Global Land Surface Fractional Vegetation Cover Estimation Using General Regression Neural Networks From MODIS Surface Reflectance

Kun Jia, Shunlin Liang, *Fellow, IEEE*, Suhong Liu, Yuwei Li, Zhiqiang Xiao, Yunjun Yao, Bo Jiang, Xiang Zhao, Xiaoxia Wang, Shuai Xu, and Jiao Cui

***Abstract*—Fractional vegetation cover (FVC) plays an important role in earth surface process simulations, climate modeling, and global change studies. Several global FVC products have been generated using medium spatial resolution satellite data. However, the validation results indicate inconsistencies, as well as spatial and temporal discontinuities of the current FVC products. The objective of this paper is to develop a reliable estimation algorithm to operationally produce a high-quality global FVC product from the Moderate Resolution Imaging Spectroradiometer (MODIS) surface reflectance. The high-spatial-resolution FVC data were first generated using Landsat TM/ETM+ data at the global sampling locations, and then, the general regression neural networks (GRNNs) were trained using the high-spatial-resolution FVC data and the reprocessed MODIS surface reflectance data. The direct validation using ground reference data from validation of land European Remote Sensing instruments sites indicated that the performance of the proposed method ($\mathbf{R^2 = 0.809}$, $\mathbf{RMSE = 0.157}$) was comparable with that of the GEOV1 FVC product ($\mathbf{R^2 = 0.775}$, $\mathbf{RMSE = 0.166}$), which is currently considered to be the best global FVC product from SPOT VEGETATION data. Further comparison indicated that the spatial and temporal continuity of the estimates from the proposed method was superior to that of the GEOV1 FVC product.**

***Index Terms*—Estimation, fractional vegetation cover (FVC), general regression neural networks (GRNNs), global, Moderate Resolution Imaging Spectroradiometer (MODIS).**

I. Introduction

VEGETATION is the basic component of the terrestrial ecosystem and plays an important role in energy exchange and biogeochemical and hydrological cycling processes on the earth surface [1], [2]. Fractional vegetation cover (FVC), which refers to the fraction of green vegetation as seen from the nadir of the total statistical area, is an important parameter in surface process models for numerical weather prediction, regional and global climate modeling, and global change studies [1], [3]–[5]. Accurate estimation of FVC on the regional and global scales are required for land surface processes and climate change studies as well as for its extensive application in agriculture, forestry, environment management, disaster risk monitoring, and drought monitoring [6]–[9].

Remote sensing provides the only feasible way to generate FVC products at the regional and global scales because of its ability to quickly provide broad, impartial and easily available data on the land surface [7], [10]–[12]. There are three main types of methods for estimating FVC from remote sensing data: the empirical methods; pixel unmixing model; and physical methods. The empirical methods are based on statistical relationships between the FVC and vegetation indices or specific spectral bands [13], [14]. Generally, the normalized difference vegetation index (NDVI) derived from the reflectance of the red and near-infrared (NIR) bands is the most frequently used index for regression models development of FVC estimation. The empirical methods are computationally efficient in operating with large amounts of data and widely used in FVC estimation on a regional scale. However, the empirical methods are highly dependent on the specific vegetation types in specific regions; thus, the expanded application of these methods at a large scale may be invalid.

The pixel unmixing model assumes that each pixel is composed of several components and estimates FVC at a subpixel level. The linear pixel unmixing model is the commonly used method, which decomposes each pixel into a linear component of a reference spectrum and considers the proportion of the vegetation component as FVC [12], [13], [15], [16]. Among the various linear pixel unmixing models, the simplest and most widely used model is the dimidiate pixel model, which is based on the assumption that pixels are composed of two parts: vegetation and nonvegetation [6], [17]. Landsat data have been used to estimate FVC with the dimidiate pixel model and achieved reliable results at the region scale [18]. However, it is difficult to determine the endmembers and the spectral of the endmembers at the global scale for FVC estimation because the land surface is complex and the spectral characteristics of objects are varied.

Manuscript received June 17, 2014; revised December 4, 2014; accepted February 16, 2015. This work was supported in part by the National High Technology Research and Development Program of China under Grant 2013AA122801, by the National Natural Science Foundation of China under Grant 41301353 and Grant 41331173, by the Special Foundation for Free Exploration of the State Key Laboratory of Remote Sensing Science under Grant 14ZY-06, and by the Fundamental Research Funds for the Central Universities.

K. Jia, S. Liu, Y. Li, Z. Xiao, Y. Yao, B. Jiang, X. Zhao, X. Wang, S. Xu, and J. Cui are with the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing 100875, China (e-mail: jiakun@bnu.edu.cn).

S. Liang is with the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing 100875, China, and also with the Department of Geographical Sciences, University of Maryland, College Park, MD 20742 USA (e-mail: sliang@umd.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2015.2409563

The physical methods are based on the inversion of canopy radiative transfer models (RTMs), which contain the physical relationships between the FVC and the spectral reflectance of the vegetation canopy. Because physical methods can be adjusted for a wide range of situations, they are more widely applicable in the theoretical sense [19]. However, due to the complexity of physical models, direct inversion is difficult and artificial neural networks (ANNs), which are based on a precomputed reflectance database, are the most popular inversion technique [8], [20]. ANNs are well known for their good performance in land surface parameter estimation using remote sensing data [21] and are very efficient from a computational point of view, particularly for operational applications with long time series of global data. Several global and regional FVC products have been generated using ANNs and RTM, such as the POLDER FVC product, which uses the ANNs and the Kuusk model [8]; the MERIS and CYCLOPES FVC products, which use ANNs and the PROSPECT + SAIL model [22], [23]; and the GEOV1 FVC product, which is an improvement of CYCLOPES [3].

ANNs are the reasonable choice for the operational global FVC estimation using remote sensing data. However, due to the complex radiative transfer mechanism and the limitation of the capability of current physical models, the validation accuracy of global FVC products is not satisfactory [23], [24]. The GEOV1 FVC product corrected the systematic underestimation of the CYCLOPES FVC product using a scaling factor and then trained the ANNs; the validation accuracy was improved [3]. Therefore, continued developing high-accuracy global FVC estimation algorithm with remote sensing data is required for providing a complete and reliable FVC product.

This paper aims to develop a reliable estimation algorithm to operationally produce a high-quality global FVC product from MODIS surface reflectance data. This method attempts to generate finer spatial resolution FVC samples using Landsat TM/ETM+ data, which are different because current samples are usually generated by model simulation. Then, the samples are used for training general regression neural networks (GRNNs) with reprocessed MODIS surface reflectance, which have good quality and spatiotemporal continuity. The trained GRNNs will be used to estimate the time series global FVC using the reprocessed MODIS data.

II. Algorithm Development

This paper employed GRNNs to estimate global land surface FVC from MODIS surface reflectance data with the training samples generated from global sampled Landsat TM/ETM+ data (see Fig. 1). First, global sampling locations were selected based mainly on the Benchmark Land Multisite Analysis and Intercomparison of Products (BELMANIP) sites and high-quality Landsat TM/ETM+ data at each location were selected for generating the finer spatial resolution FVC samples. Then, the Landsat data were preprocessed and used to estimate FVC using the dimidiate pixel model with the help of the terrestrial ecoregions of the world and 30-m finer resolution global land cover data. The MODIS reflectance data were reprocessed to remove the remaining effects of cloud contamination and other factors. Training samples were extracted using Landsat FVC and reprocessed MODIS reflectance data at each global sampling location. Finally, the GRNNs were trained and used to estimate global land surface FVC. Detailed descriptions of the method are given in the following subsections.

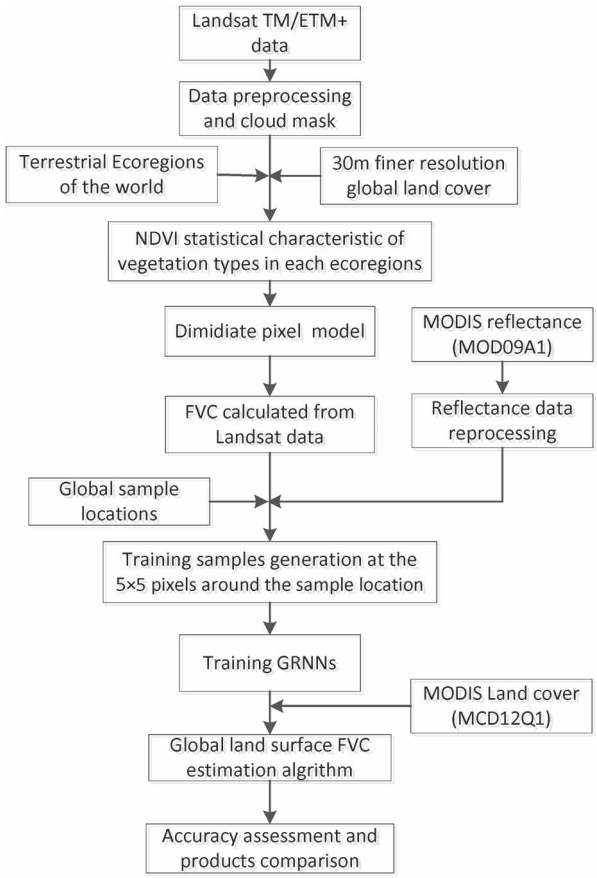


Fig. 1. Flowchart of the global land surface FVC estimation method.

A. MODIS Products and Preprocess

MODIS reflectance (MOD09A1, collection 5) and land cover (MOD12Q1, collection 5) products are used for global FVC estimation algorithm development. The MOD09A1 product has provided the surface reflectance for each of the MODIS land spectral bands with 8 day temporal sampling intervals since the year 2000. The quality of the MOD09A1 product is influenced by many factors, e.g., clouds, aerosols, water vapor, and ozone. Although most of the effects have been removed through atmospheric corrections, the remaining effects can sometimes be very large and require further processing [11], [25]. In this paper, the MODIS reflectance reprocessing method developed by Tang *et al.* [26], which is used for generating the Global LAnd Surface Satellite (GLASS) products, is used [27]. The method first identifies the data contaminated by undetected and fallout clouds in the MODIS reflectance data using MODIS snow and cloud mask data, the spectral characteristics of temporal and spatial continuity and correlation as well as other auxiliary information. Then, the contaminated data are removed using temporal-spatial filtering methods, and the missing data

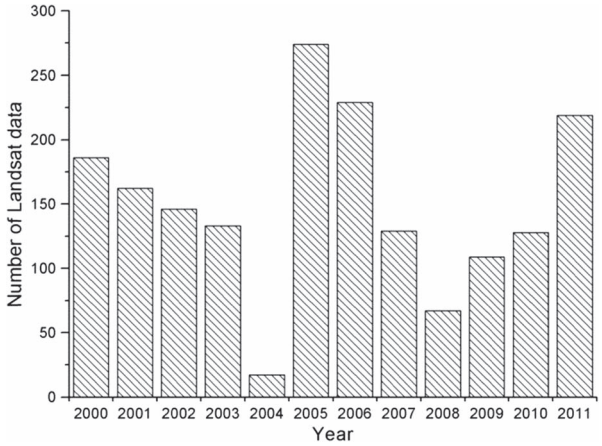


Fig. 2. Number of selected Landsat images for different years.

are filled using an optimum interpolation algorithm to obtain the final continuous and smooth surface reflectance values for earth surface parameters estimation [26].

The MOD12Q1 land cover product (layer 3) is used as the priori information to distinguish vegetation and nonvegetation regions. The FVC values in vegetation regions, which include eight main biomes (grasses and cereal crops, shrubs, broadleaf crops, savannahs, evergreen broadleaf forests, deciduous broadleaf forests, evergreen needleleaf forests, and deciduous needleleaf forests), are estimated using the GRNNs method. The FVC values in nonvegetation regions, which include three nonbiomes (unvegetated, urban, and water), are set to zero.

B. Landsat TM/ETM+ Data and Preprocess

Archived Landsat 5 TM and Landsat 7 ETM+ data were selected and downloaded from the USGS GLOVIS website (http://glovis.usgs.gov) for training sample generation. Because the scan line corrector (SLC) of the Landsat 7 ETM+ sensor failed in 2003 and resulted in approximately 22% of the pixels per scene not being scanned, the SLC-off images of Landsat 7, which referred to images acquired after the SCL failure, were not selected in this paper. Other Landsat data selection criteria included images that were not influenced by the cloud and the selection of one image in each season from each sample location from the years 2000 to 2011. Finally, 1800 scenes Landsat TM/ETM+ data were selected in this paper (see Fig. 2). The number of selected Landsat data for different years was close to a balanced distribution.

Geometric registration and atmospheric correction were critical for the application of Landsat data, as they facilitate the comparison of data through time. In this paper, only Landsat L1T data were used because of their precise registration. Atmospheric correction was performed using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) atmospheric correction tool, in which the raw digital number values were converted into surface reflectance [28]. Clouds, cloud shadows, and snow were masked using an object-based algorithm called Fmask [29].

C. Dimidiate Pixel Model

The dimidiate pixel model, which had been widely used to derive FVC [1], [12], was selected to estimate FVC using the Landsat surface reflectance data for training sample data generation. It assumed that a pixel consists of only vegetation and nonvegetation components, and its value was a linear combination of these two components. If the NDVI, which was the suitable indicator for plant growth and density of vegetation spatial distribution [1], [30], was used to represent the spectral response, the mathematical expression of the mixed pixel model would be

$$NDVI = f \times NDVI_v + (1 - f) \times NDVI_s \tag{1}$$

and then

$$f = \frac{NDVI - NDVI_s}{NDVI_v - NDVI_s} \tag{2}$$

where f was the proportion of vegetation area in the mixed pixel (FVC), $NDVI$ was the NDVI of the mixed pixel, and $NDVI_v$ and $NDVI_s$ were the NDVI of the fully vegetated and bare soil pixel, respectively.

Ideally, $NDVI_v$ and $NDVI_s$ would not change with time and space. However, the determination of the two parameters was affected by many factors, such as soil type, soil moisture, vegetation type, and chlorophyll content. Nevertheless, $NDVI_v$ and $NDVI_s$ could be determined through statistical analysis of spatial and temporal NDVI data, assuming that fully vegetated and bare soil could be found in the time and space. For example, $NDVI_s$ and $NDVI_v$ were defined as the lower and upper 2%–5% NDVI for each biome in [31]. In this paper, terrestrial ecoregions of the world [32] and 30-m finer resolution global land cover data [33] were used to determinate $NDVI_v$ and $NDVI_s$ for each biome using the statistical analysis method for FVC estimation of Landsat data, with the assumption that the soil characteristics and vegetation characteristics of each vegetation type in each ecoregion were similar and resulted in similar $NDVI_v$ and $NDVI_s$. Vegetation types were divided into forest, grass-shrub, and crop based on 30-m finer resolution global land cover in each terrestrial ecoregion.

D. Training Sample Generation

Selection of global sampling locations representing the global distribution of vegetation types and conditions was one of the most important steps for global FVC estimation method development. In this paper, the BELMANIP sites, which were located in relatively flat and homogeneous areas (at a kilometric resolution over 10×10 km^2 domains) with global representativeness and homogeneity, were selected as the basic global sampling locations [3], [20]. To prevent the sample number being insufficient, some sites in FLUXNET, which were not overlaid and were very near to the BELMANIP sites, and validation of land European Remote Sensing instruments (VALERI) sites with ground-measured FVC data were added to enrich the global sampling locations. Finally, approximately 500 global

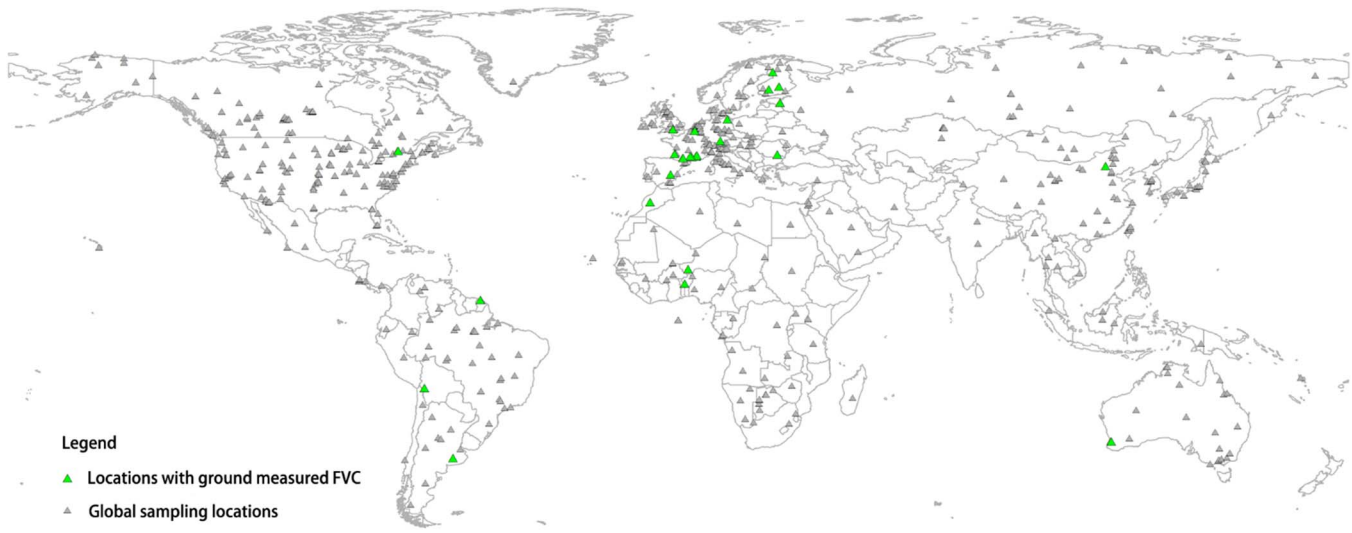


Fig. 3. Global sampling locations used to generate the global FVC estimation method.

sampling locations were selected and the spatial distribution of the locations is shown in Fig. 3.

Pixels ($5 \times 5 = 25$) in the reprocessed MODIS reflectance data around each of the global sampling locations were extracted, and then, the corresponding FVC values of the Landsat pixels compared with each extracted MODIS pixel were averaged as the sampling FVC of the MODIS reflectance data. If any of the corresponding Landsat pixels compared with one MODIS pixel was identified as a cloud by the Fmask process, the MODIS pixel would be removed from the sampling data set. To further verify that the resulting sampling data set was consistent and to remove the unstable samples, the sampling FVC values were plotted against the NDVI values, which were computed from the MODIS reprocessed surface reflectance in the red and NIR bands. Then, for each class of NDVI value (20 classes over the [0, 1] domain of variation), the cases with FVC values that were lower (respectively higher) than the 5% percentile (respectively 95%) were rejected for training the GRNNs [3]. This process further improved the consistency of outputs with input reflectance values through NDVI, as shown in Fig. 4. Finally, 16 980 cases with consistent MODIS surface reflectance were paired with refined sampling FVC values and were confirmed as the sampling data set of GRNNs. The large number of sampling samples used for GRNNs allowed for the generation of better representativeness of the surface types and conditions. The sampling data set was randomly split into a training data set composed of 90% of the data available, and the remaining 10% of the sampling data set was used to evaluate the theoretical performances of GRNNs.

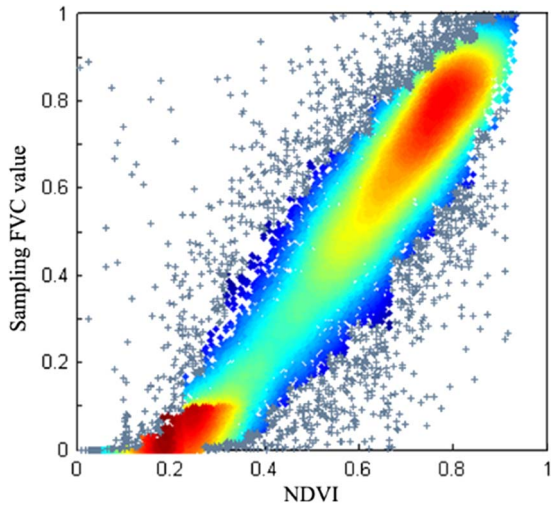


Fig. 4. Relationship between NDVI and the sampling FVC values. The density map presents the densities of the points, which are used in this paper, and the gray points correspond to the rejected cases that are outside the [5%, 95%] percentile range.

E. FVC Inversion Using GRNNs

GRNNs were the generalization of radial basis function networks and probabilistic neural networks, which were developed by [34] and extensively applied in pattern recognition, adaptive control, and surface parameter prediction [11]. The GRNNs approximated the map inherent in any sample data set, and the GRNNs process did not require iterative training; the functional estimate was computed directly from the training data. The general architecture of the GRNNs included four layers, which were the input layer, the pattern layer, the summation layer, and the output layer. The input layer provided the input variables to the GRNNs for retrieving FVC in this paper, including the reprocessed MODIS reflectance values in the red and NIR bands. The output layer computed the GRNNs predicted value of the output variable, which was the corresponding FVC for the input red and NIR reflectance. A Gaussian function was selected as the kernel function of the GRNNs, and the fundamental formulation was expressed as follows:

$$\mathrm{Y}'(\mathrm{X}) = \frac{\sum_{i=1}^{n} \mathrm{Y}^i \exp\left(-\frac{D_i^2}{2\sigma^2}\right)}{\sum_{i=1}^{n} \exp\left(-\frac{D_i^2}{2\sigma^2}\right)} \tag{3}$$

$$D_i^2 = (\mathrm{X} - \mathrm{X}^i)^T (\mathrm{X} - \mathrm{X}^i) \tag{4}$$

where D_i^2 represents the squared Euclidean distance between the input vectors X and the ith training input vector X^i, Y^i is the output vector corresponding to X^i, $Y'(X)$ is the estimation corresponding to X, n is the number of samples, and σ is a smoothing parameter that controls the size of the receptive region. The training of the GRNNs was essentially the optimization of the smoothing parameter σ because the architecture and weights of the GRNNs were determined by the input. The smoothing parameter significantly affected the GRNNs prediction accuracy and [34] suggested the holdout method to find a suitable smoothing parameter. The holdout method for a particular σ consisted of removing one sample from the training data at a time and then constructing GRNNs based on all of the other training samples. The GRNNs were then used to estimate Y for the removed sample. By repeating this process for each sample and storing each estimate, the mean-squared error between the actual sample values Y^i and the estimates could be evaluated. The value of σ giving the smallest error should be used in the final GRNNs [11], [34]. The training processes were completed when the minimum of the cost function of the smoothing parameter was reached

$$\mathrm{f}(\sigma) = \frac{1}{n}\sum\nolimits_i^n \left(\hat{Y}_i(X_i) - Y_i\right)^2 \tag{5}$$

where $\hat{Y}_i(X_i)$ is the estimation corresponding to X_i using the GRNNs trained over all of the training samples, except the ith sample. The widely used shuffled complex evolution method, developed by the University of Arizona, was selected to obtain the optimal smoothing parameter of the GRNNs, which was not susceptible to being trapped by small pits and bumps on the function's surface [35], [36].

According to the MODIS land cover product, FVC values in the vegetation regions, which included eight main biomes, were estimated using the trained GRNNs with the reprocessed MODIS surface reflectance. The FVC values in nonvegetated regions were set to zero.

F. Accuracy Assessment and Product Comparison

Assessment and validation of the moderate-resolution FVC products was usually difficult to achieve because ground point measurements were not suitable for direct comparisons due to the surface heterogeneity. Using high-spatial-resolution remote sensing data to scale the ground measurements up to moderate spatial resolution pixels for comparing and evaluating was always a suitable choice [37]. VALERI (accessed at: http://w3.avignon.inra.fr/valeri), which provided a set of validation sites with high-resolution FVC maps between 2000 and 2008, was selected to compare and evaluate the proposed method and the GEOV1 FVC product [3] with values derived from the aggregated high-resolution FVC maps. The location of the validation sites, land cover types, dates of ground measurement, and mean FVC values from 3 km $\times$ 3 km site areas are provided in Table I. These selected sites contained many land cover types, including grassland, cropland, shrubs, and forest. Therefore, the validation covered various land cover types and had representativeness. The proposed method in this paper and the GEOV1 FVC product were linearly interpolated to the acquisition date of the ground measurements for direct ground validation.

TABLE I
CHARACTERISTICS OF THE SELECTED SITES FOR ACCURACY ASSESSMENT

Site name	Lat (°)	Lon (°)	Land cover	DOY	Year	FVC
Barrax	39.06	-2.10	Cropland	193	2003	0.236
Camerons	-32.60	116.25	Broadleaf forest	63	2004	0.414
Chilbolton	51.16	-1.43	Crops and forest	166	2006	0.647
Concepcion	-37.47	-73.47	Mixed forest	9	2003	0.455
Counami	5.35	-53.24	Tropical forest	269	2001	0.838
Counami	5.35	-53.24	Tropical forest	286	2002	0.858
Demmin	53.89	13.21	Crops	164	2004	0.586
Donga	9.77	1.78	Grassland	172	2005	0.420
Fundulea	44.41	26.58	Crops	128	2001	0.341
Fundulea	44.41	26.58	Crops	144	2002	0.374
Fundulea	44.41	26.59	Crops	144	2003	0.319
Gilching	48.08	11.32	Crops and forest	199	2002	0.676
Gnangara	-31.53	115.88	Grassland	61	2004	0.221
Gourma	15.32	-1.55	Grassland	244	2000	0.236
Gourma	15.32	-1.55	Grassland	275	2001	0.126
Haouz	31.66	-7.60	Cropland	71	2003	0.248
Hirsikangas	62.64	27.01	Forest	226	2003	0.644
Hirsikangas	62.64	27.01	Forest	190	2004	0.537
Hirsikangas	62.64	27.01	Forest	159	2005	0.442
Hombori	15.33	-1.48	Grassland	242	2002	0.200
Hyytiälä	61.85	24.31	Evergreen forest	188	2008	0.461
Jarvselja	58.29	27.29	Boreal forest	188	2000	0.705
Jarvselja	58.30	27.26	Boreal forest	165	2001	0.783
Jarvselja	58.30	27.26	Boreal forest	178	2002	0.793
Jarvselja	58.30	27.26	Boreal forest	208	2003	0.803
Jarvselja	58.30	27.26	Boreal forest	180	2005	0.842
Jarvselja	58.30	27.26	Boreal forest	112	2007	0.535
Jarvselja	58.30	27.26	Boreal forest	199	2007	0.731
Laprida	-36.99	-60.55	Grassland	311	2001	0.722
Laprida	-36.99	-60.55	Grassland	292	2002	0.534
Larose	45.38	-75.22	Mixed forest	219	2003	0.847
Le Larzac	43.94	3.12	Grassland	183	2002	0.300
Les Alpilles	43.81	4.71	Crops	204	2002	0.349
Plan-de-Dieu	44.20	4.95	Crops	189	2004	0.172
Puechabon	43.72	3.65	Forest	164	2001	0.540
Rovaniemi	66.46	25.35	Crops	161	2004	0.423
Rovaniemi	66.46	25.35	Crops	166	2005	0.497
Sonian forest	50.77	4.41	Forest	174	2004	0.903
Sud_Ouest	43.51	1.24	Crops	189	2002	0.352
Turco	-18.24	-68.18	Shrubs	208	2001	0.106
Turco	-18.24	-68.19	Shrubs	240	2002	0.020
Turco	-18.24	-68.19	Shrubs	105	2003	0.044
Wankama	13.64	2.64	Grassland	174	2005	0.036
Zhang Bei	41.28	114.69	Pastures	221	2002	0.353

*DOY: Day of Year.

Intercomparison with existing FVC products was another important way to confirm the reasonability of the proposed method. The newly released GEOV1 FVC product [3], [24], which was an improved version of the CYCLOPES FVC product, was selected to compare with the result of the proposed method. The monthly averaged global FVC maps in January and July of 2003 were generated using the GEOV1 FVC product and the proposed method to investigate the spatial consistency and spatial continuity of the two products through visual analysis. To investigate the temporal consistency and temporal continuity of the GEOV1 FVC product and the

TABLE II
$NDVI_v$ AND $NDVI_s$ VALUES IN THE DIMIDIATE PIXEL MODELS FOR DIFFERENT ECOREGION AND VEGETATION TYPES

Ecoregions (see in [32])	$NDVI_s$	$NDVI_v$ for crop	$NDVI_v$ for forest	$NDVI_v$ for grass-shrub
1	0.249	0.885	0.891	0.882
2	0.245	0.878	0.894	0.889
3	0.232	0.856	0.828	0.828
4	0.226	0.883	0.883	0.877
5	0.206	0.889	0.900	0.886
6	0.243	0.881	0.901	0.891
7	0.229	0.875	0.896	0.870
8	0.183	0.868	0.885	0.857
9	0.240	0.822	0.841	0.822
10	0.164	0.844	0.863	0.837
11	0.192	0.790	0.804	0.794
12	0.203	0.861	0.882	0.856
13	0.212	0.847	0.861	0.801

1: Tropical and subtropical Moist Broadleaf Forests; 2: Tropical and Subtropical Dry Broadleaf Forests; 3: Tropical and Subtropical Coniferous Forests; 4: Temperate Broadleaf and Mixed forests; 5: Temperate Coniferous Forests; 6: Boreal Forests/Taiga; 7: Tropical and Subtropical Grasslands, Savannas, and Shrublands; 8: Temperate Grasslands, Savannas and Shrublands; 9: Flooded Grasslands and Savannas; 10: Montane Grasslands and Shrublands; 11: Tundra; 12: Mediterranean Forests, Woodlands and Scrub; 13: Deserts and Xeric Shrublands.

proposed method, the FVC temporal profiles at each validation site were extracted from 2000 to 2010. For a better evaluation of the quality of the two FVC products, the ground measurements are also shown in the figure of the FVC temporal profiles.

III. RESULT AND DISCUSSION

The determined $NDVI_v$ and $NDVI_s$ values for the dimidiate pixel models of FVC estimation using Landsat data in each ecoregion and for different vegetation types are shown in Table II. Then, the Landsat FVC was estimated using the dimidiate pixel models based on the determined $NDVI_v$ and $NDVI_s$ values. The GRNNs for global FVC estimation were trained using the training samples extracted from reprocessed MODIS surface reflectance and Landsat FVC values. The derived optimal smoothing parameter of the GRNNs in this paper was 0.0042 using the shuffled complex evolution method in the training procedure. The theoretical performance of the trained GRNNs was evaluated over the randomly selected independent samples (see Fig. 5). The results indicated that the training of the GRNNs was very efficient for the FVC estimation ($R^2 = 0.963$, RMSE $= 0.064$). Moreover, the scatters were all distributed around the 1:1 line and nearly no bias was observed over the whole range of variation of the FVC values. In addition, the randomly selected validation samples were independent from the training samples. Therefore, the theoretical validation could be considered to be a direct validation and a good supplement of the low number of ground truth data, although it would not be as reliable as the ground truth data. Thus, the theoretical performance of the trained GRNNs indicated a stable method for global land surface FVC estimation using the MODIS surface reflectance data. In total, 44 ground-measured samples from VALERI sites were used to

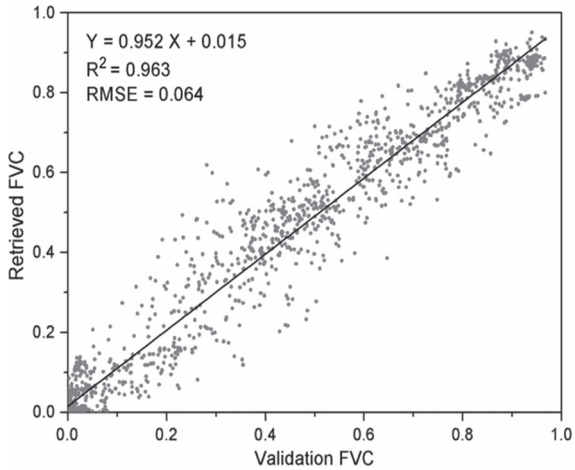


Fig. 5. Comparison between the independent validation samples and the values estimated from the trained GRNNs.

directly validate the result of the proposed method, as well as to compare to the result from the GEOV1 FVC product (see Fig. 6). A good agreement between the FVC estimated from the GRNNs method and from the ground measurements was observed, as well as that between the GEOV1 FVC and the ground measurements. The performance of the GRNNs method ($R^2 = 0.809$, RMSE $= 0.157$) in this paper was comparable with that of the GEOV1 FVC product ($R^2 = 0.775$, RMSE $= 0.166$), noting that the number of GEOV1 FVC samples was less because of invalid data existing in the period of the ground measurements. Moreover, there was a much better agreement between the FVC estimated from GRNNs method and the GEOV1 FVC product ($R^2 = 0.895$), which indicated a good consistency between the two moderate-resolution FVC estimating results, and FVC estimation at an equal spatial resolution might weaken the influence from the upscale of the ground measurements.

The monthly averaged global FVC maps of the proposed method and the GEOV1 FVC product were generated in January and July of 2003 to visually compare the spatial continuity and consistency of the two data sets (see Fig. 7). A good agreement between the FVC estimations obtained from GRNNs method and from the GEOV1 FVC product was observed, which indicated the good spatial consistency between the FVC estimated from the proposed method and the GEOV1 FVC product. However, similar to what was presented by the GEOV1 product validation report [24], the GEOV1 FVC product presented missing data across the spatial range. The percentage of the invalid data of the GEOV1 FVC product in land surface pixels for January and July 2003 was approximately 43.1% and 7.9%, respectively. For example, the GEOV1 FVC product presented a higher percentage of missing values at the high latitudes in the northern hemisphere in the map of January, and the equatorial region also presented a large fraction of gaps in both the maps of January and July (see Fig. 7), which was similar with the validation report of the GEOV1 products [24]. The main cause of the data missing for the GEOV1 FVC product was attributed to the use of the preprocessed SPOT/VGT data, which contained missing reflectance data [24], [38]. The

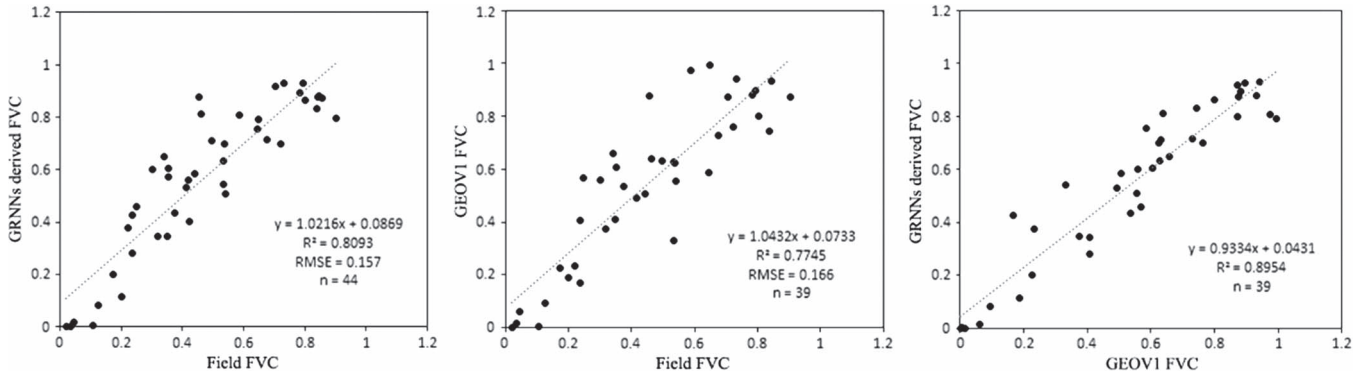


Fig. 6. Comparison of the GRNNs derived FVC, GEOV1 FVC product and field FVC. Because there was invalid data in the GEOV1 FVC product during the period of field data acquisition, the samples of the GEOV1 FVC product were less than the field FVC.

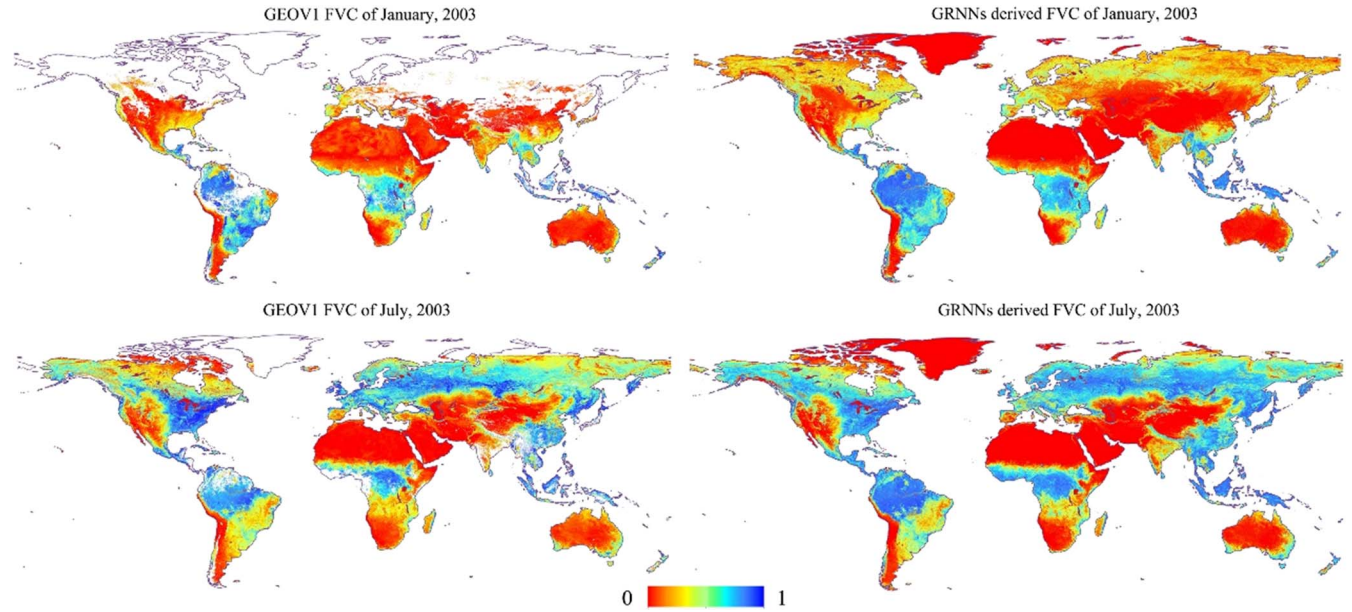


Fig. 7. Monthly averaged global maps of the GRNNs derived FVC and GEOV1 FVC products corresponding to January and July of 2003. The Antarctic Territories are not included in the map.

detailed information regarding the percentage of the missing data in the GEOV1 FVC product could be seen in [24]. In contrast, the FVC maps generated by the GRNNs method had a better performance of spatial continuity, and no missing data were observed. The reprocessed MODIS data with continuous and smooth surface reflectance values [26], [27] made a great effort in generating the spatial continued FVC data using the GRNNs method. Therefore, the global FVC maps generated from GRNNs method had a comparable spatial consistency with the GEOV1 FVC data, and the spatial continuity of the GRNNs derived FVC data were superior to that of the GEOV1 FVC product.

The FVC temporal profiles of the validation sites in Table I were generated to compare the temporal consistency and continuity between the GRNNs derived FVC and GEOV1 FVC products, and some representative FVC temporal profiles are shown in Fig. 8. The temporal consistency of the proposed method and the GEOV1 FVC product was good, and the similar magnitude and dynamic range was observed. Furthermore, there was a clear seasonality of different vegetation types in both data sets, such as the crop growth characteristics, which were clearly described by the FVC seasonality changes in crop regions (Fundulea and Romania sites), and the seasonality characteristics of forest and grassland at the Jarvselja and Gourma sites. The ground-measured data presented in Fig. 8 were also very close to the FVC temporal profiles, except for a few validation sites with distinct discrepancies that might be caused by the different measure scales of ground measurement and satellite observation. The presented vegetation seasonality changes and the temporal consistency between the proposed method and the GEOV1 FVC product indicated that these products were reliable and could reveal the actual earth surface variations. In regard to the temporal continuity of FVC products, there were also some missing data in the GEOV1 FVC product (see Fig. 8), the same as with those missing in the spatial continuity analysis of the GEOV1 FVC product. The temporal continuity of FVC derived by the proposed method was good and no missing data were observed. These results indicated that the proposed FVC

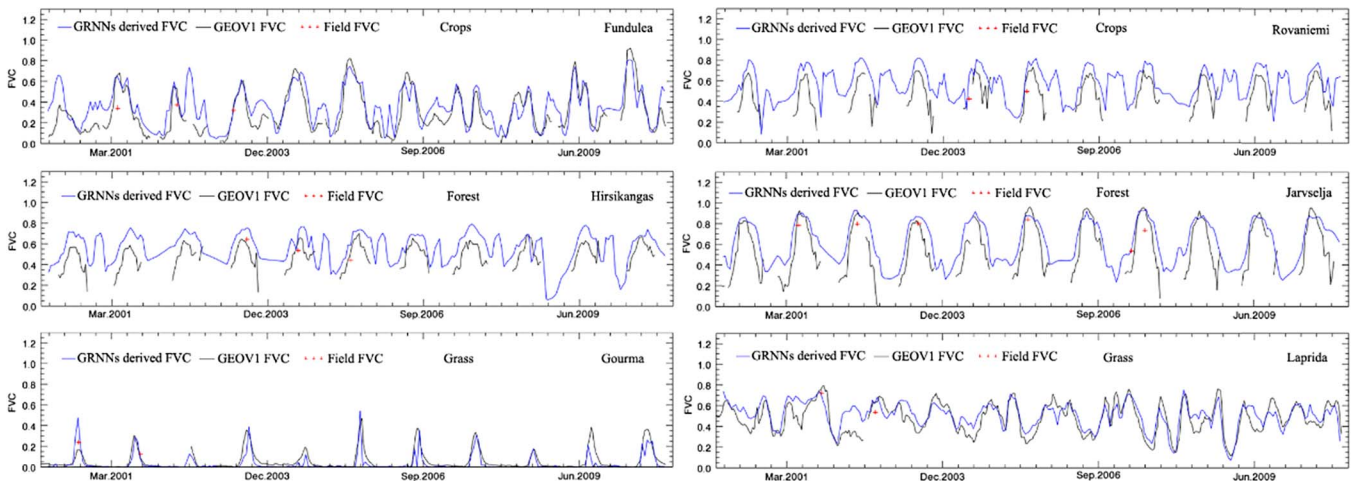


Fig. 8. Temporal profiles of the GRNNs derived FVC and GEOV1 FVC products over validation sites from 2000 to 2010.

estimation method in this paper could generate temporal and spatial continuous FVC product.

The quality assessment of the proposed FVC estimation method indicated that the method could produce reliable global FVC maps, and the spatial and temporal continuity was superior to that of the GEOV1 FVC product. The direct validation results using the VALERI sites data also showed the consistency between the proposed method and the GEOV1 FVC product, and the performance of the proposed method was comparable with that of the GEOV1 FVC product. Therefore, there is a plan to produce the global FVC product using this method based on the reprocessed MODIS reflectance data from the year 2000 to the present. This data set will be freely available to the public at a temporal resolution of 8 days and a spatial resolution of 0.5 km in the sinusoidal projection.

Although the performance of the proposed method is comfortable, much more ground reference data that represents different biomes and geographical areas are required to better assess the uncertainty at a regional scale and in areas across the world. Validation of the moderate-resolution FVC products is difficult to perform because ground measurements are difficult to match to the pixel of the remote sensing product due to the surface heterogeneity. An appropriate strategy is to use high resolution remote sensing data as a bridge to link the ground FVC measurement and the moderate spatial resolution FVC product. However, acquiring and processing high-spatial-resolution remote sensing data at a large scale is costly, which is also why there is limited ground reference data for validating the moderate spatial resolution FVC product. Thus, the further work will locate and generate more ground reference data for a better and extensive assessment of the existing moderate spatial resolution FVC products. Furthermore, the GRNNs were built for FVC estimation using MODIS data, when applied to other data sets, the architecture and parameters should be changed.

IV. Conclusion

A method using the GRNNs has been developed to estimate FVC from MODIS reflectance data in this paper. The GRNNs were trained using the samples from Landsat TM/ETM+ data derived FVC by using the dimidiate pixel model at the global sampling locations. The direct validation and comparison with the GEOV1 FVC product indicated that the performance of the GRNNs derived FVC was comparable with the GEOV1 FVC product. Furthermore, the initial results from GRNNs retrieval presented a much better spatial and temporal continuity, which was convenient for users from different research fields. In summary, the proposed method is reliable for FVC estimation and is being implemented to generate a long time series of global FVC product from MODIS reflectance data. Further work should focus on an extensive assessment of the performance of the proposed method using more ground data.

References

[1] X. Zhang, C. Liao, J. Li, and Q. Sun, "Fractional vegetation cover estimation in arid and semi-arid environments using HJ-1 satellite hyperspectral data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 21, pp. 506–512, Apr. 2013.

[2] K. Jia, B. F. Wu, Y. C. Tian, Y. Zeng, and Q. Z. Li, "Vegetation classification method with biochemical composition estimated from remote sensing data," *Int. J. Remote Sens.*, vol. 32, no. 24, pp. 9307–9325, Dec. 2011.

[3] F. Baret *et al.*, "GEOV1: LAI and FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 1: Principles of development and production," *Remote Sens. Environ.*, vol. 137, pp. 299–309, Oct. 2013.

[4] A. A. Gitelson, Y. J. Kaufman, R. Stark, and D. Rundquist, "Novel algorithms for remote estimation of vegetation fraction," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 76–87, Apr. 2002.

[5] D. Wu *et al.*, "Evaluation of spatiotemporal variations of global fractional vegetation cover based on GIMMS NDVI data from 1982 to 2011," *Remote Sens.*, vol. 6, no. 5, pp. 4217–4239, May 2014.

[6] G. Gutman and A. Ignatov, "The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models," *Int. J. Remote Sens.*, vol. 19, no. 8, pp. 1533–1543, May 1998.

[7] X. B. Zeng *et al.*, "Derivation and evaluation of global 1-km fractional vegetation cover data for land modeling," *J. Appl. Meteorol.*, vol. 39, no. 6, pp. 826–839, Jun. 2000.

[8] J. L. Roujean and R. Lacaze, "Global mapping of vegetation parameters from POLDER multiangular measurements for studies of surface-atmosphere interactions: A pragmatic method and its validation," *J. Geophys. Res. Atmos.*, vol. 107, no. D12, pp. ACL 6-1–ACL 6-14, Jun. 2002.

[9] H. Godinez-Alvarez, J. E. Herrick, M. Mattocks, D. Toledo, and J. Van Zee, "Comparison of three vegetation monitoring methods: Their relative utility for ecological assessment and monitoring," *Ecol. Indicators*, vol. 9, no. 5, pp. 1001–1008, Sep. 2009.

[10] K. Jia *et al.*, "Automatic land-cover update approach integrating iterative training sample selection and a Markov random field model," *Remote Sens. Lett.*, vol. 5, no. 2, pp. 148–156, Feb. 2014.
[11] Z. Q. Xiao *et al.*, "Use of general regression neural networks for generating the GLASS leaf area index product from time-series MODIS surface reflectance," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 209–223, Jan. 2014.
[12] G. Jiapaer, X. Chen, and A. Bao, "A comparison of methods for estimating fractional vegetation cover in arid regions," *Agricultural Forest Meteorol.*, vol. 151, no. 12, pp. 1698–1710, Dec. 2011.
[13] J. Xiao and A. Moody, "A comparison of methods for estimating fractional green vegetation cover within a desert-to-upland transition zone in central New Mexico, USA," *Remote Sens. Environ.*, vol. 98, no. 2/3, pp. 237–250, Oct. 2005.
[14] T. N. Carlson and D. A. Ripley, "On the relation between NDVI, fractional vegetation cover, and leaf area index," *Remote Sens. Environ.*, vol. 62, no. 3, pp. 241–252, Dec. 1997.
[15] P. R. J. North, "Estimation of f(APAR), LAI, and vegetation fractional cover from ATSR-2 imagery," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 114–121, Apr. 2002.
[16] J. C. Jimenez-Munoz *et al.*, "Comparison between fractional vegetation cover retrievals from vegetation indices and spectral mixture analysis: Case study of PROBA/CHRIS data over an agricultural area," *Sensors*, vol. 9, no. 2, pp. 768–793, Feb. 2009.
[17] B. Wu, M. Li, C. Yon, W. Zhou, and C. Yan, "Developing method of vegetation fraction estimation by remote sensing for soil loss equation: A case in the Upper Basin of Miyun Reservoir," in *Proc. IEEE IGARSS*, 2004, pp. 4352–4355.
[18] J. Qi *et al.*, "Spatial and temporal dynamics of vegetation in the San Pedro River basin area," *Agricultural Forest Meteorol.*, vol. 105, no. 1–3, pp. 55–68, Nov. 2000.
[19] D. S. Kimes, Y. Knyazikhin, J. L. Privette, A. A. Abuelgasim, and F. Gao, "Inversion methods for physically-based models," *Remote Sens. Rev.*, vol. 18, no. 2–4, pp. 381–439, Sep. 2000.
[20] F. Baret *et al.*, "Evaluation of the representativeness of networks of sites for the global validation and intercomparison of land biophysical products: Proposition of the CEOS-BELMANIP," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 7, pp. 1794–1803, Jul. 2006.
[21] K. Jia *et al.*, "Fractional forest cover changes in Northeast China from 1982 to 2011 and its relationship with climatic variations," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 2, pp. 775–782, Feb. 2015.
[22] F. Baret *et al.*, "LAI, fAPAR and fCover CYCLOPES global products derived from VEGETATION—Part 1: Principles of the algorithm," *Remote Sens. Environ.*, vol. 110, no. 3, pp. 275–286, Oct. 2007.
[23] F. J.García-Haro, F. Camacho, and J. Meliá, "Inter-comparison of SEVIRI/MSG and MERIS/ENVISAT biophysical products over Europe and Africa," in *Proc. 2nd MERIS/(A)ATSR User Workshop*, Frascati, Italy, 2008, pp. 1–8.
[24] F. Camacho, J. Cernicharo, R. Lacaze, F. Baret, and M. Weiss, "GEOV1: LAI, FAPAR essential climate variables and FCOVER global time series capitalizing over existing products. Part 2: Validation and intercomparison with reference products," *Remote Sens. Environ.*, vol. 137, pp. 310–329, Oct. 2013.
[25] J. M. Chen, F. Deng, and M. Z. Chen, "Locally adjusted cubic-spline capping for reconstructing seasonal trajectories of a satellite-derived surface parameter," *IEEE Trans. Geosci. Remote Sens.*, vol. 44, no. 8, pp. 2230–2238, Aug. 2006.
[26] H. R. Tang *et al.*, "A cloud detection method based on a time series of MODIS surface reflectance images," *Int. J. Digit. Earth*, vol. 6, no. 1, pp. 157–171, Dec, 2013.
[27] S. L. Liang *et al.*, "A long-term Global LAnd Surface Satellite (GLASS) data-set for environmental studies," *Int. J. Digit. Earth*, vol. 6, pp. 5–33, Dec. 2013.
[28] J. G. Masek *et al.*, "A Landsat surface reflectance data set for North America, 1990-2000," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 68–72, Jan. 2006.
[29] Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sens. Environ.*, vol. 118, pp. 83–94, Mar. 2012.
[30] A. A. Gitelson, "Remote estimation of crop fractional vegetation cover: The use of noise equivalent as an indicator of performance of vegetation indices," *Int. J. Remote Sens.*, vol. 34, no. 17, pp. 6054–6066, Sep. 2013.
[31] P. J. Sellers *et al.*, "A revised land surface parameterization (SiB2) for atmospheric GCMs.2. The generation of global fields of terrestrial biophysical parameters from satellite data," *J. Climate*, vol. 9, no. 4, pp. 706–737, Apr. 1996.
[32] D. M. Olson *et al.*, "Terrestrial ecoregions of the worlds: A new map of life on Earth," *Bioscience*, vol. 51, no. 11, pp. 933–938, Nov. 2001.
[33] P. Gong *et al.*, "Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data," *Int. J. Remote Sens.*, vol. 34, no. 7, pp. 2607–2654, Apr. 2013.
[34] D. F. Specht, "A general regression neural network," *IEEE Trans. Neural Netw.*, vol. 2, no. 6, pp. 568–576, Nov. 1991.
[35] Q. Y. Duan, S. Sorooshian, and V. Gupta, "Effective and efficient global optimization for conceptual rainfall-runoff models," *Water Res. Res.*, vol. 28, no. 4, pp. 1015–1031, Apr. 1992.
[36] Z. Q. Xiao *et al.*, "Variational retrieval of leaf area index from MODIS time series data: Examples from the Heihe river basin, north-west China," *Int. J. Remote Sens.*, vol. 33, no. 3, pp. 730–745, 2012.
[37] S. L. Liang *et al.*, "Validating MODIS land surface reflectance and albedo products: Methods and preliminary results," *Remote Sens. Environ.*, vol. 83, no. 1/2, pp. 149–162, Nov. 2002.
[38] O. Hagolle *et al.*, "Quality assessment and improvement of temporally composited products of remotely sensed imagery by combination of VEGETATION 1 and 2 images," *Remote Sens. Environ.*, vol. 94, no. 2, pp. 172–186, Jan. 2005.

Kun Jia received the B.S. degree in surveying and mapping engineering from Central South University, Changsha, China in 2006, and the Ph.D. degree in cartography and geographic information system from Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 2011.

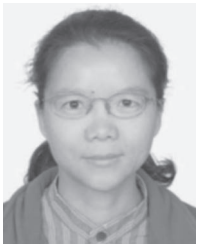
He is currently with the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing. His main research interests include estimation of fractional vegetation cover, land cover classification, and agriculture monitoring using remote sensing data.

Shunlin Liang (M'94–F'13) received the Ph.D. degree from Boston University, Boston, MA, USA.

He is currently a Professor with the Department of Geographical Sciences, University of Maryland, College Park, MD, USA, and the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing, China. His main research interests focus on estimation of land surface variables from satellite data, Earth energy balance, and assessment of environmental impacts of vegetation changes.

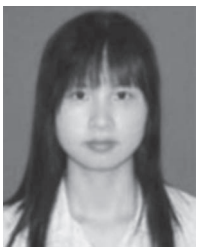
He published over 220 peer-reviewed journal papers, authored the book *Quantitative Remote Sensing of Land Surfaces* (Wiley, 2004), coauthored the book *Global LAnd Surface Satellite (GLASS) Products: Algorithms, Validation and Analysis* (Springer, 2013), edited the book *Advances in Land Remote Sensing: System, Modeling, Inversion and Application* (Springer, 2008), and coedited the books *Advanced Remote Sensing: Terrestrial Information Extraction and Applications* (Academic Press, 2012) and *Land Surface Observation, Modeling, Data Assimilation* (World Scientific, 2013).

Dr. S. Liang was an Associate Editor of the IEEE TRANSACTION ON GEOSCIENCE AND REMOTE SENSING and also a guest editor of several remote sensing related journals.



Suhong Liu received the B.S. degree in computer science from Southwest Jiaotong University, Chengdu, China, in 1988, the M.S. degree in geophysical well-logging from Jianghan Petroleum University, Jingzhou, China, in 1991, and the Ph.D. degree in cartography and remote sensing from the Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China, in 1999.

She is currently an Associate Professor with the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing. Her current research interests include spatiotemporal analysis of remotely sensed data, and retrieval of land biophysical parameters from satellite data.



Yuwei Li received the B.S. degree in cartography and geographic information system from Wuhan University, Wuhan, China, in 2013. She is currently working toward the M.S. degree with the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing, China.

Her main research interests include estimation of fractional vegetation cover using remote sensing.

Zhiqiang Xiao received the Ph.D. degree in geophysical prospecting and information technology from Central South University, Changsha, China, in 2004.

From 2004 to 2006, he was a Postdoctoral Research Associate with Beijing Normal University, Beijing, China. He is currently with the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University. His research interests include retrieving land biophysical parameters from remotely sensed data and assimilating radiometric observations into dynamic models.

Yunjun Yao received the Ph.D. degree from Peking University, Beijing, China, in 2010.

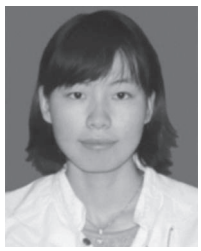
From October 2008 until October 2009, he was with the Department of Geographical Sciences, University of Maryland, College Park, MD, USA, as a joint Ph.D. student. He is currently with the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing. His main research interests include estimation of evapotranspiration and retrieval of surface biophysical parameters by remote sensing.

Bo Jiang received the B.S. degree from Central South University, Changsha, China, in 2006, the M.S. degree from Beijing Normal University, Beijing, China, in 2009, and the Ph.D. degree from Beijing Normal University in 2012. From September 2010 until May 2012, she was with the Department of Geographical Sciences, University of Maryland, College Park, MD, USA, as a joint Ph.D. student.

She is now with the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University. Her main research interests include assessing the impacts of the land cover change on the climate from various observations, and application of remote sensing products.

Xiang Zhao received the Ph.D. degree in cartography and geographic information system from the School of Geography, Beijing Normal University, Beijing, China, in 2006.

During the years 2008–2010, he was a Postdoctoral Fellow with the College of Resources Science and Technology, Beijing Normal University. He is currently with the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University. His main research interests focus on high-performance computing system construction and quantitative remote sensing application. He also did some research about long time series remote sensing data trend analysis.



Xiaoxia Wang received the B.S. degree in surveying and mapping engineering from Central South University, Changsha, China in 2014. She is currently working toward the M.S. degree in the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing, China.

Her main research interests include estimation of fractional vegetation cover using remote sensing.

Shuai Xu received the B.S. degree in electronic information engineering from Hebei University of Science and Technology, Hebei, China, in 2009.

He is currently with the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing, China. His main research interests focus on high-performance computing system construction.

Jiao Cui received the B.S. degree in computer science from Sun Yat-sen University, GuangZhou, China, in 2011.

She is currently with the State Key Laboratory of Remote Sensing Science, School of Geography, Beijing Normal University, Beijing, China. Her main research interests focus on high-performance computing system construction.