Global land use extent and dispersion within natural land cover using Landsat data

Matthew C Hansen*, Peter V Potapov, Amy H Pickens, Alexandra Tyukavina, Andres Hernandez-Serna, Viviana Zalles, Svetlana Turubanova, Indrani Kommareddy, Steve V Stehman, Xiao-Peng Song and Anil Kommareddy

1 University of Maryland, College Park, United States of America
2 State University of New York, College of Environmental Science and Forestry, Syracuse, NY, United States of America
3 Texas Tech University, Lubbock, TX, United States of America
* Author to whom any correspondence should be addressed.
E-mail: mhansen@umd.edu

Keywords: land cover and land use, global change, remote sensing

Abstract

The conversion of natural land cover into human-dominated land use systems has significant impacts on the environment. Global mapping and monitoring of human-dominated land use extent via satellites provides an empirical basis for assessing land use pressures. Here, we present a novel 2019 global land cover, land use, and ecozone map derived from Landsat satellite imagery and topographical data using derived image feature spaces and algorithms suited per theme. From the map, we estimate the spatial extent and dispersion of land use disaggregated by climate domain and ecozone, where dispersion is the mean distance of land use to all land within a subregion. We find that percent of area under land use and distance to land use follow a power law that depicts an increasingly random spatial distribution of land use as it extends across lands of comparable development potential. For highly developed climate/ecozones, such as temperate and sub-tropical terra firma vegetation on low slopes, area under land use is contiguous and remnant natural land cover have low areal extent and high fragmentation. The tropics generally have the greatest potential for land use expansion, particularly in South America. An exception is Asian humid tropical terra firma vegetated lowland, which has land use intensities comparable to that of temperate breadbaskets such as the United States’ corn belt. Wetland extent is inversely proportional to land use extent within climate domains, indicating historical wetland loss for temperate, sub-tropical, and dry tropical biomes. Results highlight the need for planning efforts to preserve natural systems and associated ecosystem services. The demonstrated methods will be implemented operationally in quantifying global land change, enabling a monitoring framework for systematic assessments of the appropriation and restoration of natural land cover.

1. Introduction

Human modification of the global land surface reflects economic activity in the expansion of built infrastructure for residential, commercial, industrial, transportation and energy demands, agricultural production in feeding increasing numbers of people with rising standards of living, and land uses such as forestry that similarly support human material needs. Land use increases in extent through the appropriation of natural land cover into human production systems with significant impacts on earth system functioning [1-6]. Considerable effort in the form of international policies have attempted to reduce such impacts, with goals of mitigating climate change, maintaining fresh water supplies, and reducing habitat loss, including initiatives such as the Millennium Ecosystem Assessment, the United Nations Framework Convention on Climate Change, and the Convention on Biological Diversity.

Existing studies that seek to track the appropriation of natural lands into human-dominated landscapes rely on sourcing input layers that are inconsistent in terms of characterization and timing, as
with the human footprint of Venter et al. [7]. When attempting to infer the impacts of global land change on the earth system, the limitations are obvious if inputs are of different scales and provenance. Moving forward, continuity of data and methods for land monitoring will facilitate improved quantification of downstream impacts on the environment, whether modeling changes to biodiversity, water quality, or carbon stocks. Consistent, available global-scale earth observation data are required to support such initiatives [8], such as the Landsat and Sentinel 2 series of medium spatial resolution imagery. The systematic global-scale characterization of such data into land cover and land use will facilitate the move to operational monitoring capabilities of human appropriation and associated impact to key ecosystem services.

Here, we present a novel global land cover and land use map for 2019, with results synthesized to quantify land use impacts globally by climate domain and ecozone. Map classes were generated from Landsat time-series imagery and topographic data using per-class variants of training data, spectral features, and algorithms (table 1), with results for all classes in the resulting global land cover and land use map validated using good practice methods [9, 10]. Land cover classes included maximum vegetation cover (mapped as a percent per pixel), woody vegetation height (for vegetation \( \geq 3 \) m), surface water and permanent ice [11–13]. Land use classes included built-up land, cropland, and land use associated with tree cover loss and gain (mainly forestry and shifting cultivation, but excluding tree cover loss due to fire) [14]. The maps are generated in a consistent methodological framework with some having been published independently and updated for our study period, and others newly created (table 1). In this study, the terms ‘land cover’ and ‘land use’ are used to refer to land under the land cover and land use classes listed above. We characterized pastures if associated with climate, vegetation and landform will feature similar land use areal extent and dispersion, with variations explained by differing types and stages of land use change. We posit that climate domains/ecozones of similar climate, vegetation and landform will feature similar land use areal extent and dispersion, with variations explained by differing types and stages of land use change. We also expect that regions with high potential for economic gain will tend towards random land use expansion with resulting limited and fragmented remnant natural land cover.

Actual land use expansion into natural land cover is initially spatially non-random, with the expansion of infrastructure facilitating further land use change [19], and land use extent becoming more spatially uniform over time, a reality somewhere between the two aforementioned scenarios. This study seeks to determine empirically where different regions reside on this continuum, enabling a relative assessment of human impact on natural lands, where natural land is defined as land cover absent of our target land uses. We posit that climate domains/ecozones of similar climate, vegetation and landform will feature similar land use area extent and dispersion, with variations explained by differing types and stages of land use change. We also expect that regions with high potential for economic gain will tend towards random land use expansion with resulting limited and fragmented remnant natural land cover.

The methods applied here will be back cast over the Landsat record and implemented for forward processing in systematically tracking the extent and rates of expansion of human impact on the global land surface. Additional analyses incorporating population, economic, biodiversity, or other ancillary information may be readily implemented using the global land cover and land use data. All data from this study may be accessed at https://gland.umd.edu/dataset/global-land-cover-land-use-v1.
Table 1. Land cover, landform and land use themes derived for this study along with data inputs and algorithms. Throughout this study, the term ‘land cover’ is used to refer to land in the categories in green and ‘land use’ is used to refer to land in the categories in orange.

| Land cover/landform/land use class | Data availabilitya | Data type | Image feature space | Algorithm |
|------------------------------------|--------------------|-----------|---------------------|-----------|
| Percent maximum vegetation cover (Ying et al [11]) | Updated for 2019 | Continuous: 0%–100% cover | 2019 annual Landsat metrics | Global regression tree |
| Vegetation height (Potapov et al [22]) | Existing https://glad.umd.edu/dataset/land-use-vegetation-height | Continuous: 3–25 m | 2019 annual Landsat metrics | Locally calibrated regression trees |
| Inland water and permanent ice (Pickens et al [13]) | Existing https://glad.umd.edu/dataset/global-surface-water-dynamics | Categorical: yes/no | 2019 annual Landsat metrics | Regional QA models of water and snow/ice |
| Wetland (Pickens et al in preparation [46]) | Newly available | Categorical: yes/no | 2013–2018 Landsat and topographical metrics | Global classification tree |
| Built-up area (Hernandez-Serna et al in preparation [47]) | Newly available | Categorical: yes/no | 2019 annual Landsat metrics | Regional deep learning convolution neural networks |
| Cropland (Potapov et al [23]) | Existing https://glad.umd.edu/dataset/croplands | Categorical: yes/no | 2016–2019 multi-year Landsat metrics | Regional classification trees |
| Tree cover loss (Hansen et al [14]) | Existing https://glad.earthengine.app/view/global-forest-change | Categorical: yes/no | 2000–2019 annual Landsat metrics | Regional classification trees |
| Tree cover gain (Hansen et al [14]) | Updated for 2019 | Categorical: yes/no | 2000–2019 annual Landsat metrics | Global/regional hybrid decision tree |

a All data is freely accessible and available for download at https://glad.umd.edu/dataset/global-land-cover-land-use-v1.
2. Methods

2.1. Land cover and land use mapping

Earth observations from satellite-based sensors offer a systematically acquired, freely accessible global-scale data input for monitoring the land surface [8]. New computing capabilities enable the efficient processing of such data. Given appropriate computing and data inputs, advanced methods for characterizing the land surface can be applied. Such methods require: (a) systematically defined land cover and land use classes, defined by their relevant physiognomic-structural properties and discernible by time-series multispectral data; (b) remote sensing and image processing techniques, including the creation of purpose-built features that facilitate characterization of the land surface; (c) application of advanced algorithms appropriate for land surface mapping; (d) domain expertise to train the algorithm, specifically geographic/environmental knowledge for calibrating, trouble-shooting, and iterating model runs; and (e) rigorous validation in assessing product accuracy and confirming area estimation. This is the consistent methodological framework that we adhered to in generating global maps across different land cover/land use classes.

Table 1 lists the land cover and land use classes that were used to create the global land cover and land use map presented in this study and details the methods for characterization of each class. All spectral inputs consisted of multi-temporal metrics, which have been repeatedly employed by our team in global mapping applications [12, 14, 20, 21]. Metrics are statistical measures derived from time-sequential image composites, for example mean annual red reflectance, that retain salient phenological information without regard to time of year. In this way, they represent a generic global feature space that facilitates extrapolation of algorithms across large areas. To generate metrics for this study, we employ Landsat 16 d composite imagery from the Global Land Analysis and Discovery lab’s analysis ready data (GLAD ARD) set [22]. The 16 d data are aggregated over annual or multi-year intervals and are employed to calculate the metrics used as inputs for mapping land cover, land use, and landform. For wetlands, we added a set of topographical metrics, including slope, curvature and a series of relative elevation layers from variably sized catchments per previous mapping studies of the Democratic Republic of Congo [24] and Indonesia [25].

Algorithms consisted of bagging (bootstrap aggregation) decision tree ensembles [26] using entropy-based classification trees for categorical variables and sum of squares-based regression trees for continuous variables (table 1 shows which variables were categorical and which were continuous). For themes that employed bagged decision trees, a median estimate was taken as the map result. Training data were derived principally from expert-interpreted labeling. For percent bare ground, we employed previously derived percent cover data for global bare ground gain assessments [11, 27]. We also employed ancillary data such as height metrics from the Global Ecosystem Dynamics Investigation (GEDI) space-borne lidar [28] to map tree height [12] and roads from Open Street Maps to map built-up land. For some themes, such as cropland, we employed continental-scale models, while for
Figure 2. Flow chart of data inputs in quantifying land use extent and dispersion measures globally. Slope was calculated using Shuttle Radar Topography Mission (SRTM), TerraSAR-X add-on for Digital Elevation Measurement (TanDEM-X), and Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) digital elevation datasets. Colored outlines indicate source data for the figures listed in the legend. Climate domains are from the FAO. All other land cover, land use and wetland extent layers were developed by the authors and may be viewed and accessed here: [https://glad.umd.edu/dataset/global-land-cover-land-use-v1](https://glad.umd.edu/dataset/global-land-cover-land-use-v1).

other themes, such as bare ground, a generic global algorithm was applied.

Due to the contextual nature of built-up lands, particularly settlements, we employed a deep learning convolution neural network algorithm. We utilized the U-Net convolutional neural network architecture, which has proven to work robustly over a variety of tasks in remote sensing [29, 30]. To train this algorithm, we used a rasterized version of the roads dataset from Open Street Maps from which we randomly sampled patches with a size of 128 × 128 pixels (each pixel corresponds to 30 × 30 m). Weight decay and data augmentation were utilized in the form of down sampling, rotations, and flips for the gradient-based optimization of the U-Net parameters.

### 2.2. Climate domains/ecozones

Ancillary data consisting of climate domains from the United Nations Food and Agricultural Organization [15] ([http://foris.fao.org/static/data/fra2010/ecozones2010.jpg](http://foris.fao.org/static/data/fra2010/ecozones2010.jpg)) were used to disaggregate our land cover and land use data. FAO climate domains include tropical, sub-tropical, temperate, boreal and polar (‘arctic’ in our study). We use all these domains, and further disaggregate the ‘tropical’ domain into ‘humid tropical’ (‘rainforest’ in the FAO data) and ‘dry tropical’ (not ‘rainforest’ in the FAO data) sub-domains for our study. Each climate domain was subdivided into ecozones, which in the FAO’s system consist of desert and semi-arid vegetation types, mountain systems, and woodland and forest systems.

We employed an analogous approach to divide climate domains into remote sensing-derived ecozones including (a) bare ground, and terra firma vegetated lands on (b) low slopes (<4% slope) and (c) high slopes (≥4% slope), and (d) wetlands, which are absent from FAO ecozones. Our ecozones are clearly defined by their physical and structural characteristics and have the advantage of having the same spatial resolution as the map (30 m). Combining the FAO climate domains with the remote sensing-derived ecozones resulted in 24 regions, which are referred to as climate/ecozones in this paper and are listed in figure 2.
Table 2. Strata for sample allocation, derived from global land cover and land use map (see section 3), with 50 sample points selected per stratum. The strata are initially divided into seven analogous classes for (1) wetland and (2) non-wetland land forms. There is an additional set of strata that supersede the previous two categories, including ice, water, cropland, built-up, and ocean. Every pixel in the map can only belong to one of the 19 strata.

| Strata                        | Physiognomic/structural thresholds                                      |
|-------------------------------|-------------------------------------------------------------------------|
| **Non-wetland strata**        |                                                                         |
| 1. True desert                | $\geq 90\%$ bare ground                                                |
| 2. Semi-arid                  | $\geq 25\%$ to $< 90\%$ bare ground                                    |
| 3. Dense short vegetation     | $0\%$ to $< 25\%$ bare ground                                          |
| 4. Open tree cover            | $\geq 3$ and ($< 10$ m or $< 70\%$ tree cover)                         |
| 5. Dense tree cover           | $\geq 10$ m and $\geq 70\%$ tree cover                                |
| 6. Recent tree cover gain     | Recent $\geq 3$ m and $\geq 10\%$ tree cover                           |
| 7. Non-fire loss, no trees, no cropland, no built-up |                                                                         |
| **Wetland strata**           |                                                                         |
| 8. Salt pan                   | $\geq 90\%$ bare ground                                                |
| 9. Semi-arid                  | $\geq 25\%$ to $< 90\%$ bare ground                                    |
| 10. Dense short vegetation    | $0\%$ to $< 25\%$ bare ground                                          |
| 11. Open tree cover           | $\geq 3$ and ($< 10$ m or $< 70\%$ tree cover)                         |
| 12. Dense tree cover          | $\geq 10$ m and $\geq 70\%$ tree cover                                |
| 13. Recent tree cover gain    | Recent $\geq 3$ m and $\geq 10\%$ tree cover                           |
| 14. Non-fire loss, no trees, no cropland, no built-up |                                                                         |
| **Superseding strata**       |                                                                         |
| 15. Ice                       | Permanent ice                                                           |
| 16. Water                     | Permanent surface water                                                 |
| 17. Cropland                  | Cropland land use                                                       |
| 18. Built-up                  | Human-built surfaces and structures                                     |
| 19. Ocean                     | Ocean                                                                   |

Figure 2 illustrates how climate/ecozone labels were combined with land cover and land use. No land uses were assigned to bare ground or wetlands. Built-up and cropland land uses were allocated as either vegetated terra firma lowland or vegetated montane landforms. Treed land uses were allowed in wetlands as well as vegetated terra firma lowland and vegetated montane landforms. Built-up area and croplands were mapped for 2019 and treed land uses derived from tree cover gain and tree cover loss since 2000, excepting loss due to fire. A slope threshold of $> 4^\circ$ was selected to define montane landforms. This threshold was applied to the minimum of the derived slope from SRTM and TanDEM-X elevation datasets for $\leq 60^\circ$ N and from TanDEM-X and ASTER for $> 60^\circ$ N [31–33].

2.3. Validation
Validation of our global land cover and land use map consisted of a probability sample of reference data for the strata listed in Table 2. The 19 strata were derived from our 30 m continuous and discrete land covers and land uses and reflect typical classes for global mapping [34]. (The stratification was different from the climate/ecozones because it was designed to assess the accuracy of the mapped classes, whereas the climate/ecozone regions serve as a useful categorization of land for reporting results.) A stratified random sample of points was obtained by looping through a selection protocol until 50 sample points per stratum were found. First, we created an extensive set of random points globally weighted by the area of a $1^\circ$ square centered at that location to ensure that every point on the surface of the earth had an equal initial inclusion probability.

Each sample in this random list was attributed to the corresponding stratum in which it was located. We then selected the first 50 sample points for each stratum from the list. For each point, the bounding $0.00025$ lat lon $-1$ pixel from our ARD inputs [22], and the 30 m Universal Transverse Mercator (UTM) pixel from original terrain-corrected Landsat imagery were recorded to obtain the final stratified random sample of pixels. The selection protocol enables the selection of a pixel sample from various projections and resolutions. Global accuracy and area estimates were produced from the stratified sample following conventional practice [9, 10] and the accompanying standard errors quantify the variability of the estimates. Reference data associated with each sample point consisted of Google Earth imagery and Moderate Resolution Imaging Spectroradiometer (MODIS) time-series data. We recorded the polygonal extents of both the lat/lon and UTM grid cells in which our sample point fell and converted them to kmls to be uploaded to GoogleEarth for interpretation. Google Earth imagery were the primary source for assigning land use categories and general land cover.
distinctions such as forest versus non-forest. We also generated 250 m MODIS time-series of red, near-infrared, and both shortwave infrared bands, as well as normalized difference vegetation and water indices, all of which were graphically displayed as 16 d time-series composites. MODIS data were useful in land cover discriminations that were discernible through phenological profiles, for example, desert from semi-arid categories.

2.4. Percent land use versus distance to land use modeling

To examine the relationship of percent land use and distance to land use, we modeled both random and frontal expansion, using the two relationships as an envelope of possible expansion dynamics. For both scenarios, we employed a 1000 by 1000 cell grid. For the random expansion scenario, we incrementally increased percent land use within the grid with each increment randomly allocated. The grid was set to a spatial resolution of 500 m and the mean distance to land use for all pixels was calculated for each increment (for land use pixels, distance to land use is zero). For frontal expansion, we expanded land use as a contiguous block from one edge of the grid to the other, calculating distance to land use at each increment. To derive our empirical relationship, our 0.000 25° lat lon−1 map was reprojected and percent land use was averaged to a 500 m spatial resolution in an equal area sinusoidal projection where land use included our built-up, cropland and treed land use categories.

Distance to land use, which we also refer to as land use dispersion, was calculated as a measure of natural land cover intactness, similar to our study of humid tropical primary forests [35]. The 30 m aggregate land use extent layer (aggregate of built up, cropland, treed land uses) was converted to a binary yes/no-land use 500 m spatial resolution layer. Each new 500 m pixel was categorized as ‘land use’ if ≥10% of it was land use in the 30 m layer. If less than 10% of the 500 m pixel was land use, the pixel was labeled ‘not land use’. Using this 10% threshold, 96% of the 30 m resolution land use layer was included in the new 500 m resolution land use layer. However, despite the fact that the 500 m land use layer missed 4% of the 30 m land use, nearly equal amounts of land use omission and commission error were included in the new 500 m layer, meaning that the 500 m land use mask almost equaled (99.5%) the 30 m area totals. The benefit of including the 10% threshold was that the new 500 m layer was not as affected by commission errors in the form of noise or stray pixels in the product, which would deleteriously impact distance calculations. The ‘distance to land use’ results presented in this study consist of the mean distance from all 500 m pixels to the nearest land use pixel for each climate/ecozone (land use pixels were taken into consideration to calculate the mean, with distance to land use for these pixels equaling zero.).

3. Results

3.1. Map accuracy

Validation results are shown in table 3 and figure 3, with reference data accessible at https://glad.umd.edu/dataset/global-land-cover-land-use-v1. Accuracies for aggregate land use consisting of built-up land, cropland and treed land uses yielded a user’s accuracy of 93.6% (SE ± 2.1%) and a producer’s accuracy of 66.2% (SE ± 5.0%) and for land cover, 94.3% (SE ± 1.3%) and 99.2% (SE ± 0.3%). Results support the utility of the map for further spatial analyses.

3.2. Global climate/ecozones

Figure 4 shows (a) the global land cover/land use map, (b) the climate/ecozone sub-regions of study, and (c) the aggregate land use extent and distance to aggregate land use layers used to calculate the extent and dispersion of land use for sub-regions. Figure 5 illustrates results of area and distance to land use for the 24 climate/ecozone regions consisting of per climate domain vegetated terra firma lands on low slopes (hereafter referred to as lowlands), vegetated lands on high slopes (hereafter referred to as montane lands), wetlands, and bare ground. The overall relationship of proportional land use area and distance to land use follows a power law with a high coefficient of determination. At very low land use percentages, land use is nonrandom in its distribution and average distance to land use is many times greater than random allocations (figure 6). As percent land use increases, the map-based empirical model corresponds more closely to the randomly dispersed land use expansion scenario than to the frontal land use expansion scenario (figure 7). From our empirical relationship, if land use accounts for 40% of total area of an ecoregion, the mean distance of all land to land use is less than one kilometer, with implications for sustaining ecosystem functioning. In the aggregate, the regions are largely clustered in affinity groups. Terra firma vegetated lowlands in the temperate, sub-tropical, dry tropical, and humid tropical climate domains feature the largest proportional areas of land use. The shortest distances to land use are found in the sub-tropical and temperate domains. The ongoing land conversion processes in the tropics drive contemporary global land change, and the greater availability of unconverted terra firma vegetated lowland in these biomes indicates likely future extensification of land use. Boreal terra firma vegetated lowland is less developed due to climate limitations that for now preclude large-scale agriculture, but with climate change may lead to rapid land use conversions [36]. The Arctic is an undeveloped outlier.
Table 3. Accuracy assessment based on 950 sample points with 50 sample points selected from each of 19 map strata (table 2) and interpreted using GoogleEarth and MODIS time-series imagery. The ‘other land use’ label was used for sample points associated with human-induced disturbance or recovery, but not attributable to a specific land use.

| Land cover/land use                  | Users Producers SE (users) | SE (prod) | Reference area (km²) | SE (km²) | Map area (km²) |
|--------------------------------------|-----------------------------|-----------|----------------------|----------|----------------|
| Bare ground                          | 92.00%                      | 91.84%    | 3.87%                | 3.53%    | 195 282 77     | 131 113 | 194 940 42     |
| Semi-arid vegetation                 | 66.00%                      | 81.62%    | 6.77%                | 5.91%    | 155 767 10     | 167 169 | 192 629 35     |
| Dense short vegetation               | 64.00%                      | 70.45%    | 6.86%                | 4.29%    | 251 627 43     | 366 470 | 276 995 61     |
| Open or short tree cover             | 70.00%                      | 58.90%    | 6.55%                | 5.88%    | 179 745 64     | 218 966 | 151 244 83     |
| Dense and tall tree cover            | 78.00%                      | 87.36%    | 5.92%                | 3.45%    | 173 767 63     | 146 034 | 194 627 05     |
| Wetland                              | 62.40%                      | 59.58%    | 4.54%                | 8.52%    | 881 455 95     | 725 70  | 100 222 41     |
| Permanent surface water              | 96.09%                      | 99.45%    | 0.62%                | 0.40%    | 296 367 53     | 414 91  | 300 444 42     |
| Permanent ice                        | 94.00%                      | 80.19%    | 3.39%                | 14.56%   | 395 948 18     | 182     | 337 764        |
| Built-up land                        | 84.00%                      | 42.80%    | 5.24%                | 10.64%   | 387 810 8      | 235 21  | 197 595 8      |
| Cropland                             | 92.00%                      | 70.11%    | 3.87%                | 6.49%    | 152 185 99     | 140 985 | 115 969 57     |
| Treed land use                       | 76.73%                      | 57.82%    | 4.59%                | 9.89%    | 437 268 8      | 208 66  | 329 489 9      |
| Other land use                       | 0.00%                       | 0.00%     | 0.00%                | 0.00%    | 379 325 4      | 371     | 0              |
| Overall accuracy                     |                             |           |                      |          | 78.35          | SE = 1.86 |

Figure 3. Sample-based area estimates from reference data versus area estimates from map pixel counts. The dashed line is 1:1 and the values of the plotted data, along with the standard errors of the area estimates, are shown in table 3. Land cover/landform classes are in green and land use classes are in orange.

Vegetated montane lands cluster in a zone of ~10% land use and have considerably lower percent land use than lowland counterparts, particularly for crop-intensive temperate and sub-tropical climate domains. Boreal and arctic montane zones again lag other montane regions in land use extensification due to climates limiting both forest growth and exploitation. As shown in table 4, wetland extent across climate domains is highly variable and is inversely related to land use extent within climate domains, indicating historical wetland loss [37] for temperate, sub-tropical and dry tropical biomes, and nascent wetland loss in the humid tropics. The arctic has 56% of its non-montane vegetated land in wetlands, the boreal 28%, and the humid tropics 19%; these three climate domains account for 71% of total global wetland extent. Only 6% of the temperate climate domain, 8% of the sub-tropics and 9% of the dry tropics consists of wetlands. Wetlands are more challenging to appropriate and thus less suitable for land use expansion as compared to moderately sloped terra firma land. The inverse relationship between wetland extent and land use extent suggests that wetland extent may serve as a barometer of the level of human impact of natural lands in a region, since draining of wetlands for conversion to land use...
becomes an attractive option only when more easily appropriated natural land cover is scarce. Bare ground lands, per our method, are absent of land use across all biomes. Taken individually, the respective built-up, cropland and treed land uses have different area/distance relationships and lower correlations compared to aggregate land use (figure 8). Built-up land use has much less area compared to cropland and treed land uses, but is more spatially distributed, resulting in shorter mean distances of land to built-up areas per climate/ecozone.

3.3. Continental climate/ecozones

A further disaggregation of climate/ecozone by continent illustrates significant variations (figure 9). The most intensive land use systems are found in Eurasian dry tropical and sub-tropical terra firma vegetated lowlands, followed by a cluster that also includes Eurasian temperate and humid tropical terra firma vegetated lowlands. Europe and Asia are home to high population densities and corresponding built-up land, high production agriculture, and mature forestry practices. Vegetated terra firma

Figure 4. (a) Global land cover and land use for 2019, where cropland, built-up area, and tree cover loss and gain represent land use activity; (b) global climate domains intersected with remote sensing-derived ecozones; (c) global land use extent (percent area in land use) and dispersion (distance to land use). Distance to land use was calculated from 500 m grid cells with \( \geq 10\% \) land use (see section 2).
Figure 5. Log–log plot of land use extent and dispersion per global climate/ecozone. Climate: arc = arctic, bor = boreal, tem = temperate, sub = sub-tropical, dry = dry tropical, hum = humid tropical; ecozones: low = terra firma vegetated land on low slopes (lowland), mon = vegetated land on high slopes (montane), wet = wetland, and bar = bare ground. For the plotted data, $r^2 = 0.85$. For only those climate/ecozones with at least at least 1000 km$^2$ of land use, which excludes arctic montane and wetland climate/ecozones, $r^2 = 0.80$. The concentric circles are proportional to the area of each climate/ecozone (outer circle) and the area of land use for that climate/ecozone (inner circle/pie chart). Land use area per climate/ecozone is subdivided by land use class (built up, cropland, treed land use). By definition, bare ecozones do not have land use and are plotted separately for size comparison only in the upper right corner.

Figure 6. Log–log plot of climate/ecozone regions from figure 5 in the context of possible land use expansion pathways.

lowlands in North American temperate and sub-tropical climates as well as South American sub-tropical and African sub-tropical climates are on a par with similar regions in Eurasia. Of the humid tropical terra firma lowlands, or rainforest environments, Mainland and Insular Southeast Asia feature levels of human-dominated land use equal to that of temperate breadbaskets, such as the United States’
Figure 7. (a) Modeled distance to land use calculated as a function of percent of land use area for our map-based empirical model and for a randomly allocated model. Black circles represent climate domain/ecozones. The solid line is 1:1 for reference; (b) ratio of distances from (a) illustrating the transition of spatial patterns with increasing affinity to random distributions as percent area in land use increases.

Figure 8. Log–log plot of land use extent and dispersion per land use type per global climate/ecozone for only those climate/ecozones with at least 1000 km$^2$ of land use. Climate: arc = arctic, bor = boreal, tem = temperate, sub = sub-tropical, dry = dry tropical, hum = humid tropical; ecozones: low = terra firma vegetated land on low slopes (lowland), mon = vegetated land on high slopes (montane), wet = wetland. Size of outer circles is proportional to area of climate/ecozone with the portion of land use shown in gray.

Table 4. Built-up, cropland and treed land use in terra firma vegetation lowlands and wetlands by climate domain.

| Climate domain | Terra firma vegetation on low slopes + wetland (km$^2$) | Land use extent (km$^2$) | Wetland extent (km$^2$) | % land use | % wetland |
|----------------|--------------------------------------------------------|--------------------------|--------------------------|------------|-----------|
| Arctic         | 264 6084                                               | 1241                     | 147 7249                 | 0.05%      | 55.83%    |
| Boreal         | 124 213 28                                             | 564 255                  | 351 0649                 | 4.54%      | 28.26%    |
| Temperate      | 120 643 82                                             | 531 1100                 | 679 051                  | 44.02%     | 5.63%     |
| Subtropical    | 642 7868                                               | 275 5243                 | 519 570                  | 42.86%     | 8.08%     |
| Dry tropical   | 192 989 39                                             | 485 0131                 | 163 2575                 | 25.13%     | 8.46%     |
| Humid tropical | 993 6022                                               | 133 7852                 | 184 1700                 | 13.46%     | 18.54%    |

Midwest. For both terra firma vegetated lowland and montane humid tropical lands, South America has the most intact and isolated natural lands, followed by Africa, and then Asia. The plot in figure 9 may be viewed from a temporal perspective as rainforests transition from low percent land use and high distance to land use to a converted state of high percent land use and low distance to land use, as one moves from South America to Africa to Asia (black line in plot).

3.4. Climate domain/ecozone area and distance measures by land use type

Different land uses dominate in different regions, for example treed land uses in montane regions (figure 8). Built-up area, particularly road and rail networks, are present in all climate domains and ecozones, serving to generically connect nodes of population and associated economic activity, resulting in shorter overall distances to land use per sub-region. Cropland, on the other hand, covers the...
largest area, but is more spatially clustered, resulting in larger distances of land to cropland land use, except for the most intensively developed climate/ecozones, such as temperate and sub-tropical terra firma vegetated lowlands. Treed land uses fall between built-up and cropland land uses in both measures. Built-up and cropland land uses have higher correlations ($r^2 = 0.67$ and $r^2 = 0.65$, respectively) for percent area in land use and distance to land use than treed land uses ($r^2 = 0.29$). Vegetated montane landforms are generically exploited globally where possible, albeit at a magnitude lesser than vegetated terra firma lowlands. The cropland relationship has two outliers, boreal and humid tropical terra firma vegetated lowlands; each of these climate/ecozones feature high distance to land use compared to other climate domain/ecozones. Boreal terra firma vegetated lowland exhibits clustered cropland land use along its southern fringe, resulting in significantly larger distances of land to cropland, a spatial pattern of cropland land use that is unlike that of lower latitudes. Humid tropical terra firma vegetated lowland is the site of considerable cropland expansion. For regions such as the Amazon Basin, cropland expansion occurs largely along its southeastern frontier, resulting in higher distance to land use due to the frontal nature of conversion dynamics.

4. Discussion

The strength of the empirical relationship of percent land use area and distance to land use indicates that land use expands in an increasingly diffuse fashion for regions of nominally uniform development potential. The most developed climate/ecozones, such as the temperate and sub-tropical terra firma vegetated lowland zones, with over 40% of land in intensive land use, have crossed into an infinitely connected land use cluster, indicating low areal extent and highly fragmented remnant natural land cover. The location of any climate/ecozone along this relationship (figures 5 and 6) likely reflects either the stage of ongoing land use expansion, intrinsic limitations to land use expansion, or extrinsic factors such as land use planning efforts that slow expansion, particularly the existence of protected area networks or other government lands not subject to economic drivers.

The comparatively low percentages of protected areas globally [38] coupled with significant remaining areas of natural land cover implicitly acknowledges future land use expansion. Along the range of ambitions for increasing protected area extent, from the 17% per nation goal of the Aichi targets to the half-Earth initiative, results here indicate the need for large, contiguous protected area planning. However, existing networks are more often small in areal extent and highly fragmented. While considerable areas of natural land cover may remain, even for highly developed climate/ecozones such as sub-tropical terra firma vegetated lowlands, the vast majority of natural land cover is in close proximity to intensive land uses.

Furthermore, while protected area expansion is a priority in maintaining key ecosystem services, protected area status is not guaranteed [39]. For example, a recent study on tropical forests quantified the

Figure 9. Log–log plot of land use extent and dispersion per ecozone per continent global climate/ecozones. Climate/ecozones of $\geq 1$ Mkm$^2$ in size underlined. Climate: arc = arctic, bor = boreal, tem = temperate, sub = sub-tropical, dry = dry tropical, hum = humid tropical; nam = North America, eas = Eurasia, sam = South America, afr = Africa, aus = Australia. Ecozones: low = terra firma vegetated land on low slopes (lowand), mon = vegetated land on high slopes (montane), wet = wetland. The concentric circles are proportional to the area of each continent climate/ecozone (outer circle) and the area of land use for that climate/ecozone (inner circle, in gray). The black line connects South American, African and Asian humid tropical terra firma vegetated lowlands.
Figure 10. Relationship between wetland extent and within-wetland land use for level three global watersheds [42]. Black circles represent area of 2019 wetland extent, and blue circles the portion of wetland impacted by treed land uses as measured from tree cover loss and gain since 2000. Only treed land uses are measured inside of wetlands, as built-up and cropland land uses are assumed to represent complete conversions of wetlands. We do not account for pasture land uses in this study, which may often be found in wetlands, but do not typically involve land cover conversion processes.

Figure 11. Sumatra, Indonesia, site of intensive land use within terra firma vegetated lowlands and wetlands.

seemingly inexorable fragmentation of humid tropical forests into ever smaller patches which are in turn more likely to be further fragmented, including forests in protected areas [40]. Montane environments are much less appropriated than lowland ones, and often serve as a kind of refugia for natural biota [41], with results of this study confirming lower land use pressure on sloped landforms. The dearth and continued exploitation of global wetlands is starkly evident, with the wetlands of Insular Southeast Asia
and the Southeast USA experiencing land change pressures unlike other wetlands globally (figure 10). Over 40% of global land use in wetlands occurs in three regions, the Southeast USA Atlantic Coast, and the islands of Sumatra (figure 11) and Borneo.

Overall, temperate and sub-tropical environments, home to advanced economies and high per capita wealth, have been more extensively appropriated for land use. Our analysis reveals land use development of terra firma vegetated lowlands within temperate and sub-tropical climate domains to have passed the critical area threshold to the point where natural lands exist as small, isolated patches embedded in nearly continuous land use. Figure 12 is an example of such a landscape in the Southeast USA, where natural lands persist in remnant wetlands and in high mountain retreats, and government facilities such as military installations account for a substantial portion of total stable land cover on terra firma vegetated lowlands. The advanced development of the temperate and sub-tropical domains poses a problem for policies that aim for lower latitude developing countries to sacrifice the same development opportunity in the cause of climate change mitigation [43].

Tropical deforestation, a significant contributor to anthropogenic forcing of climate warming, is the focus of international policies aimed at maintaining the high carbon stocks and high biodiversity of natural tropical forests. However, dry tropical and humid tropical terra firma vegetated lowland regions are the site of both the most extensive remaining natural terra firma vegetation (figure 13) and of the most rapid conversion of forests and woodlands to cropland land uses, with an eventual fate that may be very similar to temperate and sub-tropical terra firma lowland vegetated zones.

For this study, land cover is defined as the absence of built-up land, permanent cropland, and land use associated with tree cover loss and gain. We do not account for historical built-up, crop and treed land uses that have returned to a stable land cover or for low-intensity land uses (e.g. grazing or agroforestry) on mapped land covers. Land cover totals for 2019, disaggregated by three vegetation structure categories, are shown in figure 13 and indicate that the tropics are the future of land use expansion (discounting climate changes that may alter current vegetation distributions). Globally, nearly half of remaining vegetated land cover is found in the tropics, 63% if excluding high latitude boreal and arctic lands. Over half of remaining terra firma lowland tall, dense tree cover is located in the tropics, 74% if excluding the boreal and arctic. Nearly two-thirds of global terra firma lowland open/short tree cover is found in the tropics, 86% if excluding the boreal and arctic. Regarding emissions from land use change, loss of biodiversity due to habitat loss, and other ecosystem service impacts, the tropics clearly remain the climate domain facing the greatest near-term threats.

While pastures are the largest global land use in areal extent [44], they are not included in this study except for pastures on recently deforested lands. Pastures are highly variable in intensity, from planted pastures and associated intensive management schemes that enable dramatic increases in stocking rates [45], to the grazing activities of pastoralists who cover wide areas of natural grasslands in search of forage. The varying management regimes, stocking
rates, and other contexts in which pastures occur makes them challenging to map from space. For our study, regions of productive pasture that are not co-located with croplands will represent an overestimate of the distance of natural lands to land use. However, for many such regions, whether the northern Great Plains of the United States, northwest Uruguay, or northern Queensland, Australia, pasture land use is often practiced on natural grasslands. Our inclusion or exclusion of pastures can be considered from a land cover conversion perspective, and our decision was to focus on more intensive cropland, built-up, and treed land uses that by definition include the conversion of land cover. Intensively managed pastures in regions such as Europe and more recently Brazil, are typically co-located with our mapped land use themes. The result is a conservative, but largely consistent characterization of the footprint of intensive global land uses.

Only treed land uses were allowed in wetlands, as built-up or cropland land uses were assumed to have replaced wetland ecosystems, if co-located. However, the pressures from built-up and cropland land uses are evident in the study results. Figure 8 plots the pressures within climate/ecozones by individual land use, while figures 5 and 9 plot pressures by combined land uses. Wetlands in figures 5 and 9 exhibit shorter distances to land use compared to those in figure 8, as built-up and cropland land uses often form a boundary along remnant wetlands, reducing their overall mean distance to land use. Large wetland complexes such as the Everglades in the USA, the Parana Delta in Argentina, the Sundarbans on the India/Bangladesh border, or remaining coastal peatlands in Indonesia, are hemmed in by intensive land uses.

5. Conclusion

The slowing, much less reversing, of the seemingly inexorable expansion of land use remains a daunting challenge, principally in the domain of policy and governance. Land cover and land use extent and change data help to quantify the extent of the challenge, and can confirm the success or failure of policies designed to balance economic development and the continued provisioning of key ecosystem services, such as maintaining biodiversity, sequestering carbon, regulating climate, and sustaining hydrologic systems. Observational data on land use are less and less a limiting factor, given their increasing quality and abundance, while the lack of progress in policy implementation and enforcement is more and more the primary constraint. The layers presented here may be operationally deployed in assessing continued appropriation of natural lands for economic use, and may also provide evidence of efforts to restore ecosystems. Without successful policy interventions, regions where development is relatively favorable, but not yet realized, are likely to be converted. As favorable land becomes scarce, land that would not normally be developed may be converted through human engineering, for example wetland conversion,
or by degazetting protected areas. Finally, unmitigated climate change may play a role in modifying the development potential of existing natural lands. The method presented here is planned for time-series implementation and may serve as an input to assessments of policies designed to balance economic development with the maintenance of ecosystem services.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://glad.umd.edu/dataset/global-land-cover-land-use-v1.

Acknowledgments

This study was supported by the Government of Norway’s NICFI program, the Gordon and Betty Moore Foundation, the NASA Land Cover and Land Use Change program, the NASA/USGS Landsat Science Team, and National Geographic.

Author contributions

M C Hansen developed methods and wrote manuscript; P V Potapov developed and implemented ARD pre-processing and mapping methodologies; A Pickens implemented mapping and sampling methodologies; A Tyukavina designed sampling methodology; A Hernandez Serna implemented mapping methodologies, V Zalles performed reference data interpretations; S Turubanova performed reference data interpretations; I Komareddy developed tools for validation and data access; S V Stehman performed accuracy and area estimation analyses; X-P Song reviewed/confirmed results; and A Komareddy managed software and hardware tools for data processing.

ORCID iDs

Matthew C Hansen ○ https://orcid.org/0000-0003-0042-2767
Alexandra Tyukavina ○ https://orcid.org/0000-0003-3872-9844
Viviana Zalles ○ https://orcid.org/0000-0002-9532-4643
Xiao-Peng Song ○ https://orcid.org/0000-0002-5514-0321

References

[1] Vitousek P M, Mooney H A, Lubchenco J and Melillo J M 1997 Human domination of Earth’s ecosystems Science 277 494–9
[2] Sanderson E W, Jaithe M, Levy M A, Redford K H, Wannebo A V and Woolmer G 2002 The human footprint and the last of the wild the human footprint is a global map of human influence on the land surface, which suggests that human beings are stewards of nature, whether we like it or not BioScience 52 891–904
[3] Foley J A et al 2005 Global consequences of land use Science 309 570–4
[4] Turner B L, Lambin E F and Reenberg A 2007 The emergence of land change science for global environmental change and sustainability PNAS 104 20666–71
[5] Ellis E C 2011 Anthropogenic transformation of the terrestrial biosphere Phil. Trans. R. Soc. A 369 1010–35
[6] Lambin E F and Meyfroidt P 2011 Global land use change, economic globalization, and the looming land scarcity PNAS 108 3465–72
[7] Venter O et al 2016 Sixteen years of change in the global terrestrial human footprint and implications for biodiversity conservation Nat. Commun. 7 1–15
[8] Wulder M A, Masek J G, Cohen W B, Loveland T R and Woodcock C E 2012 Opening the archive: how free data has enabled the science and monitoring promise of Landsat Remote Sens. Environ. 122 2–10
[9] Stehman S V 2014 Estimating area and map accuracy for stratified random sampling when the strata are different from the map classes Int. J. Remote Sens. 35 4923–39
[10] Olofsson P, Foody G M, Herold M, Stehman S V, Woodcock C E and Wulder M A 2014 Good practices for estimating area and assessing accuracy of land change Remote Sens. Environ. 148 42–57
[11] Ying Q, Hansen M C, Potapov P V, Tyukavina A, Wang L, Stehman S V, Moore R and Hancher M 2017 Global bare ground gain from 2000 to 2012 using Landsat imagery Remote Sens. Environ. 194 161–76
[12] Potapov P et al 2021 Mapping global forest canopy height through integration of GEDI and Landsat data Remote Sens. Environ. 253 112165
[13] Pickens A H, Hansen M C, Hancher M, Stehman S V, Tyukavina A, Potapov P, Marroquin B and Shereani Z 2020 Mapping and sampling to characterize global inland water dynamics from 1999 to 2018 with full Landsat time-series Remote Sens. Environ. 243 11792
[14] Hansen M C et al 2013 High-resolution global maps of 21st-century forest cover change Science 342 850–3
[15] FAO 2012 Global Ecological Zones for FAO Forest Reporting: 2020 Update (Rome: FAO) (available at: www.fao.org/forest-resources-assessment/remote-sensing/global-ecological-zones-gez-mapping/en/)
[16] Boswell G P, Britton N F and Franks N R 1998 Habitat fragmentation, percolation theory and the conservation of a keystone species Proc. R. Soc. B 265 1921–5
[17] Walsh K A 2002 Using percolation theory to assess landscape connectivity and effects of habitat fragmentation Applying Landscape Ecology in Biological Conservation (New York: Springer) pp 105–30
[18] Scher H and Zallen R 1970 Critical density in percolation theory J. Chem. Phys. 53 3759–61
[19] Kasraian D, Maat K, Stead D and van Wee B 2016 Long-term impacts of transport infrastructure networks on land-use change: an international review of empirical studies Transp. Rev. 36 772–92
[20] DeFries R, Hansen M and Townshend J 1995 Global discrimination of land cover types from metrics derived from AVHRR pathfinder data Remote Sens. Environ. 54 209–22
[21] Hansen M C, DeFries R S, Townshend J R G, Carroll M, Dimiceli C and Sohldberg R A 2003 Global percent tree cover at a spatial resolution of 500 meters: first results of the MODIS vegetation continuous fields algorithm Earth Interact. 7 1–15
[22] Potapov P, Hansen M C, Komareddy I, Komareddy A, Turubanova S, Pickens A, Adusei B, Tyukavina A and Ying Q 2020 Landsat analysis ready data for global land cover and land change mapping Remote Sens. 12 426
[23] Potapov P, Turubanova S, Hansen M C, Tyukavina A, Zalles V, Khan A, Song X, Pickens A, Shen Q and Cortez J 2022 Global maps of cropland extent and change show
accelerated cropland expansion in the twenty-first century

[24] Bwangoy J-R B, Hansen M C, Roy D P, de Grandi G and Justice C O 2010 Wetland mapping in the Congo Basin using optical and radar remotely sensed data and derived topographical indices Remote Sens. Environ. 114 73–86

[25] Margomo B A, Bwangoy J-R B, Potapov P V and Hansen M C 2014 Mapping wetlands in Indonesia using Landsat and PALSAR data-sets and derived topographical indices Geo-Spatial Inf. Sci. 17 60–71

[26] Breiman L, Friedman J H, Olsens R A and Stone C J 1984 Classification and Regression Trees (Wadsworth Statistics/Probability)

[27] Hansen M C et al 2014 Monitoring conterminous United States (CONUS) land cover change with web-enabled Landsat data (WELD) Remote Sens. Environ. 140 466–84

[28] Dubayah R et al 2020 The global ecosystem dynamics investigation: high-resolution laser ranging of the Earth's forests and topography Sci. Remote Sens. 1 100002

[29] Brandt M et al 2020 An unexpectedly large count of trees in the West African Sahara and Sahel Nature 587 78–82

[30] Feng W, Sui H, Huang W, Xu C and An K 2018 Water body extraction from very high-resolution remote sensing imagery using deep U-Net and a superpixel-based conditional random field model IEEE Geosci. Remote Sens. Lett. 16 618–22

[31] CGIAR-CSI S 90m Digital Elevation Database v4.1

[32] Rizzioli P et al 2017 Generation and performance assessment of the global TanDEM-X digital elevation model ISPRS J. Photogr. Remote Sens. 132 119–39

[33] Tachikawa T et al 2011 ASTER Global Digital Elevation Model Version 3—summary of validation results (NASA) p 27 (Available at: http://pubs.er.usgs.gov/publication/70003960) (Accessed 16 November 2021)

[34] Rasool S I 1992 Requirements for terrestrial biospheric data for IGBP Core Projects. IGBP-DIS working paper, # 2 International Geosphere Biosphere Programme-Data and Information Systems (Paris, France)

[35] Turubanova S, Potapov P V, Tuukivaina A and Hansen M C 2018 Ongoing primary forest loss in Brazil, Democratic Republic of the Congo, and Indonesia Environ. Res. Lett. 13 074028

[36] Parfenova E, Tchebakova N and Soja A 2019 Assessing landscape potential for human sustainability and ‘attractiveness’ across Asian Russia in a warmer 21st century Environ. Res. Lett. 14 065004

[37] Finlayson M et al 2005 Millennium Ecosystem Assessment: Ecosystems and Human Well-Being: Wetlands and Water Synthesis (Washington, DC: Island Press) (Available at: https://library.wur.nl/WebQuery/wurpubs/344031) (Accessed 16 November 2021)

[38] Jufve-Bignoli D, Bingham H, MacSharry B, Deguignet M, Milam A and Kingston N 2015 World Database on Protected Areas—User Manual 1.4 (United Nations Environment Programme-World Conservation Monitoring Centre) p 219

[39] Mascia M B and Pailler S 2011 Protected area downgrading, downsizing, and degazettement (PADDD) and its conservation implications Conserv. Lett. 4 9–20

[40] Hansen M C, Wang L, Song X-P, Tuukivaina A, Turubanova S, Potapov P V and Stehman S V 2020 The fate of tropical forest fragments Sci. Adv. 6 eaax8574

[41] Sandel B and Svenning J-C 2013 Human impacts drive a global topographic signature in tree cover Nat. Commun. 4 2474

[42] Lehner B and Grill G 2013 Global river hydrography and network routing: baseline data and new approaches to study the world’s large river systems Hydro. Process. 27 2171–86

[43] Tanentzap A J 2017 The costs of saving nature: does it make ‘cents’? PLoS Biol. 15 e2003292

[44] Ramankutty N, Evan A T, Monfreda C and Foley J A 2008 Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000 Glob. Biogeochem. Cycles 22

[45] Barretto A G O P, Berndes G, Sparovek G and Wirsenius S 2013 Agricultural intensification in Brazil and its effects on land-use patterns: an analysis of the 1975–2006 period Glob. Change Biol. 19 1804–15

[46] Pickens A H, Global wetland extent from satellite data In prep.

[47] Hernandez-Serna A Global increase in built-up land, 2000 to 2020 In prep.