Agricultural and Forest Meteorology 189–190 (2014) 60–70

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

journal homepage: www.elsevier.com/locate/agrformet

Improvement of spatially continuous forest LAI retrieval by integration of discrete airborne LiDAR and remote sensing multi-angle optical data

Han Ma[a,c], Jinling Song[a,b,c,*], Jindi Wang[a,b,c], Zhiqiang Xiao[a,b,c], Zhuo Fu[d]

[a] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and the Institute of Remote Sensing Applications, CAS, Beijing 100875, China
[b] Beijing Key Laboratory for Remote Sensing of Environment and Digital Cities, Beijing Normal University, Beijing 100875, China
[c] School of Geography and Remote Sensing Science, Beijing Normal University, Beijing 100875, China
[d] Satellite Environment Center, Ministry of Environmental Protection, Beijing 100094, China

ARTICLE INFO

Article history:
Received 8 May 2012
Received in revised form 13 August 2013
Accepted 8 January 2014

Keywords:
Forest
LAI
Canopy height
LiDAR
MODIS
MISR

ABSTRACT

Forest leaf area index (LAI) is a critical variable in modeling climates and ecosystems, and is required on regional and global scales for models. However, forest LAI has proven to be difficult to obtain. In this study, we sought to improve forest LAI retrieval in a large study area in the Dayekou forest, Gansu province, by combining airborne discrete LiDAR, MODIS, and MISR data. In our retrieval scheme, canopy height is the key parameter, and the canopy height precision is of great importance when estimating LAI. To address this issue, we introduced LiDAR data and combined it with the MODIS and MISR products. First, the canopy height for the LiDAR data coverage was calculated using a local maximum filtering algorithm with a variable window size. Then, a multivariate linear regression model was developed to extrapolate the LiDAR-derived canopy height to the whole study area using the MODIS BRDF/Albedo product. In addition, the bi-directional reflectances from MODIS and MISR were used to invert the geometric-optical mutual-shadowing (GOMS) model structural parameters (nR^2, b/R, h/b, $\Delta h/b$) of the forest. These structural parameters were then combined with the forest canopy height and field measurements to retrieve the LAI of the continuous forest area at a 500-m resolution. After comparison with the true LAI measured by LAI-2000 combined with TRAC, and by TRAC alone, the highest R^2 values of the estimated LAI were 0.73 and 0.69, respectively. The results indicate that the LiDAR canopy height derived from the optical multi-angle remote sensing data can be used to retrieve the large-scale forest LAI when combined with the canopy structure information derived from GOMS model.

1. Introduction

Forest structure estimation is critical in studies of global carbon energy-cycle systems and is required by many terrestrial ecosystem models. Forest canopy leaf area index (LAI), a critical forest structural parameter for accurately modeling energy, carbon, water, and climate, is required on regional, continental, and global scales for models (Neilson and Drapek, 1998). LAI is defined as half of the total green leaf area per unit of horizontal ground surface area (Chen and Black, 1992). Remotely estimating forest structure potentially provides an ecologically significant advance over laborious ground-based estimation methods, and has become widely used to characterize forest structure (Lefsky et al., 2002). There are two main types of LAI inversion methods from remote sensing data. One is to develop the empirical or semi-empirical relationship between the vegetation indices and the LAI; the other type is based on radiation transfer models using a look up table (LUT) or neural network (NN) (Liang, 2004).

LiDAR, which is rapidly emerging as an active remote sensing technique, has been successfully used to estimate the canopy height, LAI, biomass, and other variables (Riaño et al., 2004; Lefsky et al., 2005, 2007; Dubayah et al., 2010); however, current airborne and satellite-borne LiDAR systems can only provide sample data and cannot provide data on the complete horizontal coverage because of the high cost and lack of mapping capability (Sun et al., 2006; Swatantran et al., 2011).

Though passive optical sensors cannot directly measure the vertical vegetation structure, multi-angle directional reflectance data from sensors such as MODIS and MISR can provide information on

Abbreviations: LAI, leaf area index; FAVD, foliage area volume density; GOMS, geometric-optical mutual shadowing; BRDF, bidirectional reflectance distribution function; DEM, digital elevation model; DSM, digital surface model; CHM, canopy height model; USM, uncertainty and sensitivity matrix.

* Corresponding author at: State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and the Institute of Remote Sensing Applications, CAS, Beijing 100875, China. Tel.: +86-10-58805452.
E-mail address: songjl@bnu.edu.cn (J. Song).

0168-1923/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.agrformet.2014.01.009

List of variables

G	radiation exitance from sunlit background surfaces
C	radiation exitance from sunlit crown surfaces
P	radiation exitance from shaded backgrounds
b	vertical half axis of ellipsoid
R	horizontal radius of ellipsoid
h	height at which a crown center is located
n	number of crown centers per m^2 distributed in a pixel
nR^2	parameter that describes the crown coverage density in the nadir observation
b/R	crown shape parameter that has a significant effect on the coverage density in the non-nadir direction
h/b	parameter to describe the height between the crown and the ground; it mainly affects the outward width of the hot spot
$\Delta h/b$	parameter that describes the dispersed degree of the distribution of the crown height and affects the bowl-shape of the BRDF
$H_{\text{LiDAR\_indi}}$	individual tree height extracted from LiDAR-built CHM
$H_{\text{LiDAR\_mean}}$	mean of individual tree heights extracted from CHM at the plot level
$H_{\text{field\_indi}}$	individual tree height measured *in situ*
$H_{\text{field\_mean}}$	mean of individual tree heights measured *in situ* at the plot level
H_{model}	mean canopy height derived from regression model

the BRDF characteristics of vegetation, and they have the imaging capability to make up for the limitations of LiDAR systems (Kimes et al., 2006). Finding an approach that combines optical remote sensing data and LiDAR datasets to retrieve regional and global scale vegetation parameters is of great importance. Fu et al. (2011) used several forest pixels in Gansu, China, and successfully retrieved forest LAI by combining optical multi-angle and airborne LiDAR data based on canopy reflectance model inversion. Wang et al. (2011) estimated the regional forest height at three forest sites in the United States by developing regression models that use spectrally independent directional escape probabilities combined with MODIS BRDF/Albedo product and LVIS data. Lefsky (2010) obtained a global forest canopy height map from MODIS and GLAS data using regression analysis; Chen et al. (2012) mapped a large area of fine-scale tree heights by integrating LiDAR data and Landsat imagery using a geometric optical model and spectral mixture analysis.

Based on former studies, the objectives of our study are to retrieve forest structural parameters, particularly forest LAI, by combining LiDAR and multi-angle remote sensing data, and to generate regional scale forest structure distribution information at a 500-m resolution.

To retrieve the forest structural parameters, four main steps were involved in this study. First, the canopy height was extracted from the LiDAR data coverage; secondly, the LiDAR-derived canopy height was extrapolated to the whole study area, using the MODIS BRDF/Albedo product; thirdly, based on a GOMS model, some of the canopy structural parameters were retrieved by combining the MODIS and MISR multi-angle reflectance data; finally, the LAI of the continuous forest area was estimated based on the parameters retrieved in the previous steps. These four steps are illustrated in Fig. 1.

2. Data and pre-processing

2.1. Study sites

The study area, Dayekou forest (38.6° E, 100° W) (Fig. 2), is located in the Qilian Mountain area of Gansu province in the arid region of northwest China. The forest is influenced by the temperate continental mountain climate, and the annual precipitation takes place mainly in summer. Prevalent vegetation types at this site are mountain pastures and forests, and the dominant forest type is *Piceacrassifolia*. MODIS land cover product (MCD12Q1) was used to select forest pixels, shown in green in Fig. 2, and the number of forest pixels was 5173.

2.2. Field measurements

Field measurements were performed during June 2008 in 15 sample plots A, of size of 20 m × 20 m, and 16 sample plots B, of size of 25 m × 25 m. Fig. 3 shows the distribution of sample plots on the CHM image; the red squares represent the MODIS pixels and the green and yellow crosses represent the locations of sample plots A and B, respectively. There are five sample A plots inside the green square belonging to a MODIS pixel whose land cover type is open shrubland. Because non-forest pixels were not considered in this study, the LAI value of this pixel was not estimated.

Two types of instruments, a LAI-2000 Plant Canopy Analyzer (LAI-2000) (Gower et al., 1999) and Tracing Radiation and Architecture of Canopies (TRAC) (Chen and Cihlar, 1995) were used to measure the effective LAIs (L_e) and foliage clumping indices (Ω) of the sample plots (Chen, 1996). The true LAI was calculated as the ratio between L_e and Ω. Thus, two sets of sample LAI were obtained from different combinations of L_e and Ω: L_e measured by LAI-2000 and Ω measured by TRAC, and L_e and Ω both measured by TRAC.

In addition, the geometrical structural parameters of each tree in sample B, including tree height ($H_{\text{field\_indi}}$), clear bole height (h_{clear}), horizontal radius of the tree crown (R), and the position of each tree were measured (1656 trees in total). Field measurements were used to estimate the accuracy of the LAI inversion results and the canopy height obtained from airborne LiDAR.

2.3. Airborne LiDAR data

The airborne discrete LiDAR data for the study area were acquired on June 28, 2008, using a Riegl LMS-Q560 laser scanner with a maximum scan angle of ±0.5 mrad from nadir. The point density was approximately 2–4 points/m^2. The raw point cloud data were processed using Terrasolid software to classify the data into two types, *i.e.*, ground points and non-ground points, which can produce a digital elevation model (DEM) and a digital surface model (DSM), respectively. A canopy height model (CHM) was created at a resolution of 0.5 m by subtracting the DEM from DSM, as shown in Fig. 3. The CHM represents a three-dimensional surface that characterizes the canopy height.

2.4. MODIS BRDF/Albedo product

The global land surface MODIS BRDF/Albedo product (MCD43) used multi-day, cloud-free, atmospherically corrected surface reflectance data to establish the best-fit BRDF models in seven spectral bands at 500-m resolution every 16 days on an 8-day cycle. The algorithm of the product is a kernel-driven linear BRDF model, known as the Ross-Thick/Li-Sparse Reciprocal (RTLSR) kernel-driven BRDF model (Roujean, 1992; Wanner et al., 1995; Lucht et al., 2000; Schaaf et al., 2002). This model uses a constant term for

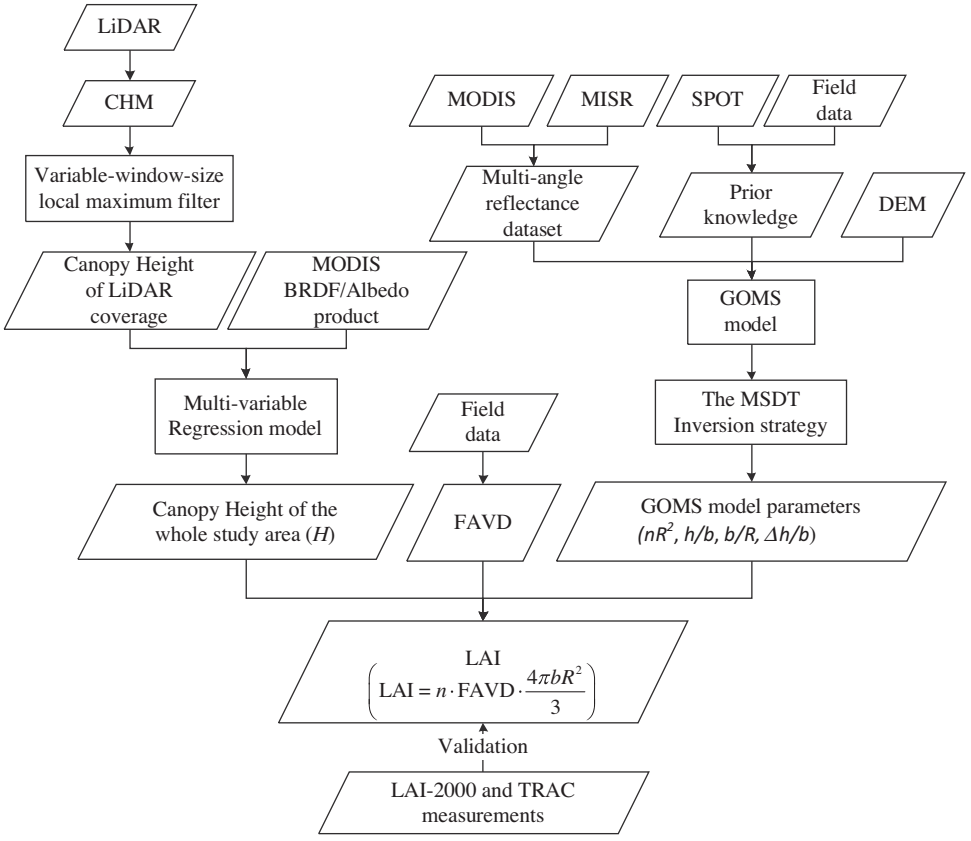

Fig. 1. Flow chart of LAI retrieval method.

isotropic scattering, the Ross-Thick kernel for volumetric scattering, and the Li-Sparse-Reciprocal kernel for geometric scattering.

The collection 005 standard operational MODIS BRDF/Albedo product includes the best-fit and well-sampled RTLSR model parameters (MCD43A1), which contains the first seven spectral bands of MODIS (centered at 648 nm, 858 nm, 470 nm, 555 nm, 1240 nm, 1640 nm, and 2130 nm) and three additional broad bands (0.3–0.7 μm, 0.7–5 μm, 0.3–5 μm). The nadir (view-angle-corrected) surface reflectances at each location (MCD43A4), giving extensive quality information (MCD43A2), are also included.

The RTLSR model parameters (MCD43A1) were used to calculate the forest pixel reflectances along the principal planes at different angles to fit the model for estimating tree heights, as will be described later.

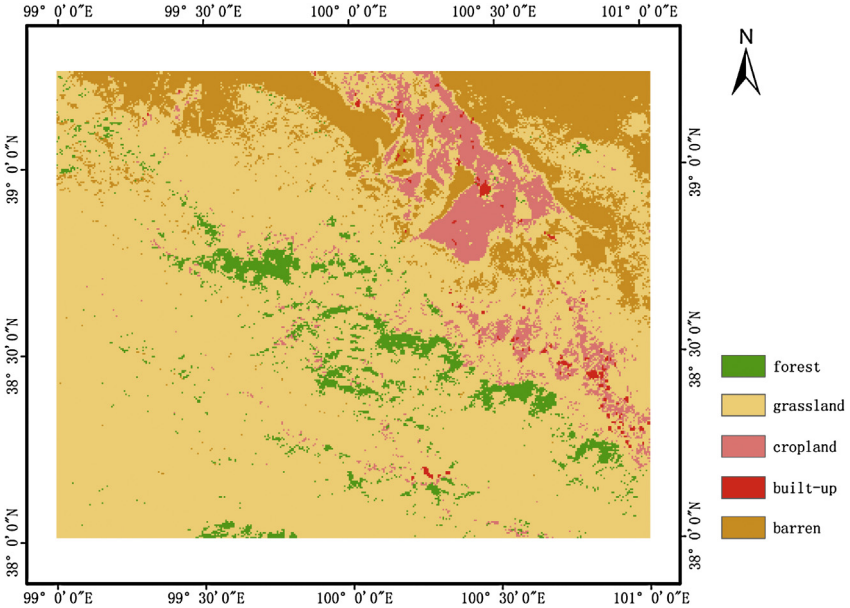

Fig. 2. Map of Dayekou forest from MODIS land-cover product (Pixels in green are forest pixels). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

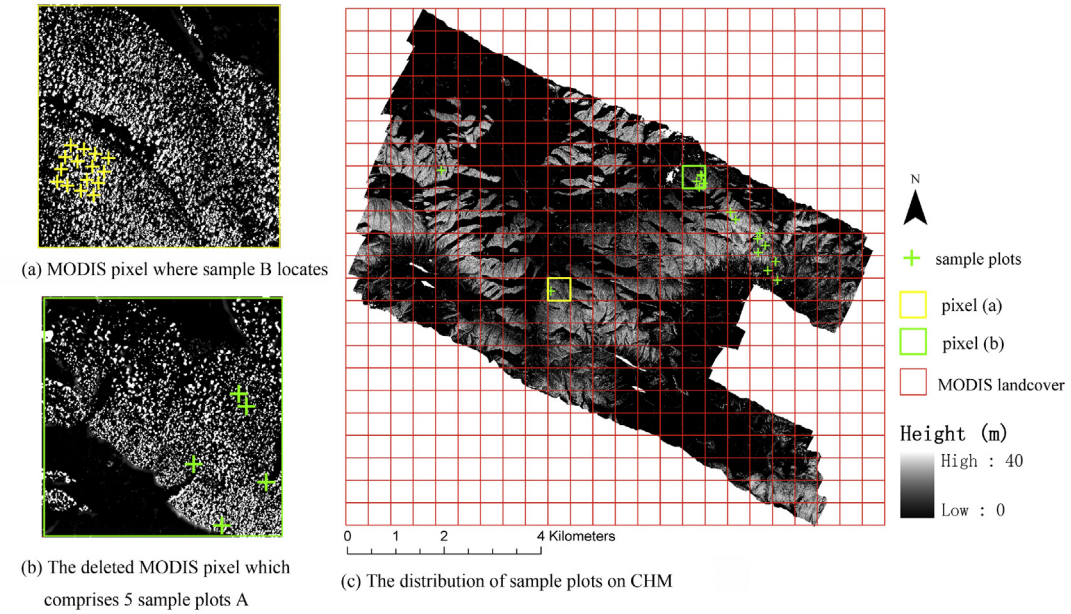

Fig. 3. Distribution of sample plots on CHM (a) MODIS pixel where sample B is located, (b) deleted MODIS pixel, which comprises five sample A plots, and (c) distribution of sample plots on CHM (red squares represent MODIS pixels, green crosses represent the locations of sample A plots, and yellow crosses represent the locations of sample B plots). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.5. MODIS, MISR, and SPOT reflectance data

The MODIS and MISR reflectance products were used to build a multi-angle bi-directional reflectance dataset of the study area forest pixels. The observed azimuth of the two sensors is approximately perpendicular, so the combination of the two products can add to the observation information. The MODIS 500-m bi-directional surface reflectance data (MOD09A1) were obtained from the NASA Distributed Active Archive Center (DAAC) on ISIN tile h25 v05 for the dates spanning May 1–July 28, 2008, a total of 12 scenes. The MOD09A product is produced in 8-day maximum-quality composites in 500-m pixels and with observations with minimal cloud cover, low solar zenith angles, and near-nadir views being chosen (Vermote and Vermeulen, 1999). The MISR 1.1-km reflectance data are acquired with view angles of 0°, ±26.1°, ±45.6°, ±60.0°, and ±70.5°. Radiometrically calibrated reflectance data at each angle are obtained at four spectral bands 446, 558, 672, and 866 nm (Diner et al., 1998).

One of the input parameters of GOMS model, the multi-angle reflectance dataset, was constructed using the following procedure. First, each of the MISR pixels was divided into four 500-m pixels with the same reflectance values and observation geometries, to ensure that MODIS and MISR had the same spatial resolution. Secondly, the projections of both MODIS and MISR were transferred into UTM, WGS84. Finally, the geometric information (sun zenith angle and azimuth angle, viewing zenith angle and azimuth angle) of each pixel and the reflectance values of the red and near-infrared bands were extracted to construct the multi-angle reflectance dataset of each pixel. The MRT tool was used for batch projection transformation of the MODIS data, and the GCTP library function was used for projection transformation of MISR and co-registration with the MODIS image, in which the MODIS image was taken as the base image (Demcsak, 1997). Fig. 4 shows the MISR and MODIS false color RGB composite images of the study area. The MISR image (left) is from path 134, orbit 44777, and the MODIS image (right) is from the 121st day of 2008. By comparing the textures of these two images, we can conclude that the processing was successful because their textures coincide well; inspection of a set of feature points shows that the geographic registration error was less than half MISR pixel (500 m).

SPOT-5 multi-band image was obtained on August 10, 2008, with the spatial resolution of 10 m. The FLASSH model was used for atmospheric correction of the SPOT-5 image. The SPOT data were mainly used to extract prior knowledge of three spectral parameters (G, C, and P) of the forest canopy.

3. Methods

3.1. Extracting canopy height from LiDAR data

LiDAR can directly measure the three-dimensional structures of forests, and studies have shown that airborne LiDAR provides a high measurement accuracy for terrain elevation and vegetation heights, even on sloped terrain or for dense forests (Popescu et al., 2011).

Based on the CHM built in Section 2.3, the individual tree heights ($H_{\mathrm{LiDAR\_indi}}$) were derived using a local maximum filtering with a variable square window. This approach has been described by Popescu et al. (2002). The local maximum algorithm operates on the assumption that the highest laser value in a spatial neighborhood represents the apex of a tree crown. The window size used to search for treetops is derived from the relationship between a tree's height and its crown size. The tree height and crown size data from the field inventory in sample B were used to derive the relationship

$$Crown\ \ size = 0.209 \times H + 1.2 \quad (1)$$

where H is the tree height (m); the R^2 value of this linear relationship is 0.60.

The field inventory data show that the minimum tree height is 1.9 m. Thus, the first step in the algorithm is to threshold the CHM values to eliminate shrubs and understory vegetation, using this minimum tree-height value. Then, the height value at each pixel of the CHM is read, and the window size is calculated using Eq. (1), to search for the local maximum; this is performed if the current pixel corresponds to the local maximum and is identified as the treetop. Once the location of each identified tree crown has been estab-

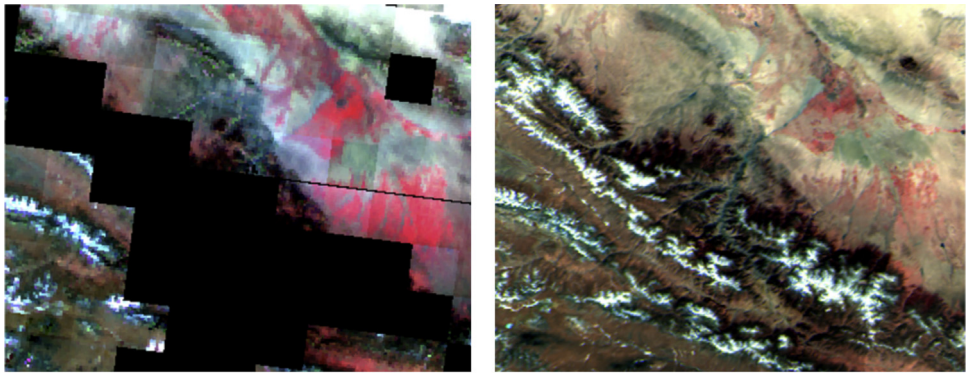

Fig. 4. (a) MISR false color RGB composite image of path 134, orbit 44777 and (b) MODIS false color RGB composite image on the 121st day of 2008.

lished, the CHM is sampled only at the positions of the tree apex to determine the height of each tree. Finally, we calculate the mean tree height for each corresponding MODIS pixel at a resolution of 500 m ($H_{\text{LiDAR\_mean}}$). To evaluate this process, the LiDAR-derived canopy height was compared with the tree field measurements in 16 subplots of sample B (see Section 4.1).

3.2. Extrapolation of canopy height to the whole study area

A major limitation of LiDAR is that only sample data for a region can be acquired. In this study, a multivariate linear regression model was developed to extrapolate the LiDAR-derived canopy height to the whole study area. Reflectance values calculated from the MODIS BRDF model parameters (MCD43A1), which were along the principal plane at seven view zenith angles (0°, ±26.1°, ±45.6°, ±60°), seven MODIS bands, and a 38° sun zenith angle, were used as independent variables, and the mean canopy heights were response variables.

To give an estimate of the predictive error in extrapolating the tree height, a bootstrapping method was used. The plot-level-height dataset derived from LiDAR ($H_{\text{LiDAR\_mean}}$), which has 142 (500 m × 500 m) pixels, was randomly partitioned into training and testing sets; half of the pixels were chosen for training and the other half for testing. The RMSE (root-mean-square error of the predicted *versus* true target value of the model) and R^2 (coefficient of determination of the predicted *versus* true target value of the model) were calculated to document the model accuracy (see Section 4.2). To find the best regression model parameters, this process was repeated 500 times; the regression results are given in Section 4.2.

3.3. Extraction of GOMS model structural parameters

As stated above, our third scheme was to extract the GOMS model structural parameters from MODIS and MISR reflectances. Based on the Li–Strahler geometric-optical model, GOMS model solves the problems of hotspot and mutual-shadowing effects of crowns at large zenith angles, and is more suitable for forests with high crown densities (Li and Strahler, 1992). GOMS model treats the basic crown shape as a spheroid and assigns four parameters to describe the spheroid (b, R, h, n); it also incorporates three multi-spectral reflectance vectors (G, C, and Z). Assuming that there are three types of scene components in a pixel, *i.e.*, sunlit background, sunlit crown, and shadowed surface, the reflectance of a pixel can be modeled as a sum of the reflectances of its individual scene components, weighted by their respective areas within the pixel. The area proportion can be expressed by a combination of the forest crown structural parameters:

$$\text{Reflecntance} = f(\theta_i, \phi_i, \theta_v, \phi_v, \theta_s, \phi_s, nR^2, b/R, h/b, \Delta h/b, G, C, Z) \quad (2)$$

whereθ_i and φ_i are the sun zenith angle and azimuth angle, respectively, θ_v and φ_v are the viewing zenith angle and azimuth angle, respectively, andθ_s and φ_s are the slope and aspect, respectively, within the pixel.

3.3.1. Input parameters

The input parameters of GOMS model include the multi-angle reflectance dataset constructed in Section 2.3, and terrain information (slope and aspect of each pixel) extracted from the DEM. In addition, prior knowledge of the canopy structure and spectral parameters is also needed to reduce the uncertainty of the inversion process (Li et al., 1998). In the present study, prior knowledge of the three spectral parameters was extracted from the atmosphere-corrected image SPOT-5, and the four structural parameters were obtained from the *in situ* measurement statistics of each tree. The initial estimates and uncertainties of the seven parameters are given in Table 1. RG, RC, and RZ are G, C, and Z, respectively, in the red band, and NIRG, NIRC, and NIRZ are G, C, and Z, respectively, in the near-infrared band.

3.3.2. Model inversion strategy

The multi-stage, sample-direction-dependent, target-decisions (MSDT) inversion strategy proposed by Li et al. (1997) was adopted in this study. This method realizes segmentation of the observation data and the parameters to be inverted. The most sensitive observation data are used to invert the most sensitive parameters. The previous inversion result is taken as the prior knowledge in the next inversion stage.

The MSDT inversion strategy is based on the uncertainty and sensitivity matrix (USM) of the parameters in a given direction or bands of observation. The USM is an objective expression of the prior knowledge and is used to determine the order of the parameters to be inverted, as well as the segmentation of the observation data in each order. The USM elements can be expressed as

$$\text{USM}(i,j) = \frac{\Delta\text{BRDF}(i,j)}{\text{BRDF}_{\text{exp}}(i)} \quad (3)$$

where ΔBRDF(i,j) is the maximum difference of BRDF as a function of only the jth parameter within its uncertainty, given the ith geometry of illumination and viewing, and $\text{BRDF}_{\text{exp}}(i)$ is the BRDF value predicted by the model at the ith geometry, with all parameters at their expected values.

Sensitivity analysis of the USM is illustrated in detail in Fu et al. (2011), and our inversion process is performed accordingly (see

Table 1
Prior knowledge of GOMS model parameters.

	Structural parameters				Spectral components (R)			Spectral components (NIR)		
	nR^2	b/R	h/b	$\Delta h/b$	RG	RC	RZ	NIRG	NIRC	NIRZ
Initial estimate	0.37	1.6	3.02	5	0.107	0.044	0.015	0.42	0.38	0.09
Upper boundary	0.1	0.74	1	0	0	0	0	0	0	0
Lower boundary	0.8	7.5	10	10	1	1	0.07	1	1	0.22

Section 4.1). The essence of the iterative inversion process is adjusting the model parameters to make the model output reflectance as close to the input observation reflectance data as possible. To quantify this closeness, a cost function is adopted (Tarantola, 2005)

$$S(g) = \frac{1}{2}\left\{\sum_{n=1}^{N}\left[\frac{f_n(g) - y_n^{\text{obs}}}{s_n}\right]^2 + \sum_{l=1}^{L}\left[\frac{x_l - x_l^{\text{prior}}}{s_l}\right]^2\right\} \quad (4)$$

where y_n^{obs} and $f_n(g)$ are the observed reflectance and corresponding modeled reflectance values, respectively. The variables σ_n^2 and σ_l^2 are the variances of the observation data and the prior distribution of parameters, respectively. The variables x_l and x_l^{prior} are the parameter value used in the model and the initial value of the parameter, respectively. The number of observation samples is N, and L is the number of parameters. Sequential quadratic programming (SQP), which is an excellent optimization algorithm for solving nonlinear programming problems (Boggs and Tolle, 1995), is adopted to search for the cost function minimum.

According to the sensitivity analysis, the steps for inversion and segmentation of the BRF data in each step are as follows:

(1) use the red band reflectance dataset with backward-viewing directions and large viewing zenith angles to invert RC;
(2) use the red band reflectance dataset with small viewing zenith angles to invert RG;
(3) use the red band reflectance dataset with forward-viewing directions and small viewing zenith angles to invert RZ;
(4) use the near-infrared band reflectance dataset with backward-viewing directions and large viewing zenith angles to invert NIRC;
(5) use the near-infrared band reflectance dataset with large viewing zenith angles to invert b/R, $\Delta h/b$, and NIRZ;
(6) use the near-infrared band reflectance dataset with small viewing zenith angles to invert NIRG;
(7) use the reflectance datasets for May, June, and July to invert nR^2(May), nR^2(June), and nR^2(July), respectively;
(8) use the reflectance datasets of both bands at small viewing zenith angles to invert h/b.

3.4. LAI estimation

In the present study, the structural parameters inverted by GOMS model combined with the canopy heights derived from LiDAR data only provide the tree crown structure information, so the volume of the tree crown, namely the ellipsoid, can be calculated as (Fu et al., 2011)

$$V = \frac{4bR^2\pi}{3} \quad (5)$$

However, the foliage distribution information inside the crown is still unknown. Because the structural parameters are obtained for each tree in sample B, and the LAI is obtained for the sample, the average foliage area volume density (FAVD), μ, of this sample can be defined as

$$\mu = \text{LAI}\frac{A}{\sum_{i=1}^{i=m} V_i} \quad (6)$$

where LAI is the LAI of the sample B forest canopy, A is the area of the sample, and V_i is the ith crown volume inside the sample. Note that it is assumed that the average FAVD is uniformly distributed across the crowns of the sample plot. If most of the trees in a study area, including the sample plot, belong to the same species, the LAI of the forest canopy of the pixel can be calculated as

$$\text{LAI} = n \cdot V \cdot \mu = \frac{\mu 4bnR^2\pi}{3} \quad (7)$$

where n is the number of crown centers distributed in a pixel per m^2, and nR^2 can be directly obtained by GOMS model. The vertical half axis of the ellipsoid b is related to h/b and H.

4. Results

4.1. Canopy height of LiDAR coverage

To evaluate the variable window-sized local maximum filtering approach, the LiDAR-derived tree heights were compared with the field-surveyed tree heights of 16 subplots within sample B. Fig. 5 shows the locations of the extracted single trees from sample B, which are basically located at the center of the highlighted circles of the CHM. According to the sample B field data, there are 1656 trees in sample B, and the number of extracted trees is 513. This difference is a result of the low point density of the LiDAR system, so trees with a low height or that are beneath a higher tree cannot be detected. As found by Popescu et al. (2002), the number of trees proved to be one of the most difficult parameters to estimate with

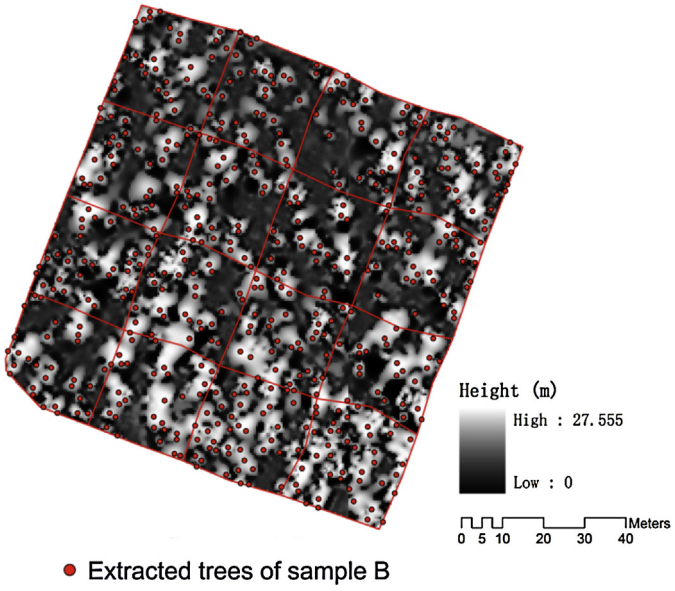

Fig. 5. Locations of extracted single treetops in sample B on CHM.

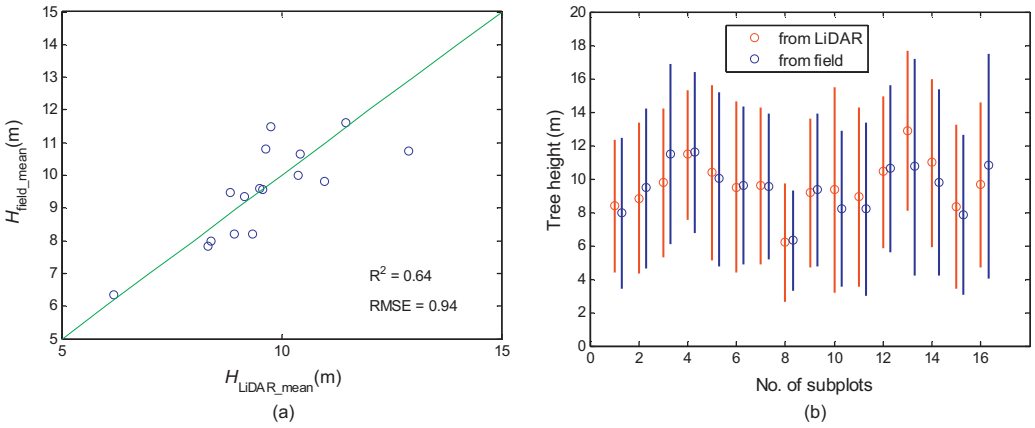

Fig. 6. (a) Comparison of H_{LiDAR_mean} and H_{field_mean} and (b) diagram of canopy height uncertainties of LiDAR-derived and field-surveyed values for the 16 subplots. The red and blue dots represent H_{LiDAR_mean} and H_{field_mean}, respectively, and the bars represent their standard deviations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

airborne LiDAR, and the results show similar R^2 values for pine and deciduous plots (0.26).

The accuracy of H_{LiDAR_mean} was validated using the field measurements of sample B. By dividing sample B into 16 subplots, we calculated the field-surveyed mean canopy height (H_{field_mean}), and its standard error for each subplot, from the field individual tree heights (H_{field_indi}). A comparison of these two types of mean height shows that H_{LiDAR_mean} corresponds well with H_{field_mean}, with an R^2 of 0.64 and an RMSE of 0.94 (Fig. 6a). Fig. 6b shows the mean canopy heights and corresponding standard deviations of LiDAR-derived and field-survey data for the 16 subplots. We find that within each subplot, not only the mean height, but also the height distribution, are quite similar.

4.2. Canopy height of whole study area

The MODIS BRDF/Albedo products of different days were tested to achieve the best results in extrapolating the LiDAR-derived height to the whole study area. Here, the BRDF/Albedo products for eight continuous days near the day of the LiDAR data acquisition were selected (DOY = 177, 185, 193, 201, 209, 217, 225, and 233). Table 2 shows the performances of these 8-day products in estimating canopy height; the RMSE and R^2 of the bootstrapping procedure are given. The probability of correct retrieval is the percentage of tree-height differences between LiDAR-derived and model-estimated heights that are less than 2 m, and the high-quality pixels here are the albedo product pixels, which are flagged as 0 in MCD43A2.

Table 2 shows that the best result, with the highest R^2 and lowest RMSE, was achieved using the day-185 albedo product, and it also shows that the probability of correct retrieval can reach 88.7%. The lowest R^2 (0.13) was calculated when the albedo product quality was at its worst. However, the albedo product of the highest-quality day only achieved an R^2 of 0.39, mainly because the further away from the day on which the LiDAR data were acquired (DOY179), the more uncertainties could arise and decrease the extrapolation accuracy. From Table 2, we can conclude that there are two main factors that influence the extrapolating ability of the albedo product: the quality, and the closeness to the day on which the LiDAR data were acquired.

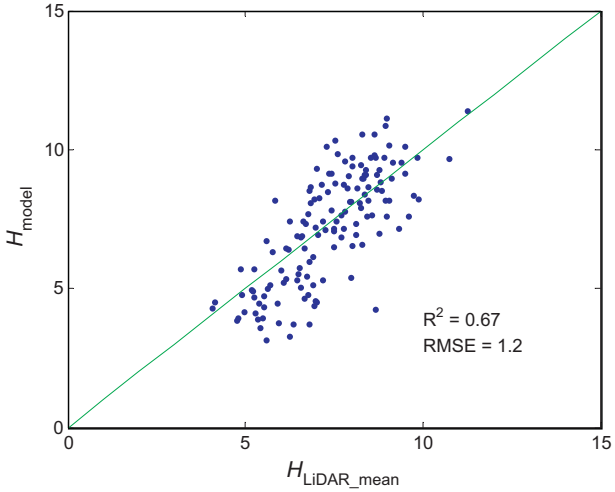

Fig. 7. Model-predicted heights *versus* LiDAR-derived heights.

By repeating the random selection of training and testing data 500 times, the range of the R^2 of the regression is between 0.1 and 0.67 when using the day-185 albedo product, and the probability that R^2 is greater than 0.40 is 42.8%. Fig. 7 shows a graph of H_{LiDAR_mean} *versus* H_{model}, obtained using the best-performing albedo product, that of day 185. This model was then applied to the whole study area to yield the canopy height of the Dayekou forest (see Fig. 8). As shown in Fig. 8, the canopy heights of most of the pixels are less than 20 m, and the apex appears when the canopy height reaches 7.47 m.

5. Inversion of GOMS model

A total of 5173 forest pixels with a resolution of 500 m were selected using the MODIS land-cover product (MCD12Q1), and

Table 2
Tests using MODIS BRDF/Albedo products from different days to estimate canopy height.

DOY	177	185	193	201	209	217	225	233
RMSE	1.74	1.2	1.86	1.94	1.84	1.57	1.90	2.14
R^2	0.47	0.67	0.35	0.13	0.38	0.48	0.39	0.36
Correct retrieval (%)	83.1	88.7	81.7	82.3	88.7	87.3	86.6	82.3
High-quality pixels (%)	92	94	67	49	81	100	88	99

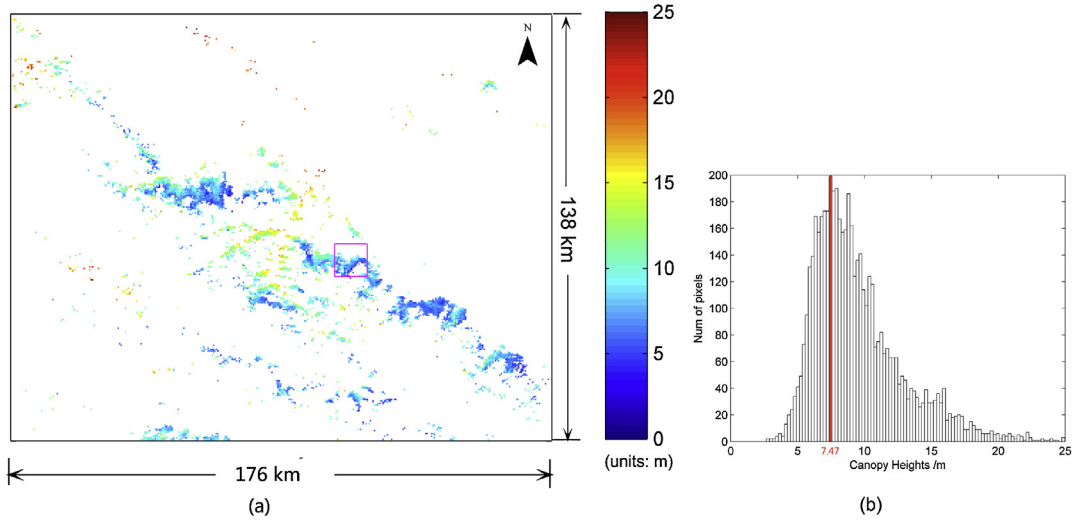

Fig. 8. (a) Canopy height distribution of Dayekou forest and (b) histogram of Dayekou forest canopy heights.

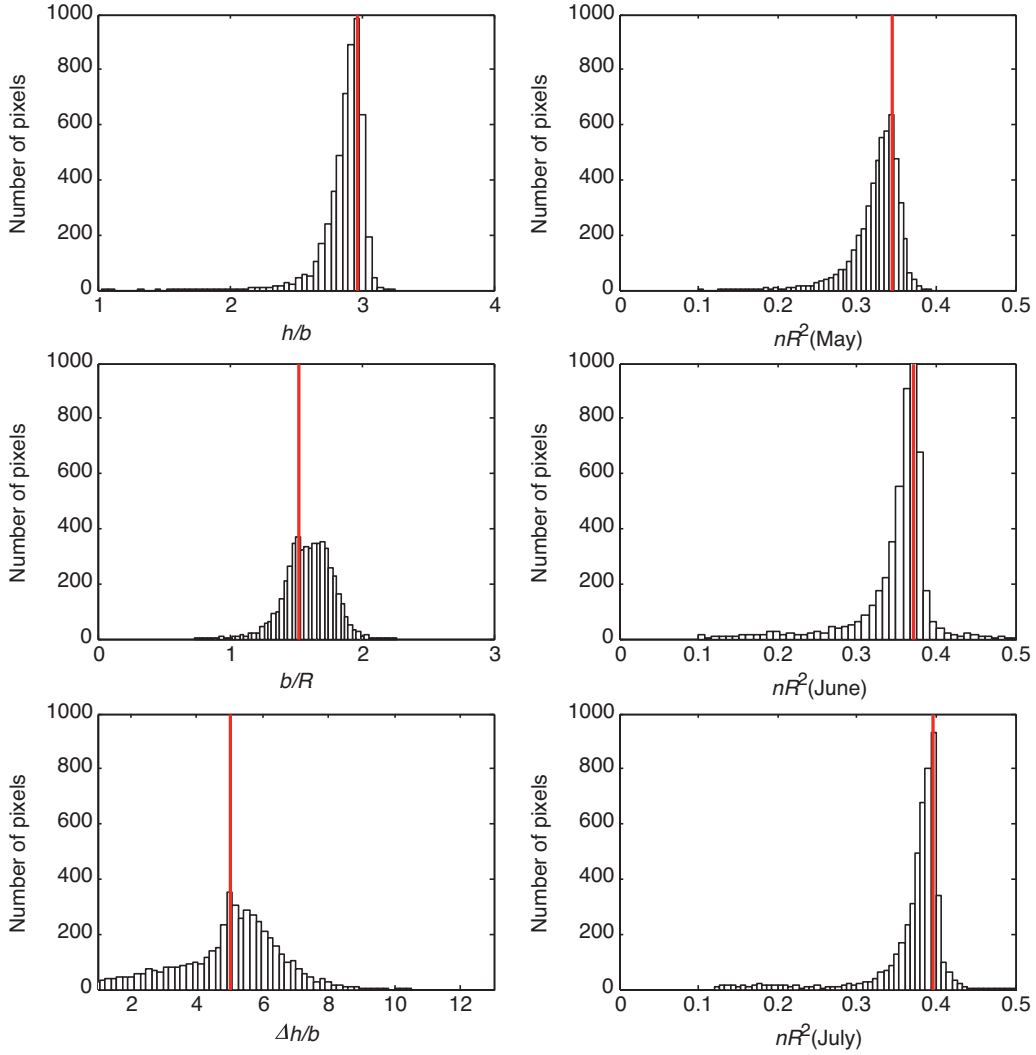

Fig. 9. Histograms of structural parameters inverted by GOMS model (red lines represent the mode of each histogram). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

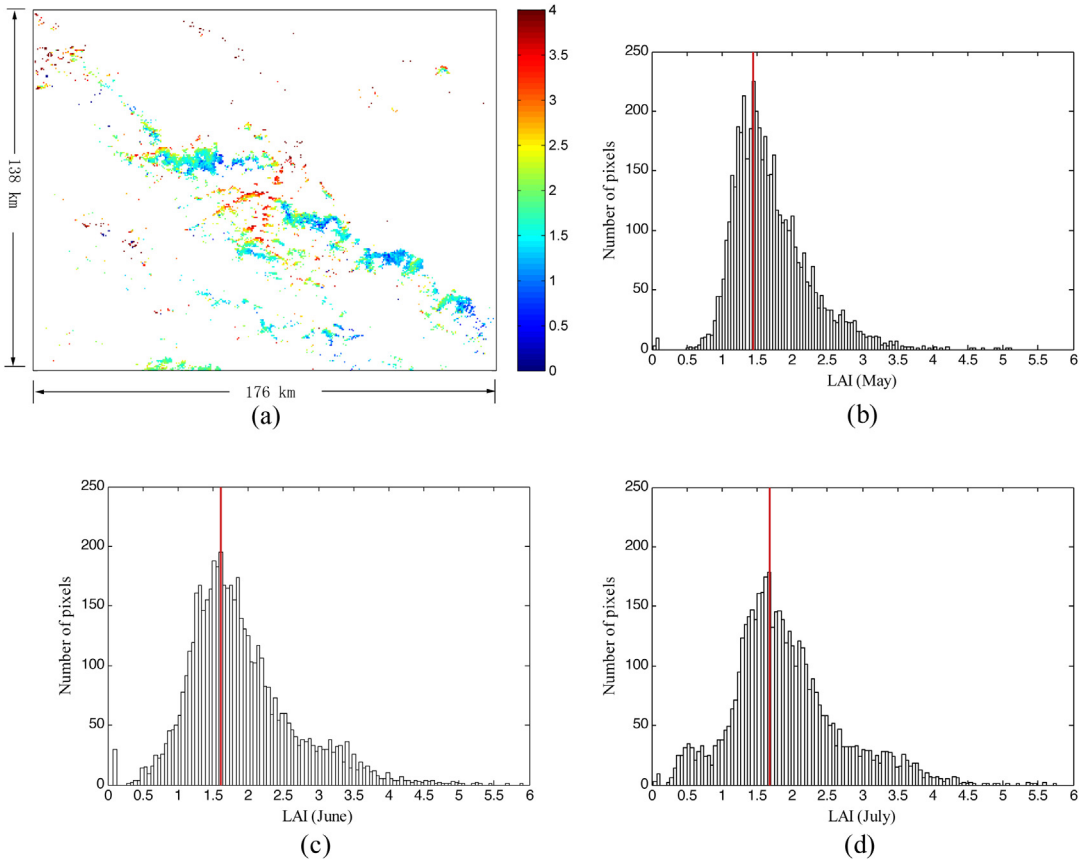

Fig. 10. (a) LAI distribution maps of Dayekou forest in June; (b)–(d) histograms of LAI distributions in May, June, and July respectively. (For interpretation of the references to color in text, the reader is referred to the web version of this article.)

the statistical results for the structural parameters of these pixels inverted by GOMS model are shown in Fig. 9. As shown in the figure, the statistics for the four structural parameters are confined to a reasonable range compared with the Dayekou forest field-survey (Table 1). The red line in each subplot represents the mode of the histogram. Comparing the position of the red line in each subplot of the figure's second column shows that the line shifts to the right with each month. This is caused by the growth habit of *Piceacrassifolia*, which thrives from May to June.

5.1. Estimation and validation of forest canopy LAI

According to the field-inventory structural parameters of each tree in sample plot B, a strong linear regression between the tree height (H) and clear height (h) can be derived, with an R^2 value of 0.97:

$$H = 0.6337h + 0.2692 \tag{8}$$

Therefore, the vertical half axis of ellipsoid b can be calculated as:

$$b = \frac{h}{h/b} = \frac{(H - 0.2692)/0.6337}{h/b} \tag{9}$$

In addition, by using the LAI field measurements and the structural parameters for sample plot B, the average FAVD can be derived according to Eqs. (5) and (6); the result is 0.59 m^{-1}. As in Eq. (7), because the variable nR^2 (for May, June, and July) was retrieved using GOMS model, the average FAVD was calculated. Additionally, b was calculated *from* H and h/b (also retrieved by GOMS model). Therefore, the LAI (for May, June, and July) of the whole study area could be calculated. The results are shown in Fig. 10; the range of LAI values is from 0 to 6. For May, June, and July, a mode (the red line) appears when the LAI value reaches 1.45, 1.61, and 1.67, respectively.

Note that there are two types of tree height (H) in the LiDAR-covered area. One is directly retrieved from the LiDAR data, and the other type is retrieved from the regression model. Here, we call the process of calculating LAI using the LiDAR-derived height Method 1, and that using the model-derived height Method 2.

Based on the statistics for the MODIS LAI product in our study area, 98.8% of the MODIS LAI values are less than 2, whereas most of the estimated LAI values range from 0 to 6. There are two possible reasons for these results: one is that, according to a previous study (Hu and Zhang, 2005), MODIS LAI underestimated the true Dayekou forest LAI value; the other is that they could arise from estimation errors.

For further validation, the LAI measured *in situ* was compared with the estimated LAI and the MODIS LAI product. As described in Section 2.2, the LAI-2000 and TRAC measurements of forest LAI in sample plots A (20 m × 20 m) and B (100 m × 100 m) were obtained. Given the scale mismatch between the *in situ* measurement and the estimated LAI, the estimated LAI must first be decomposed to the *in situ* scale to give a robust validation. According to Eq. (7), from which the LAI is calculated, the relationship between LAIs at different scales (subpixel and pixel) is as follows:

$$\frac{\mathrm{LAI}_{\mathrm{subpixel}}}{\mathrm{LAI}_{\mathrm{pixel}}} = \frac{(b \cdot \pi n R^2)_{\mathrm{subpixel}}}{(b \cdot \pi n R^2)_{\mathrm{pixel}}} \tag{10}$$

where $\mathrm{LAI}_{\mathrm{subpixel}}$ is the LAI value at the subpixel scale, namely the field-sample level, and $\mathrm{LAI}_{\mathrm{pixel}}$ is the LAI value we estimated. The parameters on the right side of Eq. (10) can be obtained from the LiDAR image: b can be calculated from H, and $\pi n R^2$ is considered

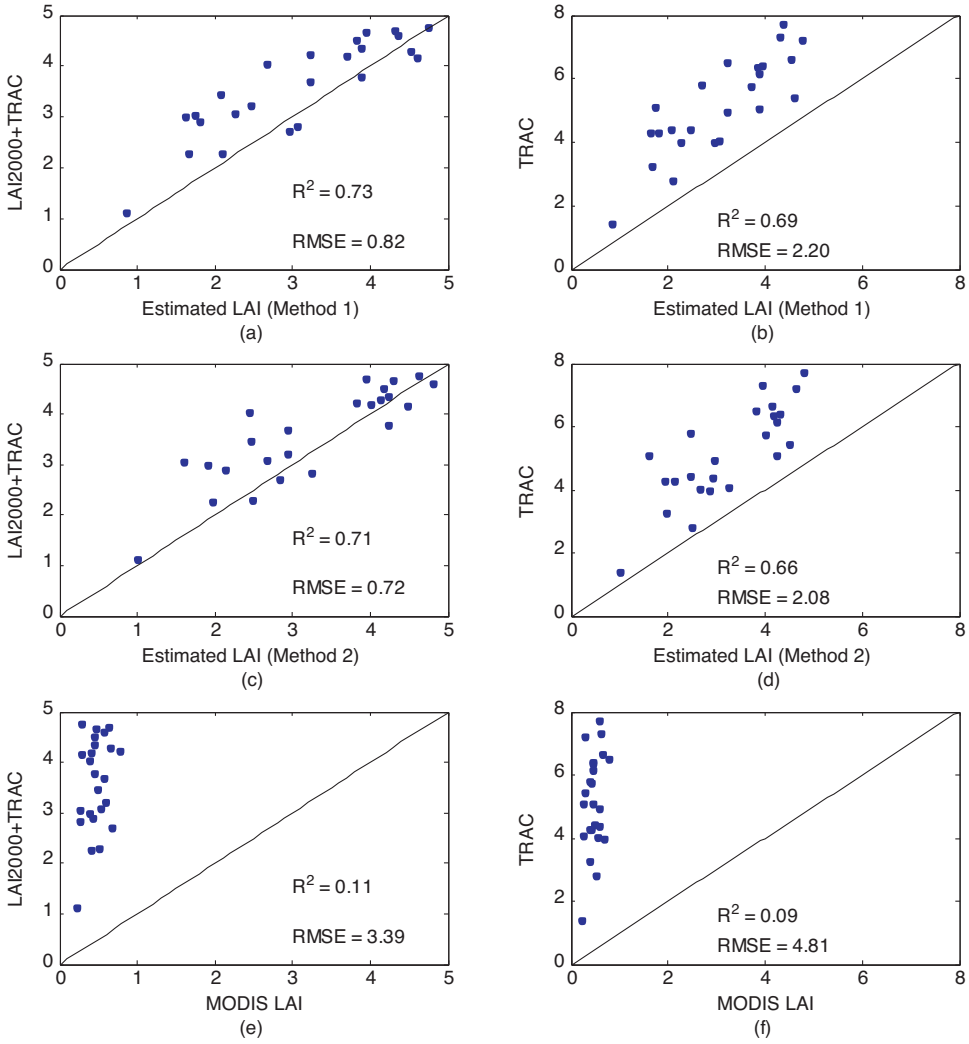

Fig. 11. Validation results for the LAI. (a, c, and e) The true LAI measured by LAI-2000 and TRAC, *versus* the LAI estimated by Method 1, the LAI estimated by Method 2, and the MODIS LAI, respectively. (b, d, and f) The true LAI measured by TRAC *versus* the LAI estimated by Method 1, the LAI estimated by Method 2, and the MODIS LAI, respectively.

to be approximately equal to the canopy cover factor, which can be calculated as the ratio between the number of canopy pixels and the number of corresponding CHM pixels classified with a zero threshold.

The 21 sample plots with their corresponding estimated LAI and MODIS LAI were retrieved by the decomposition process. Fig. 11 shows the validation results for LAI. Fig. 11(a), (c), and (e) show the true LAI measured by LAI-2000 and TRAC, *versus* LAI estimated by Method 1, LAI estimated by Method 2, and MODIS LAI, respectively. Fig. 11(b), (d), and (f) show the true LAI measured by TRAC, *versus* LAI estimated by Method 1, LAI estimated by Method 2, and MODIS LAI, respectively.

Results show that although Method 2 achieves a slightly lower R^2 than Method 1, both of these methods can achieve a high accuracy in retrieving LAI, with highest R^2 values of 0.73 and 0.69, respectively, compared with the true LAI measured by LAI-2000 and TRAC combined, and with LAI measured by TRAC alone. However, MODIS LAI show very low R^2 values compared with the field-measured LAIs (Fig. 11e and f), and the RMSE value between the MODIS LAI and measured LAI (3.39) is higher than that between the Method 1 and Method 2 LAI values (0.82 and 0.72). Fig. 11 indicates that MODIS LAI product underestimated the forest area of our study site. The validation results also show that the combined measurements of LAI-2000 with TRAC are closer to the estimated LAI than the TRAC measurements alone; this result agrees with previous studies showing that TRAC tends to slightly overestimate the LAI for dense canopies. It is recommended that LAI-2000 and TRAC combined measurements be used, to obtain the most reliable LAI measurements (Chen et al., 2006; Xing et al., 2010).

6. Discussion and conclusion

In this study, we investigated forest LAI retrieval, with continuous areal coverage, using LiDAR combined with optical multi-angle directional reflectance data (MODIS, MISR). A unique feature of our study is the capability for continuous LAI retrieval by combining the extrapolated LiDAR forest canopy heights from MODIS data and GOMS model structural parameters. Results show that this method achieved a high accuracy in our study area, with highest R^2 values of 0.73 and 0.69, respectively, compared with the true LAI measured by LAI-2000 and TRAC combined, and by TRAC alone. These results are a significant improvement in the forest LAI estimation compared with those obtained using the MODIS LAI product.

However, it must be noted that the method of extrapolating the forest canopy height using the MODIS multi-angle data only applies to pixels that have relatively high forest coverage. When the pixel also includes bare soil or other surface features, the accuracies of the estimated heights for this pixel worsen and, thus, the accuracy of the forest LAI retrieval is worsened. Moreover, the heterogeneity inside the MODIS or MISR pixels, and the co-registration error

between different sources of data, are other sources of uncertainty in this study (Tan et al., 2006).

Further research is needed to consider the following issues. First, the Dayekou forest study area is a coniferous forest with *Piceacrassifolia* as the predominant tree species; the presented method needs to be tested on more forests with different tree species. Secondly, the FAVD inside the canopy crown is assumed to be constant in the forest; actually, it can vary both vertically and horizontally. Further study on extracting information about the vertical and horizontal distributions of the FAVD from the full-waveform LiDAR data, such as that obtained by the space-borne Geoscience Laser Altimeter System (GLAS), could help to improve LAI estimates. Finally, in this study, LAI was validated using *in situ* measurements by LAI-2000 and TRAC; recent studies have shown that LiDAR-derived high-resolution LAI map can give reliable reference data for validation of a coarse-resolution LAI map (Zhao and Popescu, 2009; Jensen et al., 2011); it is recommended that future work should take this into consideration.

Acknowledgments

This research was supported in part by the Special Funds for Major State Basic Research Project (2013CB733403 and 2007CB714407), the National Natural Science Foundation of China (No. 41171263), the National 863 Program (2013AA12A301, 2012AA12A303), the 2013 Independent Scientific Research Fund Project (104-105582GK) and the project funded by Key Laboratory of Geo-Informatics of National Administration of Surveying, Mapping and Geoinformation of China (201120). The authors would also like to thank anonymous reviewers who gave valuable suggestions that has helped to strengthen the paper.

References

Boggs, P.T., Tolle, J.W., 1995. Sequential quadratic programming. Acta Numerica 4, 1–51.

Chen, G., Wulder, M.A., White, J.C., Hilker, T., Coops, N.C., 2012. LiDAR calibration and validation for geometric-optical modeling with Landsat imagery. Remote Sensing of Environment 124, 384–393.

Chen, J.M., 1996. Optically-based methods for measuring seasonal variation of leaf area index in boreal conifer stands. Agricultural and Forest Meteorology 80, 135–163.

Chen, J.M., Black, T.A., 1992. Defining leaf area index for non-flat leaves. Plant, Cell & Environment 15, 421–429.

Chen, J.M., Cihlar, J., 1995. Plant canopy gap-size analysis theory for improving optical measurements of leaf-area index. Applied Optics 34, 6211–6222.

Chen, J.M., Govind, A., Sonnentag, O., Zhang, Y., Barr, A., Amiro, B., 2006. Leaf area index measurements at Fluxnet-Canada forest sites. Agricultural and Forest Meteorology 140, 257–268.

Demcsak, M., 1997. HDF-EOS Library Users Guide for the ECS Project. Function Reference Guide. Document 170-TP-006-003, vol. 2. Hughes Applied Information Systems, Upper Marlboro, MD, USA.

Diner, D.J., Beckert, J.C., Reilly, T.H., Bruegge, C.J., Conel, J.E., Kahn, R.A., Martonchik, J.V., Ackerman, T.P., Davies, R., Gerstl, S.A.W., 1998. Multi-angle Imaging SpectroRadiometer (MISR) instrument description and experiment overview. IEEE Transactions on Geoscience and Remote Sensing 36, 1072–1087.

Dubayah, R., Sheldon, S., Clark, D., Hofton, M., Blair, J., Hurtt, G., Chazdon, R., 2010. Estimation of tropical forest height and biomass dynamics using LiDAR remote sensing at La Selva, Costa Rica. Journal of Geophysical Research 115, G00E09.

Fu, Z., Wang, J., Song, J.L., Zhou, H.M., Pang, Y., Chen, B.S., 2011. Estimation of forest canopy leaf area index using MODIS, MISR, and LiDAR observations. Journal of Applied Remote Sensing 5, 053530.

Gower, S.T., Kucharik, C.J., Norman, J.M., 1999. Direct and indirect estimation of leaf area index, f_{apar}, and net primary production of terrestrial ecosystems. Remote Sensing of Environment 70, 29–51.

Hu, S., Zhang, W., 2005. A quality assessment of MODIS LAI product in Heihe and Hanjiang River Basins. Remote Sensing Information 4, 22–27.

Jensen, J.L., Humes, K.S., Hudak, A.T., Vierling, L.A., Delmelle, E., 2011. Evaluation of the MODIS LAI product using independent LiDAR-derived LAI: a case study in mixed conifer forest. Remote Sensing of Environment 115, 3625–3639.

Kimes, D., Ranson, K., Sun, G., Blair, J., 2006. Predicting LiDAR measured forest vertical structure from multi-angle spectral data. Remote Sensing of Environment 100, 503–511.

Lefsky, M.A., 2010. A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. Geophysical Research Letters 37, 1–5.

Lefsky, M.A., Cohen, W.B., Parker, G.G., Harding, D.J., 2002. LiDAR remote sensing for ecosystem studies. BioScience 52, 19–30.

Lefsky, M.A., Harding, D.J., Keller, M., Cohen, W.B., Carabajal, C.C., Espirito-Santo, F.D.B., Hunter, M.O., de Oliveira Jr., R., 2005. Estimates of forest canopy height and aboveground biomass using ICESat. Geophysical Research Letters 32, L22S02.

Lefsky, M.A., Keller, M., Pang, Y., De Camargo, P.B., Hunter, M.O., 2007. Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms. Journal of Applied Remote Sensing 1, 013537.

Li, X., Strahler, A.H., 1992. Geometric-optical bidirectional reflectance modeling of the discrete crown vegetation canopy: Effect of crown shape and mutual shadowing. IEEE Transactions on Geoscience and Remote Sensing 30, 276–292.

Li, X., Wang, J., Hu, B., Strahler, A.H., 1998. On utilization of a priori knowledge in inversion of remote sensing models. Science in China Series D: Earth Sciences 41, 580–585.

Li, X., Yan, G., Liu, Y., Wang, J., Zhu, C., 1997. Uncertainty and sensitivity matrix of parameters in inversion of physical BRDF models. Journal of Remote Sensing S1, 113–122.

Liang, S., 2004. Quantitative Remote Sensing of Land Surfaces. John Wiley & Sons, New Jersey.

Lucht, W., Schaaf, C.B., Strahler, A.H., 2000. An algorithm for the retrieval of albedo from space using semiempirical BRDF models. IEEE Transactions on Geoscience and Remote Sensing 38, 977–998.

Neilson, R.P., Drapek, R.J., 1998. Potentially complex biosphere responses to transient global warming. Global Change Biology 4, 505–521.

Popescu, S.C., Wynne, R.H., Nelson, R.F., 2002. Estimating plot-level tree heights with LiDAR: local filtering with a canopy-height based variable window size. Computers and Electronics in Agriculture 37, 71–95.

Popescu, S.C., Zhao, K., Neuenschwander, A., Lin, C., 2011. Satellite LiDAR vs. small footprint airborne LiDAR: comparing the accuracy of aboveground biomass estimates and forest structure metrics at footprint level. Remote Sensing of Environment 115, 2786.

Riaño, D., Valladares, F., Condés, S., Chuvieco, E., 2004. Estimation of leaf area index and covered ground from airborne laser scanner (LiDAR) in two contrasting forests. Agricultural and Forest Meteorology 124, 269–275.

Roujean, J.L., 1992. A bidirectional reflectance model of the earth's surface for the correction of remote sensing data. Journal of Geophysical Research: Atmospheres 97 (20), 20455–20468.

Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Strugnell, N.C., Zhang, X., Jin, Y., Muller, J.P., 2002. First operational BRDF, albedo nadir reflectance products from MODIS. Remote Sensing of Environment 83, 135–148.

Sun, G., Ranson, K., Zhang, Z., 2006. Forest vertical parameters from LiDAR and multi-angle imaging spectrometer data. Journal of Remote Sensing – Beijing 10, 523.

Swatantran, A., Dubayah, R., Roberts, D., Hofton, M., Blair, J.B., 2011. Mapping biomass and stress in the Sierra Nevada using LiDAR and hyperspectral data fusion. Remote Sensing of Environment 115 (11), 2917–2930.

Tan, B., Woodcock, C., Hu, J., Zhang, P., Ozdogan, M., Huang, D., Yang, W., Knyazikhin, Y., Myneni, R., 2006. The impact of gridding artifacts on the local spatial properties of MODIS data: Implications for validation, compositing, and band-to-band registration across resolutions. Remote Sensing of Environment 105, 98–114.

Tarantola, A., 2005. Inverse Problem Theory and Methods for Model Parameter Estimation. Society for Industrial Mathematics, Albert.

Vermote, E., Vermeulen, A., 1999. Atmospheric correction algorithm: spectral reflectances (MOD09). ATBD version 4.

Wang, Z., Schaaf, C.B., Lewis, P., Knyazikhin, Y., Schull, M.A., Strahler, A.H., Yao, T., Myneni, R.B., Chopping, M.J., Blair, B.J., 2011. Retrieval of canopy height using moderate-resolution imaging spectroradiometer (MODIS) data. Remote Sensing of Environment 115 (6), 1595–1601.

Wanner, W., Li, X., Strahler, A., 1995. On the derivation of kernels for kernel-driven models of bidirectional reflectance. Journal of Geophysical Research 100, 21077–21090.

Xing, B., Ju, W., Zhu, G., Li, X., Liu, Y., Zhou, Y., 2010. The comparison of different methods to measure leaf area index of forests in Maoershan Mountain, Northeastern China. In: IEEE, pp. 1–5.

Zhao, K., Popescu, S., 2009. LiDAR-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA. Remote Sensing of Environment 113, 1628–1645.