Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error

Joseph O. Sexton , Xiao-Peng Song , Min Feng , Praveen Noojipady , Anupam Anand , Chengquan Huang , Do-Hyung Kim , Kathrine M. Collins , Saurabh Channan , Charlene DiMiceli & John R. Townshend

Taylor & Francis
Taylor & Francis Group

Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error

Joseph O. Sexton[a]*, Xiao-Peng Song[a], Min Feng[a], Praveen Noojipady[a], Anupam Anand[a], Chengquan Huang[a], Do-Hyung Kim[a], Kathrine M. Collins[a], Saurabh Channan[a], Charlene DiMiceli[b] and John R. Townshend[a]

[a]*Global Land Cover Facility, Department of Geographical Sciences, University of Maryland, College Park, MD, USA;* [b]*Department of Geographical Sciences, University of Maryland, College Park, MD, USA*

We developed a global, 30-m resolution dataset of percent tree cover by rescaling the 250-m MOderate-resolution Imaging Spectroradiometer (MODIS) Vegetation Continuous Fields (VCF) Tree Cover layer using circa- 2000 and 2005 Landsat images, incorporating the MODIS Cropland Layer to improve accuracy in agricultural areas. Resulting Landsat-based estimates maintained consistency with the MODIS VCF in both epochs (RMSE = 8.6% in 2000 and 11.9% in 2005), but showed improved accuracy in agricultural areas and increased discrimination of small forest patches. Against lidar measurements, the Landsat-based estimates exhibited accuracy slightly less than that of the MODIS VCF (RMSE = 16.8% for MODIS-based vs. 17.4% for Landsat-based estimates), but RMSE of Landsat estimates was 3.3 percentage points lower than that of the MODIS data in an agricultural region. The Landsat data retained the saturation artifact of the MODIS VCF at greater than or equal to 80% tree cover but showed greater potential for removal of errors through calibration to lidar, with post-calibration RMSE of 9.4% compared to 13.5% in MODIS estimates. Provided for free download at the Global Land Cover Facility (GLCF) website (www.landcover.org), the 30-m resolution GLCF tree cover dataset is the highest-resolution multi-temporal depiction of Earth's tree cover available to the Earth science community.

Keywords: tree cover; continuous fields; Landsat; MODIS; lidar

Introduction

Tree cover – defined structurally as the proportional, vertically projected area of vegetation (including leaves, stems, branches, etc.) of woody plants above a given height – affects terrestrial energy and water exchanges, photosynthesis and transpiration, net primary production, and carbon and nutrient fluxes (DeFries et al. 1995). Tree cover also affects habitat quality and movements of wildlife (Conde et al. 2010; Trainor et al. in press), residential property value for humans (Mansfield, et al. 2005), and numerous other ecosystem services. Importantly for monitoring, reporting, and verification (MRV) efforts to reduce carbon dioxide emissions from deforestation and forest degradation and to foster conservation and sustainable management of

---

forests (REDD +), tree cover provides a measurable attribute upon which forest cover may be defined. Further, changes in tree cover over time can be used to monitor and retrieve site-specific histories of forest disturbance, succession, and degradation (Huang et al. 2009).

In geospatial analyses, tree cover is most commonly inferred from categorical maps, whose schema represent ranges of cover through such classes as 'woodland', 'sparse savanna', 'woody savanna', and 'forest' (Bennett 2001). Although it is the most common scheme used for mapping and change detection, this categorical approach inadequately represents within-class heterogeneity for many analyses (DeFries et al. 1995, 1999) and can underestimate forest canopy loss by as much as 50% (Asner et al. 2005). Alternatively, tree cover may be represented more directly as a 'continuous field', in terms of fractions or proportions of pixel area (DeFries et al. 1999).

The MODerate-resolution Imaging Spectroradiometer (MODIS) Vegetation Continuous Fields (VCF) Tree Cover dataset, currently in Version 5, is produced at 250-m resolution globally from 2000 to 2010 (DiMiceli et al. 2011). In contrast to methods based on linear mixture models (e.g. DeFries et al. 1999; Asner et al. 2009), the MODIS VCF is based on a flexible regression tree algorithm, which is more capable of incorporating empirical information to improve correlation of estimates to measured tree cover. The MODIS Tree Cover VCF has been used for a wide range of continental- to global-scale assessments (e.g. DeFries et al. 2005; Miles et al. 2006; Lawrence and Chase 2007; Hansen et al. 2008; Simard et al. 2011; Harris et al. 2012). However, many land cover changes occur in patches beneath its 250-m resolution (Townshend and Justice 1988). Higher-resolution continuous-field datasets have been generated for limited areas, primarily in the United States from Landsat data (e.g. Homer et al. 2004; Rollins 2009; Hansen et al. 2011), but there are currently no global datasets representing tree cover at resolutions finer than that of the MODIS sensor.

The spatial and thematic scale of the MODIS VCF and other continuous-field datasets (e.g. Asner et al. 2009) have made reference data difficult to acquire and so quantitative error estimatesof these datasets are quite limited. Hansen et al. (2002) provided the first *de facto* – although not independent – estimates of MODISC VCF accuracy by comparing an experimental version of the dataset to the Landsat data used to train the generating model. Later, White et al. (2005) compared the MODIS VCF Version 1 to independently gathered field data across the arid southwestern United States, and Montesano et al. (2009) validated the Version 4 MODIS VCF against independent reference data derived from photo-interpreted high-resolution images across the boreal-taiga ecotone. Also, Heiskanen (2008) and Song et al. (2011) compared the MODIS VCF to other remotely sensed global datasets. Across all biomes and types of reference data, these independent assessments found that saturation of the optical signal, phenological noise, and confusion with dense herbaceous vegetation led to errors in the MODIS VCF between 10–31% Root-Mean-Square Error (RMSE), over-estimation in areas of low cover, and under-estimation in areas of high cover.

With increasing coverage worldwide, light detection and ranging (lidar) sensors now offer an additional and potentially superior means of reference data collection. Although *in situ* measurements remain the preferred solution in regions where lidar acquisition and processing are prohibited by cost, lidar provides highly accurate and

consistent measurements of vertical and horizontal canopy structure, including cover (Hopkinson and Chasmer 2009; Hudak et al. 2009; Sexton et al. 2009; Smith et al. 2010). However, no lidar-based evaluations of the MODIS or other continuous-field tree cover datasets yet exist.

We produced a global, 30-m resolution tree cover dataset for circa-2000 and 2005 epochs by rescaling the 250-m MODIS VCF Tree Cover dataset using Landsat images and ancillary data. We assessed the new, Landsat-based tree cover dataset's accuracy relative to lidar measurements and its consistency with its parent dataset, the MODIS VCF. To serve as a baseline of comparison and to help fill the dearth of error assessments of the MODIS VCF, we also performed a parallel accuracy assessment of the MODIS VCF against lidar measurements. In this paper, we describe our rescaling method and asses the 30-m, global tree cover product in comparison to the MODIS estimates and reference lidar measurements.

Methods

Generation of a Landsat-resolution tree cover dataset

Model

Tree cover (C) was estimated as a piecewise linear function of surface reflectance and temperature:

$$C_{i,t} = f\left(X_{i,t}\right) + \varepsilon, \tag{1}$$

where X is a vector of surface reflectance and temperature estimates; ε is error in the estimates produced by $f()$ applied to X; subscript i denotes the pixel's location in space, indexed by pixel; and t refers to its location in time, indexed by year. Continuous measurements, such as percent cover and surface reflectance, are robust to changes in resolution (Hilker et al. 2009; Gao et al. 2010; Feng et al. in press); although the data (X) were derived from Landsat, the model therefore makes no specification of scale and thus may be calibrated and applied at arbitrary, even different, resolutions between those of Landsat (30 m) and MODIS (250 m).

To estimate tree cover at 30-m resolution in 2000 and 2005, MODIS-based, 250-m tree cover estimates were overlaid on rescaled Landsat surface reflectance layers in each year, and a joint sample of cover and reflectance variables was drawn to generate a training dataset for each Landsat scene in each epoch (Figure 1). (Throughout, we refer to the data used to estimate model parameters as 'training' data, and we refer data whose accuracy is assumed as 'reference' data.)

The model was thus fit locally to each scene of the Landsat tiling system, the World Reference System 2 (WRS-2), in each epoch. The model was fit using the Cubist™ regression tree algorithm and applied using CubistSAM, an open-source parser for Cubist (Quinlan, 1993). Except for an allowance for extrapolation within the range [0,100], our application of regression trees was standard (i.e. neither sample boosting or bagging nor ensemble 'random forests' or 'committee models' were employed). Cubist – as well as regression trees in general – has been found to provide accurate estimates of percent-scale land cover attributes in numerous studies (e.g. Sexton et al. 2006, 2013a). Because regression trees can over-fit the data and there are often few data points at the extremes of the range of the response variable (e.g. tree cover), Cubist gives an option for either estimating within the range of the response
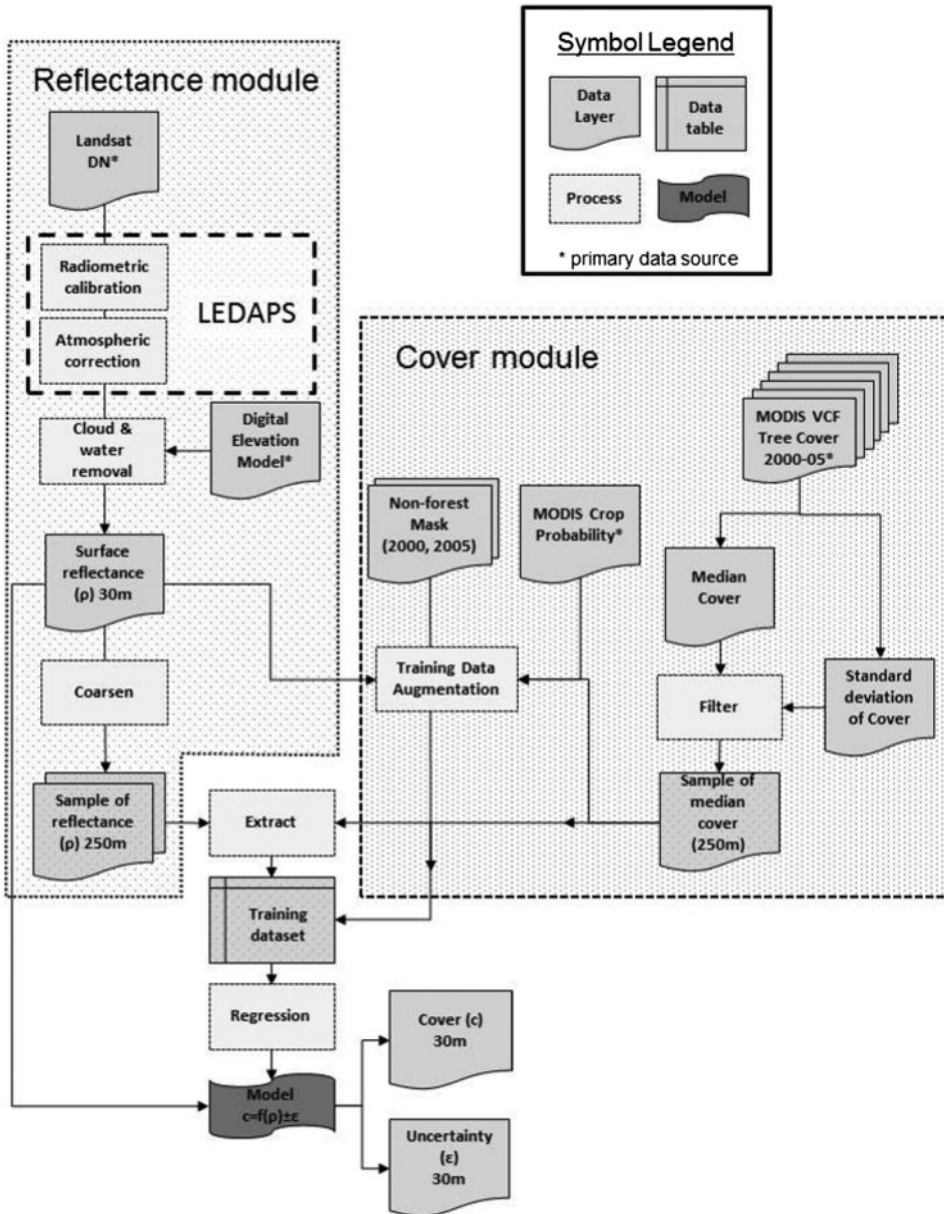
Figure 1. Flowchart of tree-cover rescaling algorithm. Arrows represent the directionality of the process, which flows generally from the top of the figure to the bottom. Shaded regions delineate distinct modules for preparation of reflectance and tree-cover data.

variable at each node (the default) or extrapolating within a specified range. To avoid over-fitting to the sometimes small samples at terminal nodes with extreme cover values, we allowed for extrapolation within the range of 0–100% tree cover. The fitted model was then applied to the original, 30-m Landsat data in order to estimate tree cover at the Landsat spatial resolution.

Data

Tree cover training data. 'Training' tree cover data for model fitting were derived primarily from the 250-m MODIS VCF Tree Cover layer (DiMiceli et al. 2011) from 2000 to 2005. Random errors (i.e. those which were not systematic, e.g. bias) were minimized by using the six-year median of cover for each pixel. Land cover changes between 2000 and 2005 were removed by calculating the standard deviation of annual tree cover estimates for each pixel over that interval and removing pixels in the top 10% of the distribution of standard deviations of each Landsat scene. Because only six years of MODIS VCF data were available, we used the median, which is a better representation of central tendency than the mean in small samples such as the six values of cover from 2000 to 2005.

Pure (i.e. 0% or 100%) and near-pure pixels are rare in the MODIS data, and tree cover tends to be over-estimated in areas of low cover, especially agricultural fields. To ameliorate under-representation of low tree cover in the training sample, we augmented the MODIS-derived reference data with information from the Training Data Automation and Support Vector Machines (TDA-SVM) automated classification algorithm (Huang et al. 2008) and the MODIS Cropland Probability Layer (Pittman et al. 2010). Cropland Probability and Tree Cover images were overlaid within each Landsat scene, and Landsat pixels with crop probability greater than 0.5 and tree cover less than 50% were selected. This selection comprised Landsat pixels with either crop or sparse vegetation cover. Within the selection, Landsat pixels identified by TDA-SVM as 'non-forest' in both 2000 and 2005 were assumed to be sparsely vegetated and were labeled as 0% tree cover. The remaining (i.e. crop) pixels in the selection were ranked by their NDVI values and divided into three sub-strata: high, medium, and low NDVI. Pixels from each of these sub-strata were randomly sampled such that the maximum proportion of Landsat 'crop' pixels was the proportion of MODIS pixels within the scene whose crop probability was greater than 60%. All of the sparsely vegetated pixels and the sample of crop pixels were then pooled with the MODIS-based reference data to form an ensemble training sample of tree cover and reflectance.

Landsat reflectance data. Surface reflectance for each epoch was retrieved from the 2000 and 2005 Global Land Survey (GLS) Landsat datasets. The GLS is a selection of Landsat and other images chosen to provide wall-to-wall, orthorectified, maximally cloud-free coverage of Earth's land area at 30-meter resolution in nominal 'epochs' of 1975, 1990, 2000, and 2005, and 2010 (Gutman et al. 2008; Franks et al. 2009). The GLS 2000 is composed of 8756 Landsat-7 Enhanced Thematic Mapper Plus (ETM +) images from 1999 to 2002, and the GLS 2005 is composed of 7284 gap-filled Landsat-7 images and 2424 Landsat-5 TM images acquired between 2003 and 2008. (We refer to these epochs throughout simply as their nominal years, '2000' and '2005'.)

GLS Landsat images were atmospherically corrected by the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (Masek et al. 2006) to estimate surface reflectance in each pixel. Atmospheric inputs and parameterization of LEDAPS are described by Feng et al. (in press). Clouds were identified and removed by optical-thermal detection, and cloud shadows by solar-geometry projection (Huang and Thomas 2010), surface reflectance pixels free of clouds, and cloud-shadows were

aggregated from 30-m to 250-m resolution by averaging all 30-m pixels within the extent of each 250-m pixel.

Error estimation

Uncertainty in every pixel was assessed relative to the training data by 10-fold cross-validation. Pixel-level uncertainty was quantified at each terminal node of the regression tree and assigned to pixels identified with that node. Because these pixel-level uncertainties were assessed relative to the reference data, errors between the reference data and actual cover are not accounted for at the pixel level. Error estimation relative to independent reference data derived from lidar is described in a later section. To assess each set of estimates relative to more direct measurements of actual cover, we compared each to approximately coincident measurements derived from small-footprint lidar measurements. (We use the term 'measurement' to refer to lidar-derived values of cover – which are calculated without statistical inference – and the more general 'estimate' to refer to values derived statistically from MODIS and Landsat images.) All comparisons were made at 250-m resolution, using MODIS estimates from 2005 and Landsat estimates from the 2005 epoch. Preliminary analyses comparing Landsat estimates to lidar measurements at 30-m resolution were not appreciably different than those reported here, although there was a small reduction of correlation believed to be due to spatial misregistration of Landsat data.

Uncertainty metrics were based on average differences between paired model and reference (or training) values (Willmott 1982), quantified by Mean Bias Error (MBE), Mean Absolute Error (MAE), and Root-Mean-Squared Error (RMSE):

$$\text{MBE} = \sum_{i=1}^{n} \frac{M_i - R_i}{n} \tag{2}$$

$$\text{MAE} = \sum_{i=1}^{n} \frac{|M_i - R_i|}{n} \tag{3}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n} (M_i - R_i)^2}{n}} \tag{4}$$

where M_i and R_i are estimated and reference tree cover values at a location i in a sample of size n.

After modeling the relationship between M and R by linear regression, their (squared) difference was disaggregated into systematic error (MSE_S) and unsystematic error (MSE_U) based on the modeled linear relationship (Willmott 1982):

$$MSE_S = \sum_{i=1}^{n} \frac{\left(\widehat{M}_i - R_i\right)^2}{n} \tag{5}$$

$$MSE_U = \sum_{i=1}^{n} \frac{\left(M_i - \widehat{M}_i\right)^2}{n} \tag{6}$$

where $\widehat{M}_i$ is the cover value predicted by the modeled relationship (Willmott 1982). Accuracy is thus quantified by the difference between the trend of model over

reference cover, and precision is quantified by the variation surrounding that trend. MSE_S and MSE_U sum to Mean-Squared Error (MSE), and therefore:

$$\text{RMSE} = \sqrt{MSE_s + MSE_u} \qquad (7)$$

(Willmott 1982). To maintain consistency, we report the square roots of MSE_S and MSE_U, i.e. RMSE_S and RMSE_U, in units of percent cover.

Tree cover reference data

For comparison to the 2005-epoch estimates, we collected small-footprint, discrete-return lidar measurements at four sites in a range of biomes (Figure 2): (1) La Selva Biological Station and its vicinity, Costa Rica (CR) in 2006; (2) the Wasatch Front in central Utah (UT), USA in 2008; (3) the Sierra National Forest in northern California (CA), USA in 2008; and (4) the Chequamegon-Nicolet National Forest, Wisconsin (WI), USA in 2005.

The Costa Rica site is dominated by tropical moist broadleaf evergreen forest surrounded by livestock pastures. The Utah site is an ecotone of temperate evergreen needle-leaf conifer forest, deciduous broadleaved shrubland, and annual grasses. The California site is dominated by tall, mixed-species temperate evergreen conifer forests of varying cover. The Wisconsin site is dominated by a mixture of temperate deciduous broadleaf hardwood and coniferous needle-leaf tree species with significant coverage of herbaceous agriculture, including corn. All lidar measurements were acquired during the growing season of each respective site, with mean point densities greater than 1 return/m^2. The Costa Rica dataset, collected in 2006, is described by Kellner et al. (2009), and the Wisconsin dataset is described by Cook et al. (2009). Figure 3 shows an example of the 3-dimensional distribution of lidar measurements in the California site. All sites were assessed visually for obvious
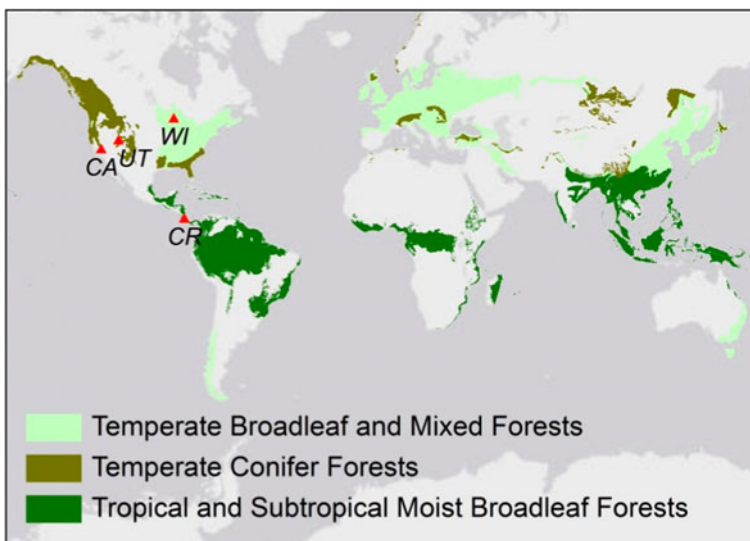

Figure 2. Distribution of lidar-based reference sites, overlaid on global biomes (Olson et al. 2001). Only the major habitat types intersecting reference sites are shown.
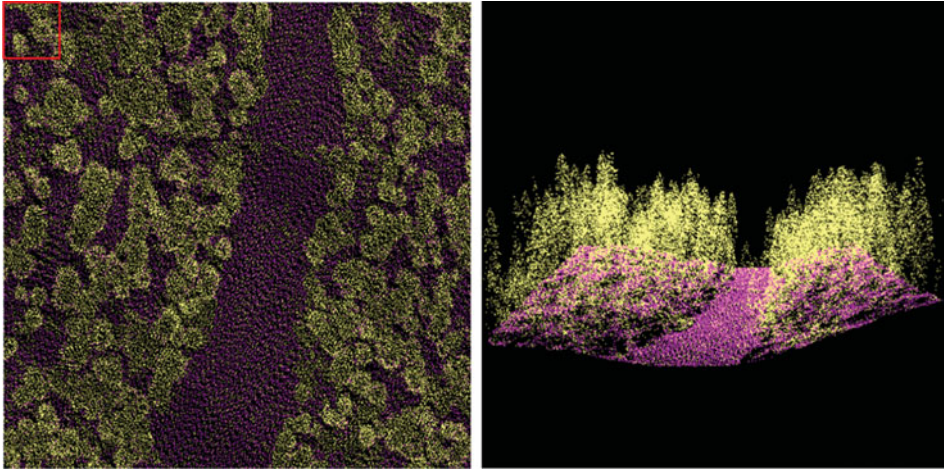
Figure 3. Three-dimensional distribution of a 250×250-m subset of the lidar measurements from the California reference site in nadir (left) and oblique (right) perspectives. Data points, which were sampled with intensity of approximately 13 points/m², are classified by height into tree (pink) and non-tree (yellow) classes. The red box in the upper-right corner shows the area of one 30-m Landsat pixel.

changes in cover between data acquisitions; in the WI dataset, obvious cover changes due to forest harvesting between Landsat and lidar acquisitions (totaling 21 pixels) were delineated manually and removed.

Tree cover (C) was calculated from lidar returns by dividing the number of returns above a criterion height by the total number of returns within a 10-m radius:

$$C = \frac{n_h}{n} \qquad (8)$$

where n is the number of returns and n_h is the number of returns above the specified height (h) (Korhonen et al. 2011). In accordance with the International Geosphere-Biosphere definition of forests, we specified the criterion $n_h = 5$ meters. Following calculation of tree cover at 10-m resolution, rasters were aggregated to 250-m resolution by averaging the values within the extent of each 250-m pixel. In pixels with steep underlying terrain (as might be likely especially in CA and UT), the varying ground elevation in large pixels can cause spurious detection of tree cover as lidar returns above 5-m height, first computing cover in small, 10-m pixels and then aggregating to 250-m pixels avoided this possibility. Also note that Relative Height (i.e. RH100) and other waveform-based metrics (Hyde et al. 2005; Dubayah et al. 2010) were not used; only height of the (discrete-return) lidar posts was used to calculate canopy height.

Results

Accuracy of the Collection-5 MODIS VCF tree cover estimates

Across the four biomes, MODIS estimates showed a positive, linear correspondence to lidar measurements of tree cover (Figure 4). Aggregating the sample across the
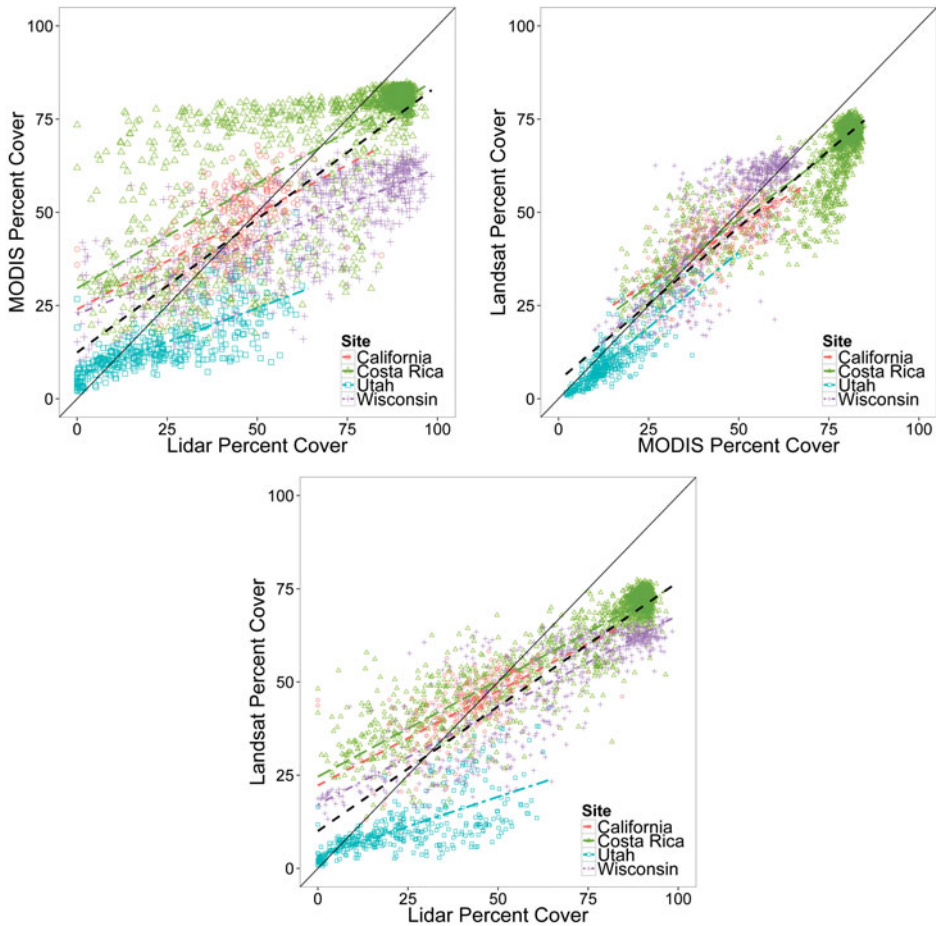
Figure 4. Scatterplots of estimated vs. reference and training tree-cover data: MODIS-based estimates vs. lidar-based measurements (top), Landsat-based vs. MODIS-based estimates (middle), and Landsat-based estimates vs. lidar-based measurements (bottom). Points and (dashed) regression lines are identified with sites by color, the overall (across-site) regression is in black, and the 1:1 line is solid black.

four sites, RMSE between MODIS estimates and lidar measurements was 16.83%, with cross-site MAE of 13.16% (Table 1).

RMSE and other metrics at each site followed this overall pattern, although there was variation among biomes (Table 2).

Some of the disagreement between MODIS estimates and lidar measurements was due to a slight negative bias (MBE = −6%) of the MODIS relative to lidar values. However, the difference was better expressed as a linear trend between the MODIS and lidar data, such that MODIS values exhibited positive bias at low cover and negative bias at high cover. This linear relationship was fairly strong – with R^2 of 0.70 and slope and intercept coefficients significantly different from zero (Table 3) – and consistent across sites.

Table 1. Across-site comparison of tree-cover estimates from MODIS, Landsat, and lidar.

Landsat	−5.57	−10.97
10.28 (8.42)	MODIS	−5.68
17.40 (15.23)	16.83 (13.16)	Lidar

Note: Values in the upper-right triangle of the matrix are Mean Bias Error (MBE). Values in the lower-left triangle are Root-Mean-Square Error (RMSE), with Mean Absolute Error (MAE) in parentheses. Biases between pairs of measurements (e.g. Landsat vs. lidar) are reported as the difference of the first element of the pair along the diagonal over the second – e.g. cover (Landsat) – cover (lidar).

The strength of the trend resulted in an approximately equal partition of uncertainty between systematic and random components, with $RMSE_S$ of 10.10% and $RMSE_U$ of 13.46% cover. Spatially, MODIS estimates replicated the pattern of lidar estimates with reasonable fidelity in the Costa Rica and Wisconsin sites, where patches of forest cover were large and discrete (Figure 5), although saturation of the MODIS values was clear in the Costa Rica site. However, the spatial correspondence was less clear in the California site and visibly poor in the Utah site, where tree cover is distributed as smooth, continuously varying gradients with shorter shrubs. Although the pattern was blurred greatly in the CA and UT ecotones, the loss of spatial pattern was slight in the CR and WI sites.

Consistency of Landsat- and MODIS-based (VCF) tree cover estimates

The relationship between Landsat estimates of tree cover and the MODIS data on which they were based was very strongly linear, near parity, and consistent among biomes (Figure 4). Relative to the MODIS estimates, Landsat estimates exhibited MBE of −6%, MAE of 8%, and RMSE of 10% cover (Table 1) in the biome samples of 2005 data. The modeled linear relationship explained 88% of the variation between the two datasets, and RMSE was equally partitioned between systematic and random components, with both $RMSE_S$ and $RMSE_U$ equaling approximately 7% cover (Table 3). Although significantly different from zero, the intercept of the linear relationship was relatively small (4.5%).

The global Landsat-MODIS VCF comparison for 2000 and 2005 epochs corroborated the aggregated site-specific results, with little difference between epochs (Figure 6). Paired Landsat- and MODIS-based estimates were distributed predominantly along the 1:1 line, with a slight under-estimation of Landsat- relative to

Table 2. Site-specific comparisons of tree-cover estimates from MODIS, Landsat, and lidar.

CR	Landsat	−8.37	−11.71	CA	Landsat	−2.16	1.31
	11.50 (10.13)	MODIS	−3.34		7.30 (5.90)	MODIS	3.48
	17.47 (16.33)	15.43 (12.06)	Lidar		8.38 (5.82)	10.55 (8.00)	Lidar
UT	Landsat	−3.85	−12.29	WI	Landsat	0.36	−13.247
	5.62 (4.47)	MODIS	−8.44		9.74 (6.79)	MODIS	−14.95
	17.64 (13.02)	14.63 (10.86)	Lidar		19.81 (17.39)	23.15 (20.13)	Lidar

Note: Values in the upper-right triangle of each sub-matrix are Mean Bias Error (MBE). Values in the lower-left triangle are Root-Mean-Square Error (RMSE), with Mean Absolute Error (MAE) in parentheses. Mean bias (MBE) between pixel-level canopy cover estimates (e.g. Landsat vs. lidar) is reported as the difference of the first element of the pair along the diagonal over the second – e.g. the MBE (Landsat, lidar) is reported as cover (Landsat) – cover (lidar).

Table 3. Linear regression summaries for pixel-level canopy cover estimates in four study areas.

Regression	Intercept (SE)	Slope (SE)	R^2	RMSE$_S$	RMSE$_U$
All sites					
MODIS–lidar	12.429 (0.549)	0.714 (0.008)	0.705	10.097	13.462
Landsat–MODIS	4.530 (0.323)	0.825 (0.005)	0.882	7.063	7.473
Landsat–lidar	10.016 (0.384)	0.668 (0.006)	0.811	14.637	9.406
Costa Rica ($n = 2044$)					
MODIS–lidar	29.621 (0.756)	0.561 (0.010)	0.628	11.242	10.573
Landsat–MODIS	12.477 (0.572)	0.710 (0.008)	0.804	9.765	6.066
Landsat–lidar	24.593 (0.380)	0.517 (0.004)	0.850	16.640	5.312
California ($n = 289$)					
MODIS–lidar	23.963 (1.835)	0.517 (0.042)	0.348	6.610	8.226
Landsat–MODIS	16.031 (1.548)	0.603 (0.033)	0.539	4.583	5.687
Landsat–lidar	22.248 (1.328)	0.506 (0.030)	0.494	5.893	5.955
Utah ($n = 425$)					
MODIS–lidar	6.069 (0.453)	0.365 (0.016)	0.552	13.556	5.500
Landsat–MODIS	−1.066 (0.372)	0.807 (0.022)	0.755	4.160	3.784
Landsat–lidar	3.316 (0.453)	0.318 (0.016)	0.483	16.766	5.492
Wisconsin ($n = 655$)					
MODIS–lidar	22.759 (0.888)	0.390 (0.013)	0.561	21.456	8.708
Landsat–MODIS	3.128 (1.384)[a]	0.941 (0.028)	0.619	0.856	9.699
Landsat–lidar	17.119 (0.809)	0.508 (0.012)	0.728	18.185	7.849

Note: RMSE$_U$ is mean 'unsystematic', or 'residual' error between original and calibrated measurements, and RMSE$_S$ is the 'systematic' error remaining between calibrated and reference measurements (see text for full explanation). Unless otherwise noted, all coefficients are significant at $Pr(>|t|) < 0.01$.
[a]$Pr(>|t|) = 0.024$.

MODIS-derived values of cover. Errors were slightly greater in the 2005 data than in the 2000 data (RMSE = 8.9% in 2000; RMSE = 11.9% in 2005), and the greatest differenceswere confined largely to the humid tropics, suggesting their origin might lie in the effects of remnant clouds in the Landsat images. The 2000 GLS 'epoch' of image collection was before the 2003 failure of the Scan-Line Corrector (SLC) of the ETM+ instrument, and so the quality of the GLS 2000 dataset likely benefitted from a greater selection of high-quality images from which to choose cloud-free data.

Error estimation of Landsat-based tree cover estimates

Across the four sampled biomes, the correspondence of Landsat estimates of tree cover to reference lidar measurements was similar to the relationship between MODIS and lidar data (Figure 4). Across the biomes, RMSE of Landsat estimates relative to lidar-measured cover was 17%, with MAE of 15% and MBE of −11% cover (Table 1). However, the overall linear relationship between Landsat estimates and lidar measurements was stronger ($R^2 = 0.81$) than that of MODIS estimates relative to lidar measurements ($R^2 = 0.71$). This strong linear trend resulted in a greater dominance of systematic (RMSE$_s = 15\%$), over unsystematic, or random noise (RMSE$_U = 9\%$) in the Landsat estimates compared to MODIS, suggesting a greater
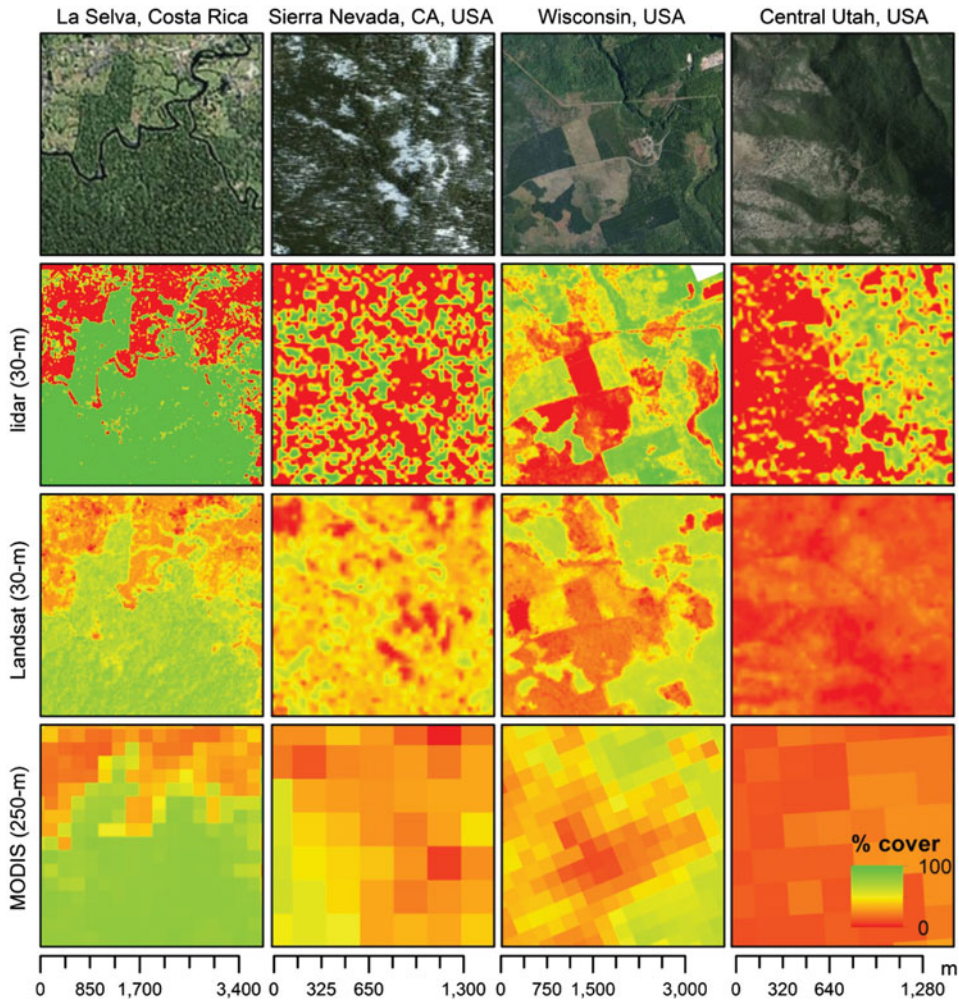
Figure 5. Spatial representation of tree cover by lidar measurements and Landsat and MODIS estimates in four sites. Imagery in top row was obtained from high-resolution, true color images provided by Microsoft Bing Maps.

potential for empirical calibration of Landsat estimates than is possible for the MODIS dataset. Although still present, saturation of Landsat estimates relative to lidar measurements was reduced slightly compared to the saturation seen in MODIS estimates.

Landsat estimates reproduced the spatial pattern of tree cover in most sites with greater fidelity than did MODIS estimates (Figure 5). The exception to this was the UT site, where there was no clear correspondence between either Landsat or MODIS estimates and the lidar measurements. Another shared artifact – i.e. recorded phenomenon caused by the algorithm but not reality – in both the Landsat and MODIS data was the slight compression of the actual frequency distribution of values, such that there were more intermediate values and correspondingly fewer values near the extremes of cover (i.e. 0% and 100%). It should be stressed, however,
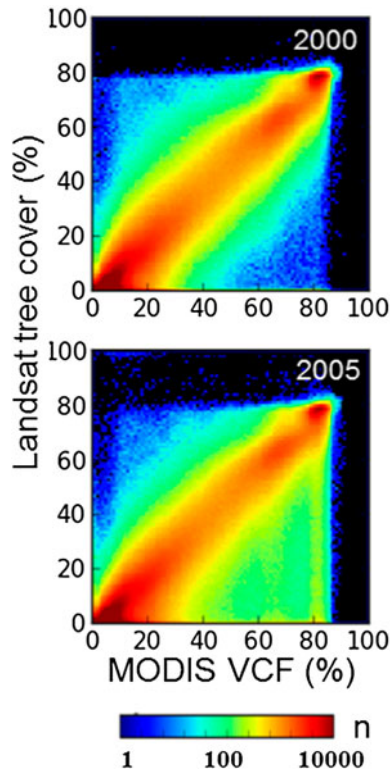
Figure 6. Joint distribution of a global sample of Landsat-based vs. MODIS-based (VCF) estimates of forest cover in 2000 (top) and 2005 (bottom).

that even considering the minor artifacts, Landsat estimates resolved greater spatial variation in tree cover than did the relatively coarse MODIS estimates.

Discussion

Figure 7 shows the first global, multi-temporal mapping of Earth's tree cover from Landsat-class data. Homer et al. (2004), Rollins (2009), and Hansen et al. (2011) describe continuous-field, Landsat-based datasets covering the United States, but there has been noconsistent, continuous-scale mapping of Earth's tree cover at Landsat (i.e. 30-m resolution and global extent) prior to this effort, nor has such a dataset been produced for multiple time periods. The greater than 8-fold increase in resolution from MODIS- to Landsat-class data – i.e. from 250-m to 30-m pixel side-length, or from 6.25-ha to 0.09-ha pixel area – greatly enhances discrimination of spatial patterns of forest cover, especially in highly fragmented landscapes (Figure 8). In agricultural regions, the increased accuracy of the Landsat-based estimates at the low end of tree cover relative to the parent MODIS data results in improved discrimination between forests and densely vegetated herbaceous crops (Figure 9). Further, as a longer-term consideration, the shifting of uncertainty from approximately equal distribution between systematic and unsystematic errors in the MODIS data toward dominance by systematic errors in the Landsat product will result in
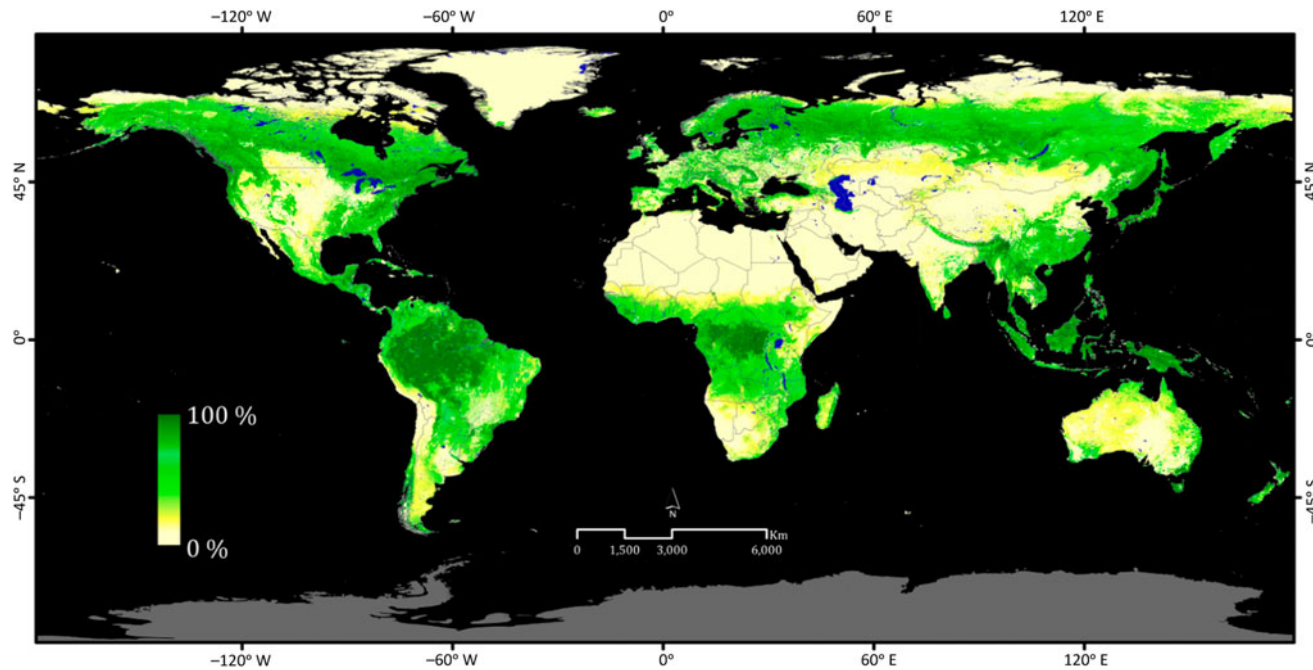
Figure 7. Global mosaic of Landsat-based estimates of tree cover. Map represents 2000 data. Gaps due to clouds and/or snow in each scene were filled first with Landsat-based data from overlapping paths, and the gaps still remaining were filled with data from the MODIS VCF Tree Cover layer in 2000.
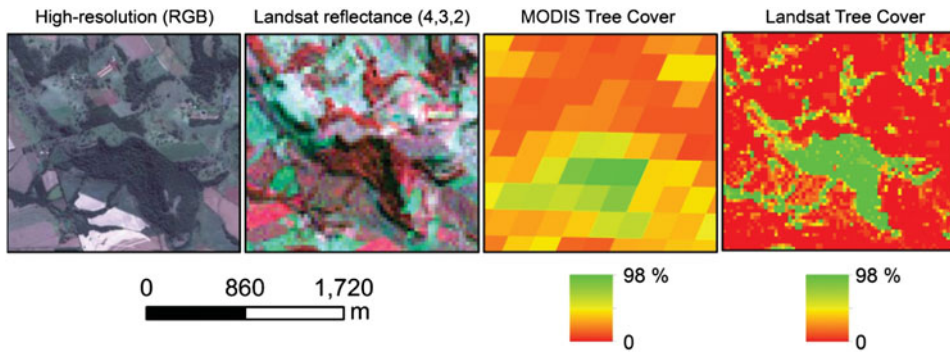
Figure 8. Resolution differences between MODIS-based and Landsat-based tree cover estimates in a highly fragmented landscape. Site shown is in Paraná, Brazil (Landsat WRS2 Path 224, Row 78; MODIS tile H13v11). Landsat and MODIS data are from images acquired in 2000.

greater potential for calibration to high-accuracy datasets such as lidar as they become available.

Accuracy of the MODIS VCF

Independent quantifications of the accuracy of the MODIS VCF (Hansen et al. 2003; DiMiceli et al. 2011) and other continuous-field datasets (e.g. Homer et al. 2004; Asner et al. 2009; Rollins 2009; Hansen et al. 2011) have been scarce. However, a global picture of the MODIS product's accuracy is beginning to emerge across this and previous studies. Our results corroborate those of earlier studies showing that the MODIS VCF dataset over-estimates tree cover in sparsely treed areas and under-estimates cover in dense forests. The latter is likely due to the saturation of the MODIS VCF tree cover estimates at approximately 80% cover, and the former is likely due to poor discrimination between trees and dense herbaceous vegetation.
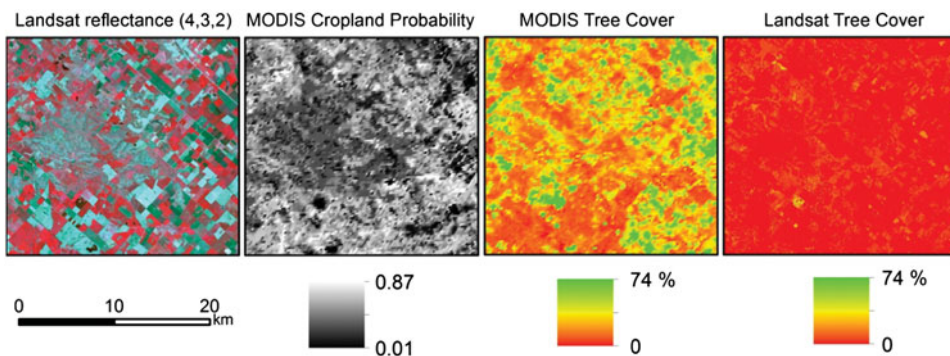

Figure 9. Comparison of accuracy between MODIS-based and Landsat-based tree cover estimates in an agricultural region. All data shown are from images acquired in 2000. Site shown is in the Buenos Aires Province, Argentina (Landsat WRS2 Path 225, Row 86; MODIS tile H13, V12) where tree cover is sufficiently low that any estimate greater than 10% is likely erroneous, the result of confusion between tree and crop cover.

Across the biomes studied, we found accuracies ranging between those of earlier studies (White et al. 2005; Montesano et al. 2009), and our results show that a large portion of the uncertainty of the MODIS VCF is systematic and thus can be removed by linear calibration to accurate reference data. There appears to be a limit to the precision with which tree-cover estimates can be inferred from passive-optical measurements. The accuracies we found for MODIS-based estimates relative to lidar-based 'truth' data were only slightly poorer than the 13% precision achieved among multiple observers visually interpreting tree cover from sub-meter-resolution imagery (Montesano et al. 2009).

Consistency of Landsat-based tree cover estimates with the MODIS VCF

Rescaling tree cover estimates from MODIS (250-m) to Landsat (30-m) resolution retained consistency with the original MODIS product and greatly improved resolution of spatial detail of surface features. Our 10% RMSE between Landsat and MODIS estimates was approximately double the 5% RMSE reported by Hansen et al. (2002) between MODIS estimates and the Landsat data on which they were based in western Zambia (a dry forest biome). However, the correspondence between our Landsat-derived and MODIS-derived estimates was greater than that of MODIS-based to lidar-based measurements, whether quantified as the coefficient of determination (R^2), uncertainty (RMSE), or either of the components of RMSE – inaccuracy ($RMSE_S$) or imprecision ($RMSE_U$). Landsat-based estimates also showed a strong potential for linear calibration to MODIS, although direct calibration to lidar – where available – would provide more accurate estimates of actual tree cover than calibration to MODIS estimates.

Accuracy of Landsat-based tree cover estimates

Although this initial assessment was based on a sample confined to North and Central America, where the density of the Landsat archive allows for optimal image selection, some general statements may be made concerning the accuracy of the product. Rescaling of MODIS estimates to Landsat resolution resulted in little loss of thematic information relative to actual cover when compared to the MODIS VCF overall (from RMSE = 16.8% for MODIS VCF to RMSE = 17.4% for Landsat-based estimates) and resulted in marked improvements in some cover types. Landsat estimates inherited uncertainties from their parent MODIS training data: all regression intercepts were greater than zero and all regression slopes were less than unity, reflecting the inability of current, uncalibrated MODIS and Landsat estimates to accurately resolve tree cover greater than 80%. However, incorporation of ancillary information (i.e. crop probability and non-forest masks) did reduce over-estimation of tree cover in agricultural regions (Figure 9). The relationship of errors among estimates from lidar, MODIS, and Landsat fit a general root-sum-of-squares error relationship:

$$\sqrt{(x-z)^2} = \sqrt{(x-y)^2+(y-z)^2} \tag{9}$$

Where z denotes an absolute, or 'truth' measurement, y denotes an estimate of z, and x denotes an estimate of y. Although overall errors relative to lidar were slightly

greater in the Landsat estimates, there was a stronger linear relationship between Landsat and lidar data ($R^2 = 0.811$) than between MODIS and lidar data ($R^2 = 0.705$), suggesting an improvement in the potential for calibration of Landsat-based estimates over that of the MODIS VCF. This general pattern held in all sites individually except UT ($R^2 = 0.552$ for MODIS, 0.483 for Landsat), which was characterized by steep topography and a complex ecotone of trees, shrubs, and grasses. Over all sites, RMSE after linear calibration ($RMSE_U$) could be expected to be 10% in Landsat, compared to 13% in MODIS, and when calibrated on a site-by-site basis, remnant uncertainty was less than 11% in MODIS-derived estimates and less than 6% in Landsat-derived estimates in each site.

A novel approach to global tree cover mapping

Challenges to global land cover mapping at coarse resolution have been summarized by a number of review papers (e.g. Townshend et al. 1994; Cihlar 2000; Fritz et al 2011), and issues surrounding global land cover monitoring at Landsat-class resolution are reviewed by Townshend et al. (2012). Much progress has been made in overcoming many challenges – e.g. automation and corrections for atmospheric, terrain, and phenological noise (Kim et al. 2011; Townshend et al. 2012; Sexton et al. 2013b). Nevertheless, generating continuous-field land cover datasets at sub-hectare resolution and global extent still faces one major obstacle: the difficulty of acquiring representative samples of reference data.

The traditional approach to continuous-field mapping has been to spatially aggregate sparse samples of high-resolution, categorical data to train models for estimating cover at spatially coarse, thematically continuous scales (e.g. Hansen et al. 2003; Homer et al. 2004; Hansen et al. 2011). Assuming that data quality is inversely proportional to spatial resolution, this general methodology gains its strength from large volumes of human-interpreted reference data. Although this approach has successfully produced the most widely adopted continuous-field datasets, generating continuous fields of land cover at Landsat-class resolution forlarge parts of the world – especially where reference data or human effort are limiting – requires a different strategy.

We have demonstrated an alternative approach for estimating tree cover at Landsat (30-m) resolution. Inverting the traditional methodology, our approach assumes the central tendency of coarse-scale, global data and uses this information to model a locally consistent, scale-insensitive relationship between reflectance and cover. The mathematical properties of continuous data have been known for some time (Stevens 1946), and recent studies have demonstrated consistent scaling behavior of remotely sensed continuous fields (Gao et al. 2006, 2010; Hilker et al. 2009; Feng et al. in press). Known errors in the training data – e.g. over-estimation of tree cover by the MODIS VCF in dense herbaceous cover – are ameliorated by augmenting the training sample with ancillary data – e.g. MODIS Crop Probability. Thus a coarse, global, continuous-field dataset is used to infer cover at finer resolution, and the highest-quality reference data are reserved for model accuracy assessment.

The rarity of high-quality reference data is likely responsible in large part for the lack of published studies validating the MODIS VCF. However, this need is increasingly being served by a piecemeal, but growing global patchwork of

small-footprint lidar datasets from which vegetation structural attributes can be measured or calculated with high reliability. Insofar as errors in land cover data transmit to all downstream products and analyses, these quality assessments provide the critical foundation for tracking uncertainty from Earth observation to modeling and prediction. We are optimistic that the increasing availability of intensive, local lidar datasets will result in improved accuracy of mapping and monitoring tree cover at global scales.

Conclusions

The MODIS VCF tree layer, now in version 5, is among the most useful and reliable global datasets representing Earth's woody vegetation. However, it suffers from a coarse spatial scale relative to many land cover patterns and changes, and there have been extremely few attempts to validate the MODIS VCF against independent reference data. In this study, we rescaled the 250-m MODIS estimates to 30-m resolution using the 2000 and 2005 Global Land Survey Landsat dataset and validated both this new Landsat-based global tree cover dataset as well as its parent MODIS data against measurements derived from lidar. The MODIS VCF exhibits uncertainty of approximately 17 percentage points RMSE across all biomes studied and between 11 and 23 points in each biome individually. Much of this uncertainty is systematic, due to over-estimation in sparsely treed (e.g. agricultural) regions and under-estimation in dense forests. The Landsat-based tree cover dataset exhibited high thematic fidelity to the MODIS VCF – including saturation at approximately 80% tree cover and uncertainty where woody cover is a gradient of trees and shrubs – but showed improved accuracy in agricultural regions. At 30-m resolution, the Landsat-based dataset also greatly improved resolution of spatial patterns of tree cover in highly fragmented landscapes. Both the MODIS- based and Landsat-based tree cover datasets show strong potential for empirical calibration to lidar measurements, but the potential for improved accuracy through calibration is slightly greater in the Landsat dataset. Calibration coefficients for Landsat-based and MODIS-based tree cover datasets are provided, both globally and for sites representing each biome. The Landsat-based tree cover dataset – global estimates of percent tree cover at 30-m resolution in 2000 and 2005 epochs – is available to the Earth observation, modeling, and analysis community for free download at the Global Land Cover Facility (www.landcover.org).

References

Asner, G. P., D. E. Knapp, E. N. Broadbent, P. J. C. Oliviera, M. Keller, and J. N. Silva. 2005. "Selective Logging in the Brazilian Amazon." *Science* 310 (5747): 480–482. http://www.sciencemag.org/content/310/5747/480.short

Asner, G. P., D. E. Knapp, A., Balaji, and G. Pez-Acosta. 2009. "Automated Mapping of Tropical Deforestation and Forest Degradation: CLASlite." *Journal of Applied Remote Sensing* 3 (1): 033543.

Bennett, B. 2001. "What is A Forest? On the Vagueness of Certain Geographic Concepts." *Topoi* 20: 189–201. http://link.springer.com/article/10.1023%2FA%3A1017965025666?LI=-true#page-1

Cihlar, J. 2000. "Land Cover Mapping of Large Areas from Satellites: Status and Research Priorities." *International Journal of Remote Sensing* 21 (6–7): 1093–1114.

Conde, D. A., F. Colchero, H. Zarza, N. L. Christensen, Jr., J. O. Sexton, C. Manterola, C. Chávez, et al. 2010. "Sex Matters: Modeling Male and Female Habitat Differences for Jaguar Conservation." *Biological Conservation* 143 (9): 1980–1988.

Cook, B. D., P. V. Bolstad, E. Naesset, R. S. Anderson, S. Garrigues, J. T. Morissette, J. Nickeson, and K. J. Davis. 2009. "Using LiDAR and quickbird Data to Model Plant Production and Quantify Uncertainties with Wetland Detection and Land Cover Generalizations." *Remote Sensing of Environment* 113 (11): 2366–2379.

DeFries, R.S., C.B. Field, I. Fung, and C.O. Justice. 1995. "Mapping the Land Surface for Global Atmosphere-biosphere Models: Toward Continuous Distributions of Vegetation's Functional Properties." *Journal of Geophysical Research* 100 (D10): 20867–20882.

DeFries, R. S., J. R. G. Townshend, and M. C. Hansen. 1999. "Continuous Fields of Vegetation Characteristics at the Global Scale at 1-km Resolution." *Journal of Geophysical Research* 104 (D14): 16911–16923.

Defries, R., A. Hansen, A. C. Newton, and M. C. Hansen. 2005. "Increasing Isolation of Protected Areas in Tropical Forests over the Past Twenty Years." *Ecological Applications* 15 (1): 19–26.

DiMiceli, C. M., M. L. Carroll, R. A. Sohlberg, C. Huang, M. C. Hansen, and J. R. G. Townshend. 2011. Annual Global Automated MODIS Vegetation Continuous Fields (MOD44B) at 250 m Spatial Resolution for Data Years Beginning Day 65, 2000 – 2010, Collection 5 Percent Tree Cover, University of Maryland, College Park, MD, USA. http://www.landcover.org/data/vcf/

Dubayah, R. O., S. L. Sheldon, D. B. Clark, M. A. Hofton, J. B. Blair, G. C. Hurtt, and R. L. Chazdon. 2010. "Estimation of Tropical Forest Height and Biomass Dynamics Using lidar Remote Sensing at La Selva, Costa Rica." *Journal of Geophysical Research* 115: G00E09. http://onlinelibrary.wiley.com/doi/10.1029/2009JG000933/full

Feng, M., J. O. Sexton, C. Huang, J. G. Masek, E. F. Vermote, F. Gao, R. Narasimhan, S. Channan, R. E. Wolfe, and J. R. Townshend. Forthcoming. "Global, Long-term Surface Reflectance Records from Landsat: A Comparison of the Global Land Survey and MODIS Surface Reflectance Datasets." *Remote Sensing of Environment*.

Franks, S., J. G. Masek, J. G. Headley, J. Gasch, and T. Arvidson. 2009. "Large Area Scene Selection Interface (LASSI): methodology of selecting Landsat imagery for the Global Land Survey 2005." *Photogrammetric Engineering & Remote Sensing* 75: 1287–1296. http://www.asprs.org/a/publications/pers/2009journal/november/2009_nov_1287-1296.pdf

Fritz, S., L. See, I. McCallum, C. Schill, M. Obersteiner, M.Van der Velde, H. Boettcher, et al. 2011. "Highlighting Continued Uncertainty in Global Land Cover Maps for the User Community." *Environmental Research Letters* 6 (4): 044005.

Gao, F., J. Masek, M. Schwaller, and F. Hall. 2006. "On the Blending of the Landsat and MODIS Surface Reflectance: Predicting Daily Landsat Surface Reflectance." *IEEE Transactions on Geoscience and Remote Sensing* 44 (8): 2207–2218.

Gao, F., J. G. Masek, R. E. Wolfe, and C. Huang. 2010. "Building a Consistent Medium Resolution Satellite Data Set Using Moderate Resolution Imaging Spectroradiometer Products as Reference." *Journal of Applied Remote Sensing* 4 (1): 043526.

Gutman, G., R. Byrnes, J. Masek, S. Covington, C. Justice, S. Franks, and R. Headley. 2008. "Towards Monitoring Land-cover and Land-use Changes at a Global Scale: The Global Land Survey 2005." *Photogrammetric Engineering & Remote Sensing* 74 (1): 6–10. http://gls. umd.edu/documents/GLS_2005_PERS_Jan2008_Gutman.pdf

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, L. Marufu, and R. Sohlberg. 2002. "Development of a MODIS Tree Cover Validation Data Set for Western Province, Zambia." *Remote Sensing of Environment* 83 (1–2): 320–335.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, M. Carroll, C. Dimiceli, and R. A. Sohlberg. 2003. "Global Percent Tree Cover at a Spatial Resolution of 500 Meters: First Results of the MODIS Vegetation Continuous Fields Algorithm." *Earth Interactions* 7 (10): 1–15.

Hansen, M. C., S. V. Stehman, P. V. Potapov, T. R. Loveland, J. R. G. Townshend, R. S. Defries, K. W. Pittman, et al. 2008. "Humid Tropical Forest Clearing from 2000 to 2005 Quantified by Using Multitemporal and Multiresolution Remotely Sensed Data SCIENCE." *Proceedings of the National Academy of Sciences of the United States of America* 105 (27): 9439–9444.

Hansen, M. C., A. Egorov, D. P. Roy, P. Potapov, J. JuP., S. Turubanova, I. Kommareddy, et al. 2011. "Continuous Fields of Land Cover for the Conterminous United States Using Landsat Data: First Results from the Web-Enabled Landsat Data (WELD) Project." *Remote Sensing Letters* 2 (4): 279–288.

Harris, N. L., S. Brown, S. C. Hagen, S. S. Saatchi, S. Petrova, W. Salas, M. C. Hansen, et al. 2012. "Baseline Map of Carbon Emissions from Deforestation in Tropical Regions." *Science* 336: 1573–1576. http://www.sciencemag.org/content/336/6088/1573.short

Heiskanen, J. 2008. "Evaluation of Global Land Cover Data Sets of the Tundra-taiga Transition Zone in Northernmost Finland." *International Journal of Remote Sensing* 29 (13): 3727–3751.

Hijmans, R. J., and J. V. Etten. 2012. raster: Geographic analysis and modeling with raster data. R package version 2.0-05. http://CRAN.R-project.org/package=raster

Hilker, T., M. A. Wulder, N. C. Coops, N. Seitz, J. C. White, F. Gao, J. G. Masek, and G. Stenhouse. 2009. "Generation of Dense Time Series Synthetic Landsat Data through Data Blending with MODIS through a Spatial and Temporal Adaptive Reflectance Fusion Model." *Remote Sensing of Environment* 113 (9): 1988–1999.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan. 2004. "Development of a 2001 National Land-Cover Database for the United States." *Photogrammetric Engineering & Remote Sensing* 70 (7): 829–840. http://129.115.102.107/LRSG/Teaching/ES6973/July04_PERS.pdf

Hopkinson, C., and L. Chasmer. 2009. "Testing LiDAR Models of Fractional Cover across Multiple Forest Ecozones." *Remote Sensing of Environment* 113 (1): 275–288.

Huang, C., K. Song, S. Kim, J. R. G. Townshend, P. Davis, J. G. Masek, and S. N. Goward. 2008. "Use of A Dark Object Concept and Support Vector Machines to Automate Forest Cover Change Analysis." *Remote Sensing of Environment* 112 (3): 970–985.

Huang, C., S. N. Goward, K. S. N. Schleeweis, N. Thomas, J. G. Masek, and Z. Zhu. 2009. "Dynamics of National Forests Assessed Using the Landsat Record: Case Studies in Eastern United States." *Remote Sensing of Environment* 113 (7): 1430–1442.

Huang, C., and N. Thomas. 2010. "Automated Masking of Cloud and Cloud Shadow for Forest Change Analysis Using Landsat Images." *International Journal of Remote Sensing* 31: 37–41. http://www.tandfonline.com/doi/abs/10.1080/01431160903369642

Hudak, A. T., J. S. Evans, and A. M. S. Smith. 2009. "LiDAR Utility for Natural Resource Managers." *Remote Sensing* 1 (4): 934–951.

Hyde, P., R. Dubayah, B. Peterson, J. Blair, M. Hofton, C. Hunsaker, R. Knox, and W. Walker. 2005. "Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems." *Remote Sensing of Environment* 96 (3–4): 427–437.

Kellner, J. R., D. B. Clark, and M. A. Hofton. 2009. "Canopy Height and Ground Elevation in a Mixed-land-use Lowland Neotropical Rain Forest Landscape." Ecological Archives E090-233.

Kim, D.-H., R. Narasimhan, J. O. Sexton, C. Huang, and J. R. Townshend. 2011. "Methodology to Select Phenologically Suitable Landsat Scenes for Forest Change Detection." *International Geoscience and Remote Sensing Symposium (IGARSS)*: 2613–2616, July 24–29. http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=6049738&tag=1

Korhonen, L., I. Korpela, J. Heiskanen, and M. Maltamo. 2011. "Airborne Discrete-return LIDAR Data in the Estimation of Vertical Canopy Cover, Angular Canopy Closure and Leaf Area Index." *Remote Sensing of Environment* 115 (4): 1065–1080.

Lawrence, P. J., and T. N. Chase. 2007. "Representing a New MODIS Consistent Land Surface in the Community Land Model (CLM 3.0)." *Journal of Geophysical Research* 112 (G1): G01023.

Mansfield, C., S. K. Pattanayak, W. Mcdow, R. Mcdonald, and P. Halpin. 2005. "Shades of Green : Measuring the Value of Urban Forests in the Housing Market." *Journal of Forest Economics* 11 (3): 177–199.

Masek, J. G., E. F. Vermote, N. E. Saleous, R. Wolfe, F. G. Hall, K. F. Huemmrich, F. Gao, et al. 2006. "A Landsat Surface Reflectance Dataset for North America, 1990–2000." *IEEE Geoscience and Remote Sensing Letters* 3 (1): 68–72.

Miles, L., A. C. Newton, R. S. DeFries, C. Ravilious, I. May, S. Blyth, V. Kapos, et al. 2006. "A Global Overview of the Conservation Status of Tropical Dry Forests." *Journal of Biogeography* 33 (3): 491–505.

Montesano, P. M., R. Nelson, G. Sun, H. Margolis, A. Kerber, and K. J. Ranson. 2009. "MODIS Tree Cover Validation for the Circumpolar Taiga–Tundra Transition Zone." *Remote Sensing of Environment* 113 (10): 2130–2141.

Olson, D. M., E. Dinerstein, E. D. Wikramanayake, N. D. Burgess, G. V. N. Powell, E. C. Underwood, J. A. D'amico, et al. 2001. "Terrestrial Ecoregions of the World: A New Map of Life on Earth." *BioScience* 51: 933–938.

Pittman, K., M. C. Hansen, I. Becker-Reshef, P. Potapov, and C. O. Justice. 2010. "Estimating Global Cropland Extent with Multi-year MODIS Data." *Remote Sensing* 2 (7): 1844–1863.

Quinlan, J. R. 1993. C4.5: Programs for Machine Learning. San Francisco, CA: Morgan Kaufmann.

R Core Team. 2012. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.r-project.org/

Rollins, M. G. 2009. "LANDFIRE: A Nationally Consistent Vegetation, Wildland Fire, and Fuel Assessment." *International Journal of Wildland Fire* 18 (3): 235–249.

Sexton, J. O., R. D. Ramsey, and D. L. Bartos. 2006. "Habitone Analysis of Quaking Aspen in the Utah Book Cliffs: Effects of Site Water Demand and Conifer Cover." *Ecological Modelling* 198 (3–4): 301–311.

Sexton, J. O., T. Bax, P. Siqueira, J. J. Swenson, and S. Hensley. 2009. "A Comparison of lidar, Radar, and Field Measurements of Canopy Height in Pine and Hardwood Forests of Southeastern North America." *Forest Ecology and Management* 257 (3): 1136–1147.

Sexton, J. O., X.-P. Song, C. Huang, S. Channan, M. E. Baker, and J. R. Townshend. 2013a. "Urban Growth of the Washington, D.C. – Baltimore, MD Metropolitan Region from 1984 to 2010 by Annual Estimates of Impervious Cover." *Remote Sensing of Environment* 129: 42–53.

Sexton, J. O., D. L. Urban, M. J. Donohue, and C. Song. 2013b. "Long-term Landcover Dynamics by Multi-temporal Classification Across the Landsat-5 Record." *Remote Sensing of Environment* 128: 246–258.

Simard, M., N. Pinto, J. B. Fisher, and A. Baccini. 2011. "Mapping forest canopy height globally with spaceborne lidar." *Journal of Geophysical Research* 116 (G4): G04021.

Smith, A. M. S., M. J. Falkowski, A. T. Hudak, J. S. Evans, A. P. Robinson, and C. M. Steele. 2010. "A Cross-comparison of Field, Spectral, and lidar Estimates of Forest Canopy Cover." *Canadian Journal of Remote Sensing* 36 (5): 447–459. http://pubs.casi.ca/doi/abs/10.5589/m09-038

Song, X., C. Huang, J. O. Sexton, M. Feng, R. Narasimhan, S. Channan, and J. R. Townshend. 2011. "An Assessment of Global Forest Cover Maps Using Regional Higher-resolution Reference Datasets." *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, 752–755.

Stevens, S. S. 1946. "On the Theory of Scales of Measurement." *Science* 103 (2684): 677–680. http://www.sciencemag.org/content/103/2684/677.full.pdf?sid=736d96d6-36b3-4a8b-8067-4540365b2473

Townshend, J. R. G., and C. O. Justice. 1988. "Selecting the Spatial Resolution of Satellite Sensors Required for Global Monitoring of Land Transformations." *International Journal of Remote Sensing* 9 (2): 187–236.

Townshend, J. R. G., C. O. Justice, D. Skole, J.-P. Malingreau, J. Cihlar, P. Teillet, F. Sadowski, et al. 1994. "The 1 km Resolution Global Data Set: Needs of the International Geosphere Biosphere Programme." *International Journal of Remote Sensing* 15 (17): 3417–3441.

Townshend, J. R., J. G. Masek, C. Huang, E. F. Vermote, F. Gao, S. Channan, J. O. Sexton, et al. 2012. "Global Characterization and Monitoring of Forest Cover Using Landsat Data: Challenges and Opportunities." *International Journal of Digital Earth* 5 (5): 373–397.

Trainor, A. M., J. R. Walters, W. F. Morris, J. Sexton, and A. Moody. Forthcoming. "Empirically Estimating a Dispersal Resistance Surface: A Case Study with Red-cockaded Woodpeckers." *Landscape Ecology*. http://link.springer.com/article/10.1007/s10980-013-9861-5#page-1

White, M. A., J. D. Shaw, and R. D. Ramsey. 2005. "Accuracy Assessment of the Vegetation Continuous Field Tree Cover Product Using 3954 Ground Plots in the South-western USA." *International Journal of Remote Sensing* 26 (12): 2699–2704.

Willmott, C. J. 1982. "Some Comments on the Evaluation of Model Performance." *Bulletin of the American Meteorological Society* 63 (11): 1309–1313.