Remote Sensing of Environment 165 (2015) 203–215

Contents lists available at ScienceDirect

Remote Sensing of Environment

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

Forest cover change in Miombo Woodlands: modeling land cover of African dry tropical forests with linear spectral mixture analysis

Marc T. Mayes [a,b,*], John F. Mustard [a], Jerry M. Melillo [a,b]

[a] Brown University, Department of Earth Environmental and Planetary Sciences, 324 Brook Street Box 1846, Providence, RI 02912, USA
[b] The Ecosystems Center, Marine Biological Laboratory, 7 MBL Street, Woods Hole, MA 02543 USA

ARTICLE INFO

Article history:
Received 20 September 2014
Received in revised form 5 March 2015
Accepted 16 May 2015
Available online 28 May 2015

Keywords:
Dry tropical forests
Africa
Miombo Woodlands
Land cover change
Land cover variability
Spectral mixture analysis

ABSTRACT

Dry tropical forests are experiencing some of the highest rates of change among the globe's forests. In sub-Saharan Africa, gross (loss, gain) and net changes in dry tropical forest areas are difficult to quantify at sub-national scales because of high spatio-temporal variability in land cover conditions due to vegetation phenology and land use practices. In this project, we developed new, field-validated remote sensing characterizations of dry season surface components to separate forest from non-forest land cover, and assessed forest changes from the 1990s–2010s in a Tanzanian Miombo Woodland landscape. Using a linear spectral mixture analysis (LSMA) approach with Landsat 5–8 data, we examined the hypothesis that higher proportions of substrate and non-photosynthetic vegetation (NPV) at non-forest regions distinguished them from forest cover against seasonally variable land cover conditions. Subsequently we evaluated the efficacy of multi-temporal classification and single-date image thresholding for identifying forest from non-forest cover. We found significantly greater proportions of substrate and NPV over non-forest compared to forest areas that enabled identification of forest cover across dry season images. Single-date, forest/non-forest maps based on an LSMA-derived metric attained overall accuracies of 81.0–85.3%, which approached multi-temporal unsupervised classifications (86.5% for forest/non-forest maps). Applying the LSMA-derived metric to study forest changes, our study region experienced a net 15.0% loss of 1995 forest area, and a 7.0% overall reduction in the total forest-occupied land cover from 1995–2011. Areas of gross forest gain were substantial, totaling 13.6% of the 1995 forest area. We found differing patterns in gross forest losses and gains among sub-regions and through time in our Tabora study area, which provide bases for testable hypotheses in future research on regional and localized drivers affecting forest cover. Our finding that non-green surface components distinguished forest from non-forest via an LSMA approach may be widely applicable to studying forest conversions in Miombo Woodlands and other dry tropical forests. This approach may also be useful for evaluating how land cover conditions change in response to potential land use or climate driving variables, or the impact of land changes for carbon balance and other ecosystem processes.

1. Introduction

Forest losses, gains and agricultural land changes in seasonally dry tropical ecosystems are among the least well quantified and understood land cover changes globally (Lambin, Geist, & Lepers, 2003; Lepers et al., 2005). These ecosystems, including tropical deciduous forests and savannas, occupy about 14% of Earth's terrestrial area (18.6×10^6 km^2) and comprise 15% of global vegetation carbon stocks (137.5 Pg) at the third-highest biomass carbon (C) density (200 g C m^{-2} (Aber & Melillo, 2001)). They have among the largest reserves of the globe's arable land that is not yet cultivated (Lambin & Meyfroidt, 2011). Among nations with the highest deforestation rates, significant proportions of their forest losses come from dry tropical forests, which have few areas under legal protected status (FAO, 2010; Green et al., 2013). Satellite remote sensing assessments of forest conversions (deforestation, forest regrowth or afforestation) in the dry tropics have been limited by the high temporal and spatial variability in land cover conditions, which stem from complex regional vegetation ecology, land use practices and disturbance regimes. Currently, poor knowledge of land cover variability across land cover types, which includes characterization of land cover components such as green vegetation or substrate (i.e. soil) at landscape scales, hinders differentiation of forest and non-forest areas and evaluation of forest change in a "wall-to-wall" fashion (Bodart et al., 2013; Grainger, 2008).

This study develops a model to quantitatively characterize and differentiate dry tropical forest cover from non-forest areas using analyses of Landsat satellite data, and assesses forest conversions in the Miombo

* Corresponding author at: Brown University Department of Earth Environmental and Planetary Sciences, 324 Brook Street Box 1846, Providence, RI 02912, USA.
E-mail address: marc_mayes@brown.edu (M.T. Mayes).

http://dx.doi.org/10.1016/j.rse.2015.05.006
0034-4257/

Woodlands of western Tanzania. The Miombo Woodlands span 2.8 million km^2 across sub-Saharan Africa, and contain the largest contiguous dry tropical forests globally (Campbell, Frost, & N. B., 1996). Many Miombo regions are home to small land-holding agricultural communities with high population growth rates and increasing land change pressures. Since the 1990s, cultivation of resource-intensive cash crops such as tobacco, and natural resource demands of urbanization such as fuel wood are contributing to high rates of forest changes (Geist, Chang, Etges, & Abdallah, 2009; Kutsch et al., 2011; Yanda, 2010). Recent studies have shown that large net loss of Miombo woodland has been almost 40% in eastern Tanzania since 1975. This rate of loss has far exceeded that of Africa's humid tropical forests (5.1% of 1975 area) (Green et al., 2013). In western Tanzania, rates of woodland habitat loss to agriculture have been documented at 4.7% from 1984–1995 and 11.2% from 1995–2000 (Yanda, 2010). Regional and plot-scale studies have also observed forest regrowth (Chidumayo, 2013; Prins & Kikula, 1996; Yanda, 2010). Beyond documented net forest cover losses, little data exists on the patterns of gross forest loss and regrowth that are resulting in the net changes.

Past assessments of regional forest change in the Miombo have relied on supervised classification or image-interpretation approaches with remotely sensed data (Cabral, Vasconcelos, Oom, & Sardinha, 2011; Green et al., 2013; Yanda, 2010). While locally informative, these are difficult to repeat for regular, long-term monitoring over large areas. These local approaches have been necessary because remote sensing-derived land cover datasets produced at global scales often fail to distinguish forest cover or lack calibration in dry tropical ecosystems and particularly in the Miombo (Lepers et al., 2005; Sedano, Gong, & Ferrao, 2005).

Forest cover in the Miombo includes areas with tree cover as low as 30%, with many trees between 3 and 5 m tall (Chidumayo, 2013; Eamus & Prior, 2001; Frost, 1996). An accuracy assessment of the MODIS Global Land Cover Product (MOD12Q1) in a Miombo region of Mozambique found that nearly all forest and agriculture training sites were inaccurately described as Savanna or Woody Savanna (Sedano et al., 2005). The global land cover product did not distinguish tree-dominated from grass-dominated areas, limiting their usefulness for forest monitoring for purposes such as carbon accounting. More recent products, such as the global forest-change maps of Hansen et al. (2013), evaluate forest conversion only in areas with greater than 50% canopy cover and with trees greater than 5 m height (Hansen et al., 2013). Such definitions may exclude many chronically disturbed and re-growing Miombo forest regions (Chidumayo, 2013; Prins & Kikula, 1996). Physically-based, regionally-calibrated measures are needed to identify forest cover and conversions in the Miombo using satellite data.

Satellite-based, quantitative characterizations of forest cover and conversions in the Miombo face two challenges. First, inter-annual variability in precipitation and fire confounds the temporal variability of land cover components among forest and non-forest land cover (Serneels, Linderman, & Lambin, 2007). Precipitation drives vegetation production and 85% of annual rainfall falls in a single wet season in Miombo regions (Fuller, 1999; Zhang, Friedl, Schaaf, Strahler, & Liu, 2005). As forests in the Miombo are comprised of 85% deciduous species and agriculture is rain-fed, forest and non-forest areas have similar phenologies (Frost, 1996). In the dry season, fires occur on 15% or more of the landscape as part of either forest clearing, or other land-use activities such as understory burning to drive game animals or to spur growth of forage for livestock (Frost, 1996; Williams et al., 2008). The spectral properties of burned vegetation in satellite data are the same across different land cover categories. Thus, in burned areas, there is limited ability to differentiate whether the burn constitutes land conversion (e.g. forest cleared for agriculture) or variability in land cover conditions without conversion.

The second challenge is the high spatial variability of land cover conditions. Field sizes in the Miombo are small (0.5–2.0 ha) and irregularly shaped (Palm et al., 2010). Long-term shifting cultivation and chronic ecological disturbances have created patchy forest structures. Non-forest areas often have some degree of tree cover (Frost, 1996) (Fig. 2b). Forest clearing is done largely by hand; the largest trees are left standing due to limited labor and for their use as sources of fruit, medicine or other products (Campbell et al., 1996).

Though complex, the seasonal variability of Miombo landscape components presents a physical basis for differentiating forest and non-forest areas. Forest and non-forest surface components change with different patterns during transitions between wet and dry seasons. From the late wet–mid-dry season, forest areas retain green vegetation cover and have lower albedo than non-forest areas; senescence occurs over many weeks, and so canopy cover remains and shades underlying areas (Frost, 1996). Meanwhile, in non-forest areas such as agriculture or grassland regions, harvest and grazing of herbaceous forage occurs rapidly at the start of the dry season. These processes remove green vegetation material, leaving behind crop residues and other non-photosynthetic vegetation (NPV) and exposing substrate. On the ground, the physical differences in land-cover components between forest and non-forest areas are largest in the early to mid-dry season, though there are large variations due to the patchiness of forest cover and distribution of cultivated fields, homesteads and low-lying grassland regions. Past studies of forest change using satellite imagery in the Miombo have targeted the early to mid-dry season as an optimal time for forest change detection using measures of greenness such as NDVI (Prins & Kikula, 1996). However, no prior studies have attempted to physically model a broader suite of surface components, or develop and test metrics to distinguish forest cover that account for the seasonal variability of land cover conditions.

With this project, we develop and assess quantitative methods for distinguishing forest from non-forest land cover in the Tanzanian Miombo and apply them to study regional-scale forest changes. We use field data, multi-temporal and single-date analyses of 30 m Landsat data in the 2008 dry season to test how satellite observations of land cover correspond with surface components of forest and non-forest land cover. We examine the hypothesis, based on ground observations, that higher senesced/non-photosynthetic vegetation and substrate exposure at non-forest sites will distinguish non-forest from forest during the early–mid-dry season at scales of Landsat observations. We compare the accuracies of a multi-date classification and single date approaches to distinguish forest and non-forest areas, and qualitatively relate analyses to 250 m MODIS-scale observations. Subsequently we assess forest change patterns from 1995–2011 at regional scales and their variability across sub-regions with different land use pressures.

2. Methods

2.1. Study area and field data

Our study region is in central Tabora province, Tanzania (Fig. 1a). The landscape is dominated by a complex dry Miombo woodland ecosystem with locally diverse land uses (Fig. 1b–e) (Palm et al., 2010). Mean annual temperature is 23.9 °C and mean annual precipitation (2000–2009) is 770 mm with 90% falling between mid-November and early May. Agro-pastoral communities have lived in Tabora for centuries; political changes have strongly affected land use since the mid-19th century (Kjekshus, 1977). Since the 1990s, land change pressures have accelerated. Major drivers of land changes include population change (both migration and local growth), resource demands from urbanization including timber and charcoal, economic reforms leading to increased demand for land for cash crops such as tobacco, and road building (Chidumayo & Gumbo, 2013; Geist et al., 2009; Lambin et al., 2003; Yanda, 2010). In local sub-regions (Fig. 1d), Tabora includes the provincial capital Tabora Town (Fig. 5a), which has a

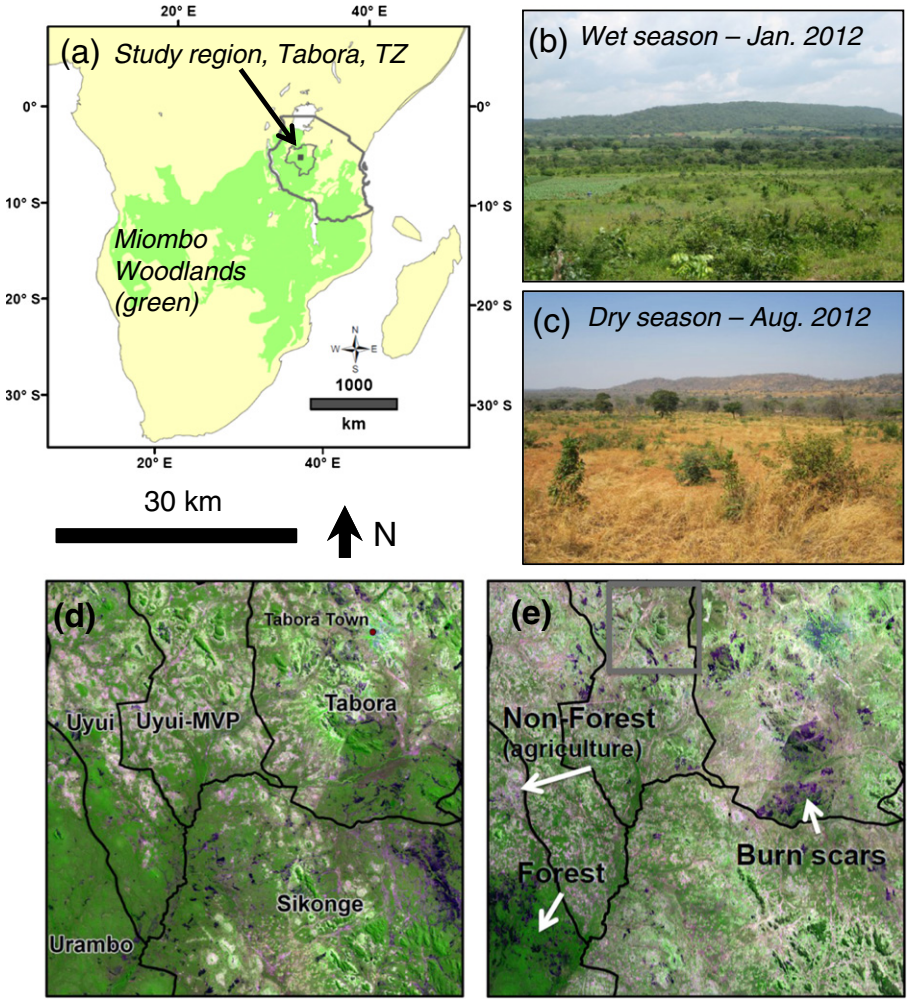

Fig. 1. Location, photos and Landsat data of the dry Miombo Woodland landscape at Tabora, Tanzania. (a) The study location is indicated by a square within central Tabora Province (small gray-bordered region) in western Tanzania, in the northeastern extent of the Miombo Woodlands (The Nature Conservancy). Photos show (b) wet and (c) dry season conditions across woodland and agriculture from the same viewpoint in 2012 (photos M. Mayes), illustrating confounded phenology and patchiness of the landscape. Landsat images from (d) 22 June 1990 (Landsat 5) and (e) 8 July 2013 (Landsat 8) [scale bar above (d)], colors mapped as red = short-wave infrared (~1.65 μm), green = near-infrared (~0.85 μm), and blue = short-wave infrared (~2.2 μm) show widespread inter-annual variability in land cover conditions and land cover changes. Dark green areas correspond to regions of dense green vegetation comprising forest canopy, while pale red and pink areas correspond with regions of higher substrate exposure comprising non-forest regions. Sub-districts studied for differences in local forest changes are indicated in (d). White arrows (e) indicate example forest, non-forest and burn scars. A gray box in (e) indicates the region over which 2007 Quickbird data was obtained for accuracy assessment.

rapidly growing population. Since 2008, a subset of villages in Uyui district has been involved with the Millennium Villages Project (MVP) (Uyui-MVP, Fig. 1d), a multi-faceted rural development program advocating increased use of mineral fertilizers and agroforestry to improve crop productivity and soil fertility (Palm et al., 2010). Elsewhere, a group of villages in Urambo district organized a local forest conservation management plan in 2001 to reduce forest clearing by pastoralists, following reform in national land tenure codes defined in The Village Land Act of 1999 that allowed such plans (Understanding Land Investment Deals in Africa, 2011). There is little baseline data available to assess impacts of these local land management efforts and change drivers on forest cover.

In Tabora, forest land cover includes closed and open-canopy tree-dominated areas, where trees are defined as woody plants taller than 3 m and canopy cover is greater than 30%, corresponding with Miombo ecological and forest research (Chidumayo, 2013; Frost, 1996) (Fig. 1e). Mature forest sites typically have not experienced clearing or fires destructive of large trees in 40 years or more and young forests are those that have regrown within the last 20–25 years. Non-forest land cover has tree canopy cover less than 30% and includes agriculture, seasonally flooding grasslands (called mbuga), village and urban centers, and open water (Fig. 1e). For agriculture, dominant crops are maize, cassava, groundnuts and tobacco (Palm et al., 2010; Yanda, 2010). Grazing occurs throughout the landscape but particularly in mbuga. To deal with confounding effects of fire, we designate a "Burn/Transition" land cover type for burn scars (Fig. 1e).

We collected field data in Tabora during the dry seasons of 2012 and 2013. We conducted site surveys, informal interviews about land-use practices, and established validation points for sites representing two forest classes (mature forest, young forest) and two non-forest land cover classes (agriculture and mbuga). We chose field sites considering stratification across mature forests, regrowth (young) forests, agriculture and mbuga classes, replication (at least 30 sites per class) and accessibility. These four classes broadly represent dominant land cover types and conditions save urban areas, which were not a focus of our study and comprise a very small proportion of regional land cover. Validation points were GPS points centered in areas at least 15 m from roads or footpaths. Points were used from 189 sites: 34 mature forest sites, 36 young (regrowth) forest sites, 85 agriculture sites and 34 mbuga sites.

2.2. Satellite data and pre-processing

We primarily used Landsat data (30 m) to study land cover dynamics, and used MODIS and Quickbird as ancillary data for pre-processing, qualitative comparisons and accuracy assessment. We obtained 12 Landsat scenes from the USGS-EROS open archive from 1990–2013 to assess land cover dynamics and forest change (Table 1). We obtained MODIS 500 m surface reflectance data for pre-processing and the MODIS 250 m Enhanced Vegetation Index (EVI) product from USGS to characterize greenness phenology across our field sites (Table 1). A July 2007 Quickbird image obtained from MVP was used for accuracy assessment.

Pre-processing of Landsat-based data for studying dry season land cover variability involved calibration to top-of-atmosphere reflectance (*p*TOA) and surface reflectance (*p*SUR) of 5 Landsat images from May–August 2008 (Table 1) (see Supplementary Data for a diagram of the pre-processing workflow). To retrieve surface reflectance (*p*SUR) we used dark-object subtraction (DOS) prior to calibration to *p*TOA (Song, Woodcock, Seto, Lenney, & Macomber, 2001). To improve the spectral response we used relative radiometric calibration techniques to scale Landsat surface reflectance values to those of MODIS 500 m data (Furby & Campbell, 2001). We scaled Landsat surface reflectance to MODIS product band-wise, using invariant targets of low, medium and high albedo of equal area between the images (Table 1). For the forest change analysis, seven 1990–2013 Landsat scenes were calibrated to *p*TOA and inter-calibrated to one *p*SUR-corrected scene from the seasonal land cover analysis (10 July 2008), similar to scene inter-calibration methods used by Elmore, Mustard, Manning, and Lobell (2000) and Sabol, Gillespie, Adams, Smith, and Tucker (2002) (Table 1 and Supplementary Data S3). Topographic correction was unnecessary because the study area was essentially flat, save a few scattered small hills.

2.3. Linear spectral mixture analysis of Landsat data to model dry season land cover components and qualitative comparisons to MODIS EVI

Linear spectral mixture analysis (LSMA) characterizes satellite pixel reflectance spectra as a multiplicative sum of spectral endmembers corresponding with known materials, or endmembers (Adams & Gillespie, 2006) (Fig. 2a). LSMA has been used effectively to characterize land cover variability and change globally, from North America (Elmore et al., 2000; Roberts et al., 1998; Yang, Weisberg, & Bristow, 2012) to the Amazon (Adams et al., 1995) to Sudan (Dawelbait & Morari, 2012). Steps in LSMA involve (1) selecting physically reasonable spectral endmembers either from direct spectroscopic sampling or from satellite images themselves, (2) evaluating the physical and mathematical reasonability of the mixture model result, and (3) using images of endmember proportions, or fraction images, to derive landscape data.

A key part of successful image endmember selection with LSMA at Landsat scales is selecting endmembers from locations where the ground coverage of materials (components) of interest is as homogeneous as possible. In Miombo woodland ecosystems, the purest pixels of different surface components occur at different times in the dry season. For example, pure pixels of green vegetation cover occur over un-harvested farm fields or forest tree canopies during the wet season and through the early dry season. Pure pixels of substrate or NPV occur during the mid or late dry season after harvest exposes larger areas of bare soil, or senescence in forest areas results in woody materials and senesced leaves as dominant surface components.

As such, we used a multi-temporal endmember selection strategy, wherein we mined five Landsat 5 images, from May–August 2008, to develop a single set of endmembers for LSMA analyses to assess seasonal land cover variability and change. We searched for the purest spectra of three dominant surface components comprising land cover across May–August 2008 image dates: green vegetation, non-photosynthetic vegetation, and substrate (Fig. 2a). We chose endmembers in 3×3 pixel areas near field sites where we observed surface components were as homogeneous as possible relative to Landsat spatial scales (30 m). We confirmed that candidate locations for endmembers had experienced minimal land cover changes between 2008 and the time of field data collection. We also used a "synthetic shade" endmember to account for low albedo at sub-pixel scales from surface components (e.g. charred leaves and woody material) or shading (Adams & Gillespie, 2006). We evaluated LSMA models with 19 candidate multi-temporal endmember sets to find the most reasonable solution. The choice of our final endmembers was based on optimization using two quantitative criteria 1) minimization of root mean square error (RMSE) and minimum % of pixels modeled with endmember fractions <0% or >100% across May–August 2008 Landsat images. After experimental trials of unmixing with endmembers across the fives Landsat dates, we checked spatial patterning in the root mean square image visually to identify patterns of model success or failure over known areas of forest, fields, mbuga grasslands, urban areas, or rock outcrops. The final endmember set included a bright vegetation spectrum from a farm field in the 7 May 2008 image, a substrate spectrum from a harvested field in the 10 July 2008 image, and an NPV spectrum from senesced woodland in the 10 July 2008 image (Fig. 2a). Although global-scale analyses of the Landsat spectral mixing space have found substrate and NPV spectra often difficult to distinguish (Small & Milesi, 2013), spectral differences in the near-infrared (NIR) (Landsat 5 Band 4, 0.83 μm broadband median) and short-wave infrared (SWIR) (Landsat 5 Band 5, 1.65 μm broadband median) were consistent

Table 1
Landsat and MODIS data obtained for (a) the quantification of dry season surface feature variability and (b) forest change analyses. Surface reflectance scaling dataset indicates the surface reflectance-corrected image data to which analysis images were normalized.

Analysis	Satellite images(s)	Date	Surface reflectance scaling dataset
Seasonal variability of Miombo land cover components	Landsat 5	7 May 2008	MOD09A1 (Surface reflectance) 500 m, 8-day (8 May)
	Landsat 5	24 June 2008	MOD09A1 (Surface reflectance) 500 m, 8-day (8 May)
	Landsat 5	10 July 2008	MOD09A1 (Surface reflectance) 500 m, 8-day (8 May)
	Landsat 5	11 Aug. 2008	MOD09A1 (Surface reflectance) 500 m, 8-day (8 May)
	Landsat 5	27 Aug. 2008	MOD09A1 (Surface reflectance) 500 m, 8-day (8 May)
	MODIS EVI (Enhanced vegetation index, MOD13Q1)		Not applicable
Forest cover change, 1995–2011	Landsat 5	22 June 1990	Landsat 5–10 July 2008 (scaled to MOD09A1)
	Landsat 5	7 July 1995	Landsat 5–10 July 2008 (scaled to MOD09A1)
	Landsat 5	26 June 1997	Landsat 5–10 July 2008 (scaled to MOD09A1)
	Landsat 5	1 July 2001	Landsat 5–10 July 2008 (scaled to MOD09A1)
	Landsat 5	13 June 2004	Landsat 5–10 July 2008 (scaled to MOD09A1)
	Landsat 5	24 June 2008	Landsat 5–10 July 2008 (scaled to MOD09A1)
	Landsat 5	3 July 2011	Landsat 5–10 July 2008 (scaled to MOD09A1)
	Landsat 5	8 July 2013	Landsat 5–10 July 2008 (scaled to MOD09A1)

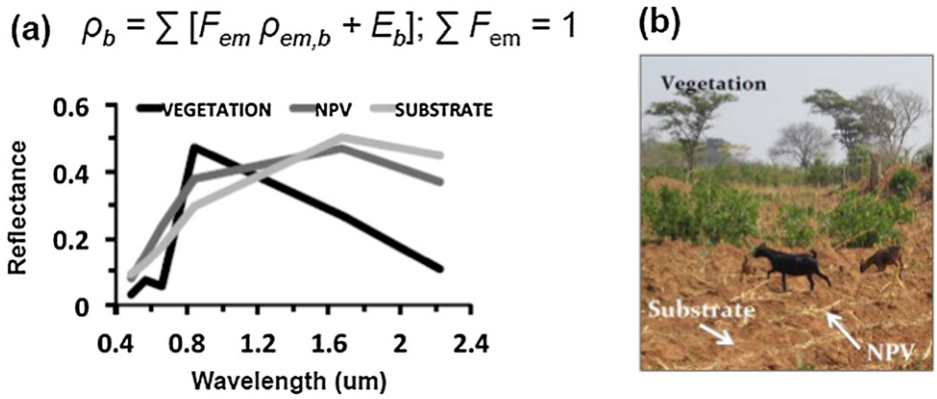

Fig. 2. Spectral endmembers used for linear spectral mixture analysis (LSMA) of Landsat data. (a) The governing equation for LSMA and the multi-temporal set of endmember reflectance spectra used to study surface components of forest and non-forest land cover. LSMA models pixel reflectance as multiplicative fractions (*Fem*) of endmember reflectance for each satellite band (*pem,b*) plus an error term (Eb), under the constraint that the multiplicative sum of all endmember fractions must equal 1. Endmembers include green vegetation (7 May 2008), non-photosynthetic vegetation (NPV) (10 July 2008), substrate (10 July 2008) and a shade endmember of reflectance = 0 for all bands to represent sub-pixel shading and dark albedo objects (not shown). (b) Photo of a typical agricultural field in Tabora, where surface components (labeled) are heterogeneous at the spatial resolution of Landsat (30 m) and MODIS Enhanced Vegetation Index (EVI) data (photos M. Mayes).

between substrate and NPV endmembers, and substrate and NPV-rich sites in Tabora. The ratio of SWIR(1.55 μm):NIR at endmember sites, and field validation sites with dominant substrate and NPV cover, highlights how the spectral differences between members, if subtle, persist with some consistency in this landscape. Assessing means (±standard errors), substrate-rich sites have SWIR–1NIR ratios of [1.50 (±0.08), n = 3], much closer to the substrate endmember ratio [1.70], while NPV-rich sites have SWIR-1:NIR ratios of [1.24 (±0.08), n = 3] which are close to the NPV-rich endmember [1.28] (see Supplementary Data S4.) Upon validation of the unmixing analyses, we used the same endmember set from 2008 for LSMA analyses of seasonal land cover variability in 2008 and forest changes from 1995–2011.

To compare Landsat analyses against MODIS land cover measures, we downloaded MODIS EVI data at 250 m for 2000–2009 over our study area (Table 1, Fig. 5). EVI greenness time series data and Landsat LSMA model endmember fractions were extracted and plotted at field site GPS points.

2.4. Comparison of multi-temporal classification versus single-date metrics for identifying forest cover across the 2008 dry season with Landsat images

We used two approaches to evaluate seasonal patterns in surface components between forest and non-forest regions quantified by Landsat LSMA: (1) multi-temporal, unsupervised classification of fraction images (Fig. 3a), and (2) metrics calculated for single image dates, derived from LSMA analyses (Fig. 3b). First, we used an unsupervised Isodata classification algorithm to study fraction image variability across the 2008 dry season Landsat images, and identify groups of multi-temporal patterns that we could subsequently relate to land cover types. Others have used unsupervised classification to identify groups of multi-temporal patterns in fraction images of semi-arid ecosystem surface features, and relate them to different land cover types or disturbances in cases where land cover dynamics are complex and poorly characterized (Elmore, Mustard, & Manning, 2003). For our analysis we used an Isodata classifier in ENVI 5.1 (Exelis) stipulating 10–30 classes, 10 iterations, a minimum of 1000 pixels per class, and default settings for merge pairs and class standard deviations, following earlier uses of this method (Elmore et al., 2003). As input to the classifier, we included endmember fraction images from all reasonable LSMA results in a time-stacked data cube. In early experiments low-albedo burned scars confounded all land cover groups. To deal with these effects, we developed a burn/transition area mask for each image date using thresholds based on the shade endmember and the Normalized Burn Index (NBR) (van Wagtendonk, Root, & Key, 2004). NBR is a normalized ratio of the difference between reflectance at near infrared (broad-band center at 0.83 μm for Band 4, Landsat 5) and short-wave infrared (broad-band center at 2.22 μm, Band 7, Landsat 5), and identifies burned areas on the basis of their high shortwave relative to near-infrared reflectance. NBR-based masks were conservative and meant to eliminate areas that confounded land cover types rather than map burn areas rigorously. We performed final unsupervised classifications with a composite burn/transition area mask including the masks for dates of all valid Landsat LSMAs. We displayed classes against Landsat *p*SUR and LSMA images, and grouped raw classes into condensed super-classes that matched land cover types. We used a median filter (3 × 3) to eliminate spurious pixel variations. We assessed the accuracy of forest, agriculture and mbuga super-classes via stratified random sampling of the multi-temporal classification result and confusion matrix analysis (see Supplementary Data S2 for confusion matrices) over an area covered by a Quickbird image from July 2007 (Fig. 1e). A total of 392 points were generated for use in accuracy assessments (see Supplementary Data).

Second, we experimented with algorithms to identify forest from non-forest areas using LSMA fractions from imagery taken at single dates (Fig. 3b). Based on our hypothesis, we focused on developing a metric using NPV and substrate fraction images, and took a thresholding approach to identify forest and non-forest land cover. We evaluated the metric's performance for differentiating forest and non-forest across the dry season using analysis of variance (ANOVA) and non-parametric Kruskal–Wallis analyses of values extracted at pixels corresponding with field sites (see Supplementary Data S1 in Supporting Information). Land cover class and image date (time) were used as the independent variables. For the ANOVA, field site data was log-transformed to satisfy assumptions of normality.

Forest and non-forest areas were identified finally using thresholds calculated based on multiples of standard deviations of metrics from field site data and checked with visual validation. Final single-date land cover maps had three classes: forest, non-forest and transition/burned area. We used a median filter (3 × 3 pixels) to eliminate spurious pixel variations and assessed single-date forest cover estimates for accuracy against the same set of stratified randomly sampled points taken for the multi-temporal classification image.

2.5. Estimates of forest cover loss and gain patterns by sub-district from 1995–2013

We produced land cover maps for seven dry season Landsat scenes from 1990–2013 using thresholding of an LSMA-based metric from single image dates (Fig. 3c). To minimize spurious identification of land conversions due to inter-annual variability, we used inter-annual

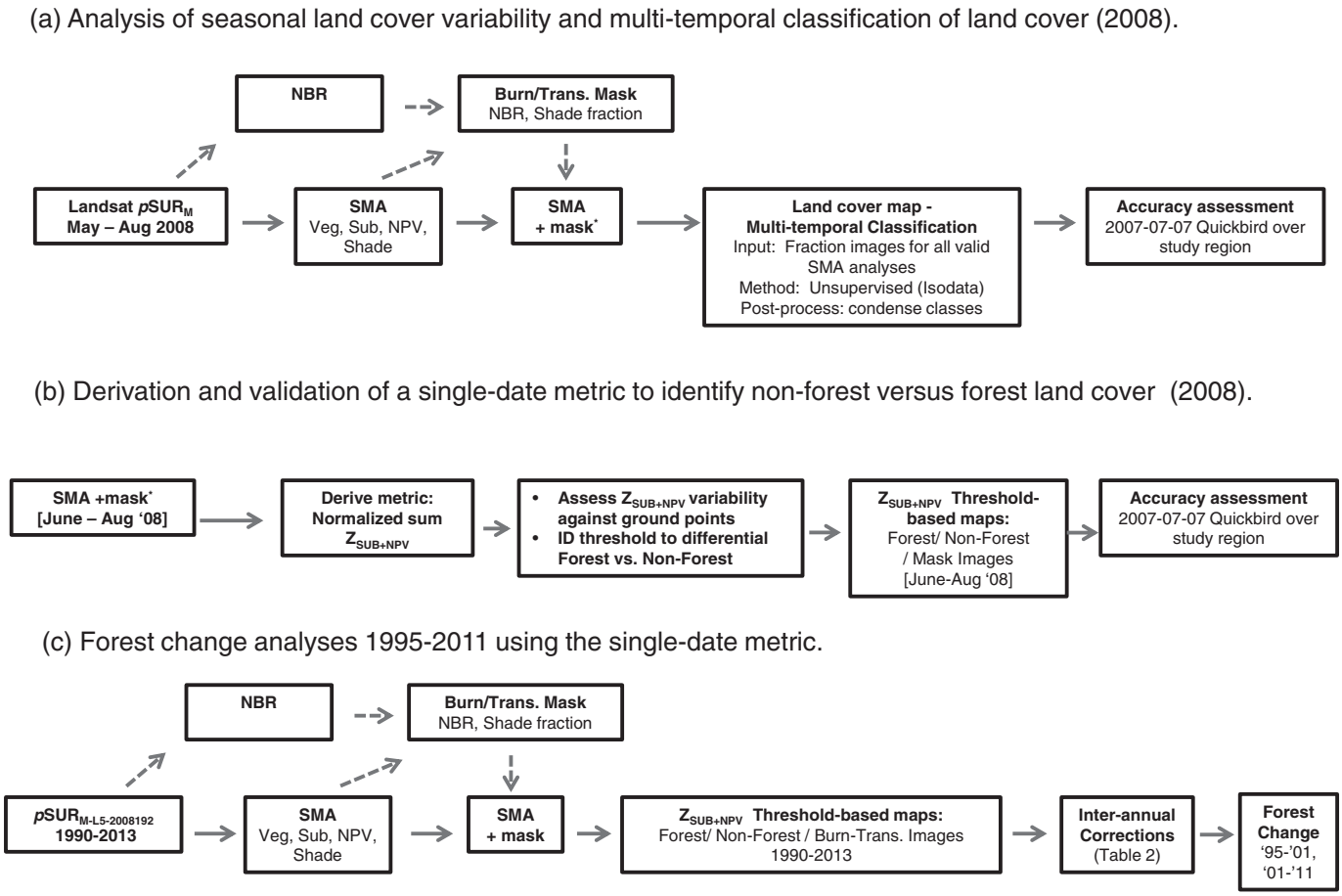

Fig. 3. Workflow for image processing methods, including (a) assessment of seasonal patterns of land cover variability in 2008, (b) derivation and validation of single-date metrics to identify forest vs. non-forest land cover, and (c) forest change analyses.

corrections to define sustained forest loss and gain similar to Toomey et al. (2013) (Table 2). The extra time requirement for forest gain was selected considering results from long-term observations of Miombo tree succession that have shown that tree recovery to minimum definitions of forest cover (3 m height) takes 7–10 years (Chidumayo, 2013). Given the data requirements of inter-annual corrections, we assessed forest changes in two periods, from 1995–2001 and 2001–2011, with the 1990 and 2013 Landsat images used as part of inter-annual corrections but not as dates for evaluating forest change. For 1995–2001, forest losses and gains were assessed at 1997 and 2001. For 2001–2011, forest losses were assessed at 2004, 2008 and 2011, and gains at 2004 and 2008; gains were not assessed at 2011 because given our correction protocol, not enough time has passed since 2008 (only 5–6 years) to evaluate sustained forest gain. Gains and losses were summed and evaluated region-wide and by sub-district. To normalize forest land conversion data for comparison across sub-districts, we calculated two additional metrics. The first, aggregate forest change, quantifies the percentage of forest area experiencing conversion (loss or gain) relative to total forest area. It is an "activity gauge" for forest conversion, with larger percentages indicative of more dynamic changes (Hansen et al., 2013). The second metric is the ratio of forest loss area to gain area (Loss:Gain), which shows the balance of land conversion processes.

3. Results

3.1. Patterns in surface components and their seasonal variability at forest and non-forest sites in Landsat LSMA analyses and MODIS EVI data

Forest and non-forest sites had differing proportions of surface components that had similar patterns of seasonal variability during the 2008 dry season. LSMA models for June–August Landsat images were within ranges of reasonable values, with greater than 95% of pixels across all fraction images modeled between 0 and 100% (Table 3). RMS errors varied between 0.013 and 0.019 and showed few spatial patterns. The May LSMA model was invalid because 9.4% of the substrate and 46% of non-photosynthetic vegetation (NPV) fraction image pixels were

Table 2
Inter-annual correction algorithms used to assess forest losses and gains. T(0) refers to the date at which change assessments are made. Negative numbers in parentheses indicate image dates two (−2) and one (−1) date prior to the date of change assessment. Positive numbers in parentheses indicate conditions for image dates following the date of change assessment. Trans stands for "transition class" comprising burned areas or other regions whose forest/non-forest status is unclear.

Inter-annual corrections for forest change analyses					
	Time point and land cover class status				
Change type	T(−2)	T(01)	T(0)	T(+1)	T(+2)
Forest loss	Forest	Forest	Non-forest burn/trans	Non-forest burn/trans	None
Forest gain	Non-forest burn/trans	Non-forest burn/trans	Forest	Forest	Forest

Table 3
Dry season variability of surface components comprising total Landsat pixel reflectance from May–August 2008 in Tabora, Tanzania, assessed by linear spectral mixture analysis (LSMA). Proportions (mean ± standard deviation) describe the average contributions of spectral endmembers to pixel reflectance across the study area. The percentage of pixels with spectral endmembers modeled with negative proportions (% < 0) or proportions greater than 1 (% > 1) should equal less than 5% of all pixels analyzed in reasonable LSMA models (Adams & Gillespie, 2006).

Date	Endmember	Mean (± σ)	% < 0	% > 1
7 May	Vegetation	0.436 (0.089)	0.133	0.004
7 May	Substrate	0.174 (0.140)	9.455	0.004
7 May	NPV	0.053 (0.133)	45.967	0.035
7 May	Shade	0.337 (0.100)	0.344	0.000
7 May	RMS	0.008 (0.003)		
24 June	Vegetation	0.227 (0.068)	0.315	0.000
24 June	Substrate	0.210 (0.117)	3.202	0.001
24 June	NPV	0.209 (0.127)	1.017	0.012
24 June	Shade	0.354 (0.088)	0.016	0.000
24 June	RMS	0.013 (0.003)		
10 July	Vegetation	0.199 (0.071)	0.590	0.000
10 July	Substrate	0.250 (0.122)	2.108	0.003
10 July	NPV	0.235 (0.143)	1.085	0.026
10 July	Shade	0.316 (0.092)	0.035	0.000
10 July	RMS	0.013 (0.003)		
11 August	Vegetation	0.128 (0.058)	2.871	0.000
11 August	Substrate	0.283 (0.122)	1.792	0.003
11 August	NPV	0.291 (0.141)	0.484	0.074
11 August	Shade	0.298 (0.089)	0.030	0.000
11 August	RMS	0.019 (0.004)		
27 August	Vegetation	0.133 (0.069)	3.583	0.000
27 August	Substrate	0.276 (0.162)	4.018	0.025
27 August	NPV	0.377 (0.167)	0.532	0.252
27 August	Shade	0.213 (0.102)	1.342	0.000
27 August	RMS	0.015 (0.005)		

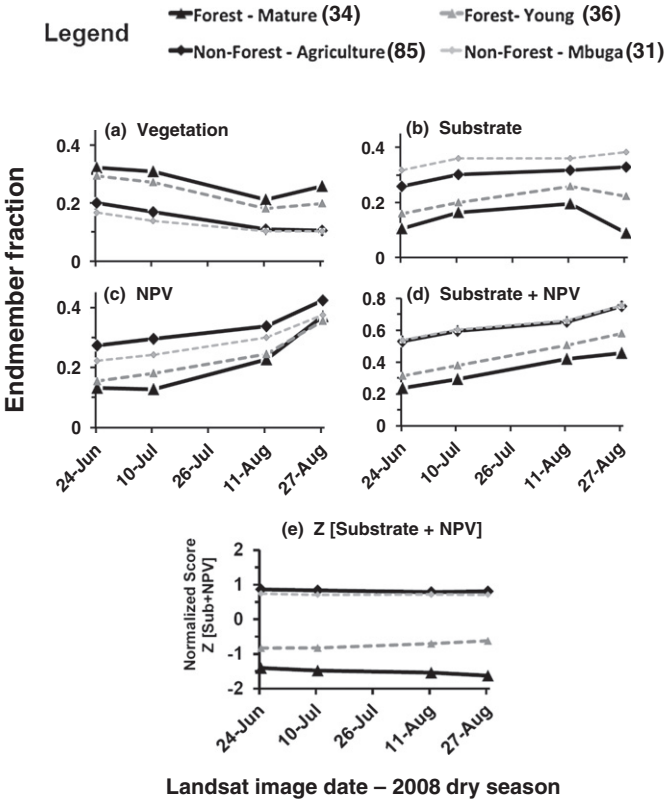

Fig. 4. Linear spectral mixture model (LSMA) results at 189 field sites, assessed in four dry season Landsat images (June–August 2008) with valid LSMA models. The field sites include mature forest ($n = 34$, verified in the field as forest sites for 40 years or more), young forest ($n = 36$, verified in the field as secondary forest regrown since the mid-1990s), agriculture ($n = 85$, annual crops regularly grown) and mbuga ($n = 34$). (a) Green vegetation endmember fractions among land cover types from June–August '08. (b) Substrate endmember fractions, (c) non-photosynthetic vegetation endmember fractions, and (d) Summed substrate and NPV fractions among land cover types from June–August '08. (e) Normalized sum of substrate and NPV fractions among field sites from June–August '08. Note the difference in vertical axis scale.

modeled with less than 0% endmember proportions (Table 3). In the May image as well as others, most pixels with negative fractions were either (1) substrate-dominated with negative NPV fractions, (2) NPV-dominated with negative substrate fractions, or (3) high shade pixels with negative substrate fractions. Many pixels in categories (1) and (3) were ultimately masked as likely burn-affected areas by combined thresholding of the shade fraction images and Normalized Burn Index images (Fig. 3), so were not used to distinguish forest versus non-forest area at single image dates. The standard deviations of the substrate and NPV fraction images showed most of the sub-0% proportions were not far from 0%, with most equaling less than −5% (substrate) and −10% (NPV).

Proportions of vegetation, substrate and NPV surface components showed consistent differences between forest and non-forest sites at Landsat scales from June–late August (Fig. 4). Shade fractions showed high sensitivity to burned areas across all image dates; however relative to burned areas, patterns among non-burned forest and non-forest sites were inconsistent. Vegetation fractions (Fig. 4a) were 0.10–0.15 greater at forest than non-forest sites (Fig. 4a), while substrate fractions (Fig. 4b) and NPV fractions (Fig. 4c) were 0.07–0.14 greater at non-forest than forest sites. Between mature and young forests, mature forests had slightly greater vegetation fractions (Fig. 4a). Among non-forest land cover types, mubga sites showed slightly greater substrate fractions (Fig. 4b) while agriculture sites showed greater NPV fractions (Fig. 4c). A sum of substrate and NPV fractions (Fig. 4d) accentuated quantitative differences between forest and non-forest land cover types.

The seasonal variability of surface components was similar across sites at Landsat and MODIS scales. At Landsat scales, vegetation fractions decreased while substrate and NPV fractions generally increased across all sites from June–August (Fig. 4a, c). Forest sites showed a slight deviation from the general pattern in late August, where vegetation fractions increased slightly while substrate fractions decreased. At MODIS scales, EVI greenness phenologies were similar across land cover types (Fig. 5). Forest sites showed slightly higher peak EVI during the peak of the wet season (Jan–Feb) and non-forest sites showed slightly lower minimum EVI near the end of the dry season (Sept–Oct). Increases in EVI began slightly earlier for forest than non-forest sites (one to two 16-day observations), though the pattern showed high inter-annual variability.

3.2. Identifying forest versus non-forest land cover: evaluating the performance of multi-temporal classification of LSMA fractions versus thresholding of LSMA analyses from single dates

Multi-temporal classification and single image date-thresholding approaches gave comparable results for differentiating forest versus non-forest land cover across the 350.7 thousand hectare (kha) analysis region. For multi-temporal classification (see Methods, Section 2.4), we used all valid Landsat LSMA fraction images sets – four in total from 24 June–27 August – as inputs for the analysis. The classifier produced 10 raw classes of pixels, which we grouped into four coherent super-classes: forest, agriculture, mbuga and other (including urban areas, large quarries, water). To verity classes we used overlays and visual validation with raw Landsat and Google Earth images. The area of the burn/transition mask derived from pre-processing (Fig. 3a) was included as a fifth class to produce a final land cover map representing general land cover conditions for the 2008 dry season (Fig. 6a). With separate classes for forest, agriculture and mbuga, the classification-based map had an overall accuracy of 82.1% (Fig. 6c), producer accuracies >80% for forest and agriculture, and high consumer accuracy for forest (93.4%) (Table 4). With a non-forest class combining agriculture and mbuga, a binary, forest/non-forest land cover map achieved an overall accuracy of 86.5% (Fig. 6c) with all producer and consumer accuracies

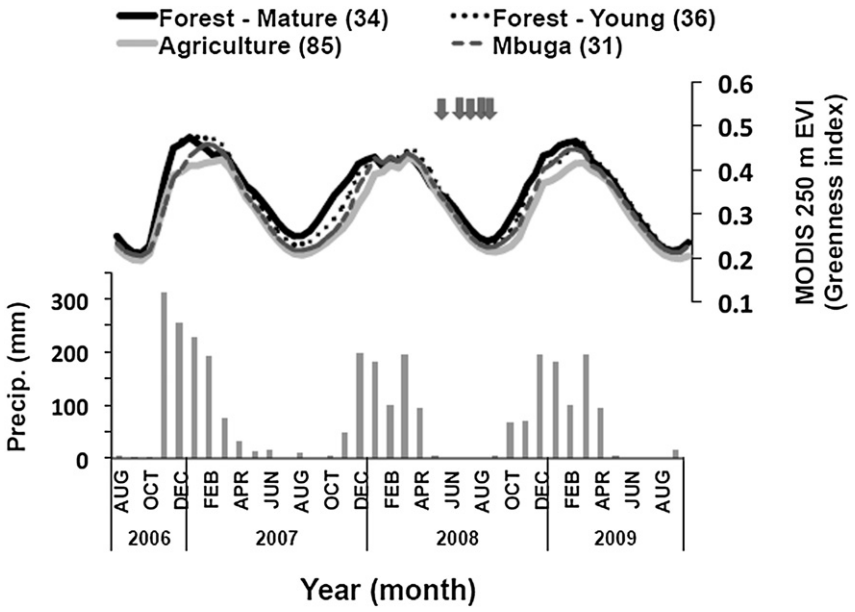

Fig. 5. Greenness phenology from 2006–2009 across 186 field sites (top) and averaged monthly precipitation data (bottom) from the Tabora Airport [US National Climate Data Center Global Station (GHCND) TZ000063932]. Greenness phenology is measured with the MODIS Enhanced Vegetation Index (EVI) product (MOD13Q1), at 250 m resolution in 16-day composited intervals. The number of field sites includes three fewer mbuga sites compared to Fig. 3 because multiple mbuga field sites fell within the same MODIS 250 m pixel footprint. Five gray arrows indicate the timing of the 2008 Landsat images used to study surface components through the dry season. Precipitation and greenness phenology are closely correlated, with the growing season beginning with the onset of the rains in October–November and ending around April–May.

near 80% or above (Table 4). Burn/transition areas across all image dates summed to 15.3% of the analysis area.

The thresholding approach, based on LSMA fractions from single dates of Landsat imagery, achieved similar performance for mapping forest versus non-forest land cover from June–August 2008. Building on our finding that non-forest sites had larger substrate and NPV fractions than forest sites (Results, 3.1), we created an index to separate non-forest and forest cover based on the sum of substrate and NPV fraction images (Fig. 3b). The relative difference of summed substrate and NPV fractions stayed consistent between forest and non-forest sites from June–August but the absolute values of these sums gradually increased (Fig. 4d), complicating use of the raw summed fractions as a threshold robust against seasonal variability. To derive an index robust against seasonal variability, we calculated normalized sum images (Z-score images) for the summed substrate and NPV fractions: $(\Sigma[\text{SUB} + \text{NPV}]\text{pixel} - \Sigma[\text{SUB} + \text{NPV}]\text{image}_{\mu}) / \Sigma[\text{SUB} + \text{NPV}]\text{image}_{\sigma}$; hereafter, Z(Sub + NPV) (Fig. 4e). The normalized sums of substrate and NPV fractions showed consistent absolute values by forest versus non-forest classes across field sites from June–August (Fig. 4e). Analysis of variance (ANOVA) and non-parametric tests validated that

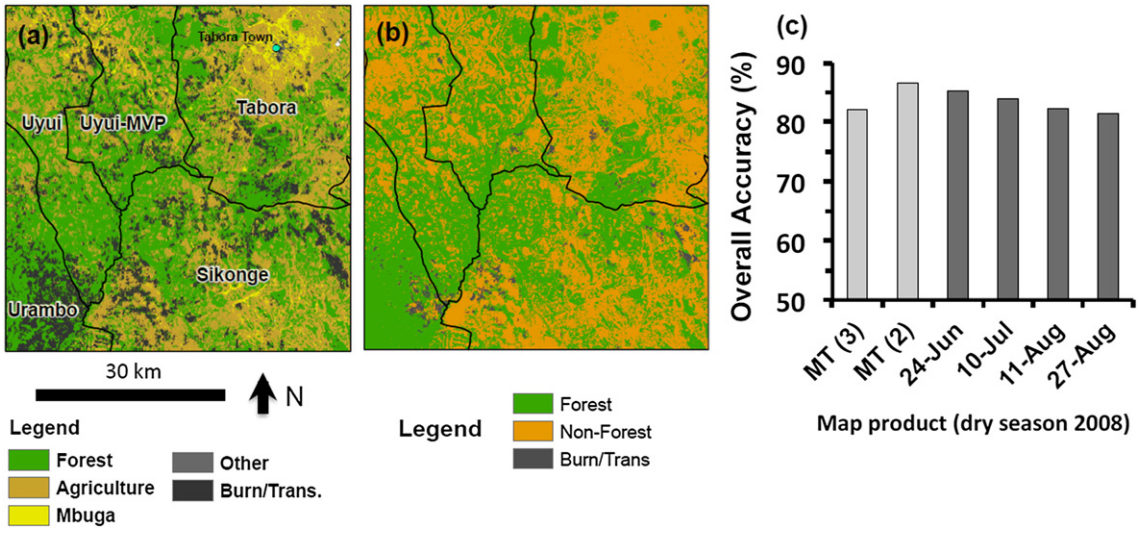

Fig. 6. Comparisons of multi-date unsupervised classification and single-date threshold-based analyses of 2008 land cover at Tabora, Tanzania. The scale bar under (a) corresponds to both panels (a) and (b). (a) Land cover map derived from multi-temporal, unsupervised Isodata classification of a time-stacked data cube of LSMA results from June–August 208, depicting forest (mature + young), agriculture, mbuga, other (urban, water) and the burn/transition mask area. Overall accuracy of the multi-temporal, unsupervised classifications separating forest, agriculture and mbuga (MT-3 in Fig. 6c) was 82.1%. (b) Land cover map derived from thresholding of an LSMA-based metric (Z[Sub + NPV]) from a single image date (24 June 2008), depicting forest, non-forest and burn/transition classes with overall accuracy of 85.0% (24-Jun in Fig. 6c). The non-forest class includes mbuga and "other" classes in (a). (c) Overall accuracies of multi-temporal classifications (light gray bars, MT(3) separating agriculture and mbuga, MT(2) combining them) and single-date assessments (dark gray bars) show that single-date, threshold-based analyses approach the accuracy of the multi-temporal unsupervised classifications (light gray bars to the left). In the single-date analyses, accuracies are highest in the early dry season.

Table 4
Producer and consumer accuracies for multi-temporal classification and single-date LSMA metric-derived maps of Tabora, Tanzania land cover for 2008.
Please refer to supplementary materials for full accuracy assessments and confusion matrices.

Land cover	Accuracy	Multi-temporal class.		Single-date LSMA metric (Z[sub + npv])			
Class	Metric	3-class	2-class	24 June	10 July	11 August	27 August
Forest	Producer	83.0	83.0	83.0	78.9	78.5	75.3
Non-forest	Producer		92.2	89.3	91.7	88.2	90.5
Agriculture	Producer	83.9					
Mbuga	Producer	69.6					
Forest	Consumer	93.4	93.4	90.2	91.7	88.8	90.1
Non-Forest	Consumer		79.7	79.9	76.7	75.6	73.4
Agriculture	Consumer	74.1					
Mbuga	Consumer	53.3					

Z(Sub + NPV) significantly differentiated forest and non-forest across field sites. Among ANOVA models, the one that best accounted for Z(Sub + NPV) variability included only land cover as an independent variable ($r2 = 0.5838$, $p < 0.0001$) (see Supplementary Data S1). Additive and interactive effects of land cover and time on Z(Sub + NPV) were not significant. Non-parametric Kruskal–Wallis models found significant land cover effects ($p < 0.0001$) without significant time effects (see Supplementary Data S1). Forest sites, grouping mature and young sites, had Z(Sub + NPV) means which were smaller and significantly different from agriculture and mbuga site means ($p < 0.05$) (see Supplementary Data S1). We calculated a single Z(Sub + NPV) threshold to separate forest from non-forest using pooled forest site data from June–August Landsat analyses ($n = 280$ for forest sites, $n = 476$ for non-forest sites). Verified by visual validation, we used the following threshold as a cutoff for forest versus non-forest area, based on multiples of the standard deviation of pooled mature and young forest site points (FOREST, below) across June–August:

$$\text{Forest}: \; Z(\text{Sub}+\text{NPV})_{\text{pixel}} \le \text{FOREST}\,\mu[Z(\text{Sub}+\text{NPV})] + 1.5 * \text{FOREST}\,\sigma[Z(\text{Sub}+\text{NPV})].$$

$$\text{Non-Forest}: \; Z(\text{Sub}+\text{NPV})_{\text{pixel}} > \text{FOREST}\,\mu[Z(\text{Sub}+\text{NPV})] + 1.5 * \text{FOREST}\,\sigma[Z(\text{Sub}+\text{NPV})]$$

where

$$\text{FOREST}\,\mu[Z(\text{Sub}+\text{NPV})] + 1.5 * \text{FOREST}\,\sigma[Z(\text{Sub}+\text{NPV})] = -0.068.$$

Values below the threshold, having low proportions of substrate and NPV relative to the image mean, were considered forest and values above this threshold non-forest.

For the June–August 2008 Landsat images we applied Z(Sub + NPV) thresholds, combined with *each date's individual* transition/burn maps (Fig. 3c, methods) to produce land cover maps of forest, non-forest and burn/transition areas (e.g. 24 June 2008, Fig. 6b). Map accuracies assessed via confusion matrices (see Supplementary Data S2) were highest for the June analysis (85.3%) and declined to 81.4% in late August (Fig. 6c). Across the total Landsat analysis area, from June–August burn/transition areas increased from 1.9–12.8% of the analysis area. From 24 June–10 July maps, which had the highest accuracies, there was a −4.7% change in the forest area estimate, a +3.2% change in non-forest area estimate and +0.77% change in burn area estimate (relative to 24 June areas). For the forest class, consumer accuracies were greater across all dates (Table 4). For non-forest, producer accuracies were higher across all dates (Table 4).

3.3. Tabora forest cover changes from 1995–2011

We found large gross areas of forest loss and gain across our study region. Single-date LSMA and inter-annual corrections identified a physical signature of land conversions between forest and non-forest: sustained increases in normalized proportions of substrate and NPV components indicated conversion from forest–non-forest, and a sustained decrease indicated conversion from non-forest–forest (Fig. 7). Overall, forest cover decreased from 163.5 kha (46.6% of the land cover) in 1995– 139.0 kha (39.6%) in 2011 (Table 5, Fig. 8). This change is a net 15.0% loss of 1995 forest area, and a 7.0% overall reduction in the total forest-occupied land cover. For 1995–2001, gross forest losses were 22.5 kha at an annualized rate of −3.75 kha and gains were 9.0 kha at an annual rate of 1.45 kha yr^{-1} (Table 5). For 2001–2011, gross forest losses were 29.2 kha at a rate of −2.92 kha yr^{-1} and gains were 13.3 kha at a rate of 1.33 kha yr^{-1} (Table 5). Between decades, rates of gross forest loss decreased while rates of gross forest gain were similar for the total study area, with many variations by sub-district (Table 5). The total area of gross forest gain was substantial, totaling 13.6% of the 1995 forest area and 43.1% of the gross forest loss area. Aggregate forest change, the proportion of total forest area experiencing change in either direction (loss or gain), was 19.2% for 1995–2001 and increased to 27.4% for 2001–2011 for the whole 350 kha study area, while the ratio of forest loss to gain areas was 2.51 during 1995–2001 and 2.20 during 2001–2011.

To normalize forest land conversion data for comparison across sub-districts, we calculated two additional metrics. The first, aggregate forest change, quantifies the percentage of forest area experiencing conversion (loss or gain) relative to total forest area. It is an "activity gauge" for forest conversion, with larger percentages indicative of more dynamic changes (Hansen et al., 2013). For our entire study area, aggregate forest change was 19.2% for 1995–2001 and increased to 27.4% for 2001–2011. The second metric is the ratio of forest loss area to gain area (Loss:Gain), which shows the balance of land conversion processes. Loss: Gain was 2.51 during 1995–2001 and 2.20 during 2001–2011 for the total study area.

Among sub-districts, patterns of forest loss and gain dynamics showed distinct variations (Table 5, Fig. 8). Sikonge had the largest absolute areas of forest conversion, but also the largest land and forest area so that percentage of forest area affected by conversions was moderate. The forest loss:gain ratio dropped from 3.4–2.6 indicating a decrease in forest losses relative to gains from 1995–2001 to 2001–2011. In Tabora, forest losses nearly doubled from 4630 ha–8900 ha and gains also increased substantially (Fig. 8, Table 5). Given that Tabora had the smallest initial forest area, aggregate changes affected a much larger percentage of sub-district forest in 2001–2011 (19.4%) compared to the prior decade. The Uyui-MVP sub-district area, by contrast, had slightly lower rates of forest loss and slightly higher rates of forest gain than non-MVP areas, such that the region had lower Loss:Gain in 1995–2001 (1.4) and 2001–2011 (1.6) than the overall landscape. Urambo showed the largest shift in forest change trajectory: gains increased by an order of magnitude from 174 ha during 1995–2001 to 2050 ha during 2001–2011, resulting in a dramatic drop in forest Loss:Gain from 19.4–1.4 and a reduction in percent forest area loss from 10.2%–3.0%.

4. Discussion

4.1. Lower proportions of substrate and NPV components differentiate forest from non-forest land cover throughout the dry season, while the temporal variability of components is similar across land cover types

Spectral mixture modeling proved to be a sharp lens through which to analyze dry season land cover variability in Tabora. Forest sites had lower fraction image proportions of substrate and NPV than non-forest sites from June–August, during the early and mid-dry season.

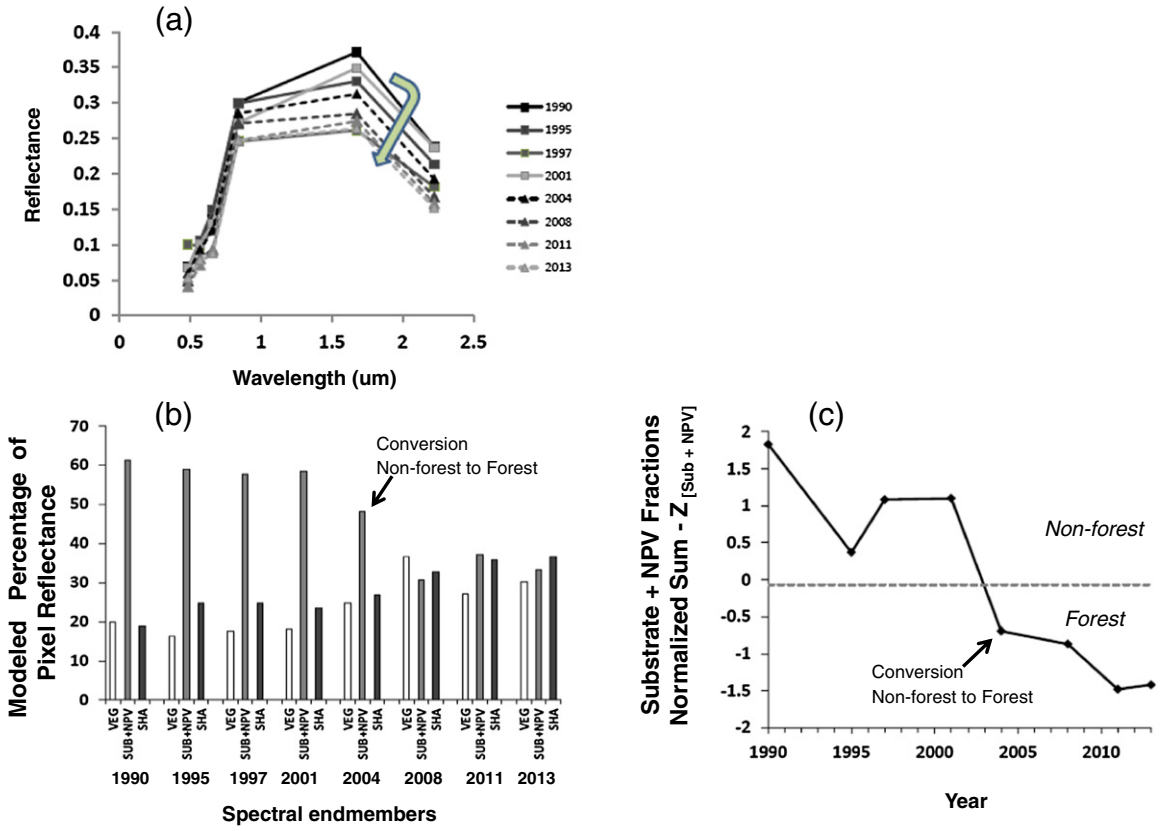

Fig. 7. Illustration of forest change detection (regrowth) using the LSMA-derived Z[Sub + NPV] thresholding metric. (a) Surface reflectance 1990–2013 of an example forest regrowth site. As forest regrows reflectance evolves in the direction of the green arrow; overall albedo decreases, and reflectance at the 1.65 μm and 2.22μm shortwave infrared broad-band centers (Landsat 5) decreases relative to the 0.83 μm band (near infrared). (b) LSMA-modeled endmember fractions (bars) for the surface reflectance spectra at the regrowth forest site, 1990–2013. Vegetation fraction increases while the summed substrate and NPV fractions decrease, with the largest changes occurring from 2001–2008. (c) Line plot of the Z [Sub + NPV] metric used as an indicator of forest land cover status. Change from non-forest–forest was assigned to year 2004 based on the inter-annual corrections (Table 2).

The SMA models across the June 2008 dry season (June, July, early and late August) were within reasonable values. The primary source of error in the model, the tradeoff of substrate and NPV fractions when either was at particularly high % cover, did not significantly impact differentiation of forest from non-forest. Non-forest areas had high overall percentages and normalized sums of substrate and NPV fractions, while forest areas had low normalized percentages of substrate and NPV fractions. The high accuracy of the multi-date, unsupervised classification showed from a "top-down," image-based perspective that differences in seasonal patterns of variability in substrate and NPV components between forest and non-forest land cover types held across the analysis region. ANOVA and Kruskal–Wallis analyses of Z(Sub + NPV) metrics at field validation sites provided "bottom-up," ground-truth statistical verification that differences in substrate and NPV surface components were stable between forest and non-forest land cover with no confounding effects of image observation. The low variations in forest areas mapped

Table 5
Tabora forest cover area estimates and rates of change from 1995–2011 in two periods (1995–2001, 2001–2011) reported in thousands of hectares (1 kha = 1000 ha). Baseline forest areas were quantified by analyses of forest cover using the Landsat LSMA fraction image-derived metric from a single image date (1995, or 2001). Averaged annual forest area changes were accounted by adding gross forest change areas (loss, gain) identified with LSMA metrics and inter-annual correction factors given in Table 2. Averaged (annualized) rates of forest loss and gain for 1995–2001 include losses and gains assessed in 1997 and 2001. Rates of loss and gain for 2001–2011 include loss areas assessed at 2004, 2008 and 2011, and gain areas assessed at 2004 and 2008.

Sub-districts (area kha)	Analysis period (Baseline — end)	Baseline forest cover area (kha)	Averaged annual forest area change (kha yr-1)		End forest cover area (kha)
			Loss	*Gain*	
Sikonge (112.2)	1995–2001	56.0	−1.51	0.45	49.7
Tabora (95.7)	1995–2001	25.6	−0.77	0.39	23.3
Urambo (35.7)	1995–2001	30.0	−0.56	0.03	26.8
Uyui (56.6)	1995–2001	29.6	−0.43	0.28	28.7
Uyui-MVP (50.5)	1995–2001	22.3	−0.48	0.35	21.6
TOTAL (350.7)	1995–2001	163.5	−3.75	1.45	150.0
Sikonge (112.2)	2001–2011	53.6	−1.48	0.57	44.5
Tabora (95.7)	2001–2011	24.9	−1.48	0.68	16.9
Urambo (35.7)	2001–2011	26.7	−0.47	0.34	25.4
Uyui (56.6)	2001–2011	28.9	−0.84	0.26	23.1
Uyui-MVP (50.5)	2001–2011	20.8	−0.60	0.37	18.5
TOTAL (350.7)	2001–2011	154.9	−2.92	1.33	139.0

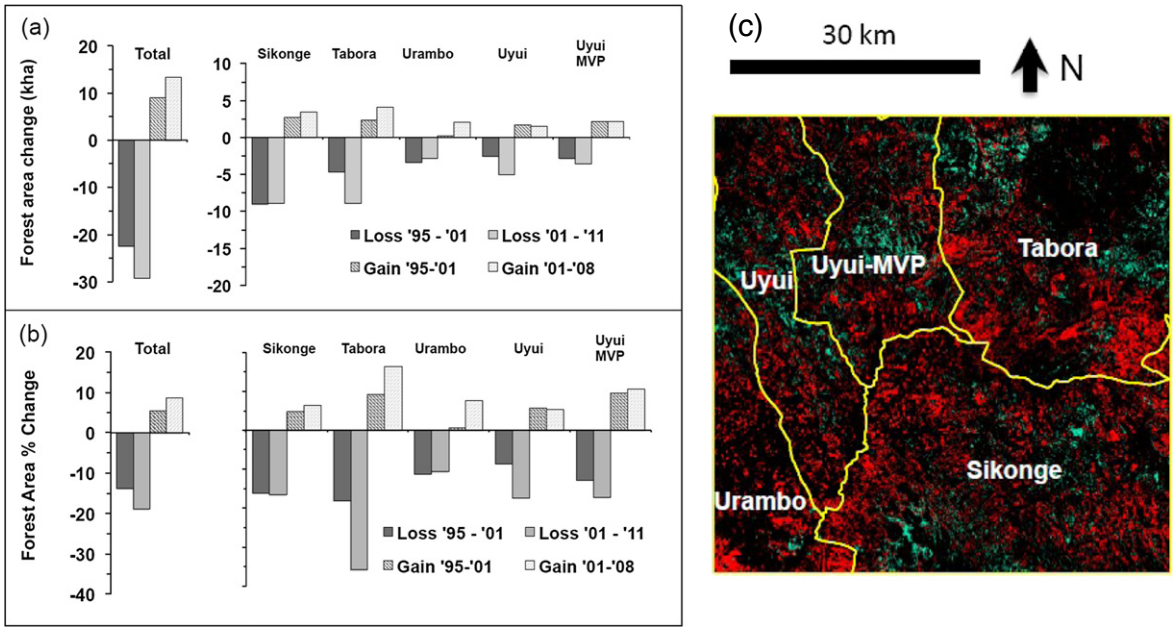

Fig. 8. Forest changes across the Tabora, Tanzania study region from 1995–2011. (a) Forest area changes shown in terms of area, thousands of hectares (kha). (b) Forest changes expressed as a percentage of forest area in the baseline year (starting year of the analysis period). (c) Map of forest changes 1995–2011 (combined). Red indicates regions of forest loss; cyan shows regions of forest gain. Sub-districts are labeled in white and bordered in yellow.

in June and July single-date analyses, and their high consumer accuracies (>90%) – comparable to the multi-date classification – showed that those dates are the most effective for mapping forest cover in this region with single images. The results support the hypothesis that early in the dry season at landscape scales, substrate and NPV fractions of non-forest areas should be high immediately following harvest (agriculture) or as a result of high grazing pressure on crop residues and grasses in mbuga regions. In contrast, substrate and NPV fractions would comprise lower proportions of forest areas, which retain green leaf cover further into the dry season before senescence progresses.

What was not apparent from Landsat LSMA analyses, or MODIS EVI observations were differences in the temporal variability of surface components between forest and non-forest. The ground-observed pattern of converging surface material proportions through the dry season was only observed for the NPV fraction images, and different rates of change in surface components, such as senescence leading to accelerating loss of vegetation components in forests through the dry season, were not apparent at Landsat scales. At MODIS scales, EVI greenness patterns were similar across land cover types. Our finding of similar patterns of temporal variability between forest and non-forest land cover components matches those of past research on Miombo region phenology using satellite data of coarse spatial/high temporal resolution (e.g. MODIS 250 m, 500 m or AVHRR), which show satellite-observed land cover phenology varies dominantly with climate (Serneels et al., 2007; Zhang et al., 2005). This result also echoes those of recent work that studied impacts of using 30 m Landsat versus 250 m MODIS data to map forest changes in a small-holder dominated region of Brazil, which found that up to 80% of forest change in areas smaller than 2 ha was missed by MODIS data (Toomey et al., 2013). In future work, Landsat analyses may be useful as masks to isolate land cover types such as forest and non-forest in MODIS data to study differences phenology, productivity or other ecosystem functions.

To differentiate forest from non-forest cover in the dry season across small-holder Miombo landscapes, the higher spatial resolution of Landsat imagery and the use of LSMA to identify patterns in non-green surface components offer an approach that is reliable and robust against variations in the status of land cover phenological cycles. Other LSMA approaches using multiple endmembers may be able to quantify surface components in such a way to distinguish land cover conditions over a wider seasonal timeframe where the landscape phenology is at different points (Quintano, Fernandez-Manso, & Roberts, 2013; Roberts et al., 1998). Our use of thresholds to classify forest versus non-forest areas discretely precluded study of forest land cover conditions in greater detail. There is ample room for further research to use LSMA fraction images as continuous metrics to evaluate forest and non-forest land cover conditions, beyond their use to map forest versus non-forest areas discreetly. For example, the shade fraction was difficult to relate straightforwardly to physical land cover features aside from identifying burned areas due to the different types of burning in forests (sub-canopy and whole stand burns), heterogeneous canopy structures (canopy cover < 30%) within non-burnt forests, and the variety of land use practices involving fire and resulting in burned areas on non-forest land cover types. These include burning of crop residues on fields and fires set to spur forage growth for livestock in mbuga grassland areas (Frost, 1996). Future work might enable use of LSMA fraction images to study variations in forest condition and structure relating to regrowth, land use pressures or management policies, as has been effective in regions such as coniferous forests in the Pacific Northwest of North America (Sabol et al., 2002; Yang et al., 2012). There is also an opportunity to validate burned areas and study the patterns of surface components they affect. Our approach for mapping and masking burned areas using the shade fraction images and normalized-burn index (NBR) was meant to be a conservative method for excluding regions whose land cover class was unclear at the time of observation, rather than a rigorously mapping burned areas themselves. The proportion of areas confounded by burns increased significantly through the dry season (from <2% in June–over 12% of the analysis area by the end of August). Thus for mapping land cover change with single-date analyses, use of later season images results in more ambiguous forest loss and regrowth estimates. For example, a regrowth forest that experiences burns in adjacent observational periods might not be designated regrowth because it is masked out as burn/transition. With our methods, the larger the burned area at a given time point, the more likely forest losses will be over-estimated or forest regrowth under-estimated. Future work to improve our methods could compare burned areas mapped with LSMA to global scale products such as the MODIS burned area product (MCD45A1).

4.2. Single-date forest cover maps approached the accuracy of multi-temporal classifications and comparisons quantify measurement uncertainties in land cover areas due to seasonal variability

Single-date forest cover maps using a constant threshold of the Z(Sub + NPV) metric performed comparably to multi-temporal classification in the early dry season. For change analyses, the differences in forest area mapped between sets of image dates such as June–July (4.7% of June forest cover) gives a quantitative estimate for an uncertainty in land cover mapping due to seasonal variability of surface components. However, many factors (e.g. variations in vegetation phenology, soil moisture) contribute to the differences in SMA model performance and fraction image models for given surface components and land cover classes. Our seasonal uncertainty of approximately 5% difference between forest areas does not identify which of such factors may contribute most to uncertainty in mapped forest areas across seasonal conditions within or between years. Greater consumer accuracies than producer accuracies for forest, and the opposite trend for non-forest areas, suggest that forest areas were slightly underestimated. Forest monitoring methods that yield conservative estimates of forest areas may be desirable for two reasons. First, conservative estimates of forest area align to a "precautionary principle" approach for estimating the status of available forest resources and habitat areas. Second, as potential inputs for coupled land–climate models, lower estimates of forest area could parameterize maximum-effect scenarios for studying impacts of forest change for climate through mechanisms such as reduced leaf-area and evapotranspiration, changes in albedo or surface roughness affecting land–atmosphere moisture exchange (Runyan, D'Odorico, & Lawrence, 2012).

4.3. Regional aggregate forest changes 2001–2011 are high relative to Tanzanian national data, and forest loss rates declined, while gain rates remained steady between decades

From 1995–2001 to 2001–2011, land cover transitioned in Tabora from 46.6% forested–39.6% forested. Forest losses were due to a mix of clearing for agriculture and tree harvesting for wood products for timber and wood-derived fuels (firewood, charcoal). Gains were due to both forest succession on abandoned cultivated plots, and recovery on sites cleared for tree harvests, but never cultivated. Our measured aggregate forest changes (27.4% for 2001–2011) are over twice as high as Tanzanian national aggregate changes (10.3%) (Hansen et al., 2013). Our findings of regional net forest losses during 1995–2001 comprising 8.6% of 1995 forest area were slightly lower, but comparable to those of Yanda (2010) who reported 11.2% forest losses for 1995–2000 in a larger region of analysis overlapping Tabora. Between decades we calculated a decrease in the net rate of forest conversion to non-forest from 1995–2001 (2.25 kha yr^{-1}) to 2001–2011 (1.59 kha yr^{-1}). This contrasts a coarse prediction by Yanda (2010) of unbounded, continuous linear increases in land clearing due to tobacco demands from the 2000s onward. We found that rates of gross forest loss were lower for 2001–2011 than 1995–2001. This drove an increased contribution of forest gain to overall forest changes, shown by the lower forest loss:gain ratio of 2.1 in 2001–2011 compared to 2.5 for the prior decade throughout the study area. Our results show that drivers affecting forest gains are interacting with pressures driving forest losses in Tabora and that both must be understood to estimate the trajectory of future forest area changes.

4.4. Variations in sub-district forest losses and gains suggest differing local drivers of forest change

Though it is beyond the scope of this work to attribute forest change patterns to drivers, the patterns in gross forest loss and gains we report serve as evidence for hypotheses for land-use drivers that are testable when the appropriate field data is available. For example, in Tabora sub-district, both forest loss and gain areas nearly doubled from 1995–2001 to 2001–2011. This pattern could relate to processes such as forest cutting for timber or charcoal without subsequent cultivation. Studies on Miombo forest succession after cutting without cultivation show that tree biomass and canopy cover can recover quickly (7–20 years) after disturbance when stems are cut, if other physical disturbances to roots or soil are minor (Chidumayo, 2013; Williams et al., 2008). Urambo's changes in forest loss and grain trajectories could be consistent with effects of a new local management plan since 2001 (after the Village Land Act of 1999) promoting forest conservation (Understanding Land Investment Deals in Africa, 2011). We can identify that a particular target region, Uyui-MVP, had multi-decadal forest change trends similar to the landscape at large. Future work to assess driving factors of forest change such as charcoal production, agricultural expansion or local management will need data collected at highly localized scales to explain these trends quantitatively, or, analyses will require up-scaling to spatial regions where reliable statistical data is available.

5. Conclusions

Complex landscape structure and the high spatio-temporal variability of phenology and land-use patterns have complicated identification of forest area and the study of land cover changes in seasonally dry tropical ecosystems worldwide. In a regional case study based in the Tanzanian Miombo Woodlands, we examined the hypothesis that non-forest sites would have higher proportions of non-green surface components than forest sites in the early to mid-dry seasons, due to differing processes affecting removal of green vegetation (harvest/grazing at non-forest sites versus senesce at forest sites). Using spectral mixture analyses of Landsat data, we found that the proportion of non-green surface components was significantly greater at non-forest than forest sites throughout the early to mid-dry season, substantiating our hypothesis. A multi-temporal classification of LSMA fraction images combining data from four image dates through a dry season, and threshold-based analyses using data from images at single dates in the same dry season, attained similar overall accuracies (86.5% multi-temporal, 81.0–85.3%) for mapping forest and non-forest land cover. The difference in mapped forest areas between single dates of imagery served as a "seasonal uncertainty" term for change detection analyses using single-date imagery — in our case the term was 5% of forest area between June and July images. Our methods may be applicable for other Miombo and dry tropical landscapes.

Our study region in Tabora showed net forest losses from 1995–2011 of 24.5 kha, representing 15% of 1995 forest area, a 7.0% reduction in forest land cover, and a transition from a state of dominantly forest to non-forest land cover. Rates of forest loss declined from 1995–2001 to 2001–2011 while forest gain rates were similar, leading to increases in forest gain areas relative to losses. Our change detection methods were sensitive to variable patterns in the balances of forest loss and gain areas by sub-district. This detection capability enables monitoring of forest changes at regional scales, but with spatial resolution to detect different localized trajectories of change where patterns may be indicative of varying land change pressures. Future work to improve satellite monitoring for dry tropical forest changes could examine use of spectral mixture analysis methods to study forest land cover conditions incorporate improved methods for characterizing land cover amidst high burn frequencies in the dry season. Quantitative satellite monitoring of land changes in remote dry tropical forest regions has high potential for improving understanding of drivers and ecosystem impacts of forest changes, and managing forest conservation efforts.

Author contributions

M. Mayes conceived the ideas and led the research project, including remote sensing method development, field data collection, analysis and

writing of the paper. J. Mustard advised the development of remote sensing methods and contributed to data analysis and writing. J. Melillo contributed to research ideas and manuscript composition. The authors have no conflicts of interest with regard to any project elements.

Acknowledgments

Thanks to Chris Neill, Cheryl Palm, Gerson Nyadzi, Kagya Eliezer and the Millennium Villages Project-Tabora staff for their help with introductions and logistics in Tabora. Thanks to Niwaeli Kimambo for her help in the field and with translations. This work was funded by the US National Science Foundation Partnerships for International Research and Education (PIRE) program, project title "Ecosystems and Human Well-Being" (Award # 0968211) PI Chris Neill.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2015.05.006.

References

Aber, J.D., & Melillo, J.M. (2001). *Terrestrial ecosystems* (2nd ed.). San Diego: Academic Press.

Adams, J.B., & Gillespie, A.R. (2006). *Remote sensing of landscapes with spectral images: A physical modeling approach*. Cambridge; New York: Cambridge University Press.

Adams, J.B., Sabol, D.E., Kapos, V., Almeida, R., Roberts, D.A., Smith, M.O., et al. (1995). Classification of multispectral images based on fractions of endmembers — Application to land-cover change in the Brazilian Amazon. *Remote Sensing of Environment*, *52*, 137–154.

Bodart, C., Brink, A.B., Donnay, F., Lupi, A., Mayaux, P., & Achard, F. (2013). Continental estimates of forest cover and forest cover changes in the dry ecosystems of Africa between 1990 and 2000. *Journal of Biogeography*, *40*, 1036–1047.

Cabral, A.I.R., Vasconcelos, M.J., Oom, D., & Sardinha, R. (2011). Spatial dynamics and quantification of deforestation in the central-plateau woodlands of Angola (1990–2009). *Applied Geography*, *31*, 1185–1193.

Campbell, B.M., Frost, P., & N. B. (1996). Miombo woodlands and their use: Overview and key issues. In B.M. Campbell (Ed.), *The Miombo in transition: Woodlands and welfare in Africa* (pp. 1–10). Bogor, Indonesia: Center for International Forestry Research (CIFOR).

Chidumayo, E.N. (2013). Forest degradation and recovery in a miombo woodland landscape in Zambia: 22 years of observations on permanent sample plots. *Forest Ecology and Management*, *291*, 154–161.

Chidumayo, E.N., & Gumbo, D.J. (2013). The environmental impacts of charcoal production in tropical ecosystems of the world: A synthesis. *Energy for Sustainable Development*, *17*, 86–94.

Dawelbait, M., & Morari, F. (2012). Monitoring desertification in a Savannah region in Sudan using Landsat images and spectral mixture analysis. *Journal of Arid Environments*, *80*, 45–55.

Eamus, D., & Prior, L. (2001). Ecophysiology of trees of seasonally dry tropics: Comparisons among phenologies. *Advances in Ecological Research*, *32*(32), 113–197.

Elmore, A.J., Mustard, J.F., & Manning, S.J. (2003). Regional patterns of plant community response to changes in water: Owens Valley, California. *Ecological Applications*, *13*, 443–460.

Elmore, A.J., Mustard, J.F., Manning, S.J., & Lobell, D.B. (2000). Quantifying vegetation change in semiarid environments: Precision and accuracy of spectral mixture analysis and the Normalized Difference Vegetation Index. *Remote Sensing of Environment*, *73*, 87–102.

FAO (2010). *Global forest resources assessment 2010: Main report*. Rome: Food and Agriculture Organization of the United Nations.

Frost, P. (1996). The ecology of Miombo Woodlands. In B.M. Campbell (Ed.), *The Miombo in transition: Woodlands and welfare in Africa* (pp. 11–57). Bogor, Indonesia: Center for International, Forestry Research (CIFOR).

Fuller, D.O. (1999). Canopy phenology of some mopane and miombo woodlands in eastern Zambia. *Global Ecology and Biogeography*, *8*, 199–209.

Furby, S.L., & Campbell, N.A. (2001). Calibrating images from different dates to 'like-value' digital counts. *Remote Sensing of Environment*, *77*, 186–196.

Geist, H.J., Chang, K.T., Etges, V., & Abdallah, J.M. (2009). Tobacco growers at the crossroads: Towards a comparison of diversification and ecosystem impacts. *Land Use Policy*, *26*, 1066–1079.

Grainger, A. (2008). Difficulties in tracking the long-term global trend in tropical forest area. *Proceedings of the National Academy of Sciences of the United States of America*, *105*, 818–823.

Green, J.M.H., Larrosa, C., Burgess, N.D., Balmford, A., Johnston, A., Mbilinyi, B.P., et al. (2013). Deforestation in an African biodiversity hotspot: Extent, variation and the effectiveness of protected areas. *Biological Conservation*, *164*, 62–72.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., et al. (2013). High-resolution global maps of 21st-century forest cover change. *Science*, *342*, 850–853.

Kjekshus, H. (1977). *Ecology control and economic development in East African history : The case of Tanganyika 1850–1950*. London: Heinemann.

Kutsch, W.L., Merbold, L., Ziegler, W., Mukelabai, M.M., Muchinda, M., Kolle, O., et al. (2011). The charcoal trap: Miombo forests and the energy needs of people. *Carbon Balance and Management*, *6*, 5.

Lambin, E.F., Geist, H.J., & Lepers, E. (2003). Dynamics of land-use and land-cover change in tropical regions. *Annual Review of Environment and Resources*, *28*, 205–241.

Lambin, E.F., & Meyfroidt, P. (2011). Global land use change, economic globalization, and the looming land scarcity. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 3465–3472.

Lepers, E., Lambin, E.F., Janetos, A.C., DeFries, R., Achard, F., Ramankutty, N., et al. (2005). A synthesis of information on rapid land-cover change for the period 1981–2000. *Bioscience*, *55*, 115–124.

Palm, C.A., Smukler, S.M., Sullivan, C.C., Mutuo, P.K., Nyadzi, G.I., & Walsh, M.G. (2010). Identifying potential synergies and trade-offs for meeting food security and climate change objectives in sub-Saharan Africa. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 19661–19666.

Prins, E., & Kikula, I.S. (1996). Deforestation and regrowth phenology in miombo woodland — Assessed by Landsat Multispectral Scanner System data. *Forest Ecology and Management*, *84*, 263–266.

Quintano, C., Fernandez-Manso, A., & Roberts, D.A. (2013). Multiple Endmember Spectral Mixture Analysis (MESMA) to map burn severity levels from Landsat images in Mediterranean countries. *Remote Sensing of Environment*, *136*, 76–88.

Roberts, D.A., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R.O. (1998). Mapping chaparral in the Santa Monica Mountains using multiple endmember spectral mixture models. *Remote Sensing of Environment*, *65*, 267–279.

Runyan, C.W., D'Odorico, P., & Lawrence, D. (2012). Physical and biological feedbacks of deforestation. *Reviews of Geophysics*, *50*.

Sabol, D.E., Gillespie, A.R., Adams, J.B., Smith, M.O., & Tucker, C.J. (2002). Structural stage in Pacific Northwest forests estimated using simple mixing models of multispectral images. *Remote Sensing of Environment*, *80*, 1–16.

Sedano, F., Gong, P., & Ferrao, M. (2005). Land cover assessment with MODIS imagery in southern African Miombo ecosystems. *Remote Sensing of Environment*, *98*, 429–441.

Serneels, S., Linderman, M., & Lambin, E.F. (2007). A multilevel analysis of the impact of land use on interannual land-cover change in East Africa. *Ecosystems*, *10*, 402–418.

Small, C., & Milesi, C. (2013). Multi-scale standardized spectral mixture models. *Remote Sensing of Environment*, *136*, 442–454.

Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., & Macomber, S.A. (2001). Classification and change detection using Landsat TM data: When and how to correct atmospheric effects? *Remote Sensing of Environment*, *75*, 230–244.

Toomey, M., Roberts, D.A., Caviglia-Harris, J., Cochrane, M.A., Dewes, C.F., Harris, D., et al. (2013). Long-term, high-spatial resolution carbon balance monitoring of the Amazonian frontier: Predisturbance and postdisturbance carbon emissions and uptake. *Journal of Geophysical Research—Biogeosciences*, *118*, 400–411.

Understanding Land Investment Deals in Africa (2011). Country Report: Tanzania. In F. Mousseau, & A. Mittal (Eds.), Oakland, California USA: The Oakland Institute.

van Wagtendonk, J.W., Root, R.R., & Key, C.H. (2004). Comparison of AVIRIS and Landsat ETM+ detection capabilities for burn severity. *Remote Sensing of Environment*, *92*, 397–408.

Williams, M., Ryan, C.M., Rees, R.M., Sarnbane, E., Femando, J., & Grace, J. (2008). Carbon sequestration and biodiversity of re-growing Miombo woodlands in Mozambique. *Forest Ecology and Management*, *254*, 145–155.

Yanda, P.Z. (2010). Impact of small scale tobacco growing on the spatial and temporal distribution of Miombo woodlands in Western Tanzania. *Journal of Ecology and the Natural Environment*, *2*, 010–016.

Yang, J., Weisberg, P.J., & Bristow, N.A. (2012). Landsat remote sensing approaches for monitoring long-term tree cover dynamics in semi-arid woodlands: Comparison of vegetation indices and spectral mixture analysis. *Remote Sensing of Environment*, *119*, 62–71.

Zhang, X.Y., Friedl, M.A., Schaaf, C.B., Strahler, A.H., & Liu, Z. (2005). Monitoring the response of vegetation phenology to precipitation in Africa by coupling MODIS and TRMM instruments. *Journal of Geophysical Research—Atmospheres*, *110*.