SUPPLEMENTAL MATERIALS

Supplemental Text: Descriptions of global forest (land) cover and change datasets

1. Land cover datasets (forest classification)

1.1. Copernicus Global Land Cover (version 3) (CGLC)

The Copernicus Global Land Cover (CGLC) [S1] dataset offers high resolution (100-m) annually available data on forest cover types and fractional tree cover (0-100%) since 2015. Currently, the CGLC data is available for the years 2015-2019, with the intent to continue releasing new annual data periodically. CGLC is fully free and open access to all users.

CGLC uses 23 discrete classes for the land cover classification scheme, 12 of which are forest cover classes. CGLC defines forest classes as open or closed for each: evergreen needleleaf, evergreen broadleaf, deciduous needleleaf, deciduous broadleaf, mixed, and unknown forest types. Open forests are defined by a top-layer tree canopy cover of 15-70%, with a second layer of mixed shrubs and grasslands. Closed forests are defined by a tree canopy cover of >70%. Along with the discrete land cover classification data layer, CGLC provides a "classification probability" data layer, which provides a quality indicator of classification accuracy, and a "change confidence" data layer, which indicates confidence in detecting discrete class changes across years. The fractional land cover maps provide continuous (0-100%) land cover classification for ten different land classes, including a tree cover fraction layer. These fractional land cover layers allow users to tailor land cover definitions for specific applications, e.g., heterogeneous land cover areas, areas with low density tree cover, or agroforestry.

The earliest version of CGLC used imagery collected by the PROBA-V satellite, but CGLC now uses Sentinel-2 data to achieve higher global accuracy. CGLC uses a supervised classification algorithm. The process is broken into two parts: the generation of epoch-wise land cover classifications for each reference year (based on the data range +/- 1 year from the reference year), followed by temporal post-processing to generate consistent land cover classifications across all years [S2]. Ancillary datasets are included via expert rules to generate the annual land cover maps and make them temporally consistent. Importantly, the classification algorithm was eco-region specific (biome clustered), meaning that the training data were clustered by biome to optimize the models for each eco-region. Unlike several of the other datasets we review, CGLC does not include a definition of minimum tree height in defining forest classes or tree cover since it relies on optical remotely sensed data.

The current version achieves a reported overall accuracy of just over 80% (80.6% in 2015, 80.3% in 2019) when compared to 28,000 independent validation points [S3]. By region, the overall accuracy was highest in Asia (83.7%) and lowest for North America (77.6%) [S3]. The overall accuracy of the land cover change/no-change map was 99.6% [S3]. The user's accuracy for closed forests was 82% and for open forests was 62%. With optical remotely sensed data as used in CGLC, classification accuracy in areas with high yearly cloud cover can be lower, which means that not all world regions are classified accurately. Additionally, highly fragmented landscapes, such as mosaiced areas of forest and small cropland fields, have lower classification accuracy. Depending on how dense or sparse the small cropland fields in a region are, these misclassifications can lead to an overestimation or underestimation of cropland areas for the region. The classifier also had difficulty distinguishing between open forests and shrublands, especially where neighboring pixels were also fuzzy between these two classes.

1.2 ESA WorldCover (ESA-WC)

The European Space Agency (ESA) generated WorldCover [S4,5] (ESA-WC) as a global, 10-m resolution baseline land cover dataset for the years 2020 and 2021. ESA-WC is published open access and was derived from Sentinal-1 and Sentinal-2 imagery. Note that ESA has also published 300-m resolution land cover maps since 1992, as described in the following section. The 'tree cover' class is defined as areas dominated by trees with a cover of 10% or more, possibly with other land cover classes present below the canopy. This class includes afforested areas, plantations, and areas seasonally or permanently flooded with fresh water. However, the 'tree cover' class excludes mangroves, which are classified as a separate land cover class.

The ESA-WC classification model was based on the CGLC model, and it combines Sentinel-1 C-band Synthetic Aperture Radar, Sentinel-2 multi-spectral image data, and multiple ancillary datasets to classify land cover. ESA-WC was generated using a supervised classification, specifically the gradient boosting decision tree algorithm (CatBoost). ESA-WC uses a single methodology applied over all world regions, which means that accuracy varies by location. The algorithms were different for the 2020 and 2021 maps, so there are concerns about using these data for land cover change since changes could be due to either real change or to algorithm differences. There are noted issues with classification confusion between tree cover and shrub cover, especially for the USA, Argentina, Indonesia, and all of Africa and Europe [S6].

The overall accuracy against the validation set was 74.4% for 2020 and 76.7% for 2021. For 2020, the producer's accuracy for the tree cover class was 89.9% and user's error was 80.8% [S6]. For 2020, The overall accuracy was highest for Asia (80.7%), followed by Europe (76.8%), South America (76.1%), Africa (73.6%), North America (72.2%), Eurasia (70.2%), and Oceania (67.5%) [S6]. For 2021, the producer's accuracy for the tree cover class was 91.9% and user's error was 80% [S6]. For 2020, The overall accuracy was highest for Asia (82.1%), followed by Europe (77.9%) and South America (77.9%), Africa (76.5%), North America (74.6%), and Eurasia (72.5%) and Oceania and Australia (72.5%) [S6].

1.3 ESA-CCI-LC/C3S-LC (ESA-C3S-LC)

The European Space Agency Climate Change Initiative Land Cover (ESA-CCI-LC) [S7] and Copernicus Climate Change Service global Land Cover (C3S-LC) [S8] provides global land cover maps at 0.00278° (about 300-m at the equator) resolution annually from 1992-2015 and 2016-2020, respectively. C3S-LC was generated to be consistent with ESA-CCI-LC. The data are freely available and published open access. The ESA more recently released the 100-m resolution Copernicus Global Land Cover dataset (CGLC) and the 10-m resolution WorldCover dataset, both described in previous sections.

The land cover classes are defined by the United Nations (UN) Food and Agriculture Organization (FAO) Land Cover Classification System (LCCS). This classification system classifies pixels as land with tree cover >15% of types (1) broadleaf evergreen, (2) broadleaf deciduous, (3) needleleaf evergreen, (4) needleleaf deciduous, and (5) mixed leaf type (broadleaf and needleleaf). The system also includes classes for (i) mosaic of cropland (>50%) with natural vegetation (tree, shrub, herbaceous cover) (<50%), (ii) mosaic of natural vegetation (tree, shrub, herbaceous cover) (>50%) with cropland (<50%), (iii) mosaic of trees and shrubs (>50%) with herbaceous cover (<50%), and (iv) mosaic of herbaceous cover (>50%) with trees and shrubs (<50%). Separate classes are made for tree cover in (a) flooded (fresh or brackish water) and in (b) flooded (saline water). For land cover change detection, forests are defined as (1)-(5), (iii), (a), and (b).

The classification algorithm consists of two classifiers: a supervised machine learning spectral classification (Gaussian Maximum Likelihood) and an unsupervised classification (Iterative Self-Organizing Data Analysis Technique (ISODATA) clustering technique). The results from the two classifiers are merged to generate the land cover maps [S9,10]. The overall accuracy of ESA-C3S-LC for each year from 2016-2020 ranged between 70.5%-71.1% [S11]. The accuracy assessment for 2016-2020 used 2,945 independent validation datapoints. The user's accuracy for forest classes ranged from 67-90%. However, there are some known areas that have potentially lower accuracy due to limited coverage, including the western part of the Amazon basin, Chile, the southern part of Argentina, the western part of Congo basin, the Gulf of Guinea, the eastern part of Russia, the eastern coast of China, and Indonesia. Additionally, the land cover change dataset does not capture change between forest classes (e.g., deciduous to mixed forest type).

1.4 ESRI Land Cover (ESRI-LC)

The ESRI Land Cover (ESRI-LC) dataset, also known as the Sentinel-2 10m land cover time series of the world from 2017-2022 [S12] offers 10-m spatial resolution land cover data annually for 2017-2022. The dataset classifies land cover using 10 classes. ESRI-LC is published open access, with plans to continue releasing updates annually. The ESRI-LC dataset defines forests as any significant clustering of trees ≥15 m, typically with closed or dense canopy. Examples include wooded vegetation and dense clusters of tall vegetation in savannas, plantations, swamps, or mangroves [S12]. Areas with very sparse wooded vegetation are captured by the "scrub/shrub" land cover class, rather than the forest class.

ESRI-LC is derived from Sentinal-2 imagery using a supervised classification deep learning algorithm, specifically a fully convolutional neural network with a UNet architecture [S13]. The training data consists of 24,000 5km x 5km images, which yields about 6 billion labeled 10m image pixels, generated by the National Geographic Society. The algorithm is run on Microsoft's Planetary Computer, which allows the annual data layers to be generated rapidly and with a high resolution. ESRI's land cover classification model achieved an overall accuracy of 76-91% on the validation dataset [S12]. In their analysis, tree cover classification had 82-97% accuracy for case analyses of California, Costa Rica, Belgium, and Laos. The 2017 land cover data are based on fewer training images, so this year may have lower accuracy than 2018-2022.

1.5 Global Land Cover and Land Use Change, 2000-2020 (GLAD-LCLUC)

The GLAD Global Land Cover and Land Use Change (GLAD-LCLUC) maps forest cover extent in 2000 and 2020 as well as the total loss and gain in forest cover over this period [S14]. Forests were defined using the Landsat-based global tree height model calibrated for GEDI observations. Forests were defined as pixels with ≥ 5 m forest height.

The forest extent change (net forest extent loss and gain) was derived directly from the year 2000 and 2020 map comparison. Net forest height increase was classified only if a pixel (i) demonstrated net forest extent gain, or (ii) has net forest height increase from the year 2000 to 2020 by ≥ 100%. Net forest height loss was calculated only for forest loss pixels or in case the net height reduction was by ≥ 50% of the year 2000 value. Areas with small differences between the years 2000 and 2020 forest height data that have no indication of forest disturbance were considered stable forests. This dataset also includes data indicating forest disturbances, where forests experienced significant or stand-replacement disturbances between 2000-2020.

1.6 JAXA Forest/Non-Forest map (JAXA-FNF)

The Global PALSAR-2/PALSAR Forest/Non-Forest map (JAXA-FNF) [S15] is a 25-m resolution dataset produced by the Japan Aerospace Exploration Agency (JAXA) Earth Observation Research Center (EORC). The FNF dataset is free to use under their terms of use, but JAXA retains the copyright ownership of the data.

FNF is generated using the backscattering intensity to classify 'forest' and 'non-forest' areas globally annually between 2007-2010 and 2015-2020, with additional years forthcoming. JAXA-FNF is derived from the Japanese L-band Synthetic Aperture Radars (PALSAR and PALSAR-2) on the Advanced Land Observing Satellites (ALOS and ALOS-2). FNF defines forest as areas larger than 0.5 hectares and forest cover ≥10% with trees 5-m or taller, based on the Food and Agriculture Organization (FAO) definition of forests [S16]. It does not include land predominantly under agricultural or urban land use. For 2007-2017, the classification algorithm uses a conventional threshold method, meaning that it uses region-dependent thresholds for the backscatter parameter that defines forests to account for variation between biomes. For 2017-2020, the classification algorithm uses a random forest model to improve accuracy by 5-10% and divide forests into two categories: forest cover ≥90% and forest cover ≥10% and <90% [S16]. The random forest classifier was applied on a continental basis.

The average classification accuracy was found to be 84.9% when compared to validation data derived from Google Earth Imagery and 91.3% when compared to the Degree Confluence Project validation data [S15].

1.7 MODIS Land Cover Type (collection 6) (MODIS-LC)

MODIS Land Cover Type (MCD) [S17] provides global land cover types at 500 m resolution annually since the year 2001 through 2020, with continued annual releases expected. MCD is fully free and open access to all users, and it is derived from NASA Terra and Aqua satellite imagery. MCD data layers are available for six different land cover classification schemes – (1) the Annual International Geosphere-Biosphere Programme (MODIS-IGBP) classification, (2) the Annual University of Maryland (MODIS-UMD) classification, (3) the Annual Leaf Area Index (MODIS-LAI) classification, (4) the Annual BIOME-Biogeochemical Cycles (MODIS-BGC) classification, (5) the Annual Plant Functional Types (MODIS-APFT) classification, and (6) the FAO-Land Cover Classification System. The FAO-Land Cover Classification System includes a land cover layer (MODIS-FAO), a land use layer (MODIS-FAO-LCC2), and a surface hydrology layer (MODIS-FAO-LCC3), as well as classification confidence layers for each of these three.

In relation to forest land cover, the different MCD classification schemes use different forest definitions. The MODIS-IGBP, MODIS-UMD, MODIS-LAI, and MODIS-FAO schemes define forests as evergreen, deciduous, or mixed, and needleleaf or broadleaf, with tree cover > 60% and canopy height > 2 m. The MODIS-IGBP, MODIS-UMD, and MODIS-LAI schemes define savannahs separately with tree cover 10-60% and canopy height > 2 m, and MODIS-IGBP and MODIS-UMD include a class for cropland/natural vegetation mosaics, defined as mosaics of small-scale cultivation 40-60% with natural tree, shrub, or herbaceous vegetation. The MODIS-FAO scheme includes classes for open forests, with tree cover 30-60% and canopy height >2 m, and sparse forests, with tree cover 10-30% and canopy height >2 m. The MODIS-FAO-LCC2 scheme includes dense forests (>60% tree cover, >2 m canopy height), open forests (10-60% tree cover, >2 canopy height), and forest/cropland mosaics (mosaics of small-scale cultivation 40-60% with >10% natural tree cover). The MODIS-BGC and MODIS-APFT schemes define forests as evergreen, deciduous, or mixed, and needleleaf or broadleaf, with tree cover >10% and canopy height >1 m or >2 m, respectively.

MODIS MCD is derived using a supervised classification algorithm on MODIS Terra and Aqua reflectance data. The method applies post-processing to add prior probability knowledge and additional information to further refine the classes. The creators of MODIS MCD emphasize that the product should not be used to determine post-classification land cover change because the uncertainty in land use classification makes it difficult to determine land use change for similar classes [S18]. The overall accuracy of the IGBP layer is estimated to be 67% globally [S19]. The MODIS-FAO/FAO-LCC2/FAO-LCC3 schemes had moderately higher accuracy – the MODIS-FAO layer had an overall accuracy of 74%, MODIS-FAO-LCC2 81%, and MODIS-FAO-LCC3 87% [S19]. The user's accuracy for the forest classes ranged from about 60-88%. There are known issues with misclassification of temperate evergreen needleleaf forests as broadleaf evergreen forests in Japan, Chile, and the Pacific Northwest of the US. Similarly, areas of evergreen broadleaf forests are misclassified as evergreen needleleaf forests in Australia and parts of South America.

2. Forest cover change datasets

2.1 Forest Cover Change in the Humid Tropics (Vancutsem-TMFCC)

The Forest Cover Change in Tropical Moist Forests (Vancutsem-TMFCC) dataset is a 30-m resolution long-term forest cover change dataset available for the humid tropics from 1990-2022 [S20]. The European Commission's Joint Research Centre developed this dataset using 41 years of Landsat time series. The dataset is available for much of the land within the band of the tropics, but their classification algorithm is specifically for tropical moist forests. The *transition map* includes multiple classes, including (1) Undisturbed tropical moist forests (TMF), (2) degraded TMF, (3) TMF regrowth, (4) deforested land – forest converted to tree plantations, (5) deforested land – forest converted to water, (6) deforested fand – forest converted to other land cover, (7) recent deforestation or degradation (2019-2021), (8) permanent or seasonal water, and (9) other land cover (including afforestation). For a pixel to be classified as TMF regrowth, the pixel needed to show regrowth over a period of three years before it is included in the regrowth class. Deforestation indicates that the pixel had a change in land cover from forest to non-forest, and degradation indicates that the pixel had some type of temporary disturbance (e.g., logging, fires, weather events).

The *annual change collection* contains data on TMF extent, disturbances (deforestation and degradation), and forest regrowth (post-deforestation) for each year between 1990 and 2021. The year of deforestation or degradation is the first year when TMF was disturbed, which can be followed by regrowth or not. The duration indicates the number of days of the disturbance, the annual number of disruption observations is the number of disturbances for each year between 1982 and 2021, and the intensity is the total number of disturbances.

2.2 Hansen Global Forest Change (version 1.8) (Hansen-GFC)

The Hansen Global Forest Change (Hansen-GFC) [S21] is a 30-m resolution dataset that characterizes forest extent and change over time using timeseries Landsat imagery. GFC is free and open access to all users. GFC includes layers for tree cover extent for the year 2000, forest loss year for 2001-2021, and total forest gain for the period 2000-2012. Trees were defined as all vegetation taller than 5 m, and the 2000 tree cover data layer estimates the percent tree cover (0-100%) in 2010. The reference point for forest classification is 50% tree cover. Forest loss year is defined as the year of gross forest cover loss (a stand-replacement disturbance or a change from a forest to non-forest state, i.e., >50% canopy cover to nearly 0% canopy cover). It is encoded as either a 0 representing no loss or a value in the range 1-20, representing the loss detection year between 2001-2020. Forest gain is defined as a non-forest to forest change entirely within the period between 2000 and 2012.

In a validation assessment, the false positivity rate (commission errors) was estimated at 13% and false negativity rate (omission errors) at 12%. The accuracy is known to vary by biome. For example, the accuracy is lower in sub-Saharan Africa where disturbances in smallholder landscapes are often missed by the model. Additionally, there is the problem of accurately classifying loss year correctly. The validation assessment concluded that there was 75% confidence that forest loss occurred in the classified year and 97% confidence that it occurred within one year before or after the classified loss year. A full validation assessment has not yet been conducted for the model implemented from 2011 onwards using Landsat 8 data.

Importantly, the tree cover, loss, and gain data sets cannot be accurately combined or compared against each other due to variation in the research methodologies. Users are therefore unable to compute net forest cover change based on these data. Additionally, the methodology for producing forest cover loss data was changed after 2010, so comparisons between the original 2000-2010 data and the updated 2011-2020 data are not recommended without further investigation of the data for the specific application.

3. Additional datasets, historical land cover datasets, and other useful resources

3.1 Global Forest Cover Change Tree Cover

The Global Forest Cover Change Tree Cover dataset (GFCC) [S22] provides 30-m resolution tree canopy cover (0-100%) data from 2000, 2005, 2010, and 2015. Tree cover is defined as woody vegetation greater than 5 meters in height. It is derived from a Landsat-based rescaling of the MODIS Vegetation Continuous Fields 250-m dataset.

3.2 GlobeLand30

GlobeLand30 is a global 30 m resolution land cover data set available for the years 2000, 2010, and 2020, and it may continue to be released every ten years in the future. The dataset is derived from Landsat, HJ-1 (the China Environment and Disaster Reduction Satellite), and GF-1 (China High Resolution Satellite) multispectral image GlobeLand30 defines forests as lands covered with trees with canopy cover >30%, and it includes classes for deciduous broadleaf forests, evergreen broadleaf forests, deciduous coniferous forests, evergreen coniferous forests, and mixed forest. GlobeLand30 also includes a class for sparse woodlands, defined by tree canopy cover of 10%-30%.

GlobeLand30 2000 and GlobeLand30 2010 achieved an overall classification accuracy of 80% in eight selected areas, based on an assessment using 154,586 pixel samples [S23]. The analysis found that the classification accuracy of forest areas was slightly higher, at 83.6% [S23]. The total accuracy of

GlobeLand30 2010 was found to be 83.5% and for GlobeLand30 2020 was found to be 85.7% based on a validation test set of over 230,000 points [S24].
GlobeLand30 is not included in our following analyses since it is not currently available on Google Earth Engine, and we are unable to view the registration agreement at the data download portal.

3.3 Terra MODIS Vegetation Continuous Fields (collection 6)

Terra MODIS Vegetation Continuous Fields (VCF) [S25] provides data on basic vegetation surface cover annually since the year 2000 at 250 m resolution. VCF is fully free and open access to all users. VCF includes data layers for percent tree cover, percent non-tree cover, and percent bare land. Also available are layers for the standard deviation of pixel values for the percent tree cover and percent non-tree cover datasets, based on the 30 models used to generate the surface cover values, and “quality” data layers based on pixel cloudiness, aerosol levels, and view zenith. VCF is derived from NASA Terra and Aqua satellite imagery.

The percent tree cover data layer can be used to identify forested areas based on user definitions. VCF had limited validated against ground truth data in Maryland and Mato Grasso, Brazil. Based on this validation, the most recent VCF data had a Root Mean Square Error (RMSE) of 9.47% and a Mean Absolute Error (MAE) of 7.87% in Maryland, and a 10.46% RMSE and 9.40% MAE in Mato Grosso [S26]. Another study found that MODIS VCF could be underestimating tree cover by 9-15% in tropical forests and savannahs [S27].

3.4 NASA’s GEDI Forest Canopy Height map

The NASA GEDI Forest Canopy height dataset is a 30-m resolution global forest canopy height map for 2019 [S28]. The NASA GEDI lidar instrument was used to measure the vegetation structure, from which the forest canopy height was derived.

3.5 Spatial Database of Planted Trees (SDPT Version 1.0), 2015

The Spatial Database of Planted Trees (SDPT) is a compiled database of planted trees (plantations) for 82 countries in 2015. The country-level planted tree maps are derived from supervised classification or manual polygon delineation of Landsat, SPOT or RapidEye satellite imagery [S29].

3.6 FAO Global Land Cover - SHARE (GLC-SHARE)

The FAO Global Land Cover - SHARE (GLC-SHARE) is a global land cover dataset published in the year 2014 and available at 30 arc-second (approximately 1-km) resolution based on data collected from 1998-2012. The United Nations Food and Agriculture Organization (UN FAO) partnered with countries globally to collate the best available data for each nation. Some countries had much higher spatial resolution data available, and there are inconsistent methodologies for data collection across countries. The tree cover class is defined by areas dominated by tree plants with cover ≥10% [S30]. Afforestation and plantation areas are included in this class as well as tree covered areas that are permanently or seasonally flooded (though mangroves are excluded). The overall accuracy is reported at 80.2%, with the tree cover class achieving an accuracy between 91.8%-94.9%.

3.7 Finer Resolution Observation and Monitoring of Global Land Cover

Finer Resolution Observation and Monitoring of Global Land Cover (FROM-GLC) is a 30-meter resolution global land cover map available for 2010, 2015, and 2017 [S31]. FROM-GLC s derived from Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) data. FROM-GLC10, a 10-meter resolution land cover map derived from Sentinel-2 imagery, is also available for 2017 only. Forest classes are defined as broadleaf forests, needleleaf forests, mixed forest, and orchards. The three forest classes (excluding orchards) are limited to areas with >15% tree cover and >3-meter tree height. The overall accuracy of FROM-GLC was reported as 64.9-72.4% [S31,32], and the user’s accuracy for the aggregated forest classes was 80.5-84.2% [S31,32].

3.8 Global Land Surface Satellite Land Cover

Global Land Surface Satellite Land Cover (GLASS-GLC) is a 5-km resolution global land cover dataset spanning 34 years, from 1982-2015 [S33]. The "forest" land cover class includes broadleaf, needleleaf, and mixed leaf with tree cover ≥10% and tree height > 5-meters. The overall accuracy reported was 82.8% [S33].

3.9 GlobCover

The European Space Agency (ESA) released GlobCover, a 300-m landcover map derived from the MERIS full resolution surface reflectance time series data, for two periods: December 2004 – June 2006 and January-December 2009 [S34]. GlobCover defines forests based on the UN LCCS as: (1) broadleaved evergreen or semi-deciduous forest (>15% tree cover >5m); (2) broadleaved deciduous forest (>40% tree cover >5m); (3) broadleaved deciduous forest/woodland (15-40% tree cover >5m); (4) needleleaved evergreen forest (>40% tree cover >5m); (5) needleleaved deciduous or evergreen forest (15-40% tree cover >5m); (6) mixed broadleaved and needleleaved forest (>15% tree cover >5m); or (7) mixed forest (>15% tree cover >5m). GlobCover also includes mosaic land classes of cropland/vegetation (including forests) and of forest/grassland as well as classes for regularly flooded forests with fresh/brackish water, saline/brackish water, or fresh/brackish/saline water.

The overall accuracy of GlobCover is 70.7%, based on 1,408 homogenous independent validation datapoints, and 58%, based on 2,190 heterogenous validation points [S35]. As with ESA-CCI-LC/C3S-LC, which also uses the MERIS data, GlobCover has issues with limited data coverage in regions such as the Amazon basin. There are known issues with misclassification of tropical forests, flooded forests, and forest type based on leaf type and phenology. The GlobCover maps cannot be used for land cover change detection.

3.10 Global Land Survey

The Global Land Survey (GLS) is a 30-m resolution land cover dataset generated in partnership between the U.S. Geological Survey (USGS) and the National Aeronautics and Space Administration (NASA) [S36]. GLS is available for the years 1975, 1990, 2000, 2005, and 2010, though there are significant spatial gaps for the years 1975 (12.6% missing), 1990 (6.55% missing); for the years 2000-2010 only 0.21-0.59% of land area is not covered [S37]. It is derived from Landsat 1 through Landsat 7 ETM+ data collected from 1972-2011.

3.11 Global Land Cover Characterization

The USGS Global Land Cover Characterization (GLCC) is a global land cover dataset derived from imagery collected from April 1992 – March 1993 at 1-km resolution. GLCC is documented to have a 66.9% accuracy [S38]. Forests are defined as deciduous broadleaf, deciduous needleleaf, evergreen broadleaf, evergreen needleleaf, and mixed forests [S39].

3.12 IGBP-DISCover

The International Geosphere-Biosphere Program Data and Information System's DISCover (IGBP-DIS) land cover dataset is derived from imagery collected from April 1992 – March 1993 at 1-km resolution [S40,41]. Forests are defined as deciduous broadleaf, deciduous needleleaf, evergreen broadleaf, evergreen needleleaf, and mixed forests. IGBP-DISCover is documented to have a 66.9% accuracy [S38].

3.13 GLC 2000

GLC 2000 is a 1-km resolution land cover dataset available for November 1999-December 2000. GLC 2000 is documented to have a 68.6% accuracy [S38,42]. Forests are defined as deciduous broadleaf, deciduous needleleaf, evergreen broadleaf, evergreen needleleaf, mixed forests, regularly flooded (fresh and brackish water or saline water), mosaic tree cover/other natural vegetation, and burnt forest. Forests are tree covered areas with tree height > 3m and tree cover >15% (closed cover >40%) [S42].

3.14 University of Maryland Land Cover

The University of Maryland land cover (UMD) map is a 1-km resolution land cover dataset available for 1992-1993 [S43]. Forests are defined as deciduous broadleaf, deciduous needleleaf, evergreen broadleaf, evergreen needleleaf, mixed forests, and woodlands. Forests are tree covered areas with tree height > 5m and tree cover >60%, and woodlands are defined as tree covered areas with canopy cover >40% and less than 60% and tree height >5m [S43]. UMD is documented to have a 65% accuracy [S38].

3.15 Global Land Cover by National Mapping Organizations

Global Land Cover by National Mapping Organizations (GLCNMO) used MODIS (Terra and Aqua) data to map land cover for 2003 [S44], 2008 [S45], and 2013 [S46] at a resolution of 30 arcseconds (~1-km) for 2003 and 15 arcseconds (~500-m) for 2008 and 2013. The land cover classification system is based on the FAO LCCS, with forest classes of broadleaf evergreen, broadleaf deciduous, needleleaf evergreen, needleleaf deciduous, mixed forest, and open trees. Countries cooperated in providing training data (44 countries in 2003, 14 countries in 2008, and 22 countries in 2013). The overall accuracy reported was 74.8-77.9%, and the user's accuracy for the aggregated forest classes was 88-88.6% [S44-46].

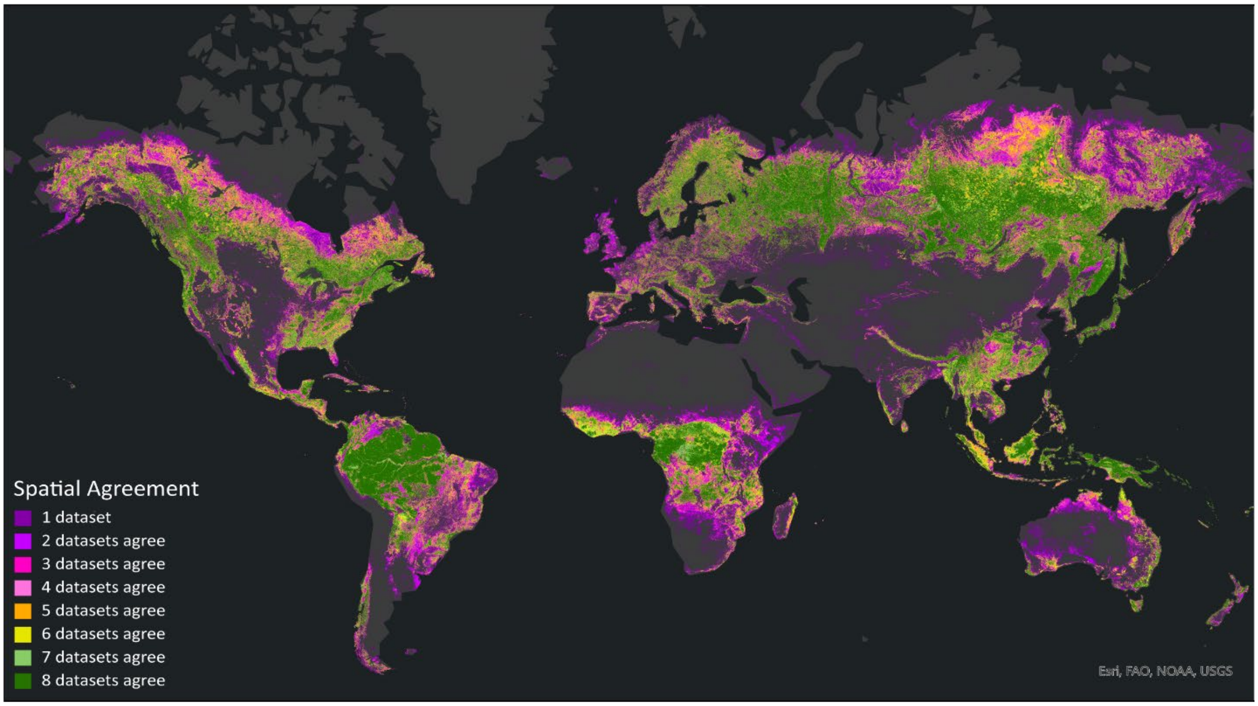

Fig. S1. Spatial agreement of forest cover classifications between eight land cover datasets, including the savanna land cover class for MODIS-IGBP. Spatial agreement is defined as the number of datasets that define a pixel as forest, between 1 and 8. Full agreement among all seven datasets corresponds to a value of eight (dark green), and no agreement among the datasets corresponds to a value of 1 (dark purple). No color (gray) indicates that none of the datasets classified the area as forest. All data are from the year 2019 except for ESA-WC, which is from 2020.

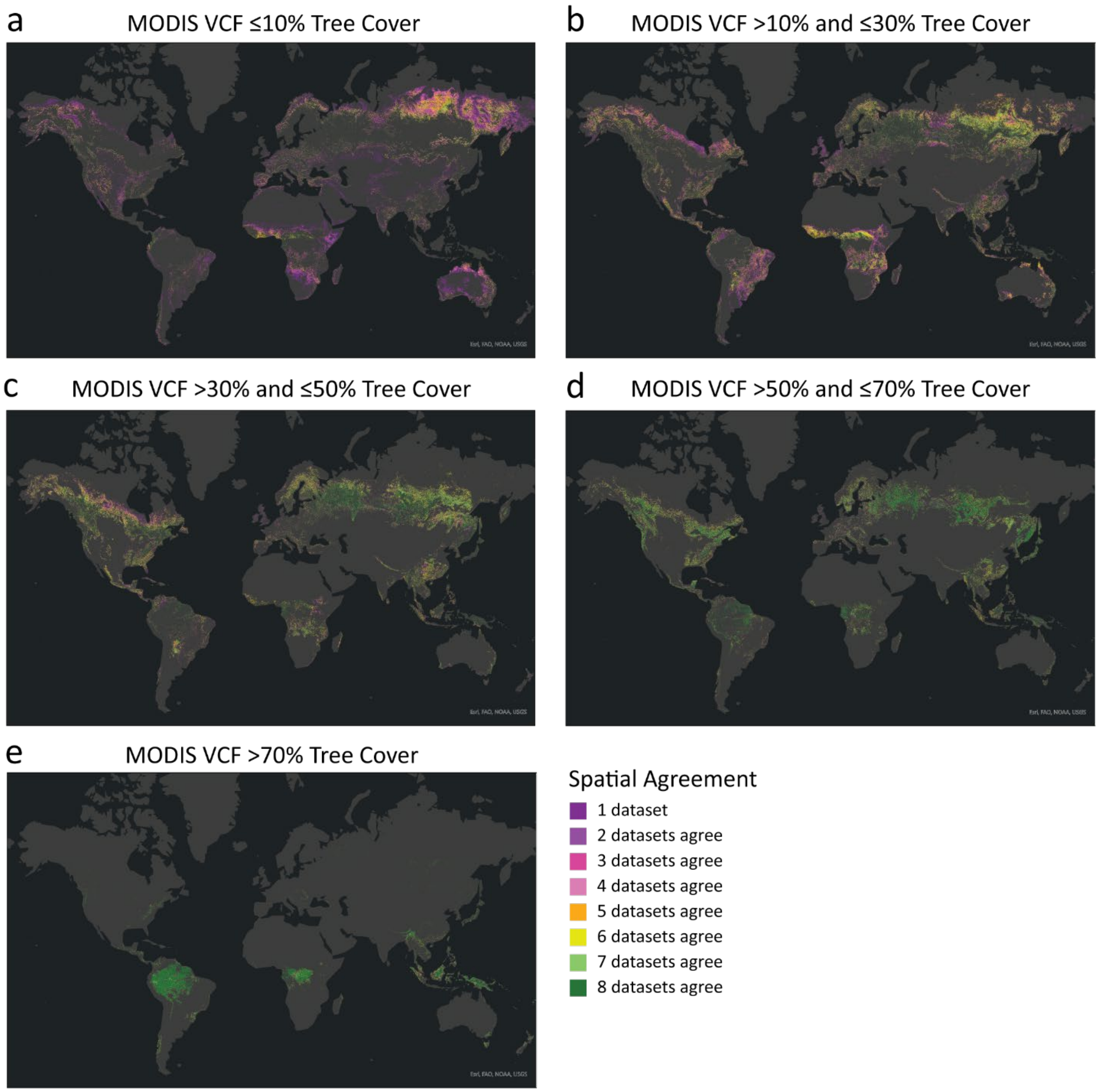

Fig S2. Spatial agreement for regions defined by different levels of tree canopy cover, as defined by the Terra MODIS Vegetation Continuous Fields (VCF) product [S47] using five thresholds: (a) ≤10% canopy cover, (b) >10% and ≤30%, (c) >30% and ≤50%, (d) >50% and ≤70%, and (e) >70%. Dark green (primarily shown for high levels of canopy cover in panel e) indicates full spatial congruence. Dark purple (primarily shown for very low levels of canopy cover in panel a) indicates no spatial congruence (only one dataset defines land as forested). Panel a includes pixels classified as 0% tree cover by MODIS VCF.

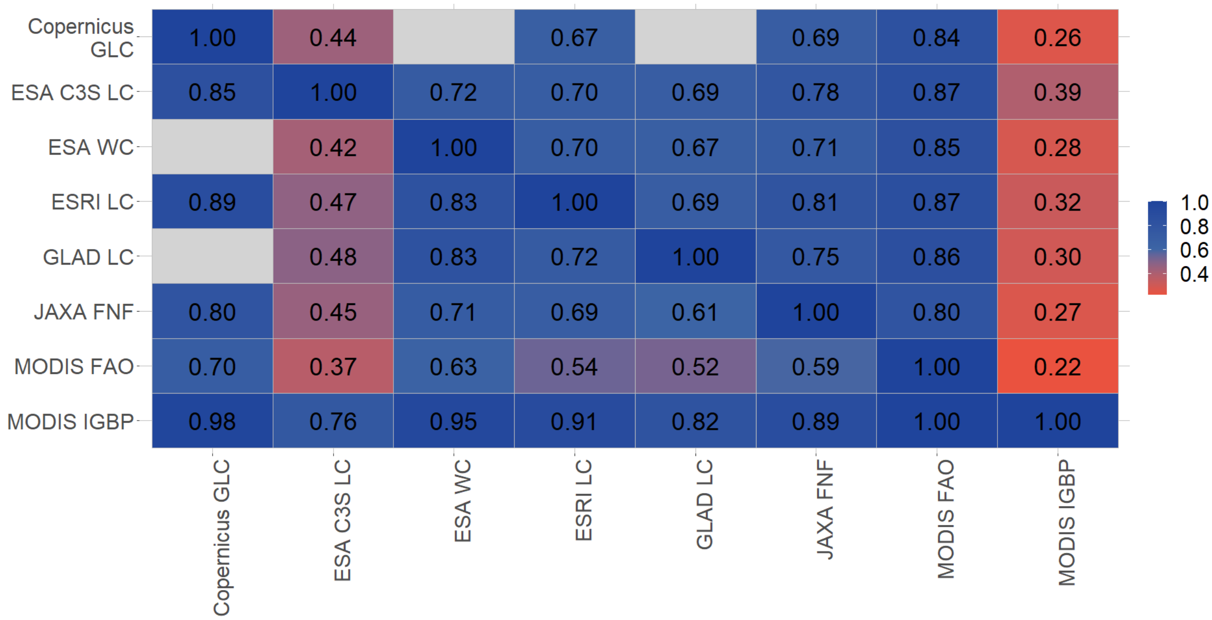

Fig. S3. Pairwise agreement between each dataset using alternative definition (showing which forest cover datasets relatively over/under-estimate forest cover compared to each other). Pairwise agreement is defined as the area of intersection where both datasets agree is forest (**F1 ∩ F2**) divided by the total forest area defined by the dataset in the first column (**F1**).

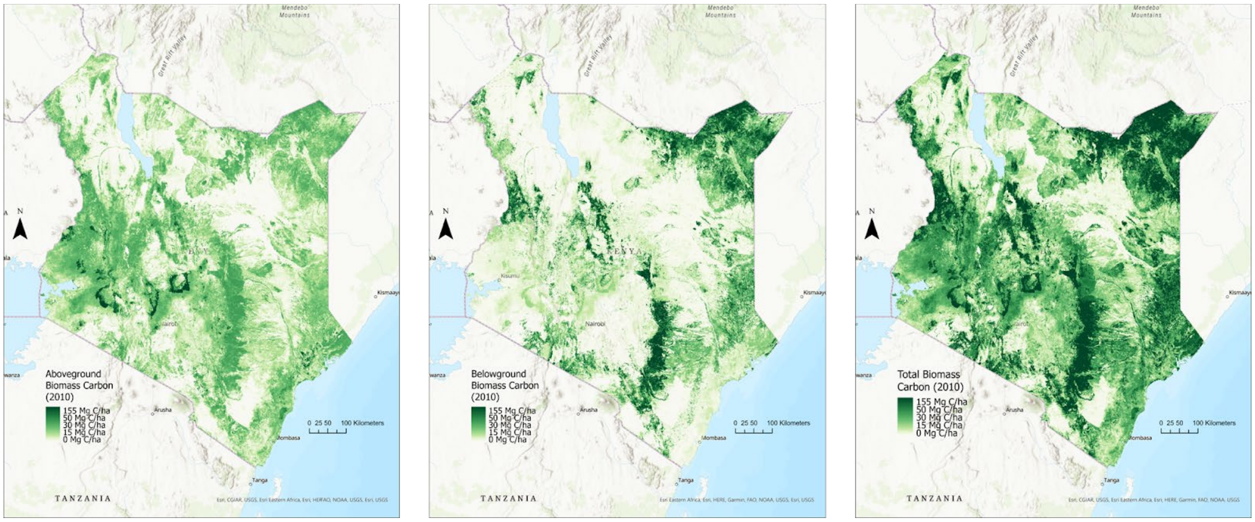

Fig. S4. Above-ground (left), below-ground (center), and total (right) biomass carbon in Kenya for the year 2010 reproduced from [S48].

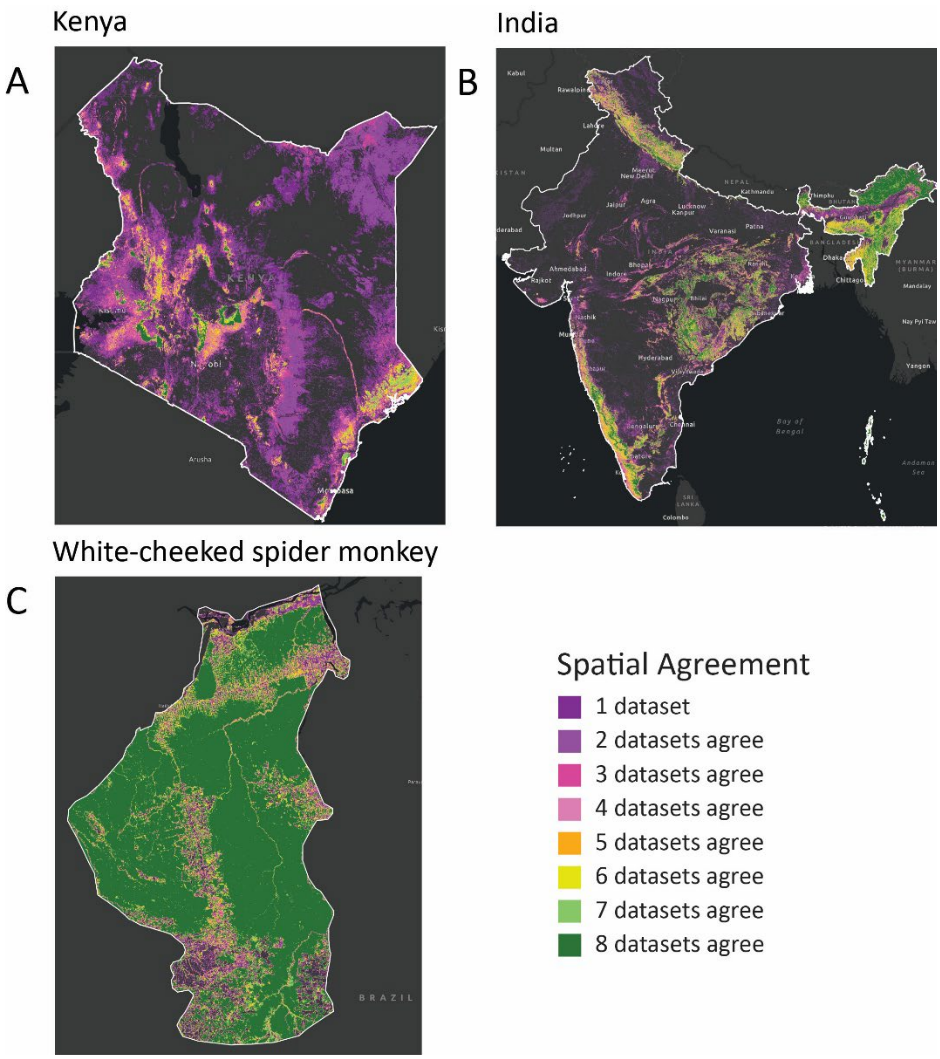

Fig. S5. Spatial agreement among the eight datasets for (A) Kenya, (B) India, and (C) the white-cheeked spider monkey's habitat range in Brazil. (A) Spatial agreement on forested lands in Kenya was low across the eight GFDs. Of the pixels that were classified as a forest by at least one dataset, only 2% were classified as a forest by all eight datasets, and only 11% of pixels had high agreement (≥5 datasets). The highest agreement occurred in the tropical wet regions (coastal, central, and west central Kenya), whereas disagreement was higher for the arid and temperate dry regions (northern and southern Kenya). For example, some of the datasets (e.g., ESRI-LC, JAXA-FNF) captured trees and shrubs in arid regions in their classifications, whereas these areas were excluded by other datasets. Forest-cropland mosaics in the temperate central part of Kenya also had high variability in classification across the datasets. (B) We found moderately high spatial agreement between the eight forest cover datasets for India. At least five datasets agreed on 50% of the total area defined as forests across all datasets, while all eight datasets agreed on 17% of pixel area classified as a forest. The highest agreement was in the northeastern region near Bhutan, while there was very low agreement in the drier regions of central India. (C) There was relatively high spatial agreement of forest area between the eight datasets for the white-cheeked spider monkey's region. For example, in 2020, the total forest cover estimate was between 401,943 km2 (GLAD-LCLUC) and 471,462 km2 (MODIS-FAO) (Table S8). However, GFD agreement within species' ranges was highly dependent on the biomes the range intersects.

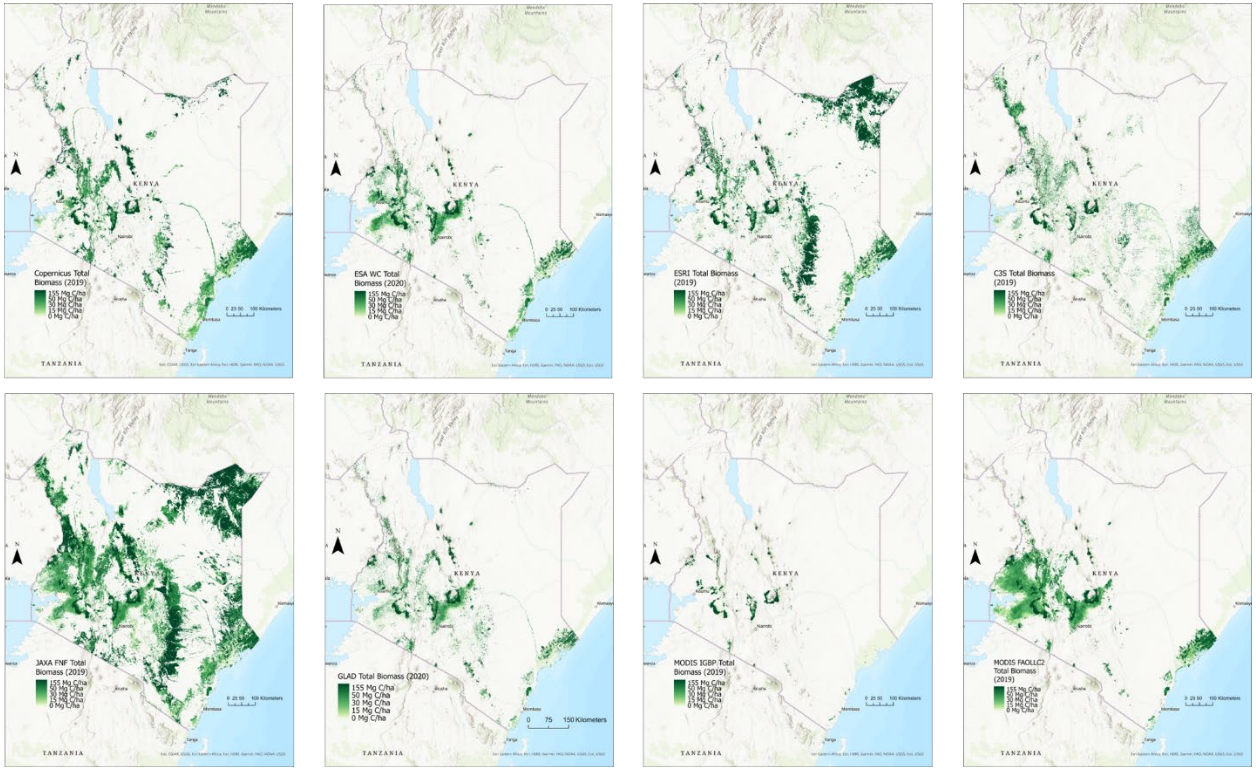

Fig. S6. Total biomass carbon estimates on forested lands using (clockwise from top right): CGLC, ESA-WC, ESRI-LC, ESA-C3S-LC, GLAD-LCLUC, MODIS-FAO, MODIS-IGBP, GLAD-LCLUC, and JAXA-FNF.

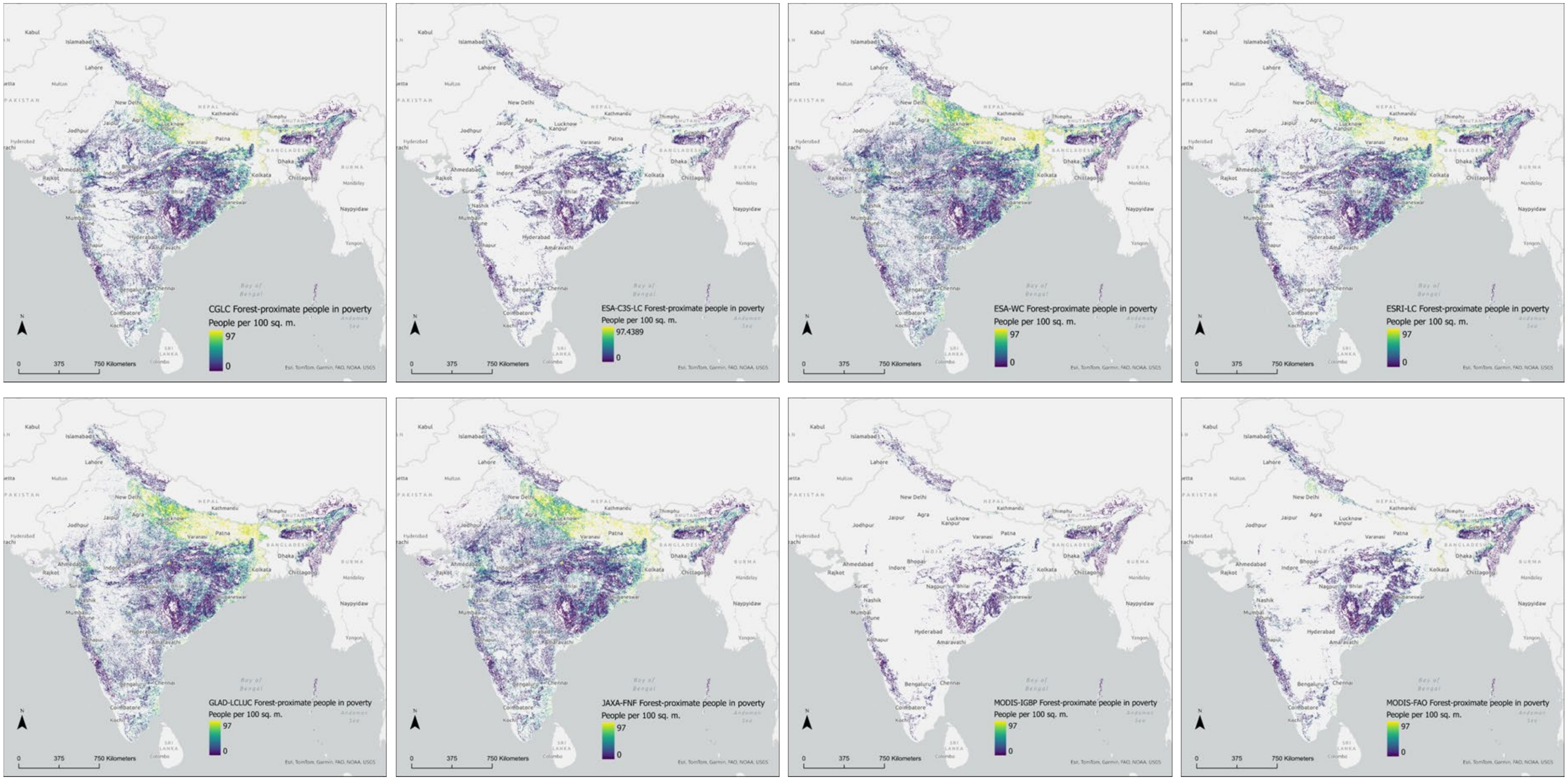

Fig. S7. Forest-proximate people in poverty in India using (clockwise from the top tight): CGLC, ESA-C3S-LC, ESA-WC, ESRI-LC, MODIS-FAO, MODIS-IGBP, JAXA-FNF, and GLAD-LCLUC. All figures are on the same scale, where yellow indicates high population density of forest-proximate populations living in poverty and purple indicates low population density of forest-proximate populations living in poverty.

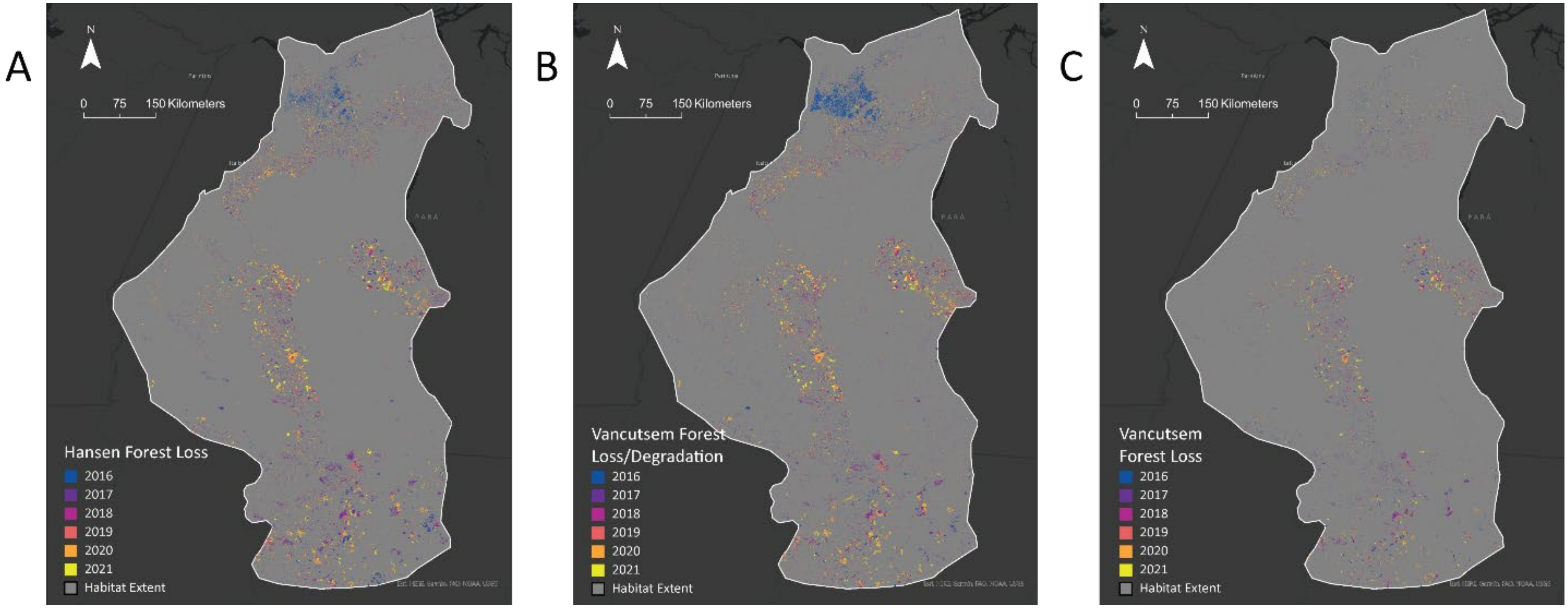

Fig. S8. (A) Hansen-GFC forest cover loss, (B) Vancutsem-TMFCC forest cover loss and degradation, and (C) Vancutsem forest cover loss for 2016-2021. Figures S6A and S6B are reproduced from Figure 6 in the main text, for comparison to Figure S6C.

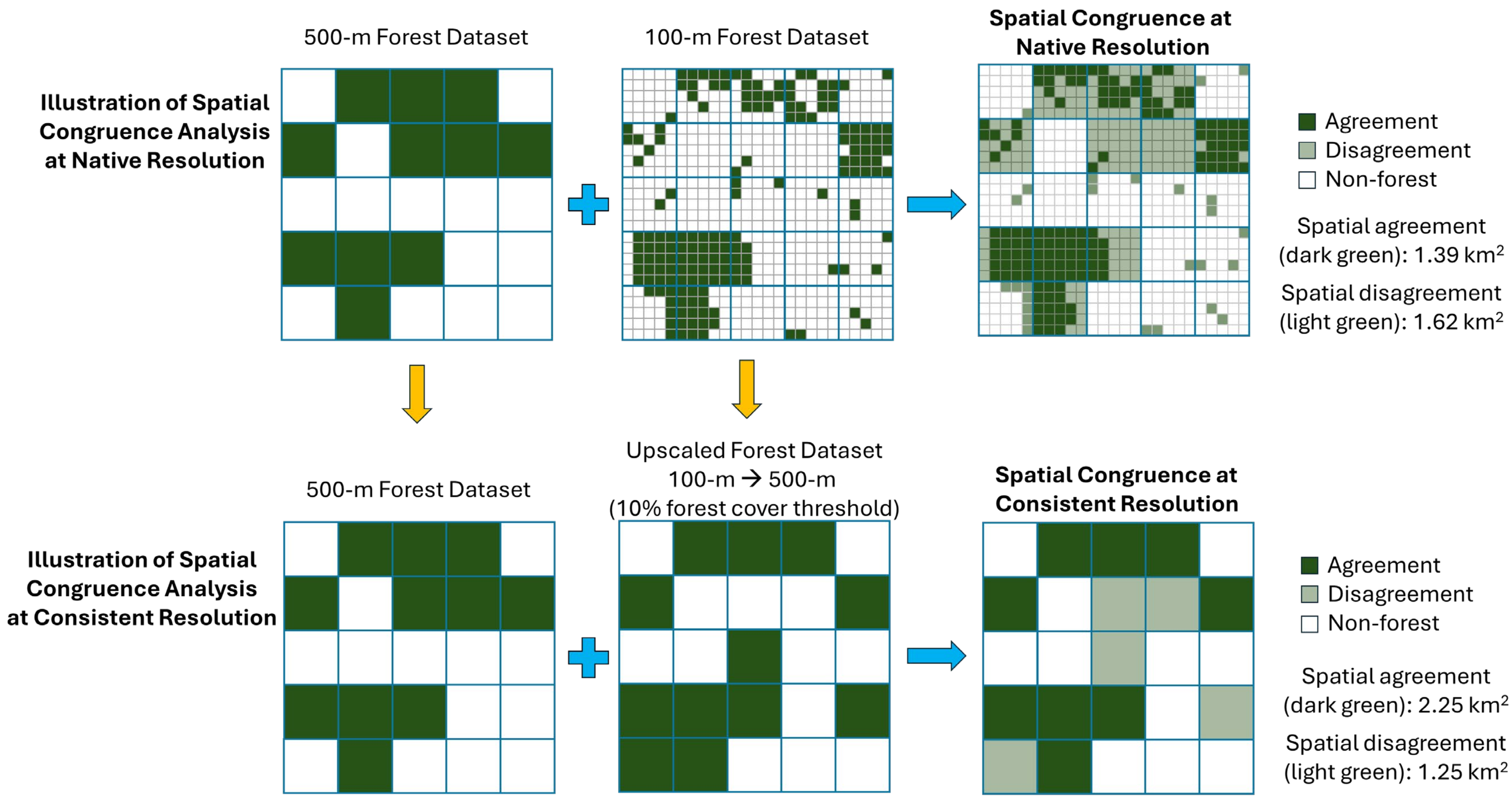

Fig. S9. Illustration of the two methods used to calculate spatial congruence between forest datasets. **Top row:** Example of comparing datasets at their original resolutions. A coarse-resolution product (left) and a fine-resolution product (center) are directly overlaid (right) to assess agreement at the native pixel level. While this approach reflects how users may apply available datasets in practice, the scale mismatch may influence results. **Bottom row:** Harmonized-resolution comparison. Both the fine-resolution and coarse-resolution datasets are resampled to a common 500-meter grid. For fine-resolution data, the proportion of forest pixels in each coarse cell is calculated and then thresholded (e.g., at 10%, as shown) to produce a binary forest map. Spatial congruence is then calculated from these harmonized-resolution binary maps. This method ensures valid cross-dataset comparison and avoids downscaling artifacts. Dark green: forest agreement; light green: disagreement; white: no forests.

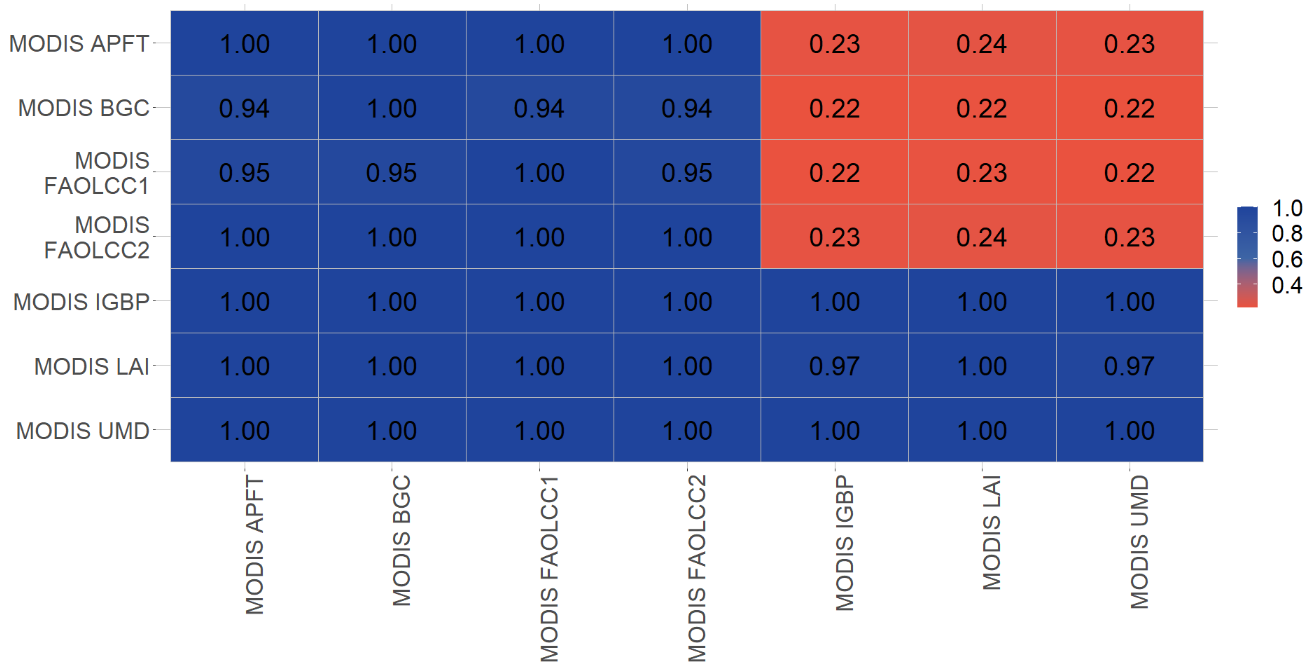

Fig. S10. Pairwise agreement for full set of MODIS datasets (showing which forest cover datasets relatively over/under-estimate forest cover compared to each other). Pairwise agreement is defined as the area of intersection where both datasets agree is forest ($\mathbf{F1} \cap \mathbf{F2}$) divided by the total forest area defined by the dataset in the first column ($\mathbf{F1}$). Blue indicates high overall spatial agreement between two datasets; red indicates low overall spatial agreement between two datasets.

Table S1. Summary of the characteristics of the reviewed global forest datasets (see inclusion criteria in methods section). All of the datasets are publicly available at no cost to users. See extended text for more detailed descriptions of each dataset, including the land cover classification methodology, reported accuracy, and known issues. We included descriptions of all seven MODIS land cover datasets, but our analysis focuses only on the IGBP and the FAOLCC1 classification systems since there was otherwise significant overlap between these dataset's classifications (see pairwise agreement tables). The excluded datasets are shown in light gray instead of black text. Another important source of variance between GFDs was in reported accuracy, which ranged from 67-91%, though these assessments vary in detail and were almost always conducted by the dataset producers rather than by independent observers. Detailed descriptions of each dataset, including the land cover classification methodology, reported accuracy, and any known challenges are presented in the extended text below.

Dataset name	Version	Geograp-hical coverage	Spatial resolution	Temporal resolution	Time span	Source imagery	Classification method	Reported accuracy	Forest/tree cover definition(s)
Copernicus Global Land Service (CGLS) [S1]	3	Global	100-meter	Every 1 year	2015-2019	PROBA-V	Random forest classifier and regression	77.6-83.7%	(1) Closed forests (tree canopy cover >70%): Evergreen needleleaf, evergreen broadleaf, deciduous needleleaf, deciduous broadleaf, mixed forest, other forest type (2) Open forests (tree canopy cover 15-70%): Evergreen needleleaf, evergreen broadleaf, deciduous needleleaf, deciduous broadleaf, mixed forest, other forest type
ESRI Land Cover (ESRI-LC) [S12]	1	Global	10-meter	Every 1 year	2017-2023	Sentinel-2	Deep learning (fully convolutional neural network with a UNet architecture)	85-90%	Trees - Any significant clustering of tall (~15-m or higher) dense vegetation, typically with a closed or dense canopy (examples: wooded vegetation, clusters of dense tall vegetation within savannas, plantations, swamp or mangroves (dense/tall vegetation with ephemeral water or canopy too thick to detect water underneath).
European Space AgencyWorld Cover (ESA-WC) [S4,5]	V100	Global	10-meter	Every 1 year	2020-2021	Sentinel-1-SAR and Sentinel-2 and others	Gradient boosting decision tree algorithm	67.5-80.7%	Areas dominated by trees with a cover ≥ 10%, possibly with other land cover classes present below the canopy. (Includes afforested areas, plantations, and areas seasonally or permanently flooded with fresh water, but excludes mangroves.)

European Space Agency Global Land Cover (ESA-C3S-LC) [S7,8]	v2.1.1	Global	300-meter	Every 1 year	1992-2022	MERIS, AVHRR, SPOT-VGT, PROVA-V, Sentinel-3	Gaussian Maximum Likelihood classifier and ISODATA clustering technique	70.5-71.1%	Tree cover >15% of types broadleaf evergreen, broadleaf deciduous, needleleaf evergreen, needleleaf deciduous, and mixed leaf type. Also includes classes for landscape mosaics that include some level of tree cover.
Global Land Cover and Land Use Change, 2000-2020 (GLAD-LCLUC) [S14]	v1	Global	30-meter	20 years	2000-2020	Landsat ARD, GEDI	Regression tree ensemble (bootstrap aggregation)	91.40%	Forest defined by pixels dominated by woody vegetation ≥ 5 m in height. The forest extent change (net loss and gain) was derived directly from the year 2000 and 2020 map comparison. Net forest gain classified only if a pixel had net forest height increase from the year 2000 to 2020 by ≥ 100%. Net forest loss estimated if the net height reduction was by ≥ 50% of the year 2000 value. Areas with small differences between the years 2000 and 2020 forest height data that have no indication of forest disturbance are considered stable forests.
JAXA Global PALSAR-2/PALSAR Forest/Non-Forest map (JAXA-FNF) [S15]	2.0.0	Global	25-meter, 100-meter, 1-km/0.25 deg	Annually 2007-2010, annually 2015-2020	2007-2020	PALSAR-2/PALSAR SAR	Random forest classifier	84.9-91.3%	"Forest" is defined as the tree covered land with the area larger than 0.5 ha and canopy cover over 10%. - Forests with a canopy cover of 90% or more. - Forests with a canopy cover of 10% to 90%.

MODIS Land Cover Datasets [S17]:									
MODIS Land Cover Type 1: Annual International Geosphere-Biosphere Programme	6	Global	500-meter	Every 1 year	2001-2020	MODIS Terra and Aqua reflectance	Decision tree and neural network classifier	67%	Forests defined by canopy height >2m and tree cover >60%. - Forest types: evergreen needleleaf, evergreen broadleaf, deciduous needleleaf, deciduous broadleaf, mixed forests (40-60% of each deciduous and evergreen tree type).

(MODIS-IGBP) classification									Savannas defined by canopy height >2m and tree cover 10-60%. (Savanna types: woody savannas (tree cover 30-60%), savannas (tree cover 10-30%)).
MODIS Land Cover Type 2: Annual University of Maryland (MODIS-UMD) classification								67%	Forests defined by canopy height >2m and tree cover >60%. - Forest types: evergreen needleleaf, evergreen broadleaf, deciduous needleleaf, deciduous broadleaf, mixed forests (40-60% of each deciduous and evergreen tree type). Savannas defined by canopy height >2m and tree cover 10-60%. (Savanna types: woody savannas (tree cover 30-60%), savannas (tree cover 10-30%)).
MODIS Land Cover Type 3: Annual Leaf Area Index (MODIS-LAI) classification								n/a	Forests defined by canopy height >2m and tree cover >60%. - Forest types: evergreen needleleaf, evergreen broadleaf, deciduous needleleaf, deciduous broadleaf. Savannas defined by canopy height >2m and tree cover 10-60%.
MODIS Land Cover Type 4: Annual BIOME-Biogeochemical Cycles (MODIS-BGC) classification								n/a	Trees and shrubs defined by canopy height >1m and woody vegetation cover >10%. - Vegetation types: evergreen needleleaf vegetation, evergreen broadleaf vegetation, deciduous needleleaf vegetation, deciduous broadleaf vegetation.
MODIS Land Cover Type 5: Annual Plant Functional Types classification								n/a	Trees defined by canopy height >2m and tree cover >10%. - Tree area types: evergreen needleleaf trees, evergreen broadleaf trees, deciduous needleleaf trees, deciduous broadleaf trees.

MODIS FAO-Land Cover Classification System 1 (MODIS-FAO) land cover layer								74-81%	Forests defined by canopy height >2m and tree cover >60%. - Forest types: evergreen needleleaf, evergreen broadleaf, deciduous needleleaf, deciduous broadleaf, mixed broadleaf/ needleleaf (40-60% of each broadleaf deciduous and evergreen needleleaf tree), mixed broadleaf evergreen/deciduous forests: (40-60% of each broadleaf evergreen and deciduous tree. Open forests defined by canopy height >2m and tree cover 30-60%. Sparse forests defined by canopy height >2m and tree cover 10-30%.
MODIS FAO-Land Cover Classification System 2 (MODIS-FAO-LCC2) land use layer								74-81%	Dense forests defined by canopy height >2m and tree cover >60%. Open forests defined by canopy height >2m and tree cover 10-60%. Forest/cropland mosaics defined as mosaics of small-scale cultivation 40-60% with >10% natural tree cover.

Forest Cover Change Datasets									
Forest Cover Change in the Humid Tropics (Vancutsem-TMFCC) [S20]	1	Tropical moist forests. Global Ecological Zones: “tropical rain forest,” “tropical moist forest,” “tropical mountain system,” and “tropical dry forest”.	30-meter (0.09-ha)	Every 1 year	1990 - 2022	Landsat 4,5, 7, and 8	Procedural sequential decision tree	Disturbance mapping accuracy: 91.4%	Forests are defined by Landsat image spectral signal and detection of disruptions based on spectral signals at the pixel scale (not by a tree cover percent). Undisturbed forest cover: closed tropical moist (evergreen or semi-evergreen) forest coverage without any disturbance (degradation or deforestation) observed over the Landsat historical record Forest degradation: Short term tree canopy cover loss (<2.5 years) Deforestation: Long term conversion of forested land into non-forested land (>2.5 years) Forest regrowth: tree cover regrowth (for >3 years) after deforestation
Hansen Global Forest Change (Hansen-GFC) [S21]	1.8	Global (except Antarctica and many Arctic islands)	30-meter	Forest cover: 2000 only. (Forest cover loss: every 1 year, forest gain: 2000-2012 only)	2000 - 2023	Landsat 4,5, 7, and 8	Bagged decision tree	Loss year accuracy: 75.2% (and was correct within one year before or after 96.7% of the time)	Trees were defined as all vegetation taller than 5m in height.

Table S2. Spatial congruence estimates from withholding one dataset at the time. Forest match is the area where all seven remaining datasets agree land is classified as forested; forest totals are the area where one or more datasets classify land as forested. Spatial congruence is the proportion of land with full agreement on forest classification of all the land classified as forested by at least one dataset.

Scenario	**Forest Match Area (sq. km.)** *Seven datasets agree classed as forest*	**Total Forest Area (sq. km.)** *One or more datasets classed as forest*	**Spatial Congruence (%)**
Withhold ESA-C3S-LC	18,660,165	67,014,856	27.8%
Withhold CGLC	17,371,492	66,801,184	26.0%
Withhold ESA-WC	17,444,952	67,477,371	25.9%
Withhold ESRI-LC	17,707,487	66,996,039	26.4%
Withhold MODIS-FAO	17,348,420	63,490,248	27.3%
Withhold GLAD-LCLUC	18,422,361	67,736,145	27.2%
Withhold MODIS-IGBP	23,656,642	67,960,829	34.8%
Withhold JAXA-FNF	17,649,671	64,410,370	27.4%

Table S3. Spatial congruence estimates for regions defined by different levels of tree canopy cover, as defined by the Terra MODIS Vegetation Continuous Fields (VCF) product [S47] using five thresholds: (a) ≤10% canopy cover, (b) >10% and ≤30%, (c) >30% and ≤50%, (d) >50% and ≤70%, and (e) >70%. Forest match is the area where all seven remaining datasets agree land is classified as forested; forest totals are the area where one or more datasets classify land as forested. Spatial congruence is the proportion of land with full agreement on forest classification of all the land classified as forested by at least one dataset.

MODIS VCF tree canopy cover	**Forest Match Area (sq. km.) Eight datasets agree classed as forest**	**Total Forest Area (sq. km.) One or more datasets classed as forest**	**Spatial Congruence (%)**
≤10%	102,712	12,559,462	0.8%
>10% and ≤30%	467,471	21,287,568	2.2%
>30% and ≤50%	2,294,027	14,140,727	16.2%
>50% and ≤70%	7,361,368	11,797,287	62.4%
>70%	6,889,250	7,923,369	86.9%

Table S4: Spatial congruence estimates generated when normalizing datasets by spatial resolution to 500-meters and selecting consistent cutoffs for defining forests (10% or 60%). One MODIS dataset was withheld for each threshold since MODIS is already at 500-meters, and MODIS-FAO used a 10% canopy cover threshold for defining forests; MODIS-IGBP used a 60% canopy cover threshold for defining forests. These scenarios provide evidence on the influence of spatial resolution and forest definition, though the original forest classification definitions are used when normalizing.

	Scenario *After normalizing to 500m*	**Forest Match Area (sq. km.)** *All datasets agree classed as forest*	**Total Forest Area (sq. km.)** *One or more datasets classed as forest*	**Spatial Congruence (%)**
1	Cutoff 10% canopy cover	20,715,447	76,818,461	27.0%
2	Cutoff 10% canopy cover and withhold MODIS-IGBP	34,425,076	76,818,449	44.8%
3	Using only sparse forest classes (CGLC, ESA-WC, ESA-C3S-LC, GLAD-LCLUC, JAXA-FNF, MODIS-FAO)	35,688,929	75,894,292	47.0%
4	Cutoff 60% canopy cover	17,525,550	64,242,197	27.3%
5	Cutoff 60% canopy cover and withhold MODIS-FAO	17,539,986	58,158,190	30.2%
6	Using only dense forest classes (CGLC, ESRI-LC, JAXA-FNF, MODIS-IGBP)	17,498,209	41,839,511	41.8%

Table S5. Cumulative percent of forest pixels in agreements for each biome. Percent agreement is defined as area where the number of datasets agree that pixels are forests divided by the area where at least one dataset classifies pixels as forest. MODIS-IGBP consistently had lower estimates of forest cover relative to the other datasets (by at least one standard deviation for 12 out of 14 biomes), while MODIS-FAO often had higher estimates of forest cover relative to the other datasets (by at least one standard deviation for 7 out of 14 biomes). The JAXA-FNF dataset had lower estimates of forest cover relative to the other datasets for drier forest biomes and grasslands. ESA-WC is the only dataset that did not deviate beyond one standard deviation of the mean estimate for all biomes.

Percent of forest pixels in agreement for each biome														
	1. Tropical and subtropical moist broadleaf forests	2. Tropical and subtropical dry broadleaf forests	3. Tropical and subtropical coniferous forests	4. Temperate broad-leaf and mixed forests	5. Temperate conifer forests	6. Boreal forests or taiga	7. Tropical and subtropical grasslands, savannas, and shrublands	8. Temperate grasslands, savannas, and shrublands	9. Flooded Grasslands and Savannas	10. Montane grasslands and shrub-lands	11. Tundra	12. Mediterranean forests, woodlands, and scrub	13. Deserts and xeric shrub-lands	14. Mangroves
8	50%	12%	14%	33%	32%	23%	7%	9%	8%	9%	2%	5%	1%	7%
≥7	62%	21%	14%	47%	54%	45%	19%	20%	12%	15%	11%	13%	2%	16%
≥6	72%	31%	51%	57%	65%	57%	30%	30%	19%	20%	18%	22%	4%	26%
≥5	79%	41%	65%	63%	73%	67%	40%	31%	25%	24%	26%	32%	7%	43%
≥4	84%	50%	75%	69%	79%	75%	50%	47%	32%	29%	35%	43%	10%	58%
≥3	88%	61%	84%	75%	84%	83%	62%	56%	42%	36%	45%	55%	17%	70%
≥2	93%	76%	92%	83%	90%	90%	78%	69%	61%	46%	59%	72%	33%	83%
≥1	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Table S6: Forest extent for each dataset-biome combination (in square kilometers). Values indicate the area of forested pixels estimated by each dataset within each biome extent.

	Forest Extent (sq. km.) for each dataset-biome combination													
	1. Tropical and subtropical moist broadleaf forests	2. Tropical and subtropical dry broadleaf forests	3. Tropical and subtropical coniferous forests	4. Temperate broadleaf and mixed forests	5. Temperate conifer forests	6. Boreal forests or taiga	7. Tropical and subtropical grasslands, savannas, and shrublands	8. Temperate grasslands, savannas, and shrublands	9. Flooded Grasslands and Savannas	10. Montane grasslands and shrublands	11. Tundra	12. Mediterranean forests, woodlands, and scrub	13. Deserts and xeric shrublands	14. Mangroves
CGLC	15,300,750	271,087	929,152	913,015	522,677	150,927	1,394,725	480,896	6,303,311	2,541,116	10,786,790	8,024,518	1,387,089	220,027
ESA-C3S-LC	11,560,628	270,587	1,060,288	413,280	261,288	65,912	1,004,130	428,211	4,689,734	2,293,592	10,292,276	5,570,779	770,758	151,665
ESA-WC	14,592,927	212,132	797,269	703,380	269,284	90,810	1,157,824	387,528	5,943,169	2,328,066	10,087,711	4,919,081	1,263,645	183,188
ESRI-LC	14,035,568	221,044	930,931	608,424	390,345	146,608	874,208	251,150	5,242,753	2,082,317	8,996,465	6,642,308	1,101,084	251,844
GLAD-LCLUC	12,807,857	248,201	690,067	446,021	165,629	108,048	990,997	407,578	5,028,369	2,029,416	7,709,706	5,409,335	824,396	142,864
JAXA-FNF	14,402,879	598,040	1,461,536	943,240	1,306,007	128,006	1,669,239	513,008	5,828,481	2,486,293	9,422,244	8,626,034	1,239,263	263,859
MODIS-FAO	16,685,224	251,813	760,204	828,740	203,292	203,230	1,577,521	525,590	7,233,557	2,417,699	11,765,003	7,758,065	1,283,696	472,602
MODIS-IGBP	11,039,448	96,842	87,458	132,581	21,386	47,955	450,744	131,071	3,735,140	1,247,192	3,781,501	1,246,506	315,989	68,949

Table S7. Percent of total biomass carbon for the total land area in Kenya that is defined as forested lands for each of the eight land cover datasets. Total biomass carbon is generated from the Spawn *et al.* [S48] dataset for the year 2010 and is held constant. The forested lands are identified for the years 2019 and 2020 (where available) for the CGLC, GLAD-LULCC, ESA-WC, ESA-C3S-LC, ESRI-LC, JAXA-FNF, MODIS-FAO, and MODIS-IGBP datasets. Four datasets had particularly low estimates of Kenya's forest carbon biomass: MODIS-IGBP (2%), ESA-C3S (5%), ESA-WC (6%), GLAD-LCLUC (6%). On the other hand, MODIS-FAO, CGLC, ESRI-LC, and JAXA-FNF resulted in more moderate to high estimates (8-9%, 11%, 16-22%, and 7-37%, respectively).

Forest cover Dataset	**% of total 2010 biomass in Kenya**	
	2019	**2020**
CGLC	11%	
GLAD-LCLUC		6%
ESA-WC		6%
ESA-C3S-LC	5%	5%
ESRI-LC 2017-2021	16%	22%
ESRI- 2020		26%
JAXA-FNF (dense + sparse classes)	34%	37%
JAXA-FNF (dense class only)	7%	8%
MODIS-FAO	8%	9%
MODIS-IGBP	2%	2%

Table S8. Estimated forest-proximate people in poverty in India in 2016. Estimated population data [S49], population in poverty [S50], land area and forest area, and forest-proximate people also are recorded. The map of urban and rural areas were generated using the methodology presented in Newton et al. [S51]. ESA-WC uses a similar tree cover definition to CGLC and generates very similar forest area estimates, but much higher estimates of forest-proximate people living in poverty (by almost 88 million people). ESRI-LC defines forest cover more strictly than CGLC and ES-WC, including only densely forested areas, but this definition resulted in an estimate of only 16 million fewer forest-proximate people living in poverty in India than the CGLC estimate. In comparison, using the much lower resolution MODIS or ESA-C3S-LC data yielded significantly lower estimates of forest-proximate people living in poverty.

India Forest-Proximate People in Poverty (2018)	CGLC	GLAD-LCLUC	ESA-WC	ESA-C3S-LC	ESRI-LC	JAXA-FNF	MODIS-FAO	MODIS-IGBP
Total Population	1,309,789,184							
Total Urban Population	345,267,515							
Total Rural Population	964,521,669							
Total Poor Population	324,748,418							
Total Urban Poor Pop.	16,092,267							
Total Rural Poor Pop.	308,656,151							
Forest Proximate People	578,057,515	763,864,850	822,868,940	149,430,115	495,166,902	741,489,667	240,737,712	52,791,664
Forest Proximate People in Poverty	182,025,897	233,176,650	252,164,042	56,100,789	162,094,558	233,997,686	61,928,129	23,062,210
Total Land Area (sq. km.)	3,093,580							
Urban Land Area (sq. km.)	67,167							
Rural Land Area (sq. km.)	3,026,413							
Forest Area (sq. km.)	778,725	609,858	752,452	461,949	619,563	684,695	656,405	302,273
Forest Cover Area on Rural Lands (sq. km.)	767,262	596,383	734,154	456,333	611,811	676,815	636,416	298,378

Table S9. White-cheeked spider monkey forested habitat area (sq. km.) in 2020 (2019 for CGLC).

Forest cover Dataset	White-cheeked spider monkey forested habitat area (sq. km.)
CGLC	456,909
GLAD-LCLUC	401,943
ESA-WC	442,089
ESA-C3S-LC	433,162
ESRI-LC	434,368
JAXA-FNF	456,357
MODIS-FAO	471,462
MODIS-IGBP	423,173

Table S10. Agreement between the Hansen-GFC dataset and the Vancutsem-TMFCC dataset for forest loss in the white-cheeked spider monkey habitat area. Values are in sq. km. for each year.

	Year						
	2015-2016 (2016)	**2016-2017 (2017)**	**2017-2018 (2018)**	**2018-2019 (2019)**	**2019-2020 (2020)**	**2020-2021 (2021)**	**Total**
Hansen-GFC and Vancutsem-TMFCC agree (sq. km.)	3,044	2,347	2,121	2,256	3,332	2,544	15,644
Hansen-GFC forest, Vancutsem-TMFCC not forest (sq. km.)	2,065	3,269	1,636	1,278	1,764	1,466	11,478
Vancutsem-TMFCC forest, Hansen-GFC not forest (sq. km.)	4,861	2,415	2,001	1,589	2,857	1,233	14,956
% agreement area of total forest change area	31%	29%	37%	44%	42%	49%	37%

Supplemental Information References

S1. Buchhorn, M., Lesiv, M., Tsendbazar, N.-E., Herold, M., Bertels, L., and Smets, B. (2020). Copernicus global land cover layers—collection 2. Remote Sensing *12*, 1044. https://doi.org/10.3390/rs12061044.

S2. Buchhorn, M., Bertels, L., Smets, B., De Roo, B., Lesiv, M., Tsendbazar, N.-E., Masiliunas, D., and Li, L. (2021). Copernicus Global Land Service: Land Cover 100m: Version 3 Globe 2015-2019: Algorithm Theoretical Basis Document (Dataset v3.0, doc issue 3.4).

S3. Buchhorn, M., Smets, B., Bertels, L., de Roo, B., Lesiv, M., Tsendbazar, N.-E., Li, L., and Tarko, A. (2021). Copernicus Global Land Service: Land Cover 100m: version 3 Globe 2015-2019: Product User Manual.

S4. Zanaga, D., Van De Kerchove, R., Daems, D., De Keersmaecker, W., Brockmann, C., Kirches, G., Wevers, J., Cartus, O., Santoro, M., Fritz, S., et al. (2022). ESA WorldCover 10 m 2021 v200. https://doi.org/10.5281/zenodo.7254221.

S5. Zanaga, D., Van De Kerchove, R., De Keersmaecker, W., Souverijns, N., Brockmann, C., Quast, R., Wevers, J., Grosu, A., Paccini, A., Vergnaud, S., et al. (2021). ESA WorldCover 10 m 2020 v100.

S6. Tsendbazar, N., Li, L., Koopman, M., Carter, S., Herold, M., Georgieva, I., and Lesiv, M. (2021). WorldCover Product Validation report (D12-PVR), version 1.1.

S7. ESA (2017). Land Cover CCI Product User Guide Version 2.

S8. Defourny, P., Lamarche, C., Flasse, C., Kirches, G., Böttcher, M., and Brockmann, C. (2019). C3S - Product User Guide and Specification ICDR Land Cover 2016 to 2018, v1.2.

S9. Defourny, P., Lamarche, C., Flasse, C., Brockmann, C., Böttcher, M., and Kirches, G. (2019). Algorithm Theoretical Basis Document ICDR Land Cover 2016.

S10. Defourny, P., Lamarche, C., Marissiaux, Q., Brockmann, C., Böttcher, M., and Kirches, G. (2021). Algorithm Theoretical Basis Document ICDR Land Cover 2020.

S11. Defourny, P., Lamarche, C., Flasse, C., Kirches, G., Böttcher, M., and Brockmann, C. (2020). Product Quality Assurance Document ICDR Land Cover 2016-2020.

S12. Karra, K., Kontgis, C., Statman-Weil, Z., Mazzariello, J.C., Mathis, M., and Brumby, S.P. (2021). Global land use/land cover with Sentinel 2 and deep learning. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 4704-4707.

S13. Impact Observatory (2022). Methodology & Accuracy Summary 10m Global Land Use Land Cover Maps.

S14. Potapov, P., Hansen, M.C., Pickens, A., Hernandez-Serna, A., Tyukavina, A., Turubanova, S., Zalles, V., Li, X., Khan, A., Stolle, F., et al. (2022). The Global 2000-2020 Land Cover and Land Use Change Dataset Derived From the Landsat Archive: First Results. Frontiers in Remote Sensing *3*. 10.3389/frsen.2022.856903.

S15. Shimada, M., Itoh, T., Motooka, T., Watanabe, M., Shiraishi, T., Thapa, R., and Lucas, R. (2014). New global forest/non-forest maps from ALOS PALSAR data (2007–2010). Remote Sensing of Environment *155*, 13-31. https://doi.org/10.1016/j.rse.2014.04.014.

S16. JAXA (2022). Global 25 m Resolution PALSAR-2 Forest/Non-Forest Map (FNF) (Ver.2.0.0) Dataset Description.

S17. Friedl, M., and Sulla-Menashe, D. (2015). MCD12Q1 MODIS/Terra+ aqua land cover type yearly L3 global 500m SIN grid V006 [data set]. NASA EOSDIS Land Processes DAAC *10*, 200.

S18. Sulla-Menashe, D., and Friedl, M.A. (2018). User guide to collection 6 MODIS land cover (MCD12Q1 and MCD12C1) product. USGS: Reston, VA, USA *1*, 18.
S19. Sulla-Menashe, D., Gray, J.M., Abercrombie, S.P., and Friedl, M.A. (2019). Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product. Remote Sensing of Environment *222*, 183-194. https://doi.org/10.1016/j.rse.2018.12.013.
S20. Vancutsem, C., Achard, F., Pekel, J.-F., Vieilledent, G., Carboni, S., Simonetti, D., Gallego, J., Aragão, L.E.O.C., and Nasi, R. (2021). Long-term (1990-2019) monitoring of forest cover changes in the humid tropics. Science Advances *7*, eabe1603. 10.1126/sciadv.abe1603.
S21. Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A., Tyukavina, A., Thau, D., Stehman, S., Goetz, S.J., and Loveland, T.R. (2013). High-resolution global maps of 21st-century forest cover change. science *342*, 850-853.
S22. Sexton, J.O., Song, X.-P., Feng, M., Noojipady, P., Anand, A., Huang, C., Kim, D.-H., Collins, K.M., Channan, S., DiMiceli, C., and Townshend, J.R. (2013). Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. International Journal of Digital Earth *6*, 427-448. 10.1080/17538947.2013.786146.
S23. Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., et al. (2015). Global land cover mapping at 30m resolution: A POK-based operational approach. ISPRS Journal of Photogrammetry and Remote Sensing *103*, 7-27. https://doi.org/10.1016/j.isprsjprs.2014.09.002.
S24. Banzhaf, E., Wu, W., Luo, X., and Knopp, J. (2021). Integrated Mapping of Spatial Urban Dynamics—A European-Chinese Exploration. Part 1—Methodology for Automatic Land Cover Classification Tailored towards Spatial Allocation of Ecosystem Services Features. Remote Sensing *13*, 1744.
S25. Dimiceli, C., Carroll, M., Sohlberg, R., Kim, D., Kelly, M., and Townshend, J. (2015). MOD44B MODIS/terra vegetation continuous fields yearly L3 global 250m SIN grid V006. NASA EOSDIS Land Processes Distributed Active Archive Center *10*.
S26. Townshend, J., Hansen, M., Carroll, M., DiMiceli, C., Sohlberg, R., and Huang, C. (2020). User Guide for the MODIS Vegetation Continuous Fields product Collection 6 Version 1.
S27. Adzhar, R., Kelley, D.I., Dong, N., Torello Raventos, M., Veenendaal, E., Feldpausch, T.R., Philips, O.L., Lewis, S., Sonké, B., and Taedoumg, H. (2021). Assessing MODIS Vegetation Continuous Fields tree cover product (collection 6): performance and applicability in tropical forests and savannas. Biogeosciences Discussions, 1-20.
S28. Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C.E., et al. (2021). Mapping global forest canopy height through integration of GEDI and Landsat data. Remote Sensing of Environment *253*, 112165. https://doi.org/10.1016/j.rse.2020.112165.
S29. Harris, N., Goldman, E.D., and Gibbes, S. (2019). Spatial database of planted trees (SDPT) version 1.0. Technical Note, World Resources Institute.
S30. Latham, J., Cumani, R., Rosati, I., and Bloise, M. (2014). Global land cover share (GLC-SHARE) database beta-release version 1.0-2014. FAO: Rome, Italy.
S31. Gong, P., Wang, J., Yu, L., Zhao, Y., Zhao, Y., Liang, L., Niu, Z., Huang, X., Fu, H., Liu, S., et al. (2013). Finer resolution observation and monitoring of global land cover:

first mapping results with Landsat TM and ETM+ data. International Journal of Remote Sensing *34*, 2607-2654. 10.1080/01431161.2012.748992.
S32. Gong, P., Liu, H., Zhang, M., Li, C., Wang, J., Huang, H., Clinton, N., Ji, L., Li, W., Bai, Y., et al. (2019). Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017. Science Bulletin *64*, 370-373. https://doi.org/10.1016/j.scib.2019.03.002.
S33. Liu, H., Gong, P., Wang, J., Clinton, N., Bai, Y., and Liang, S. (2020). Annual dynamics of global land cover and its long-term changes from 1982 to 2015. Earth Syst. Sci. Data *12*, 1217-1243. 10.5194/essd-12-1217-2020.
S34. Arino, O., Ramos Perez, J.J., Kalogirou, V., Bontemps, S., Defourny, P., and Van Bogaert, E. (2012). Global land cover map for 2009 (GlobCover 2009).
S35. Bontemps, S., Defourny, P., Van Bogaert, E., Arino, O., Kalogirou, V., and Perez, J.R. (2011). GlobCover 2009: Products description and validation report.
S36. Gutman, G., Byrnes, R.A., Masek, J., Covington, S., Justice, C., Franks, S., and Headley, R. (2008). Towards monitoring land-cover and land-use changes at a global scale: The Global Land Survey 2005. Photogrammetric Engineering and Remote Sensing *74*, 6-10.
S37. Gutman, G., Huang, C., Chander, G., Noojipady, P., and Masek, J.G. (2013). Assessment of the NASA–USGS Global Land Survey (GLS) datasets. Remote Sensing of Environment *134*, 249-265. https://doi.org/10.1016/j.rse.2013.02.026.
S38. Grekousis, G., Mountrakis, G., and Kavouras, M. (2015). An overview of 21 global and 43 regional land-cover mapping products. International Journal of Remote Sensing *36*, 5309-5335. 10.1080/01431161.2015.1093195.
S39. Loveland, T.R., Reed, B.C., Brown, J.F., Ohlen, D.O., Zhu, Z., Yang, L., and Merchant, J.W. (2000). Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. International Journal of Remote Sensing *21*, 1303-1330. 10.1080/014311600210191.
S40. Loveland, T.R., and Belward, A.S. (1997). The International Geosphere Biosphere Programme Data and Information System global land cover data set (DISCover). Acta Astronautica *41*, 681-689. https://doi.org/10.1016/S0094-5765(98)00050-2.
S41. Loveland, T.R., Brown, J., Ohlen, D., Reed, B., Zhu, Z., Yang, L., and Howard, S. (2009). ISLSCP II IGBP DISCover and SiB Land Cover, 1992-1993. 10.3334/ORNLDAAC/930.
S42. Bartholomé, E., and Belward, A.S. (2005). GLC2000: a new approach to global land cover mapping from Earth observation data. International Journal of Remote Sensing *26*, 1959-1977. 10.1080/01431160412331291297.
S43. Hansen, M.C., Defries, R.S., Townshend, J.R.G., and Sohlberg, R. (2000). Global land cover classification at 1 km spatial resolution using a classification tree approach. International Journal of Remote Sensing *21*, 1331-1364. 10.1080/014311600210209.
S44. Tateishi, R., Uriyangqai, B., Al-Bilbisi, H., Ghar, M.A., Tsend-Ayush, J., Kobayashi, T., Kasimu, A., Hoan, N.T., Shalaby, A., Alsaaideh, B., et al. (2011). Production of global land cover data – GLCNMO. International Journal of Digital Earth *4*, 22-49. 10.1080/17538941003777521.
S45. Tateishi, R., Hoan, N.T., Kobayashi, T., Alsaaideh, B., Tana, G., and Phong, D.X. (2014). Production of global land cover data-GLCNMO2008. Journal of Geography and Geology *6*, 99.

S46. Kobayashi, T., Tateishi, R., Alsaaideh, B., Sharma, R.C., Wakaizumi, T., Miyamoto, D., Bai, D.L., Gegentana, G., Maitiniyazi, A., Cahyana, D., et al. (2017). Production of global land cover data–GLCNMO2013. Journal of Geography and Geology *9*.
S47. DiMiceli, C., Sohlberg, R., and Townshend, J. (2022). MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250m SIN Grid V061 [Data set]. https://doi.org/10.5067/MODIS/MOD44B.061.
S48. Spawn, S.A., Sullivan, C.C., Lark, T.J., and Gibbs, H.K. (2020). Harmonized global maps of above and belowground biomass carbon density in the year 2010. Scientific Data *7*, 112. 10.1038/s41597-020-0444-4.
S49. WorldPop (2020). www.worldpop.org.
S50. Chi, G., Fang, H., Chatterjee, S., and Blumenstock, J.E. (2022). Microestimates of wealth for all low- and middle-income countries. Proceedings of the National Academy of Sciences *119*, e2113658119. doi:10.1073/pnas.2113658119.
S51. Newton, P., Castle, S.E., Kinzer, A.T., Miller, D.C., Oldekop, J.A., Linhares-Juvenal, T., Madrid, M., Pina, L., and de Lamo Rodriguez, J. (2022). The number of forest- and tree-proximate people: a new methodology and global estimates. FAO.