The expansion of tree plantations across tropical biomes

Matthew E. Fagan [1,7 ✉], Do-Hyung Kim[2,7], Wesley Settle[1], Lexie Ferry[1], Justin Drew[1], Haven Carlson[1], Joshua Slaughter[1], Joshua Schaferbien[1], Alexandra Tyukavina [3], Nancy L. Harris [4], Elizabeth Goldman[4] and Elsa M. Ordway[5,6]

Across the tropics, recent agricultural shifts have led to a rapid expansion of tree plantations, often into intact forests and grasslands. However, this expansion is poorly characterized. Here, we report tropical tree plantation expansion between 2000 and 2012, based on classifying nearly 7 million unique patches of observed tree cover gain using optical and radar satellite imagery. The resulting map was a subsample of all tree cover gain but we coupled it with an extensive random accuracy assessment ($n = 4,269$ points) to provide unbiased estimates of expansion. Most predicted gain patches (69.2%) consisted of small patches of natural regrowth (31.6 ± 11.9 Mha). However, expansion of tree plantations also dominated increases in tree cover across the tropics (32.2 ± 9.4 Mha) with 92% of predicted plantation expansion occurring in biodiversity hotspots and 14% in arid biomes. We estimate that tree plantations expanded into 9.2% of accessible protected areas across the humid tropics, most frequently in southeast Asia, west Africa and Brazil. Given international tree planting commitments, it is critical to understand how future tree plantation expansion will affect remaining natural ecosystems.

Over the last several decades, the expansion of commercial agriculture has become one of the key drivers of rainforest loss[1,2]. Concurrently, agricultural production in tropical biomes has undergone a dramatic shift that has led to increases in the production of crops and wood products from tree plantations[3,4], defined here as productive monocultures composed of erect trees. Numerous studies have reported regional expansions in agricultural tree crops (for example, banana, oil palm and rubber[5,6]) and in timber and pulp plantations (for example, pine and eucalyptus[7,8]). This expansion of tree plantation area has frequently come at the expense of intact forest and grassland ecosystems[5–7,9]. Despite this risk, many countries have prioritized expanding tree plantations as part of their international commitments to restore degraded tropical habitats[10,11]. For example, 45% of national commitments to the Bonn Challenge, an international goal to restore 350 million ha of land by 2030, consist of expanding tree plantations[10].

Despite the availability of national statistics to document these trends, consistent, spatially explicit estimates of global increases in tree plantation area are lacking[12–14]. Many efforts to map tree plantations rely on manual delineation of tree plantation boundaries from high-resolution satellite imagery and/or field-collected data[1,13,15–17]. These expert delineations cannot consistently detect changes in tree plantation area across regions because they focus on areas of intensive production and on a single time period. For example, although a recent study of tropical humid forest change ostensibly tracked plantation expansion over time, its delineation of tree plantations largely relied on single-date expert interpretation and the plantation map accuracy was not reported[17].

The absence of systematic monitoring of plantation expansion complicates assessments of both forest restoration efforts and the impacts of tree plantations on natural ecosystems. Confusion between natural forests and tree plantations in forest change maps leads to 'cryptic forest loss'[18] and/or overestimates of natural forest loss and recovery[19,20]. This is especially true in the tropics, with its rapid tree growth and expansion of tree plantations. National reporting indicates net increases in tree plantation cover of 1–2% a year[21] but tracking net changes in the area of frequently disturbed land covers like tree plantations and natural forest regrowth can underestimate the occurrence of expansion in new areas as existing areas are harvested[12]. According to the UN Food and Agriculture Organization (FAO), the area of planted forests (294 million ha) and arborescent (treelike) crops (102 million ha) in 2020 was much smaller than that of natural forest cover regenerating from disturbances (3.75 billion ha)[21,22]. However, given that the FAO definition of naturally regenerating forest includes existing, selectively logged forests, it remains unclear whether natural forest recovery or tree plantations are driving expansions in global tree cover.

To address these uncertainties, we undertook a pantropical assessment of increases in tree plantation area, with tree plantations defined here as monocultures of agricultural or industrial arborescent species established and managed by humans for fruit, wood, fibre and other products. We focused on the tropics (25° N to 25° S) due to the prevalence of natural forest conversion to commodity tree crops across tropical latitudes[1] and high rates of potential carbon sequestration from tropical tree regrowth[23]. Our aim was to use remote sensing data to accurately distinguish plantation expansion from natural forest regrowth and assess recent expansion across tropical biomes, biodiversity hotpots and protected areas.

Monitoring plantation expansion

Mapping tree plantations consistently using satellite data is challenging, especially using the moderate spatial resolution imagery

[1]Department of Geography and Environmental Systems, University of Maryland Baltimore County, Baltimore, MD, USA. [2]United Nations Children's Fund, New York, NY, USA. [3]Department of Geographical Sciences, University of Maryland, College Park, MD, USA. [4]World Resources Institute, Washington, DC, USA. [5]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA. [6]Department of Ecology and Evolutionary Biology, University of California Los Angeles, Los Angeles, CA, USA. [7]These authors contributed equally: Matthew E. Fagan, Do-Hyung Kim. ✉e-mail: mfagan@umbc.edu

Table 1 | Estimated expansion area of tree plantations and natural regrowth, 2000–2012

Region	Estimated gain area (Mha)	Area 95% CI (Mha)	Percentage of mapped patch area	Mapped patch number	Mean patch size (ha)
Africa					
Tree plantations	0.4	0.1	25.0	77,046	5.3
Natural regrowth	7.2	8.5	73.5	1,296,932	0.9
Non-gain	2,584.8	8.5	1.5	23,456	1.1
Latin America					
Tree plantations	5.8	2.5	58.6	499,042	6.6
Natural regrowth	14.8	7.2	37.7	1,549,391	1.4
Non-gain	1,522.8	7.6	3.7	88,832	2.3
Australasia					
Tree plantations	26.0	9.1	75.6	1,415,054	5.7
Natural regrowth	9.5	4.1	24.1	1,928,709	1.3
Non-gain	1,088.0	9.9	0.3	23,219	1.4
Tropics					
Tree plantations	32.2	9.4	65.7	1,991,142	5.9
Natural regrowth	31.6	11.9	32.8	4,775,032	1.2
Non-gain	5,195.6	15.1	1.5	135,507	1.9

The estimated area is accuracy-corrected (left two columns: mean ± 95% CI), while the mapped patch area is a subsample of total estimated area and not accuracy-corrected (right three columns). The non-gain areaestimate (left two columns) includes both the background area and the area of mapped tree cover gain estimated to be not forested ('open'). The non-gain estimates in the right three columns refer to mapped open patches.

(10–100 m) needed for comprehensive regional coverage. This challenge arises in part from spectral and structural similarities between regrowing natural forest and plantations of trees[24–26]. Tree plantations are spectrally diverse, with substantial variation in spectral signatures across species, age classes, planting pattern and density, nutrition and disease status, soil types, understory cover and disturbance intensities[24,27–29]. Regrowing forests are also spectrally and structurally diverse, for similar reasons, especially across distinct forest types[30,31]. These similarities, coupled with geographic variation in spectral reflectance and phenology and persistent tropical cloudiness, make it difficult to consistently distinguish tree plantations from natural forest using satellite imagery[5].

Although intra- and inter-regional plantation variability has constrained accuracy at larger scales, maps based on automated classification of remotely sensed imagery have successfully monitored plantation expansion at regional scales (for example, ref. [5]). However, the accuracy of regional tree plantation maps differs widely across different geographies and plantation species[5,13]. In general, extensive monocultures of non-native plantation species have been more readily distinguished, especially if they possess distinct spacing or phenology. For example, rubber and oil palm have been mapped with intermediate to high accuracy (80–90%+) at local to near-global scales[5,32–35]. However, other tree crop and timber species often are mapped with lower accuracy: in Brazil and Chile, country-level mapping of non-native eucalyptus and pine plantations has had intermediate levels of accuracy (70–89%)[15,36].

Research design
In this study, we integrated two different types of satellite imagery, optical and microwave, to better characterize the spectral (optical) and structural (microwave) properties of regrowing natural forests and tree plantations. To estimate recent increases in plantation area and in natural forest regrowth, we reclassified a widely used map of gains in tree cover between 2000 and 2012 (ref. [37]), the Global Forest Change (GFC) product. From this dataset, we selected mapped

gain patches ≥0.45 ha in size that persisted for 4 years after 2012, at least through the end of 2015 ($n = 6,901,681$). Over each resulting patch of gain pixels, we extracted spectral and microwave satellite imagery data (Landsat, ALOS PALSAR-2 and Sentinel-1) and other ancillary data (for example, GFC tree cover, patch size and shape; Supplementary Table 1) for a total of 32 metrics per patch (Supplementary Table 1).

We then used labelled training data and machine learning to predict patch-level land use in 2015. The output was a classification of gain patches as either 'tree plantation', 'natural regrowth' or 'open' (<10% tree cover); these classes characterize increases in tree cover between 2000 and 2012 that persisted through 2015. Then, to assess the extent of plantation expansion across the tropics, classified gain footprints were intersected with available spatial data on biomes, biodiversity hotspots, national borders and protected areas.

Product accuracy
We assessed the classification accuracy of both our machine learning model and the resulting map of plantation expansion. First, model accuracy was independently evaluated by random sampling of 2,000 gain patches across the tropics. Next, map accuracy was evaluated by stratified random sampling of points across the tropics ($n = 4,269$); sampling strata included biomes, continents and predicted land-use classes (including non-gain areas). For each reference patch and point, high-resolution imagery was used to determine a single land-use class. Reference patches with mixed land uses were assigned to a class using per cent cover rules (for example, majority per cent cover; Supplementary Table 5) and the estimated overall model accuracy was 90.6% (±0.1 (95% confidence interval, CI); Supplementary Tables 6–8). The map accuracy assessment indicated that the GFC gain patches were a subsample of actual tree cover increases, for both tree plantations (producer's accuracy of 33.3% (±9.8)) and natural regrowth (producer's accuracy of 15.9% (±6.0)). Class accuracies were highest in the humid tropical biome and lower and more variable in non-humid biomes (Supplementary Table 9a–d).
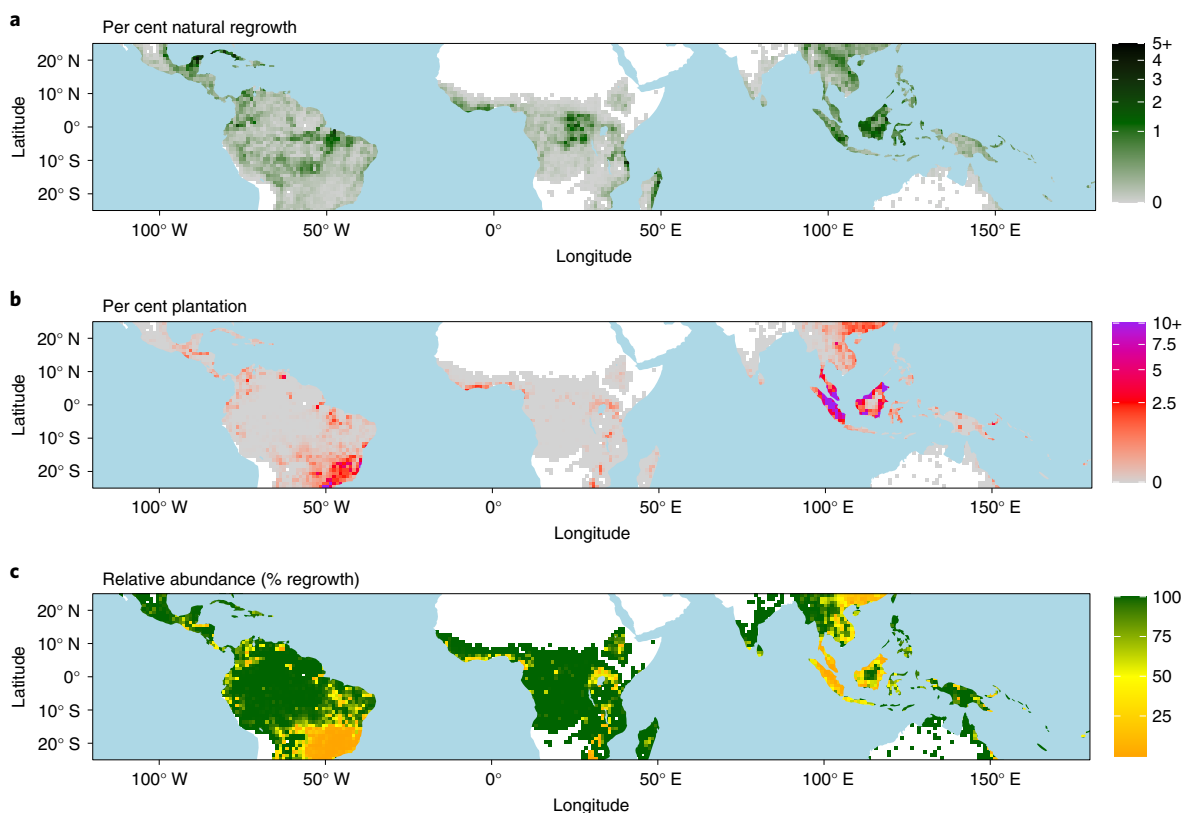
Fig. 1 | Pantropical distribution of natural regrowth and tree plantations. a,b, The percentage land area occupied by each land use—natural regrowth (**a**) and tree plantation (**b**); higher-resolution (30 m) data were summarized across one-degree grid cells for display. **c**, The relative abundance of natural regrowth, as a proportion of total observed gains in tree cover.

Despite the conservative nature of the GFC product, low rates of commission error permitted relative comparisons of the distribution of tree plantations and robust adjusted estimates of the overall area of tree plantations (user's accuracy of 90.7% ($\pm$5.1)) and natural regrowth (user's accuracy of 84.5% ($\pm$4.6)). This far exceeds the accuracy of a recent tropical moist forest product[17], which coupled comparable omission errors with very high rates of commission error (mean gain class user's accuracy of 33.9%, Supplementary Table 10).

Global expansion patterns

Between 2000 and 2012, persistent gain in tree cover associated with tree plantations was comparable to the persistent gain in tree cover associated with natural forest regrowth. Pantropically, there were ~32.2 ($\pm$9.4) Mha of additional tree plantations and 31.6 ($\pm$11.9) Mha of additional natural regrowth (Table 1, Fig. 1 and Supplementary Figs. 4–9). Although observed patches of natural regrowth far outnumbered plantation patches, individual plantation patches were ~4.9$\times$ larger than areas of forest regrowth (Table 1). Across continental regions, tree plantations were predicted to increase across the most area in Asia, followed by Latin America (Table 1 and Fig. 2). The concentration of plantation expansion there is probably driven largely by the production of relatively few tree crops (oil palm, rubber and acacia in Australasia; oil palm, eucalyptus and pine in Latin America)[12]. Africa had the smallest predicted plantation expansion but plantations still made up a sizeable fraction of the area of several African countries (Supplementary Table 11) and Africa is a nascent frontier for oil palm expansion[35].

Our findings show that expansion of tropical tree plantations was significantly higher in the humid tropical biome (Fig. 2 and Supplementary Table 9a–d). Although area estimates ranged widely in non-humid tropical biomes because of the relative rarity of tree

plantations there, predicted plantation areas in non-humid biomes were robust due to high user's accuracy. Matching concerns about recent afforestation in arid regions[38,39], new plantations were a common (14% of predicted global plantation area; Fig. 2) form of increasing tree cover in arid biomes, particularly in tropical grasslands, savannas and shrublands in eastern Africa and southeast Latin America. Furthermore, predicted plantation expansion was concentrated in designated biodiversity hotspots, particularly the Sundaland (southeast Asia), Cerrado (southern Brazil) and Atlantic Forest (southeast Brazil) hotspots (collectively 92.8% of total plantation area; Supplementary Fig. 11). By contrast, natural regrowth was relatively evenly distributed among biomes and hotspots, although natural regrowth was slightly more abundant and plantations less abundant, in ecoregions with high remnant natural forest cover (Supplementary Table 11, $P < 0.0001$).

We show that predicted plantation expansion was concentrated (82.8%) in four large countries (Indonesia, Brazil, Malaysia and China) but tree plantations were widespread, occurring in 78 of 112 tropical countries with detected increases in tree cover (Supplementary Table 12). Relative to natural regrowth, predicted tree plantation patches were distributed in a distinct geographic manner. Tree plantation expansion was more likely than natural regrowth to be found near navigable waterways ($P < 0.0001$, Supplementary Fig. 12a), probably reflecting the role of global trade in influencing the expansion of plantations. Similarly, new tree plantations were more likely to be found in highly human-dominated landscapes[40] than was natural regrowth ($P < 0.0001$, Supplementary Fig. 12b), reflecting the global distribution of natural regrowth along deforestation frontiers (Fig. 1). Locations where plantation expansion accompanied natural regeneration were relatively spatially constrained (for example, Indonesia and western Africa).
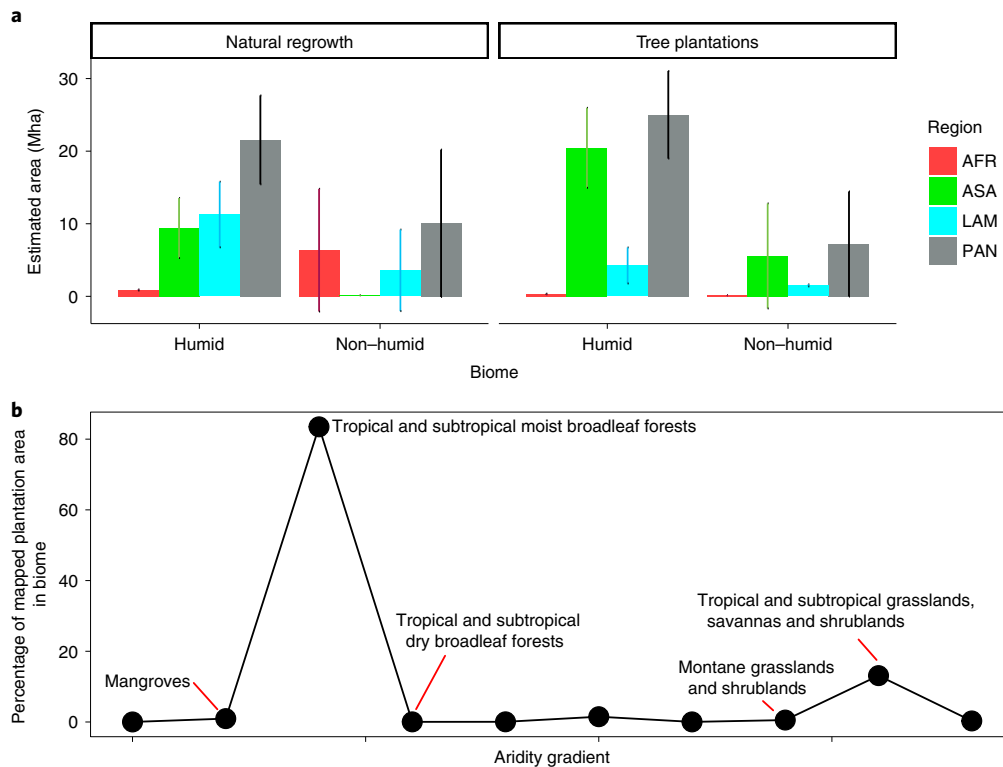
Fig. 2 | Estimated expansion of tree plantations and natural regrowth into terrestrial regions and biomes. a, Estimated gain area is shown for each land use, biome and region (AFR, Africa; ASA, Australasia; LAM, Latin America; PAN, Pantropical total). Error bars are 95% CIs. **b**, Normalized mapped plantation area shown for each biome as a point along an ordinal aridity gradient. Biomes with >0.5% of mapped global plantation area are labelled.

Expansion in protected areas

Tree plantation expansion into protected areas (PAs) was not commonly observed across the tropics, affecting 4.2% of PAs (PAs with ≥5 ha of expansion) (Fig. 3 and Supplementary Table 13). However, expansion was more common in PAs located in the humid tropical biome (5.5% of PAs) especially in PAs there with above-average accessibility and human influence[40] (9.2% of affected PAs). Furthermore, plantation expansion often occurred in close proximity to tropical PAs (<1 km outside), affecting 7.8% of all PAs, 11.4% of humid tropical biome PAs and 16.2% of humid PAs with high human influence (Supplementary Tables 13 and 14). Notable concentrations of plantation expansion into PAs include southeast Asia, southeast Brazil and western and eastern Africa (Fig. 2). Plantation expansion was rarely widespread within PAs, with only 2.1% of PAs experiencing ≥50 ha of tree plantation expansion. But where it occurred inside PAs, plantation expansion was a dominant mode of increasing tree cover, making up 49.5% (±38.5%, s.d.) of observed gains in cover within affected PAs. In Africa, parks with stricter protected status (IUCN category) were less likely to experience plantation expansion ($P < 0.0001$; Supplementary Table 15). However, in Latin America and Asia, protected status was unrelated to the presence of plantations within parks ($P > 0.05$), potentially reflecting greater encroachment pressure in these regions.

Product limitations

There are several known limitations of this product. First, its accuracy in characterizing increases in tree cover depends on the accuracy of the original GFC product[37]. The original product contained relatively high errors of omission in predicting tropical tree cover gain (capturing ~50% of observed gain; Supplementary Table 10). Our map product shares that trait, resulting in variable estimates of the area of tropical tree cover gain. Therefore, it is possible that either tree plantations or natural regrowth are relatively more

abundant, with tree plantations estimated to make up between 34.4% and 67.8% of total tree cover gain (95% CI, Table 1). Regardless, our results clearly indicate that tree plantation expansion is widespread across all tropical regions and biomes. The estimated relative dominance of plantations does not arise solely from differential persistence of detected plantations and natural regrowth over time or from patch size thresholds, as predicted plantation expansion is at least 40.4% of the original GFC gain area (Supplementary Fig. 13).

The GFC product also tends to underestimate tree cover in arid biomes[41,42]. Consequently, outside the humid tropical biome our product may misestimate both total tree cover gain area and the relative importance of plantations. Our product was also not designed to distinguish rarer land uses intermediate between tree plantations and natural regrowth, such as assisted natural regeneration. Further, because our product was derived from both optical and microwave data, local variation in topography and forest structure affected classification accuracy. Our product tended to have lower accuracy in mountainous areas (where microwave returns vary) and in natural regrowth patches where structure or colour resembled row-planted monocultures. For example, visual inspection showed more frequent errors in open natural regrowth affected by fire or selective logging and in single-species dominated natural regrowth patches with a smooth canopy (for example, riparian regrowth in the western Amazon). In certain cases, natural regrowth within plantations was underestimated: when mixed gain patches occurred, containing both tree plantations and natural regrowth, whole gain patches were labelled according to their majority land-use class.

When we compared our estimates of increases in plantation cover to those reported by tropical countries to the FAO between 2000 and 2010, we found that our accuracy-adjusted estimates of total plantation expansion area (32.2 ± 9.4 Mha) were higher than the total net increase in plantations reported for this period (21.1 Mha). As FAO-reported plantation area is a net (losses subtracted
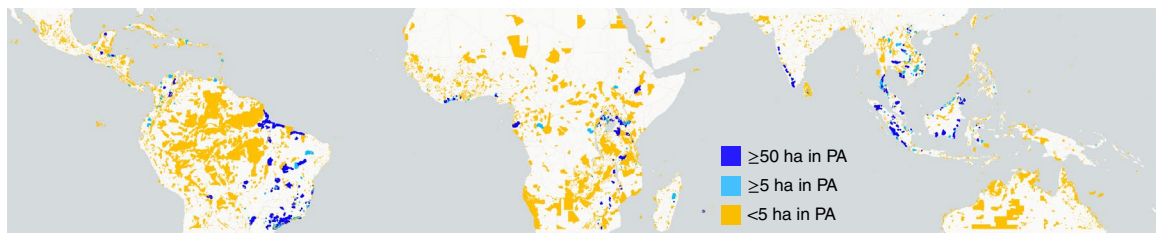
Fig. 3 | Expansion of plantations into tropical protected areas. PAs are colour-coded by estimated plantation expansion area (2000–2012).

from gains), our corrected estimate of plantation expansion should be higher than the FAO estimate. However, for most countries, our unadjusted map predictions were actually lower than the FAO-reported area of timber and agricultural plantations, most notably in India, Bangladesh and Africa (Supplementary Fig. 13). This is probably due to both high errors of omission in the original GFC gain dataset (ref. [37], Supplementary Table 10) and definitional differences (for example, orchard crop fields can be <5 m in height and not meet the GFC tree cover definition). Thus, our map predictions of the relative abundance of plantation expansion across countries are likely to be conservative.

Discussion

Tree plantations are a key element of ambitious policy proposals to restore ecosystem services and address climate change, including the Bonn Challenge, the Trillion Tree initiatives and the UN Decade on Ecosystem Restoration[43]. Given their potential impacts on biodiversity, fire risk and human well-being, tree plantations are controversial[38,44]. This controversy has largely proceeded in the absence of global data on plantation expansion, and the net impact of plantation expansion has thus been difficult to assess, especially in understudied regions like Africa. In this context, our findings provide a few useful insights for future policy and research.

First, tree plantations make up a large proportion of recent increases in tropical tree cover. This indicates that many global- and country-level estimates of tropical forest loss and gain are potentially biased by the presence of tree plantations, which have disturbance rates distinct from those of natural forests. Distinguishing tree plantations from natural regrowth enables assessments of their relative persistence over time[45], the spatiotemporal clustering of expansion and replanting[12] and the displacement of previous land uses by plantation expansion[46]. In the absence of these analyses, countries dominated by plantations will be unable to forecast production or limit its impact on biodiversity or smallholder agriculture.

Second, the abundance of natural forest regeneration relative to plantations was much lower than predicted by nationally reported statistics, suggesting that they potentially do not accurately track forest expansion through natural regrowth. This is most probably because the area of selectively logged tropical forest far exceeds that of natural regrowth[47] but may also result in part from rapid reclearing of natural regrowth[45] and inconsistent national reporting of regrowth in agricultural landscapes[48]. Understanding the fate of the large number of natural regrowth patches identified here is critical to determining the long-term climate mitigation potential of tropical forests[49]. Countries making international climate and restoration commitments would benefit from annual monitoring of natural regeneration. Using data on tree height[50] and carbon uptake[23,51], future research could estimate annual carbon sequestration rates for tree plantations and natural regeneration during this period.

Third, the widespread expansion of plantations into arid biomes and tropical PAs indicate that economic considerations frequently took precedence over conservation policies and interests during this time period. The net benefits of planting trees—for carbon, biodiversity and food security—entirely depend on the types of ecosystem they replace[7]. Thus, tracking plantation expansion is essential to improve estimates of net global carbon sequestration and available agricultural area and to assess the net impacts of progress towards restoration commitments. Using new data on tropical forest loss and regrowth[17], products like this one could be annually updated. If plantation expansion continues at the observed pace in the coming decades, achieving tropical conservation and food production goals will be increasingly difficult. Given current widespread international interest in tree planting, it is critical to monitor going forward how proposed expansions in plantation cover will affect remaining natural ecosystems.

Methods

Study region and patches. We focus on recent areas of increase in tree cover or 'gain', between the years 2000 and 2012, delineated in a previous study of global forest change (ref. [37], GFC data v.1.5). We examined all mapped patches of contiguous gain pixels (30 m resolution) between 25° N and 25° S that met our minimum patch size criteria (≥0.45 ha). Then, because our preliminary analysis indicated that false-positive identification of tree cover gain was more common in small patches and in patches that showed ephemeral vegetative gain (gains followed by losses), we set criteria for minimum forest persistence. We selected only gain patches >0.45 ha in size that persisted at least 4 years after 2012. These minimum patch size and persistence criteria acted as a conservative forest filter, eliminating non-forest error patches, and their derivation is described in more detail in the Supplementary Methods. Contiguous gain patches were generated, updated by removing forest loss and converted to polygons in Google Earth Engine, with a unique ID number assigned.

Patch-level data. For each resulting gain patch ($n = 6,904,335$), several characteristics were calculated, including patch area, patch perimeter length, the patch perimeter:area (P:A) ratio and a patch compactness index based on area and perimeter[52]. In addition, two additional patch metrics were derived: the distance to the nearest patch with a high P:A ratio (in the top 1% of patches) and the distance to the nearest patch with a high compactness index (in the top 1% of patches). Patch characteristics were calculated in the native projection of the original raster dataset (WGS-84) using the sf and geosphere packages in R v.3.5.1.

To distinguish natural regrowth and tree plantations, we used two different types of coincident satellite remote sensing data: optical and radar data. Moderate-resolution Landsat optical imagery have been used for many studies of forest and land cover due to the availability of free, consistent observations over several decades[13,37]. However, as an optical sensor, Landsat data are severely influenced by the presence of clouds, particularly in the tropics[53,54]. Furthermore, the spectral resolution of single-date Landsat data has often not been sufficient by itself to separate forest and tree plantations[13,55]. Radar data, in contrast, are less sensitive to cloud cover and potentially more sensitive to differences in structure between natural and anthropogenic tree cover[56,57]. Although until recently radar data had relatively limited availability compared to Landsat[56], global mosaics of PALSAR-1 (2007–2010), PALSAR-2 (2015+) and Sentinel-1 (2014+) data are now freely available.

In this analysis, we used the global Landsat spectral mosaics available from the GLAD laboratory (v.1.3)[37], covering the entire globe for the nominal year of 2015. For coincident radar data that were as close as possible to the 2012 end date for gain patches, we used the 2015 global mosaic of L-band PALSAR-2 data available from JAXA[58] and a 2015 global mosaic of C-band Sentinel-1 data from the ESA[59]. Details on optical and radar data band selection and processing are available in the Supplementary Methods.

In addition to spectral, radar and patch-characteristic data, we included several satellite-derived land cover datasets as input data in this analysis. These data consisted of GFC tree cover in the year 2000, GFC loss year and a 1 km dataset on crop extent (https://developers.google.com/earth-engine/datasets/catalog/USGS_GFSAD1000_V1). For each patch, we calculated the patch-level mean and standard deviation of all pixels that intersected with the patch boundary.

The patch-level mean and standard deviation calculations for each band included pixels that overlapped with individual patch boundaries but all calculations weighted individual pixels according to their area, following a standard Google Earth Engine algorithm.

The resulting dataset included 35 different variables for each individual patch (Supplementary Table 1); all variables were used in further analyses. Missing Sentinel-1 or ALOS Palsar data values were detected in 0.038% of patches, mainly due to lack of data coverage over small, remote islands. These patches were omitted from further analysis and are mapped as no-data in the final dataset (final patch $n = 6,901,681$).

Reference data labelling. Tree plantation reference data came from two main sources: the World Resource Institute (WRI) and manual delineation of plantations. We used the WRI tree plantation polygons to select all Hansen gain polygons that intersected their boundaries, with postintersection filtering described in the Supplementary Methods. Due to the lack of quality plantation training data in Africa, Australia and mainland southeast Asia, we also manually delineated a variety of plantation species across these regions using freely available time series of high-resolution imagery (Google Earth, Bing and ArcGIS basemap). Gain patches were selected in this process by intersection with manually created polygons, including only patches dominated by obvious plantation species.

Criteria for distinguishing plantation species included (1) domination by a single commercial tree species (consistent colour and uniform canopy shapes) and (2) the clear presence of rows of trees at some point in the imagery time series (pattern). Clear evidence of human disturbance over time, extractive infrastructure and regular patch edges often facilitated identification of potential plantation patches but by itself was not diagnostic of plantations due to farm and plantation abandonment. Common commercial plantation species (oil palm, rubber, eucalyptus, pine) were recognizable and distinct in high-resolution imagery, as were many unknown plantation species in Africa and Australia. Diverse polyspecies plantations and agroforestry stands were omitted by this methodology. We did not manually delineate plantation systems with a mixed set of species or those that lacked a row structure (for example, some agroforestry). In addition, globally rare plantation types with distinct species, cultivars and/or conditions (for example, insect infestation) are likely to have been omitted.

Natural forest regrowth reference data were derived from three main sources: the Intact Forest Landscapes dataset (IFL, ref. [60]); the World Database on Protected Areas (WDPA v.1.4, ref. [61]); and manual delineation of natural regrowth. The WDPA polygons were edited to include only patches clearly dominated by natural regrowth, as is described in more detail in the Supplementary Methods. Gain patches were then selected and labelled in this process by intersection with IFL, WDPA and manually created polygons. Our initial reference data sample, based on intact forest landscapes and PAs, had gaps in spatial coverage of secondary forests in southeastern Amazonia, southern Mexico, Africa and southeast Asia. To address these gaps, we manually delineated natural regrowth patches across these regions using freely available time series of high-resolution imagery (Google Earth, Bing and ArcGIS basemap). Natural regrowth was distinguished by several criteria, including a diversity of tree crown shapes, colours and/or sizes, irregular patch edges, the absence of human infrastructure and/or no evidence of tree planting or rows in the imagery time series (Supplementary Fig. 1). Some natural stands did not meet several of these criteria but could be distinguished from plantations by one or more attribute.

Training data processing. The resulting training data included 729,092 gain patches, consisting of natural regrowth, 15 different known plantation species and several unknown plantation species (Supplementary Table 2). Oil palm and eucalyptus dominated the initial training sample, making up 82% of all plantation training data. Plantation training data were unevenly distributed across continents, with Africa having the smallest number of samples (Supplementary Table 2). By contrast, the training data for natural regrowth were more evenly distributed across continents (Supplementary Table 2). With no a priori information on the relative abundance of different plantation species and natural regrowth across the globe, the initial training dataset was an unbalanced mixture of different plantation species and natural regrowth patches. To balance the plantation training data, we first resampled the existing data, which is described in more detail in the Supplementary Methods. We then randomly withheld 10% of the processed training data as internal validation data ($n = 72,908$).

Second, given that the relative global proportion of plantation and regrowth was unknown, we created 11 different final training datasets through subsampling (and where necessary to increase sample size of a particular class, sampling with replacement). Each training sample was created with different proportions (balancing) of plantations and regrowth (total sample $n = 559,580$–$839,370$), covering a range of proportions that preliminary analysis indicated changed the predicted outcomes of classification models (Supplementary Table 3 and Supplementary Methods).

Machine learning classification models. For each of the 11 different balancings of the training data, we fit 5 types of binary classification models, totalling 55 classification models. The binary classes predicted were natural regrowth and tree

plantation and the independent predictors are summarized in Supplementary Table 1. All classification models were implemented in the R interface to the H2O machine learning environment (https://www.h2o.ai, v.3.26.0.5), with default model parameter and grid search settings. The five types of classification models were: (1) gradient boosting machines (GBM), (2) logistic general linear models (GLM), (3) distributed random forest (DRF), (4) extremely randomized trees (XRT) and (5) feed-forward deep-learning neural networks (DL). To facilitate grid searches for parameter values and conduct direct model intercomparison, we used the H2O automated machine learning (h2o.autoML) algorithm with default values[62]. At the end of each autoML run, the most accurate model (highest AUC) for each model type was saved (Supplementary Methods).

The resulting best models in terms of AUC in each of the five model families, for each of the 11 training datasets (55 models in total), were each used to predict the binary class of the internal validation data ($n = 72,908$). The resulting predictions included the binary class prediction and the likelihood of that predicted class, which ranged from 0 to 1. The likelihood of the plantation class was selected as a variable for further analysis, with 55 total classification model predictions. All of these 55 classification model predictions, along with the internal test class labels, were used as input data to a stacked ensemble machine learning model.

Stacked ensemble classification. Stacked ensemble machine learning models[63] were used to predict the binary land-use class (natural regrowth or tree plantation) of each patch, using all of the individual classification model predictions as input data. Stacked ensemble predictions used the random forest algorithm as a final classifier, with model parameters set to the H2O modelling environment defaults (ntrees $= 50$, mtries $= 7$ (the square root of the number of predictor variables)). The internal validation data were used as the input labelled training data. A separate stacked ensemble model was developed for three main tropical regions: Latin America, Africa and Australasia (Asia, Australia and Oceania), leading to different input training data proportions for each region (Supplementary Table 4). The resulting stacked ensemble models, one for each tropical region, were used to predict the binary land-use class of all patches within their region.

Masking and postprocessing. Postclassification analysis of the predicted product indicated two main quality issues. First, because GFC gain patches in mangrove patches were quite rare globally[64] and thus rare in the training data, misclassification of low-diversity mangrove gain patches as tree plantations was common in southeast Asia. To address this issue, a high-quality global mangrove extent map (ref. [65], 90% overall accuracy) was used to correct our product. All gain patches that intersected with the 2015 mangrove extent were reclassified as the natural regrowth class, unless they exceeded 20 ha in size and had less than 5% of their area overlapping the mangrove map. This low per cent area threshold excluded the majority of tree plantation patches that bordered mangroves.

Second, visual inspection of the final product indicated that a small proportion of gain patches (1.3% of the independent testing data) were non-forest by 2015, with low levels of tree cover (<10% cover). Patches with low tree cover in 2015 were generally typified by low levels of greenness (patch-level mean NDVI < 0.48) and low biomass (patch-level mean PALSAR HV decibel return <1,800). Thus, these two manually determined thresholds were used jointly to classify gain patches with very low tree cover as a separate third land-use class, 'open'. These thresholds readily captured bare ground and open water but sometimes included short plantations and secondary forests in arid regions with seasonal deciduousness.

Accuracy assessment. To permit assessment of how accurately patches of tree plantations and natural regrowth were distinguished across the tropics, we generated two different sets of labelled testing data. We first assessed model accuracy using randomly selected reference polygons ($n = 2,000$; Supplementary Fig. 2) sampled with respect to stratum weight (patch area), with replacement. Next, we assessed map accuracy (the gain polygons together with the background, non-gain area) using randomly located reference points ($n = 4,269$; Supplementary Fig. 3). Random points were stratified by biome, continent and land-use class, with distinct sampling densities within the humid ($n = 2,981$) and non-humid ($n = 1,288$) biomes. In each polygon and at each point, land use for the year 2015 was separately identified by two trained analysts using freely available high-resolution imagery (Google Earth, Bing and ArcGIS basemap).

Reference polygons (which consisted of whole gain polygons) were categorized into one of three broad land-use categories (natural regrowth, tree plantations and open) for accuracy assessment, on the basis of a set of rules evaluating per cent tree cover, land use and land cover (Supplementary Table 5) for the year 2015. See Supplementary Methods for details. We then calculated the area-weighted accuracy of our final model for the global study area (Supplementary Table 6)[66]. Derived confusion matrices show the count of reference data in a particular class (Supplementary Table 6) but the associated per cent accuracy statistics are polygon area-corrected estimates, following ref. [66]. We report overall model accuracy both including and excluding very large polygons (Supplementary Tables 6–8).

To allow for observed geolocation error in high-resolution imagery, the occurrence of tree plantations or natural regrowth within 30 m (one Landsat pixel) of the reference points was recorded (with 'no-gain' as the third possible land-use class). We then calculated area-corrected estimates of map accuracy and class

area, following best practices for stratified random sampling[67,68]. Sampling strata included both tree cover gain classes and the non-gain class and sampling was further stratified by biome and continent. To limit the impact of the large no-gain class on our estimates, we defined a regrowth-possible 'buffer class'[69] for accuracy assessment (described further in the Supplementary Methods).

In a separate analysis to assess the comparative accuracy of the GFC and TMF[17] products, we assessed land use at random points ($n = 2,536$) across the humid tropical biome for the period 2000–2012. See Supplementary Methods for details.

Patch-level analysis. Using the centroid of each observed patch of tree cover gain to locate patches in space, we examined the distribution of tree plantation and natural regrowth patches (Supplementary Figs. 5–10) with respect to (1) navigable bodies of water and (2) a metric of human influence (the human impact index (HII)[40]). Both the 2009 HII and water map rasters were created by Venter et al.[40] and had a spatial resolution of 1 km²; the water data included navigable rivers, lakes and the ocean. A raster of Euclidian distance to the nearest navigable water was calculated using Guidos Toolbox. For each patch, we extracted (1) distance to water and (2) HII for the centroid point as response variables. We then tested for significant differences in the distribution of the two focal land-use classes (tree plantations and natural regrowth) using a separate Kruskal–Wallis rank sum test for each response variable.

Country-level analysis. We summarized the predicted plantation and regrowth area in each tropical country by intersecting the centroids of predicted land-use polygons with a national border database[70]. Both polygon area and country area between 25° N to 25° S were calculated using the WGS-84 projection in the geosphere R package[71]. We compared our predicted plantation expansion (plantation gain patches between 2000 and 2012, surviving to 2015) in each country with the area of increase in plantations and tree crops reported to the FAO for the period 2000–2010.

Biome- and hotspot-level analysis. We also summarized the estimated plantation and regrowth area in each tropical biome[72] and biodiversity hotspot (ref. [73], v.2016.1). We intersected the centroids of predicted land-use polygons with published biome and hotspot datasets. Terrestrial hotspot areas were selected for analysis and the biome dataset contained ecoregions as a biome subunit. Polygon area, country area and biome area between 25° N to 25° S were calculated using the WGS-84 projection in the geosphere R package[71]. For each biome, ecoregion and hotspot, we calculated the total and percentage area of observed natural regrowth and tree plantation expansion. Additionally, in each ecoregion, we calculated the per cent of area occupied by intact forests in 2013 (ref. [74], 'intact forest dominance'). To assess how well ecoregion area and intact forest dominance predicted the total area of tree plantations and natural regrowth in each ecoregion, we conducted a multiple linear regression, with all area variables log-transformed. The type of tree cover gain was a categorical predictor which interacted with ecoregional area and intact forest dominance, respectively.

Protected area analysis. Protected area data (boundary, location and status) were derived from the 2018 World Database on Protected Areas (WDPA v.1.4, ref. [61]) and details on WDPA data preprocessing and analysis are available in the Supplementary Methods. To examine how tree plantation expansion and natural forest recovery affected PAs across our study region, we quantified the area of plantation and regrowth located inside PAs and outside but near PAs (≤1 km from the border). To limit the impact of misclassification (commission) error on estimates of plantation area within PAs, only PAs containing >5 ha of plantation area were considered to be affected by plantations; in our dataset, 5 ha is the approximate estimated mean size of tree plantations in Africa. We then used high-resolution imagery (see above for methods) to manually inspect all PAs with >5 ha of plantation expansion and removed incorrectly predicted plantation polygons from the affected PA's area estimate. To quantify the effect of human influence and accessibility on PAs, we calculated the HII[40] for each PA and compared the degree of plantation occurrence across PAs. The effect of IUCN PA ranking on the occurrence of plantations in PAs was also examined, across three plantation area thresholds (0, 5 and 50 ha). Due to marked differences in mean IUCN rank across continents and data nonlinearity, we compared the mean IUCN rank between parks with and without plantations (binomial response variable) using Kruskal–Wallis one-way analysis of variance tests. Tests were done separately for each continent and threshold.

Reporting summary. Further information on research design is available in the Nature Research Reporting summary linked to this article.

Data availability
All data needed to replicate our results are available in the article, online or the supplementary information. Manually generated training data are available from the corresponding author, M.E.F., upon reasonable request. Predicted map outputs can be downloaded from the Global Forest Watch data repository: https://data.globalforestwatch.org/content/pantropical-tree-plantation-expansion-2000-2012/about

Code availability
All Python code needed to replicate our input data from Google Earth Engine are available on github at https://github.com/dohyung-kim/plantation. All R code for data analysis are available from the corresponding author, M.E.F., upon reasonable request, with the main R scripts available on github.

References
1. Curtis, P. G., Slay, C. M., Harris, N. L., Tyukavina, A. & Hansen, M. C. Classifying drivers of global forest loss. *Science* **361**, 1108–1111 (2018).
2. Gibbs, H. K. et al. Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s. *Proc. Natl Acad. Sci. USA* **107**, 16732–16737 (2010).
3. Payn, T. et al. Changes in planted forests and future global implications. *Ecol. Manag.* **352**, 57–67 (2015).
4. Pendrill, F., Persson, U. M., Godar, J. & Kastner, T. Deforestation displaced: trade in forest-risk commodities and the prospects for a global forest transition. *Environ. Res. Lett.* **14**, 055003 (2019).
5. Hurni, K. & Fox, J. The expansion of tree-based boom crops in mainland Southeast Asia: 2001 to 2014. *J. Land Use Sci.* **13**, 198–219 (2018).
6. Vijay, V. et al. The impacts of oil palm on recent deforestation and biodiversity loss. *PLoS ONE* **11**, e0159668 (2016).
7. Heilmayr, R., Echeverría, C. & Lambin, E. F. Impacts of Chilean forest subsidies on forest cover, carbon and biodiversity. *Nat. Sustain.* **3**, 701–709 (2020).
8. le Maire, G., Dupuy, S., Nouvellon, Y., Loos, R. A. & Hakamada, R. Mapping short-rotation plantations at regional scale using MODIS time series: case of eucalypt plantations in Brazil. *Remote Sens. Environ.* **152**, 136–149 (2014).
9. Wang, M. M. H., Carrasco, L. R. & Edwards, D. P. Reconciling rubber expansion with biodiversity conservation. *Curr. Biol.* **30**, 3825–3832 (2020).
10. Lewis, S. L., Wheeler, C. E., Mitchard, E. T. A. & Koch, A. Restoring natural forests is the best way to remove atmospheric carbon. *Nature* **568**, 25–28 (2019).
11. Dave, R. et al. *Second Bonn Challenge Progress Report: Application of the Barometer in 2018* (IUCN, 2019).
12. Sloan, S., Meyfroidt, P., Rudel, T. K., Bongers, F. & Chazdon, R. The forest transformation: planted tree cover and regional dynamics of tree gains and losses. *Glob. Environ. Change* **59**, 101988 (2019).
13. Petersen, R. et al. *Mapping Tree Plantations with Multispectral Imagery: Preliminary Results for Seven Tropical Countries* (WRI, 2016).
14. Erb, K.-H. et al. Land management: data availability and process understanding for global change studies. *Glob. Change Biol.* **23**, 512–533 (2017).
15. Souza, C. M. et al. Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat Archive and Earth Engine. *Remote Sens.* **12**, 2735 (2020).
16. Miettinen, J. et al. Extent of industrial plantations on Southeast Asian peatlands in 2010 with analysis of historical expansion and future projections. *GCB Bioenergy* **4**, 908–918 (2012).
17. Vancutsem, C. et al. Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. *Sci. Adv.* **7**, eabe1603 (2021).
18. Puyravaud, J.-P., Davidar, P. & Laurance, W. F. Cryptic destruction of India's native forests. *Conserv. Lett.* **3**, 390–394 (2010).
19. Fagan, M. E. et al. Mapping pine plantations in the southeastern U.S. using structural, spectral, and temporal remote sensing data. *Remote Sens. Environ.* **216**, 415–426 (2018).
20. Tropek, R. et al. Comment on "High-resolution global maps of 21st-century forest cover change". *Science* **344**, 981 (2014).
21. *Global Forest Resources Assessment 2020* (FAO, 2020).
22. *FAOSTAT Agricultural Statistics Database* (FAO, 2019); http://faostat.fao.org/site/291/default.aspx
23. Cook-Patton, S. C. et al. Mapping carbon accumulation potential from global natural forest regrowth. *Nature* **585**, 545–550 (2020).
24. Hurni, K., Schneider, A., Heinimann, A., Nong, D. H. & Fox, J. Mapping the expansion of boom crops in mainland Southeast Asia using dense time stacks of Landsat data. *Remote Sens.* **9**, 320 (2017).
25. Miettinen, J., Shi, C. & Liew, S. C. 2015 Land cover map of Southeast Asia at 250 m spatial resolution. *Remote Sens. Lett.* **7**, 701–710 (2016).
26. Torbick, N., Ledoux, L., Salas, W. & M. Zhao, M. Regional mapping of plantation extent using multisensor imagery. *Remote Sens.* **8**, 236 (2016).
27. Azizan, F. A., Kiloes, A. M., Astuti, I. S. & Abdul Aziz, A. Application of optical remote sensing in rubber plantations: a systematic review. *Remote Sens.* **13**, 429 (2021).
28. Bégué, A. et al. Remote sensing and cropping practices: a review. *Remote Sens.* **10**, 99 (2018).

29. Bey, A. & Meyfroidt, P. Improved land monitoring to assess large-scale tree plantation expansion and trajectories in Northern Mozambique. *Environ. Res. Commun.* **3**, 115009 (2021).

30. Jucker, T. et al. Topography shapes the structure, composition and function of tropical forest landscapes. *Ecol. Lett.* **21**, 989–1000 (2018).

31. Féret, J.-B. & Asner, G. P. Spectroscopic classification of tropical forest species using radiative transfer modeling. *Remote Sens. Environ.* **115**, 2415–2422 (2011).

32. Poortinga, A. et al. Mapping plantations in Myanmar by fusing Landsat-8, Sentinel-2 and Sentinel-1 data along with systematic error quantification. *Remote Sens.* **11**, 831 (2019).

33. Gutiérrez-Vélez, V. H. et al. High-yield oil palm expansion spares land at the expense of forests in the Peruvian Amazon. *Environ. Res. Lett.* **6**, 044029 (2011).

34. Descals, A. et al. High-resolution global map of smallholder and industrial closed-canopy oil palm plantations. *Earth Syst. Sci. Data.* **13**, 1211–1231 (2021).

35. Ordway, E. M., Naylor, R. L., Nkongho, R. N. & Lambin, E. F. Oil palm expansion and deforestation in Southwest Cameroon associated with proliferation of informal mills. *Nat. Commun.* **10**, 114 (2019).

36. Heilmayr, R., Echeverría, C., Fuentes, R. & Lambin, E. F. A plantation-dominated forest transition in Chile. *Appl. Geogr.* **75**, 71–82 (2016).

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

38. Bond, W. J., Stevens, N., Midgley, G. F. & Lehmann, C. E. R. The trouble with trees: afforestation plans for Africa. *Trends Ecol. Evol.* **34**, 963–965 (2019).

39. Veldman, J. W. et al. Where tree planting and forest expansion are bad for biodiversity and ecosystem services. *Bioscience* **65**, 1011–1018 (2015).

40. Venter, O. et al. Global terrestrial Human Footprint maps for 1993 and 2009. *Sci. Data* **3**, 160067 (2016).

41. Fagan, M. E. A lesson unlearned? Underestimating tree cover in drylands biases global restoration maps. *Glob. Change Biol.* **26**, 4679–4690 (2020).

42. Bastin, J. F. et al. The extent of forest in dryland biomes. *Science* **356**, 635–638 (2017).

43. Fagan, M. E., Reid, J. L., Holland, M. B., Drew, J. G. & Zahawi, R. A. How feasible are global forest restoration commitments? *Conserv. Lett.* **13**, e12700 (2020).

44. Malkamäki, A. et al. A systematic review of the socio-economic impacts of large-scale tree plantations, worldwide. *Glob. Environ. Change* **53**, 90–103 (2018).

45. Schwartz, N. B., Aide, T. M., Graesser, J., Grau, H. R. & Uriarte, M. Reversals of reforestation across Latin America limit climate mitigation potential of tropical forests. *Front. For. Glob. Change* **3**, 85 (2020).

46. Noojipady, P. et al. Managing fire risk during drought: the influence of certification and El Niño on fire-driven forest conversion for oil palm in Southeast Asia. *Earth Syst. Dynam.* **8**, 749–771 (2017).

47. Bullock, E. L., Woodcock, C. E., Souza, C. Jr. & Olofsson, P. Satellite-based estimates reveal widespread forest degradation in the Amazon. *Glob. Change Biol.* **26**, 2956–2969 (2020).

48. Sloan, S. & Sayer, J. A. Forest Ecology and Management Forest Resources Assessment of 2015 shows positive global trends but forest loss and degradation persist in poor tropical countries. *Ecol. Manag.* **352**, 134–145 (2015).

49. Heinrich, V. H. A. et al. Large carbon sink potential of secondary forests in the Brazilian Amazon to mitigate climate change. *Nat. Commun.* **12**, 1785 (2021).

50. Potapov, P. et al. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sens. Environ.* **253**, 112165 (2021).

51. Bernal, B., Murray, L. T. & Pearson, T. R. H. Global carbon dioxide removal rates from forest landscape restoration activities. *Carbon Balance Manag.* **13**, 22 (2018).

52. Li, W., Goodchild, M. F. & Church, R. An efficient measure of compactness for two-dimensional shapes and its application in regionalization problems. *Int. J. Geogr. Inf. Sci.* **27**, 1227–1250 (2013).

53. Asner, G. P. Cloud cover in Landsat observations of the Brazilian Amazon. *Int. J. Remote Sens.* **22**, 3855–3862 (2001).

54. Wilson, A. M. & Jetz, W. Remotely sensed high-resolution global cloud dynamics for predicting ecosystem and biodiversity distributions. *PLoS Biol.* **14**, e1002415 (2016).

55. Gutiérrez-Vélez, V. H. & DeFries, R. Annual multi-resolution detection of land cover conversion to oil palm in the Peruvian Amazon. *Remote Sens. Environ.* **129**, 154–167 (2013).

56. Reiche, J. et al. Combining satellite data for better tropical forest monitoring. *Nat. Clim. Change* **6**, 120–122 (2016).

57. Erinjery, J. J., Singh, M. & Kent, R. Mapping and assessment of vegetation types in the tropical rainforests of the Western Ghats using multispectral Sentinel-2 and SAR Sentinel-1 satellite imagery. *Remote Sens. Environ.* **216**, 345–354 (2018).

58. Shimada, M. et al. New global forest/non-forest maps from ALOS PALSAR data (2007–2010). *Remote Sens. Environ.* **155**, 13–31 (2014).

59. Torres, R. et al. GMES Sentinel-1 mission. *Remote Sens. Environ.* **120**, 9–24 (2012).

60. Potapov, P. et al. The last frontiers of wilderness: tracking loss of intact forest landscapes from 2000 to 2013. *Sci. Adv.* **3**, e1600821 (2017).

61. *World Database on Protected Areas User Manual 1.4* (UNEP-WCMC, 2016).

62. *AutoML: Automatic Machine Learning* (H2O.ai, 2020); https://h2o-release.s3.amazonaws.com/h2o/rel-yau/5/docs-website/h2o-docs/automl.html

63. Healey, S. P. et al. Mapping forest change using stacked generalization: an ensemble approach. *Remote Sens. Environ.* **204**, 717–728 (2018).

64. Lagomasino, D. et al. Measuring mangrove carbon loss and gain in deltas. *Environ. Res. Lett.* **14**, 25002 (2019).

65. Bunting, P. et al. The global mangrove watch—a new 2010 global baseline of mangrove extent. *Remote Sens.* **10**, 1669 (2018).

66. Pickens, A. H. et al. Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series. *Remote Sens. Environ.* **243**, 111792 (2020).

67. Olofsson, P. et al. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **148**, 42–57 (2014).

68. Stehman, S. V. Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes. *Int. J. Remote Sens.* **35**, 4923–4939 (2014).

69. Olofsson, P. et al. Mitigating the effects of omission errors on area and area change estimates. *Remote Sens. Environ.* **236**, 111492 (2020).

70. *Database of Global Administrative Areas (GADM) v.3.6* (GADM, 2018); https://gadm.org/download_country_v3.html

71. Hijmans, R. J., Williams, E., Vennes, C. M. & Hijmans, M. R. J. Package 'geosphere' version 1.5-10. Spherical trigonometry (2017).

72. Olson, D. M. et al. Terrestrial ecoregions of the world: a new map of life on Earth: a new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *Bioscience* **51**, 933–938 (2001).

73. Mittermeier, R. A., Turner, W. R., Larsen, F. W., Brooks, T. M. & Gascon, C. in *Biodiversity Hotspots: Distribution and Protection of Conservation Priority Areas* (eds Zachos, F. E. & Habel, J. C.) 3–22 (Springer, 2011).

74. Potapov, P. et al. The last frontiers of wilderness: tracking loss of intact forest landscapes from 2000 to 2013. *Sci. Adv.* **3**, e1600821 (2017).

Acknowledgements

Author contributions

M.E.F. and D.H.K. were responsible for conceptualization and formal analysis. M.E.F. undertook visualization, project administration and supervision. M.E.F., D.H.K. and A.T. developed the methodology. M.E.F., L.F., J.D., H.C., W.S., J. Slaughter, J. Schaferbien and A.T. conducted validation. N.L.H., E.G. and E.M.O. provided resources (datasets). M.E.F., D.H.K., N.L.H., A.T., E.G. and E.M.O. were responsible for writing.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41893-022-00904-w.

Correspondence and requests for materials should be addressed to Matthew E. Fagan.

Peer review information *Nature Sustainability* thanks Sean Sloan and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature portfolio

Corresponding author(s): Matthew E. Fagan

Last updated by author(s): 03-24-2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Google Earth Engine and QGIS version 3.10 were used for data collection in this study, along with several publicly available data sources listed in the Methods and Supplementary Information.
Data analysis	Data analysis was completed using Python 2.7 and R 4.0.3. Our code is available following our code availability statement: "All Python code needed to replicate our input data from Google Earth Engine are available on github at https://github.com/dohyung-kim/plantation. All R code for data analysis are available from the corresponding author, MEF, upon reasonable request, with the main R scripts available on github."

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data needed to replicate our results are available in the article, online, or the supplementary information. Manually-generated training data are available from the corresponding author, MEF, upon reasonable request. Predicted map outputs can be downloaded from the Global Forest Watch data repository: https://data.globalforestwatch.org/content/pantropical-tree-plantation-expansion-2000-2012/about

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences   ☐ Behavioural & social sciences   ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This is a machine learning study conducted with training, testing, and validation data to output and assess a map. The training and testing sample sizes are very large (>10,000 samples), and the independent validation data was large relative to the region of interest (n=4269). All statistical analyses were conducted using either standard unbiased estimators (for map area and accuracy) or nonparametric ANOVAs (analyzing trends across the map, or protected areas).
Research sample	For our analyses, the map units we predicted and sampled from consist of mapped patches of tree cover gain identified by Hansen et al. (2013). Sample size varied depending on the analysis, but generally exceeded 10,000 samples across all analyses.
Sampling strategy	Only two sets of analyses used statistical inference. For our estimates of map area, we based our sample size on the resulting confidence intervals around our area estimates, as well as estimates of further improvements if we increased sample size. For our statistical analyses, we used all available data (not sub-samples).
Data collection	Our training and testing data was derived by trained analysts. Our validation data was derived independently, by a team of at least two analysts, with the PI making the final call in cases of confusion. See our Supplementary Information for more details.
Timing and spatial scale	The spatial scale is the global tropics, between 25 degrees North and South. The timing is the years 2000-2012, persisting to 2015.
Data exclusions	No data were excluded from our analyses.
Reproducibility	Our global-scale predictions of tree plantation area were consistent across numerous different machine learning algorithms.
Randomization	Our validation data was a stratified random sample across continents, biomes, and land cover classes, with sampling density proportional to the area of the strata.
Blinding	Our validation data labels were derived without reference to the predictions of the machine learning model.

Did the study involve field work?   ☐ Yes   ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒ ☐	Antibodies
☒ ☐	Eukaryotic cell lines
☒ ☐	Palaeontology and archaeology
☒ ☐	Animals and other organisms
☒ ☐	Human research participants
☒ ☐	Clinical data
☒ ☐	Dual use research of concern

Methods

n/a	Involved in the study
☒ ☐	ChIP-seq
☒ ☐	Flow cytometry
☒ ☐	MRI-based neuroimaging