Science of Remote Sensing 8 (2023) 100096

Contents lists available at ScienceDirect

Science of Remote Sensing

journal homepage: www.sciencedirect.com/journal/science-of-remote-sensing

Using Landsat and Sentinel-2 spectral time series to detect East African small woodlots

Niwaeli E. Kimambo [a,b,c,*], Volker C. Radeloff [b]

[a] Department of Geography, Middlebury College, McCardell Bicentennial Hall, 287 Bicentennial Way, Middlebury, VT, 05753, USA
[b] SILVIS Lab, Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, 1630 Linden Drive, Madison, WI, 53706, USA
[c] Department of Geography, University of Wisconsin-Madison, Science Hall, 550 N Park St, Madison, WI, 53706, USA

ARTICLE INFO

Keywords:
Plantations
Google earth engine
Tree cover map
Trend analysis
Tanzania
Timber rush
Temporal signature
Tree planting
LandTrendr

ABSTRACT

Accurate maps of gains in tree cover are necessary to quantify carbon storage, wildlife habitat, and land use changes. Satellite-based mapping of emerging smallholder woodlots in heterogeneous landscapes of sub-Saharan Africa is challenging. Our goal was to evaluate the use of time series to detect and map small woodlots (<1 ha) in Tanzania. We distinguished woodlots from other land cover types by woodlots' distinct multi-year spectral time series. Woodlots exhibit greening from planting to maturity followed by browning at harvest. We compared two time series approaches: 1) a linear model of Tasseled Cap Wetness (TCW) and other indices, and 2) LandTrendr temporal segmentation metrics. The approaches had equivalent woodlot detection accuracy, but LandTrendr segments had lower accuracy for characterizing woodlot age. We tested the effect of the following factors on woodlot detection and mapping accuracy: the length of the time series (2009–2019), frequency of observations (all Landsat vs. only Landsat-8), spatial resolution (30-m Landsat vs. 10-m Sentinel-2), and woodlot age and size. Woodlot mapping accuracies were higher with longer time series (54% at 3-yrs vs 77% at 7-yrs). The accuracies also improved with more observations, especially when the time series was short (3-yrs Landsat-8 only: 54% vs. all-Landsat: 64%, p-value <0.001). Sentinel-2's higher spatial resolution minimized commission errors even for short time series. Finally, less than half of young and small (<0.4 ha) woodlots were detected, suggesting considerable omission errors in our and other woodlot maps. Our results suggest that the accurate detection of woodlots is possible by analyzing multi-year time series of Landsat and Sentinel-2 data. Given the region's woodlot boom, accurate maps are needed to better quantify woodlots' contribution to carbon sequestration, livelihoods enhancement, and landscape management.

1. Introduction

Mapping tree cover in different kinds of landscapes is crucial for understanding the climatic and economic benefits of trees (Crowther et al., 2015; Hansen et al., 2013), and remote sensing is an essential tool for such mapping. While mapping forests versus non-forests is generally highly accurate, separating native forests from forest plantations is more difficult, especially when plantations are small (Sexton et al., 2016). Accordingly, while the Food and Agriculture Organization estimates that there are 278 million hectares of tree plantations worldwide, it also notes high uncertainty of this estimate because of difficulties in quantifying smallholder woodlots (Payn et al., 2015). That is problematic because smallholder woodlots in Africa and southeast Asia contribute substantially to those regions' timber needs (Jacovelli, 2014; Mather, 2007; Rudel, 2009), and they are becoming more common (Payn et al., 2015; Kimambo et al., 2020). In addition to timber supply, smallholder woodlots are also being promoted as a valuable carbon mitigation and landscape restoration strategy, even as the carbon stocks associated with woodlots remain poorly quantified (Fagan et al., 2020; Veldman et al., 2015; Dave et al., 2017). Thus, suitable remote sensing approaches are needed to accurately map smallholder woodlots.

Whereas natural forests are declining globally, the extent of woodlots has been increasing, especially in developing countries (e.g., in Vietnam (Nawir et al., 2007), India (Mather, 2007), Indonesia (Torbick et al., 2016), Uganda (L'Roe and Naughton-Treves, 2016), and Ethiopia (Jenbere et al., 2012)). We define a woodlot as smallholder-planted trees, typically on their own private land and in a small area (<5 ha), to be used for firewood, timber, or fruit (Ngaga, 2011; Kimambo and

* Corresponding author. Department of Geography, McCardell Bicentennial Hall, 287 Bicentennial Way, Middlebury College, Middlebury, VT, 05753, USA.
E-mail address: nkimambo@middlebury.edu (N.E. Kimambo).

https://doi.org/10.1016/j.srs.2023.100096
Received 15 March 2023; Received in revised form 3 August 2023; Accepted 3 August 2023
Available online 5 August 2023

Naughton-Treves, 2019). The increase in woodlots in East Africa is partly due to increased demand for tree products like timber and fuel wood due to rapid urbanization and population growth (Held et al., 2017; Indufor, 2011; Jacovelli, 2009). In some countries, including in India, Vietnam, and China, smallholder woodlots have increased partly due to government promotion and subsidies (Nawir et al., 2007; Borah et al., 2018; Frayer et al., 2014; Arvola et al., 2020). Despite strong evidence for the increasing contribution of smallholder woodlots to tree cover gain trends, there is a lack of systematic methods to detect woodlots and quantify their extent for both local and global level statistics (Torbick et al., 2016; Kröger, 2012).

Woodlot expansion has been especially rapid in East Africa after 2010 (Jacovelli, 2014), coincident with a period of increased image availability from medium-resolution sensors such as Landsat and Sentinel-2. However, three factors make woodlot detection and mapping from satellite images challenging: 1) woodlots are established asynchronously – smallholders plant their woodlots in different years and manage them differently, which means that woodlots as a land cover class are highly heterogeneous; 2) woodlots are small – the majority are <1 ha, and their average size is 0.45 ha (Kimambo et al., 2020); and 3) woodlots are part of heterogeneous landscapes and are often planted next to annual crops, shrubland, and native forests.

The first challenge for woodlot detection and mapping – asynchronous establishment and variability – is both a challenge and a potential methodological advantage because woodlots' spectral time series and visual characteristics may separate woodlots from other land covers (Deng et al., 2020). Spectral time series have been used to distinguish land cover that have intra-annual variability (e.g., rice paddies (Kontgis et al., 2015) and maize (Lobell et al., 2003), or rubber plantations (Hurni et al., 2017)). These studies relied on an annual and synchronous time series signature. Woodlots' spectral trajectory, however, is multi-year, and each woodlot exhibits an increase in greenness over 2–7 years after planting (Fig. 1). A multi-year spectral time series can identify a woodlot's greening trend (Deng et al., 2020; Gao et al., 2016), but not distinguish it from shrub encroachment or forest regeneration. Furthermore, woodlots will have a wide range of time series trajectories due to asynchronous establishment. Therefore, woodlot's spectral time series may require additional data such as visual characteristics (Kimambo et al., 2020) in order to accurately distinguish woodlots from other greening trends.

To chart any multi-date spectral time series, multispectral reflectance for each date is typically reduced to an index value, such as the

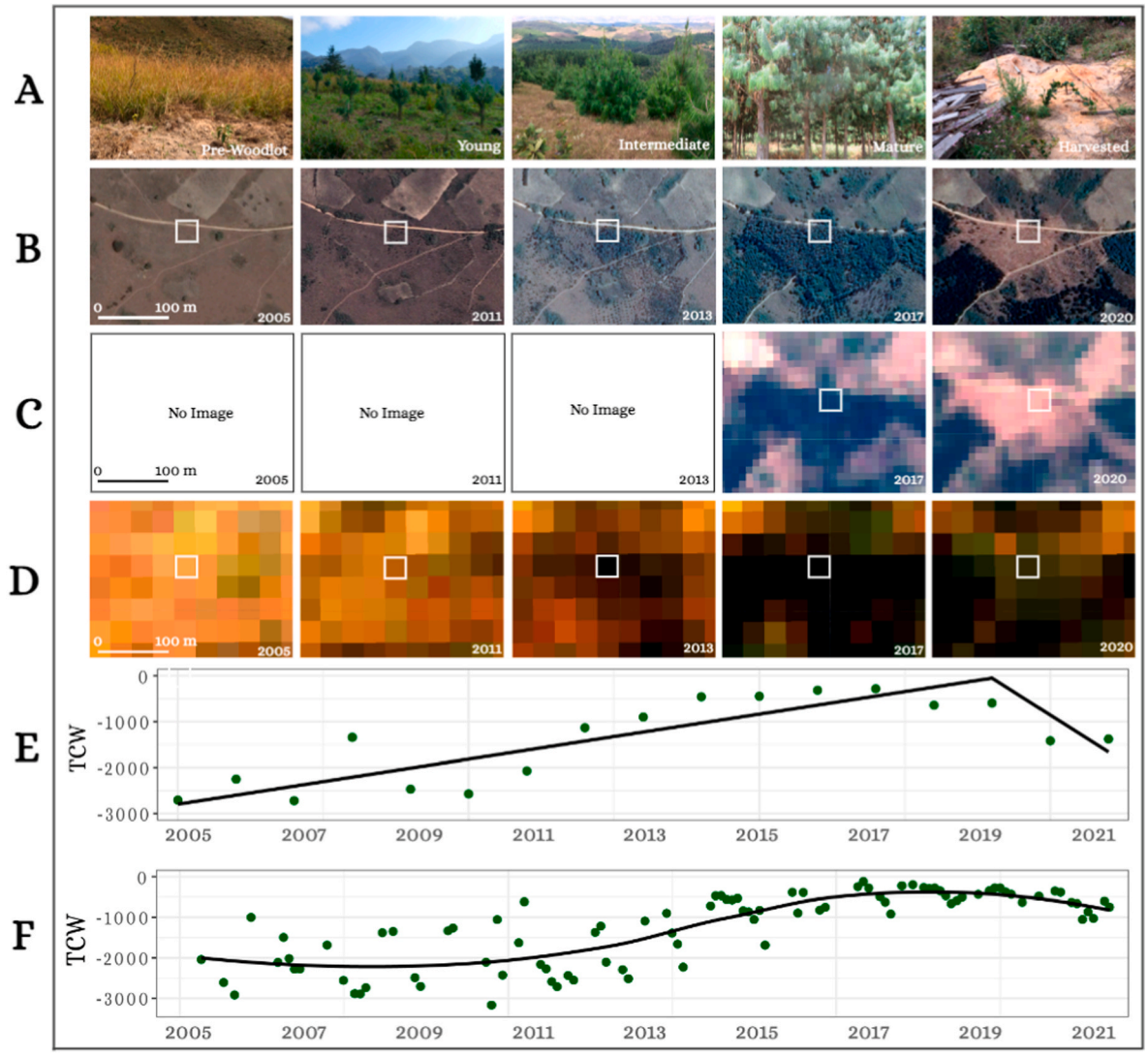

Fig. 1. Woodlot growth as seen A) on the ground, B) in high-resolution (1–3 m) Google Earth Pro (CNES/Airbus) images, C) in Sentinel-2 images (RGB/B4, B3, B2), D) in Landsat Images (RGB/B4,B3,B2) E) in temporal segments generated by LandTrendr, with fitted lines for different LandTrendr segments and F) in all the Tasseled Cap Wetness values for all available images (2005–2020), with a fitted trendline generated from Locally Estimated Scatterplot Smoothing (*LOESS(TCW ~ date)*). Pictures (Row A) are illustrative and do not represent the same location. Rows B – F are the same location (35.319463°E, 8.522063°S).

Normalized Difference Vegetation Index (NDVI) or the Enhanced Vegetation Index (EVI) (Deng et al., 2020; Hurni et al., 2017; Qiao et al., 2016). Tasseled Cap Transformation converts multi-band reflectance into three components (brightness, greenness, and wetness), thereby reducing noise from senesced vegetation and soil characteristics (Sonnenschein et al., 2011). Comparisons of NDVI and the Tasseled Cap Transformation show the later to be less susceptible to noise and the components represent biophysically interpretable aspects of land cover (Dymond et al., 2002; Healey et al., 2005). A woodlot could be captured with a positive slope or cumulative value of Tasseled Cap greenness or wetness (Fig. 1). The magnitude of the values and may allow distinguishing woodlots that are at different points in their establishment trajectories but may fail to capture a woodlot that has experienced a harvest cycle.

Landsat-based Detection of Trends in Disturbance and Recovery (LandTrendr) algorithm performs temporal segmentation of spectral time series data that could potentially identify a woodlot's growth and harvest cycle (Kennedy et al., 2010). LandTrendr identifies periods of stability versus periods of change based on seasonal or annual satellite observations. LandTrendr was originally designed to map forest disturbance and recovery, and a spectral time series of woodlot establishment and subsequent harvest presents an analogous temporal pattern. Compared to related time-series analysis algorithms such as Continuous Change Detection Classification (CCDC) or Breaks for Additive Seasons and Trend (BFAST), LandTrendr requires fewer images in a given year (Zhu, 2017) and has been tested for disturbance and recovery mapping in a wide range of ecoregions (Kennedy et al., 2018). However, LandTrendr has only be applied once in the smallholder systems of sub-Saharan Africa (Schneibel et al., 2017). Detecting small woodlots would be an excellent test case for LandTrendr's applicability, given East Africa's limited image availability (Goward et al., 2006), heterogeneous landscapes, variable field sizes, and flexible cropping systems (Xu et al., 2018). The algorithm can be useful for characterizing woodlots at different points in their establishment trajectory, including harvest and replanting cycles.

The second challenge for woodlot detection and mapping – small size of smallholder woodlots – may limit ability to separate woodlots from adjacent land cover with similar spectral signature. Spectral similarity may differ at different points in woodlot's establishment trajectory (Fig. 1). For example, a newly planted woodlot may resemble annual cropland or grassland, a mature woodlot may resemble woody vegetation, and a newly harvested woodlot may resemble bare ground. This challenge may be overcome by higher spatial resolution of the input images. The availability of ESA's Sentinel-2 10-m images provides 9 times more pixels per woodlot compared to Landsat, but Sentinel-2 has a shorter time series Fig. 1C). In single-date analyses, including Sentinel-2 imagery improved woodlot mapping accuracies in Tanzania (Koskinen et al., 2019) and in south Asia (Hurni et al., 2017; Nomura and Mitchard, 2018; Torbick et al., 2016). The utility of Sentinel-2's higher spatial resolution in woodlot mapping needs to be evaluated while controlling for time series length.

The third challenge for woodlot detection and mapping – limited class separability due to within-class and landscape heterogeneity – stems from the asynchronous establishment and the small field sizes. Ultimately, accurate mapping of any land cover type depends on whether it is separable from the surrounding context (Lu and Weng, 2007; Ozdogan and Woodcock, 2006). Spectral distinctness of a woodlot might depend on its characteristics such as age and size: mature and large woodlots would be more spectrally distinct (Fassnacht et al., 2016; Grabska et al., 2019; Sheeren et al., 2016). One-time woodlot maps rarely explicate the heterogeneity of woodlots as a landcover by indicating characteristics like woodlot age (Koskinen et al., 2019). Without sub-characterization, the maps could be favoring spectral signals from large and mature woodlots. Analyses are needed that determine whether woodlot detection and mapping are sensitive to woodlot characteristics (e.g., size and age).

Our overall goal was to develop and test two time series-based methods for mapping smallholder woodlots. Our first objective was to quantify woodlot mapping accuracy given: a) length of the satellite images time series; b) frequency of observations; and c) spatial resolution of the images. Our second objective was to compare woodlot detection accuracy given: a) time series described by a linear model of tasseled cap wetness index versus LandTrendr temporal segments; and b) woodlot's age and size characteristics.

2. Methods

2.1. Study area

To develop and test time series approach for woodlot detection, we performed our study on an area of 10,000 km^2 in the Southern Highlands of Tanzania, centered on 8.9°S, 34.7°E (Fig. 2). The Southern Highland (ca. 202,770 km^2) is a high-elevation (avg. elevation: 1800 m) region with high rainfall (up to 2000 mm per year at 2000 m asl) and moderate temperatures (Fick and Hijmans, 2017). These climatic conditions are well-suited for tree plantations, which is why the region produces most of Tanzania's timber and contains 15 government and privately-owned large timber plantations (Kimambo et al., 2020; Koskinen et al., 2019). The region is experiencing a substantial smallholder tree planting boom since c. 2010 (Kimambo et al., 2020; Ngaga, 2011; Indufor, 2011) due to market demand for construction timber, electric poles, firewood, and other tree products (Koskinen et al., 2019; Arvola et al., 2019). The smallholder woodlots average 0.45 ha in size and are of variable age, with the majority established between 2012 and 2015 (Kimambo et al., 2020). These smallholder woodlots typically contain fast-growing species of pine and eucalyptus and have very short rotation cycles (7–10 years) because smallholders harvest the woodlots early to meet cash needs (FDT, 2015). Woodlots are interspersed with other smallholder land uses, including rain-fed annual crops, tea, and pasture. Other land cover types in the study area include natural grasslands, wet montane forests, miombo woodlands, shrublands, and urban areas.

2.2. Analysis

We conducted most of our analysis for the two objectives (Fig. 3) in Google Earth Engine (Gorelick et al., 2017). For each objective, we first prepared the input image stacks, the training data, and independent validation data (see 2.2.1). For Objective 1, we assessed woodlot mapping accuracy given imagery input parameters (see 2.2.2 and Fig. 3). For Objective 2, we compared woodlot detection success between our mapping approach from Objective 1 and LandTrendr temporal segmentation (see 2.2.3 and Fig. 3). Finally, we analyzed woodlot detection patterns given woodlot size and age characteristics.

2.2.1. Data

2.2.1.1. Input image stacks. We created input image stacks with varied time series length, frequency of observations, and spatial resolution (Table 1). The image stacks were based on Landsat and Sentinel-2 data. For Landsat data, we analyzed all the available Landsat scenes from Google Earth Engine's Landsat Surface Reflectance Tier 1 data (2010–2019, accessed January 2021). We used the CFMask-generated pixel quality band to remove clouds and cloud shadows. We harmonized Landsat-8 (OLI) to Landsat-5 (TM) and Landsat-7(ETM+) following Roy et al., 2016. For Sentinel-2 data, we downloaded Sentinel-2A and 2B MSI (Multi-Spectral Instrument) Level 1C data from Sentinel's scihub for 2017–2019. Our study area corresponds to tile T36LYR. We used Sen2Cor (version 2.9) to process the images to Level 2A, including atmospheric correction and cloud detection, and uploaded the corrected imagery to Google Earth Engine for subsequent analysis. We masked clouds using pixel quality information produced by Sen2Cor

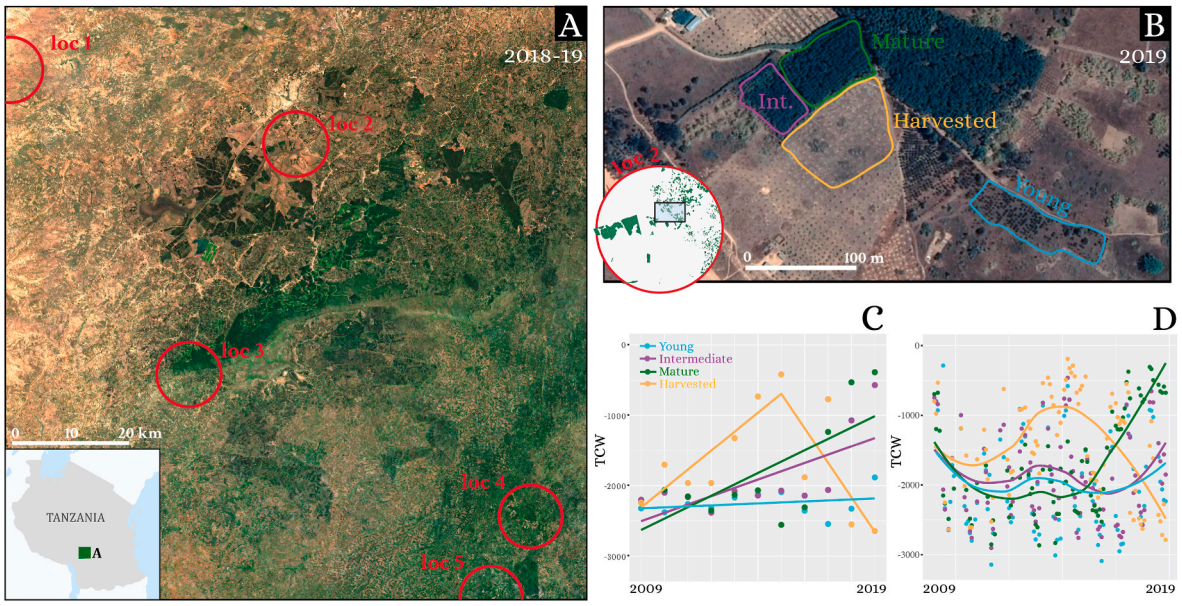

Fig. 2. A) The study area extent. Circles labeled loc 1 – loc 5 are manual digitization locations. The underlying image is a 1-year (Jan–Dec 2019) Sentinel-2 true-color median composite (RGB: B4, B3, B2). Context map shows Tanzania and the study area extent. B) Sample woodlots that we manually digitized at location 2 in Google Earth Pro (CNES/Airbus), with four woodlot classes: "Young" (aqua), "Intermediate" (purple), "Mature" (green) and "Harvested" (tan). C) LandTrendr segments (2009–2019) for each of the woodlots shown in panel B. The fitted lines are LandTrendr-generated segments D) Tasseled Cap Wetness (TCW) spectral time series (2009–2019) for the four woodlots. The fitted trendline is a Locally Estimated Scatterplot Smoothing (*LOESS(TCW ~ date)*). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

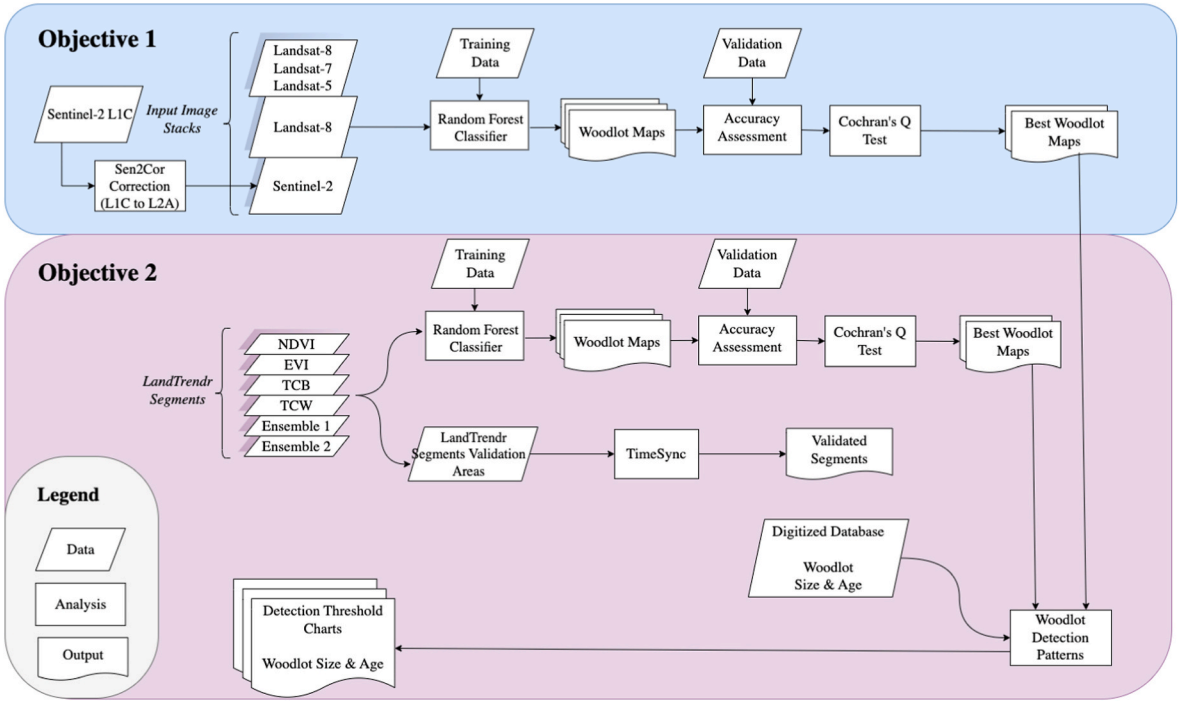

Fig. 3. Summary of analysis steps for woodlot mapping and detection. Shape of text boxes indicates input data, analysis steps, and intermediate and final outputs (see legend). The Data are fully described in Section 2.2.1. The Analysis are described under each objective. In Objective 1 (section 2.2.2) we performed woodlot mapping under different time series constraints. In Objective 2 (section 2.2.3), we analyzed woodlot detection patterns given time series analysis approach and woodlot characteristics.

(Appendix A).

To summarize the reflectance data of Landsat images, we calculated four indices for each image: the Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI) ((B3 – B5)/(B3 + B5)), Bare Soil Index (BSI) ((B7+B4) –(B5 +B2))/((B7 +B4) +(B5+B2)), and the wetness index from the Tasseled Cap Transformation (TCW). For Sentinel-2, we selected the 10-m bands (B2, B3, B4, B5, and B8) and a 20-m band (B11) and resampled B11 to 10-m to match the spatial resolution of the other bands. We calculated the same four indices as Landsat. To capture spectral time series that could distinguish woodlots

Table 1
The characteristics of input image stack used to map woodlots. The input image stacks varied by the length of the time series, observation density, and spatial resolution.

	Calendar Range	Time series length (years)	Spatial Resolution (m)	Monthly TCW averages (bands)	Slope + Intercept (bands)				Total no. of Bands
					NDVI	BSI	NDWI	TCW	
Landsat-8, 7, & 5[a]	2010–2019	10	30	113	2	2	2	2	121
	2011–2019	9	30	108	2	2	2	2	116
	2012–2019	8	30	96	2	2	2	2	104
	2013–2019	7	30	84	2	2	2	2	92
	2014–2019	6	30	72	2	2	2	2	80
	2015–2019	5	30	60	2	2	2	2	68
	2016–2019	4	30	48	2	2	2	2	56
	2017–2019	3	30	36	2	2	2	2	44
Landsat-8	2013–2019	7	30	80	2	2	2	2	89
	2014–2019	6	30	72	2	2	2	2	80
	2015–2019	5	30	60	2	2	2	2	68
	2016–2019	4	30	48	2	2	2	2	56
	2017–2019	3	30	36	2	2	2	2	44
Sentinel-2	2016–2019	4	10	48	2	2	2	2	56
	2017–2019	3	10	36	2	2	2	2	44

Abbreviations: NDVI: Normalized Difference Vegetation Index; BSI: Bare Soil Index; TCW: Tasseled Cap Wetness.
[a] Images for Landsat-5 only available up to year 2013, Landsat-8 starting in 2013.

from other land cover classes, we calculated the slope and intercept of a linear model for each pixel using the four indices and varying the time series length (Table 1). We also calculated the average monthly TCW for a given time period.

We tested the effects of the length of the time series, frequency of observation, and spatial resolution on woodlot detection accuracy. Our shortest time series was three years (2017–2019) and then we increased the length in reverse chronology by one-year increments. To examine frequency of observation effect, we created five image stacks from Landsat-8 only (2013–2019) and eight from Landsat-8, Landsat-5, and Landsat-7 combined (2010–2019). We also created two image stacks based on Sentinel-2 images (2017–2019) to test the advantages of 10-m resolution versus Landsat's 30-m resolution (Table 1).

2.2.1.2. LandTrendr temporal segments. We used the Google Earth Engine implementation of the LandTrendr algorithm to generate temporal segments (Kennedy et al., 2018). The input images were from Landsat and atmospherically corrected, harmonized, and masked for clouds with procedures outlined above. The images covered a ten-year time window (2010–2019), all available months (01-Jan – 31-Dec) and used LandTrendr's default medoid function to create a single image per year. For a given pixel and image index, this medoid implementation picks a pixel value that is numerically closest (in Euclidian spectral distance) to the median of all the pixels under consideration, calculated annually (Kennedy et al., 2018). By creating a median image for each year, consistency among images, which is necessary for temporal segmentation, is ensured.

In LandTrendr terms, woodlot establishment is equivalent to fast tree cover gain (Figs. 1, 2), so we parameterized LandTrendr to optimally capture this type of change. We followed Kennedy et al. (2010) guidelines for parameterization to better identify timing of woodlot establishment: setting higher maximum segments (*max_segments* = 6) and turning off the recovery threshold (*recovery_threshold* = 1.0). We did not apply a minimum mapping unit because woodlots can be very small. For the other ten LandTrendr parameters, we conducted a sensitivity analysis following Rodman et al. (2021). The final set of parameters maximized detection of rapid tree cover gain consistent with how woodlots establish and minimized sensitivity to within-year variability (Appendix B). We generated segments from four indices (NDVI, TCW, TCB, and EVI, (Table 2)) because they have been shown to be the best-performing indices in LandTrendr (Cohen et al., 2018), and compared their woodlot detection accuracies.

2.2.1.3. Training and validation data. Our training sample generation emphasized distinguishing woodlots from other classes based on both time series (Deng et al., 2020) and visual characteristics (Kimambo et al., 2020; Koskinen et al., 2019). We selected 778 samples across 12 land cover classes ("Cropland", "Forest", "Grassland", "Tea", "Urban", "Wetland", "Woodland", "Water", "Intermediate", "Mature", "Young", and "Harvested"). To label samples, we interpreted a sample's time series via a publicly available time series viewing tool based on Landsat and Sentinel-2 images (Yin, 2019)(Appendix C). Concomitantly, we visually interpreted high-resolution Google Earth Pro images (Appendix C). If no Google Earth Pro image was available, we checked Sentinel-2 10-m images from 2019. For the four woodlot classes ("Harvested", "Young", "Intermediate", and "Mature"), we recorded woodlot age by approximating planting date visually (Fig. 2 and next section for details) and by inspecting the 2010–2019 time series. In the time series, "Harvested" woodlots had a greening-then-loss signal with no re-greening. "Young" woodlots were planted within four years of target map date (2019), so the time series showed a greening trend after year 2015. "Intermediate" woodlots were planted four to six years prior, and their time series greening began between years 2013 and 2015. "Mature" woodlots were more than six years old, and their greening began prior to

Table 2
The characteristics of LandTrendr segments used to detect woodlots. The segments were of the same time series length but from different indices.

Index[a]	Calendar Range	Time series length (years)	Spatial Resolution (m)	Landtrendr segmentation output (bands)						Total no. of Bands
				Magnitude	DSNR	Year-of-detection	Segment Duration	pre-value	gain rate	
NDVI	2010–2019	10	30	1	1	1	1	1	1	6
TCW	2010–2019	10	30	1	1	1	1	1	1	6
EVI	2010–2019	10	30	1	1	1	1	1	1	6
TCG	2010–2019	10	30	1	1	1	1	1	1	6

[a] Abbreviations: NDVI: Normalized Difference Vegetation Index; TCW: Tasseled Cap Wetness, EVI Enhanced Vegetation Index; TCG: Tasseled Cap Greenness; DSNR: Disturbance Signal to Noise Ratio.

year 2012. The final training dataset contained 326 samples for the four woodlot classes and 452 samples for the other seven classes.

For validation, we aggregated the classified map into five classes ("Other", "Harvested", "Young", "Intermediate", and "Mature"). We performed a stratified random sample with the Area Estimation & Accuracy Assessment toolbox (Bullock et al., 2019), following Olofsson et al. (2014) accuracy assessment recommendations. We randomly generated 415 validation samples by balancing sample distribution across strata and maintaining a target standard error on the overall accuracy of 0.02. The final validation sample distribution was: "Other (118)", "Harvested (48)", "Young (90)", "Intermediate (80)", and "Mature (79)".

2.2.1.4. Independent database on woodlot size and age data from manual digitization. We required independent information on woodlots' characteristics in order to analyze woodlot detection patterns. We manually digitized woodlots to generate independent woodlot characteristics, namely woodlots' age and size. We randomly selected five subregions that were each 10,000 ha (100 km^2), they covered 5% of the study area (Fig. 2). In each subregion we manually digitized all woodlots based on visual interpretation in the most recent (year 2018/2019) high-resolution (1-m) Google Earth Pro images (Fig. 1B). The availability of up-to-date high-resolution Google Earth images varied by subregion. Some subregions' most recent image dated back to the early 2000s. Thus, we recorded the date the image was acquired and the age category for the woodlot. Each woodlot was delineated following borders such as fire breaks and farm boundaries, while ensuring uniform texture within the woodlot (See Fig. 2 for examples, and (Kimambo et al., 2020) for details). We assigned each the woodlot an age category of either: "Young", "Intermediate", "Mature" or "Harvested" based on its tree density and crown size. The "Young" category represents woodlots where tree planting is evident, but trees are very sparse with the bare ground visible between tree rows. The "Intermediate" category are woodlots with round tree crowns that are almost touching, but the tree rows are still separated. The "Mature" category are woodlots where the tree canopy is dense and fully closed, and bare ground and planting rows are no longer visible.

2.2.2. Woodlot mapping

To map woodlots, we performed supervised classification with random forest because woodlots are a heterogeneous land cover. Our input image stacks differed in a) the length of the time series; b) frequency of observations; and c) image resolution. We performed parameter tuning for random forest (Breiman, 2001) and tested sensitivity of the classifier to the size of the training sample and input image stack (Appendix C). We used 500 trees, the square root of the number of variables at each split and set no minimum leaf population. Each input image stack was classified with the same training points into 12 land cover classes (Fig. 3). We then simplified the classification output from 12 classes to two ("Other" and "Woodlot") and used the independent validation data to assess overall accuracy in separating woodlots from all other land covers. Subsequently, we repeated the accuracy assessment with five classes ("Other", "Harvested", "Young", "Intermediate", and "Mature") and determined mapping accuracy for each woodlot age class. For each output map, we estimated overall accuracy as well as commission and omission errors for each class.

To determine which input image characteristics best mapped woodlots, we tested whether any of the accuracies were significantly different. We applied Crochan's Q test, an omnibus test that extends the McNemar's test for more than two comparisons (Foody, 2004). Crochan's Q test indicated if any of the 15 output maps were different from the rest but did not identify which one. For this, we applied a post-hoc pairwise Wilcoxon test. We selected the three best woodlot mapping outputs based on the overall accuracies and the p-values of the pairwise Wilcoxon test. We compared the three best mapping outputs from Objective 1 to LandTrendr woodlot detection from Objective 2.

2.2.3. Woodlot detection

To analyze whether the time series analysis approach affected woodlot detection, we performed LandTrendr segmentation and compared results to the linear model approach from Objective 1. LandTrendr segmentation output contained: year of detection, duration of segment, magnitude of change, pre-value, rate of change, and change–signal-to-noise ratio. We validated the LandTrendr segments using the independent validation samples and TimeSync (Cohen et al., 2010). In TimeSync, we checked the segmentation outcome from Landtrendr for each validation location. We recorded whether the segments correctly identified woodlots (i.e., pixels with gain) and the timing (year of detection) and duration (number of years) of woodlot gain. We compared TimeSync validation metrics to validation metrics from Objective 1.

To analyze the effect of woodlot characteristics (i.e., age and size) on woodlot detection, we conducted agreement/disagreement evaluation. Our benchmark was the independent database of manually digitized woodlots in random locations. We calculated agreement/disagreement on woodlot presence between the database and in maps produced by the linear model approach and by the LandTrendr segmentation approach. We report woodlot detection rates as a function of the age of the woodlot and the size of the woodlot.

3. Results

3.1. Woodlot mapping

We accurately mapped woodlots (up to 77.2 ± 6.7% accuracy) and estimated wide-spread woodlot coverage (36 ± 21%) in the study area. Our mapping results were robust to random forest parameterization and training sample size (Appendix C.1 – C.3). Spatially, more woodlots were in the higher elevation areas and near regions with large government and private timber plantations (Fig. 5A). Visual assessment of the maps showed that we mapped "Mature" woodlots accurately even when the woodlots were mixed in with other land uses (Fig. 5, Loc 1 & Loc 2). However, we had high omission errors for "Young" woodlots.

3.1.1. Time series length

Woodlot mapping accuracy increased with the length of the time series of the input data. For two-class maps ("Woodlot" vs "Other"), there was no statistically significant difference in the accuracies of the maps produced with at least a 5-year time series if using Landsat-8 images alone, or least a 4-year time series if using combined Landsat-8, 7, and 5 images (Table 3). However, peak accuracies were attained with a 7-year time series. In other words, woodlots were adequately distinguished from other land covers with a time series of 4–5 years, but additional years further improved mapping accuracies. These results were robust to different training sample sizes (Appendix C, Fig. 1C).

When we classified woodlots by age ("Young", "Intermediate", "Mature", and "Harvested"), the errors also generally decreased with time series length. Omission and commission errors generally decreased with longer time series particularly for the "Intermediate", "Mature", and "Harvested" woodlots (Fig. 4). Furthermore, when mis-classified, "Young" woodlots were mis-classified as "Other"; while the rest of the woodlot classes tended to be mis-classified as a woodlot of a different age. "Young" woodlots had highest omission errors, and these errors were not improved by time series length (Fig. 4). "Mature" woodlot class, on the other hand, had the lowest commission and omission errors even in very short time series (Fig. 4). Omission and commission errors for "Harvested" woodlots sharply declined with longer time series, for up to 7 years (Fig. 4).

3.1.2. Frequency of observations

Woodlots were mapped more accurately when input data combined

Table 3
Overall accuracies and p-values for the pairwise comparisons using Wilcoxon sign test for all mapping outputs. The maps had two classes: "Woodlots" versus "Other". They differed by sensor and length of the time series (columns/rows 1–3). The overall accuracies and margins of error are shown in column 4 and row 4. The bolded cells indicate comparisons that are statistically significantly different at p = 0.05. For similar comparison with all woodlot classes see Appendix D.

Sensor				Landsat-8, 7, & 5								Landsat-8					Sentinel-2
	Calendar Range			'10-'19	'11-'19	'12-'19	'13-'19	'14-'19	'15-'19	'16-'19	'17-'19	'13-'19	'14-'19	'15-'19	'16-'19	'17-'19	'16-'19
		Time series (yrs)		10	9	8	7	6	5	4	3	7	6	5	4	3	4
			Accuracy	0.70 (± 0.07)	0.69 (± 0.07)	0.73 (± 0.07)	0.77 (± 0.07)	0.74 (± 0.06)	0.70 (± 0.06)	0.70 (± 0.06)	0.65 (± 0.07)	0.71 (± 0.07)	0.72 (± 0.06)	0.68 (± 0.06)	0.66 (± 0.06)	0.54 (± 0.07)	0.59 (± 0.06)
Landsat-8, 7, & 5	'11-'19	9	0.69 (± 0.07)	0.91													
	'12-'19	8	0.73 (± 0.07)	0.50	0.23												
	'13-'19	7	0.77 (± 0.07)	0.15	**0.04**	0.38											
	'14-'19	6	0.74 (± 0.06)	0.37	0.29	0.79	0.94										
	'15-'19	5	0.70 (± 0.06)	1.00	0.91	0.78	0.33	0.30									
	'16-'19	4	0.70 (± 0.06)	1.00	1.00	0.65	0.29	0.33	0.94								
	'17-'19	3	0.65 (± 0.07)	0.30	0.40	0.13	**0.03**	**0.02**	0.16	0.21							
Landsat-8	'13-'19	7	0.71 (± 0.07)	0.86	0.68	0.86	0.30	0.58	0.96	0.86	0.22						
	'14-'19	6	0.72 ± 0.06)	0.58	0.47	1.00	0.70	0.86	0.65	0.51	0.06	0.76					
	'15-'19	5	0.68 (± 0.06)	0.86	0.96	0.49	0.20	0.19	0.71	0.89	0.40	0.63	0.23				
	'16-'19	4	0.66 (± 0.06)	0.47	0.59	0.22	0.05	**0.04**	0.29	0.33	0.84	0.30	0.06	0.52			
	'17-'19	3	0.54 (± 0.07)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**		
Sentinel-2	'16-'19	4	0.59 (± 0.06)	**0.00**	**0.01**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	0.11	**0.00**	**0.00**	**0.01**	0.06	0.49	
	'17-'19	3	0.57 (± 0.06)	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.00**	**0.01**	**0.00**	**0.00**	**0.00**	**0.00**	0.97	0.29

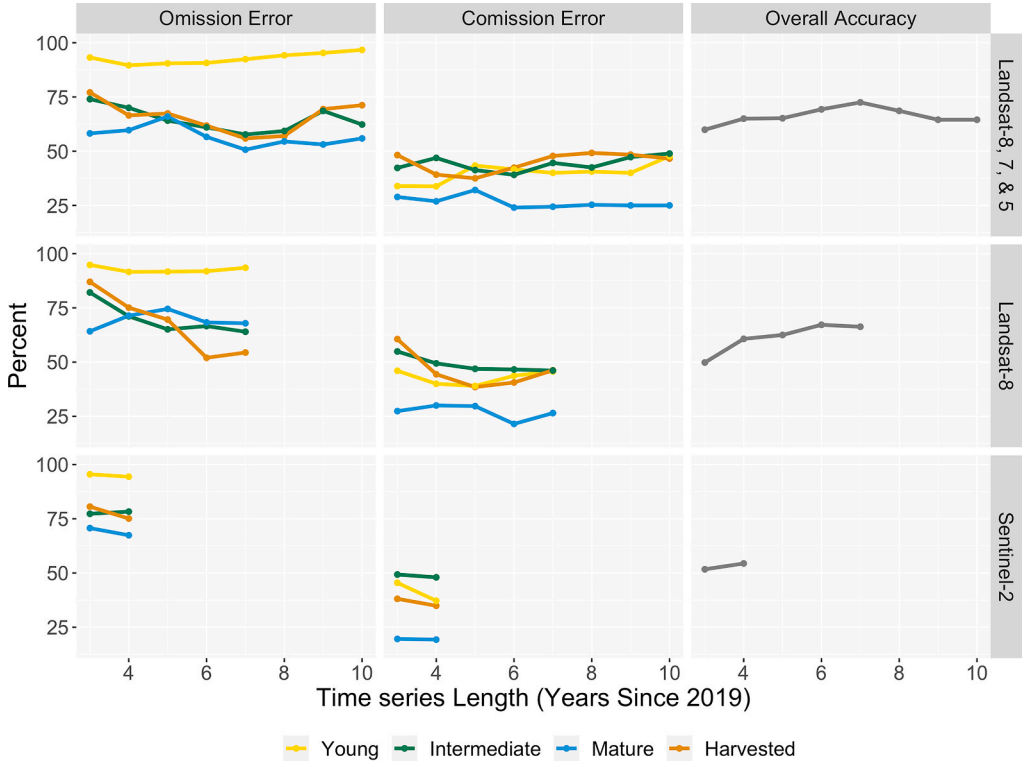

Fig. 4. Commission and omission errors and overall accuracy given input image characteristics. For each sensor, the input images varied by time series length, ranging from a 3- to a 10-year time series. For Sentinel-2, the maximum time series length used was 4 years, and for Landsat-8, the maximum length used was 7 years.

multiple sensors, thus increasing the frequency of observations. Two-class output maps ("Woodlot" vs "Other") that reached highest overall accuracies resulted from time series inputs that combined Landsat-8, 7, and 5 (77%, 7-yr time series). Woodlots mapped from only Landsat-8 time series of the same length had lower overall accuracies, but this difference was not statistically significant (71%, p-value 0.30; Table 3). For time series longer than 7 years, the overall accuracy of the combined Landsat-8, 7, and 5 started to decline, but most of these maps are not statistically significantly different from each other (7-year: 77%; 8-year: 73%; 9-year: 69%; 10-year: 70% Table 3).

3.1.3. *Spatial resolution*

Sentinel-2's higher spatial resolution had the lowest overall accuracies for the two-class output maps ("Woodlot" vs "Other"). For the shortest time series of 3 years, woodlots were mapped more accurately with Sentinel-2's 10-m images compared to Landsat-8's 30 m input images (3-year time series: Sentinel-2: 57% vs Landsat-8: 54%; p-value 0.97). For the 4-year time series, Landsat-8 had higher accuracy (4-year

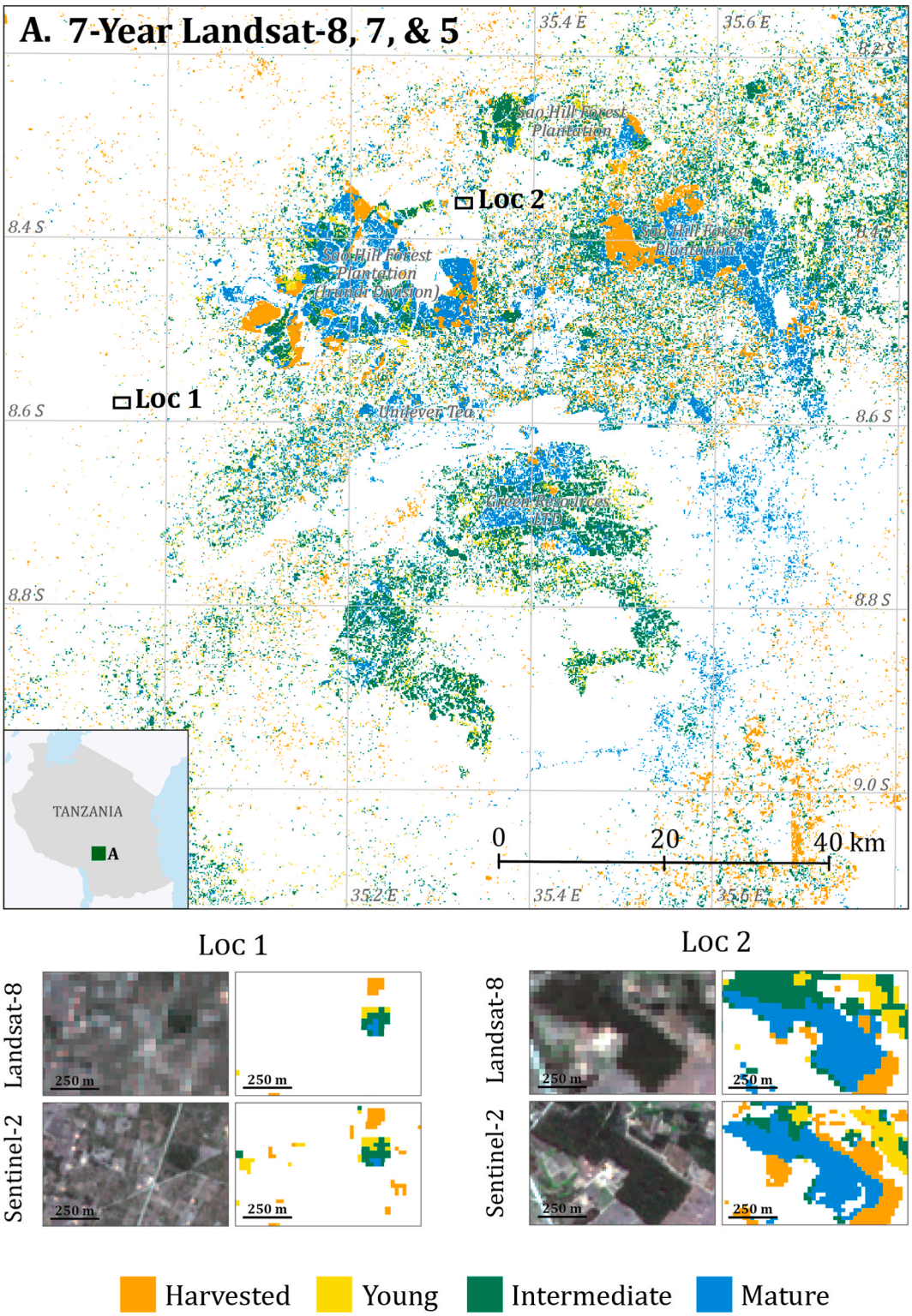

Fig. 5. A) Woodlots of different ages detected in our study area. Loc 1 and Loc 2 show differences in woodlot sizes as well as differences in detection for different spatial resolutions (Landsat-8 30-m versus Sentinel-2 10-m resolution). Lower spatial resolution resulted in more confusion among woodlot age classes, while higher spatial resolution minimized commission errors especially for mature woodlots. Labeled areas are large pine and eucalyptus plantations.

time series: Sentinel-2: 59% vs Landsat-8: 66%; p-value 0.06). These differences in accuracy were not statistically significantly different though (Table 3). However, the higher spatial resolution of Sentinel-2 images resulted in the lowest commission errors, especially for the "Mature" woodlot class (20% and 19% for 3-year and 4-year time series respectively, Fig. 4).

3.2. Woodlot detection

3.2.1. Time series analysis approach

All LandTrendr time series segmentation based on the different indices (EVI, TCB, TCW, NDVI) accurately identified woodlots. For maps with two land cover classes ("Woodlot" and "Other"), the overall

accuracy was 81.4–83.6%. The best index for woodlot identification was TCW. However, there was no statistically significant difference in woodlot detection accuracy between the LandTrendr segments and our time series linear model approach (see Appendix D Table D2 for pairwise comparison).

LandTrendr temporal segments were less accurate in characterizing woodlot age via generation of a breakpoint year and calculation of segment length. LandTrendr segments generated from different indices placed the breakpoint for year of woodlot establishment (i.e., breakpoint preceding any gain segment) correctly in only 34–38% of the samples when compared to the independent validation sample. The proportion increased 55.7–63.6% if the tolerance was ±2 years. Similarly, LandTrendr segments generated from different indices correctly estimated the age of the woodlot based on the length of the gain segment in only 15.9–24.8% of the validation samples. The proportion of accurately aged woodlots increased to 44.3% if the tolerance was ±2 years.

3.2.2. *Woodlot age and size*

There was a strong correspondence regarding presence vs. absence of woodlots between the best woodlot maps from both time series approaches, and the independent database of woodlots characteristics generated from manual digitization. 50–80% of the woodlots in the database were correctly identified as woodlots in the two-class maps ("Other" and "Woodlot"). 4–10% of the woodlots in the maps were false positives, meaning they showed up on the woodlot maps but were absent in the digitized woodlot database.

Woodlot size had a strong effect on woodlot detection. Woodlots <1 ha had higher omission errors than larger ones. On average, only 50% of woodlots <0.4 ha present in the digitized woodlot database were detected in woodlot maps. The proportion increased to 60% for woodlots measuring 0.4–1 ha. Highest detection (75–80%) was of woodlots >1 ha. Essentially, a woodlot >1 ha had an approximately 1.5 times higher likelihood of being detected compared to a woodlot <1 ha (Fig. 6).

Woodlot age also limited detectability. "Mature" woodlots were more readily detected than "Young" woodlots (Fig. 6). Furthermore, only ca. 50% of "Young" woodlots that were in the digitized woodlot database were detected in woodlot maps, and this was true for all woodlot sizes <1 ha. For woodlot sizes >1 ha, even the "Young" woodlots in database were detected in woodlot maps at a high proportion (~75%). However, detectability varied more strongly by woodlot size rather than age (Fig. 6).

Detected woodlots had commission errors across age classes, meaning the different ages were confused with each other. "Mature" woodlots were most consistently correctly identified. "Young" woodlots, when detected, were often confused with "Intermediate" woodlots. Similarly, when "Mature" woodlots were misclassified, they were most frequently identified as "Intermediate" woodlots (Fig. 6). "Harvested" woodlots were less likely to be identified as "Young", while "Young" woodlots were less likely to be identified as "Mature" (Fig. 6).

4. Discussion

We compared different time series approaches to detect and map

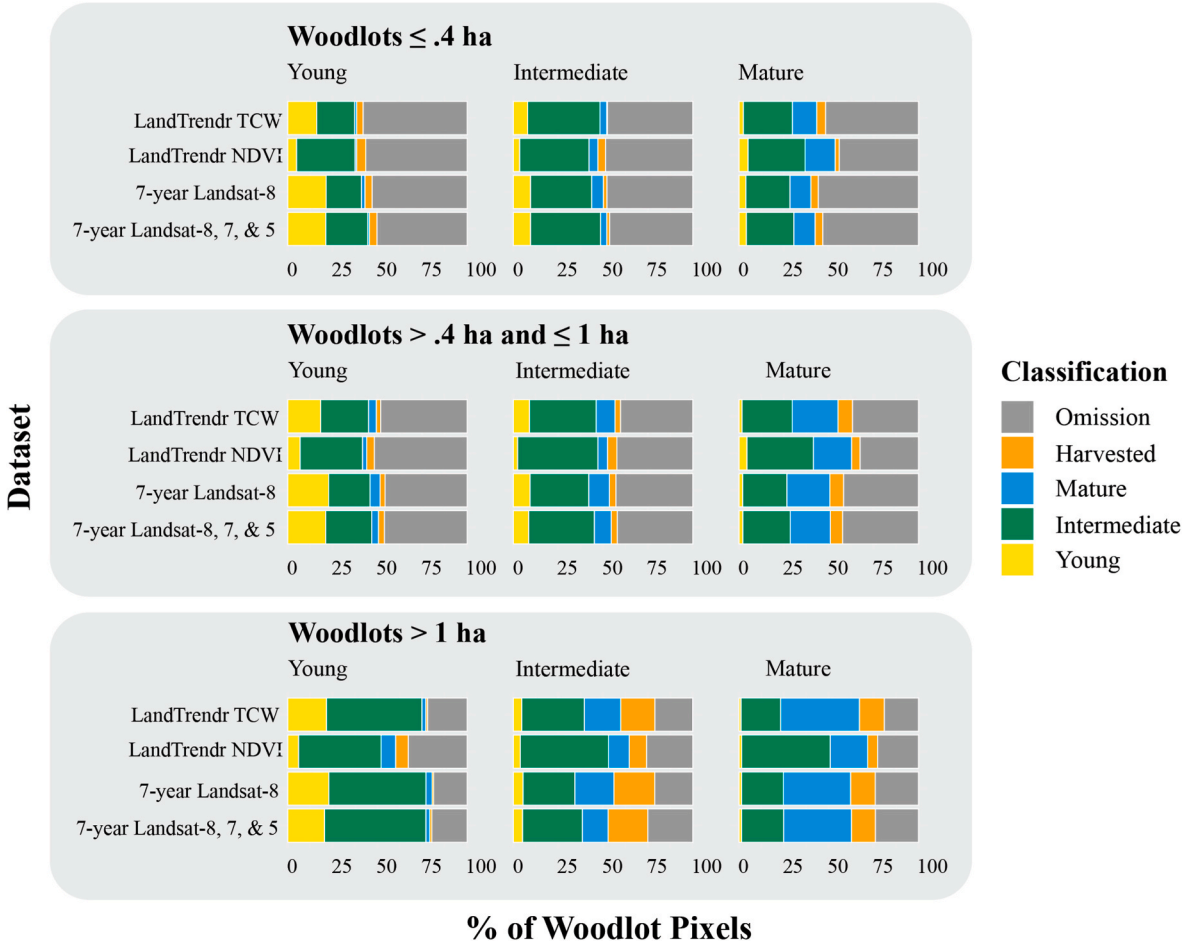

Fig. 6. Accurate woodlot detection depends on the age and size characteristics of the woodlot. Top: woodlots <0.4 ha, Middle: woodlots 0.4–1 ha, Bottom: woodlots >1 ha. The columns distinguish woodlots by their age. The woodlots with highest omission errors are those that are both "Young" and small (top-left). The charts also show misclassification within the woodlot class, with "Young", >1 Ha woodlots often misclassified as "Intermediate"-aged woodlots.

smallholder woodlots in a highly heterogeneous East African landscape. When the time series had sufficient duration and observation frequency, we found comparable performance in woodlot detection between our linear model of woodlots' spectral time series and LandTrendr temporal segmentation. Woodlot size and age limited accurate detection especially when a woodlot was both small and young. Our results are encouraging, because they indicate that spectral time series derived from a pixel-level linear model can be used to map woodlots accurately. However, omission errors were high, especially for very small woodlots.

4.1. Mapping woodlots accurately

The overall accuracy of the woodlot maps generally improved with a longer time series (5–7 years). Our finding differs from previous studies that suggested focusing on a single cloud-free and dry-season image for mapping woodlots (Indufor, 2011; Mankinen et al., 2017). We found that time series can capture woodlot's greening as the canopy fills in (Fig. 1), thereby capturing woodlots that are still establishing (i.e., "Young" and "Intermediate" woodlots). If woodlots are mapped with a single-date image, only the fully established woodlots (i.e., "Mature") are spectrally distinct enough to capture accurately (Fassnacht et al., 2016). The nascent (i.e., "Young" age class) are likely mis-classified as grassland or cropland in single-date image classifications. "Harvested" woodlots require time series information to be correctly identified as woodlots.

We also found that combining Landsat-8, 7, and 5 images improved woodlot detection, and that was the case for time series of any length. In fact, with more images, a shorter time series of only 4 years can adequately separate woodlots from other land covers, but longer time series minimize commission and omission errors across woodlot ages. Other plantation detection studies with different mapping approaches have similarly recognized the advantage of more frequent observations (Hurni et al., 2017; Qiao et al., 2016; Gao et al., 2015). Unfortunately, most of Africa has limited image availability prior to Landsat 8 due to a sparse network of ground-receiving stations and image acquisition strategies (Goward et al., 2006). However, number of images available for analysis almost doubled for some months when we combined all Landsat sensors (Appendix E). More frequent observations may improve classification accuracy by providing additional information about class separability to a classifier (Carrão et al., 2008; Carrasco et al., 2019). In our linear model approach, more frequent observations may have minimized errors in the linear model and in the estimates of the mean monthly index value.

Despite the accuracy gains from combining Landsat-8, 7, and 5, Landsat-8 was the most important data source in the spectral time series. This is likely why lengthening the time series beyond 7 years resulted in lower accuracies, simply because Landsat-8 images were not available for the earlier years. In analyses where the time series is long enough, woodlot detection could be performed with a time series that favor Landsat-8 as suggested in other land cover mapping studies with multi-date image composites (Azzari and Lobell, 2017; Griffiths et al., 2013).

Increasing the spatial resolution from Landsat's 30-m to Sentinel-2's 10-m resolution had limited effect on overall classification accuracy. This surprised us, given that a median-sized woodlot of 0.45 ha would be comprised of 45 Sentinel-2 pixels compared to five Landsat-8 pixels. In heterogeneous landscapes, the higher spatial resolution reduces mixed pixels, which improves class separability (Ozdogan and Woodcock, 2006; Li et al., 2014; Xie et al., 2008). Furthermore, Sentinel-2 images are equipped with additional red edge bands, which we incorporated in our analysis, and are known to be helpful in tree cover mapping (Grabska et al., 2019). However, none of this significantly improved the overall classification accuracy.

Despite the lack of overall accuracy improvements, Sentinel-2 maps had the lowest commission errors particularly for the "Mature" woodlot class. An increase in spatial resolution improved "Mature" woodlots class separability likely by reducing mixed pixels. Furthermore, "Mature" woodlots benefited the least from time series length, likely due to their sufficient spectral distinctness from other land cover classes (Fassnacht et al., 2016). For other woodlot classes, the gains in spatial resolution could not make up for the short time series. As the Sentinel-2 mission accrues more years of observation, its higher spatial resolution, coupled with longer time series, may improve woodlot and other land cover mapping (Carrasco et al., 2019).

4.2. Woodlot detection with time series analysis

We found that the two different time series analysis approaches (linear model of spectral time series versus temporal segmentation algorithm, LandTrendr) were equivalent in their woodlot detection. LandTrendr segmentation can be advantageous because the segments contain information about timing and duration of change that could be used to determine a woodlot's age (Cohen et al., 2010). However, while LandTrendr segments could accurately identify woodlots based on whether the pixel had any greenness gain, the accuracies for characterizing woodlots' age were low.

Methods that take advantage of time series analysis are needed for mapping land cover trajectories, such as newly planted woodlots in sub-Saharan Africa (Bey and Meyfroidt, 2021). LandTrendr is a well-established approach with a user-ready tool implemented in Google Earth Engine (Kennedy et al., 2018). However, LandTrendr segmentation produces more accurate segments when available input images are well-timed, cloud-free, and of sufficient observation frequency for the time window of interest (Kennedy et al., 2010). LandTrendr analysis can better detect land cover change when the disturbance is at a scale appropriate to the image resolution (e.g., forest clearing as opposed to selective logging (Rodman et al., 2021; Nguyen et al., 2020)). Taking full advantage of the LandTrendr algorithm proved challenging in our study area. Similar to other parts of sub-Saharan Africa, our study area had few high-quality images available, especially for earlier Landsat missions (Ju and Roy, 2008; Yu et al., 2015). Therefore, even with careful parameterization (Appendix B), LandTrendr did not outperform the linear model with supervised classification approach (Appendix D).

Despite LandTrendr's limitations in our study, its segments can substantially improve disturbance detection when used as inputs to a secondary classifier (Cohen et al., 2018, 2020). A supplemental analysis of an ensemble of segments from multiple indices as inputs showed that the approach reduced commission errors (Appendix D, Table D2). Our analysis corroborates previous findings for how ensemble approaches can improve forest disturbance detection (Cohen et al., 2020; Hislop et al., 2019).

Woodlot characteristics (i.e., age and size) also affected detectability. We found that about half of the small woodlots were missed, and that raises the question if previous woodlot presence/absence studies underestimated woodlot extent (Koskinen et al., 2019). This is a cautionary finding because the presence/absence studies are viewed as a way to meet national timber assessment needs (Mankinen et al., 2017; Mauya et al., 2019). We suggest that woodlot mapping via remote sensing need to be paired with field inventories targeting woodlots smaller than 0.4 ha. Alternatively, studies can be supplemented with wall-to-wall manual digitization of sample areas in order to estimate margins of error.

Overall, we found that woodlots could be detected, but determining their age posed a challenge. Smallholders tend to plant woodlots in neighboring plots of land (Fig. 1), thereby placing woodlots of various ages and sizes next to each other. As the effective size of a tree planted area increases due to spatial contiguity, woodlot presence/absence detection improves. Essentially, the spatial pattern of a target land cover interacts with the resolution of the input image in determining the accuracy of the final map (Ozdogan and Woodcock, 2006; Wu, 2004). Whereas contiguous planting of woodlots may have improved overall detection, it caused confusion when mapping woodlot age class by mixing pixels whose woodlots were of different ages. The most common

misclassification was inclusion of other age categories in the "Intermediate" woodlot age category, especially for smaller woodlots. Additionally, "Young" woodlots are challenging to detect because of their confusion with surrounding landscape (Dong et al., 2013), while "Harvested" woodlots require a time series in order to capture the transition from trees to no trees.

4.3. Methods transfer to map fine-scale tree gain

Our woodlot mapping approach was successful because we analyzed time series which improved both training sample generation and woodlot age separability. Previous woodlot detection analyses generated training data from visual interpretation of high-resolution Google Earth Pro images (Koskinen et al., 2019; Mankinen et al., 2017). However, Google Earth Pro images do not have consistent image coverage over time, thus some training samples could be out of date relative to the classification image. We minimized this challenge by collecting training samples via simultaneous visual interpretation in Google Earth Pro and in a spectral time series viewer tool (Yin, 2019) (Appendix A). This combined approach improved the quality of the training samples and enabled assignment of woodlot age classes. However, our overall accuracies were lower compared to mapping of large-scale plantations in sub-Saharan Africa (Bey and Meyfroidt, 2021) or in South Asia (Torbick et al., 2016; Deng et al., 2020; Hurni et al., 2017). Those difference are to be expected given the small size of woodlots and the heterogeneity of smallholder landscapes.

Our analysis is timely because we assessed methods for quantifying woodlots amidst the documented trend of tree cover gain in sub-Saharan Africa's drylands and farmland (Brandt et al., 2018; Miller et al., 2017). Some studies have observed tree gain by mapping new large plantations (Torbick et al., 2016; Bey and Meyfroidt, 2021). Plantation mapping may systematically miss smallholder woodlots (Koskinen et al., 2019; Mankinen et al., 2017). Though there is evidence for emergence of smallholder woodlots in Southern Highlands of Tanzania (Held et al., 2017; Etongo et al., 2015; Friis-Hansen and Pedersen, 2016), western Uganda (L'Roe and Naughton-Treves, 2016; Bailey et al., 2021), and Ethiopian highlands (Jenbere et al., 2012; Telila et al., 2015), most of the evidence is not based on remote sensing. Our time series mapping approach can be used to fill this knowledge gap by quantify smallholder woodlots more routinely in these regions.

There is renewed impetus and funding for the use of tree planting for landscape restoration, habitat improvement, and carbon sequestration (Hess, 2021). New initiatives like the Bonn Challenge (Fagan et al., 2020) increase the need for scalable approaches for monitoring fine-scale tree cover gain. Remote sensing could be a solution, but new approaches are needed for monitoring restoration when trees are newly planted and where restoration is undertaken by many distributed actors in small land areas. Our study highlights how time series approaches can be used to map and monitor activities of many distributed tree planters. By implementing a mapping approach like ours, restoration monitors can provide a more accurate assessment of how stakeholders are meeting their restoration pledges.

Finally, there is a need for accurate and up-to-date woodlot maps in order to resolve some accounting challenges in restoration implementation. Woodlot establishment is a core restoration approach and reported as such (Dave et al., 2017; FLR, 2015). Unfortunately, woodlots are counted as landscape restoration even when they are established for non-restoration purposes. In landscape restoration pledges, such as the Bonn Challenge, there is no clear distinction between woodlot establishment that occurred independent of restoration pledges and establishment that occurred because of it (Fagan et al., 2020; Dave et al., 2017; Pistorius et al., 2017). Furthermore, short-rotation plantations may provide miniscule carbon benefits especially if they are used for biofuel (Kongsager et al., 2013; Johnston and Radeloff, 2019). As more smallholders undertake woodlot establishment independent of global pledges for restoration, their activities need to be distinguishable from restoration-driven fine-scale tree cover gain. Mapping approaches like ours provide tools for better characterization woodlots and consequently better attribution of the role of smallholders in global landscape restoration pledges.

5. Conclusion

We used time series to map smallholder woodlots and characterize their age. Our results suggest that accurate detection and mapping of woodlots is possible by analyzing multi-year time series of Landsat and Sentinel-2 data. Though our approach was overall successful, there were widespread omission errors for small and young woodlots. Furthermore, time series analysis with our supervised classification outperformed automated temporal segmentation with LandTrendr particularly in assigning woodlot age. These findings are timely given the region's woodlot boom and continued promotion of tree planting in global environmental policies. Accurate maps are needed to better quantify the contribution of woodlots and other fine-scale tree cover gain to carbon sequestration, livelihoods enhancement, and landscape management.

Author responsibilities

Conceptualization, N.E.K. and V.C.R.; methodology, N.E.K. and V.C.R.; investigation, N.E.K.; data curation, N.E.K.; writing—original draft preparation, N.E.K.; writing—review and editing, V.C.R. and N.E.K.; funding acquisition, V.C.R. and N.E.K.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

We thank three anonymous reviewers for their constructive feedback. We thank K. Lewińska for advising on analytical design and for feedback on the manuscript. I. Rojas and K. Tucker commented on earlier drafts of the manuscript. K. Udoh, Y. Xu, L. Sovell-Fernandez, and V. Falardeau assisted with figures through Middlebury's Undergraduate Research Program. N.E.K received a 1-year National Geographic Society's Early Career Grant (EC-51238R-19) supporting this work, and we gratefully acknowledge support by the NASA Land Cover and Land Use Change Program.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.srs.2023.100096.

References

Arvola, A., Malkamäki, A., Penttilä, J., Toppinen, A., 2019. Mapping the future market potential of timber from small-scale tree farmers: perspectives from the southern highlands in Tanzania. Small-Scale For 18, 189–212. https://doi.org/10.1007/s11842-019-09414-8.

Arvola, A., Brockhaus, M., Kallio, M., Pham, T.T., Chi, D.T.L., Long, H.T., Nawir, A.A., Phimmavong, S., Mwamakimbullah, R., Jacovelli, P., 2020. What drives smallholder tree growing? Enabling conditions in a changing policy environment. For. Policy Econ. 116, 102173 https://doi.org/10.1016/j.forpol.2020.102173.

Azzari, G., Lobell, D.B., 2017. Landsat-based classification in the cloud: an opportunity for a paradigm shift in land cover monitoring. Remote Sens. Environ. 202, 64–74. https://doi.org/10.1016/j.rse.2017.05.025.

Bailey, K., Salerno, J., Newton, P., Bitariho, R., Namusisi, S., Tinkasimire, R., Hartter, J., 2021. Woodlot management and livelihoods in a tropical conservation landscape. Ambio 50, 1351–1363. https://doi.org/10.1007/s13280-020-01484-9.

Bey, A., Meyfroidt, P., 2021. Improved land monitoring to assess large-scale tree plantation expansion and trajectories in Northern Mozambique. Environ. Res. Commun. 0–12. https://doi.org/10.1088/2515-7620/ac26ab.

Borah, B., Bhattacharjee, A., Ishwar, N., 2018. Bonn Challenge and India, New Delhi, India.

Brandt, M., Rasmussen, K., Hiernaux, P., Herrmann, S., Tucker, C.J., Tong, X., Tian, F., Mertz, O., Kergoat, L., Mbow, C., David, J.L., Melocik, K.A., Dendoncker, M., Vincke, C., Fensholt, R., 2018. Reduction of tree cover in West African woodlands and promotion in semi-arid farmlands. Nat. Geosci. 11, 328–333. https://doi.org/10.1038/s41561-018-0092-x.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Bullock, E., Olofsson, P., Arevalo, P., 2019. AREA2: Tools and Guidance for Estimating Map Accuracy and Area on the Google Earth Engine Platform. https://github.com/bullocke/AREA2.

Carrão, H., Gonçalves, P., Caetano, M., 2008. Contribution of multispectral and multitemporal information from MODIS images to land cover classification. Remote Sens. Environ. 112, 986–997. https://doi.org/10.1016/j.rse.2007.07.002.

Carrasco, L., O'Neil, A.W., Daniel Morton, R., Rowland, C.S., 2019. Evaluating combinations of temporally aggregated sentinel-1, sentinel-2 and landsat 8 for land cover mapping with Google Earth engine. Rem. Sens. 11 https://doi.org/10.3390/rs11030288.

Cohen, W.B., Yang, Z., Kennedy, R., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync - tools for calibration and validation. Remote Sens. Environ. 114, 2911–2924. https://doi.org/10.1016/j.rse.2010.07.010.

Cohen, W.B., Yang, Z., Healey, S.P., Kennedy, R.E., Gorelick, N., 2018. A LandTrendr multispectral ensemble for forest disturbance detection. Remote Sens. Environ. 205, 131–140. https://doi.org/10.1016/j.rse.2017.11.015.

Cohen, W.B., Healey, S.P., Yang, Z., Zhu, Z., Gorelick, N., 2020. Diversity of algorithm and spectral band inputs improves landsat monitoring of forest disturbance. Rem. Sens. 12, 1–15. https://doi.org/10.3390/rs12101673.

Crowther, T.W., Glick, H.B., Covey, K.R., Bettigole, C., Maynard, D.S., Thomas, S.M., Smith, J.R., Hintler, G., Duguid, M.C., Amatulli, G., Tuanmu, M.N., Jetz, W., Salas, C., Stam, C., Piotto, D., Tavani, R., Green, S., Bruce, G., Williams, S.J., Wiser, S.K., Huber, M.O., Hengeveld, G.M., Nabuurs, G.J., Tikhonova, E., Borchardt, P., Li, C.F., Powrie, L.W., Fischer, M., Hemp, A., Homeier, J., Cho, P., Vibrans, A.C., Umunay, P.M., Piao, S.L., Rowe, C.W., Ashton, M.S., Crane, P.R., Bradford, M.A., 2015. Mapping tree density at a global scale. Nature 525, 201–205. https://doi.org/10.1038/nature14967.

Dave, R., Saint-Laurent, C., Moraes, M., Simonit, S., Raes, L., Karangwa, C., 2017. Bonn Challenge Barometer of Progress: Spotlight Report, p. 36. https://portals.iucn.org/library/sites/library/files/documents/2017-060.pdf.

Deng, X., Guo, S., Sun, L., Chen, J., 2020. Identification of short-rotation eucalyptus plantation at large scale using multi-satellite imageries and cloud computing platform. Rem. Sens. 12 https://doi.org/10.3390/rs12132153.

Dong, J., Xiao, X., Chen, B., Torbick, N., Jin, C., Zhang, G., Biradar, C., 2013. Mapping deciduous rubber plantations through integration of PALSAR and multi-temporal Landsat imagery. Remote Sens. Environ. 134, 392–402. https://doi.org/10.1016/j.rse.2013.03.014.

Dymond, C.C., Mladenoff, D.J., Radeloff, V.C., 2002. Phenological differences in Tasseled Cap indices improve deciduous forest classification. Remote Sens. Environ. 80, 460–472. https://doi.org/10.1016/S0034-4257(01)00324-8.

Etongo, D., Djenontin, I.N.S., Kanninen, M., Fobissie, K., 2015. Smallholders' tree planting activity in the ziro province, southern Burkina Faso: impacts on livelihood and policy implications. Forests 6, 2655–2677. https://doi.org/10.3390/f6082655.

Fagan, M.E., Reid, J.L., Holland, M.B., Drew, J.G., Zahawi, R.A., 2020. How feasible are global forest restoration commitments? Conserv. Lett. 1–8. https://doi.org/10.1111/conl.12700.

Fassnacht, F.E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L.T., Straub, C., Ghosh, A., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L.T., Straub, C., Ghosh, A., Lati, H., Sterenczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., Straub, C., Ghosh, A., 2016. Review of studies on tree species classification from remotely sensed data. Remote Sens. Environ. 186, 64–87. https://doi.org/10.1016/j.rse.2016.08.013.

FDT, 2015. Baseline Tree Grower Survey Report. Forestry Development Trust, Iringa, Tanzania, pp. 1–65.

Fick, S.E., Hijmans, R.J., 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37, 4302–4315. https://doi.org/10.1002/joc.5086.

FLR, I.U.C.N., 2015. The Bonn Challenge: Restoration Options. http://www.bonnchallenge.org/content/restoration-options. (Accessed 10 November 2018).

Foody, G.M., 2004. Thematic map comparison: evaluating the statistical significance of differences in classification accuracy. Photogramm. Eng. Rem. Sens. 70, 627–633. https://doi.org/10.14358/PERS.70.5.627.

Frayer, J., Sun, Z., Müller, D., Munroe, D.K., Xu, J., 2014. Analyzing the drivers of tree planting in Yunnan, China, with Bayesian networks. Land Use Pol. 36, 248–258. https://doi.org/10.1016/j.landusepol.2013.08.005.

Friis-Hansen, E., Pedersen, R.H., 2016. Timber Rush - Private Forestry on Village Land. https://www.diis.dk/en/projects/timber-rush-private-forestry-on-village-land. (Accessed 15 November 2018).

Gao, T., Zhu, J., Zheng, X., Shang, G., Huang, L., Wu, S., 2015. Mapping Spatial Distribution of Larch Plantations from Multi-Seasonal Landsat-8 OLI Imagery and Multi-Scale Textures Using Random Forests, pp. 1702–1720. https://doi.org/10.3390/rs70201702.

Gao, T., Zhu, J., Deng, S., Zheng, X., Zhang, J., 2016. Timber production assessment of a plantation forest : an integrated framework with field-based inventory , multi-source remote sensing data and forest management history. Int. J. Appl. Earth Obs. Geoinf. 52, 155–165. https://doi.org/10.1016/j.jag.2016.06.004.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth engine: planetary-scale geospatial analysis for everyone. Remote Sens. Environ. 202, 18–27. https://doi.org/10.1016/j.rse.2017.06.031.

Goward, S., Arvidson, T., Williams, D., Faundeen, J., Irons, J., Franks, S., 2006. Historical record of landsat global coverage: mission operations, NSLRSDA, and international cooperator stations. Photogramm. Eng. Rem. Sens. 72, 1155–1169. https://doi.org/10.14358/PERS.72.10.1155.

Grabska, E., Hostert, P., Pflugmacher, D., Ostapowicz, K., 2019. Forest stand species mapping using the sentinel-2 time series. Rem. Sens. 11, 1–24. https://doi.org/10.3390/rs11101197.

Griffiths, P., van der Linden, S., Kuemmerle, T., Hostert, P., 2013. A pixel-based landsat compositing algorithm for large area land cover mapping. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 6, 2088–2101. https://doi.org/10.1109/jstars.2012.2228167.

Hansen, M.C., V Potapov, P., Moore, R., Hancher, M., a Turubanova, S., Tyukavina, a, Thau, D., V Stehman, S., Goetz, S.J., Loveland, T.R., Kommareddy, a, Egorov, a, Chini, L., Justice, C.O., Townshend, J.R.G., 2013. High-resolution global maps of 21st-century forest cover change. Science 342, 850–853. https://doi.org/10.1126/science.1244693.

Healey, S.P., Cohen, W.B., Zhiqiang, Y., Krankina, O.N., 2005. Comparison of Tasseled Cap-based Landsat data structures for use in forest disturbance detection. Remote Sens. Environ. 97, 301–310. https://doi.org/10.1016/j.rse.2005.05.009.

Held, C., Jacovelli, P., Techel, G., Nutto, L., Wathum, G., Wittmann, N., 2017. Tanzanian Wood Product Market Study, p. 134.

Hess, L., 2021. AFR100 initiative gets a boost as USD 2 billion funding goal before next COP set. Glob. Landsc. Forum. https://news.globallandscapesforum.org/55716/afr100-initiative-gets-a-boost-as-usd-2-billion-funding-goal-before-next-cop-set/?utm_campaign=GLF_at_COP26_Newsletter_3rd&utm_medium=email&utm_source=Mailchimp_Emailblast_November&utm_source=General+contacts&u.

Hislop, S., Jones, S., Soto-Berelov, M., Skidmore, A., Haywood, A., Nguyen, T.H., 2019. A fusion approach to forest disturbance mapping using time series ensemble techniques. Remote Sens. Environ. 221, 188–197. https://doi.org/10.1016/j.rse.2018.11.025.

Hurni, K., Schneider, A., Heinimann, A., Nong, D., Fox, J., 2017. Mapping the expansion of boom crops in mainland southeast Asia using dense time stacks of landsat data. Rem. Sens. 9, 320. https://doi.org/10.3390/rs9040320.

Indufor, 2011. Timber Market Dynamics in Tanzania and in Key Export Markets: A Market Study. Dar-es-Salaam, Tanzania.

Jacovelli, P.A., 2009. Uganda's sawlog production grant scheme: a success story from Africa. Int. For. Rev. 11, 119–125. https://doi.org/10.1505/ifor.11.1.119.

Jacovelli, P.A., 2014. The future of plantations in Africa. Int. For. Rev. 16, 144–159. https://doi.org/10.1505/146554814811724748.

Jenbere, D., Lemenih, M., Kassa, H., 2012. Expansion of eucalypt farm forestry and its determinants in Arsi Negelle District, south Central Ethiopia. Small-Scale For 11, 389–405. https://doi.org/10.1007/s11842-011-9191-x.

Johnston, C.M.T., Radeloff, V.C., 2019. Global mitigation potential of carbon stored in harvested wood products. Proc. Natl. Acad. Sci. U.S.A. 116, 14526–14531. https://doi.org/10.1073/pnas.1904231116.

Ju, J., Roy, D.P., 2008. The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally. Remote Sens. Environ. 112, 1196–1211. https://doi.org/10.1016/j.rse.2007.08.011.

Kennedy, R.E., Yang, Z., Cohen, W.B., 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - temporal segmentation algorithms. Remote Sens. Environ. 114, 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008.

Kennedy, R.E., Yang, Z., Gorelick, N., Braaten, J., Cavalcante, L., Cohen, W.B., Healey, S., 2018. Implementation of the LandTrendr algorithm on Google Earth engine. Rem. Sens. 10, 1–10. https://doi.org/10.3390/rs10050691.

Kimambo, N.E., Naughton-Treves, L., 2019. The role of woodlots in forest regeneration outside protected areas: lessons from Tanzania. Forests 10. https://doi.org/10.3390/f10080621.

Kimambo, N.E., L'Roe, J., Naughton-Treves, L., Radeloff, V.C., 2020. The role of smallholder woodlots in global restoration pledges – lessons from Tanzania. For. Pol. Econ. 115 https://doi.org/10.1016/j.forpol.2020.102144.

Kongsager, R., Napier, J., Mertz, O., 2013. The carbon sequestration potential of tree crop plantations. Mitig. Adapt. Strategies Glob. Change 18, 1197–1213. https://doi.org/10.1007/s11027-012-9417-z.

Kontgis, C., Schneider, A., Ozdogan, M., 2015. Mapping rice paddy extent and intensification in the Vietnamese Mekong River Delta with dense time stacks of Landsat data. Remote Sens. Environ. 169, 255–269. https://doi.org/10.1016/j.rse.2015.08.004.

Koskinen, J., Leinonen, U., Vollrath, A., Ortmann, A., Lindquist, E., d'Annunzio, R., Pekkarinen, A., Käyhkö, N., Annunzio, R., 2019. Participatory mapping of forest plantations with open foris and Google Earth engine. ISPRS J. Photogrammetry Remote Sens. 148, 63–74. https://doi.org/10.1016/j.isprsjprs.2018.12.011.

Kröger, M., 2012. Global Tree Plantation Expansion : a Review. ICAS Review Paper Series No. 3.

Li, M., Zang, S., Zhang, B., Li, S., Wu, C., 2014. A review of remote sensing image classification techniques: the role of Spatio-contextual information. Eur. J. Remote Sens. 47, 389–411. https://doi.org/10.5721/EuJRS20144723.

Lobell, D.B., Asner, G.P., Ortiz-Monasterio, J.I., Benning, T.L., 2003. Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties. Agric. Ecosyst. Environ. 94, 205–220. https://doi.org/10.1016/S0167-8809(02)00021-X.

Lu, D., Weng, Q., 2007. A survey of image classification methods and techniques for improving classification performance. Int. J. Rem. Sens. 28, 823–870. https://doi.org/10.1080/01431160600746456.

L'Roe, J., Naughton-Treves, L., 2016. Forest edges in western Uganda: from refuge for the poor to zone of investment. For. Pol. Econ. 1–41. https://doi.org/10.1016/j.forpol.2016.12.011.

Mankinen, U., Kayhko, N., Koskinen, J., Pekkarinen, A., 2017. Forest Plantation Mapping of the Southern Highlands. Iringa, Tanzania.

Mather, A.S., 2007. Recent Asian forest transitions in relation to foresttransition theory. Int. For. Rev. 9, 491–502. https://doi.org/10.1505/ifor.9.1.491.

Mauya, E.W., Koskinen, J., Tegel, K., Hämäläinen, J., Kauranne, T., Käyhkö, N., 2019. Modelling and predicting the growing stock volume in small-scale plantation forests of Tanzania using multi-sensor image synergy. Forests 10, 279. https://doi.org/10.3390/f10030279.

Miller, D.C., Muñoz-Mora, J.C., Christiaensen, L., 2017. Prevalence, economic contribution, and determinants of trees on farms across Sub-Saharan Africa, for. Pol. Econ. 84, 47–61. https://doi.org/10.1016/j.forpol.2016.12.005.

Nawir, A.A., Kassa, H., Sandewall, M., Dore, D., Campbell, B., Ohlsson, B., Bekele, M., 2007. Stimulating smallholder tree planting - lessons from Africa and Asia. Unasylva 58, 53–58.

Ngaga, Y.M., 2011. Forest plantations and woodlots in Tanzania. African For. Forum Work. Pap. Ser. 1, 80.

Nguyen, T.H., Jones, S., Soto-Berelov, M., Haywood, A., Hislop, S., 2020. Landsat time-series for estimating forest aboveground biomass and its dynamics across space and time: a review. Rem. Sens. 12, 1–25. https://doi.org/10.3390/RS12010098.

Nomura, K., Mitchard, E.T.A., 2018. More than meets the eye: using Sentinel-2 to map small plantations in complex forest landscapes. Rem. Sens. 10 https://doi.org/10.3390/rs10111693.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57. https://doi.org/10.1016/j.rse.2014.02.015.

Ozdogan, M., Woodcock, C.E., 2006. Resolution dependent errors in remote sensing of cultivated areas. Remote Sens. Environ. 103, 203–217. https://doi.org/10.1016/j.rse.2006.04.004.

Payn, T., Carnus, J.M., Freer-Smith, P., Kimberley, M., Kollert, W., Liu, S., Orazio, C., Rodriguez, L., Silva, L.N., Wingfield, M.J., 2015. Changes in planted forests and future global implications. For. Ecol. Manage. 352, 57–67. https://doi.org/10.1016/j.foreco.2015.06.021.

Pistorius, T., Carodenuto, S., Wathum, G., 2017. Implementing forest landscape restoration in Ethiopia. Forests 8, 61. https://doi.org/10.3390/f8030061.

Qiao, H., Wu, M., Shakir, M., Wang, L., Kang, J., Niu, Z., 2016. Classification of small-scale eucalyptus plantations based on NDVI time series obtained from multiple high-resolution datasets. Rem. Sens. 8, 1–20. https://doi.org/10.3390/rs8020117.

Rodman, K.C., Andrus, R.A., Veblen, T.T., Hart, S.J., 2021. Disturbance detection in landsat time series is influenced by tree mortality agent and severity, not by prior disturbance. Remote Sens. Environ. 254, 112244 https://doi.org/10.1016/j.rse.2020.112244.

Roy, D.P., Kovalskyy, V., Zhang, H.K., Vermote, E.F., Yan, L., Kumar, S.S., Egorov, A., 2016. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. Remote Sens. Environ. 185, 57–70. https://doi.org/10.1016/j.rse.2015.12.024.

Rudel, T.K., 2009. Tree farms: driving forces and regional patterns in the global expansion of forest plantations. Land Use Pol. 26, 545–550. https://doi.org/10.1016/j.landusepol.2008.08.003.

Schneibel, A., Stellmes, M., Röder, A., Frantz, D., Kowalski, B., Haß, E., Hill, J., 2017. Assessment of spatio-temporal changes of smallholder cultivation patterns in the Angolan Miombo belt using segmentation of Landsat time series. Remote Sens. Environ. 195, 118–129. https://doi.org/10.1016/j.rse.2017.04.012.

Sexton, J.O., Noojipady, P., Song, X.P., Feng, M., Song, D.X., Kim, D.H., Anand, A., Huang, C., Channan, S., Pimm, S.L., Townshend, J.R., 2016. Conservation policy and the measurement of forests. Nat. Clim. Change 6, 192–196. https://doi.org/10.1038/nclimate2816.

Sheeren, D., Fauvel, M., Josipovíc, V., Lopes, M., Planque, C., Willm, J., Dejoux, J.F., 2016. Tree species classification in temperate forests using Formosat-2 satellite image time series. Rem. Sens. 8, 1–29. https://doi.org/10.3390/rs8090734.

Sonnenschein, R., Kuemmerle, T., Udelhoven, T., Stellmes, M., Hostert, P., 2011. Differences in Landsat-based trend analyses in drylands due to the choice of vegetation estimate. Remote Sens. Environ. 115, 1408–1420. https://doi.org/10.1016/j.rse.2011.01.021.

Telila, H., Hylander, K., Nemomissa, S., 2015. The potential of small Eucalyptus plantations in farmscapes to foster native woody plant diversity: local and landscape constraints. Restor. Ecol. 23, 918–926. https://doi.org/10.1111/rec.12257.

Torbick, N., Ledoux, L., Salas, W., Zhao, M., 2016. Regional mapping of plantation extent using multisensor imagery. Rem. Sens. 8 https://doi.org/10.3390/rs8030236.

Veldman, J.W., Overbeck, G.E., Negreiros, D., Mahy, G., Le Stradic, S., Fernandes, G.W., Durigan, G., Buisson, E., Putz, F.E., Bond, W.J., 2015. Where tree planting and forest expansion are bad for biodiversity and ecosystem services. Bioscience 65, 1011–1018. https://doi.org/10.1093/biosci/biv118.

Wu, J., 2004. Effects of changing scale on landscape pattern analysis: scaling relations. Landsc. Ecol. 19, 125–138. https://doi.org/10.1023/B:LAND.0000021711.40074.ae.

Xie, Y., Sha, Z., Yu, M., 2008. Remote sensing imagery in vegetation mapping: a review. J. Plant Ecol. 1, 9–23. https://doi.org/10.1093/jpe/rtm005.

Xu, Y., Yu, L., Zhao, F.R., Cai, X., Zhao, J., Lu, H., Gong, P., 2018. Tracking annual cropland changes from 1984 to 2016 using time-series Landsat images with a change-detection and post-classification approach: experiments from three sites in Africa. Remote Sens. Environ. 218, 13–31. https://doi.org/10.1016/j.rse.2018.09.008.

Yin, H., 2019. Landsat Sentinel Time Series Viewer. https://github.com/hyinhe/GEE-codes.

Yu, L., Shi, Y., Gong, P., 2015. Land cover mapping and data availability in critical terrestrial ecoregions: a global perspective with Landsat thematic mapper and enhanced thematic mapper plus data. Biol. Conserv. 190, 34–42. https://doi.org/10.1016/j.biocon.2015.05.009.

Zhu, Z., 2017. Change detection using landsat time series: a review of frequencies, preprocessing, algorithms, and applications. ISPRS J. Photogrammetry Remote Sens. 130, 370–384. https://doi.org/10.1016/j.isprsjprs.2017.06.013.