Remote Sensing of Environment 185 (2016) 221–232

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Mapping tree height distributions in Sub-Saharan Africa using Landsat 7 and 8 data

Matthew C. Hansen [a,*], Peter V. Potapov [a], Scott J. Goetz [b], Svetlana Turubanova [a], Alexandra Tyukavina [a], Alexander Krylov [a], Anil Kommareddy [a], Alexey Egorov [c]

[a] University of Maryland, United States
[b] Woods Hole Research Center, United States
[c] South Dakota State University, United States

ARTICLE INFO

Article history:
Received 2 August 2015
Received in revised form 26 January 2016
Accepted 11 February 2016
Available online 2 March 2016

ABSTRACT

Landsat time-series multi-spectral data, GLAS (Geoscience Laser Altimeter System) height data and a regression tree model were used to estimate tree height for a transect in Sub-Saharan Africa ranging from the Sahara Desert through the Congo Basin to the Kalahari Desert (+22 to −22° latitude and 23 to 24° longitude). Objectives included comparing the performance of Landsat 7- and 8-derived inputs separately and combined in mapping tree height at a regional scale, assessing the relative value of good observation counts and different Landsat spectral inputs for tree height estimation across a range of environments, and describing tree height distributions and discontinuities in Sub-Saharan Africa. A total of 5371 images were processed and per pixel quality assessed to create a set of multi-temporal metrics for the 2013 and 2014 calendar years for Landsat 7 only, Landsat 8 only and both Landsat 7 and 8 combined. Differences in performance were slight between different sensor inputs. However, performance generally improved with increasing numbers of good observations. Metrics derived from red reflectance data contributed most in estimating tree height. The regression tree algorithm accurately reproduced the LiDAR-derived height training data with an overall mean absolute error (MAE) for tree height estimation of 2.45 m using integrated Landsat 7 and 8 data. Significant underestimations were quantified for tall tree cover (MAE of 4.65 m for >20 m heights) and overestimations for low/no tree cover (MAE 1.61 for <5 m heights). Resulting tree distributions were found to be discontinuous with a primary dry seasonal woodlands cluster of 5–10 m in height, a second cluster of primarily dry evergreen forest tree cover from 11–17 m, and a third cluster of humid evergreen forest tree cover ≥18 m. The integration of Landsat 7 and 8 and forthcoming Sentinel 2 time-series optical data to extend the value of LiDAR forest structure measurements is recommended.

1. Introduction

The Operational Land Imager (OLI) onboard the Landsat 8 spacecraft represents a significant advance in earth observation monitoring capability (Loveland and Irons, this issue). Additional spectral bands, enhanced bandwidths, increased signal to noise, greater radiometric detail, and other technical advances make Landsat 8 the most robust in the Landsat series. While Landsat 8 data are demonstrably technically superior to preceding Landsat sensors, we believe that the integration of data from different Landsat missions in order to provide denser time-series information to be critical to advancing monitoring capabilities. The high cadence, high spatial detail information domain outlined by Cihlar (2000) will most likely be met through a virtual constellation approach. Currently, Landsat 7 and 8 are available for integrated use. The Landsat-like spectral bandwidths of the Sentinel 2 satellites (Drusch et al., 2012) promise a further enrichment of 10–30 m time-series multi-spectral data.

With the opening of the Landsat archive, numerous research efforts have begun to exploit time-series of Landsat imagery in characterizing land change (Kennedy, Cohen, & Schroeder, 2007; Huang et al., 2009; Potapov et al., 2012; Zhu & Woodcock, 2014). However, the uneven acquisition of Landsat imagery through time and across space and missions poses a significant challenge to developing turnkey land cover characterization models such as those of the MODIS Land Science Team (Justice et al., 2002). Landsat data richness varies geographically as a function of the application and variation of acquisition strategies, cloud cover, and length of orbit over land. Fig. 1 shows annual growing season coverage for the Landsat 7 record through 2012 and illustrates the historic paucity of data for many parts of Africa, especially the dry tropics. The most data rich period of Landsat starts in 2013 with the combined Landsat 7 and 8 image archive. While variations in data richness due to cloud cover will always impact passive optical sensors such as Landsat, the integration of 8-day repeat Landsat 7 and 8 and

* Corresponding author.

http://dx.doi.org/10.1016/j.rse.2016.02.023

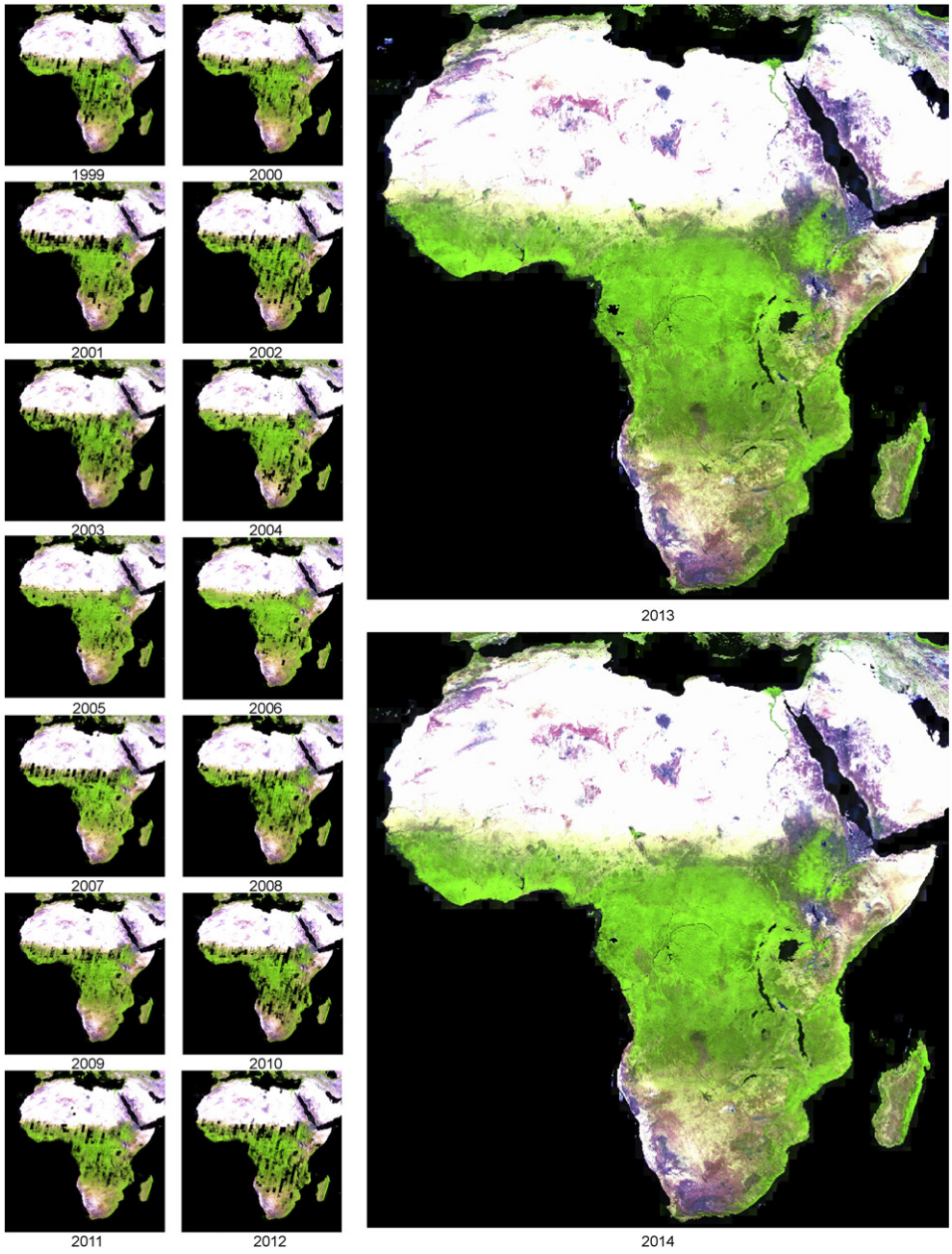

Fig. 1. Annual growing season median normalized reflectance from Landsat 7 Enhanced Thematic Mapper data (ETM+) from 2000 to 2014. Black areas in the interior of Africa represent data gaps, i.e. no viable land observations for that year.

forthcoming Sentinel 2 data will likely result in a MODIS-like turnkey large area mapping and monitoring capability.

One approach to exploiting time-series observations for mapping land cover is to convert all viable data for a given study period into multi-temporal metrics. Metrics are statistical derivatives of time-series observations that are not tied to a specific time of year and have been used with AVHRR, MODIS and Landsat data (DeFries, Hansen, & Townshend, 1995; Hansen, DeFries, Townshend, & Sohlberg, 2000; Hansen et al., 2003; Potapov et al., 2012, 2015). Examples include peak greenness or range of near-infrared reflectance. Metrics have the advantage of tolerating unequal data richness; for example, quantiles can be calculated for a population of 10 or 100 input observations. Since metrics are not tied to a specific date of year, they are appropriate for regional to global scale mapping. Africa, in particular, is a suitable region for the application of metrics due to the mirrored phenology north and south of the equator; metrics reorder input time-series observations, normalizing phenological variation without respect to the timing of phenology and have been shown to perform better than time-sequential data in mapping African tree cover (Hansen, Townshend, DeFries, & Carroll, 2005). In this study, we assess the value of metrics derived from Landsat 7 and 8 data streams to map tree height in Sub-Saharan Africa.

For large area mapping, we employ decision tree models, either classification trees for categorical land cover themes, such as forest and non-forest, or regression trees for continuously varying themes, such as percent tree cover. Tree models are robust, distribution-free algorithms that are easily implemented and transparent in their operation. Tree models have been used for the suite of land cover products from the MODIS Land Science Team, (Friedl et al., 2002; Hansen et al., 2002) and more recently in our global forest study using Landsat 7 Enhanced Thematic Mapper Plus (ETM+) data (Hansen et al., 2013). Hanan et al. recently posited that classification and regression tree (CART) models can produce misleading results due to the possibility that they "impose discontinuities" in data where there "may be none

present in reality" (Hanan, Tredennick, Prihodko, Bucini, & Dohn, 2014). This supposition was in response to the study of Staver, Archibald, and Levin (2011), who examined a CART-derived tree cover map (Hansen et al., 2003) and found that sub-Saharan African savanna and forest may be alternative stable states due to the interactions of climate and disturbance, largely fire. Sub-Saharan tree distribution differs greatly from potential woody cover when modeling tree distributions based on precipitation as a predictor (Sankaran et al., 2005). The presented study does not address the ecological factors determining tree distributions in Sub-Saharan Africa, but instead seeks to confirm if Sub-Saharan tree distributions are discontinuous. To do so, we employ Landsat 7 and 8 data using our standard approach for large area mapping, bagged regression trees using CART, to characterize tree heights in Africa. At 30 m spatial resolution, a more discrete depiction of landscape heterogeneity is enabled compared to MODIS and other coarser data sources. To ensure an independent reference, we use as our calibration data LiDAR-derived height measurements from the GLAS sensor onboard the ICESat satellite (Zwally et al., 2002).

Mapping tree heights with multi-spectral imagery is a relatively new application and is dependent on integrating synoptic coverage optical data with samples of height data, often from LiDAR-derived reference data. Optical data are considered to be sensitive to land cover properties in the horizontal plane and relatively insensitive to vertical structure (Cohen & Spies, 1992). However, shadowing effects from within canopy vertical structure lead to light extinction and the ability to accurately discriminate forests from non-forest cover. A number of very different land cover types can have the same spectral signature at any given time, confounding the identification of tree cover. For example, inundated grasslands and some kinds of intensively managed cropland can have very similar spectral responses to dense tree cover. However, when including time-series spectral information, these confusions can be overcome and tree cover discriminated. Exploiting both the temporal and spectral information domains offers the possibility of advancing the use of Landsat in the characterization of vegetation vertical structure. The key to doing so is the availability of suitable calibration and validation data, now increasingly available in the form of LiDAR (light detection and ranging) derived height metrics.

There is a growing literature on integrating samples of LiDAR data with synoptic optical data coverages. Hudak, Lefsky, Cohen, and Berterretche (2002) evaluated several methods for integrating LiDAR and Landsat data in characterizing forest structure attributes. Asner et al. advanced this type of integration to quantify forest biomass at a sub-national scale (Asner et al., 2012). Ota et al. (2014) applied similar methods in estimating change in vertical structure over time. Hilker, Wulder, and Coops (2008) related LiDAR data to very high spatial resolution optical imagery in demonstrating their utility in estimating forest stand height and cover and Lefsky (2010) and Simard, Pinto, Fisher, and Baccini (2011) integrated LiDAR and MODIS data to create global forest height maps. This paper builds on the product reported by Tyukavina et al. (2015) that employed annual Landsat 7 growing season composite images and height data from the Geoscience Laser Altimeter System (GLAS) to estimate pan-tropical forest height. Here we improve upon the Landsat inputs, moving from a single growing season composite to a set of multi-temporal metrics derived from Landsat 7 and 8 data. We test the method by characterizing tree height distributions in Sub-Saharan Africa.

2. Study area

Our study area is a latitudinal transect of Africa, from +22° to −22° latitude and from 23° to 24° longitude. This transect includes the typical Sub-Saharan variation of vegetation cover along a climatic gradient. The annual movement of the Intertropical Convergence Zone, a region of converging air masses and convective rain related to the movement of the sun's declination (Waliser & Gauthier, 1993), creates a banding of increasingly moisture-limited vegetation types as distance from the equator increases (White, 1983; Olson et al., 2001). From the equator south, banding is present but less well-defined compared to the region north of the equator due to the impacts of highland southern Africa and the Benguela Current on regional climate.

Within the transect, Congolian rainforest ecoregions dominate at low latitudes. A mosaic of rainforest and savanna overlaps the transition from rainforest to woodlands both north and south of the forest massif; this transition includes significant extents of derived secondary savanna, largely related to the population concentrations along this ecoregional transition (UNESCO, 1978). Beyond the transition zone in the north, tree cover is found within the Eastern Sudanian savanna ecoregion and is characterized by undifferentiated and *Isoberlina*-dominated woodlands. The respective woodlands south of the Congolian rainforests are more varied, dominated by Zambezian woodlands characterized by *Brachystegia*, *Julbernardia*, and *Isoberlina* genera tree species. Zambezian *Cryptosepalum* dry forests on well-drained Kalahari sands are found 10–15° south and constitute the only evergreen forests outside of the humid tropics within the study area. The southern end of the transect (>15° south) consists of a mix of the treeless Barotse Plain and southern Miombo, Baikiaea, and Mopane woodland ecoregions (Edmonds, 1976; Jeanes, 1991; Olson et al., 2001). For this study, tree heights and associated distributions are quantified within this 1° by 44° transect.

3. Data

3.1. Landsat 7 ETM + and Landsat 8 OLI data

Landsat images are available from the USGS Earth Resources Observation and Science Center (http://landsat.usgs.gov/). We queried and processed all available Landsat 7 and 8 data over our study area between 2013 and 2014 inclusive. We did not use the shorter wavelength blue and green bands due to persistent atmospheric contamination and scattering effects (Vermote, El Saleous, & Justice, 2002). Longer wavelength red (ETM + 0.630–0.690 μm and OLI 0.630–0.680 μm), near infrared (ETM + 0.775–0.900 μm and OLI 0.845–0.885 μm) and shortwave infrared (ETM + 1.550–1.750 μm and OLI 1.560–1.660 μm) (ETM + 2.090–2.350 μm and OLI 2.100–2.300 μm) bands were used as inputs. As part of our quality assessment (QA) system, we employed the blue (ETM + 0.45–0.52 μm and OLI 0.45–0.51 μm), green (ETM + 0.525–0.605 μm and OLI 0.525–0.600 μm) and thermal (ETM + 10.40–12.50 μm and Landsat 8's Thermal Infrared Sensor 10.60–11.19 μm) bands in screening viable land observations.

3.2. GLAS height data

The Geoscience Laser Altimeter System (GLAS) on the Ice, Cloud, and Land Elevation Satellite (ICESat) was launched on 12 January 2003. We acquired GLAS data from the National Snow and Ice Data Center (NSIDC, http://nsidc.org/data/icesat), Release 28, focusing on the L1A Global Altimetry Data (GLA01) and the L2 Global Land Surface Altimetry Data (GLA14). The noise and raw waveform data were obtained from GLA01 and the elevation, signal beginning, signal end, and the fitted Gaussian waveforms were obtained from GLA14. These parameters were used to screen the GLAS data sets and derive a series of metrics that were then used to characterize vertical vegetation structure.

The quality of each GLAS shot varies depending on several factors, including cloud cover, atmospheric conditions and the ground surface slope and roughness (Lefsky et al., 2002, 2005). A two-step procedure was used to identify and remove GLAS shots deemed not suitable for use as calibration data. The first was a technical analysis to identify GLAS shots that did not produce adequate waveforms for characterizing vegetation. A second screening was applied to remove outliers based on comparisons with existing tree cover and height maps.

The signal-based screening employed a set of seven criteria in screening GLAS data. Our first criterion eliminated shots lacking

Table 1
Stratum-based GLAS height training data totals, including the percent of training pixels retained after application of heuristics based on comparison to existing crown cover and height maps. NA indicates that no GLAS shots were found for this value and tree cover was used to reassign GLAS shots to 0 m.

	0 m	2–5 m	6–10 m	11–15 m	16–20 m	21–25 m	26–30 m	>30 m
Training pixels	23,844	33,326	59,844	47,518	22,463	11,623	7536	1345
Percent retained	NA	61%	82%	88%	90%	86%	77%	22%

coordinates. The second consisted of shots that contained less than two distinct peaks in the waveform, with the intention of removing shots lacking a peak associated with a vegetation canopy (as opposed to a peak associated with a ground return). A third criterion eliminated all shots where the amplitude of the raw waveform did not exceed twice the noise threshold, removing shots with atmospheric attenuation, including clouds. A fourth criterion removed those shots where the difference between the GLAS measured ground elevation and the SRTM elevation provided in the GLAS product, as distributed by the NSIDC, was greater than 25 m. The 25 m threshold represented the 95th percentile of elevation differences in a random worldwide sample. This comparison eliminated shots in which the ground surface was not adequately identified in the GLAS return. The fifth criterion removed shots where the signal beginning was over 86 m (representing the 98th percentile) or the signal end was greater than 20 m (representing the 95th percentile) or less than 1 m. This criterion was intended to exclude shots with interference before reaching any possible vegetation cover and shots where the signal end was improperly captured. The sixth criterion removed shots where the leading edge, defined here as the difference between the signal beginning and the H90 (90% energy remaining), was greater than 45% of the height (signal beginning − signal end). This threshold represented the 90th percentile and occurred in areas where there were great differences in vegetation canopy height, such as a forest next to a cleared area. The seventh and last criterion was the length of the trailing edge, defined here as the difference between signal end and H10 (10% energy remaining); differences greater than 35% of the height (signal beginning − signal end) were removed. This threshold represented the 90th percentile and occurred in areas of high slope, which often results in a longer trailing edge. The impacts of slope on GLAS screening and tree height characterization were not included in this study as the selected transect is absent of areas of high relief; Africa in general does not have extensive mountain ranges, unlike other continents. Any extension of the demonstrated method to larger areas will include a study of the impact of GLAS data screening on calibration data and resulting Landsat-driven characterizations.

The GLAS waveform structure is generally defined by the height distribution of the vegetation canopy and understory; our metric, height, was defined as the difference between signal beginning and signal

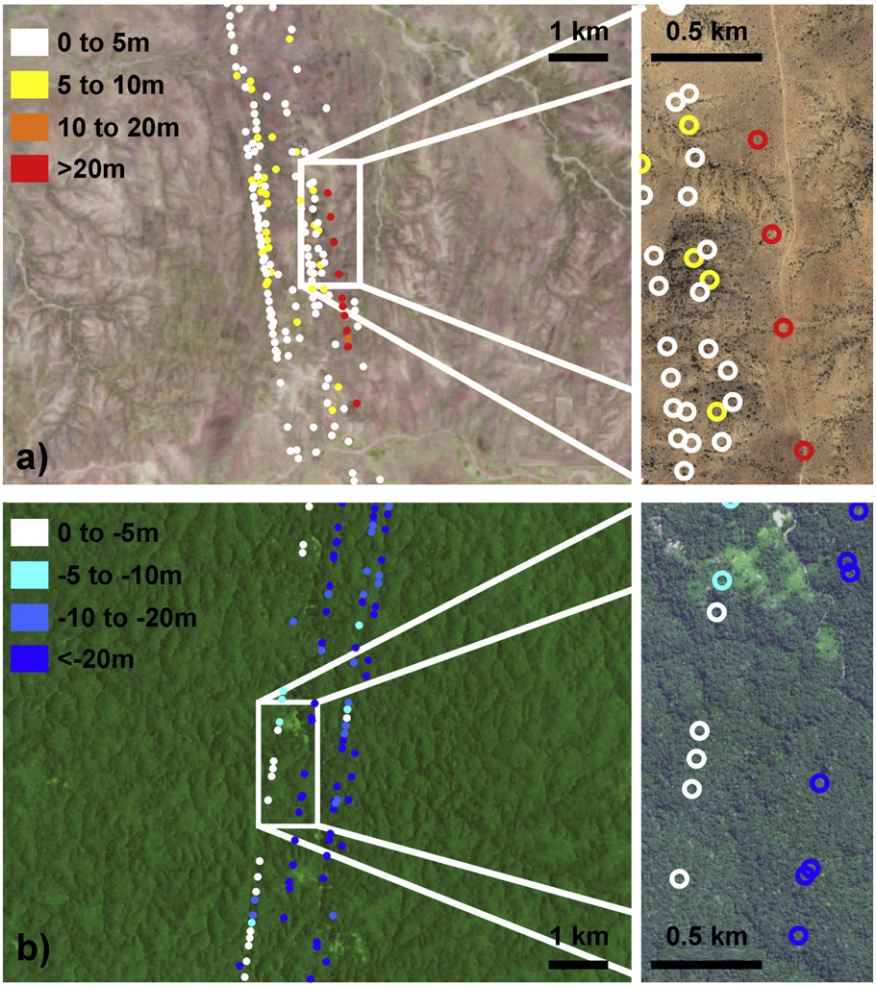

Fig. 2. GLAS training data screening for extreme outliers defined as abs(GLAS − Landsat 7 decadal tree height) ≥10 m for a) Gharb Darfur state, Sudan centered at 23.88E, 13.04N, and b) Kasai Oriental province, Democratic Republic of the Congo centered at 23.70E, 4.37S.

end. GLAS heights calculated following the initial screening process were subsequently compared to existing tree cover and tree height maps of Africa to remove remaining outliers. Two threshold-based outlier removal rules were applied to the input height data. Within the study area, all GLAS shots had values ≥2 m, even for remote areas in the central Sahara desert. To adjust these data, we used the percent tree cover layer from Hansen et al. (2013) and relabeled GLAS shots co-located with 0% tree crown cover to a height of zero meters. A second rule for filtering suspect GLAS shots employed a twelve-year Landsat 7-estimated mean height per pixel from the data of Tyukavina et al. (2015). Differences between GLAS and Landsat heights that were greater than 10 m were removed from the GLAS calibration reference data. Table 1 shows the percentage of GLAS shots that having passed the initial screening, were removed after application of these two heuristics. The rules disproportionately impact the amount of GLAS shots with extreme low and high height values.

Fig. 2 illustrates GLAS data removal due to the ≥10 m difference threshold for an area of low, short tree cover and an area of dense, tall tree cover. Suspect data appear to be sensor and overpass-related. For Fig. 2a in Sudan, GLAS shots with height estimates of over 20 m are located along an overpass that images a treeless divide. For Fig. 2b, GLAS shots 20 m lower in height than the Tyukavina estimate are found along two overpasses in an area of dense humid tropical forest cover. Both examples show adjacent overpasses that contain viable shots, indicating that the heuristics remove outliers without impacting the variability of sampled tree cover. While it is clear from Table 1 that the resulting extreme low and high end tree height calibration may be of limited value, the dominant range of tree heights are represented by valid GLAS height data; over 80% of GLAS data were retained for the 5 m to 25 m height range. A total of 207,499 GLAS shots passed all screening criteria.

4. Methods

4.1. Radiometric normalization

Our method for normalization consisted of processing each Landsat image separately per the methods outlined in Potapov et al. (2012). The processing steps included the following: 1) a per pixel quality assessment for cloud cover, haze, shadows and water, 2) a per band bulk bias radiometric adjustment, and 3) a per band radiometric adjustment for bi-directional reflectance distribution function effects as a function of sensor view zenith, solar zenith and solar azimuth angles. For step 1, we employed a series of classification tree models applied to top of atmosphere corrected Landsat data, resulting in per pixel quality flags. Each model was applied per Landsat image, yielding cloud, haze, shadow and water class probability values. Based on these values, a quality assessment (QA) code was assigned to each pixel reflecting the probability of the pixel to be a land or water cloud-free observation. For steps 2 and 3, we employed a MODIS reference image to perform a relative normalization. Annual 16-day MODIS data from the collection 5 MOD44C product from 2000 through 2011 (Carroll et al., 2010) were ranked by NDVI and associated Landsat-like red, near infrared and shortwave infrared band values averaged for the 50th to 90th percentile NDVI composite ranked images. These data were used as a global reference to normalize the Landsat imagery. The first step consisted of a bias adjustment which largely accounted for residual atmospheric effects such as water vapor content. The second step was a cross-track

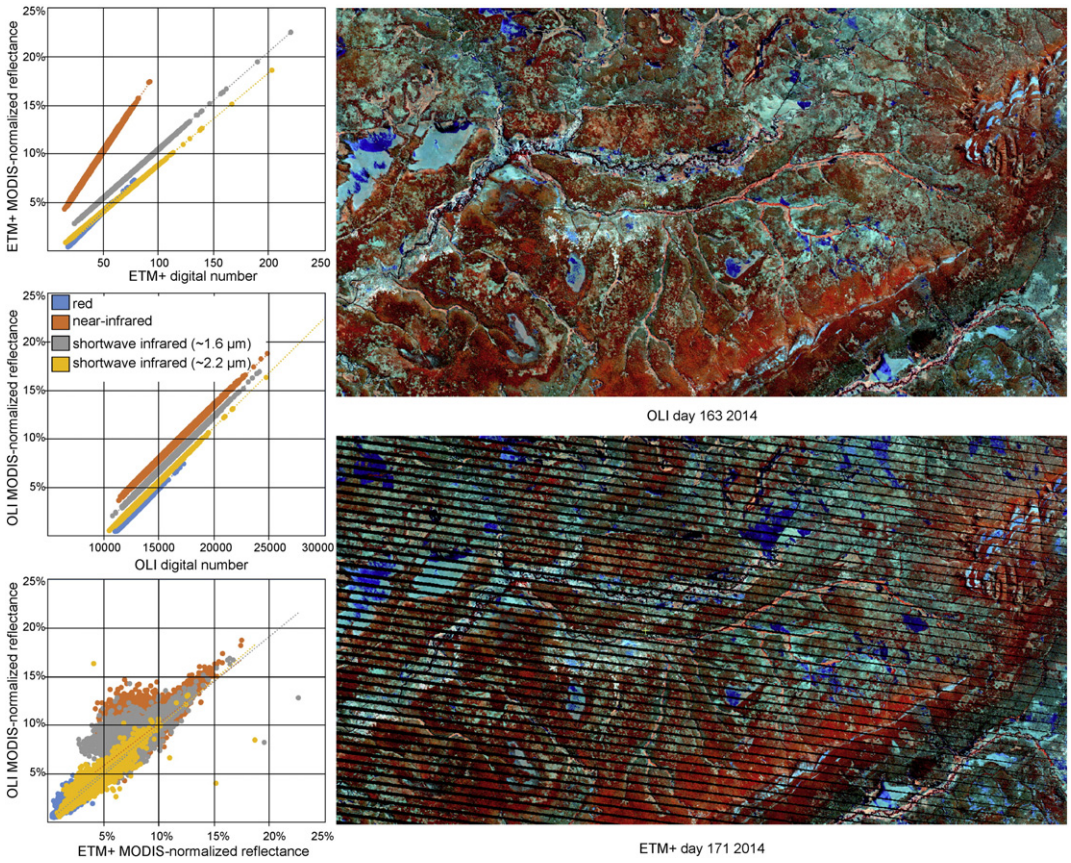

Fig. 3. On left, plots of pre-normalized digital numbers and post-normalized reflectances for two example Landsat images for path/row 175/068 acquired on June 12, 2014 and June 20, 2014 for Landsat 8 and 7, respectively. Image subsets on right illustrate consistent radiometric normalization between the two sensors. R–g–b is near infrared, shortwave infrared (~1.6 µm) and shortwave infrared (~2.2 µm). The subset is approximately 50 km by 30 km and centered at 23°36′E, 11°32′S in eastern Angola, near the borders of Zambia and the Democratic Republic of the Congo.

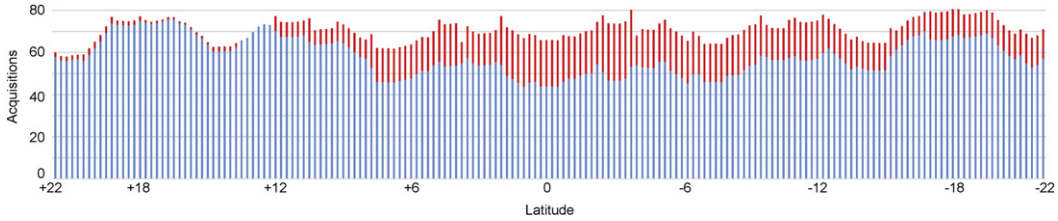

Fig. 4. Combined Landsat 7 and 8 acquisition totals per quarter degree of latitude from +22° north to −22° south and 23 to 24° east. Blue represents scenes processed, red not processed.

adjustment that used sensor view zenith angle as a dummy variable to adjust for effects of surface anisotropy. Hansen, Roy, Lindquist, Justice, and Altstaat (2008) showed that BRDF effects over Congo Basin humid tropical forests were more pronounced during the equinoxes when the solar azimuth was in line with cross-track sensor view angle.

The bias adjustment is a bulk, whole-scene correction; the BRDF adjustment operates within scene and employs MODIS and Landsat quality-vetted land pixels that are within 5% reflectance of each other to build a cross-track linear model. The BRDF correction is largely driven by land cover types that dominate within the scene footprint and is more pronounced at low, tropical latitudes and at dates near the equinoxes. Fig. 3 illustrates the before and after radiometric response for a Landsat 7 scene and a Landsat 8 scene acquired 8 days later for path/row 175/068, from digital numbers to MODIS-normalized reflectance (plots at left). The timing of these image acquisitions is during the transition from the rainy to dry season, with the onset of burning and little or no spectral change due to phenology present. Deviations in spectral response when comparing the normalized data (Fig. 3 plot at lower left) reflect land change between the two acquisition dates, in this case savanna fires. Other land areas within the respective scenes have a more consistent inter-date spectral response, indicating the successful radiometric normalization of the individual scenes.

4.2. Observation richness

Landsat acquisition volumes have improved greatly over time beginning with the Landsat 7 global acquisition strategy (Arvidson, Goward, Gasch, & Williams, 2006). Landsat 8 acquisition volumes are unprecedented and are now coupled with an enhanced Landsat 7 acquisition strategy over large land masses (Loveland pers. comm.). As such, more of the globe approaches the complete overpass coverage previously only employed over the United States; an 8-day repeat coverage is a realistic possibility for the entire globe. For this study, we included all Landsat 7 and 8 scenes for the 2013 and 2014 calendar years. A total of 5371 scenes were fully quality assessed, normalized and used as input for metrics generation. A total of 904 images were acquired but not processed to full normalized reflectance due to a variety of factors, including L1T processing failures, high GCP root mean square error, the TIRS instrument failure onboard Landsat 8 in late 2014, and other reasons. Since images for Landsat 8 were first collected in April of 2013, the first quarter of the 2013 inputs consisted solely of Landsat 7 data. Fig. 4 illustrates the number of total observations from processed and not processed scenes per quarter degree of latitude over the study area. The average total scene acquisition count across all latitudes was just over 70, which represents over three-quarters of the maximum possible given a nominal 8-day revisit for combined Landsat 7 and 8.

Moving from the per scene level to the per pixel level, a considerable reduction in data richness was observed. Fig. 5 illustrates the mean number of good quality pixels for Landsat 7 and Landsat 8 per quarter degree of latitude. While most scenes were processed, the number of pixels discarded due to poor quality assessment was high, particularly in lower latitude humid tropical Africa. Near the equator, the total number of viable land observations averaged less than 10 over the two year period for Landsat 7. Highest totals were found over the Sahara where clear sky conditions predominated.

4.3. Metrics generation

All quality-vetted and normalized observations were placed in a data pool and used to derive multi-temporal metrics. Metrics are designed to be a generic feature space that captures phenological variation in order to facilitate regional-scale mapping. Creating metrics using Landsat inputs is more complicated than with MODIS due to the increased variation in observation counts per pixel. Landsat data are not acquired systematically like the near-daily coverage of coarse spatial resolution sensors such as MODIS. In addition to acquisition frequency, cloud cover impacts data coverage differentially across space for all optical sensors, further impacting the derivation of a consistent feature space with which to characterize land cover. To ensure sufficient quality data volumes, we employed data for 2013 and 2014 from Landsat 7 and 8. Fig. 5 shows the average number of quality land observations across the transect for the separate and combined Landsat 7 and 8 records. Analysis of phenology by definition requires sufficient time-series inputs to capture the variation of vegetation growth and decline over time. All metrics were calculated using per band ranks of all good land observations from 2013 and 2014. Red, near-infrared and both shortwave infrared bands were sorted from minimum to maximum reflectance values. Only ranks between the 10th and 90th percentiles were used as inputs. Means for the inter-percentile ranges of 10 to 90, 10 to 25, 25 to 75, and 75 to 90 were calculated for all bands and the Normalized Difference Vegetation Index (NDVI) per pixel and for 3 × 3 moving window kernels. Separate and combined Landsat 7 and 8 metrics were created for testing in the characterization of tree height distribution.

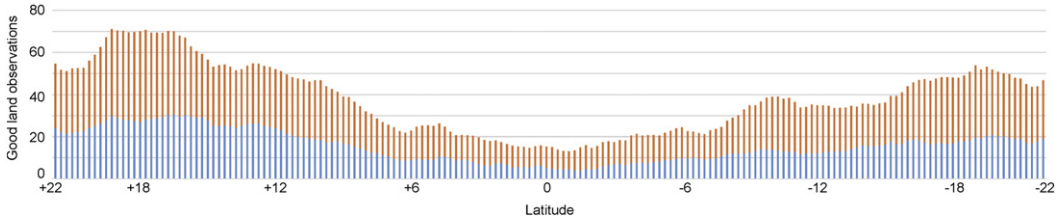

Fig. 5. Mean good land observation counts by quarter degree for 2013 and 2014 for Landsat 7 (blue) and Landsat 8 (orange) for the study area.

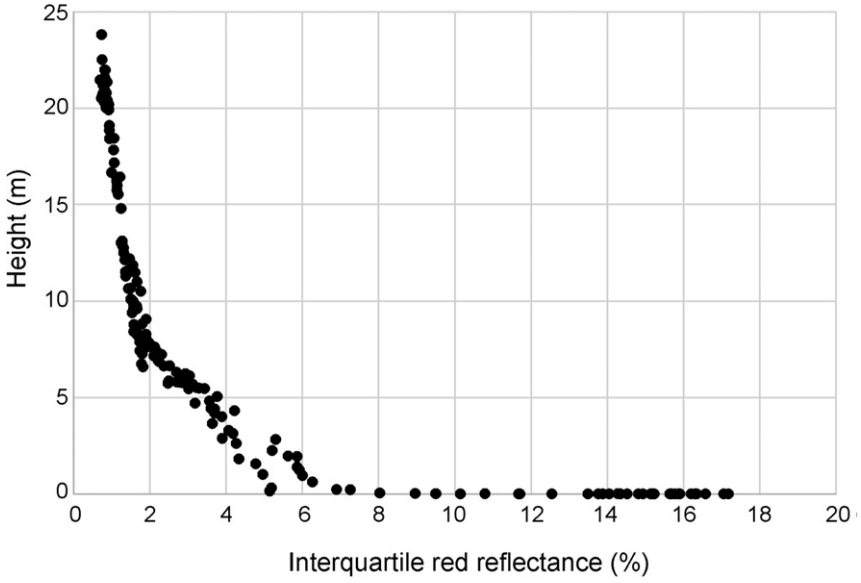

Fig. 6. Per quarter degree interquartile red reflectance for 2013 and 2014 Landsat 7 and 8 data versus estimated height of tree cover in meters.

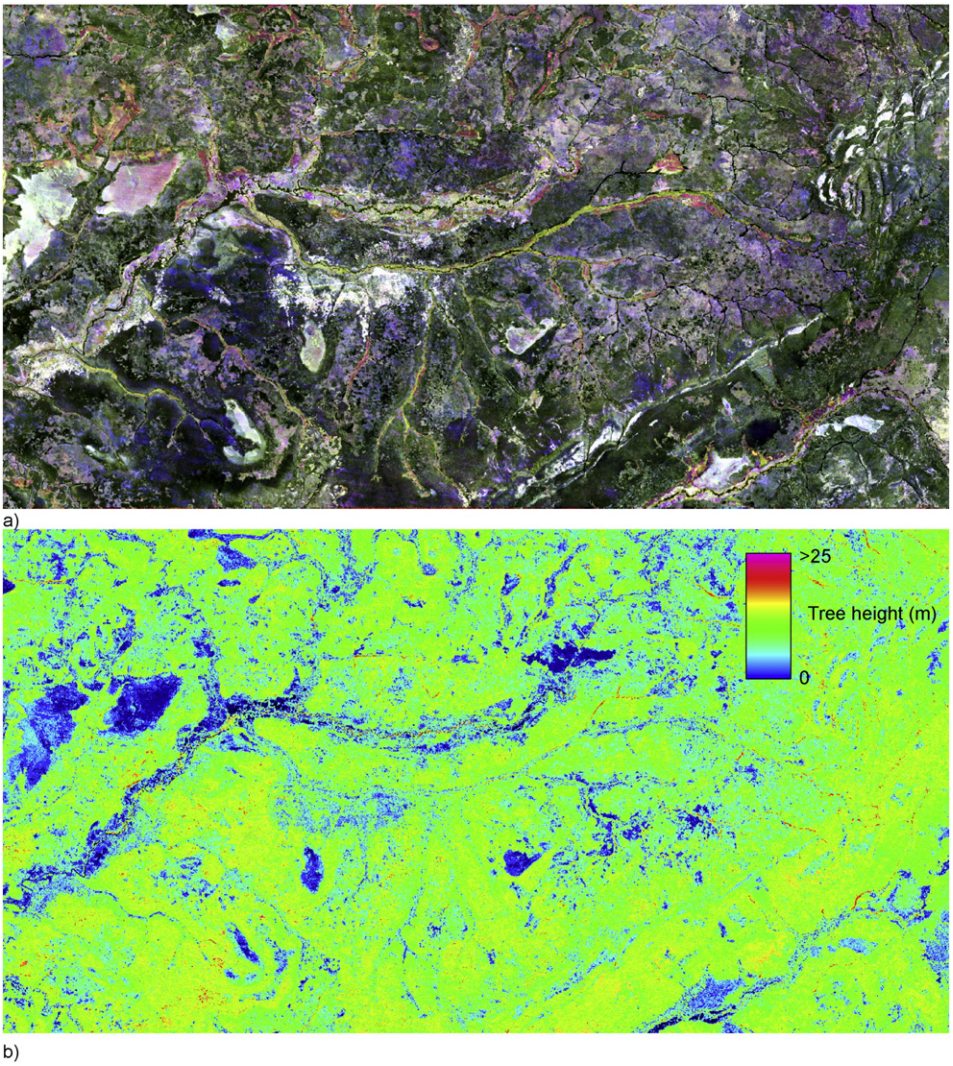

Fig. 7. a) R–g–b of the three most important metrics for estimating tree height where red is mean red reflectance of the 25th to 75th percentile ranked observations, green is mean red reflectance of the 10th to 25th percentile ranked observations, and blue is mean shortwave infrared reflectance of the 75th to 90th percentile ranked observations; b) estimated tree height.

Table 2
Mean absolute error (left) and mean error (right) from 207,499 GLAS height calibration shots and Landsat-modeled estimates reported by GLAS-defined strata and overall.

GLAS height strata	Landsat 7 and 8		Landsat 8		Landsat 7	
0–5 m	1.61	1.41	1.67	1.48	1.62	1.42
>5–10 m	2.09	0.38	2.09	0.37	2.15	0.42
>10–15 m	2.70	−0.81	2.73	−0.85	2.73	−0.81
>15–20 m	3.65	−1.58	3.70	−1.67	3.77	−1.76
>20–25 m	4.34	−3.29	4.40	−3.35	4.50	−3.42
>25–30 m	6.57	−6.47	6.64	−6.55	6.89	−6.80
>30 m	13.24	−13.24	13.35	−13.35	13.66	−13.66
Overall	2.45	−0.17	2.49	−0.17	2.52	−0.20

4.4. Algorithms and product evaluation

GLAS height data were used as training data and related to the Landsat-derived metrics using a regression tree algorithm. Regression trees employ a sum of squares criterion, referred to as deviance, to split a dependent variable, in this case height, into successively less varying subsets. To avoid overfitting, we used a set of seven bagged tree models based on 10% samples from the training data set, and report the per pixel median of the seven models. Individual tree growth was terminated when additional splits decreased model deviance by less than 0.001 of the deviance of the total training set population. The ability of the feature space to accurately estimate height was determined in a training accuracy mode where the bagged tree outputs were compared to training data values in order to evaluate different Landsat inputs. Hansen et al. (2005, 2008) employed such a method for intercomparing the relative value of spectral inputs given a single training data set. While the resulting accuracy measures were not absolute - no independent probability-based height measures were available to assess accuracy for the study area — we were able to determine the ability of the respective Landsat inputs to recreate the GLAS height calibration data, including height distributions across the study area.

We also applied an unsupervised clustering algorithm to our result in order to assess the distribution of estimated Landsat-derived heights. Unsupervised clustering is a standard method for identifying natural clusters in an *n*-dimensional feature space. We employed a *k*-means clustering algorithm (Tou & Gonzalez, 1974) to iteratively determine sampled means for a set of clusters, in this case tree height. We drew ten 0.1% samples of the height results and for each sample ran *k*-means for 2 to 20 clusters, resulting in 190 total runs. Cluster centers were weighted by their respective populations and totaled for each integer height over all runs. Totals were ranked by height of cluster center to find local minima, excluding cluster values from the tails of the population (<5 m and >25 m).

5. Results

5.1. Spectral variable importance

Red reflectance was the most important variable in estimating tree height, accounting for 58%, 65% and 63% of total reduced deviance for the Landsat 7, 8 and 7 and 8 combined models, respectively. The interquartile red reflectance metric accounted for 34%, 40% and 48% of the respective models. Fig. 6 illustrates the relationship between this metric and estimated tree height along the study area transect. Results conformed to a previous global study that employed AVHRR data and classification trees to map tree cover, finding that a 5% red reflectance threshold at the greenness time of year achieved greater than 90% accuracy in mapping tree-dominated landscapes (Hansen, Dubayah, & DeFries, 1996). The second most important contributor to the height model was mean red reflectance of the 10th to 25th percentile ranked observations and the third the mean swir reflectance of band 7 ETM+/band 8 OLI of the 75th to 90th percentile ranked observations. These three metrics accounted for 72% of the reduced deviance for the combined Landsat 7 and 8 model. Fig. 7a shows these metrics in r–g–b for the same area shown in Fig. 3. Estimated tree heights from the combined Landsat 7 and 8 metrics are shown in Fig. 7b. This area is within the Zambezian *Cryptosepalum* dry forest ecoregion with tree heights in the 15 m range, taller on average than other dry forests within the study area.

5.2. Training versus estimated height

Mean absolute model errors were calculated for each input scenario and results show a near-functional equivalency for this study for each input data set (Table 2). While combined Landsat 7 and 8 inputs performed the best, there was not a dramatic difference between the three model outputs. Errors were highest in the >20 m strata where heights were consistently underestimated. These results indicate a likely lack of sensitivity of Landsat in estimating tall canopies. For the >30 m stratum, the mean absolute and mean errors were equivalent, meaning a universal underestimation. Overestimates in the non-treed stratum (≤5) were likely due to the presence of poor quality GLAS retrievals of low cover conditions. Table 3 shows per stratum population percentages for the height training data set and the estimated heights for the transect as a whole. The proportions differ and reveal a number of training and spectral data limitations. First, the result contained a larger amount of pixels with zero height values, largely in the northern part of the transect within the northern Sahel and southern Sahara desert, areas with little training data. There was also a greater overall proportion of tall (>20 m) heights in the map output compared to the training data, as atmospheric attenuation likely impacted training data volumes in humid forests. However, mapped tall forests were largely limited to the 20–25 m range. A lack of signal to differentiate taller canopies is a likely reason, as shown in Fig. 6. The result was Landsat characterizing tall tree heights more as a class than a continuum. GLAS-training versus Landsat 7 and 8-estimated height data are plotted in Fig. 8 and illustrate the variation in training data density and areas of correspondence and disagreement.

By-latitude plots of the GLAS calibration and Landsat-estimated heights are shown in Fig. 9. These plots depict the mode of the GLAS training data per quarter degree of latitude in Fig. 9a, and the mode of the Landsat-estimated results per quarter degree of latitude in Fig. 9b. The GLAS inputs are the same 207,499 shots from Fig. 8, but distributed graphically across the study's north–south transect. The Landsat plot is calculated from the entire population of the map output. In this comparison, the underestimation by the Landsat product for tall forests is illustrated, as is a more discretized output. This effect is the one investigated by Hanan et al. where the authors posit such discontinuities in African tree cover to be solely a function of the regression tree algorithm. Figs. 8 and 9 show the accurate reproduction of the LiDAR-derived tree heights used in calibration with no evidence of introduced or artificial

Table 3
Percent distribution of height training data and mapped result.

	0 m	2–5 m	6–10 m	11–15 m	16–20 m	21–25 m	26–30 m	>30 m
Training	11.5	16.1	28.8	22.9	10.8	5.6	3.6	0.6
Result	27.6	15.9	22.3	12.0	7.8	11.8	2.7	0.002

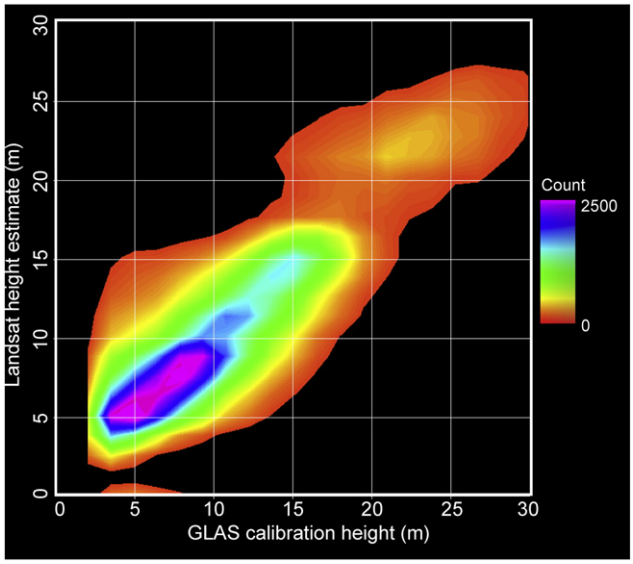

Fig. 8. Density plot of GLAS training height data (207,499 shots) versus Landsat-derived height. For these data, a simple linear regression model yielded an r^2 of 0.46 and a model of Landsat height = GLAS height * 1.07 − 0.28.

discontinuities. The height of 19 m tall tree cover is not present in these plots. Tree heights from 11 to 17 m are found largely in the area from −10 to −15° latitude south. In this latitudinal range we find the farthest southern extent of evergreen tropical forest within the Zambezian *Cryptosepalum* dry forest ecoregion (Edmonds, 1976; Jeanes, 1991; Olson et al., 2001). This ecoregion is dominated by native *Cryptosepalum exfoliatum pseudotaxus* tree cover with well-drained Kalahari sands the primary determinant of their extent. These forests represent the largest area of tropical evergreen forest in Africa outside the equatorial zone.

Apart from the *Cryptosepalum* dry forests, the two main distributions of tall (≥20 m) and medium tree cover (5–10 m) are related to evergreen humid tropical forest and dry seasonal woodlands, respectively. The Congolian rainforest ecoregions constitute the tall tree cover depicted largely within ±6° latitude. The low tree cover bands along the edge of the forest massif are part of the forest/savanna mosaic associated with human settlements and associated land use activities. The main woodland zones north of the humid tropical forests extend until approximately 10° north. The southern woodlands are more extensive and varied, with the *Cryptosepalum* woodlands clearly centered at 12° latitude south. Miombo and other woodland ecosystems constitute the remaining and more typical southern African dry tropical tree cover. Fig. 10 places these results in geographic context.

5.3. Error by observation count

Increasing observation richness, or the number of quality-passed observations per pixel, and tree height mapping accuracy were correlated, as shown in Fig. 11. The general trend in each of the Landsat 7, 8 and 7 and 8 combined examples follows the same sequence. Using combined Landsat 7 and 8 as an example, we will review this relationship. It is helpful to understand that observation richness is related largely to bioclimatic variation. As we move from left to right across the graph, different biomes are represented due to the relative impacts of cloud cover on the total number of usable observations per pixel. First, relatively high mean absolute errors were found in low observation frequency sites, mainly within the humid tropical forest domain. Mean absolute error of humid tropical forest height decreased as total observation counts increase. The humid tropical forest biome was capped at roughly 25 good observations. Good observations increased outside of the humid forest zone due to greater seasonality and a higher likelihood of acquiring a cloud-free observation in seasonally dry tropical environments. The sequence of approximately 25 to 40 observations was mainly within the woodland savannas adjacent to humid tropical rainforests. These woodlands are more extensive and varied in the southern portion of the transect. Areas with greater than 40 observations were found in semi-desert and desert areas. In the Landsat 8 and 7 and 8 combined models, an increase in error within this zone was due to the aforementioned inconsistent GLAS retrievals in low cover conditions. Landsat 7- and 8-only plots reveal less observational richness and a resulting higher overall error (large observation counts having higher mean absolute error).

5.4. Unsupervised clustering of result

The following height values ranked lowest in terms of weighted cluster totals: 11, 10, 17 and 18. Results indicated three tree height clusters: a dry seasonal woodland cluster of 5–10 m, an intermediate height cluster consisting primarily of dry evergreen forests and regrowing humid tropical forests of 11–17 m, and a third cluster of humid evergreen tropical forest of ≥18 m. The spectral responses for the tree height clusters are shown in Fig. 12. Fig. 12a shows density plots for the entire study area for two interquartile metrics and Fig. 12b a density of the respective low, medium and high/tall tree height clusters for the same two metrics. From Fig. 12b, it is clear that trees exists in a very limited

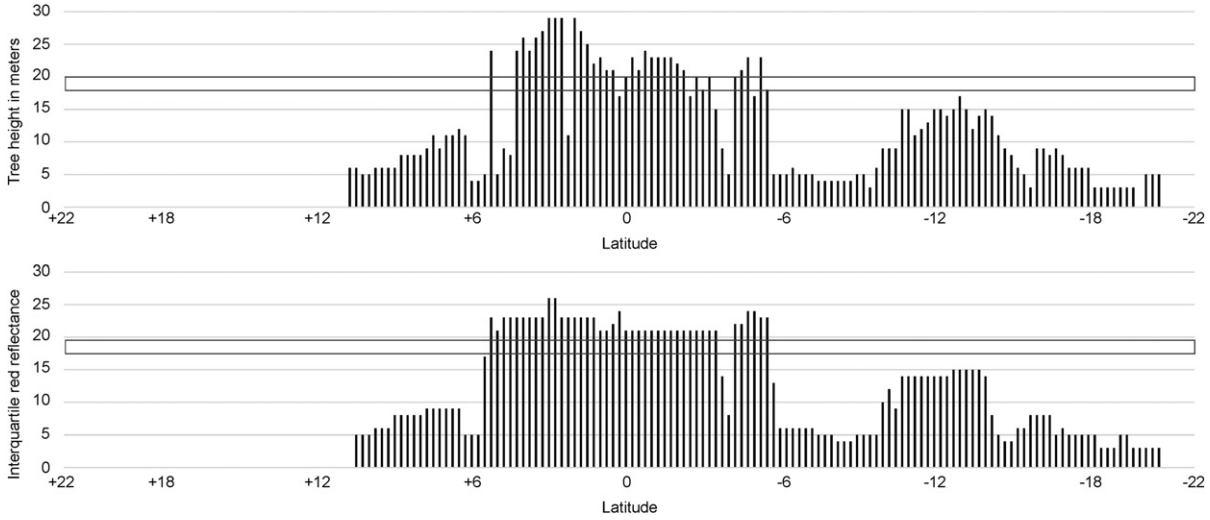

Fig. 9. Per quarter degree of latitude a) mode of GLAS height calibration data and b) mode of estimated Landsat-derived height data. The rectangle is centered on 19 m where no mode tree cover values are found.

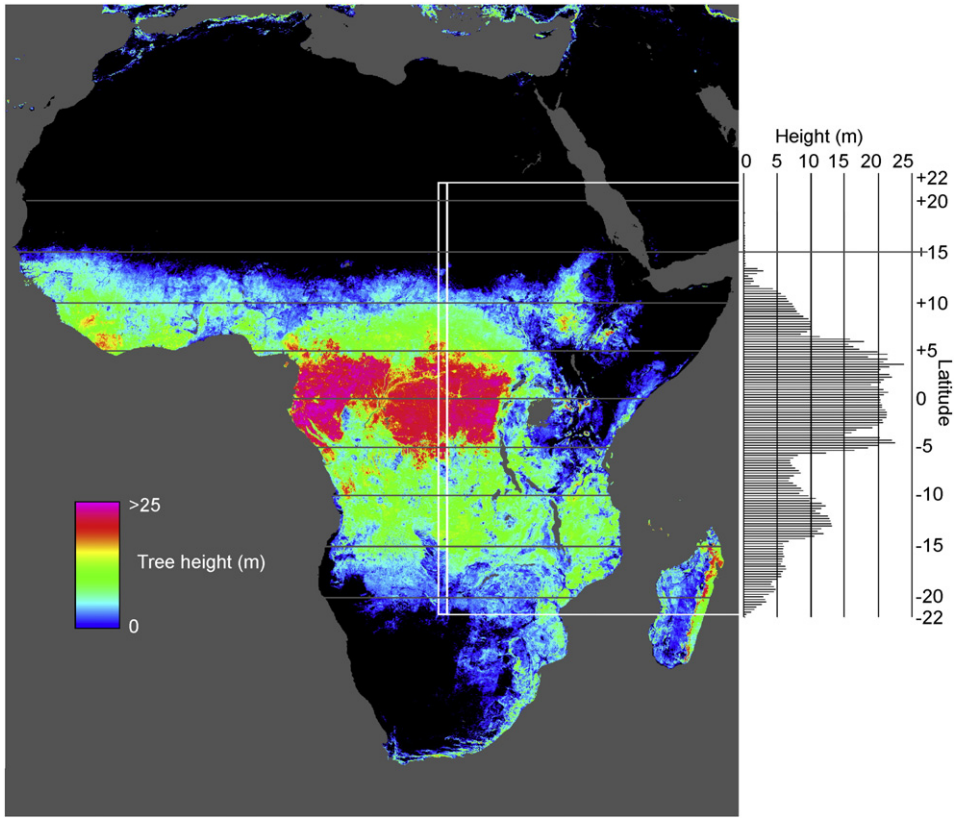

Fig. 10. Estimated tree height for Africa, from Tyukavina et al. (2015) with study area transect outlined in white and results of this study plotted per quarter degree latitude for mean height derived using combined Landsat 7 and 8-derived metrics.

spectral space, and that dominant tree populations are discreetly identifiable. The dominant clusters of tree heights are related to 1) short height dry woodlands (blue in Fig. 12b), 2) medium height dry evergreen forests (green in Fig. 12b) and 3) tall humid forests (red in Fig. 12b). Some medium height tree cover has a spectral affinity to tall tree cover (green adjacent to red, but brighter in red and nir in Fig. 12b). These pixels are located mainly in the Congo Basin in areas of secondary regrowth within the rural complex (Mayaux, Richards, & Janodet, 1999) and in natural humid tropical forests having more open canopies.

6. Discussion and conclusions

The GLAS-calibrated Landsat-derived height estimates for a transect across tropical Sub-Saharan Africa revealed three dominant tree height populations, a 5–10 m dry seasonal woodland cluster, a 11–17 m dry evergreen forest cluster, and a ≥18 m humid evergreen forest cluster. Tree heights from 17 to 18 m were relatively infrequent and largely defined the separation of dry woodlands and forests from humid forests; tree heights in this range consisted of regrowing and open canopied humid forests. Overall, the regression tree models exhibited high fidelity in reproducing the LiDAR-derived height calibration data and resulting Sub-Saharan African tree height distributions (Figs. 8 and 9).

Biases in tree height estimations were found for low and high tree heights with respective over and under estimations. For short heights, the capability of Landsat was not clear due to the likely overestimation of low heights in the GLAS data. The lack of reliable calibration for short vegetation precluded us from stating anything about Landsat's ability to discern variations in short woody cover. It may also be that in low tree cover conditions, a proportionally larger background soil and understory signal confounded accurate height or cover estimation. For tall canopies, a diminishing signal was a likely limitation. Fig. 6 illustrates the sharply declining interquartile red reflectance with increasing tree height. Interquartile red reflectance for tree cover taller than 20 m was less than 1%, indicating a lack of signal in further discrimination of tall tree cover.

A relatively simple set of multi-temporal metrics based on inter-percentile means yielded an overall mean absolute model error of 2.45 m for a transect covering the full range of variation of Sub-Saharan African tree cover. Results were encouraging and confirmed the use of Landsat in estimating vegetation structure parameters. For example, Tyukavina et al. (2015) employed the same approach used

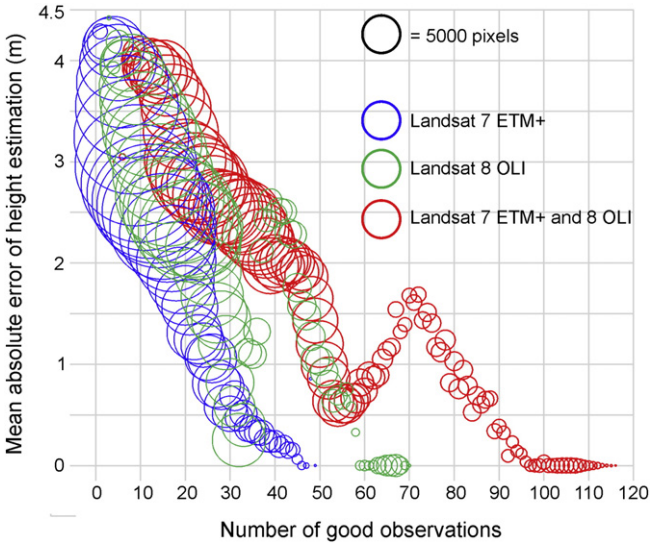

Fig. 11. Per pixel good quality-assessed observation totals per pixel versus mean absolute error for height estimates. Circle diameter is proportional to the number of pixels with the given number of observations.

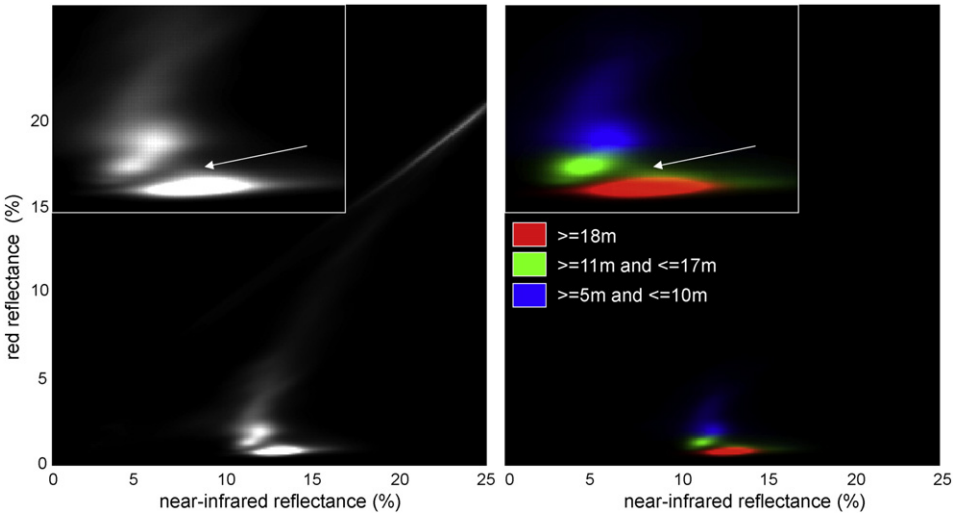

Fig. 12. Spectral plot of interquartile metrics (mean of 25–75% per band ranks) from all good observations from 2013 and 2014 for the study area from combined Landsat 7 and 8 data. At left is a density plot of the entire transect population, including treed and non-treed cover. At right is an r–g–b density plot for trees, defined as ≥5 m heights. Here, red is the high/tall cluster, green is the medium height cluster, and blue is the low height cluster. Upper left of each panel is a zoom of the tree-dominated portion of the spectral plot. Arrows highlight a discontinuity in the input spectral data that corresponds to a gap in the Landsat-based thematic output as determined by GLAS training data.

here to create a Landsat 7-derived pan-tropical height map. This height map was subsequently used to generate carbon stock strata in estimating emissions from tropical forest disturbance. The use of dense time-series optical data to extend the value of LiDAR forest structure measurements is a promising research frontier.

Results illustrate the value of increased observation frequency in reducing the uncertainty of thematic map outputs. Improved accuracy was found to be a function of observational richness as mean absolute error decreased over the two major tree cover distributions with increasing volumes of good quality-assessed data. The differences in accuracy between Landsat 7, 8 and 7 and 8 combined were real, but very small — on the order of hundredths of a meter for overall mean absolute error. In the case of African tree height mapping, Landsat 8 outperformed Landsat 7, reflecting the value of the OLI instrument in advancing land characterization. However, increasing the frequency of observations through the integrated use of both sensors proved superior to individual instrument capabilities. For land themes other than height, the requirements for data richness and metric feature space will of course differ from what was found in this study. In all cases, Landsat 7 and 8 should be routinely integrated in support of land monitoring applications as the increase in observation richness will likely improve characterization accuracies. Given the unprecedented Landsat data volumes being collected for Landsat 7 and 8 and the expected addition of Sentinel 2, more and varied large area vegetation mapping efforts that exploit improved image cadence should be attempted and formally evaluated.

Acknowledgments

Support for this study was provided by the U.S. Agency for International Development's Central Africa Regional Program for the Environment (CARPE) through NASA grants NNX14AR46G and NNX12AI75G, NASA's Terrestrial Ecology program through grant number NN12AB4G, and the Gordon and Betty Moore Foundation through grant number 3125.

References

Arvidson, T., Goward, S. N., Gasch, J., & Williams, D. (2006). Landsat-7 long-term acquisition plan: Development and validation. *Photogrammetric Engineering and Remote Sensing, 72*, 1137–1146.

Asner, et al. (2012). High-resolution mapping of forest carbon stocks in the Colombian Amazon. *Biogeosciences, 9*, 2683–2696.

Carroll, M., Townshend, J. R. G., Hansen, M. C., DiMiceli, C., Sohlberg, R., & Wurster, K. (2010). Vegetative cover conversion and vegetation continuous fields. In B. Ramachandran, C. Justice, & M. Abrams (Eds.), *Land remote sensing and global environmental change: NASA's EOS and the Science of ASTER and MODIS*. New York: Springer.

Cihlar, J. (2000). Land cover mapping of large areas from satellites: Status and research. *International Journal of Remote Sensing, 21*, 1093–1114.

Cohen, W. B., & Spies, T. A. (1992). Estimating structural attributes of Douglas-Fir/Western Hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment, 41*, 1–17.

DeFries, R., Hansen, M., & Townshend, J. (1995). Global discrimination of land cover types from metrics derived from AVHRR Pathfinder data. *Remote Sensing of Environment, 54*, 209–222.

Drusch, M., et al. (2012). ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment, 120*, 25–36.

Edmonds, A. C. (1976). *Vegetation map of Zambia, 1:500 000*. Lusaka, Zambia: Forest Department, Government of Zambia, Survey Department.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., et al. (2002). Global land cover mapping from MODIS: Algorithms and early results. *Remote Sensing of Environment, 83*, 287–302.

Hanan, N. P., Tredennick, A. T., Prihodko, L., Bucini, G., & Dohn, J. (2014). Analysis of stable states in global savannas: Is the CART pulling the horse? *Global Ecology and Biogeography, 23*, 259–263.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., Carroll, M., Dimiceli, C., & Sohlberg, R. A. (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) (15 pp., online journal).

Hansen, M. C., Townshend, J. R. G., DeFries, R. S., & Carroll, M. (2005). Estimation of tree cover using MODIS data at global, continental and regional/local scales. *International Journal of Remote Sensing, 26*, 4359–4380.

Hansen, M., Dubayah, R., & DeFries, R. (1996). Classification trees: An alternative to traditional land cover classifiers. *International Journal of Remote Sensing, 17*, 1075–1081.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., & Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. *International Journal of Remote Sensing, 21*, 1331–1364.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., Sohlberg, R., Carroll, M., & Dimiceli, C. (2002). Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data. *Remote Sensing of Environment, 83*(1&2), 303–319.

Hansen, M. C., Roy, D., Lindquist, E., Justice, C. O., & Altstaat, A. (2008). A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin. *Remote Sensing of Environment, 112*, 2495–2513.

Hansen, M. C., et al. (2013). High-resolution global maps of 21st-century forest cover change. *Science, 342*, 850–853.

Hilker, T., Wulder, M. A., & Coops, N. C. (2008). Update of forest inventory data with Lidar and high spatial resolution satellite imagery. *Canadian Journal of Remote Sensing, 34*, 5–12.

Huang, C., Goward, S. N., Schleeweis, K., Thomas, N., Masek, J. G., & Zhu, Z. (2009). Dynamics of national forests assessed using the Landsat record: Case studies in eastern United States. *Remote Sensing of Environment, 113*, 1430–1442.

Hudak, A. T., Lefsky, M. A., Cohen, W. B., & Berterretche, M. (2002). Integration of lidar and Landsat ETM+ data for estimating and mapping forest canopy height. *Remote Sensing of Environment*, *82*, 397–416.

Jeanes, K. W. (1991). *Landscapes and grasslands of Western Province, Zambia.* Western Province: Dept. of Agriculture.

Justice, C. O., Townshend, J. R. G., Vermote, E. F., Masuoka, E., Wolfe, R. E., Saleous, N., ... Morisette, J. T. (2002). *Remote Sensing of Environment*, *83*, 3–15.

Kennedy, R. E., Cohen, W. B., & Schroeder, T. A. (2007). Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sensing of Environment*, *110*, 370–386.

Lefsky, M. (2010). A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophysical Research Letters*, *37*, L15401. http://dx.doi.org/10.1029/2010GL043622.

Lefsky, M. A., Cohen, W. B., Harding, D. J., Parker, G. G., Acker, S. A., & Gower, S. T. (2002). Lidar remote sensing of above-ground biomass in three biomes. *Global Ecology and Biogeography*, *11*, 393–399.

Lefsky, M. A., Harding, D. J., Keller, M., Cohen, W. B., Carabajal, C. C., Del Bom Espirito-Santo, F., ... de Oliveira, R., Jr., (2005). Estimates of forest canopy height and above-ground biomass using ICESat. *Geophysical Research Letters*, *32*. http://dx.doi.org/10.1029/2005GL023971.

Mayaux, P., Richards, T., & Janodet, E. (1999). A vegetation map of Central Africa derived from satellite imagery. *Journal of Biogeography*, *26*, 353–366.

Olson, D. M., et al. (2001). Terrestrial ecoregions of the world: A new map of life on Earth. *Bioscience*, *51*, 933–938.

Ota, T., et al. (2014). Estimation of airborne lidar-derived tropical forest canopy height using Landsat time series in Cambodia. *Remote Sensing*, *6*, 10750–10772.

Potapov, P. V., Turubanova, S. A., Hansen, M. C., Adusei, B., Broich, M., Altstatt, A., ... Justice, C. O. (2012). Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM+ data. *Remote Sensing of Environment*, *122*, 106–116.

Potapov, P. V., Turubanova, S. A., Tyukavina, A., Krylov, A., McCarty, J. J., Radeloff, V. C., & Hansen, M. C. (2015). Eastern Europe's forest cover dynamics from 1985 to 2012 quantified from the full Landsat archive. *Remote Sensing of Environment*, *159*, 28–43.

Sankaran, M., et al. (2005). Determinants of woody cover in African savannas. *Nature*, *438*, 846–849.

Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research*, *116*, G04021. http://dx.doi.org/10.1029/2011JG001708.

Staver, A.C., Archibald, S., & Levin, S. (2011). Tree cover in sub-Saharan Africa: Rainfall and fire constrain forest and savanna as alternative stable states. *Ecology*, *92*, 1063–1072.

Tou, J. T., & Gonzalez, R. C. (1974). *Pattern recognition principles.* Addison-Wesley Publishing Co (377 pp.).

Tyukavina, A., Baccini, A., Hansen, M. C., Potapov, P. V., Stehman, S. V., Houghton, R. A., ... Goetz, S. J. (2015). Aboveground carbon loss in natural and managed tropical forests from 2000 to 2012. *Environmental Research Letters*, *10*, 074002. http://dx.doi.org/10.1088/1748-9326/10/7/074002.

UNESCO (1978). Tropical forest ecosystems: A state-of-knowledge report. *Natural resources research*, *14*, UNESCO (683 pp.).

Vermote, E. F., El Saleous, N. Z., & Justice, C. O. (2002). Atmospheric 985 correction of MODIS data in the visible to middle infrared: First results. *Remote Sensing of Environment*, *83*, 97–111.

Waliser, D. E., & Gauthier, C. (1993). A satellite-derived climatology of the ITCZ. *Journal of Climate*, *6*, 2162–2174. http://dx.doi.org/10.1175/1520-0442(1993)006,2162:ASDCOT.2.0.CO;2.

White, F. 1983. The vegetation of Africa. A descriptive memoir to accompany the UNESCO/AETFAT/UNSO Vegetation Map of Africa (3 Plates, Northwestern Africa, Northeastern Africa, and Southern Africa, 1:5,000,000), UNESCO, Paris. ISBN: 9231019554.

Zhu, Z., & Woodcock, C. E. (2014). Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, *144*, 152–171.

Zwally, J., et al. (2002). ICESat's laser measurements of polar ice, atmosphere, ocean and land. *Journal of Geodynamics*, *34*, 405–445.