical dynamic representation of subgrid spatial variability of precipitation in the two-layer variable infiltration capacity model," *Journal of Geophysical Research D: Atmospheres*, vol. 101, no. 16, pp. 21403–21422, 1996.

[19] R. D. Koster, M. J. Suarez, A. Ducharne, M. Stieglitz, and P. Kumar, "A catchment-based approach to modeling land surface processes in a general circulation model 1. Model structure," *Journal of Geophysical Research D: Atmospheres*, vol. 105, no. 20, pp. 24809–24822, 2000.

[20] I. C. Prentice, X. Liang, B. E. Medlyn, and Y.-P. Wang, "Reliable, robust and realistic: the three R's of next-generation land surface modelling," *Atmospheric Chemistry and Physics Discussions*, vol. 14, no. 17, pp. 24811–24861, 2014.

[21] L. Yang, C. Huang, C. G. Homer, B. K. Wylie, and M. J. Coan, "An approach for mapping large-area impervious surfaces: synergistic use of Landsat-7 ETM+ and high spatial resolution imagery," *Canadian Journal of Remote Sensing*, vol. 29, no. 2, pp. 230–240, 2003.

[22] C. Homer, C. Huang, L. Yang, B. Wylie, and M. Coan, "Development of a 2001 national land-cover database for the United States," *Photogrammetric Engineering & Remote Sensing*, vol. 70, no. 7, pp. 829–840, 2004.

[23] J. A. Fry, G. Xian, S. Jin et al., "Completion of the 2006 national land cover database for the conterminous united states," *Photogrammetric Engineering and Remote Sensing*, vol. 77, no. 9, pp. 858–864, 2011.

[24] M. L. Imhoff, P. Zhang, R. E. Wolfe, and L. Bounoua, "Remote sensing of the urban heat island effect across biomes in the continental USA," *Remote Sensing of Environment*, vol. 114, no. 3, pp. 504–513, 2010.

[25] L. Bounoua, A. Safia, J. Masek, C. Peters-Lidard, and M. L. Imhoff, "Impact of urban growth on surface climate: a case study in Oran, Algeria," *Journal of Applied Meteorology and Climatology*, vol. 48, no. 2, pp. 217–231, 2009.

[26] J. Hansen, R. Ruedy, M. Sato et al., "A closer look at United States and global surface temperature change," *Journal of Geophysical Research D: Atmospheres*, vol. 106, no. 20, pp. 23947–23963, 2001.

[27] M. C. Hansen, R. S. Defries, J. R. G. Townshend, and R. Sohlberg, "Global land cover classification at 1 km spatial resolution using a classification tree approach," *International Journal of Remote Sensing*, vol. 21, no. 6-7, pp. 1331–1364, 2000.

[28] B. Tan, J. T. Morisette, R. E. Wolfe et al., "An enhanced TIMESAT algorithm for estimating vegetation phenology metrics from MODIS data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 4, no. 2, pp. 361–371, 2011.

[29] J. Masek, *Earth Day-2009*, NASA Goddard Space Flight Center, Personal Communication, 2009.

[30] P. J. Sellers, C. J. Tucker, G. J. Collatz et al., "A revised land surface parameterization (SiB2) for atmospheric GCMs. Part II: the generation of global fields of terrestrial biophysical parameters from satellite data," *Journal of Climate*, vol. 9, no. 4, pp. 706–737, 1996.

[31] UT-BAll-ELLE, "Worldwide historical estimates of leaf area index, 1932–2000," ORNL 27(4-00), Oak Ridge National Laboratory for the U.S. Department of Energy, 2001.

[32] L. T. Steyaert and R. G. Knox, "Reconstructed historical land cover and biophysical parameters for studies of land-atmosphere interactions within the eastern United States," *Journal of Geophysical Research D: Atmospheres*, vol. 113, no. 2, Article ID D02101, 2008.

[33] J. T. Morisette, F. Baret, J. L. Privette et al., "Validation of global moderate-resolution LAI products: a framework proposed within the CEOS land product validation subgroup," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 7, pp. 1804–1817, 2006.

[34] J. S. Borak, M. F. Jasinski, and R. D. Crago, "Time series vegetation aerodynamic roughness fields estimated from modis observations," *Agricultural and Forest Meteorology*, vol. 135, no. 1, pp. 252–268, 2005.

[35] S. V. Kumar, C. D. Peters-Lidard, Y. Tian et al., "Land information system: an interoperable framework for high resolution land surface modeling," *Environmental Modelling and Software*, vol. 21, no. 10, pp. 1402–1415, 2006.