Article:

Wang, Y, Ziv, G orcid.org/0000-0002-6776-0763, Adami, M et al. (8 more authors) (2019) Mapping tropical disturbed forests using multi-decadal 30 m optical satellite imagery. Remote Sensing of Environment, 221. pp. 474-488. ISSN 0034-4257

https://doi.org/10.1016/j.rse.2018.11.028

Mapping tropical disturbed forests using multi-decadal 30 m optical satellite imagery

Yunxia Wang[1], Guy Ziv[1], Marcos Adami[2], Edward Mitchard[3], Sarah A. Batterman[1], Wolfgang Buermann[4], Beatriz Schwantes Marimon[5], Ben Hur Marimon Junior[5], Simone Matias Reis[5], Domingos Rodrigues[6], David Galbraith[1]

1. School of Geography, University of Leeds, Leeds, LS2 9JT, United Kingdom
2. National Institute for Space Research (INPE) – Amazon Regional Center, Av. Perimetral 2651, Belém, PA, Brazil
3. School of Geosciences, University of Edinburgh, Edinburgh, EH8 9YL, United Kingdom
4. School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, United Kingdom
5. Universidade do Estado do Mato Grosso, Campus de Nova Xavantina, Nova Xavantina, MT, 78690-000, Brazil
6. Instituto de Ciencias Naturais, Humanas e Sociais, Sinop, MT, 78577-267, Brazil

* Corresponding author: wangyx.tina@outlook.com

Keywords: Mato Grosso, Forest disturbance, Post-deforestation regrowth, Forest degradation

Abstract

Tropical disturbed forests play an important role in global carbon sequestration due to their rapid post-disturbance biomass accumulation rates. However, the accurate estimation of the carbon sequestration capacity of disturbed forests is still challenging due to large uncertainties in their spatial distribution. Using Google Earth Engine (GEE), we developed a novel approach to map cumulative disturbed forest areas based on the 27-year time-series of Landsat surface reflectance imagery. This approach integrates single date features with temporal characteristics from six time-series trajectories (two Landsat shortwave infrared bands and four vegetation indices) using a random forest machine learning classification algorithm. We demonstrated the feasibility of this method to map disturbed forests in three different forest ecoregions (seasonal, moist and dry forest) in Mato Grosso, Brazil, and found that the overall mapping accuracy was high, ranging from 81.3% for moist forest to 86.1% for seasonal forest. According to our classification, dry forest ecoregion experienced the most severe disturbances with 41% of forests being disturbed by 2010, followed by seasonal forest and moist forest ecoregions. We further separated disturbed forests into degraded old-growth forests and post-deforestation regrowth forests based on an existing post-deforestation land use map (TerraClass) and found that the area of degraded old-growth forests was up to 62% larger than the extent of post-deforestation regrowth forests, with 18% of old-growth forests actually being degraded. Application of this new classification approach to other tropical areas will provide a better constraint on the spatial extent of disturbed forest areas in Tropics and ultimately towards a better understanding of their importance in the global carbon cycle.

1. Introduction

As hotspots of global biodiversity and carbon storage, tropical forests play an important role in biodiversity conservation, climate change mitigation and the provision of multiple other ecosystem services (Foley et al. 2005). However, millions of hectares of tropical forests have been lost due to deforestation and degradation disturbances, resulting in estimated net carbon emissions of 1.4 ± 0.5 Pg yr^{1} from 1990-2010 (Houghton 2012). These emissions represent the second largest anthropogenic source of carbon dioxide to the atmosphere after burning of fossil fuels (van der Werf et al. 2009). In contrast, a significant proportion of previously disturbed tropical forests are regrowing, trapping some of the carbon we are adding to the atmosphere, and with the potential to sequester more in the future. The carbon sink due to tropical forest recovering from deforestation and logging has been reported to be up to 70% greater than that of intact tropical forests (Pan et al. 2011). However, our ability to accurately assess tropical carbon sources or sinks is hampered by the lack of precise information on the extent of disturbed forests in the tropics (Baccini et al. 2017).

Remote sensing has played a key role in identifying forest disturbances and recovery, especially with the recent proliferation of high-resolution satellite data (Hansen et al. 2013). Several approaches have previously been used to map disturbed forests in tropical regions, including optical approaches based on moderate resolution MODIS imagery (Langner et al. 2007), high-resolution Landsat imagery (Lu 2005; Vieira et al. 2003) and very high-resolution SPOT data (Carreiras et al. 2014; Kimes et al. 1999; Souza et al. 2003) , as well as Synthetic Aperture Radar (SAR) (Kuplich 2006; Trisasongko 2010) and Lidar-based approaches (Andersen et al. 2014). However, the majority of these studies have focused on local scales and have been based on single date images. For example, Vieira et al. (2003) classified forests into young, intermediate, advanced and mature forests for one municipality in the state of Pará, using Landsat spectral information and vegetation indices, and found that combining Landsat shortwave infrared band (1.55-1.75 μm) with NDVI generated a better classification than using any individual band/index. Carreiras et al. (2017) further demonstrated the use of combined Landsat spectral bands with ALOS PALSAR backscatter intensity to distinguish secondary regrowth forest and mature forest in three landscapes in Brazilian Amazon. Such multiple multi-sensor fusion approaches have yet to be applied over regional scales.

Several regional satellite-based land cover classifications that include secondary regrowth and forest degradation have become available for Neotropical regions. Two prominent examples are the TerraClass post-deforestation land use/land cover classification (Almeida et al. 2016) and the DEGRAD forest degradation product (INPE 2007-2013) , both of which were developed by Brazilian National Institute for Space Research (INPE) specifically for the Brazilian Amazon. In TerraClass, available since 2004, secondary regrowth forest is mapped on previously deforested areas larger than 6.25 ha using a semi-manual approach (Almeida et al. 2016). The DEGRAD product is produced mainly by visual interpretation of Landsat and CBERS satellite images from a single year and is annually available between 2007 and 2013 (INPE 2007-2013). Recently, another product, MapBiomas, has become available that provides annual national-level land cover and land use maps for Brazil (MapBiomas 2015). MapBiomas, available from 2000 to 2016, classifies forest land cover as dense forest, open forest, secondary forest, degraded forest, flooded forest or mangrove, using an empirical decision tree classification algorithm based on single date spectral mixture analysis. All of those single date imagery based approaches are limited in the discriminatory power they can provide as they make no use of temporal degradation/recovery signals which characterise disturbed forests. Thus, none of the existing products fully exploits the potential of existing Landsat time-series data spanning multiple decades to provide reliable maps of both forest regrowth and degradation. Furthermore, none of these products captures historical (pre-2000) disturbances. There is therefore a clear need for a product that provides a more comprehensive picture of historical disturbances over tropical regions.

Methods that exploit temporal information in satellite data (e.g. threshold approaches, trajectory fitting or segmentation) have been found to be very useful for mapping forest disturbances (Hermosilla et al. 2015; Hirschmugl et al. 2017; Huang et al. 2010; Kayastha et al. 2012; Kennedy et al. 2007; Kennedy et al. 2010; White et al. 2017). However, majority of these time-series based approaches are based on a single time-series trajectory and have mainly been implemented at local scales in extratropical regions (e.g. Canada, U.S.). For example, the recently developed LandTrendr (Kennedy et al. 2010), Vegetation Change Tracker (Huang et al. 2010) and patch-based VeRDET (Vegetation Regeneration and Disturbance Estimates through Time) (Hughes et al. 2017) algorithms have all only been extensively tested in the United States. A recent inter-comparison of disturbance detection algorithms for US forests found that different time-series analysis algorithms are sensitive to different disturbance patterns, with little agreement among these disturbance detection results (Cohen et al. 2017). Thus, when applying these algorithms elsewhere, local calibration and further secondary classification are needed to improve the algorithm's classification performance (Cohen et al. 2018). Machine learning approaches (i.e. random forest) offer the potential to harness the differential sensitivities of different time-series once provided with an appropriate training dataset, but have rarely been coupled with multiple time-series trajectories in Tropics.

In this study, we develop a novel Landsat multiple time-series based classification methodology to map cumulative disturbed forest areas in Tropics, which exploits the power of 1) time-series images relative to single date images, 2) an ensemble of reflectance bands/indices trajectories relative to single trajectories, and 3) machine learning algorithms which enhances classification power by harnessing the differential sensitivities of different time-series. The 'disturbed forests' in this study include both degraded old growth forests and post-deforestation regrowth forests. The former are characterised by a reduction of forest canopy cover (e.g. selective logging, windfall, fire) but have not been clearfelled and thus have not been included in deforestation estimates. The latter refer to areas that have been previously deforested (clearfelled) and converted to other land uses (e.g. pasture, agriculture and mining) but which have subsequently undergone a recovery process following abandonment. Our approach integrates information from six different time-series trajectories (Landsat 5/7 short-wave infrared band 5, band 7, NDVI, SAVI, $NDWI_{2130}$, $NDWI_{1640}$), extracting both statistical and temporal characteristics from each trajectory which then serve as inputs for random forest classification. It not only captures disturbances occurring within study period (1984-2010), but also areas disturbed prior to 1984 which thereafter have exhibited clear recovery patterns. Here, we apply this method to three forest ecoregions (seasonal, moist and dry forests) in the Brazilian state of Mato Grosso.

2. Study Area

Our study area (Fig. **1**), the state of Mato Grosso, is located in the southern edge of Brazilian Legal Amazon. Mato Grosso is the third largest state in Brazil, covering a total area of 903,357 km^2. According to the Terrestrial Ecoregions of the World (TEOW) from World Wildlife Fund (WWF), 43% of Mato Grosso area is covered by Cerrado (tropical savanna), 27% by seasonal forest, 18% by moist forest, 6% by dry forest and 6% by Pantanal (tropical wetlands) (Olson et al. 2001). In Mato Grosso, 139,917 km^2 have been deforested since 1988 (INPE 2017) amounting to 26.5 % of the state's intact forest in that year (Skole and Tucker 1993), most of which has been converted into pasture and agricultural land use due to demand for beef and soy beans (Barona et al. 2010). According to TerraClass (Almeida et al. 2016), herbaceous pasture and shrubby pasture cover 61.4% of the total deforested areas in Mato Grosso while 19.2% of deforested areas are under secondary regrowth (including secondary vegetation and regeneration with pasture). The combination of extensive disturbances and significant amount of remaining intact forest makes Mato Grosso an ideal testbed for the application of our newly developed disturbed forests mapping approach (see section 3).

As indicated, TerraClass is a project that maps land use/land cover on previous deforested areas provided by PRODES (Program for Deforestation Monitoring, INPE 2017) at approximately bi-annual intervals across the Brazilian Legal Amazon (Almeida et al. 2016). TerraClass classifies previously deforested areas into 12 land use categories including pasture, annual crops, secondary vegetation and urban areas. It is extensively validated via field campaigns to determine the accuracy of classification. These have been conducted across different Amazonian regions, including the state of Mato Grosso. This is the best available information on the distribution of secondary forests in any region of the Tropics. However, TerraClass involves a huge effort based largely on visual interpretation and does not map degradation.

The aim of this study is to propose a Landsat multiple time-series based approach in Tropics to 1) improve the efficiency/cost-effectiveness of mapping disturbed forests vs. intact forests, facilitating future TerraClass efforts, 2) map degraded old-growth forests (outside of TerraClass), and 3) eventually enable mapping of disturbed forests over domains for which no reliable data on forest disturbance exist. Only forest areas are considered in this study. To make sure all non-forest areas are excluded, we created a forest cover mask by merging TerraClass-2010 old-growth forest, secondary vegetation and pasture with regeneration categories (Fig. 1). The latter category effectively captures the beginning of the regenerative process containing shrubs and early successional vegetation (Almeida et al. 2016).

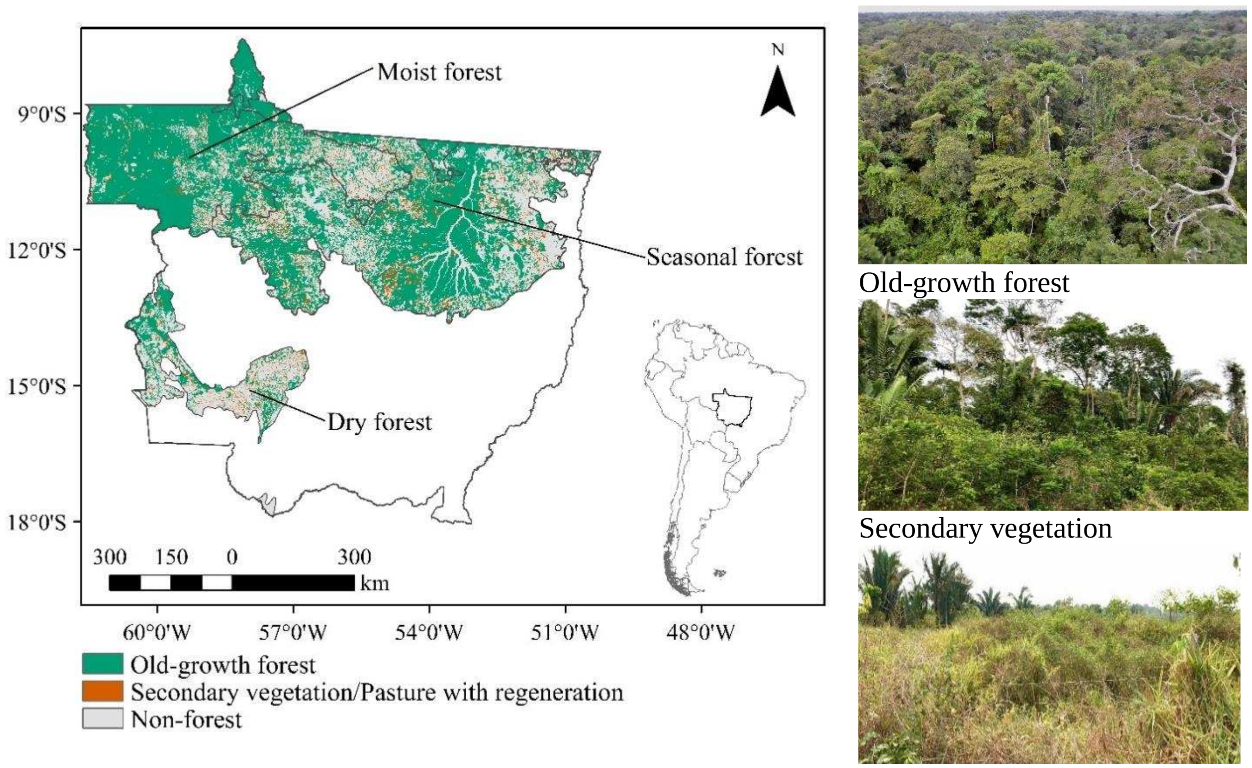


Fig. 1. TerraClass classification map for 2010 (Pasture with regeration in TerraClass is treated as young secondary vegetation). Later, we merged old-growth forest, secondary vegetation and pasture with regeneration into the forest cover mask as the forest boundry. The study area encompasses three WWF forest ecoregions (moist, seasonal and dry forest).

3. Methodology and dataset

The whole approach was developed in Google Earth Engine (GEE) (Gorelick et al. 2017). GEE is a cloud-based geospatial processing platform which consists of over 40 years of historical and current Earth observation imagery, making pixel-based land use and land cover classification feasible across large regions through its inbuilt machine learning algorithms. The overall methodology (Fig. **2**) consisted of building Landsat multiple (six) annual time-series trajectories, calculating trajectory metrics (eleven metrics divided into four groups, Table **2**), generating a training and validation database, applying a machine learning random forest classification algorithm and validating the disturbed forests vs. intact forests classification map, all of which were coded and processed in GEE. We subsequently used the post-deforestation regrowth forest mask generated from TerraClass-2010 to separate the disturbed forests identified through our classification map into post-deforestation regrowth forests and degraded forests (Table **1**). Finally, we performed a relative important analysis of trajectories and trajectory metrics used in the random forest classification to evaluate the extent to which the full suite of all trajectories/metrics enhanced discriminatory power relative to a single trajectory or individual group of trajectory metrics. To do this, ten separate classifications were performed whereby our classification procedure was repeated for each individual trajectory separately (but using all four groups of trajectory metrics), or separately for individual groups of trajectory metrics (but using all six trajectories).

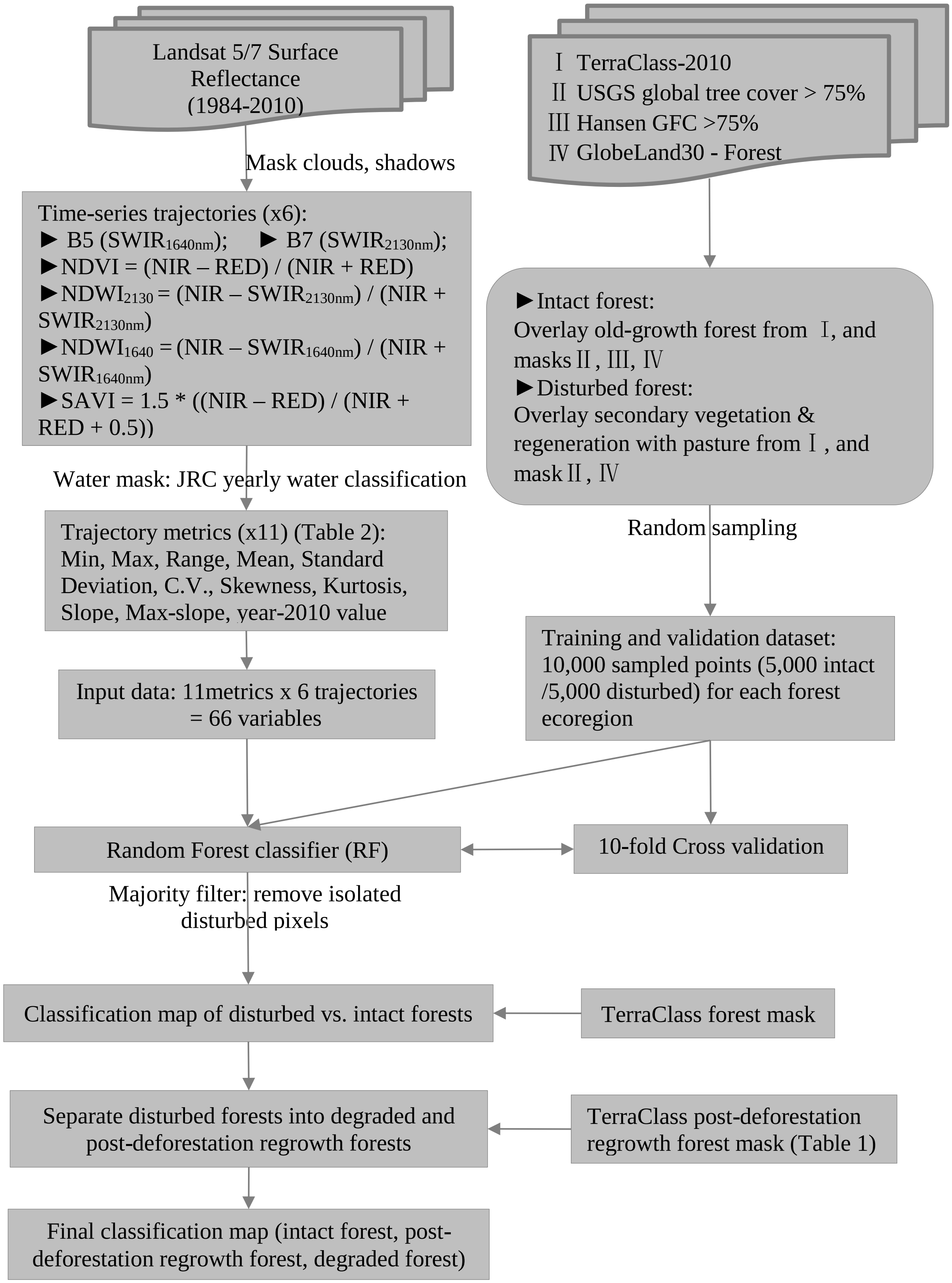


Fig. 2. Classification Methodology for discrimination of disturbed forests and intact forests

Table 1. Classification categories for forested land cover types used in this study.

Categories	Description
Total area	Total area of each ecoregion
Forest cover	Forest mask from TerraClass classification for the year of 2010, combining TerraClass categories of old-growth forest, secondary vegetation and regeneration with pasture.
Intact forest	Forests that have never been experienced any detectable disturbances during 1984-2010. Classified from this study.
Disturbed forest	Cumulative disturbed forest areas during 1984-2010. Classified from this study. Further separated into Post-deforestation regrowth forest & Degraded forest.
Post-deforestation regrowth forest	Areas that have been previously deforested (clearfelled) and converted to other land uses (e.g. pasture, agriculture and mining) but which have subsequently undergone a recovery process following abandonment. Secondary vegetaion or regeneration with pasture in TerraClass-2010.
Degraded forest	Degraded old-growth forests. Characterised by a reduction of forest canopy cover (e.g. selective logging, windfall, fire) but have not been clearfelled and thus have not been included in deforestation estimates.

3.1 Time-series trajectories

3.1.1 Landsat surface reflectance dataset

We used Landsat atmospherically corrected surface reflectance (SR) products (30 m resolution) (Masek et al. 2006; USGS 2018) to generate annual time-series trajectories. All Landsat-5 Thematic Mapper (TM) surface reflectance images aquired during the period of 1984-2010 were used except for 2001 and 2002. In 2001, most images had striping artifacts limiting their use, while in 2002, images from Landsat 5 only covered 61% of our study area. For these reasons, we used Landsat-7 Enhanced Thematic Mapper Plus (ETM+) images, which are compatible in their spectral characteristics (Claverie et al. 2015; Home et al. 2013), for these two years. In terms of spectral bands, we chose spectral bands 3 (red, 0.52 - 0.60 µm) which is sensitive to the amount of chlorophyll, 4 (near-infrared, 0.76 - 0.90 µm) which is related to leaf cellular structure, 5 (shortwave-infrared, 1.55 - 1.75 µm) and 7 (shortwave-infrared, 2.08 - 2.35 µm) which relate to leaf water content (Nelson et al. 2000). To minimize the influence of variable extent of rivers on the classification, we excluded water bodies in our analysis using the Joint Research Center (JRC) Yearly Water Classification History v1.0 product. This dataset contains maps of the location and temporal distribution of surface water from 1984 to 2015 at annual resolution, generated using more than three million scenes from Landsat 5, 7 and 8 (Pekel et al. 2016).

3.1.2 Generating time-series trajectories

We processed 11,483 images in total for our entire study period (1984-2010), ranging from 257 to 715 annual images depending on data availablity, with annual spatial coverage of 99% of our study area (see Table S1 in supplementary information). Five steps were involved to process the Landsat SR data and produce time-series image stacks for 1984-2010. First, areas covered by clouds and cloud shadows were removed based on the pixel quality and radiometric saturation attributes of the Landsat surface reflectance product. Second, original surface reflectance (16-bit signed integer) values were converted to 0-1 range values by multiplying by the scale factor of 0.0001. Third, four vegetation indices (VIs) were calculated including the Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index ($NDWI_{2130}$, $NDWI_{1640}$) (Chen et al. 2005) and Soil-Adjusted Vegetation Index (SAVI) (Huete 1988). Fourth, to minimise the influence of cloud contamination and improve the quality of input data, we selected the maximum value of individual VIs for each year (Maxwell and Sylvester 2012). For time-series of reflectance from spectral bands 5 and 7, median values were calculated for each year. In the final step, we used the JRC Yearly Water Classification History v1.0 product to mask water areas (Pekel et al. 2016). After processing, annual time-series trajectories (1984-2010) of Landsat SR spectral band 5 (1.55 - 1.75 μm), band 7 (2.08 - 2.35 μm), NDVI, $NDWI_{2130}$, $NDWI_{1640}$ and SAVI were used for the classification of disturbed forests and intact forests.

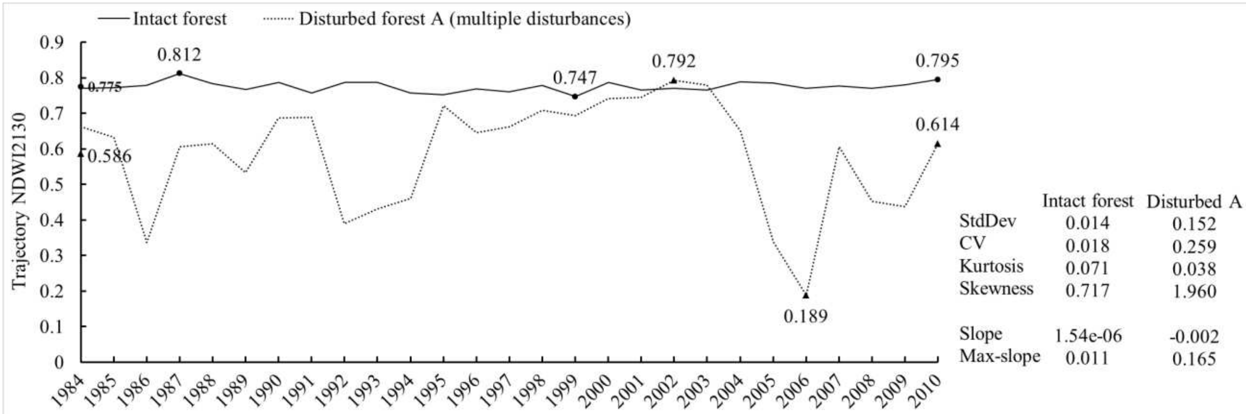
3.2 Trajectory metrics

We calculated eleven metrics divided into four groups (Table 2) for each of the six spectral trajectories to act as inputs for random forest algorithm (see section 3.4), based on a priori expectations of divergence between intact and disturbed forests. Each of these 11 metrics may capture information that is linked to a particular disturbance type. For example, the coefficient of variation (C.V.) shows the extent of variability in relation to the mean. Forests which have experienced large disturbances would be expected to have higher C.V. than undisturbed intact forests. We further hypothesized that time-series trajectories of intact forest would follow a normal distribution, while those of disturbed forest would tend not to and be much more likely to exhibit greater skewness and kurtosis. Finally, trends (based on linear regressions) were also estimated from the time-series trajectories. We hypothesized that disturbance events would likely result in either decreasing (deforestation/degradation) or increasing (regrowth) trends over time, and thus expected that the regression slopes of disturbed pixels would be much smaller/greater than undisturbed pixels where we expected that the slope value is close to zero. It has been found that regrowth secondary forests in Amazonia are cut and burned on average every 5 years (Aguiar et al. 2016). Thus, we also considered the maximum absolute regression slopes derived from individual 5-year windows within the 1984-2010 study period.

Fig. 3 demonstrates differences in trajectories and trajectory metrics between intact and disturbed forest pixels. For intact forests (undisturbed during 1984-2010), we expected trajectories to fluctuate, but to follow a normal distribution pattern, while trajectories of disturbed forests were expected to exhibit more pronounced decrease and increase patterns. Trajectories of disturbed forest pixels' can follow various patterns, depending on whether they have been disturbed once (Fig. 3 Disturbed B) or multiple times (Fig. 3 Disturbed A) within the study period (1984-2010) or disturbed before 1984 but following a clear recovery pattern within study period (Fig. 3 Disturbed C).

Table 2. Metrics for each time-series trajectory and related main GEE algorithms. The metrics were divided into location, scale, temporal and single year groups which were further used for metric important analysis (see section 4.4).

Group	Name	Description	Main GEE algorithm
Location metrics	Min	Minimum of time-series	ee.Reducer.min()
	Max	Maximum of time-series	ee.Reducer.max()
	Range	The range between maximum and minimum of time-series	Code equation ‘max-min’
	Mean	The mean of time-series	ee.Reducer.mean()
Scale metrics	StdDev	Standard deviation of time-series	ee.Reducer.stdDev()
	C.V.	Coefficient of variation of time-series	Code equation ‘mean/stdDev’
	Kurtosis	Dispersion measure related to the tails of Normality distribution test (D'Agostino 1970, see methods)	Code equations based on the reference
	Skewness	Symmetry measure related to Normality distribution test (D'Agostino 1970, see methods)	Code equations based on the reference
Temporal metrics	Slope	Linear regression slope of total time-series	ee.Reducer.linearFit()
	Max-slope	Maximum linear regression slope of every 5-year window	Function of 5-year window; ee.Reducer.linearFit(); ee.Reducer.max()
Single year	Year-2010	Time-series trajectory value at year 2010	‘FilterMetadata’ equals 2010



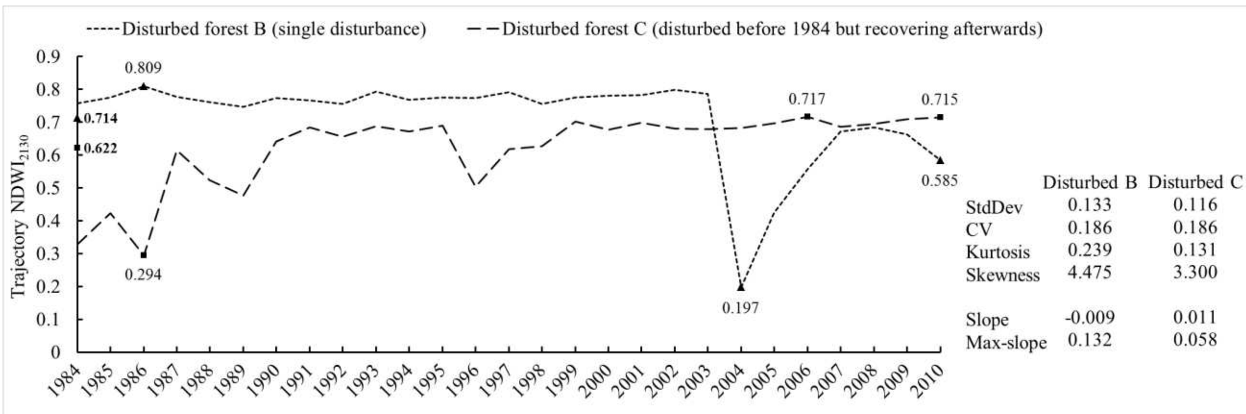


Fig. 3. Examples ($NDWI_{2130}$) of time-series trajectories for illustrative intact forest pixel and disturbed forest pixels. Values of trajectory scale and temporal metrics extracted from each trajectory (Table 2) are shown to the right of the graph. Metrics of max, min and year-2010 value are shown on the trajectory with the mean marked on y axis.

3.3 Sampling design

We used GEE random sampling to generate a set of spatially representative points of disturbed and intact forests for classification training and validation based on TerraClass-2010 map of old-growth forest, secondary vegetation and pasture with regeneration, USGS (United States Geological Survey) 30 m Global Tree Cover 2010 (Hansen et al. 2013), the Hansen Global Forest Change (GFC) product (Hansen et al. 2013), and 30 m Global Land Cover 2010 (GlobeLand30-2010) produced by National Geomatics Centre of China (Chen et al. 2015). Since TerraClass uses deforestation vector data from PRODES (INPE 2017) as input data to map subsequent land use/covers (Almeida et al. 2016), it inherited PRODES historical misalignment issues. To better align TerraClass with GFC products, we registered the TerraClass-2010 classification map using the GEE image displacement algorithm by calculating the displacement between TerraClass-2010 forest mask and GFC forest mask (Hansen et al. 2013).

For intact forests, points were randomly sampled from areas that met the following conditions: i) classified as old-growth forest in TerraClass-2010; ii) tree canopy cover > 75% in GFC in 2000 and no forest loss during 2000-2010; iii) tree cover >75% in USGS 30 m Global Tree Cover 2010; and, iv) classified as forest in GlobeLand30-2010. Similarly, disturbed forest pixels were sampled from areas that satisfied the following conditions: i) classified as secondary vegetation or regeneration with pasture in TerraClass-2010; ii) tree cover > 75% in USGS 30 m Global Tree Cover 2010; iii) classified as forest in GlobeLand30-2010. To reduce the influence of unwanted positional errors among these land cover products and avoid edge effects, we required that both intact forest and disturbed forest sampled points were located at least 100m away from the patch boundary. For each forest ecoregion (moist/seasonal/dry forest), 10000 points (5000 intact and 5000 disturbed) were randomly sampled, respectively. In total, we sampled 30000 intact and disturbed points across the study area as the training and validation database.

3.4 Random forest classifier

Mapping of disturbed forests was performed by using the GEE Random Forest classifier algorithm, which has been recently successfully applied to cropland mapping (Shelestov et al. 2017; Xiong et al. 2017), oil palm plantation detection (Lee et al. 2016), mapping urban settlement and population (Patel et al. 2015) and soil mapping (Padarian et al. 2015). Random Forest (RF) classification is a relatively well-known supervised machine leaning algorithm that iteratively produces an ensemble of decision tree classifications by using corresponding randomly selected subsets of the training dataset (Breiman 2001). It grows classification trees by splitting each node using a random selection subset of input variables, which reduces overfitting and yields a more robust classification compared to other classifiers (Breiman 2001). RF uses a voting system to classify data and the final classification category for each pixel is determined by the plurality vote of all trees generated to build the forest.

We used 66 variables comprising 11 metrics (Table 2) for each of the six time-series trajectories as input predictors for the RF classification. RF classifications were applied in moist, seasonal and dry forest ecoregions, respectively. All classifications were based on the outputs of 500 decision trees (See Fig. S1 in supplementary information). Each tree split was based on eight variables randomly selected from all 66 input variables, which was the default configuration for the GEE random forest classifier. After constructing our disturbed forest classification, we performed a post-classification filtering to reduce noise and remove spurious classification artefacts by applying a 90m x 90m majority filter.

3.5 Classification validation

To evaluate how well our classification performed, we used ten-fold cross-validation (Kohavi 1995; Schaffer 1993) based on above randomly sampled database (See section 3.3, i.e. 10000 points for each forest ecoregion), which randomly partitions our sampled database into ten equal sized subsets. Of the ten subsets, a single subset (1000 points) was retained as the validation data for testing the classification algorithm, and the remaining nine subsets (9000 points) were used as training data for RF classifier. The cross-validation process was repeated ten times. The final accuracy estimation was determined by the average of ten-fold results. The accuracy matrix included overall accuracy (OA), producer's accuracy (PA), user's accuracy (UA) and Kappa statistic (Kohavi 1995).

For an additional independent confirmation for our Landsat optical sensor based classification of disturbed forests vs. intact forests, we used another microwave radar based satellite product, ALOS/PALSAR 25 m spatial resolution mosaic imagery, as visual interpretation. ALOS PALSAR imagery consists of dual polarization HH (transmission of horizontal wave and reception of horizontal component) and HV (horizontal transmission and vertical reception), but it has been shown that the polarization mode HV is more effective in deforestation detection than HH polarization (Motohka et al. 2014), which corresponds with findings of close relations between HV backscatter and vegetation structural properties (e.g. forest height, forest cover) (Joshi et al. 2015). Thus, we visually compared the 2007-2010 ALOS/PALSAR HV backscatter change with our final classification results.

SAR data are stored as digital number (DN) in unsigned 16 bit and typified by a high degree of speckles in the image (random ‘salt and pepper’ noise). To reduce noise and improve image interpretability, a multi-temporal speckle filter (7×7) (Lee 1980; Lopes et al. 1990) was implemented in GEE and applied to 2007-2010 PALSAR images, without significant loss of spatial resolution. Filtered ALOS/PALSAR HV backscatter DN values were converted to sigma-naught (σ^0) in decibel (dB) units using the following equation:

$$\sigma^0 = 10 * log_{10}(DN^2) - 83 \tag{1}$$

σ^0 is generally negative and can vary from -35 dB in very low backscatter areas (degraded/deforested area), up to 0 dB for extremely high backscatter (dense forest area). For visual interpretation, we expected a decrease or an increase in σ^0 in forest areas that have been recently disturbed or are recovering from past disturbances (Joshi et al. 2015). However, we also expected that many disturbed areas in our classification would not be captured by PALSAR due to its short time period (2007-2010).

4. Results

4.1 Classification results

As represented in Fig. 2, the new developed disturbed forests vs. intact forests classification approach was applied to three different ecoregions in Mato Grosso. The final classification map (Fig. 4) was generated by training the random forest classifier individually for each ecoregion on the entire sampled database. Our classification results representative of the year 2010 show that disturbed forests (both post-deforestation regrowth forests and degraded forests) were widely spread across Mato Grosso, but were most prevalent along rivers and next to non-forest areas (Fig. 4). Forests in Mato Grosso covered a total area of 295,383 km^2 in 2010 (Table **3**), accounting for about 63% of the total study area. Our results show that, until 2010, 25% of the total forested area was disturbed (Table **3**). Forest cover percentage varied considerably across ecoregions, ranging from 37% in dry forest to 74% in moist forest (Table **3**). Dry forest experienced the most severe disturbances with 41% of forest cover classified as disturbed, followed by seasonal forest and moist forest where disturbed forests accounted for 28% and 20% of forest cover, respectively (Table **3**).

We further separated disturbed forests identified through our classification map into post-deforestation regrowth forests and degraded forests. It shows that the area of degraded forests was up to 62% larger than the area of post-deforestation regrowth forests across ecoregions, with degraded forests and post-deforestation regrowth forests covering a total area of 47,039 km^2 and 28,246 km^2, respectively (Table **4**). By comparing degraded forests and old-growth forests classified in TerraClass for the year of 2010, we found that 18% of areas identified as old-growth forests in TerraClass were actually degraded forests, ranging from 15% to 27% across various ecoregions (Table **4**).

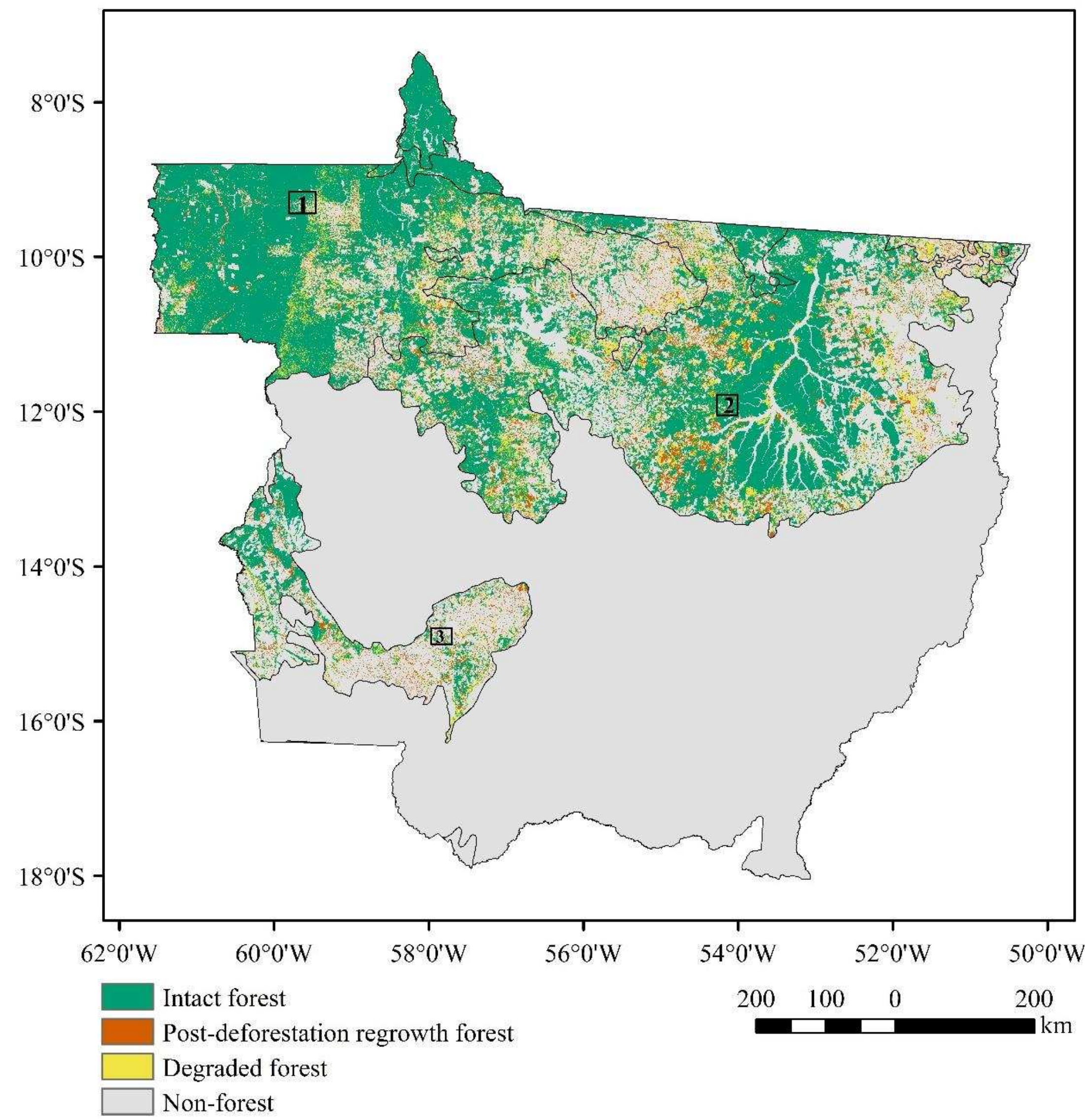


Fig. 4. Classification map of intact forest, post-deforestation regrowth and degraded forest representative of the year 2010. Non-forest areas include areas under anthropogenic use or natural savannahs/wetlands. Small areas 1 to 3 represent three focal regions within individual ecoregions, for which subsequent fine-scale visual interpretation confirmation were performed (Fig. 5-7).

Table 3. Areal extent (in km^2) of intact forest and historically disturbed forest representative of 2010.

	Moist forest	Seasonal forest	Dry forest	**Total**
Total area	170,154	245,514	54,454	**470,122**
Forest cover (% of total area)	125,474 (73.74%)	149,571 (60.92%)	20,338 (37.35%)	**295,383** **(62.83%)**
Intact forest (% of forest cover)	100,050 (79.74%)	107,991 (72.20%)	12,058 (59.29%)	**220,099** **(74.51%)**
Disturbed forest (% of forest cover)	25,424 (20.26%)	41,581 (27.80%)	8,280 (40.71%)	**75,285** **(25.49%)**

Table 4. Areal extent (km^2) of post-deforestation regrowth forest and degraded forest representative of 2010.

	Moist forest	Seasonal forest	Dry forest	**Total**
Post-deforestation regrowth (% of disturbed forest)	8,188 (32.21%)	15,950 (38.36%)	4,108 (49.62%)	**28,246** **(37.52%)**
Degraded forest (% of disturbed forest)	17,236 (67.79%)	25,631 (61.64%)	4,171 (50.38%)	**47,039** **(62.48%)**
TerraClass old-growth forest	116,226	131,703	15,622	**263, 551**
% of degraded forest within TerraClass	14.83%	19.46%	26.70%	**17.85%**

4.2 Ten-fold cross validation

Ten-fold cross validation was used as the main validation of our disturbed forests and intact forests classification map, with accuracy matrices provided in Table **5**. Overall, all the classification accuracies were above 80% with Kappa agreements above 62%. Across ecoregions, the overall accuracy was the highest in seasonal forest at 86.1%, with a producer's accuracy of 88.9% for intact forests and 83.3% for disturbed forests. In moist forest and dry forest regions, the overall accuracies were lower at 81.3% and 82.6%, respectively.

Table 5. Ten-fold cross validation accuracy based on sampled points from our study.

Regions	Overall accuracy	Producer's accuracy		User's accuracy		Kappa statistic
		Intact forest	Disturbed forest	Intact forest	Disturbed forest	
Moist forest	0.813	0.888	0.737	0.772	0.867	0.625
Seasonal forest	0.861	0.889	0.833	0.842	0.882	0.722
Dry forest	0.826	0.856	0.797	0.809	0.846	0.653

4.3 High-resolution image interpretation

To further validate our classification, we consider in detail one landscape within each biome, comparing our results to radar and other very high-resolution data. Examples in Fig. **5**-7 allow for visual comparison of our classification in selected focal areas within each forest ecoregion with corresponding ALOS PALSAR HV backscatter (σ^0) temporal (2007-2010) change composite images and very high-resolution (5 m) RapidEye true-colour composite images (Team 2017). Overall, this comparison at local scales shows a very good visual agreement between our classification and the PALSAR temporal change as well as with RapidEye images across ecoregions (Fig. **5**-7), especially those logging roads shown in Fig. **6**. As expected, there were some mismatches between our classification and the temporal change in PALSAR HV σ^0, such as several disturbed areas from our classification not appearing in PALSAR temporal change image. This is likely due to PALSAR images only being available from 2007 and thus not capturing much forests disturbed before 2007.

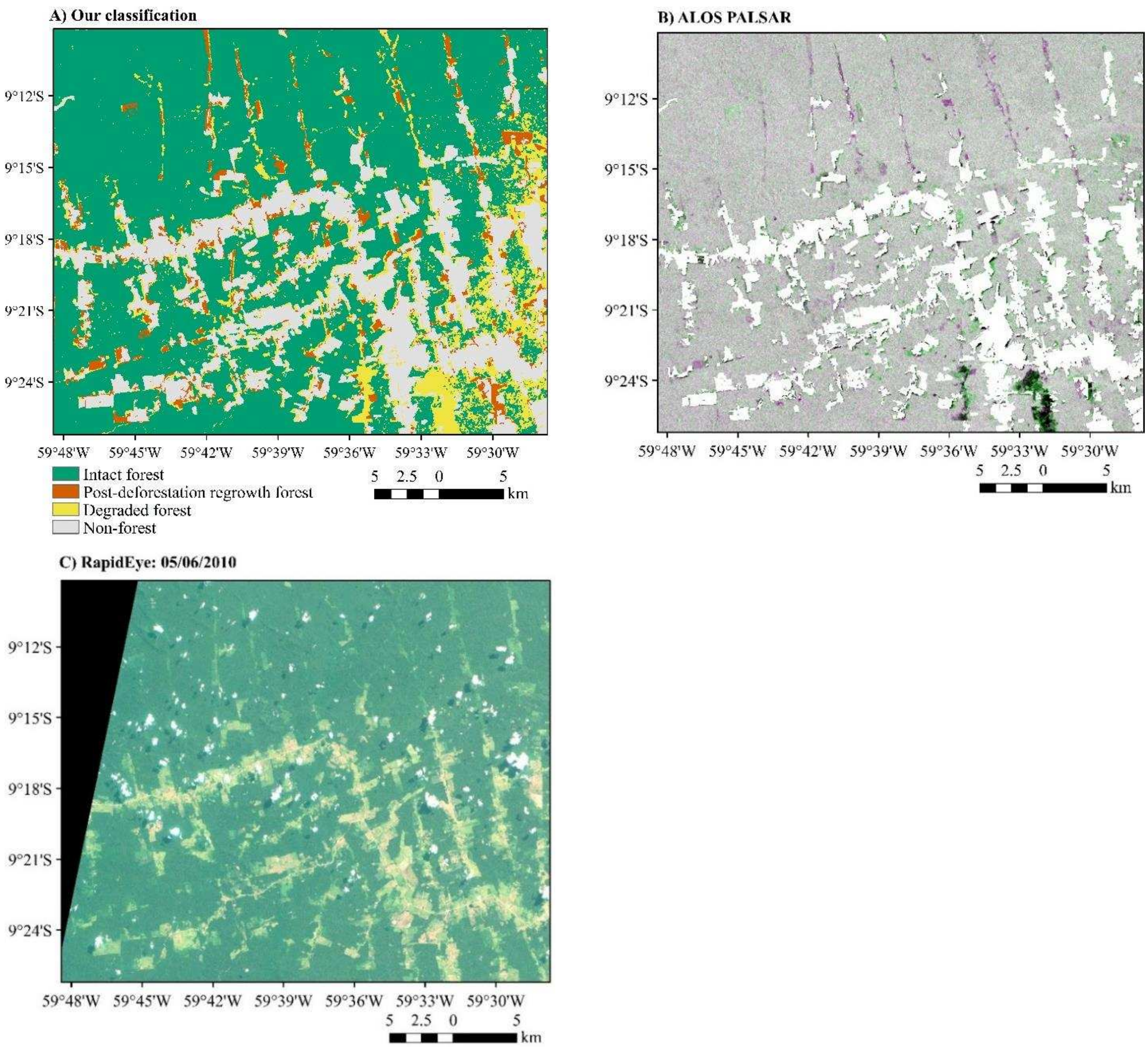


Fig. 5. Moist forest focal region (area 1 in Fig. 4). A) Detailed classification map. B) Forest masked ALOS PALSAR HV σ^0 temporal change, pink represents increase of σ^0, green represents decrease of σ^0 between 2007-2010, grey represents little/no change between 2007-2010, white areas are non-forest. C) RapidEye true-colour composite image (See Fig. S2 in supplementary information for better visualization).

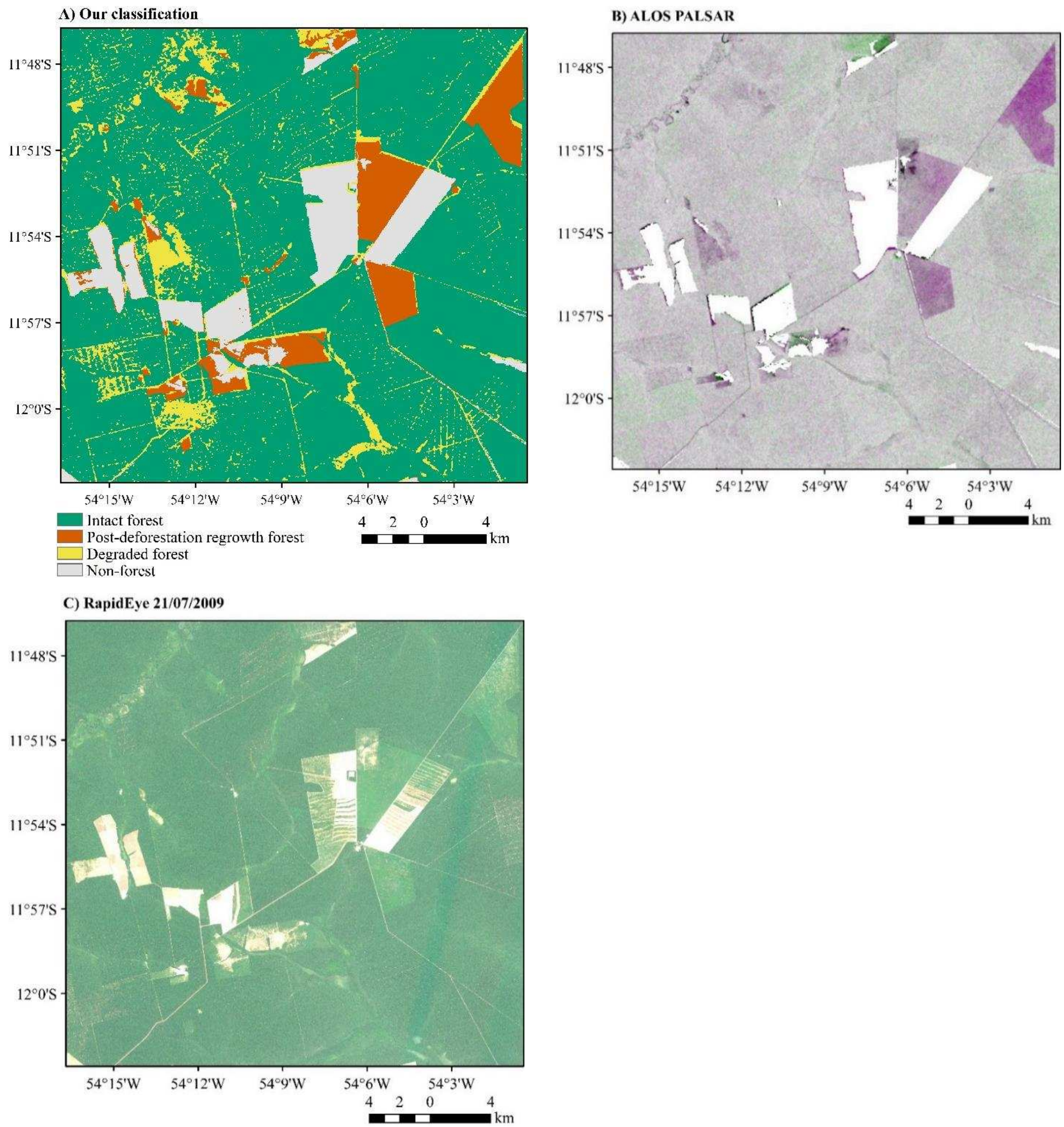


Fig. 6. Seasonal forest focal region (area 2 in Fig. 4). A) Detailed classification map. B) Forest masked ALOS PALSAR HV σ^0 temporal change, pink represents increase of σ^0, green represents decrease of σ^0 between 2007-2010, grey represents little/no change between 2007-2010, white areas are non-forest. C) RapidEye true-colour composite image (See Fig. S3 in supplementary information for better visualization).

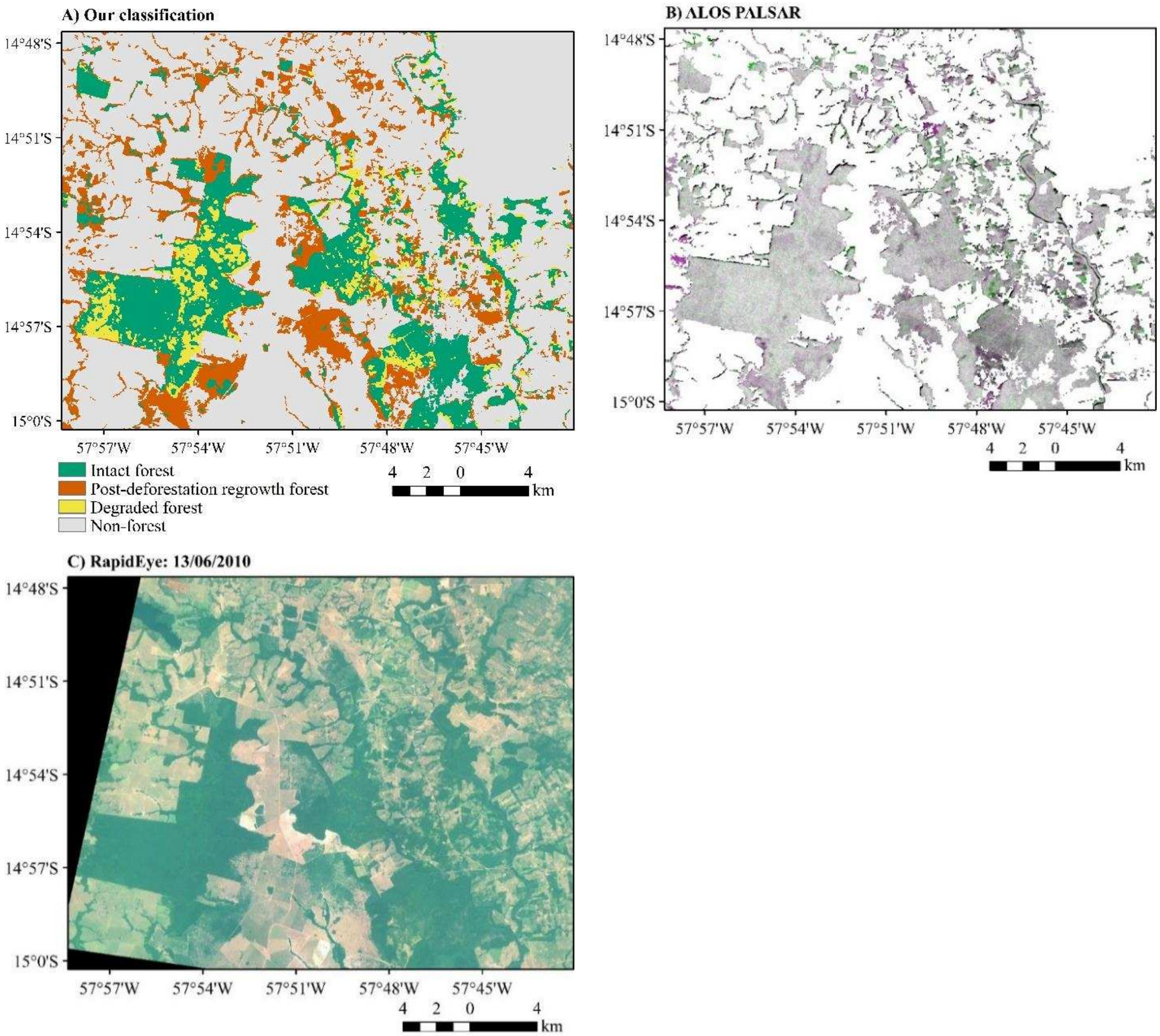


Fig. 7. Dry forest focal region (area 3 in Fig. 4). A) Detailed classification map. B) ALOS PALSAR HV σ^0 temporal change, pink represents increase of σ^0, green represents decrease of σ^0 between 2007-2010, grey represents little/no change between 2007-2010, white areas are non-forest. C) RapidEye true-colour composite image (See Fig. S4 in supplementary information for better visualization)**.**

4.4 Importance of individual trajectories and metrics

The relative importance of individual trajectories in our classification was measured by the percentage of overall accuracy change (% OAC) when running our classification for a single trajectory (but using all four groups of trajectory metrics) relative to our full suite multi-trajectory classification (Table 5). The larger the overall accuracy change, the less important an individual trajectory is in distinguishing the differences between disturbed forests and intact forests. All of the single time-series trajectories based classifications had much lower (3-15% across ecoregions) overall classification accuracy than our full suite classification (Fig. 8). In moist forest and dry forest ecoregions, Landsat shortwave spectral band 5 and 7 were the most important trajectories for distinguishing disturbed forests and intact forests, decreasing %OAC the least relative to our full suite classification. However, in the seasonal forest ecoregion, NDWI trajectories were the most important, decreasing the overall accuracy the least, followed by spectral band 7.

The important of specific groups of trajectory metrics (Table 2) was determined in an analogous manner to the importance of specific trajectories. Importance patterns for groups of metrics were similar across ecoregions (Fig. 8B), with location metrics being the most important in distinguishing disturbed and intact forests, followed by temporal metrics, scale metrics and single year (2010) values. However, single year (2010) values alone were found to have much less discriminatory power than other metrics, resulting in much lower (up to 20%) classification accuracy relative to our full suite classification with all groups of metrics included (Fig. 8B).

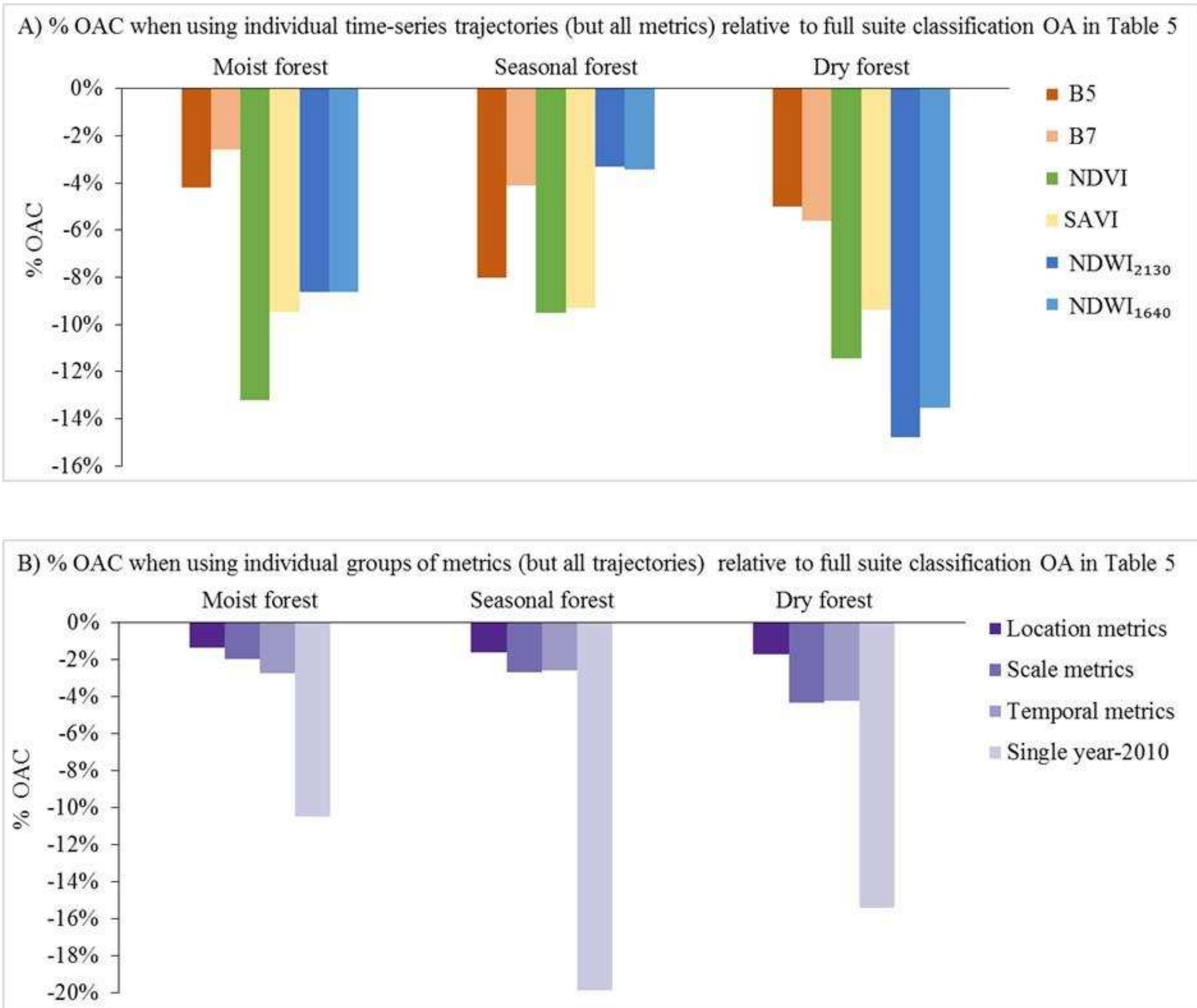


Fig. 8. The percentage of overall accuracy change (% OAC) when running our classification procedure for individual trajectories separately (but using all four groups of trajectory metrics) or separately for individual groups of trajectory metrics (but using all six trajectories) relative to our full suite classification with all trajectories/metrics included (Table 5). The larger the absolute % OAC, the less important the particular trajectory (or the group of trajectory metrics) is.

4.5 Comparing with other products

We compared our classification of disturbed forests in Mato Grosso with other relevant products which have recently become available (Fig. **9Error! Reference source not found.**). These include the MapBiomas land use/cover products (2000-2010) and the Latin American secondary forest map recently produced by Chazdon et al. (2016). The latter was derived from the map of Neotropical forest aboveground biomass of Baccini et al. (2012) for 2008. To ensure comparability in time, we only compared disturbed forests from our classification against the area of secondary forests < 24 years old from Chazdon et al. (2016). To compare against MapBiomas products (2000-2010), we reclassified open forest, degraded forest, secondary forest, and flooded forest categories from MapBiomas-2010 map into one disturbed forest class. Areas classified as non-dense forest in 2000-2009 MapBiomass products but classified as dense forest in 2010 were also considered as disturbed forests.

Our estimate of disturbed forest area in Mato Grosso was three times larger than disturbed forests from MapBiomas with corresponding spatial distribution shown in Fig. 9 (A&B). The biggest classification differences was located in moist forest ecoregion, followed by seasonal forest and dry forest. The difference relative to MapBiomas may be due to the use of different classification methods (single date based classification) and the limited time period (2000-2010) for MapBiomas. However, secondary forest area estimates from Chazdon et al. (2016) were approximately three times greater than the disturbed area from our classification (Fig. 9C), increasing to four times greater in the dry forest biome. This may be due to the coarse resolution (500 m) of forest age map, the misclassification of some anthropogenic land use areas as forest or to errors arising from interpreting the age from the forest biomass map (Chazdon et al. 2016).

The large discrepancies of estimated disturbed forests among those products highlight the importance of using high-resolution time-series images and the consideration of historical disturbances when mapping secondary forest regrowth and forest degradation. By excluding pre-2000 historical disturbances and ignoring time-series spectral characteristics, MapBiomas significantly underestimate the area of disturbed forests (Fig. 9B), and correspondingly may underestimate the impacts of disturbance on tropical biodiversity and carbon cycles.

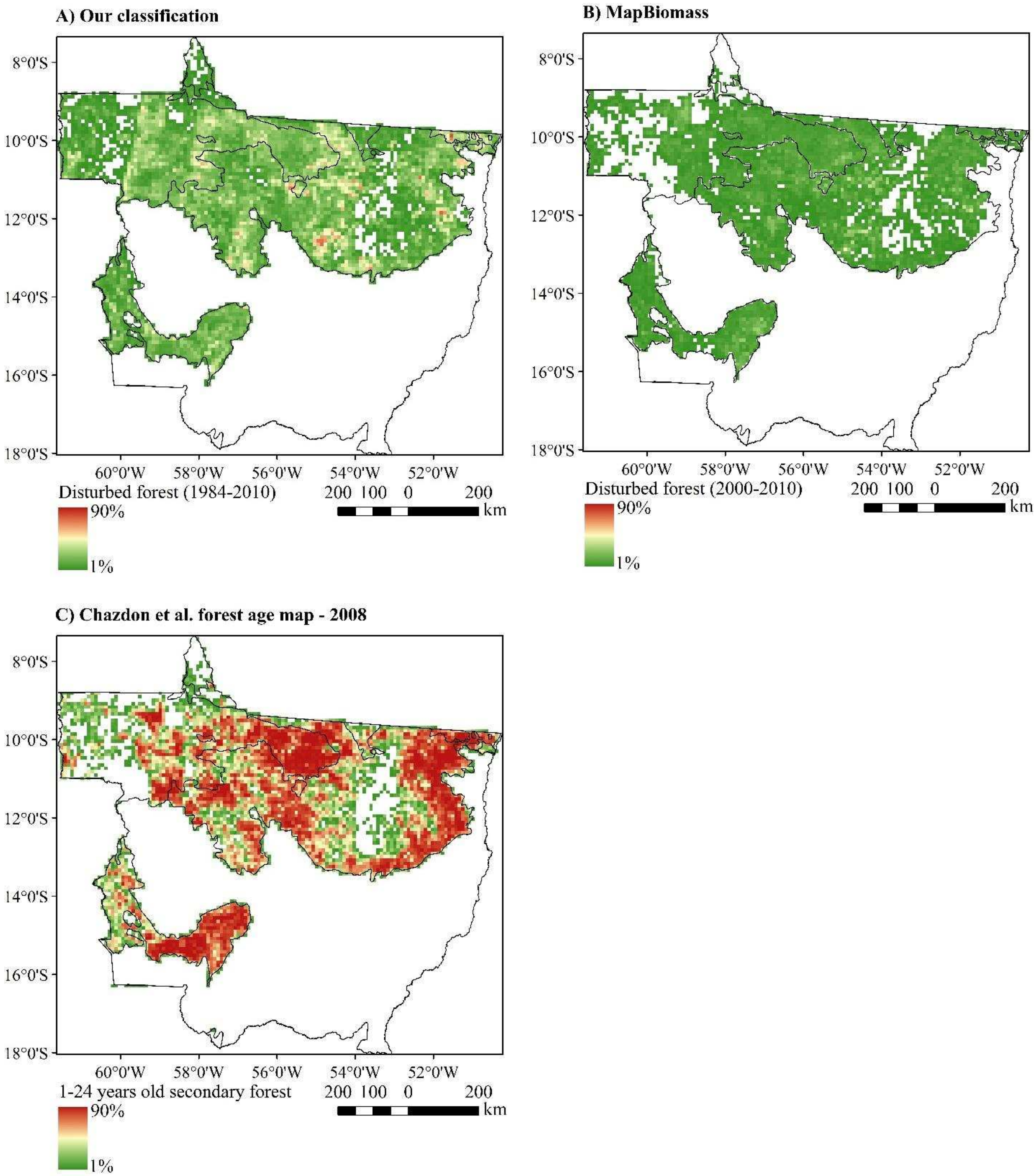


Fig. 9. Comparison of our classification with MapBiomas land use/cover 2000-2010, and Chadzon et al. 2008 secondary forest age map. Values represent the percentage of the area of disturbed forests within each grid cell (10*10km). White areas (within study area) represent no disturbed pixels were identified within that grid cell. The disturbed areas are 75285 km^2, 24577 km^2, 246829 km^2 for figure panel A, B, C, respectively.

5. Discussion

In this study, we developed a new time-series approach in GEE to map disturbed forests (both forest degradation and post-deforestation regrowth) and intact forests. This approach incorporates random forest machine learning algorithm with multiple Landsat time-series trajectories, which enhances classification power by harnessing differential sensitivities of different time-series. It is flexible with respect to the disturbance patterns it captures. It detects three different disturbances trends (Fig. 3): 1) single disturbance – time-series have a decrease then increase pattern; 2) multiple disturbances – time-series have multiple increase and decrease signatures pattern; 3) recovery on previous disturbed areas – time-series only have an increase pattern. For example, in this study, it not only maps areas that disturbed and recovering during time-series period (1984-2010), but also captures areas that disturbed before 1984 but following a recovery process after 1984, making our approach more valuable and suitable for distinguishing disturbed forests and intact forests.

Application of our approach in moist/seasonal/dry ecoregions in Mato Gross resulted in high overall classification accuracy, ranging from 81.3% to 86.1% across ecoregions. On one hand, the misclassification of disturbed forests as intact forests may relate to the fast recovery process of secondary regrowth forests whose structural and spectral characteristics could be similar to intact forests after 20-40 years recovery (Aide et al. 2000; Poorter et al. 2016). The degraded old-growth forests recover at even faster rates. For example, it has been shown that about 50% of the canopy opening caused by selective logging becomes closed within one year of regrowth (Asner et al. 2004), making it harder to capture such quick recovery process from remote sensing perspectives. On the other hand, the misclassification of intact forests as disturbed might be because of our sampling of intact forests points which may still include few disturbed old-growth forests, as TerraClass does not map degraded forests. Furthermore, the variation of classification accuracy across ecoregions might be due to the differences of land-use history, land use intensity, severity of disturbance events, soil fertility and texture (Chazdon 2003) and water availability (Poorter et al. 2016), which are highly associated with post-disturbance recovery processes and the structure of regrowth forests.

By separating disturbed forests into post-deforestation regrowth forests and degraded forests, we found that approximately two-thirds of disturbed forests were degraded forests, highlighting the importance of effective systems for detecting these. Forest monitoring system should not only focus on clear-cut forest deforestation and recovery, but also degraded forests which may release more than double the amount of carbon than released by deforestation (Baccini et al. 2017). Interestingly, our classification clearly captured straight-line patterns of disturbed forests, which also present a consistent agreement with both PALSAR HV backscatter intensity change and RapidEye very high resolution images (Fig. 6). Further development of our methodology may provide new opportunities to map selective logging activities at a large regional scale.

The methodology developed in this study dramatically exploits the power of multiple long-term Landsat time-series in the discrimination of disturbed vs. intact forests with support of GEE's massive storage and calculation capability. Unlike previously published single time-series trajectory based approaches (e.g. LandTrendr, VCT, VeRDET) (Cohen et al. 2017), this approach incorporates six different time-series trajectories which generates a much higher classification accuracy than single-trajectory based classification (Fig. 8A). Also, this approach integrates single year features with scale, location and temporal characteristics derived from time-series trajectories, which significantly enhanced the discriminatory power. Single year features were found to be the least powerful (up to 20% less) for discriminating disturbed pixels compared to the combined use of single year features and other time-series features (Fig. 8B). Thus, combination of single year and time-series features represents a significant advance on widespread single-year approaches to map previously disturbed forests.

6. Conclusion

Our study explored the feasibility of using multiple long time-series Landsat surface reflectance data to map tropical historically disturbed forests as far back as 1984. Using a case study of Mato Grosso moist, seasonal and dry forests, we found that this methodology has high potential in mapping various forested land cover types related to disturbances with an overall accuracy of up to 86.1%. The classification approach developed in this study is capable of capturing not only forest regrowth from forest deforestation (clear-cut), but also forest degradation (partially cut) due to selective logging or other small scale disturbances. Based on TerraClass-2010 forest mask, until 2010, 41% dry forest in Mato Grosso were disturbed, with 28% and 20% of seasonal forest and moist forest disturbed, respectively. By comparing classification from this study with TerraClass-2010 land cover map, we found that up to 18% of area classified as old-growth forest in TerraClass was actually degraded forests, highlighting the importance of including degradation monitoring alongside clear felling monitoring .

Our study clearly demonstrates the potential of extensive time-series of satellite imagery to map historical forest disturbances and recovery processes. More specifically, the discrimination of disturbed forests (both degraded forest and post-deforestation regrowth forest) vs. intact forests was enhanced by simultaneously combining a suite of single date features and time-series characteristics derived from multiple time series of spectral bands and vegetation indices. Our approach is readily applicable to other larger tropical areas, making pan-tropical mapping of forest disturbances and regrowth a highly tangible prospect.

Acknowledgements

This work was carried out with the support of Google Earth Engine Research Award (GZ, DG, WB). YW acknowledges support from the China Scholarship Council (201506300051) and the University of Leeds. GZ acknowledges support from ECOPOTENTIAL project funded under EU's Horizon 2020 programme grant agreement 641762. DG acknowledges support from the NERC-funded TREMOR project (NE/N004655/1). SB acknowledges support from NERC independent research fellowship NE/M019497/1. We thank Dr. Timothy R. Baker from University of Leeds for contributing to the development of ideas for this paper.

References

Aguiar, A.P.D., Vieira, I.C.G., Assis, T.O., Dalla-Nora, E.L., Toledo, P.M., Oliveira Santos-Junior, R.A., Batistella, M., Coelho, A.S., Savaget, E.K., & Aragão, L.E.O.C. (2016). Land use change emission scenarios: anticipating a forest transition process in the Brazilian Amazon. Glob Chang Biol, 22, 1821-1840

Aide, T.M., Zimmerman, J.K., Pascarella, J.B., Rivera, L., & Marcano-Vega, H. (2000). Forest regeneration in a chronosequence of tropical abandoned pastures: implications for restoration ecology. Restoration ecology, 8, 328-338

Almeida, C.A.d., Coutinho, A.C., Esquerdo, J.C.D.M., Adami, M., Venturieri, A., Diniz, C.G., Dessay, N., Durieux, L., & Gomes, A.R. (2016). High spatial resolution land use and land cover mapping of the Brazilian Legal Amazon in 2008 using Landsat-5/TM and MODIS data. Acta Amazonica, 46, 291-302

Andersen, H.-E., Reutebuch, S.E., McGaughey, R.J., d'Oliveira, M.V., & Keller, M. (2014). Monitoring selective logging in western Amazonia with repeat lidar flights. Remote Sensing of Environment, 151, 157-165

Asner, G.P., Keller, M., Pereira Jr, R., Zweede, J.C., & Silva, J.N. (2004). Canopy damage and recovery after selective logging in Amazonia: field and satellite studies. Ecological Applications, 14, 280-298

Baccini, A., Goetz, S., Walker, W., Laporte, N., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P., Dubayah, R., & Friedl, M. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. Nature Climate Change, 2, 182-185

Baccini, A., Walker, W., Carvalho, L., Farina, M., Sulla-Menashe, D., & Houghton, R. (2017). Tropical forests are a net carbon source based on aboveground measurements of gain and loss. Science, 358, 230-234

Barona, E., Ramankutty, N., Hyman, G., & Coomes, O.T. (2010). The role of pasture and soybean in deforestation of the Brazilian Amazon. Environmental Research Letters, 5

Breiman, L. (2001). Random forests. Machine Learning, 45, 5-32

Carreiras, J.M., Jones, J., Lucas, R.M., & Gabriel, C. (2014). Land use and land cover change dynamics across the Brazilian Amazon: insights from extensive time-series analysis of remote sensing data. Plos One, 9, e104144

Carreiras, J.M., Jones, J., Lucas, R.M., & Shimabukuro, Y.E. (2017). Mapping major land cover types and retrieving the age of secondary forests in the Brazilian Amazon by combining single-date optical and radar remote sensing data. Remote Sensing of Environment, 194, 16-32

Chazdon, R.L. (2003). Tropical forest recovery: legacies of human impact and natural disturbances. Perspectives in Plant Ecology, evolution and systematics, 6, 51-71

Chazdon, R.L., Broadbent, E.N., Rozendaal, D.M., Bongers, F., Zambrano, A.M.A., Aide, T.M., Balvanera, P., Becknell, J.M., Boukili, V., & Brancalion, P.H. (2016). Carbon sequestration potential of second-growth forest regeneration in the Latin American tropics. Science Advances, 2, e1501639

Chen, D.Y., Huang, J.F., & Jackson, T.J. (2005). Vegetation water content estimation for corn and soybeans using spectral indices derived from MODIS near- and short-wave infrared bands. Remote Sensing of Environment, 98, 225-236

Chen, J., Chen, J., Liao, A.P., Cao, X., Chen, L.J., Chen, X.H., He, C.Y., Han, G., Peng, S., Lu, M., Zhang, W.W., Tong, X.H., & Mills, J. (2015). Global land cover mapping at 30 m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing, 103, 7-27

Claverie, M., Vermote, E.F., Franch, B., & Masek, J.G. (2015). Evaluation of the Landsat-5 TM and Landsat-7 ETM+ surface reflectance products. Remote Sensing of Environment, 169, 390-403

Cohen, W.B., Healey, S.P., Yang, Z., Stehman, S.V., Brewer, C.K., Brooks, E.B., Gorelick, N., Huang, C., Hughes, M.J., & Kennedy, R.E. (2017). How similar are forest disturbance maps derived from different Landsat time series algorithms? Forests, 8, 98

Cohen, W.B., Yang, Z., Healey, S.P., Kennedy, R.E., & Gorelick, N. (2018). A LandTrendr multispectral ensemble for forest disturbance detection. Remote Sensing of Environment, 205, 131-140

D'Agostino, R.B. (1970). Transformation to normality of the null distribution of g1. Biometrika, 679-681

Foley, J.A., DeFries, R., Asner, G.P., Barford, C., Bonan, G., Carpenter, S.R., Chapin, F.S., Coe, M.T., Daily, G.C., Gibbs, H.K., Helkowski, J.H., Holloway, T., Howard, E.A., Kucharik, C.J., Monfreda, C., Patz, J.A., Prentice, I.C., Ramankutty, N., & Snyder, P.K. (2005). Global consequences of land use. Science, 309, 570-574

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 202, 18-27

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S.V., 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

Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., & Hobart, G.W. (2015). Regional detection, characterization, and attribution of annual forest change from 1984 to 2012 using Landsat-derived time-series metrics. Remote Sensing of Environment, 170, 121-132

Hirschmugl, M., Gallaun, H., Dees, M., Datta, P., Deutscher, J., Koutsias, N., & Schardt, M. (2017). Methods for Mapping Forest Disturbance and Degradation from Optical Earth Observation Data: a Review. Current Forestry Reports, 3, 32-45

Home, O.D.H.N., List, C.D., NPP, N.P.P., Questions, D.P., Curation, D., Visualizer, S.M., & Checker, L.-W. (2013). LEDAPS calibration, reflectance, atmospheric correction preprocessing code, version 2

Houghton, R. (2012). Carbon emissions and the drivers of deforestation and forest degradation in the tropics. Current Opinion in Environmental Sustainability, 4, 597-603

Huang, C., Goward, S.N., Masek, J.G., Thomas, N., Zhu, Z., & Vogelmann, J.E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. Remote Sensing of Environment, 114, 183-198

Huete, A.R. (1988). A soil-adjusted vegetation index (SAVI). Remote Sensing of Environment, 25, 295-309

Hughes, M.J., Kaylor, S.D., & Hayes, D.J. (2017). Patch-based forest change detection from Landsat time series. Forests, 8, 166

INPE (2007-2013). Mapping of Forest Degradation in the Brazilian Amazon. www.obt.inpe.br/degrad/. Brazilian National Institute for Space Research

INPE (2017). Projecto Prodes: Monitoramento de Floresta Amazonica Brasileira por satelite. http://www.obt.inpe.br/prodes/index.php. Instituto Nacional de Pesquisas Espaciais

Joshi, N.P., Mitchard, E.T.A., Schumacher, J., Johannsen, V.K., Saatchi, S., & Fensholt, R. (2015). L-Band SAR Backscatter Related to Forest Cover, Height and Aboveground Biomass at Multiple Spatial Scales across Denmark. Remote Sensing, 7, 4442-4472

Kayastha, N., Thomas, V., Galbraith, J., & Banskota, A. (2012). Monitoring wetland change using inter-annual landsat time-series data. Wetlands, 32, 1149-1162

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

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 Sensing of Environment, 114, 2897-2910

Kimes, D.S., Nelson, R.F., Salas, W.A., & Skole, D.L. (1999). Mapping secondary tropical forest and forest age from SPOT HRV data. International Journal of Remote Sensing, 20, 3625-3640

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In, Ijcai (pp. 1137-1145): Montreal, Canada

Kuplich, T.M. (2006). Classifying regenerating forest stages in Amazonia using remotely sensed images and a neural network. Forest Ecology and Management, 234, 1-9

Langner, A., Miettinen, J., & Siegert, F. (2007). Land cover change 2002–2005 in Borneo and the role of fire derived from MODIS imagery. Glob Chang Biol, 13, 2329-2340

Lee, J.S. (1980). Digital Image-Enhancement and Noise Filtering by Use of Local Statistics. Ieee Transactions on Pattern Analysis and Machine Intelligence, 2, 165-168

Lee, J.S.H., Wich, S., Widayati, A., & Koh, L.P. (2016). Detecting industrial oil palm plantations on Landsat images with Google Earth Engine. Remote Sensing Applications: Society and Environment, 4, 219-224

Lopes, A., Touzi, R., & Nezry, E. (1990). Adaptive Speckle Filters and Scene Heterogeneity. Ieee Transactions on Geoscience and Remote Sensing, 28, 992-1000

Lu, D. (2005). Integration of vegetation inventory data and Landsat TM image for vegetation classification in the western Brazilian Amazon. Forest Ecology and Management, 213, 369-383

MapBiomas (2015). Project MapBiomas - Collection 2 of Brazilian Land Cover & Use Map Series, accessed on 15/11/2017 through the link: http://mapbiomas.org/pages/database/mapbiomas_collection

Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Gao, F., Kutler, J., & Lim, T.-K. (2006). A Landsat surface reflectance dataset for North America, 1990-2000. IEEE Geoscience and Remote Sensing Letters, 3, 68-72

Maxwell, S.K., & Sylvester, K.M. (2012). Identification of “ever-cropped” land (1984–2010) using Landsat annual maximum NDVI image composites: Southwestern Kansas case study. Remote Sensing of Environment, 121, 186-195

Motohka, T., Shimada, M., Uryu, Y., & Setiabudi, B. (2014). Using time series PALSAR gamma nought mosaics for automatic detection of tropical deforestation: A test study in Riau, Indonesia. Remote Sensing of Environment, 155, 79-88

Nelson, R.F., Kimes, D.S., Salas, W.A., & Routhier, M. (2000). Secondary forest age and tropical forest biomass estimation using thematic mapper imagery. Bioscience, 50, 419-431

Olson, D.M., Dinerstein, E., Wikramanayake, E.D., Burgess, N.D., Powell, G.V.N., Underwood, E.C., D'Amico, J.A., Itoua, I., Strand, H.E., Morrison, J.C., Loucks, C.J., Allnutt, T.F., Ricketts, T.H., Kura, Y., Lamoreux, J.F., Wettengel, W.W., Hedao, P., & Kassem, K.R. (2001). Terrestrial ecoregions of the worlds: A new map of life on Earth. Bioscience, 51, 933-938

Padarian, J., Minasny, B., & McBratney, A.B. (2015). Using Google's cloud-based platform for digital soil mapping. Computers & Geosciences, 83, 80-88

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S., Rautiainen, A., Sitch, S., & Hayes, D. (2011). A large and persistent carbon sink in the world's forests. Science, 333, 988-993

Patel, N.N., Angiuli, E., Gamba, P., Gaughan, A., Lisini, G., Stevens, F.R., Tatem, A.J., & Trianni, G. (2015). Multitemporal settlement and population mapping from Landsat using Google Earth Engine. International Journal of Applied Earth Observation and Geoinformation, 35, 199-208

Pekel, J.-F., Cottam, A., Gorelick, N., & Belward, A.S. (2016). High-resolution mapping of global surface water and its long-term changes. Nature, 540, 418

Poorter, L., Bongers, F., Aide, T.M., Zambrano, A.M.A., Balvanera, P., Becknell, J.M., Boukili, V., Brancalion, P.H., Broadbent, E.N., & Chazdon, R.L. (2016). Biomass resilience of Neotropical secondary forests. Nature, 530, 211

Schaffer, C. (1993). Selecting a Classification Method by Cross-Validation. Machine Learning, 13, 135-143

Shelestov, A., Lavreniuk, M., Kussul, N., Novikov, A., & Skakun, S. (2017). Exploring Google Earth Engine Platform for Big Data Processing: Classification of Multi-Temporal Satellite Imagery for Crop Mapping. Frontiers in Earth Science, 5, 17

Skole, D., & Tucker, C. (1993). Tropical deforestation and habitat fragmentation in the Amazon: satellite data from 1978 to 1988. Science, 260, 1905-1910

Souza, C., Firestone, L., Silva, L.M., & Roberts, D. (2003). Mapping forest degradation in the Eastern Amazon from SPOT 4 through spectral mixture models. Remote Sensing of Environment, 87, 494-506

Team, P. (2017). Planet application program interface: In Space for life on earth. San Francisco, CA

Trisasongko, B.H. (2010). The use of polarimetric SAR data for forest disturbance monitoring. Sensing and Imaging: An International Journal, 11, 1-13

USGS (2018). PRODUCT GUIDE. LANDSAT 4-7 SURFACE REFLECTANCE (LEDAPS) PRODUCT. United States Geographical Survey, Department of the Interior

van der Werf, G.R., Morton, D.C., DeFries, R.S., Olivier, J.G.J., Kasibhatla, P.S., Jackson, R.B., Collatz, G.J., & Randerson, J.T. (2009). CO2 emissions from forest loss. Nature Geoscience, 2, 737-738

Vieira, I.C.G., de Almeida, A.S., Davidson, E.A., Stone, T.A., de Carvalho, C.J.R., & Guerrero, J.B. (2003). Classifying successional forests using Landsat spectral properties and ecological characteristics in eastern Amazonia. Remote Sensing of Environment, 87, 470-481

White, J.C., Wulder, M.A., Hermosilla, T., Coops, N.C., & Hobart, G.W. (2017). A nationwide annual characterization of 25 years of forest disturbance and recovery for Canada using Landsat time series. Remote Sensing of Environment, 194, 303-321

Xiong, J., Thenkabail, P.S., Gumma, M.K., Teluguntla, P., Poehnelt, J., Congalton, R.G., Yadav, K., & Thau, D. (2017). Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. ISPRS Journal of Photogrammetry and Remote Sensing, 126, 225-244

List of Figure Captions