nature

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41586-018-0411-9

In the format provided by the authors and unedited.

Global land change from 1982 to 2016

Xiao-Peng Song[1]*, Matthew C. Hansen[1], Stephen V. Stehman[2], Peter V. Potapov[1], Alexandra Tyukavina[1], Eric F. Vermote[3] & John R. Townshend[1]

[1]Department of Geographical Sciences, University of Maryland, College Park, MD, USA. [2]College of Environmental Science and Forestry, State University of New York, Syracuse, NY, USA. [3]NASA Goddard Space Flight Center, Greenbelt, MD, USA. *e-mail: xpsong@umd.edu

Supplementary Methods

Definitions

Vegetation continuous fields (VCF) represent land surface as a fractional combination of vegetation functional types that can be remotely sensed from satellites[9]. Consistent with previous research[32-36], the VCF product developed in this study consists of percentages of tree canopy (TC) cover, short vegetation (SV) cover and bare ground (BG) cover. Trees are defined as all vegetation taller than 5 meters in height. TC refers to the proportion of the ground covered by the vertical projection of tree crowns[37,38]. SV characterizes the proportion of the ground covered by vegetation other than trees, including shrubs, herbaceous vegetation, and mosses, while BG represents the proportion of the land surface not covered by vegetation. TC, SV and BG are quantified from nadir view at top of canopy and are mapped during the local annual peak of a growing season[26,36,39]. TC is not equivalent to forest cover, although forest cover may be defined based on TC. For example, the FAO defines forest as a parcel or unit of land of at least 0.5 hectares in size which is covered by 10% or more trees that are 5 meters or taller[5]. Gain or loss in TC, SV, or BG refers to net increase or decrease in each respective cover over the study period due to any anthropogenic or natural factors, excluding temporary changes attributable to within-year vegetation phenology or year-to-year rotations.

Generation of AVHRR VCF

The Advanced Very High Resolution Radiometer (AVHRR) instruments on-board NOAA satellites remain an important data source for studying long-term changes in land surface properties as they provide the longest time-series of global satellite measurements[40-42]. We used the version 4 Long Term Data Record (LTDR) to generate the annual VCF products[42,43]. The LTDR was compiled from AVHRR observations through a series of processing steps including radiometric calibration, geolocation correction, atmospheric correction and bi-directional reflectance effect correction[42]. The daily LTDR surface reflectance data contain 5 multi-spectral layers of AVHRR channels 1-5 and the normalized difference

vegetation index (NDVI) layer computed from channels 1 and 2[44]. Each pixel is 0.05° × 0.05° in size. We implemented an improved version of the operational Moderate Resolution Imaging Spectroradiometer Vegetation Continuous Field (MODIS VCF) approach to convert daily LTDR to yearly VCF[33].

Daily AVHRR was first aggregated into monthly composites based on the maximum NDVI value in the month. Maximum NDVI composition can minimize cloud contamination, reduce bi-directional and off-nadir viewing effects, minimize band-correlated atmospheric effects and enhance vegetation discrimination[45]. The technique has been widely adopted for generating NDVI and land-cover products from daily satellite data for sensors such as AVHRR, MODIS and VEGETATION[41,46-49].

Monthly composites were subsequently converted to annual phenological metrics[7,33,50-52] (Supplementary Table 2). Metrics are statistical transformations of pixel time-series that can capture the salient features of vegetation phenology while maintaining high spatial and temporal data consistency. Metrics thus provide a unique advantage to large-area land cover mapping and monitoring. We created a total of 735 annual metrics from a combination of 5 multi-spectral bands and one NDVI layer, each available as time-series of 12 months.

An empirical normalization procedure was applied to enhance the year-to-year consistency of the AVHRR metrics (Supplementary Fig. 1). Time-series data from AVHRR are known to have systematic discrepancies due to different satellite platforms, orbital drift, changes in sensor design and sensor degradation[40,41,53,54]. The systematic differences are particularly pronounced before and after year 2000; beginning with NOAA-16 in 2000, satellite orbits were stabilized and a major improvement was introduced in the sensor design to increase sensitivity at the low end of radiance[40]. Research has also shown that the varying observational solar zenith angle as a result of orbital drift affects reflectance more than NDVI and is negatively related to leaf area or positively related to soil exposure[55]. That is, dense vegetation is less affected than sparse vegetation. Additionally, remaining atmospheric effects in the

AVHRR surface reflectance can also cause inconsistency between years. The normalization was designed to remove these artifacts unrelated to actual surface change.

A rich literature exists on calibration of AVHRR time series. One commonly used method is to apply calibration coefficients estimated from "stable targets" such as deserts, oceans, clouds or rainforests[56-61]. For example, earlier works by Myneni *et al.*[27,59] used the Sahara desert as reference to adjust global NDVI. Gutman (1999)[60] used global deserts and rainforests to correct reflectances as well as NDVI. Recently, data from well-calibrated sensors such as MODIS and SPOT were used as reference for anchoring AVHRR-based NDVI time series[40,41].

To normalize annual metrics, we designed a two-step approach, using MODIS data as reference. The first step was to apply a dark object subtraction (DOS) to remove systematic biases for vegetated surfaces, especially forest. DOS is also a simple and effective method of removing atmospheric contamination in remotely sensed data[62-66]. We used the intact forest landscapes (IFL)[67] of the tropical rainforest biome (i.e. the minimally disturbed tropical rainforests, average tree cover 97%; Supplementary Fig. 1c) as the dark stable target, which was also considered a spectral end-member. The second step was to apply a slope-based adjustment for pixels that contained visible bare ground. This step involved the use of tropical, subtropical and temperate deserts with 100% Landsat-based bare ground cover[26] (Supplementary Fig. 1c) as the bright stable target, or the other spectral end-member. Biases over other land surfaces are assumed to be within these two extreme end members[60]. To create the MODIS reference data, an identical procedure was applied to daily MODIS LTDR[43] to derive annual metrics for years 2000 through 2016. The 17-year median values for each metric were subsequently derived and used as reference.

DOS was conducted by applying the following equations:

$$y_{m,t,i} = x_{m,t,i} - \bar{B}_{m,IFL} \quad (1)$$

$$\bar{B}_{m,IFL} = \frac{\sum_{j=1}^{n_{IFL}} (x_{m,t,j} - r_{m,j})}{n_{IFL}} \quad (2)$$

where, $x_{m,t,i}$ is the original AVHRR value of metric *m* in year *t* and pixel *i*, $y_{m,t,i}$ is the DOS-adjusted AVHRR value, $\bar{B}_{m,IFL}$ is the mean bias of metric *m* over a total of n_{IFL} IFL pixels indexed by *j*, $r_{m,j}$ is the MODIS reference value of metric *m* in IFL pixel *j*.

The soil-induced bias was then corrected relative to the desert end-member, which has maximum residual bias after DOS correction, as well as the IFL end-member, which has minimum residual bias. Dense vegetation is largely immune to this correction. The correction is summarized by the following equations:

$$z_{m,t,i} = y_{m,t,i} - \bar{B}_{m,DES} * \frac{(v_{t,i} - \bar{V}_{t,IFL})}{(\bar{V}_{t,DES} - \bar{V}_{t,IFL})} \quad (3)$$

$$\bar{B}_{m,DES} = \frac{\sum_{k=1}^{n_{DES}} (y_{m,t,k} - r_{m,k})}{n_{DES}} \quad (4)$$

$$\bar{V}_{t,IFL} = \frac{\sum_{j=1}^{n_{IFL}} v_{t,j}}{n_{IFL}} \quad (5)$$

$$\bar{V}_{t,DES} = \frac{\sum_{k=1}^{n_{DES}} v_{t,k}}{n_{DES}} \quad (6)$$

where, $z_{m,t,i}$ is the slope-adjusted AVHRR value of metric *m* in year *t* and pixel *i*, $y_{m,t,i}$ is the DOS-adjusted value from equation (1), $\bar{B}_{m,DES}$ is the mean bias of metric *m* over a total of n_{DES} desert (DES) pixels indexed by *k*, $v_{t,i}$ is the peak growing season NDVI value of pixel *i* in year *t*, $\bar{V}_{t,IFL}$ is the mean peak growing season NDVI value of all IFL pixels, $\bar{V}_{t,DES}$ is the mean peak growing season NDVI value of all desert pixels, and $r_{m,k}$ is the MODIS reference value of metric *m* in desert pixel *k*. Here we use peak growing season NDVI, which is one of the metrics and computed as the mean of all NDVI values between 75 and 100 percentiles, in the slope term instead of the annual mean NDVI as used in Gutman (1999)[60], because our annual VCF represents the vegetation state of the local peak growing season. Using this annual metric (before any correction) dynamically optimizes AVHRR data for the growing season of each year.

Adjusted annual metrics were used as input to supervised regression tree models to generate the annual TC and BG product. This non-parametric machine learning method was chosen as it can accommodate nonlinear relationships between the dependent variable (percent TC or percent BG) and independent variables (AVHRR metrics); in addition, the decision rules are easily interpretable[68-70]. Training data for TC were obtained by spatially aggregating the circa-2000 Landsat-based percent TC product from 0.00025° × 0.00025° to 0.05° × 0.05°, which was in turn trained using very-high spatial resolution images[7]. For each 0.05° × 0.05° grid cell, we computed the average value of all Landsat TC pixels that fall in the grid cell and derived the percentage of TC per grid cell. Likewise, training data for BG were obtained by spatially aggregating the circa-2000 Landsat-based percent BG product[26]. Model training and prediction were performed separately for TC and BG. We pooled two years of AVHRR metrics before and after 2000 (i.e. 1999 and 2001) as input features to train 21 bagged regression tree models to account for the remaining inter-annual bias of AVHRR metrics, if any, as well as to avoid over-fitting of the regression tree algorithm. The 21 trained models were applied to annual AVHRR metrics to generate percent TC and BG for each year. Due to missing data in years 1994 and 2000, TC and BG maps in these two years were not produced from AVHRR, but were linearly interpolated using antecedent and subsequent annual TC or BG estimates on a per pixel basis. Following the MODIS VCF approach[33], annual SV was derived as the residual term by subtracting TC and BG percentages from 100. Permanent water surfaces were excluded based on the Landsat-derived permanent surface water product[7].

Accuracy assessment

Validating a global land-cover product spanning multiple decades is a challenge. The primary obstacle is the lack of sufficient ground observations that match the spatial extent, the temporal frequency and the thematic content of a satellite-derived product. Satellite observations with higher spatial and temporal resolutions can characterize land cover and change with higher accuracy[71,72]. Thus, higher-resolution satellite or aerial imagery is often employed to replace ground observations when determining the

reference condition for validation[73]. Here we leverage the established validation protocols[74,75] and the best available reference datasets to evaluate the accuracy of our VCF product. Specifically, we used a sub-meter resolution, global land-cover validation sample developed by the United States Geological Survey (USGS)[76] as the primary reference for TC. We also used the 30-m resolution Landsat-based TC, SV and BG estimates as reference to evaluate the AVHRR-derived TC, SV and BG layers.

The USGS reference dataset is a stratified random sample of TC estimates produced from n = 475 sample blocks distributed across the globe[74-76] (Supplementary Fig. 2a). Each sample block was 5-km × 5-km (~0.05° × 0.05°) in size. Sub-meter resolution commercial images including QuickBird, WorldView, IKONOS and GeoEye between years 2002 and 2014, depending on each block, were classified to categorical land cover classes including tree cover[76]. The percent TC for each block was computed from these data to provide the reference values for comparison to the AVHRR percent TC. The USGS reference data were developed in Universal Transverse Mercator (UTM) projection and the footprints of the 5-km × 5-km reference sample blocks did not exactly overlap with AVHRR pixels, which were in Geographical Latitude / Longitude projection (Supplementary Fig. 2b-c). This geolocation mismatch inevitably introduced some error in the validation results. Thus, we also evaluated AVHRR TC using the Landsat-based TC estimates. Because the spatial units of the Landsat estimates were spatially aligned with the AVHRR pixels, this comparison is free from geolocation error. For BG and SV, due to the lack of reliable high-resolution reference data, we used Landsat-based BG and SV (computed as 100% – Landsat-based BG% – Landsat-based TC%) estimates at the USGS sample locations as reference data for estimating accuracy. These BG and SV reference data were obtained for the same stratified sample of blocks used to evaluate the AVHRR TC product[74,75].

The paired AVHRR and reference VCF values were used to calculate four accuracy metrics including root-mean-square-error (RMSE), mean absolute error (MAE), mean error (ME) and r^2 [75,77]:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} w_i*(p_i-r_i)^2}{\sum_{i=1}^{n} w_i}} \quad (7)$$

$$MAE = \frac{\sum_{i=1}^{n} w_i*|p_i-r_i|}{\sum_{i=1}^{n} w_i} \quad (8)$$

$$ME = \frac{\sum_{i=1}^{n} w_i*(p_i-r_i)}{\sum_{i=1}^{n} w_i} \quad (9)$$

$$r^2 = 1 - \frac{\sum_{i=1}^{n}(p_i-r_i)^2}{\sum_{i=1}^{n}(p_i-\bar{r})} \quad (10)$$

where p_i, r_i and w_i are estimated VCF, reference VCF and sample weight (inverse of inclusion probability of the sample block for the stratified design) at a location *i* in a sample of size *n*; $\bar{r}$ is the estimated mean of the reference values.

We also computed the conventional confusion matrices including overall accuracy (OA), user's accuracy (UA) and producer's accuracy (PA) using the paired AVHRR and reference VCF values and a general ratio estimator[75,78]:

$$\hat{R} = \frac{\sum_{h=1}^{H} N_h*\bar{y}_h}{\sum_{h=1}^{H} N_h*\bar{x}_h} \quad (11)$$

where, *H* is the total number of strata; N_h is the total number of 5-km × 5-km blocks within stratum *h*; $\bar{y}_h$ and $\bar{x}_h$ are the sample means of variables *y* and *x* in stratum *h* and the specific identity of *y* and *x* depends on the accuracy metric being estimated. To estimate OA, *y* = area of agreement between AVHRR and reference for a VCF class *c* in each sample block (i.e., overlapped area) and *x* = area of the sample block. To estimate UA, *y* = area of agreement between AVHRR and reference for a VCF class *c* and *x* = area of class *c* mapped by AVHRR. To estimate PA, *y* = area of agreement between AVHRR and reference for a VCF class *c* and *x* = area of class *c* given by reference.

The estimated variance of $\hat{R}$ is:

$$\hat{V}(\hat{R}) = \frac{1}{\hat{X}^2} \sum_{h=1}^{H} N_h^2(1 - n_h/N_h)(s_{yh}^2 + \hat{R}^2 * s_{xh}^2 - 2 * \hat{R} * s_{xyh})/n_h \quad (12)$$

where $\hat{X}^2 = \sum_{h=1}^{H} N_h * \bar{x}_h$, n_h is the number of sample blocks selected from stratum h, s_{yh}^2 and s_{xh}^2 are the sample variances of y and x for stratum h and s_{xyh} is the sample covariance of x and y for stratum h. The standard error of $\hat{R}$ is the square root of the estimated variance. As noted above, the identity of x and y depends on the accuracy metric being estimated. A summary of accuracy results for TC, SV and BG is provided in Supplementary Fig. 2.

Trend analysis

Per-pixel TC, SV and BG percentages were aggregated to a series of spatial scales including global, continental, climate zone, biome and country scales to obtain annual total area estimates at these aggregated scales. For example, for the trend analysis of Africa, the per-pixel values of each cover type were aggregated to produce a single value for each year in the time series. We used the FAO ecological zones boundary shapefile to report VCF area estimates per biome and per climate zone[30]. We also used the Global Administrative Areas (GADM) country boundary shapefile (http://www.gadm.org) to report VCF area estimates per country.

The approach to change analysis was predicated on using a linear trend (Theil-Sen estimator) to smooth the annual time series of data when determining net change[79]. Although the classification methodology (monthly compositing, annual metrics calculation, inter-annual bias adjustment and multi-year model training) was constructed to ensure year-to-year consistency to the degree possible, the smoothing approach was still necessary because of the annual variation in the percent TC, SV, and BG values attributable to a variety of sources including different weather conditions, varying vegetation phenology, and image misregistration. As such, explicitly differencing the annual layers to estimate year-to-year change is not reliable. For TC, SV and BG time series in each aggregated spatial unit (e.g., a biome or a country), we applied the Theil-Sen estimator to derive the slope (annual change) of trend and provide the estimate of net change between 1982 and 2016 (i.e., slope times 34 years). The upper and lower change estimates based on the 90% confidence interval for the slope were also derived (Extended Data Tables 1

and 2, Supplementary Table 1). It is important to point out that the derived Theil-Sen trend represents long-term land-cover changes as the effect of changes in sensor capabilities has been effectively removed.

We further imposed the objective constraint of statistical significance of the trend to define net change at the pixel level. A Mann-Kendall test was applied to the TC, SV, and BG time series in each pixel[80]. If the Mann-Kendall test was not statistically significant ($p \geq 0.05$), we defined net change as 0. If the trend test was significant ($p < 0.05$), we applied the Theil-Sen estimator to estimate the per-pixel net change between 1982 and 2016. The significance level for these individual tests was not adjusted to control the Type I error rate for the suite of multiple tests (the so-called "experimentwise error rate"). These non-parametric statistical methods were chosen due to their robustness for trend detection and insensitivity to outliers. They have been applied to detect the greenness trend of land surface using AVHRR-based NDVI and leaf area index datasets[29,81,82] as well as the microwave-based vegetation optical depth data[83]. Six global VCF gain (positive slope) and loss (negative slope) layers were derived: (i) tree canopy gain; (ii) tree canopy loss; (iii) short vegetation gain; (iv) short vegetation loss; (v) bare ground gain; and (vi) bare ground loss (Fig. 1b and Extended Data Fig. 1). Subsequently, per-pixel loss (gain) were aggregated to global, continental, climate zone, biome and country scales to derive gross loss (gain) estimates for each aggregated spatial unit (Extended Data Tables 1 and 2 , Supplementary Table 1).

Uncertainty analysis

The uncertainties of the area estimates of net land-cover change were characterized as statistical bounds (Extended Data Tables 1 and 2). Here we conducted an additional uncertainty analysis on gross change estimates to investigate whether the overall VCF trends hold true.

We first varied the statistical significance level in the Mann-Kendall trend test for defining change. Compared with change area estimates resulted from the $p < 0.05$ threshold, using $p < 0.1$ to define change, the estimated TC, BG and SV change area would differ by 6%, 2% and 14% respectively, whereas using $p < 0.01$ to define change, the estimated TC, BG and SV change area would differ by 16%, 1% and 31%

respectively. Moreover, the signs of TC, BG and SV change were consistent at all significance levels — net gain in TC, net loss in BG and net loss in SV.

We further investigated the effect of VCF mapping uncertainty on change characterization. We employed the deviance value (i.e., the sum of squared difference between predicted value and training reference value) of each leaf node of the bagged regression tree models and computed a root-mean-square-deviation (RMSD) layer as VCF prediction uncertainty[68,84]. This per-pixel uncertainty layer was produced for each year between 1982 and 2016. Since RMSD is a quantitative indicator of land-cover uncertainty, we compared it with the magnitude of land-cover change by constructing a "signal-to-noise" ratio. The uncertainty of change for a given pixel *i* is then represented by the ratio of land-cover change to RMSD, summarized using the following equations:

$$RMSD_{i,k} = \frac{\sum_{t=1}^{T} \sqrt[2]{\sum_{j=1}^{M} (p_{j,i,k} - l_{j,i,k})^2 / M}}{T} \quad (13)$$

$$\overline{RMSD_i} = \frac{\sum_{k=1}^{N} RMSD_{i,k}}{N} \quad (14)$$

$$r_i = \frac{\Delta VCF_i}{\overline{RMSD_i}} \quad (15)$$

where, for each pixel *i* in year *k*, the $RMSD_{i,k}$ is the mean value of *T* bagged regression tree models indexed by *t*, in which the corresponding leaf node consists of *M* pixels indexed by *j* and each pixel in the leaf node has a predicted value of $p_{j,i,k}$ and the training value of $l_{j,i,k}$; the annual mean $\overline{RMSD_i}$ (unit: percent land cover) is the average value of *N* years; the ratio metric r_i for each pixel *i* is computed as 1982-2016 VCF change within the pixel (ΔVCF_i in units of percent land cover) to 1982-2016 average model prediction uncertainty $\overline{RMSD_i}$.

A greater absolute value of the ratio metric r_i indicates lower uncertainty of land-cover change and vice versa (Extended Data Fig. 6). The density distributions (Extended Data Fig. 6e and 6f) suggest that for

any threshold (dashed lines), the proportion of area under the frequency curve for tree cover gain always exceeds tree cover loss and similarly the proportion of area under the frequency curve for bare ground loss always exceeds bare ground cover gain. Hence, the overall trends in r_i corroborate the main findings of our study, which are that there is a net gain in tree cover and a net loss in bare ground cover over the study period of 1982 to 2016.

Driver attribution

Drivers of land-cover and land-use change are diverse, multi-scale and interactive[1,85-89]. Different drivers can be most broadly classified into two groups: anthropogenic and natural. Anthropogenic drivers are mainly related to land-use activities (e.g., deforestation, agricultural expansion, agricultural intensification, infrastructure construction and resource extraction), which are in turn driven by a number of underlying demographic, economic, technological, institutional, and cultural factors. Natural land-change drivers also include a variety of agents such as wildfire, drought, flood, windthrow, landslide, disease, insect attack, natural vegetation growth and glacial retreat, many of which are related to long-term climatic variation. Different drivers interact with each other in complex ways and the interactions are even evident at the broadest level in the Anthropocene[90]. With substantial human perturbations to the climate system, human-induced climate change and natural climatic variation and their effect on terrestrial ecosystems are intertwined. Disentangling human-induced climate change from natural climatic variation is a challenge, which can be studied using Earth system models[91]. Our objective for the global driver attribution was to provide a statistical, observation-based estimate of the relative contribution of direct human activities versus indirect drivers (including the combined effects of natural and human-induced climate change) to the observed global land change. Regionally dominant, specific land-change drivers were not explicitly quantified, but were identified and summarized through a comprehensive literature review.

We used a global probability sample and interpretation of high resolution images from Google Earth to estimate the proportion of changes attributable to drivers[92,93], separately for each VCF change type: (i)

tree canopy gain; (ii) tree canopy loss; (iii) short vegetation gain; (iv) short vegetation loss; (v) bare ground gain; and (vi) bare ground loss. For each VCF change type, 250 sample pixels (a pixel is a 0.05° × 0.05° grid cell) were selected with probability proportional to each pixel's absolute change area (-1 * change area in the case of loss) of the target VCF change type, where the area of change was obtained from the global change layers described above. A total of 1500 sample pixels were selected (Extended Data Fig. 3a). For each sample pixel, we created a polygon feature representing its boundary and imported it in Google Earth (Extended Data Fig. 4). Each polygon was also divided into 25 0.01° × 0.01° grid cells to aid photo interpretation. We used high-resolution images and the time slider tool in Google Earth to estimate the proportion of a pixel under human land use, including forestry and agricultural landscapes, cities, villages, houses, roads and other artificial objects. This proportion value was defined as the direct human impact associated with land-cover and land-use changes within the pixel. The impact of indirect drivers was defined as the residual of direct human impact. Areas of long-term land degradation resulting from the combined effects of land use and climate change were labeled as indirect if no signs of land use, for example fence lines or grazing paddocks, were observed. We estimated the direct human impact for each VCF change type as well as all land changes, using the following equations:

$$H_c = \frac{\sum_{j=1}^{n_c} h_j}{n_c} \tag{16}$$

$$OH = \frac{\sum_{c=1}^{6} w_c * H_c}{\sum_{c=1}^{6} w_c} \tag{17}$$

where H_c is the direct human impact of each of the 6 cover change types indexed by c, h_j is the proportion of pixel j that is under human land use, n_c is sample size (n_c= 250), OH is the overall direct human impact of all land changes, and w_c is the weight of each cover change type, given by the proportion of its global area over total absolute change area of all types (Extended Data Table 1). Similarly, we also estimated the overall direct human impact for all land changes within a continent and a biome. Attribution results are summarized in Extended Data Fig. 3.

References

32 DeFries, R. S., Townshend, J. R. G. & Hansen, M. C. Continuous fields of vegetation characteristics at the global scale at 1-km resolution. *J. Geophys. Res.* **104**, 16911-16923 (1999).

33 Hansen, M. C. *et al.* Global percent tree cover at a spatial resolution of 500 meters: First results of the MODIS vegetation continuous fields algorithm. *Earth Interact.* **7**, 10 (2003).

34 Hansen, M. C. *et al.* Towards an operational MODIS continuous field of percent tree cover algorithm: examples using AVHRR and MODIS data. *Remote Sens. Environ.* **83**, 303-319 (2002).

35 Hansen, M. C., Townshend, J. R. G., DeFries, R. S. & Carroll, M. Estimation of tree cover using MODIS data at global, continental and regional/local scales. *Int. J. Remote Sens.* **26**, 4359-4380 (2005).

36 Hansen, M. C. *et al.* Continuous fields of land cover for the conterminous United States using Landsat data: First results from the Web-Enabled Landsat Data (WELD) project. *Remote Sens. Lett.* **2**, 279-288 (2011).

37 Jennings, S. B., Brown, N. D. & Sheil, D. Assessing forest canopies and understorey illumination: Canopy closure, canopy cover and other measures. *Forestry* **72**, 59-74 (1999).

38 Korhonen, L., Korhonen, K. T., Rautiainen, M. & Stenberg, P. Estimation of forest canopy cover : A comparison of field measurement techniques. *Silva Fenn.* **40**, 577-588 (2006).

39 Hansen, M. C. *et al.* Monitoring conterminous United States (CONUS) land cover change with web-enabled Landsat data (WELD). *Remote Sens. Environ.* **140**, 466-484 (2014).

40 Pinzon, J. & Tucker, C. A non-stationary 1981–2012 AVHRR NDVI3g time series. *Remote Sens.* **6**, 6929-6960 (2014).

41 Tucker, C. *et al.* An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *Int. J. Remote Sens.* **26**, 4485-4498 (2005).

42 Franch, B. *et al.* A 30+ year AVHRR land surface reflectance climate data record and its application to wheat yield monitoring. *Remote Sens.* **9**, 296 (2017).

43 Pedelty, J. *et al.* Generating a long-term land data record from the AVHRR and MODIS Instruments. *IEEE Int. Geosci. Remote Sens. Symp. Proc.*, 1021-1025 (2007).

44 Tucker, C. J. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* **8**, 127-150 (1979).

45 Holben, B. N. Characteristics of maximum-value composite images from temporal AVHRR data. *Int. J. Remote Sens.* **7**, 1417-1434 (1986).

46 Loveland, T. R. *et al.* Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data. *Int. J. Remote Sens.* **21**, 1303-1330 (2000).

47 Hansen, M. C., DeFries, R. S., Townshend, J. R. G. & Sohlberg, R. A. Global land cover classification at 1 km spatial resolution using a classification tree approach. *Int. J. Remote Sens.* **21**, 1331-1364 (2000).

48 DeFries, R. S., Hansen, M., Townshend, J. R. G. & Sohlberg, R. Global land cover classifications at 8 km spatial resolution: The use of training data derived from Landsat imagery in decision tree classifiers. *Int. J. Remote Sens.* **19**, 3141-3168 (1998).

49 Bartholomé, E. & Belward, A. S. GLC2000: A new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* **26**, 1959-1977 (2005).

50 Lloyd, D. A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. *Int. J. Remote Sens.* **11**, 2269-2279 (1990).

51 Reed, B. C. *et al.* Measuring phenological variability from satellite imagery. *J. Veg. Sci.* **5**, 703-714 (1994).

52 DeFries, R., Hansen, M. & Townshend, J. Global discrimination of land cover types from metrics derived from AVHRR pathfinder data. *Remote Sens. Environ.* **54**, 209-222 (1995).

53 Latifovic, R., Pouliot, D. & Dillabaugh, C. Identification and correction of systematic error in NOAA AVHRR long-term satellite data record. *Remote Sens. Environ.* **127**, 84-97 (2012).

54 Song, Z., Liang, S., Wang, D., Zhou, Y. & Jia, A. Long-term record of top-of-atmosphere albedo over land generated from AVHRR data. *Remote Sens. Environ.* **211**, 71-88, (2018).

55 Kaufmann, R. K. *et al.* Effect of orbital drift and sensor changes on the time series of AVHRR vegetation index data. *IEEE Trans. Geosci. Remote Sens.* **38**, 2584-2597 (2000).

56 Kaufman, Y. J. & Holben, B. N. Calibration of the AVHRR visible and near-IR bands by atmospheric scattering, ocean glint and desert reflection. *Int. J. Remote Sens.* **14**, 21-52 (1993).

57 Vermote, E. & Kaufman, Y. J. Absolute calibration of AVHRR visible and near-infrared channels using ocean and cloud views. *Int. J. Remote Sens.* **16**, 2317-2340 (1995).

58 Los, S. O. Estimation of the ratio of sensor degradation between NOAA AVHRR channels 1 and 2 from monthly NDVI composites. *IEEE Trans. Geosci. Remote Sens.* **36**, 206-213 (1998).

59 Myneni, R. B., Tucker, C. J., Asrar, G. & Keeling, C. D. Interannual variations in satellite-sensed vegetation index data from 1981 to 1991. *J. Geophys. Res. Atmos.* **103**, 6145-6160 (1998).

60 Gutman, G. G. On the use of long-term global data of land reflectances and vegetation indices derived from the advanced very high resolution radiometer. *J. Geophys. Res. Atmos.* **104**, 6241-6255 (1999).

61 Donohue, R. J., Roderick, M. L. & McVicar, T. R. Deriving consistent long-term vegetation information from AVHRR reflectance data using a cover-triangle-based framework. *Remote Sens. Environ.* **112**, 2938-2949 (2008).

62 Hansen, M. C. *et al.* A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo basin. *Remote Sens. Environ.* **112**, 2495-2513 (2008).

63 Potapov, P. V. *et al.* Quantifying forest cover loss in Democratic Republic of the Congo, 2000–2010, with Landsat ETM+ data. *Remote Sens. Environ.* **122**, 106-116 (2012).

64 Chavez Jr, P. S. Radiometric calibration of Landsat Thematic Mapper multispectral images. *Photogramm. Eng. Remote Sensing* **55**, 1285-1294 (1989).

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

66 Woodcock, C. E., Macomber, S. A., Pax-Lenney, M. & Cohen, W. B. Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. *Remote Sens. Environ.* **78**, 194-203 (2001).

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

68 Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. *Classification and regression trees*. (Chapman & Hall/CRC, 1984).

69 Hansen, M., Dubayah, R. & Defries, R. Classification trees: An alternative to traditional land cover classifiers. *Remote Sens. Lett.* **17**, 1075-1081 (1996).

70 Friedl, M. A. & Brodley, C. E. Decision tree classification of land cover from remotely sensed data. *Remote Sens. Environ.* **61**, 399–409 (1997).

71 Mascorro, V. S., Coops, N. C., Kurz, W. A. & Olguin, M. Choice of satellite imagery and attribution of changes to disturbance type strongly affects forest carbon balance estimates. *Carbon Balance Manag.* **10**, 30 (2015).

72 Song, X.-P., Huang, C., Sexton, J. O., Channan, S. & Townshend, J. R. Annual detection of forest cover loss using time series satellite measurements of percent tree cover. *Remote Sens.* **6**, 8878-8903 (2014).

73 Strahler, A. H. *et al.* Global land cover validation: recommendations for evaluation and accuracy assessment of global land cover maps. (Office for Official Publications of the European Communities, Luxembourg, 2006).

74 Olofsson, P. *et al.* A global land-cover validation data set, part I: Fundamental design principles. *Int. J. Remote Sens.* **33**, 5768-5788 (2012).

75 Stehman, S. V., Olofsson, P., Woodcock, C. E., Herold, M. & Friedl, M. A. A global land-cover validation data set, II: Augmenting a stratified sampling design to estimate accuracy by region and land-cover class. *Int. J. Remote Sens.* **33**, 6975-6993 (2012).

76 Pengra, B., Long, J., Dahal, D., Stehman, S. V. & Loveland, T. R. A global reference database from very high resolution commercial satellite data and methodology for application to Landsat derived 30m continuous field tree cover data. *Remote Sens. Environ.* **165**, 234-248 (2015).

77 Willmott, C. J. Some comments on the evaluation of model performance. *Bull. Am. Meteorol. Soc.* **63**, 1309-1313 (1982).

78 Cochran, W. G. *Sampling Techniques (3rd edition)*. (John Wiley & Sons, 1977).

79 Sen, P. K. Estimates of the regression coefficient based on Kendall's tau. *J. Am. Stat. Assoc.* **63**, 1379-1389 (1968).

80 Mann, H. B. Nonparametric tests against trend. *Econometrica* **13**, 245-259 (1945).

81 Guay, K. C. *et al.* Vegetation productivity patterns at high northern latitudes: A multi-sensor satellite data assessment. *Glob. Change Biol.* **20**, 3147-3158 (2014).

82 Fensholt, R. *et al.* Greenness in semi-arid areas across the globe 1981–2007 — an Earth observing satellite based analysis of trends and drivers. *Remote Sens. Environ.* **121**, 144-158 (2012).

83 Andela, N., Liu, Y. Y., van Dijk, A. I. J. M., de Jeu, R. A. M. & McVicar, T. R. Global changes in dryland vegetation dynamics (1988–2008) assessed by satellite remote sensing: comparing a new passive microwave vegetation density record with reflective greenness data. *Biogeosciences* **10**, 6657-6676 (2013).

84 Breiman, L. Bagging Predictors. *Machine Learning* **24**, 123-140 (1996).

85 Lambin, E. F. *et al.* The causes of land-use and land-cover change: Moving beyond the myths. *Global Environ. Change* **11**, 261-269 (2001).

86 Geist, H. J. & Lambin, E. F. Proximate Causes and Underlying Driving Forces of Tropical Deforestation. *Bioscience* **52**, 143-150 (2002).

87 DeFries, R. S., Rudel, T., Uriarte, M. & Hansen, M. Deforestation driven by urban population growth and agricultural trade in the twenty-first century. *Nature Geoscience* **3**, 178-181 (2010).

88 Song, X.-P., Huang, C., Saatchi, S. S., Hansen, M. C. & Townshend, J. R. Annual carbon emissions from deforestation in the Amazon basin between 2000 and 2010. *PLoS One* **10**, e0126754 (2015).

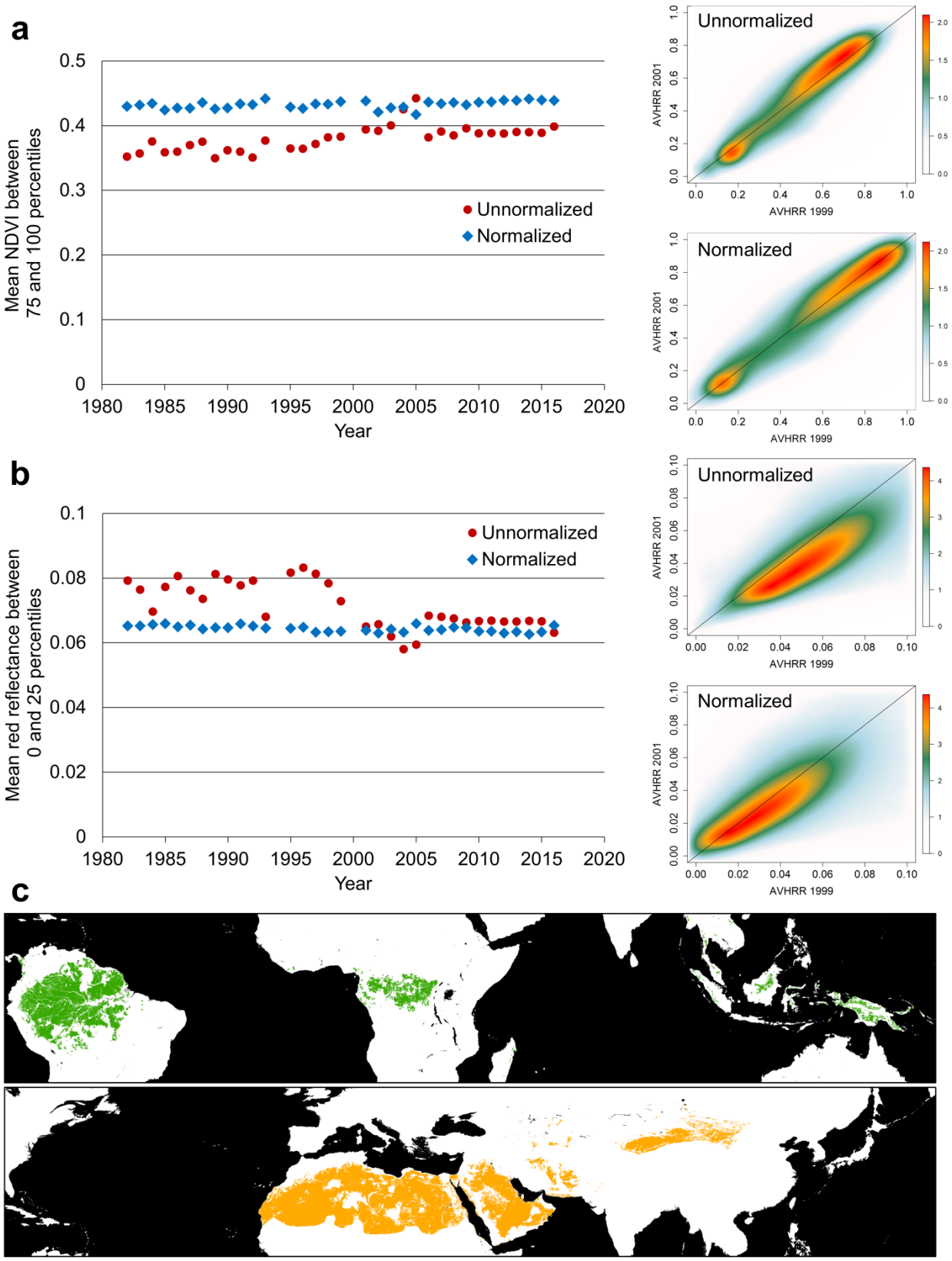
89 Liu, S. *et al.* Grand Challenges in Understanding the Interplay of Climate and Land Changes. *Earth Interact.* **21**, 1-43 (2017).

90 Steffen, W., Grinevald, J., Crutzen, P. & McNeill, J. The Anthropocene: Conceptual and historical perspectives. *Philos. Trans. A Math. Phys. Eng. Sci.* **369**, 842-867 (2011).

91 Mao, J. *et al.* Human-induced greening of the northern extratropical land surface. *Nat. Clim. Change* **6**, 959-963 (2016).

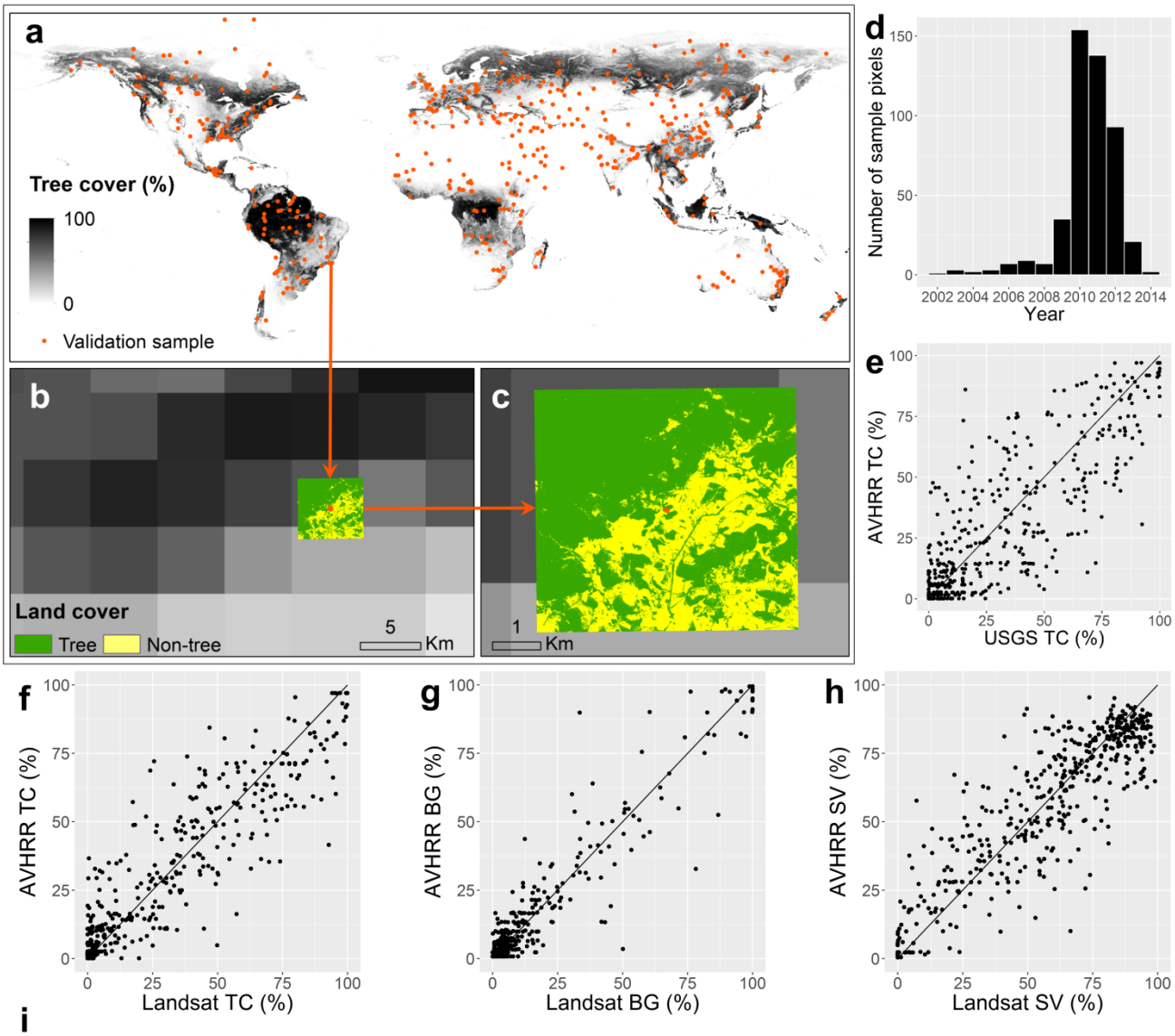
92 De Sy, V. *et al.* Land use patterns and related carbon losses following deforestation in South America. *Environ. Res. Lett.* **10**, 124004 (2015).

93 Tyukavina, A. *et al.* Types and rates of forest disturbance in Brazilian Legal Amazon, 2000–2013. *Sci. Adv.* **3**, e1601047 (2017).



Supplementary Figure 1 | Adjusting systematic biases in annual AVHRR metrics using Moderate Resolution Imaging Spectroradiometer data as reference. a, The metric displayed is the mean Normalized Difference Vegetation Index (NDVI) between the 75th and 100th percentiles. This metric is

related to the local peak growing season and was the most important variable driving the derived regression tree models for bare ground cover estimation. **b**, The metric displayed is mean red reflectance between the 0th and 25th percentiles. This metric is also related to the local peak growing season and was the most important variable for tree cover estimation. In both **a** and **b**, the dot plots on the left show the normalized and unnormalized annual global mean values (n = 5,806,001 pixels), excluding Antarctica and Greenland, and the density scatter plots on the right show pixel-level comparison between years 1999 and 2001 before (top) and after normalization (bottom). Normalizing AVHRR using the Moderate Resolution Imaging Spectroradiometer data as a reference not only changes the absolute value and data range of each individual year but also enhances year-to-year consistency. **c**, Maps of the intact forest landscape (top, green) and deserts (bottom, orange) that are used as stable targets for normalization.



i

Products	Class	User's accuracy % (SE)	Producer's accuracy % (SE)	Overall accuracy % (SE)	RMSE % (SE)	MAE % (SE)	ME % (SE)	r^2
AVHRR TC vs USGS TC	TC	84 (2)	79 (1)	90.1 (0.6)	15.6 (5.0)	9.9 (0.6)	-1.5 (0.8)	0.73
	Not TC	92 (1)	94 (1)					
AVHRR TC vs Landsat TC	TC	85 (1)	88 (1)	93.4 (0.4)	11.2 (3.9)	6.6 (0.4)	1.0 (0.5)	0.85
	Not TC	96 (1)	95 (1)					
AVHRR BG vs Landsat BG	BG	91 (2)	92 (1)	95.6 (0.5)	9.1 (5.1)	4.4 (0.5)	0.4 (0.7)	0.93
	Not BG	97 (1)	97 (1)					
AVHRR SV vs Landsat SV	SV	91 (1)	90 (1)	90.4 (0.6)	14.0 (5.4)	9.6 (0.6)	-0.5 (0.8)	0.79
	Not SV	90 (1)	91 (1)					

Supplementary Figure 2 | Accuracy assessment of the global land-cover product based on a validation sample of AVHRR pixels (n = 475). a, Spatial distribution of the validation samples (red dots) overlaid on long-term (1982–2016) mean tree cover. The United States Geological Survey (USGS) tree cover reference data (5-km × 5-km, Universal Transverse Mercator projection) have greater spatial details (coloured squares in **b** and **c**) owing to their sub-metre resolution, but have a geo-location mismatch with the AVHRR product (0.05° × 0.05°, grey-scale squares in **b** and **c**) owing to different projections. **d**, Temporal distribution of the USGS tree cover sample. **e**, Scatter plots of AVHRR TC cover against USGS reference. AVHRR and reference are matched by year and centre coordinates. **f**–**h**, Scatter plots of AVHRR TC, BG and SV cover (year 2001) against Landsat-based estimates, which are free from geo-location mismatch. **i**, Quantitative error metrics, including conventional confusion matrices as well as root-mean-square-error (RMSE), mean absolute error (MAE), mean error (ME) and r^2.