Forest canopy height from the Multiangle Imaging SpectroRadiometer (MISR) assessed with high resolution discrete return lidar

Mark Chopping a,⁎, Anne Nolin b, Gretchen G. Moisen c, John V. Martonchik d, Michael Bull d

a Department of Earth and Environmental Studies, Montclair State University, Montclair, NJ 07043, United States
b Department of Geosciences, Oregon State University, Corvallis, OR 97331, United States
c Department of Earth and Environmental Studies, Montclair State University, Montclair, NJ 07043, United States
d USDA, Forest Service, Rocky Mountain Research Station, 507 25th Street, Ogden, UT 84401, United States

⁎ Corresponding author.
E-mail address: chopping@pegasus.montclair.edu (M. Chopping).

Abstract

In this study retrievals of forest canopy height were obtained through adjustment of a simple geometric-optical (GO) model against red band surface bidirectional reflectance estimates from NASA’s Multiangle Imaging SpectroRadiometer (MISR), mapped to a 250 m grid. The soil-understory background contribution was partly isolated prior to inversion using regression relationships with the isotropic, geometric, and volume scattering kernel weights of a Li-Ross kernel-driven bidirectional reflectance distribution function (BRDF) model. The height retrievals were assessed using discrete return lidar data acquired over sites in Colorado as part of the Cold Land Processes Experiment (CLPX) and used with fractional crown cover retrievals to obtain aboveground woody biomass estimates. For all model runs with reasonable backgrounds and initial $b/r$ (vertical to horizontal crown radii) values $>2.0$, root mean square error (RMSE) distributions were centered between 2.5 and 3.7 m while $R^2$ distributions were centered between 0.4 and 0.7. The MISR/GO aboveground biomass estimates predicted via regression on fractional cover and mean canopy height for the CLPX sites showed good agreement with U.S. Forest Service Interior West map data (adjusted $R^2 = 0.84$). The implication is that multiangle sensors such as MISR can provide spatially contiguous retrievals of forest canopy height, cover, and aboveground woody biomass that are potentially useful in mapping distributions of aboveground carbon stocks, tracking disturbance, and in initializing, constraining, and validating ecosystem models. This is important because the MISR record is spatially comprehensive and extends back to the year 2000 and the launch of the NASA Earth Observing System (EOS) Terra satellite; it might thus provide a ~10-year baseline record that would enhance exploitation of data from the NASA Deformation, Ecosystem Structure and Dynamics of Ice (DESDynI) mission, as well as furthering realization of synergies with active instruments.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Mapped estimates of forest canopy height, fractional cover, and aboveground woody biomass are relevant to important science questions regarding carbon storage and cycling, susceptibility to wildfire, gaseous and particulate emissions from wildfire, changes in structure from disturbance (pathogens, insect outbreaks, wildfire, storms, forest management practices such as thinning and logging), and assessment of biodiversity and wildlife habitat. To be useful these must be available over large areas, on a regular basis, and with as long a record as possible. Globally, forest has the largest magnitude and absolute uncertainty in carbon fluxes after fossil fuel emissions and is currently a sink (CCSP, 2007). However there is concern that with rapid climate change the terrestrial forest sink might be reduced in magnitude — or even become a source on decadal timescales — as a result of both deforestation and degradation (e.g., Kurz et al., 2008).

Regular mapping of the spatial distribution of aboveground woody biomass is one way to estimate changes in carbon fluxes between the biosphere and the atmosphere. Currently estimates are often made using averages based on rates of land use change (deforestation and reforestation; Olander et al., 2008; Canadell et al., 2007); that is, reductions in forest carbon storage are often assumed to be the result of deforestation through clearing for agricultural crops or timber. However emissions from forest degradation are highly uncertain and could either offset emissions from deforestation or more than double them (Houghton and Goetz, 2008).

Changes in forest aboveground carbon storage from disturbance might be regarded as merely transient if recovery occurs over the course of a few decades and carbon gains are the same as the initial losses. However if trends in disturbance are driven by decadal changes in climate then they are likely to affect storage on timescales that matter to society. This now seems to be the case: the effects of recent climate change on the forests of the western United States have been...
seen not only through increased susceptibility to fire (Running, 2006; Westerling et al., 2006) but also from drought stress (van Mantgem et al., 2009), with related insect outbreaks, such as the explosion in the distribution of western pine beetle (Dendroctonus brevicomis LeConte) and ensuing high rates of pine mortality in both the western United States and Canada since the early 2000s (McKenzie et al., 2009; Kurz et al., 2008). It is thus important that changes in C storage in forest are tracked as accurately as possible.

Canopy height is a vegetation structure metric that is useful in obtaining more accurate estimates of aboveground woody biomass and is a key indicator of successional status (Wofsy et al., 2008). This information must be available in a spatially comprehensive form, however there are only three remote sensing technologies that are able to map canopy height continuously over large areas: lidar, radar, and multi-angle imaging. All three technologies have limitations in terms of precision, accuracy, calibration requirements, and spatial and temporal sampling and coverage. For repetitive mapping over large areas, active instruments (lidar, radar) on spaceborne platforms are the future technologies of choice. This is the motivation behind the 5-year NASA Deformation, Ecosystem Structure and Dynamics of Ice (DESDynl) mission that will use full waveform, multiple beam lidar and L-band interferometric synthetic aperture radar to map forest structure and biomass globally, with launch expected some time in the period 2010–2015.

For mapping trends in canopy height and biomass over large areas, moderate resolution multiangle imaging is one option, although the data are difficult to interpret and the results difficult to validate because of the large ground resolution element (230–250 m). While it is known that canopy structure information is encapsulated in multiangle reflectance data (Nolin, 2004) and may be accessed with physical or semi-empirical models, empirical methods have provided the highest accuracies in estimating canopy heights to date. For example, Kimes et al. (2006) used a neural network with multiangle reflectance data from the airborne version of MISR (AirMISR), trained by height data from the NASA Laser Vegetation Imaging Sensor (LVIS, a waveform lidar) to obtain accurate canopy height estimates (R² = -0.9). However, AirMISR was flown in the solar principal plane (PP) in which the canopy structure signal is strongest (as well as in the cross-PP) but the MISR viewing plane is not close to the PP at mid-latitudes. Similarly, Heiskanen (2006) employed feed-forward multilayer neural networks to estimate forest heights from MISR multiangle and multispectral data in a tundra-taiga transition zone, with good results: using biotope inventory data with a minimum mapping unit of 1 ha as reference, results for 275 m top-of-atmosphere reflectance and 1.1 km surface reflectance data gave RMSEs of 1.98 m and 1.29 m, respectively, and R² of 0.71 and 0.83, respectively (N = 61,756 and 34,999, respectively). However, empirical methods may be limited in that they rely entirely on calibration, or “training”, and may suffer from extrapolation error when applied outside the domain for which they were trained.

Physical canopy reflectance model inversion with multivariate data has been attempted but it has not been straightforward to extract information in a way that is both meaningful and measurable (Widlowski et al., 2004). Schull et al. (2007) demonstrated how AIRMS data acquired over a mixed deciduous broadleaf and evergreen needle leaf forest (Howland Forest, Maine) can be interpreted using canopy spectral invariant theory. They compared the results with those from multivariate linear regression models and estimates from LVIS data and obtained coefficients of determination of 0.4–0.8. While use of canopy spectral invariant theory holds much promise there are some assumptions that may make mapping difficult in practice; and in particular the assumption of a black (non-reflecting) background. This probably introduces only small errors for environments with dark backgrounds such as the Howland Forest site but it might severely impact estimates in regions with brighter soils and sparser and/or highly variable understories.

Geometric-optical modeling is another physical modeling approach that has some attractive features. The ability to obtain meaningful canopy structural parameters using data from MISR in a geometric-optical (GO) modeling framework was demonstrated in a previous study in which retrievals were assessed against United States Forest Service (USFS) cover and height maps over ~200,000 km² in New Mexico and Arizona (Chopping et al., 2008a). Retrieved distributions of crown cover, mean canopy height, and woody biomass for forested areas showed good matches with data from USFS Interior West (FS-IW) maps, with R² values of 0.78, 0.69, and 0.81 (N = 576) and absolute mean errors of 0.10, 2.2 m, and 4.5 tons acre⁻¹ (10.1 Mg ha⁻¹), respectively, after filtering for high root mean square error (RMSE) on model fitting, the effects of topographic shading, and the removal of a very small number of outliers. Although the matches between the USFS Interior West (FS-IW; U.S. Forest Service, 2005) and MISR/GO height and cover maps were good, there remains the question of error in the former, especially since it is based partly on combining Moderate Resolution Imaging Spectroradiometer (MODIS) data with Forest Inventory and Analysis plot data and a number of environmental and climate variables in predictive non-parametric models (e.g., Ruenflacht et al., 2004; Blackard and Moisen, 2005). Furthermore, the FS-IW maps were not intended to be used for validation and may not have high accuracies; for this reason it has been necessary to seek other means of evaluating the results from GO model inversions with MISR data. Use of USFS FIA survey data is problematic because of the difference in scale: the plots are 36 m in diameter and do not provide an adequate means of assessing moderate resolution data, especially in heterogeneous canopies. It is clearly not practical to attempt to obtain ground-based data on canopy height for moderate resolution data. For these reasons, lidar data acquired from the air provide the best means of adequately assessing MISR retrievals. In this study heights from a high resolution discrete return lidar were used.

To enable practical GO model inversion with multiangle data it is necessary to estimate background reflectance anisotropy as well as magnitude (Gemmell, 2000). Most studies that exploit GO models do not do so using multi-angle data (e.g., Peddle et al., 1999); most attempts to isolate the background response in any context do so in the spectral domain (e.g., Spanner et al., 1990, Hall et al., 1995, Schlerf & Atzberger, 2006); some are theoretical or simulation studies (e.g., Ranson et al., 1986); and many dealing with canopy reflectance modeling assume that the background is simply “soil”, or that it is uniform. There are no studies of which we are aware that deal with extracting estimates of the background BRDF (or rather, the slice of it in the instrument viewing plane) in a dynamic way in the context of model inversion, as here. For example, Peltoniemi et al. (2005) investigate the spectral and directional properties of a forest understory using a field goniometer to sample the BRDF with a view to developing an empirical and/or theoretical understory model. Canisius and Chen (2007) report a method for extraction of forest background reflectance using the Four-Scale model and MISR data and assumed a Lambertian background: their objective was to obtain isotropic spectral reflectance magnitudes. In contrast, Ni and Li (2000) recognized the need to incorporate an anisotropic background but represented this using a static BRDF; that is, one that is not spatially dynamic.

2. Methods

2.1. Study area

The MISR data used in this study correspond to the State of Colorado and parts of southern Wyoming and western Nebraska, USA, an area of ~211,000 km², with some missing data where MISR aerosol/surface retrievals failed as a result of cloud or other atmospheric anomalies (contrails, aerosols, topographic shading). For the assessment of the retrievals, lidar ground and canopy elevation data and orthoimagery for the Cold Land Processes Experiment (CLPX) sites that correspond
with the mapped MISR pixels were acquired (Fig. 1). These sites are in the Rocky Mountains in north central Colorado. Forest types present in the CLPX sites include: fir–spruce, aspen–birch, western hardwoods, pinyon–juniper, ponderosa pine, lodgepole pine, as well as non-forest (grasslands with varying shrub cover). The CLPX sites are located in three isolated sub-regions—termed North Park, Rabbit Ears, and Fraser—that have different topography and forest characteristics (Table 1). North Park is mainly sage–grassland with generally low topographic relief and shallow drainage and some scattered shrubs and trees. The Rabbit Ears area is a higher elevation region with rolling topography. Land cover types are mixed coniferous–deciduous forests, within which trees tend to form large clumps (sometimes in long ribbon patterns perpendicular to the direction of the prevailing wind), with open meadows in between. The Fraser area has the highest relief of the three areas and land cover is mostly coniferous, subalpine forest, with alpine tundra at the highest elevations. Some locations within the Fraser site were subject to clear-cutting in the 1950s that has resulted in a matrix of high and low stands in some places.

2.2. MISR data sets and processing

MISR consists of nine pushbroom cameras arranged to view along-track that acquire image data with nominal view zenith angles relative to the surface reference ellipsoid of 0.0°, ±26.1°, ±45.6°, ±60.0°, and ±70.5° (forward and aft of the Terra satellite) in four spectral bands (446, 558, 672, and 866 nm). The 672 nm (red) band images are acquired with a nominal maximum cross-track ground spatial resolution of 275 m in all nine cameras and all bands are acquired at this resolution in the nadir camera (Diner et al., 1999). This study employed a new method for extracting and processing MISR data from the Hierarchical Data Format (HDF) distribution files that are obtained using the MISR Order and Customization Tool hosted at the NASA Langley Atmospheric Science Data Center (http://lodup05.larc.nasa.gov/MISR/cgi-bin/MISR/main.cgi). Instead of performing atmospheric corrections after extraction in order to obtain surface reflectance estimates, a regression algorithm was developed by MISR scientists at NASA JPL that exploits the fact that the 1.1 km data are routinely screened and corrected as part of the Land Surface product. This efficient algorithm was installed as part of the MISR Toolkit and routines were constructed that apply the regression and reproject the data to a raster image grid of a chosen size and interval in any common map projection. The MISR Toolkit-based routines require four MISR data products: the MISR level 1B2 MI1B2 T Terrain-projected Spectral

![Fig. 1. Locations of the CLPX sites superimposed on the Forest Service Interior West (FS-IW) height map (extent shown is the black area in the inset). The gray area in the inset shows the area mapped with MISR. Only areas considered forest are included in the Forest Service map: areas shown in white either have heights < 1 m or are non-forest.](image-url)
Radiance product; the MISR Level 1B2 M11B2GEOP Geometric Parameters product; the MISR Level 2 MIL2ASLS Land Surface product; and the MIANCAPG Ancillary Geographic product. The M11B2T product is the terrain-projected top-of-atmosphere spectral radiance with a nominal 275 m spatial resolution in the nadir multispectral bands and off-nadir red band (Diner et al., 1999). The M11B2GEOP product provides grids of solar azimuth, solar zenith, and nine viewing azimuth and zenith angles at 17.6 km resolution. The Land Surface product provides surface bihemispherical and directional-bihemispherical reflectance (albedo), hemispherical directional and bidirectional reflectance factor (BRF), BRF model parameters, leaf-area index (LAI), fraction of photosynthetically active radiation, and normalized difference vegetation index on a 1.1 km grid. The MISR Ancillary Geographic product provides geographical locations, elevation data, surface azimuths, and a land/water mask, on 1.1 km and 17.6 km resolution grids.

MISR Terrain-projected spectral radiance data were acquired for an August 10, 2002 Terra overpass (orbit 014073, path 34, blocks 57–60. The standard MISR algorithm uses two separate values for the surface albedo parameter: 0.015 for land or shallow/in-land/coastal waters and 0.0 for deep ocean (Diner et al., 2008, section 3.5.4). The land/inland water value used in the standard algorithm is too high for the dark lakes in this scene, so the surface albedo parameter was set to zero across the entire scene to increase coverage in the Land product that is required to obtain surface Terrain-projected BRFs. The modified Land product was used with MISR Toolkit routines to calculate surface reflectance estimates and map these to the Universal Transverse Mercator map projection, WGS84 spheroid/datum, zone 13N, with a grid interval of 250 m.

2.3. Lidar reference data and sites

High resolution discrete return lidar data and associated orthoimagery from the 2003 CLPX-Airborne Infrared Orthophotography and LIDAR Topographic Mapping campaign — part of the Cold Land Processes Field Experiment (CLPX) — were used (Miller, 2003). The lidar instrument was an Azimuth Corporation ALS40 24 KHz system, now called the Leica ALS40. The instrument records the first and last returns per laser pulse, along with intensity, with a user-selected off-nadir scanning angle. The data for the CLPX campaigns were acquired at approximately 1372 m above mean terrain, normalized to ground controls, and processed to remove noise and redundancies. Horizontal spacing of the data is approximately 1.5 m, with approximately 0.05 m vertical tolerances. The data were obtained from the CLPX archive at the National Snow and Ice Data Center (NSIDC, Boulder, Colorado) as elevations, filtered into ground and top-of-vegetation data sets in an ASCII XYZ format. The CLPX sites for which useful lidar data are available, their codes, and their vegetation types are listed in Table 1. Each of these 107 sites completely encompasses 9 - 12 MISR cells, 15 of which were selected as background calibration sites.

In order to obtain estimates of mean canopy height corresponding to the mapped MISR imagery, a surface digital elevation model (DEM) was first constructed via interpolation on the ground elevation data. For each site the ground elevations in the DEM were then subtracted from the corresponding vegetation elevations, with checks for anomalies (e.g., where either of the elevations was out of range), to produce a 2 m spatial resolution tree height raster image. Statistics for all MISR raster cells fully within the CLPX lidar-derived tree height images were extracted using an index image in which each raster cell holds an unique, ordinal identifier.

2.4. Geometric-optical model

Geometric-optical models are able to resolve statistical distributions of discrete objects within an instrument’s instantaneous field-of-view (IFOV) (Strahler et al., 2005). Simple geometric-optical (GO) models treat the surface as an assemblage of discrete objects of equal radius, shape and height, evenly distributed within a spatial unit. A tree or shrub crown is represented by a geometric primitive (a sphere in this case) whose center is located at a specified mean height above a (nominally diffuse scattering) background. These models predict the top-of-canopy reflectance response to important canopy physical parameters (plant number density, foliature volume, mean canopy crown height, radius, and crown shape, background brightness and anisotropy) as a linear combination of the contributions from sunglint and viewed, and shaded and viewed crown and background components (Li and Strahler, 1985; Chen et al., 2000), as Eq. (1):

\[ R = G_{k_c} + C_{k_c} + T_{k_f} + Z_{k_f} \]  

where \( R \) is bidirectional spectral reflectance; \( k_c, k_f \) and \( k_g \) are the GO modeled proportions of sunglint background, sunglint crown, shaded crown and shaded background, respectively; and \( G, C, T, \) and \( Z \) are the contributions of the sunglint background, sunglint crown, shaded crown, and shaded background, respectively. GO models are particularly appropriate for the exploitation of solar wavelength remote sensing data acquired at differing viewing and/or illumination angles because the proportions of sunglint and shaded crown and background in the remote sensing instrument ground-projected IFOV vary with viewing and illumination angles and canopy configuration. The simple geometric model (SGM), a GO model incorporating a dynamic background and a crown volume scanning term, was used (Chopping et al., 2006; Chopping et al., 2008a,c). It is formulated as Eq. (2):

\[ R = G_{\text{Walthall}}(\theta_i, \varphi, \psi).k_c(\theta_i, \varphi, \psi) + C_{\text{Ross}}(\theta_i, \varphi, \psi).k_c(\theta_i, \varphi, \psi) \]  

where \( \theta_i, \varphi \), and \( \psi \) are the view zenith, solar zenith and relative azimuth angles, respectively; \( k_c \) and \( k_g \) are the calculated proportions of sunglint and viewed background and crown, respectively; \( G_{\text{Walthall}} \) is the background contribution from the modified Walthall model (Walthall et al., 1985; Nilson and Kuusk, 1989); and \( C_{\text{Ross}} \) is the simplified Ross turbid medium approximation for plane parallel canopies (Ross, 1981). The shaded components \( T \) and \( Z \) (Eq. (1)) are discarded; they are assumed black, as in the kernel-driven bidirectional reflectance distribution function (BRDF) models (Roujean et al., 1992; Wanner et al., 1995). The model’s upper canopy parameters are plant number density (\( \lambda \)), mean crown radius (\( r \)), crown vertical to horizontal radius ratio (\( b/r \)), crown center height to vertical radius ratio (\( h/b \)), and crown leaf area index (LAI). Leaf reflectance in the red wavelengths is fixed at 0.09. See Chopping et al., 2006, 2008a for further details on model derivation.

2.5. Isolation of the background contribution

For optimal GO model inversion in arid and semi-arid environments it is important that the contribution of the background can be approximated for the illumination/viewing configurations of the remote sensing observations. In this study, the background includes everything that is not part of the upper canopy (trees, or in some cases shrubs) and is typically composed of several elements (bare soil, crusts, and understory plants such as grasses, forbs, and sub-shrubs) within the mapped instantaneous field-of-view (IFOV) of the instrument. The background includes all non-tree features, not just the understory directly underneath trees; for example, it includes areas of open grassland adjacent to a forest stand where both are inside the mapped MISR cell. In order to estimate background brightness and anisotropy in the various MISR illumination/viewing configurations, linear multiple regression based on a small number of calibration sites was used (Chopping et al., 2006, 2008a,b,c,d). The independent variables are BRDF model isotropic (iso), geometric (geo), and volume scattering (vol) kernel weights obtained by adjusting the LiSparse-RossThin model
against MISR red band bidirectional reflectance factors (BRFs) in all nine cameras using the Algorithm for Modeling Bidirectional Reflectance Anisotropies of the Land Surface (AMBRALS) code (Strahler et al., 1996), with the objective the minimization of absolute RMSE. The multi-spectral reflectance estimates from the MISR nadir-viewing camera were not used as this can result in over-fitting and because light in the shorter wavelengths is more prone to scattering by the atmosphere.

Ideally it would be possible to find a suite of 250 m² calibration sites covering the domain of possible background configurations (i.e., areas with varying proportions of soil, grasses, or other low vegetation) that do not include any upper canopy and use the MISR data for these to calibrate the regression models. However, it is usually impossible to find 250 m² areas that do not contain some shrubs or trees, so background contributions must be extracted using the GO model. Provided with a set of MISR red band BRFs, estimates of upper canopy crown cover and approximate b/r ratio, and the GO model, an optimization algorithm can extract the best-matching background for each site. For this study, 15 CLPX sites out of a total of 107 were selected for use in background calibration.

The regression equations were established for a range of forests canopy/background configurations by fixing number density and setting mean shrub radius to match fractional crown cover estimated from unsupervised classification of the CLPX orthoimagery using the ISODATA algorithm, or in one or two cases, via manual digitization of the canopy area. Either snow-covered (April) or mostly snow-free (September) scenes were used in the classifications, depending on the scene characteristics (e.g., some of the April scenes were acquired as oblique images and could not be used to estimate cover). The LAI and h/b model parameters were fixed at 2.08 and 1.0, respectively, for all sites and b/r was set to 2.0 and 0.2 for forest and grassland sites, respectively. An optimization algorithm was then used to determine for each calibration site the best-fitting Walthall model parameters with respect to the MISR data. This allows the regression of each of the Walthall model parameters on the three independent variables and the resulting regression equations can be used to obtain estimates of the background response prior to adjustment of the GO model (Chopping et al., 2006, 2008a,b,c). This method has been shown to provide reasonable estimates of the background contribution for desert grasslands with and without shrubs: the relationship between the predicted background brightness and the product of understory fractional cover and mean understory grayscale values was strong ($R^2 = 0.75$, Chopping et al., 2008a). Subsequent experiments with fixed and dynamic backgrounds over New Mexico and Arizona forests showed that it is necessary to use a dynamic background for accurate GO model inversions and that extrapolation error is not limiting over surprisingly large areas (Chopping et al., 2008a,b). Three measures of background brightness — the Walthall diffuse scattering parameter and reflectance for a view zenith angle of 45° with the sun overhead in

---

Fig. 2. Predicted background brightness in the vicinity of the CLPX North Park and Rabbit Ears sites (a) Walthall isotropic parameter (b) backscattering direction: VZA = 45° in the solar principal plane with overhead sun (c) forward-scattering direction (same geometry) (d) MISR/GO fractional cover (e) MISR/GO mean canopy height (f) canopy height map for the entire area.
The backscattering and forward-scattering directions and in the solar principal plane — are shown in Fig. 2a, b, and c, respectively, together with fractional cover (d) and mean canopy heights (e) for an area surrounding the North Park and Rabbit Ears CLPX sites. It can be seen that the background under forested areas is predicted as somewhat darker than that of the non-forest areas, as expected. In the forward-scattering direction the background is in general darker and contrast between forest and non-forest background brightness is lower, as expected, owing to shadowing by understory plants.

The selection of sites for background calibration was not obvious — partly because the accuracy of the cover estimates varies — so some effort was necessary to determine the optimal subset. To do this, the MISR data and Li-Ross kernel weights were extracted for the 15 candidate sites and a series of regression and modeling tests was performed. In general, background extraction is more difficult for sites with higher crown cover because the background signal is weaker. In this study, sites with fractional cover \( N > 0.30 \) proved more challenging and one site in the Rabbit Ears, Spring Creek (RS) CLPX site with an estimated fractional cover of 0.38 was eventually excluded. Model inversion runs were then performed with different sites and settings and assessed for all calibration sites to seek the best combination with respect to RMSE and \( R^2 \) against the lidar-based estimates of height and orthoimagery-based cover estimates, prior to application over the entire area (most of the state of Colorado and parts of Wyoming and Nebraska).

2.6. Model inversion protocol

In this study, the SGM was adjusted against the MISR red band data in all nine views using the Praxis algorithm (Brent, 1973; Powell, 1964) with min\(|\text{RMSE}|\) as the objective function and no constraints or weighting of the error terms. The \( \text{LAI} \), \( \lambda \), and \( h/b \) model parameters were fixed at 2.08, 0.012, and 2.00, respectively, with \( r \) and \( b/r \) left as free parameters set to initial values in the ranges 3.0–6.0 m and 0.5–.
2.5, respectively. The \( r \) and \( \lambda \) parameters are internally coupled and the same result is obtained if fractional cover is maintained but these parameters are varied; thus it is not feasible to adjust both parameters simultaneously. A model inversion run proceeded by reading from the multi-angle red reflectance and kernel weight files, submitting these to the minimization code, and accumulating the results in an output file. Runs were first performed for the calibration sites only in order to provide rapid indications of the results with respect to the lidar heights. To determine the robustness of the inversions (i.e., the quality of the results that might be obtained with access to only minimal reference data, and to assess sensitivity to the background contribution and initial values for the free parameters \( r \) and \( b/r \)), a series of inversion runs was subsequently performed for all CLPX sites with different sets of background coefficients and systematically-varying \( r \) and \( b/r \) starting values. For two backgrounds (#5 and #7) the \( b/r \) value used for forest in the extractions was experimentally set to an unreasonable value of 0.2 (very oblate crown shape), in order to demonstrate the impact of using poor dynamic backgrounds. In all, model inversion runs were completed for 140 combinations of backgrounds (7), \( r \) (3.0–6.0 in increments of 1.0) and \( b/r \) (0.5–2.5 in increments of 0.5). To process the entire area the input data were divided into 12 sets of files and two sets processed on each of six dual-processor XServe machines, with the results files — containing model fitting RMSE, \( r \), \( b/r \), fractional cover, and mean canopy height — concatenated together at the end of the run. The inversions for the entire area were performed extremely rapidly, completing in <30 min (about the same time required to complete inversions of the Li-Ross BRDF model using AMBRALS). The results were then read into raster image layers for mapping and analysis. Estimates of aboveground woody biomass were obtained using regression of the FS-IW biomass estimates on the MISR/GO fractional cover and mean canopy height retrievals.

It is important to note that while the \( r \) and \( b/r \) parameters match known distributions of woody plant sizes and crown shapes inferred from vegetation maps (for example, the more oblate crown shape values for mesquite-dominated areas vs the more spherical crown shape values for creosotebush-dominated areas; Chopping et al.,

![Graph](image_url)

**Fig. 6.** Lidar mean height (April), MISR/GO mean height, FS-IW fractional cover (fcov), and MISR/GO fractional cover (fcov) for all sites (a) quasi-random order within CLPX sites (b) ranked by lidar mean height (c) lidar and MISR/GO heights using abs(h+b) for the [background / r / b/r] combination with the lowest RMSE and removing all retrievals where fcov>0.99. Lines are included for clarity of reading only.
2008b), they are effective parameters. Because λ is fixed and r is free, the model is adjusted to match fractional crown cover. Because \( h/b \) and \( r/b \) are fixed and r and \( b/r \) are adjusted, the inversion effectively provides an estimate of \( b \), vertical crown radius. This allows the retrieval of an estimate of \( h \), via \( h/b \times b \), noting that \( h \) refers to the height of the center of the crown; the height of the top of a tree is \( h + b \).

In the course of performing repeat model inversion runs on different computers, it was noticed that the sign of \( b/r \) and thus \( h \) was occasionally reversed in the low height grassland sites in 11 out of 37 cases. The reasons for this are not clear but inspection of the results when the absolute value of all \( h \) values is taken suggests that this is owing to a mathematical fitting anomaly; sign reversal is in both directions (when positive on machine A, \( h \) can be negative on machine B, and vice-versa); the values were identical to at least three decimal places; and heights calculated with absolute values are consistently within the expected ranges. The MISR/GO height results obtained in this way are henceforth termed MISR/GO\( \text{abs}(h + b) \), or denoted by “abs”.

### 3. Results and discussion

Since the GO model is highly simplified and all parameters except \( r \) and \( b/r \) are fixed, there are good reasons to expect low accuracy; or, if there is good accuracy that this is spurious (i.e., not reflecting model operation but merely a correlate). However the results do not support either of these contentions. The MISR/GO retrievals and CLPX lidar height estimates for the calibration sites indicate that MISR/GO height retrievals are more accurate with respect to the lidar heights than the FS-IW map estimates (Fig. 3(a)). Notably, the FS-IW map misses forest completely in sites 7 and 10 and overestimates height for sites 11 through 14. For the forest calibration sites only (\( N=7 \)), the coefficients of determination were 0.87 vs CLPX lidar-derived heights, 0.86 vs FS-IW map fractional cover estimates, and 0.23 vs fractional cover estimated from orthoimagery. This compares with 0.34 vs CLPX lidar heights and 0.28 vs fractional cover estimated from orthoimagery for the FS-IW height and cover data. There was an anomalously high MISR/GO height retrieval for the Fraser Forest Fool Creek (FF) site but this is easily isolated because fractional crown cover goes to 1.0, an infeasible result (Fig. 3(b)). Validation of the MISR/GO retrievals is not straightforward as the slope of the lidar:MISR relationship, even the sign, depends on the way the lidar heights are calculated and on the filtering used to separate ground and vegetation elements (Fig. 4).

For all available CLPX forest sites (\( N=58 \), excluding the Fool Creek sites where infeasible crown cover values were obtained but including all combinations of backgrounds and initial \( r \) and \( b/r \) values), typical \( R^2 \) values vs April and September lidar heights are \(-0.47 \) and \(-0.71 \), respectively, while typical RMSE values are \(-2 \) m and \(-5 \) m (Fig. 5). The MISR/GO height retrievals for shrubs in the North Park grassland sites do not match well with the lidar data, particularly when the sign of the result is observed (Fig. 6). This is either owing to a need to provide separate calibration coefficients optimized for either grassland or forest, or to the lidar system’s inability to resolve low woody plants, or both. The MISR/GO retrievals are much better than the FS-IW map height estimates with respect to the lidar heights (Fig. 6(a)–(b)). The MISR/GO cover retrievals are also more consistent with the lidar height retrievals, while the FS-IW map misses some forest sites completely (Fig. 6(a)–(b)). As with the single Fool Creek calibration site, anomalously high MISR/GO height values were obtained for a number of locations in several of the Fraser Forest sites (#96–107 in the plot). Again, these can easily be screened out as they correspond to fractional cover retrievals of 1.0 (Fig. 6(a)–(b)). When the absolute values of \( h \) were used, a better match was obtained with the April lidar heights, with the exception of the Fool Creek sites (Fig. 6(c)).

It is not known why the retrievals were sometimes very precise matches to the CLPX lidar data (e.g., sites 41–55 and 67–77) and less so for other sites but this may depend on the lidar data: the strength of the relationship depends on whether the April or September lidar data are used. There is also a bias in the MISR/GO results when assessed against the September lidar heights that is not apparent in the assessment against April lidar heights. Since the September lidar mean heights are consistently lower than the April mean heights and by as much as 7.7 m (Fig. 6(c)), it seems plausible that this might be a result of the filtering used to divide the lidar elevations into ground and vegetation sets. Artifacts are clearly visible when the lidar data are plotted over ortho-imagery (Fig. 7). It is likely that small, non-tree objects – rocks and/or low plants – were inadvertently included in the vegetation data set (Fig. 7(a)), resulting in lower average heights (contrary to what might be expected, since the presence of a snowpack would result in higher ground elevations and therefore somewhat lower tree heights when calculated by subtracting the ground from the vegetation elevations). Most of these small objects would be covered with snow in April and thus not be identified as vegetation by the filtering algorithm; they would therefore not contribute to the means. There are also some linear filtering artifacts in these data (Fig. 7(b)).

The mosaic of mature and regrowing stands in some of the Fraser Forest sites may be responsible for the lower lidar mean canopy

---

**Fig. 7.** Examples of non-trivial heights for non-tree features in the September lidar vegetation layer, plotted over a near-infrared/red/green false color composite from ortho-imagery (a) CLPX RS site (b) CLPX RW site.
a  
RMSE (m) 

b  
RMSE (m) 

c  
RMSE (m) 

d  
RMSE (m) 

r, per background with b/r = 0.5, 1.0, 1.5, 2.0, and 2.5
Table 2

<table>
<thead>
<tr>
<th>RMSE and $R^2$ vs lidar heights for all MISR/GO height retrievals.</th>
<th>RMSE (m)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Min</td>
</tr>
<tr>
<td>April</td>
<td>2.6</td>
<td>1.4</td>
</tr>
<tr>
<td>April (abs$^b$)</td>
<td>2.0</td>
<td>1.4</td>
</tr>
<tr>
<td>September</td>
<td>3.8</td>
<td>2.8</td>
</tr>
<tr>
<td>September (abs$^b$)</td>
<td>3.6</td>
<td>2.8</td>
</tr>
</tbody>
</table>

$^a$With $b/r<2.0$ and reasonable dynamic backgrounds.
$^b$Taking the absolute value of MISR height retrievals (see text).

heights and thus the disagreement with the MISR/GO retrievals (Fig. 6(a) and (c)). Topography may also play a role (no adjustments to angles were applied prior to GO model inversions) and the non-Poisson distribution of trees at the stand scale may violate the GO assumption of evenly distributed plants; or all three factors may be partly responsible. The towers that were installed in the CLPX sites in April result in negligible changes to means as there is only one per site; and they are only detectable in the statistics for the North Park grassland sites.

The results of the series of systematic model inversion runs indicate that the results for forest are robust with respect to the choice of initial values for the free parameters $r$ and $b/r$, as long as the starting value for $b/r$ is $<2.0$ (Fig. 8). Only 29 out of the 140 [background, $r$, $b/r$] combinations resulted in an RMSE $<$10 and without exception these resulted from the use of the experimental backgrounds ($#5$ and $#7$) and $b/r$ starting values $<$2.0. When these conditions were excluded, RMSE distributions vs April and September lidar were centered around 2.5 m and 3.7 m with maxima of 5.5 m and 5.6 m, respectively (Fig. 8(a)–(d), Table 2) and $R^2$ distributions were centered around 0.4 and 0.7 with minima of 0.10 and 0.24 (Fig. 9(e)–(h), Table 2). While the relationships are weaker than those typically achieved with lidar instruments, RMSE values are comparable over areas with marked topography or mixed crown shapes and at different times of year (e.g., Pang et al., 2008; Sun et al., 2008; Lefsky et al., 2008; Hyde et al., 2007). Note that when the MISR retrievals are poor — such as when inadequate background coefficients are used — they tend to be spectacularly poor, suggesting that inadequate model inversions are easily detectable with a minimal set of reference data.

Representations of typical 250 m$^2$ MISR/GO maps of aboveground woody biomass, mean canopy height, fractional cover, and model-fitting RMSE are given in Fig. 10. The missing data in these maps corresponds to 17.6 km$^2$ blocks where the MISR aerosol/surface retrievals failed, often because of cloud cover. However there are also obvious clouds in areas that were not screened out in the MISR processing, that can be clearly seen from their orientation in the model-fitting RMSE map (Fig. 10(d)). Including these data compromises inversions for cover and height (Fig. 10(b)–(c)); however the sensitivity of model-fitting RMSE to atmospheric perturbations provides a detection mechanism, allowing the compositing of data from multiple overpasses; this has been found to be effective in previous mapping exercises (Chopping et al., 2008a). According to the model-fitting RMSE (maximum across all sites $=0.01$) none of the CLPX validation MISR data were impacted by clouds.

The regression equation for estimating aboveground woody biomass from MISR/GO fractional cover and mean canopy height was obtained using data for CLPX forest sites with valid FS-IW estimates. The intercept ($−6.2$), cover coefficient (41.76), and height coefficient (1.06) were all significant at the 86% level or better ($p=0.0359$, $0.0008$, and 0.1375, respectively), with the large coefficient for cover reflecting the strong dependence of biomass on this attribute, as expected. Both cover and height had linear relationships with biomass, with the distributions of residuals approximately normal. MISR/GO aboveground biomass estimates predicted via regression showed good agreement with the FS-IW map values ($N=39$, adjusted $R^2=0.84$), although these may not themselves be accurate.

4. Conclusions

This study shows that multiance, mono-spectral data from MISR can be used for mapping forest canopy height over large areas and at low cost, with implications for the mapping of aboveground woody biomass. A limitation of the previous study (Chopping et al., 2008a) was that the U.S. Forest Service reference data used were not intended for validation purposes: the 2005 Interior West forest maps were released as draft data sets intended for review. This study has shown that when assessed against high resolution discrete-return lidar height estimates in central Colorado, accurate forest canopy heights can be retrieved using only MISR red band data in a GO modeling framework. It is important to note that the MISR height retrievals were obtained independently from the lidar data: the results are model-based, not empirical fits to data, or trained in any way. The multiance/GO model approach offers some attractive features that include good accuracy vs lidar-based height estimates; excellent coverage; parsimony, since only red band data are required; low cost, since it exploits existing NASA Earth Observing System data; rapidity; limited manual intervention, with one background coefficient set adequate for large areas; and the ability to map shrubs as well as forest. The approach also has several limitations: model inversion can provide infeasible values, although these can usually be identified and screened out; the method is unsuitable for closed canopies such as those of tropical and other dense forest; it is not possible to decompose fractional cover into number density and mean crown radius; and retrievals are probably less precise in the vertical dimension than those from large footprint lidar.

Application to contiguous, wall-to-wall mapping over large areas (e.g., the western United States) depends on two conditions. First, MISR data from multiple Terra orbits are required to compensate for missing surface retrievals owing to clouds, contrails (and their shadows) and aerosol retrieval failures. Previous work has shown that compositing on minimum model-fitting error is an efficient way to preferentially select the least contaminated observation because even very small deviations from a surface angular signature result in relatively large RMSE values (Chopping et al., 2008a,b). Second, a sparse grid of background calibration regression coefficient sets is required, although in the mapping experiments performed to date one set seems adequate for large areas that have similar soil and understory vegetation characteristics (e.g., Arizona and New Mexico). Such a grid could be easily constructed using data from any site with fractional canopy cover information, such as the National Science Foundation LTERs, Ameriflux sites, and/or EOS Land Validation Core Sites (http://landval.gsfc.nasa.gov), noting that this has to be performed only once for each kind of background and time of year. Given the insensitivity to error in the required estimates of fractional cover for a set of background calibration sites, estimates derived from any high-resolution imagery may be adequate.

While multiance remote sensing and associated modeling methods might be considered somewhat complex, applications have begun to proliferate in recent years (Diner et al., 2005; Chopping, 2008). Applications of the approach described here to problems in terrestrial ecology are diverse and include mapping distributions of aboveground woody carbon stocks over large areas; tracking biomass loss and recovery from fire and other disturbance; and providing a means to reduce uncertainties in satellite-derived maps of fractional snow cover. This approach might also serve to provide a year 2000-baseline crown cover, canopy height, and aboveground biomass record in support of the 5-year

Fig. 8. RMSE of MISR/GO vs lidar mean canopy heights for all combinations of backgrounds (bg-N) and starting values for $r$ indicated on the x-axis and $b/r$ (in the order 0.5, 1.0, 1.5, 2.0, and 2.5) for each value (a) April with abs$(h+b)$ (b) September with abs$(h+b)$ (c) April (d) September.
Fig. 10. MISR/GO maps of (a) aboveground woody biomass via regression on Forest Service estimates for the CLPX sites (b) mean canopy height (c) fractional crown cover (d) model-fitting RMSE, for most of Colorado and parts of Wyoming and Nebraska. The maps were made using a dynamic background using data from all calibration sites except one in the Rabbit Ears, Spring Creek area and inverting with a [$r$, $b/r$] start point of [3.0, 1.0]. Note the correlation between model-fitting error and surface retrieval failures (gaps in the coverage) that are owing to the presence of clouds (bright linear features).

Fig. 9. Frequency distributions of RMSE and $R^2$ for MISR/GO vs lidar mean canopy heights (a) April RMSE (b) September RMSE (c) April $R^2$ using abs($h+b$) (d) September RMSE using abs($h+b$) (e) April $R^2$ (f) September $R^2$ (g) April $R^2$ using abs($h+b$) (h) September $R^2$ using abs($h+b$).
NASA DESDynl mission, noting that the ability to map primary canopy structure parameters using a GO approach does not detract from the need to deploy active instruments for global mapping of vegetation structure, especially for closed canopy tropical forest. DESDynl will also provide information on the finer scale horizontal distributions of structural parameters, as it will provide maps with a much higher spatial resolution (~25 m); and a variety of metrics can be extracted from waveforms. On the other hand, the short duration of the mission limits the extent to which it might capture all important patterns of biomass loss and recovery. For example, extreme events such as the 2004 Alaska wildfires (Pfister et al., 2005) may or may not occur within the lifetime of the mission. Future multiangle imagers with IFOVs of around 100 m, together with canopy structure information from active instruments will help to provide a more complete and accurate picture of terrestrial carbon storage in the terrestrial biosphere, as well as forest dynamics. The use of multiangle red band data from MISR and MODIS together in GO model inversion may also prove useful and early results with MODIS alone have shown promise (Chopping et al., 2009, 2008c), although it remains to be seen whether including MODIS data will enhance the MISR-based results.

Acknowledgments

This research was supported by NASA Earth Observation System grant NNX08AE71G to MC (Technical Manager: Dr. William Emanuel). The MISR data were obtained from the NASA Langley Atmospheric Science Data Center. The CLPX data were provided courtesy of National Snow and Ice Data Center, Boulder. We thank Xiaohong Chopping; David Diner (MISR Scientist, NASA/JPL) and the MISR Science Team; Molly McAllister (National Snow and Ice Data Center, University of Colorado, Boulder, CO); Chad Poole (Spectrum Mapping LLC, Albuquerque, NM); Ron Tynio and Tracey Frescino (US Forest Service, Rocky Mountain Research Station, Ogden, UT); Matt Smith and the Global Land Cover Facility (University of Maryland, College Park, MD); Joseph Youn and Michael Stoppay (Computer Operations for Research and Education, College of Science and Mathematics, Montclair State University); and all participants in CLPX.

References


