Forest structure and aboveground biomass in the southwestern United States from MODIS and MISR

Mark Chopping a,⁎, Crystal B. Schaad b, Feng Zhao b, Zhuosen Wang b, Anne W. Nolin c, Gretchen G. Moisen d, John V. Martonchik c, Michael Bull e

a Earth & Environmental Studies, Montclair State University, 1 Normal Ave., Montclair, NJ 07043, United States
b Center for Remote Sensing, Boston University, 675 Commonwealth Ave., Boston, MA 02215, United States
c Department of Geosciences, 104 Wilkinson Hall, Oregon State University, Corvallis, OR 97331, United States
d USDA Forest Service, Rocky Mountain Research Station, 507 25th Street, Ogden, UT 84401, United States
e NASA/Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109, United States

Abstract

Red band bidirectional reflectance factor data from the NASA MODerate resolution Imaging Spectroradiometer (MODIS) acquired over the southwestern United States were interpreted through a simple geometric–optical (GO) canopy reflectance model to provide maps of fractional crown cover (dimensionless), mean canopy height (m), and aboveground woody biomass (Mg ha−1) on a 250 m grid. Model adjustment was performed after dynamic injection of a background contribution predicted via the kernel weights of a bidirectional reflectance distribution function (BRDF) model. Accuracy was assessed with respect to similar maps obtained with data from the NASA Multiangle Imaging Spectroradiometer (MISR) and to contemporaneous US Forest Service (USFS) maps based partly on Forest Inventory and Analysis (FIA) data. MODIS and MISR retrievals of forest fractional cover and mean height both showed compatibility with the USFS maps, with MODIS mean absolute errors (MAE) of 0.09 and 8.4 m respectively, compared with MISR MAE of 0.10 and 2.2 m, respectively. The respective MAE for aboveground woody biomass was ~10 Mg ha−1, the same as that from MISR, although the MODIS retrievals showed a much weaker correlation, noting that these statistics do not represent evaluation with respect to ground survey data. High height retrieval accuracies with respect to averages from high resolution discrete return lidar data and matches between mean crown aspect ratio and mean crown radius maps and known vegetation type distributions both support the contention that the GO model results are not spurious when adjusted against MISR bidirectional reflectance factor data. These results highlight an alternative to empirical methods for the exploitation of moderate resolution remote sensing data in the mapping of woody plant canopies and assessment of woody biomass loss and recovery from disturbance in the southwestern United States and in parts of the world where similar environmental conditions prevail.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

One of the three overarching objectives of NASA's Carbon Cycle and Ecosystems program is to quantify global vegetation productivity, biomass, carbon fluxes, and changes in land cover (NASA, 2006). However, the amount and distribution of terrestrial carbon in forests are dynamic and are still poorly known (Canadell et al., 2007; CCSP, 2007; Heimann & Reichstein, 2008). Furthermore, while much forest remote sensing has focused on deforestation (Broadbent et al., 2008; Goetz et al., 2008), it is thought that forest degradation is at least as important as deforestation in terms of carbon fluxes between the atmosphere and the biosphere (Houghton & Goetz, 2008) and assessing degradation using remote sensing is more challenging. Accurate, annual, wall-to-wall maps of aboveground carbon stocks are needed to help address these issues but cannot be constructed with traditional multispectral remote sensing methods because they do not provide explicit information on canopy three-dimensional structure. Remote sensing of vegetation over large areas using moderate spatial resolution data from across-track scanning sensors such as the NOAA Advanced Very High Resolution Radiometer (AVHRR) and the NASA MODerate resolution Imaging Spectroradiometer (MODIS) has been used mainly to estimate primary production using spectrally-derived metrics of photosynthetic activity, and primarily spectral reflectance band ratios such as the Enhanced Vegetation Index (EVI; Huete et al., 2002) or the Normalized Difference Vegetation Index (NDVI; Tucker, 1979). The advantage of this approach is that it appears to be applicable across biomes and the ratios are quickly and easily calculated. However this approach does not provide useful information on canopy structure

⁎ Corresponding author. Tel.: +1 973 655 7384; fax: +1 973 655 4072.
E-mail address: chopping@pegasus.montclair.edu (M. Chopping).

0034-4257/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2010.08.031
because vegetation indices are composite measures of foliage greenness, fractional vegetation cover, canopy depth, and background optical and physical properties and are thus insensitive to canopy three-dimensional structure; or at least, they are unable to decompose the signal into measurable structural metrics.

Empirical approaches have been used to estimate physical canopy parameters such as height for limited areas using data from NASA’s Airborne Multispectral Imaging Spectroradiometer (AirMISR) and Multispectral Imaging Spectroradiometer (MISR), with some success (Heiskanen, 2006; Kimes et al., 2006; Schull et al., 2007); however empirical methods rely entirely on calibration data that may not be available over large areas and may thus be prone to extrapolation error. The most successful methods developed to date to exploit MODIS data products to map forest aboveground live biomass at large scales use nonparametric modeling methods such as regression trees. For example, Blackard et al. (2008) developed a spatially explicit dataset of aboveground live forest biomass using USDA Forest Service Forest Inventory and Analysis (FIA) program ground measured inventory plot data for the conterminous U.S., Alaska, and Puerto Rico. While successful — correlation coefficients ranged from a high of 0.73 in the Pacific Northwest, to a low of 0.31 in the Southern region — this approach requires extensive plot data and numerous input variables that may not be available elsewhere.

Canopy reflectance modeling is a physical approach to interpreting moderate resolution remote sensing data (e.g., Flasse, 1993; Steinberg et al., 2008; Verstraete, 1994) and can provide structural as well as functional information; however it can be difficult to invert complex models with many adjustable parameters. An approach of intermediate complexity is the use of semi-empirical metrics in which canopy structural information is embedded, obtained by exploiting the directional signal (i.e., by exploiting the variation in spectral radiance with off-nadir viewing and/or illumination into account; Asner et al., 1998; Privette et al., 1997). These include the parameters of semi-empirical bidirectional reflectance distribution function (BRDF) models and ratios thereof, such as the Structural Scattering Index (Gao et al., 2003). These metrics are useful in contexts such as enhancing land cover mapping but they have two major limitations: they do not provide physical measures that can be validated in the field; and they hold less meaning for the user base (e.g., Forest Service, ecologists, and modelers). An ideal approach might therefore be one that is as simple and rapid as possible (c.f., vegetation indices), has less stringent calibration requirements than empirical nonparametric modeling, while providing first-order physical canopy structure measures such as fractional crown cover (dimensionless) and mean canopy height (m). These attributes are conceptually straightforward, can be measured in the field or using an appropriate remote sensing technology such as light detection and ranging (lidar), and are strongly related to aboveground woody biomass (Mg ha$^{-1}$).

The goal of the research described here was to assess simple, fast methods for exploiting the structural information in passive, moderate resolution remote sensing observations that have broad spatial coverage and high temporal sampling rates, to ascertain whether they can be used to map aboveground woody biomass. Retrievals of these canopy parameters were previously obtained over Arizona and New Mexico by adjusting a simple geometric-optical (GO) canopy reflectance model against MISR red band bidirectional reflectance factors (BRFs) mapped onto a 250 m grid (Chopping et al., 2008a). The same method was applied here in order to discover whether the structural signal embedded in accumulated multi-angle data from MODIS red band BRFs accumulated over multiple orbits can also be exploited to retrieve canopy fractional cover and mean height, and with what accuracies. Here, accumulated MODIS red band BRFs were used in an attempt to map woody plant crown cover, mean canopy height, and — via regression on Forest Service estimates — aboveground woody biomass over most of the southwestern United States. In this region trends in stocks and patterns of forest disturbance are known to be driven by climate and changing rapidly; increasing awareness of both direct and indirect effects of recent climate change is reflected in the literature (McKenzie et al., 2009; van Mantgem et al., 2009; Westerling et al., 2006).

2. Methods

Cloud-screened MODIS Collection 5 (V005) red band BRFs and associated solar and viewing zenith and azimuth angles with a nominal 250 m spatial resolution (mapped on a 231.7 m grid in sinusoidal projection) were accumulated from the MOD09 product over DOY 153–164 in June 2002, for MODIS tiles h08v05 and h09v05. This period represents the end of the dry season and was selected for maximum woody plant greenness with largely senescent grasses, for lower cloud cover, and for greater stability in surface conditions, important because MODIS observations must be accumulated. The area covered by these two tiles includes almost all of the southwestern United States, including the states of Arizona, New Mexico, and Nevada, most of California, and parts of Nebraska, Utah, Colorado, and Texas. This area includes large tracts of desert grassland that often show significant woody shrub encroachment, riparian and river valley woodland, and upland forest. These data were processed in the original sinusoidal map projection with a raster cell size of $231.7$ m and were screened for contamination by cloud cover.

The use of an accumulated multi-angle data set obtained via accumulation of MODIS BRFs over many orbits is not straightforward. The most important assumption made in this study is that spectral radiance measurements at the top of the atmosphere from subsequent overpasses represent the same area on the ground, as the ground-projected instantaneous field-of-view (GIFOV) can vary by a factor of three at the most extreme viewing angles. This assumption is a first-order approximation and is adopted in order to avoid the need to apply sophisticated methods for reassigning the data, or operations that could degrade quality — such as resampling the data to a coarser scale and then mapping back to a 250 m grid — and is only partly mitigated by Tobler’s first law of geography (“Everything is related to everything else, but near things are more related than distant things.”), since in the western US landscapes can be highly heterogeneous. However, the data used were processed by standard MODIS algorithms that result in the MOD09 product resampled onto a 231.7 m Sinusoidal projection grid and our results are produced on this standard grid. Any analysis that requires the use of MODIS data from more than a single overpass — such as applications that exploit the temporal domain — would be subject to the same limitations. Other limitations of the accumulation approach are the assumptions that the surface does not change importantly during the accumulation period and that the effectiveness of surface retrievals (corrections for atmospheric attenuation) is similar for all overpasses.

The MODIS red band MOD09 BRFs were interpreted through a simple GO model. GO models are able to resolve statistical distributions of discrete objects within an instrument’s GIFOV (Chen et al., 2000; Li et al., 1995; Strahler et al., 2005). Simple GO models treat the surface as an evenly distributed assemblage of discrete objects of equal radius, shape and height. A tree or shrub crown is represented by a spheroid with its center located at a specified mean height above a (nominally diffuse scattering) plane (Fig. 1). These models predict the top-of-canopy reflectance as a function of important canopy physical parameters (plant number density, foliage volume, mean canopy crown height, radius, and crown shape, background brightness and anisotropy) and illumination-viewing geometry. BRFs are modeled as a linear combination of the contributions from sunlit and viewed, and shaded and viewed crown and background components (Li & Strahler, 1985), as in Eq. (1):

$$R = G.k_z + C.k_c + T.k_d + Z.k_z$$

(1)
where \( R \) is bidirectional spectral reflectance; \( k_G, k_C, k_T \) and \( k_Z \) are the GO modeled proportions of sunlit background, sunlit crown, shaded crown and shaded background, respectively; and \( G, C, T, \) and \( Z \) are the contributions of the sunlit background, sunlit crown, shaded crown, and shaded background, respectively. Although GO models have been used with nadir-spectral remote sensing data such as those from Landsat (Peddle et al., 2003), they are particularly appropriate for exploiting solar wavelength remote sensing data acquired at differing viewing and/or illumination angles because the proportions of sunlit and shaded crown and background in the remote sensing instrument GIFOV vary with viewing and illumination geometry.

The simple geometric model (SGM), a GO model incorporating a dynamic background and an in-crown volume scattering term, was used. The functions used in this model are derived from prior work on kernel-driven bidirectional reflectance distribution function (BRDF) models (Roujean et al., 1992; Schaaf et al., 2002; Strahler et al., 1996; Walthall et al., 1985; Wanner et al., 1995) and are given in a number of previous publications (Chopping et al., 2006, 2008a, 2008b). It is formulated as Eq. (2):

\[
R = C_{\text{Walthall}}(\theta_i, \phi, \psi) k_G(\theta_i, \phi, \psi) + C_{\text{Ross}}(\theta_i, \phi, \psi) k_C(\theta_i, \phi, \psi) (2)
\]

where \( \theta_i, \phi, \psi \) are the view zenith, solar zenith and relative azimuth angles, respectively; \( k_G \) and \( k_C \) are the calculated proportions of sunlit and viewed background and crown, respectively; \( C_{\text{Walthall}} \) is the background contribution from the Walthall model (Walthall et al., 1985); and \( C_{\text{Ross}} \) is the simplified Ross turbid medium approximation for plane parallel canopies (Ross, 1981). The shaded components \( T \) and \( Z \) are discarded; they are assumed black. \( k_G \) and \( k_C \) are calculated exactly via Boolean geometry for the principal and perpendicular planes and approximated away from these; they are provided by Eqs. (3) and (4), respectively:

\[
k_G = e^{-\lambda \pi^2 \left( \sec \theta_i' + \sec \phi' - \alpha(\phi, \psi) \right)}
\]

\[
k_C = \left( 1 - e^{-\lambda \pi^2 \sec \theta_i'} \right) \frac{1}{2} \left( 1 + \cos \epsilon' \right)
\]

where \( \epsilon' \) is the transformed scattering phase angle given by Eq. (5):

\[
\cos \epsilon' = \cos \theta_i' \cos \phi' + \sin \theta_i' \sin \phi' \cos \psi.
\]

\( \lambda \) is the number density of objects; \( r \) is the average radius of these objects; \( O \) is the overlap area between the shadows of illumination and viewing (Wanner et al., 1995); Eq. (6):

\[
O = 1 / \left( r - \sin r \cos t \right) (\sec \theta_i' + \sec \phi')
\]

where \( t \) is a parameter that indirectly expresses the locations of the end points of the line that intersects the shadows of viewing and illumination; this allows \( k_G \) to be expressed in a way that depends only on the value of \( t \) (Wanner et al., 1995). These functions include the parameters \( b/r \) (vertical crown radius/horizontal crown radius) and \( h/b \) (height of crown center/vertical crown radius) which describe the shape and height of the crown. The prime indicates equivalent zenith angles obtained by a vertical scale transformation in order to treat spheroids as spheres (i.e., \( \theta' = \tan^{-1}(h/b) \tan \theta \); Wanner et al., 1995). The model's parameters are mean plant crown horizontal radius, \( r \); crown aspect ratio, \( b/r \); crown height ratio, \( h/b \); plant number density (number per unit area), \( \lambda \); and per-crown leaf area index (LAI); and
the four empirical parameters of the Walthall BRDF model (Walthall et al., 1985). The crown radius and number density parameters are coupled internally; identical results are obtained if fractional cover is maintained but either parameter is varied. However if number density is fixed and mean crown radius is left adjustable then an estimate of fractional cover can be obtained. The $h/b$ and $b/r$ parameters are also coupled, since mean canopy height depends on $b$ and the retrieval of $b$ depends on $b/r$. Similarly, mean canopy height cannot be obtained without knowledge of both $h/b$ and $b$ but if $h/b$ is fixed and $r$ and $b/r$ are left adjustable then estimates of $b$ (and therefore $h$) can be obtained. This depends on the assumption that $r$ and $b/r$ respond to different aspects of the BRF patterns presented to the minimization algorithm: ideally, $r$ should respond (mostly) to fractional cover and $b/r$ (mostly) to shadowing effects.

For the sake of simplicity the GO model makes some important assumptions: diffuse irradiance is ignored and so the contributions of shaded crown and background are discarded, as in kernel-driven BRDF models; leaf reflectance is fixed at 0.09 (rather high for most conifer species); and in-crown LAI is fixed at 2.08 (typical for mesquite shrubs). Only $r$ and $b/r$ are allowed as free parameters on model adjustment: since other parameters are fixed, these are effective rather than physical parameters. However, the coupled dependent terms (e.g., fractional cover and crown center height) are physical parameters: fractional canopy cover is obtained by using the retrieved $r$ value with the fixed $\lambda$, as $1 - \exp(-\lambda \pi r^2)$; and canopy height is obtained via $h/b \times b + b$, where $b = b/r \times r$.

For accurate retrievals of upper canopy parameters (fractional crown cover and aspect ratio) it is critical to provide a priori estimates of the contribution of the background to BRF at the viewing and illumination angles of the observations. The background is defined here as all elements in an instrument instantaneous field-of-view (IFOV) that are not part of the woody plant canopy. In the arid and semi-arid southwestern US the soil and understory background can account for a large proportion of the surface area under mapped IFOVs, typically up to 75% in desert grasslands with well established woody shrub cover. Even though the background contribution is much smaller in forest, it is not negligible (Gemell, 2000) because the mineral soils of this region are bright across the red to near-infrared wavelengths and canopies are frequently open. The background may be complex and might typically contain a mixture of bare soil, crusts, grasses, sub-shrubs, and other understory plants; the method adopted here attempts to avoid this complexity by focusing on estimation of the bulk background scattering behavior.

To enable GO model adjustment for fractional cover and crown center height, the contribution of the background should ideally be approximated for each location, for the angular illumination and viewing configurations of the observations a priori (before the $r$ and $b/r$ parameters are adjusted). This is attempted here using regression of the background model parameters on LiSparse-RossThin model red channel data (Wanner et al., 1995) isotropic, geometric, and volume-scattering kernel weights. Kernel weight maps were obtained via inversion against the same MODIS red band BRFs used for GO model inversion, using the Algorithm for Modeling Bidirectional Reflectance Anisotropies of the Land Surface (AMBRALS) code (Strahler et al., 1996), with the objective of minimizing the absolute Root Mean Square Error (RMSE) between the model and data. Obtaining the regression coefficients that will be used to predict the background model parameters for each location is not straightforward sing MODIS data. Since it is difficult or impossible to find areas of $250 \times 250$ m ($62,500$ m²) that do not include some woody plants, the background contribution in the MODIS viewing plane must be extracted for a set of locations using the GO model and estimates of $\lambda$, $r$, and $h$, or some combination thereof. To this end, the parts of the MODIS tiles corresponding to the US Department of Agriculture (USDA), Agricultural Research Service (ARS) Jornada Experimental Range were resampled using bilinear interpolation to the Universal Transverse Mercator map projection, WG84 spheroid, datum zone 13, with a grid interval of 250 m, to match shrub canopy statistics extracted from 1 m spatial resolution panchromatic IKonos imagery. These sites were chosen in shrub-dominated desert grassland because the background signal is stronger than in forest; a greater proportion is exposed. Backgrounds were extracted for a small number (10-15) of locations by finding the best-fitting set of Walthall model parameters, using numerical methods with initial values set to those corresponding to a field-measured BRDF for sand. The set of sites used was refined by assessing the significance of the regressions and removing outliers.

The relations established between LiSparse-RossThin model red band kernel weights and background model parameters allow prediction of the background contribution at the illumination and viewing angles of each MODIS observation. This method has been shown to be feasible in a first approximation using MISR BRFs, with low extrapolation error over large areas (Chopping et al., 2008a). For the V005 MODIS data used here, the extraction of backgrounds was found to be much more problematic than with the MISR data. This may be because: the MODIS viewing plane is closer to the solar principal plane at these latitudes leading to greater variability in BRDF shape; multi-angle data are accumulated; the angular sampling is not always complete and/or consistent; and/or the background model used has too many parameters (four), leading to instability in model fitting. Typical fitted Walthall model backgrounds, resulting modeled background contributions, and observed BRFs are shown for three calibration sites in Fig. 2, along with the corresponding 1 m panchromatic IKonos image subsets and derived shrub maps.

In view of these difficulties, an alternative method for background extraction was explored that uses the US Forest Service (USFS) Forest Inventory Analysis (FIA)-based Interior West forest map series (henceforth FIA-IW) instead of IKons-derived shrub canopy statistics. The FIA-IW maps are a suite of raster products with a spatial resolution of 250 m produced using a nonparametric modeling framework (using classification and regression trees) that makes use of Forest Inventory Analysis survey data, soils, topographic, MODIS vegetation index, MODIS MOD14 Vegetation Continuous Fields, and climate variables as predictor variables (Blackard & Moisen, 2005; Blackard et al., 2008; Ruenenacht et al., 2004). The outputs include maps of aboveground biomass, crown cover, and forest height that were also used as reference data in this study. In this background extraction method, the mean crown radius parameter was specified using crown cover data extracted from the FIA-IW maps for seven contrasting forest locations distributed across New Mexico and Arizona. The data were selected to represent a range of forest cover and height values (avoiding very high cover situations in which the background signal would be too weak) and screened for topographic shading using a hillshade map derived from a digital elevation model constructed through interpolation of Shuttle Radar Topography Mission elevation data (Farr et al., 2007). The background model parameters were fitted in the same way as before.

Three SGM inversion runs were performed for the two MODIS tiles: one using the V005 data with canopy statistics estimated from the FIA-IW maps (termed _mod11); one using the V005 data with IKonos-derived canopy statistics for sites in the Jornada Experimental Range (termed _mod9); and one using Collection 4 (V004) data with more complete angular sampling and the same IKonos-derived canopy statistics (termed _mod3). Each 4800×4800 MODIS tile required 23,040,000 model inversions since inversions were performed wall-to-wall (i.e., including both forest and non-forest areas); each run therefore required a total of 46,080,000 inversions. The results are referred to collectively as MODIS-SW. The inversion protocol was the same as that used in previous MISR mapping work in New Mexico/Arizona (henceforth MISR-NM/AZ) and Colorado (henceforth MISR-CO) using site-based backgrounds. In all cases, the SGM was adjusted against the red band data in all available viewing/illumination configurations (up to nine looks from MISR and more than 13 from MODIS) using the Praxis algorithm (Brent, 1973; Powell, 1964) with min(−RMSE) as the
objective function and no constraints or weighting of the error terms. The crown LAI, $\lambda$, and $h/b$ model parameters were fixed at 2.08, 0.012 (740 plants per grid cell), and 2.00, respectively, with $r$ and $b/r$ left as free parameters and set to initial values of 0.25 and 0.2, respectively, for the MODIS-SW and MISR-NM/AZ runs, and to typical ranges of values for the MISR-CO runs (140 combinations: seven different backgrounds; $r$ from 3.0–6.0 in increments of 1.0; and $b/r$ from 0.5–2.5 in increments of 0.5). Model inversion runs proceeded by submitting BRFs and kernel weights to the minimization code and accumulating the results (model-fitting RMSE, $r$, $b/r$, fractional cover, and $h$) in an output file. Aboveground woody biomass estimates were obtained via regression on retrieved cover and height, using estimates from the FIA-IW biomass map.

GO model inversion results previously obtained using MISR data over New Mexico/Arizona (MISR-NM/AZ; Chopping et al., 2008a) and over Colorado (MISR-CO; Chopping et al., 2009) are used here for comparative purposes. MISR data products used include Level 1B2 MI1B2T Terrain-projected Spectral Radiance; Level 1B2 MI1B2GEOP Geometric Parameters; Level 2 MIL2ASAE Aerosol; Level 2 MIL2ASLS Land Surface; and the MIANCAGP Ancillary Geographic products. The MI1B2T product is the terrain-projected top-of-atmosphere spectral radiance with a nominal 1.1 km spatial resolution in the off-nadir, non-red bands and a nominal 275 m spatial resolution in the nadir multi-spectral and off-nadir red bands (Diner et al., 1999). For the MISR-NM/AZ mapping, MISR spectral radiances were acquired for twelve Terra satellite overpasses in late May and early June 2002 and surface reflectance estimates were estimated using the MISR aerosol data, assuming a desert aerosol type. GO model retrievals from multiple inversions over each location were selected on the basis of minimum model fitting error, as this was found to be monotonically related to accuracy (Chopping et al., 2008a). For the MISR-CO mapping the spectral radiances for a single overpass were converted to bidirectional reflectance factor (BRF) estimates via regression on the Land Surface product BRFs using MISR Toolkit routines developed at NASA’s Jet Propulsion Laboratory.

Reference data were extracted from the FIA-IW aboveground biomass, crown cover, and canopy height maps for the same 1063 random locations used to assess the MISR-NM/AZ results (Chopping et al., 2008a). The FIA-IW maps were not intended for validation purposes but are the most comprehensive and extensive contiguous geospatial data available that include forest height estimates. For this study the maps were resampled to match the ~231.7 m MODIS Sinusoidal grid. The accuracies of the FIA-IW estimates and the MISR-CO retrievals have been assessed with respect to canopy heights calculated from discrete return lidar data and cover estimates from orthophotography acquired as part of the Cold Land Processes Experiment (CLPX) in the Rocky Mountains in Colorado (Chopping et al., 2009; Miller, 2003). The lidar canopy height estimates over MISR raster cell areas were calculated by generating a 2 m spatial resolution digital elevation model (DEM) via interpolation of the lidar ground elevations and subtracting these values from the corresponding vegetation elevations (with checks for anomalies, i.e., where either of the elevations was out of range). This produced a 2 m spatial resolution “tree height” raster image from which averages for all MISR raster cells fully within the lidar-derived “tree height” rasters were extracted with all
zero values ignored). Since the horizontal spacing of the lidar shots was ~1.5 m and crowns are sometimes more than 3 m in diameter (but rarely much more), the resulting means may on occasion be biased downwards by returns from the edges of crowns; however returns from low trees and shrubs that are quite prevalent in many places have a far greater influence on the mean. No effort was made to filter out the heights for these lower plants and this is appropriate because the GO model-derived values include all woody plants (trees and shrubs). For the CLPX area the FIA-IW maps are known to contain some anomalies but the method has been shown to be generally reliable over large areas (Blackard et al., 2008). For the _mod3 and _mod9 runs crown cover and mean height are retrieved from the GO model with no rescaling and they are retrieved independently from the reference data.

3. Results and discussion

Model fitting RMSE was generally low with means between 0.016 and 0.035 and was clearly lower for forested than non-forested regions. RMSE is relatively large (>0.05) in some places and particularly in east-central Texas (Fig. 3); this is probably owing to the need to apply a different background for this region and may reflect differences in vegetation type and soil optical and structural properties. The resulting MODIS-SW fractional cover and mean height maps are shown in Figs. 4 and 5, along with the corresponding USFS Interior West maps that only include forest. Spatial distributions of forest canopy cover and mean canopy height values retrieved using the accumulated MODIS red band multi-angle BRFs show reasonable matches with those in the USFS FIA-IW maps. It can be seen that the retrievals of fractional cover and mean canopy height were less reasonable where model fitting RMSE is high, with more extreme values for both cover and height in the _mod11 and _mod9 run outputs in east-central Texas; a different set of background prediction coefficients may therefore be required for this region.

The best MODIS-SW retrievals in terms of spatial distributions and with respect to the reference data were obtained for run _mod3, i.e., using the set of background prediction coefficients derived against Collection 4 (V004) MODIS data for a slightly earlier period than the V005 data used to invert the SGM. This is most likely because the V005 data used to calibrate the background for run _mod9 were screened very conservatively for poor aerosol retrievals so that observations in the near-nadir to 40° backscattering viewing zenith angle range were removed from the data (Fig. 2a–c). This potentially impacts the extraction of backgrounds, the quality of the LiSparse-RossThin model kernel weights used in background prediction, and the eventual SGM inversions through which retrievals of forest cover and height are made possible. The _mod11 retrievals appear to be less accurate with respect to the other MODIS-SW runs and the FIA-IW maps: cover is overestimated for parts of southern Nevada and height is underestimated in forest in northern Arizona and in the Colorado Rockies.

The MODIS-SW _mod3 (n=895) and MISR-NM/AZ (n=576) estimates of forest fractional cover and mean height both show general agreement with the reference data extracted from the FIA-IW maps, with MODIS mean absolute errors (MAE) of 0.09 and 8.4 m respectively; and MISR MAE of 10.0 and 2.2 m, respectively (Table 1). The mean absolute error values for estimates of aboveground woody biomass via regression against USFS estimates were ~10.1 Mg ha$^{-1}$ for both MODIS and MISR, although the MODIS relationships were much weaker and height retrievals were biased (Fig. 6b). MODIS fractional cover and height retrievals show linear relationships with the USFS map data, with some unexplained scatter, particularly in the height retrievals (Fig. 6b). This may be an artifact of topography, or owing to inadequate prediction of the background contribution stemming from the inconsistent angular sampling of the MODIS data that affects background extraction and the LiSparse-RossThin kernel weights used in background prediction. Extending the period over which MODIS data are accumulated and/or relaxing screening for poor observation quality (based on aerosol or cloud flags) may alleviate this problem. There is a tradeoff between ensuring data quality and providing an angular sampling that better describes surface reflectance anisotropy; further research is required to assess the optimal settings for data screening. Nevertheless, there is a strong correspondence between MODIS/GO, MISR/GO, and USFS height estimates.

There are several limiting factors in GO model inversion against multangle data sets synthesized from MODIS data. The most important one for sparse, arid environments is the requirement to provide the background red wavelength reflectance magnitude and anisotropy, since here the contribution from bright, exposed backgrounds is very large at many angular configurations. As mentioned in the Methods section, difficulties were encountered in finding optimal sets of regression coefficients to predict the background a priori from MODIS data. Recent research into retrieving forest backgrounds confirm and support the finding that retrieval of the background reflectivity is problematic at the typical range of viewing and illumination angles of MODIS data, while those of MISR provide acceptable results (Pisek & Chen, 2009). Another limitation of the approach as presented here is that changes in leaf spectral reflectance owing to

Fig. 3. Map of model-fitting absolute root mean square error (RMSE) for SGM inversions of tiles h08v05 and h09v05 (sinusoidal projection), run _mod3. Values close to 0.1 indicate very poor model fitting. Yellow lines are state boundaries. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Please cite this article as: Chopping, M., et al., Forest structure and aboveground biomass in the southwestern United States from MODIS and MISR, Remote Sensing of Environment (2011), doi:10.1016/j.rse.2010.08.031
(e.g.,) widespread mortality from pine beetle infestation will impact GO model retrievals. However, a change from living (green) to dead (red-brown) leaves is likely to result in a reduction in retrieved cover and height, as red band reflectance will increase (recent MISR/GO biomass mapping runs indicate that dead trees in affected parts of Colorado and Wyoming are treated as absent, with biomass losses of more than 150 Mg ha$^{-1}$ in 2009 over 2000; Chopping et al., 2010); moreover, there is no way to integrate easily this information a priori.

Since the concept of retrieving canopy heights from mono-spectral moderate resolution Earth observation data may seem implausible in view of the prevailing multi-spectral remote sensing paradigm and the simplicity of the GO model; and since the comparisons presented thus far show only assessments against USFS maps made using empirical methods, two lines of supporting evidence from adjustment of the model against MISR bidirectional reflectance factor data are provided here: first, the spatial correspondence between retrieved $r$ and $b/r$ parameters and known vegetation distributions; and second, a summary of assessments of MISR/GO height retrievals against estimates from high resolution discrete return lidar data.

Maps of mean crown shape and mean crown radius for the USDA, ARS Jornada Experimental Range in southern New Mexico show good matches with a 1998 Long Term Ecological Research (LTER) vegetation map that includes dominant shrub species; and with known vegetation distributions (Fig. 7). In particular, more oblate crown shapes are retrieved for mesquite-dominated areas, more spherical shapes for creosotebush-dominated areas, and more prolate shapes for woodland, while mean crown radius retrievals are highest for riparian pecan orchards near the Rio Grande ($\leq 7$ m), high for
upland juniper and mesquite-dominated areas (\( \leq 3 \) m), intermediate for creosotebush- and tarbush-dominated areas (\( \leq 2 \) m), and lowest for playa grasses with few woody plants (\( \leq 0.04 \) m) (Fig. 7). The linear boundaries between creosotebush (cyan), tarbush (blue/purple), and the town of Organ (purple) can also be seen in the south-east quadrant of the \( b/r \) ratio map, immediately to the west of the North arrow. Note that these matches are obtained even though \( r \) and \( b/r \) are retrieved as effective rather than physical parameters (since the \( \lambda \) and \( h/b \) ratio parameters are fixed); they thus improve confidence in model operation.

Recent assessments of MISR/GO height retrievals against high resolution (~2 m) discrete return lidar data acquired as part of the CLPX campaign in April and September 2003 were obtained as part of a research initiative aimed at developing methods to reduce uncertainties in satellite-derived maps of fractional snow cover by accounting for canopy cover and height. The lidar data correspond to means for mapped 250 m\(^2\) MISR observations for 57 forest sites in the Colorado Rocky Mountains. Model inversion runs were completed for 140 combinations of dynamic backgrounds and initial \( b/r \) values, with the same inversion protocols that were used in this study. For all runs with reasonable backgrounds and an initial \( b/r \) value of \(<2.0\), RMSE distributions vs. the April and September lidar mean heights were centered around means of 2.5 m and 3.7 m with maxima of 5.5 m and 5.6 m, respectively, while \( R^2 \) distributions were centered around means of 0.4 and 0.7 with minima of 0.10 and 0.24 (Table 2). With some care taken in background extraction and selection of initial values, it is possible to achieve an RMSE of \( \approx 3.0 \) m and \( R^2 \) of \( \approx 0.7 \) (Chopping et al., 2009). The absolute errors are roughly comparable to those achieved with lidar instruments over areas with marked topography or mixed crown shapes and at different times of year (Hyde et al., 2007; Lefsky et al., 2008; Pang et al., 2008; Sun et al., 2008), although the relationships are weaker. Because fractional cover and mean height are retrieved from GO model adjustment independently from the reference data with no training or rescaling, it is unlikely that these results could be obtained unless the GO model is able to explain the observed data, implying that the correspondence between MODIS-SW, MISR-NM/AZ, and USFS FIA-IW height estimates is not spurious.

### Table 1

<table>
<thead>
<tr>
<th>Fractional cover (Dimensionless)</th>
<th>Mean height (Meters)</th>
<th>Aboveground woody biomass</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODIS</td>
<td>MISR</td>
<td>MODIS</td>
</tr>
<tr>
<td>Mean relative error (%)</td>
<td>26</td>
<td>30</td>
</tr>
<tr>
<td>Mean absolute error (%)</td>
<td>0.09</td>
<td>0.10</td>
</tr>
<tr>
<td>2 Standard deviations</td>
<td>0.19</td>
<td>0.30</td>
</tr>
<tr>
<td>Means</td>
<td>0.44</td>
<td>0.48</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.55</td>
<td>0.78</td>
</tr>
<tr>
<td>RMSE</td>
<td>0.10</td>
<td>0.12</td>
</tr>
</tbody>
</table>

### Fig. 6

Retrieved fractional canopy cover and height and estimated aboveground woody biomass vs. reference data from a 2005 Forest Service map series for the Interior West (random samples). (a)–(c) MODIS-SW cover, height, and biomass, respectively (\( N = 895 \)); and (d)–(f) MISR-NM/AZ cover, height, and biomass, respectively (\( N = 576 \)). Note that statistics refer to the GO model inversion results vs. data derived from other moderate resolution remote sensing products as well as ground data.
4. Conclusions

The results presented here show that in spite of many assumptions and approximations, a simple GO model can be inverted against moderate resolution red band multi-angle reflectance data to retrieve reasonable distributions of forest canopy cover and mean canopy height over large areas. As far as is known, this is the first time that the structural signal in an accumulated MODIS multi-angle data set has been exploited in an attempt to map canopy height over large areas, or indeed at all. Retrievals using MISR data are more accurate with respect to the FIA-IW estimates than those from MODIS but the latter were achieved using data sets with part of the angular domain obscured (missing observations in the 5–40 backscattering region) and sets of background regression coefficients that are thought to be far from optimal. The difficulty of isolating the background signal at MODIS observation geometries is thought to be the major limitation for the exploitation of accumulated MODIS multiangle data sets with a GO model. This has also been noted by other researchers (Canisius & Chen, 2007), although forest background reflectance brightness has been successfully extracted over larger areas using data from MISR (Pisek & Chen, 2009; note that for GO model inversion both magnitude and anisotropy are required to provide estimates of the contribution of the background at all illumination and viewing geometries; this is somewhat more challenging than estimating only brightness). This raises the possibility that MISR-derived background reflectance magnitude and anisotropy maps could be used to enable GO model inversions with MODIS data, although it is not clear what advantages this might provide over the use of MISR for both background retrieval and GO model inversion.

Table 2
MISR/GO height retrieval RMSE and correlation vs. lidar-derived estimates.

<table>
<thead>
<tr>
<th></th>
<th>RMSE (m)</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Min</td>
</tr>
<tr>
<td>April lidar</td>
<td>2.0</td>
<td>1.4</td>
</tr>
<tr>
<td>September lidar</td>
<td>3.6</td>
<td>2.8</td>
</tr>
</tbody>
</table>
This may also have implications for practical implementation of other canopy reflectance modeling methods and particularly the very promising spectral invariants approach that is based on estimates of recollision and escape probabilities (Huang et al., 2007, 2008; Knyazikhin et al., 2005; Lewis & Disney, 2007; Schull et al., 2007; Stenberg, 2007). The original spectral invariants approach makes the assumption that the vegetation canopy is bounded from below by a non-reflecting (black) surface; this is clearly violated in the real world and so efforts must be made to address how the canopy and background signals might be isolated.

An important caveat is that the results presented here cannot be considered as validated because (1) the data extraction was for a random sample of 1063 points in Arizona and New Mexico that are not completely representative of the mapped area even though they span a large set of conditions over more than 200,000 km²; and (2) the accuracy of the FIA-IW estimates may be much lower than required. The question of the required accuracy of aboveground live woody biomass estimates from near-future Earth observation systems has largely been resolved by the NASA Vegetation Structure Working Group: an accuracy of 10% in biomass (or 10 Mg C ha⁻¹) is sought as this would yield a first estimate of the fractions of forest area disturbed, recovering, and unchanged in different regions of the earth and thus advance our understanding of forest dynamics (NASA, 2008). Uncertainty in biomass currently contributes ~100% uncertainty to estimates of carbon emissions (Houghton, 2005, 2008). Since it is not feasible to acquire ground data for the validation of moderate resolution remote sensing results, canopy height estimates from lidar instruments such as NASA’s Laser Vegetation Imaging Sensor (LVIS) provide the best means of evaluating retrievals. On the other hand, retrievals are unlikely to be completely spurious because MODIS/GO mean crown shape and radius distributions correspond well with known vegetation type distributions and match high resolution discrete return lidar heights reasonably well (Chopping et al., 2009).

The main limitation of the GO model approach is that it is unlikely to be suitable for closed canopies and where the background is dark, leading to lack of contrast (Pinty et al., 2002). However it might be applied for mapping and monitoring woodlands, savannas, and shrub-dominated zones in arid and semi-arid regions worldwide. Application over very large areas with differing backgrounds requires constructing a sparse grid of background regression coefficients. Extracting the necessary data is not straightforward, requiring at a minimum estimate of crown cover from high resolution imagery; calibration for dynamic background prediction might also benefit from the use of lidar height estimates to set the b/r parameter rather than using approximations, as here. These results have implications for using moderate resolution remote sensing data in the assessment of woody biomass loss and recovery from fire and other disturbance factors (pathogens, pests, thinning, logging, and windstorms). Applications of this approach might extend to provide a baseline crown cover, canopy height, and aboveground biomass record from 2000 onwards in support of the NASA Deformation, Ecosystem Structure and Dynamics of Ice (DESYnI) mission that is not likely to start prior to 2017. Canopy height is also a critical component for models predicting climate feedbacks. Further research is required to assess the results more rigorously against lidar-derived canopy height metrics and to determine whether it is feasible to use MISR-derived backgrounds to enable MODIS/GO and/or spectral invariant approaches, or combined MISR/MODIS data sets in model inversions to obtain a wider angular sampling — since the viewing planes are almost orthogonal (Jin et al., 2002) — in order to obtain more accurate retrievals.

Acknowledgments

This research was supported by NASA Earth Observing System grant NNX08AE71G to MC (Technical Manager: Dr. Diane Wickland). The MISR and MODIS data were obtained from the NASA Langley Atmospheric Science Data Center and the NASA EOS Data Gateway, respectively. We thank Xiaohong Chopping, David Diner (MISR Scientist, NASA/JPL), Jock Blackard and Ron Tymcio (US Forest Service, Rocky Mountain Research Station, Ogden, UT), Matt Smith and the Global Land Cover Facility (University of Maryland, College Park, MD), Barbara Nolen (Jornada Long Term Ecological Research), Joseph Youn and Michael Stoppay (Computer Operations for Research and Education, College of Science and Mathematics, Montclair State University), and the three anonymous reviewers.

References


Please cite this article as: Chopping, M., et al., Forest structure and aboveground biomass from MODIS and MISR, Remote Sensing of Environment (2011), doi:10.1016/j.rse.2010.08.031