

scientific data



OPEN

DATA DESCRIPTOR

Consistent global dataset on biodiversity intactness footprint of agricultural production from 2000 to 2020

Can Trong Nguyen[✉], Davina Vačkářová & Jan Weinzettel

Global biodiversity is rapidly declining, primarily due to agricultural production driven by both domestic and transboundary consumption. This study addresses the challenges posed by inconsistent spatiotemporal biodiversity data by developing a time series of biodiversity loss footprints based on Biodiversity Intactness Index (BII). Numerous land use, land cover, and auxiliary datasets were integrated to produce a consistent time series of high-resolution harmonized land use (HHLU) maps. These maps were utilized to quantify spatial BII using linear-mixed effect models. Biodiversity intactness loss (BII footprint) was subsequently attributed to specific crops and livestock commodities. This study provides comprehensive global datasets, including HHLU and BII maps, and synthesized BII footprints across 14 biomes, 193 countries and territories, 154 crop items, and 9 livestock categories from 2000 to 2020. These datasets facilitate spatiotemporal analyses to identify trends and patterns in global biodiversity integrity and biodiversity footprints, thereby elucidating the ecological trade-offs embedded in international trade. These insights can encourage appropriate interventions to transform consumption patterns and supply chains toward the effective conservation of global biodiversity.

Background & Summary

Despite the multiple ways humans have become disconnected from the rest of the natural world¹, people remain inherently dependent on ecosystems and the services they provide². However, human behavior continues to escalate global declines in biodiversity and ecosystem services. Recent studies consistently document systemic declines in global biodiversity with reference to multiple domains and indicators in the past decades³. These data are based on various baselines, including temporal declines in species populations⁴, shifts in the original composition and diversity of terrestrial assemblages⁵, and changes in species' threat status⁶. Without effective political action and integrated strategies, these negative trends will continue in the future^{7,8}. The widespread decline in biodiversity contributes to the homogenization of the biosphere and the transgression of planetary boundaries for biosphere integrity. This transgression is strongly related to other planetary boundaries and may trigger irreversible changes that jeopardize human well-being^{9,10}.

Agricultural production is among the leading drivers of global biodiversity loss, primarily due to the expansion of croplands and grazing areas, as well as the intensification of farming practices to satisfy the increasing consumption demands of the growing and more affluent global population^{11,12}. Agricultural activities occupy and convert natural habitats into intensively managed landscapes, reducing native biodiversity and disrupting essential ecosystem service flows^{13,14}. Notably, increasing affluence has emerged as a significant distant driver of biodiversity loss, often geographically remote from the point of consumption^{15,16}. Rising consumption has been identified as a dominant factor driving habitat conversion, climate change, and other indirect ecological pressures, as societies strive to meet the material demands of their increasingly affluent population¹⁷. Global telecouplings and displacements, which propagate hidden impacts and spillovers of product consumption on biodiversity, point to the vanishing connections between consumers and local ecosystems¹⁸. To address these hidden dynamics, biodiversity footprint analysis has emerged as a promising framework to account for the complex ecological consequences of supply chains^{19–23}. This approach captures both direct and indirect biodiversity

Charles University, Environment Centre, José Martího 407/2, Prague, 16200, Czech Republic. [✉]e-mail: can.nguyen@czp.cuni.cz

losses throughout international trade and supply chains, helping to assign demand-side responsibility remote biodiversity loss²³. Biodiversity footprint accounting aims to link existing biodiversity datasets and models on biodiversity with “footprinting” methods, such as Environmentally extended multi-regional input-output (EE-MRIO) analysis and Life Cycle Assessment (LCA). However, obtaining consistent and high-resolution time series data on the contribution or footprint of economic sectors is critical to the effectiveness of these approaches. This is because long-term assessments require harmonized spatiotemporal datasets to accurately capture trends and support the development of evidence-based decision-making.

Consistent biodiversity footprint data and modeling approaches combine biodiversity metrics with multiple anthropogenic pressures, enabling the development of spatially explicit models of biodiversity response. The development of spatially explicit modeling approaches to biodiversity footprints have been enabled by existing land-use change datasets. Chaudhary *et al.*¹² developed biodiversity characterization factors based on the countryside species-area relationship (SAR), which allowed for spatially explicit estimation of biodiversity loss from agriculture, pasture, and forest land use. Marques *et al.*¹⁵ applied this approach to quantify the biodiversity footprint of potential global bird species extinctions driven by changing land-use. Similarly, Wilting *et al.*²⁴ advanced the GLOBIO modeling framework utilizing the mean species abundance (MSA) as a function of multiple anthropogenic pressures. Boakes *et al.*²⁵ further expanded the biodiversity footprint concept by assessing both production and consumption footprints of food-related commodities, focusing on land-driven and greenhouse gas-driven species richness and rarity-weighted species richness.

Among the biodiversity metrics, Biodiversity Intactness Index (BII) stands out as a global leader in biodiversity indicators, enabling the quantification of biodiversity loss which informs our understanding of biosphere planetary boundaries^{5,26}. BII estimates the average abundance of organisms in a certain area relative to an undisturbed reference baseline to yield an improved spatiotemporal biodiversity indicator compared to many other indices²⁶. The current version of BII integrates statistical models of overall organismal abundance and compositional similarity to generate a picture of a minimally impacted (intact) assemblage²⁷. Despite the limitations in its estimation and applications, BII is generally straightforward, practical, and sensitive to various driving factors and artificial interventions^{28,29}. It has been widely applied in the assessment of biodiversity integrity in diverse contexts, including assessments of grasslands affected by fire and grazing in Africa and Amazonia³⁰, biodiversity response to agricultural intensification^{31,32}, interactions between private conservation and natural systems³³, forest biodiversity monitoring^{27,34}, soil biodiversity assessment³⁵, and global biodiversity footprint analysis^{36,37}. However, the BII model is based on the Land-Use Harmonization (LUH) framework, which classifies land cover by vegetation-based states such as primary and secondary vegetation rather than land use types observed in other land use, land cover datasets^{5,27,38}. While this categorization facilitates the estimations of BII, original LUH datasets are derived from coarser-resolution models rather than direct observations, limiting their accuracy for finer-scale and national analyses. Although some efforts have been made to overcome this limitation, a globally consistent and spatially detailed dataset based on standardized methods remains essential for robust biodiversity assessment across multiple scales^{27,39}. Most current BII applications are constrained to single time points or relatively short-term intervals, largely due to the lack of reliable and consistent datasets.

This study therefore seeks to develop consistent datasets of harmonized land use and BII to derive the BII loss footprint associated with agricultural activities (i.e., crop and livestock production) between 2000 and 2020. We generate a consistent global time series of land use and BII by integrating statistical models with land use data and other auxiliary datasets. This work provides high-resolution harmonized land use (HHLU) maps to accompany land use fraction datasets, aligned with the LUH classification system. We ultimately allocate BII loss footprints to both crop and livestock production, synthesizing the footprints across major biomes, countries and territories, and production sectors annually. The dataset of agricultural biodiversity loss footprints is an initial step toward quantifying local biodiversity impacts of agricultural production at national and territorial scales^{40,41}. This information can serve as a critical foundation for integrating biodiversity footprint analysis into international trade, helping to reveal the hidden ecological costs embedded in global supply chains, clarify demand-side responsibility, and inform more sustainable production and consumption strategies.

Methods

This section outlines the required datasets (Table 1) and primary methodological steps used to generate globally consistent datasets of harmonized land use and Biodiversity Intactness Index (BII) maps. Land-use harmonization integrates numerous data sources to produce temporally and spatially consistent datasets across the study period. BII maps are obtained through a combination of statistical modeling and spatial prediction techniques. Subsequently, allocation approaches are applied to attribute biodiversity loss footprints to specific crop and livestock production activities. An overview of the entire workflow is intuitively illustrated in Fig. 1, including all data inputs, modeling steps, and footprint allocation processes.

Datasets and processing procedures. *HILDA + global land use change dataset.* Historic Land Dynamics Assessment (HILDA+, ~0.01°) is the primary land use dataset used to define habitat environments in biodiversity modeling. The dataset consistently covers a long-term period (1960–2019) at ~1 km spatial resolution annually. It is constructed based on a data-driven reconstruction approach to provide six land use/cover categories corresponding to the definitions of FAO (Food and Agriculture Organization), including forest, grassland, pasture, cropland, urban areas, and sparse/no vegetation (Fig. 2-A and Table S1). It harmonizes remote sensing, historical land use reconstruction, and statistics to generate authentic land use maps rather than other land cover data sources that are delineated solely from remote sensing imagery⁴². Land use reflects how humans utilize and manage land, making it more directly relevant than land cover for modeling biodiversity responses to land use changes and anthropogenic pressures. Moreover, HILDA+ is among the first to distinguish pastureland used for livestock grazing from general grassland areas, which is particularly important for our analyses of biodiversity

Source	Category	Years	Data type	Pixel size
HILDA+ land use ⁴²	Land use	2000–2019	Raster image	1 km
MODIS MCD12Q1 land cover ⁴⁴	Land use	2001–2020	Raster image	500 m
Global human pasture ⁴⁵	Land use	2000	Raster image	~1 km
Global Pasture Watch (GPW) ⁴⁶	Land use	2000–2022	Raster image	30 m
Forest Management Data ⁴⁷	Land use	2015	Raster image	100 m
Intact forests landscapes (IFL) ⁴⁸	Land use	2000, 2013, 2016, 2020	Shapefile, polygons	—
Global terrestrial human footprint ⁴⁹	Human impact	2000–2022	Raster image	1 km
Population density ⁵⁴	Human impact	2000–2020	Raster image	~1 km
ESRI global road network	Human impact	—	Shapefile, lines	—
Accessibility to Cities ⁵⁷	Human impact	2015	Raster image	~1 km
PREDICTS database ^{51,52,76}	Biodiversity	1984–2018	Data Records	—
Spatial Production Allocation Model (SPAM) ⁵⁸	Agricultural production	2005, 2010, 2020	Raster image	~10 km
Gridded Livestock of the World (GLW) ⁶²	Agricultural production	2010, 2015, 2020	Raster image	~10 km

Table 1. Datasets used to develop biodiversity intactness maps and biodiversity footprints.

loss associated with livestock production. The relatively fine resolution of HILDA+ also allows a more precise detection of changes in local land use trends and more accurate assessments of how these changes are connected to biodiversity loss and associated impacts⁴³.

MODIS land cover product. MCD12Q1 version 6.1 is an annual land cover type data product derived from MODIS (Moderate Resolution Imaging Spectroradiometer) sensors on board the AQUA and TERRA satellites. It uses a supervised decision tree classification to generate global land cover maps at approximately 500-meter spatial resolution, covering the years 2001 to 2023. It provides land cover data under five different classification schemes, such as the International Geosphere-Biosphere Programme (IGBP), University of Maryland (UMD), leaf area index, BIOME-biogeochemical Cycles (BGC) and plant functional types⁴⁴. We derived the layer of land cover from the University of Maryland (UMD) classification scheme including sixteen (16) land cover categories (Table S2) to complement HILDA+ and other land use data sources, helping to refine broad land use categories and adjust for potential misclassifications.

Other auxiliary data on land use. In addition to HILDA+ and MCD12Q1, we incorporated multiple land use-related datasets to enhance the accuracy of land use maps and to better translate them into representations of habitats. These datasets include human pasture layer and cultivated grassland data from the Global Pasture Watch (GPW) dataset, which were used to correct and refine the pasture class in the HILDA+ dataset by removing natural grassland areas^{45,46}. Although the generic forest class from HILDA+ can be subdivided into different forest types, these distinctions are mainly based on vegetation characteristics (i.e., evergreen versus deciduous) rather than on forest management regimes that are more relevant for biodiversity assessment. Incorporating forest management data enhances differentiation to distinguish plantation forests, oil palm plantations, and agroforestry systems from natural and semi-natural regenerating forests⁴⁷.

To convert land use maps into habitat-focused land use maps, we further incorporated the land use maps with two critical datasets of intact forest landscape and terrestrial human footprint data. Intact forest landscape (IFL) data was adopted to identify primary forest landscape among generic forest pixels, as IFL represents unfragmented forest patches of at least 500 m² that are free from significant human disturbance⁴⁸. In addition, human footprint data combines eight human pressure variables including built environment, population, night-time lights, crop and pasturelands, roads and railways, and navigable waterways to capture cumulative human pressures on ecosystems and natural landscapes. It presents human impacts at different levels ranging from near wilderness/intactness to high modification, using a scale from 0 (intact) to 50 (most modified)^{49,50}. Therefore, human footprint data allow us to classify habitat landscapes based on the intensity of human modification. Land use with wilderness and intact status is categorized as minimally used primary and primary vegetation habitats, while more heavily modified areas are classified as secondary vegetation and other highly modified landscapes.

PREDICTS Biodiversity database. PREDICTS (Projecting Responses of Ecological Diversity in Changing Terrestrial Systems) is a global biodiversity dataset collected from individual studies worldwide. It describes information at site-level species abundance for diverse taxa (i.e., plants, fungi, invertebrates, and vertebrates), habitat characteristics, and supplementary information⁵². It describes the site-level information. The early version (v1.1) released in 2016 included 480 studies across 94 countries from 1984 to 2013⁵¹. It recorded over 3,278,056 measurements from 26,194 sampling locations with 47,089 species. The later version (v2022) adds about 1,040,752 measurements from 115 studies at 9,544 sampling locations in 46 countries (10,635 species)⁵². This update expands the data coverage to 101 countries worldwide up to the year 2018, with most studies spanning 2000 to 2010. The current data version is distributed across a wide range of landscapes and land-use habitats, providing broad ecological and geographical representation (Figure S1).

The two versions were combined and preprocessed using a series of data quality control steps, such as data filtering, data transformation, and estimation. More explicitly, the dominant land use on each site was refined to

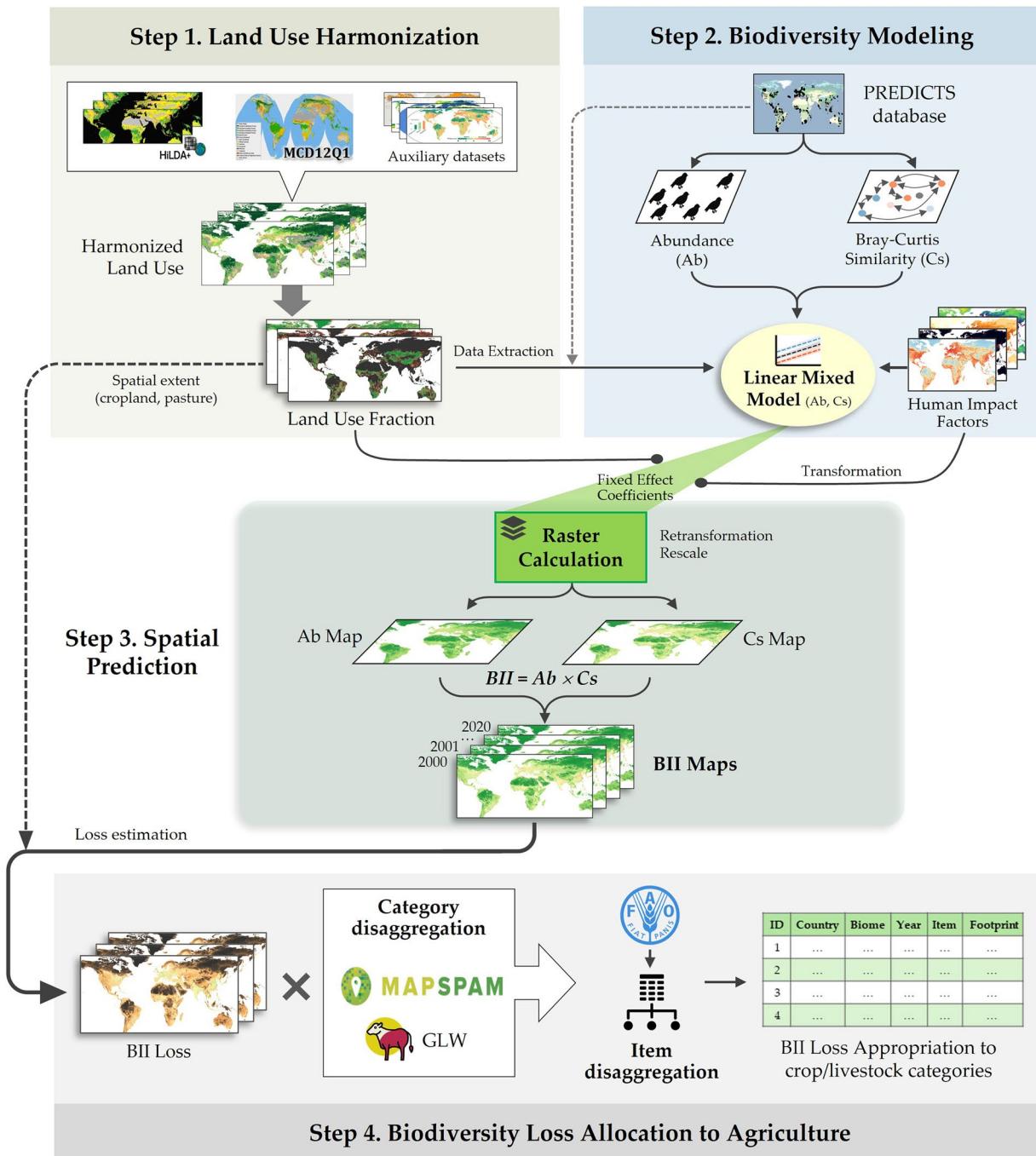


Fig. 1 Overview of the methodological framework used to generate biodiversity intactness maps and biodiversity loss footprints associated with agricultural production.

the same classification scheme as our harmonized land use data, while a site with unclear (“cannot decide”) land use was eliminated from further analyses. Biodiversity metrics were retrieved from the PREDICTS database to estimate BII including the total abundance of organisms regardless of taxonomic groups at each location and the compositional similarity of each location relative to other sites. Sampling efforts may be different across studies and sites. Therefore, total abundance values were normalized by sampling efforts and rescaled to a 0–1 range to make them comparable across studies and sites²⁷. This range reflects the intactness of species abundance compared to a natural baseline set at 100%. Compositional similarity of a site location to human impacts reflects the community structure of the considered site in terms of occurrence and quantity of species compared to the baseline sites of primary minimal-used habitats using a balanced Bray-Curtis index^{26,27,39}. Subsequently, both the rescaled total abundance and Bray-Curtis index datasets need to pass data quality checks to eliminate outliers before being used for biodiversity modeling.

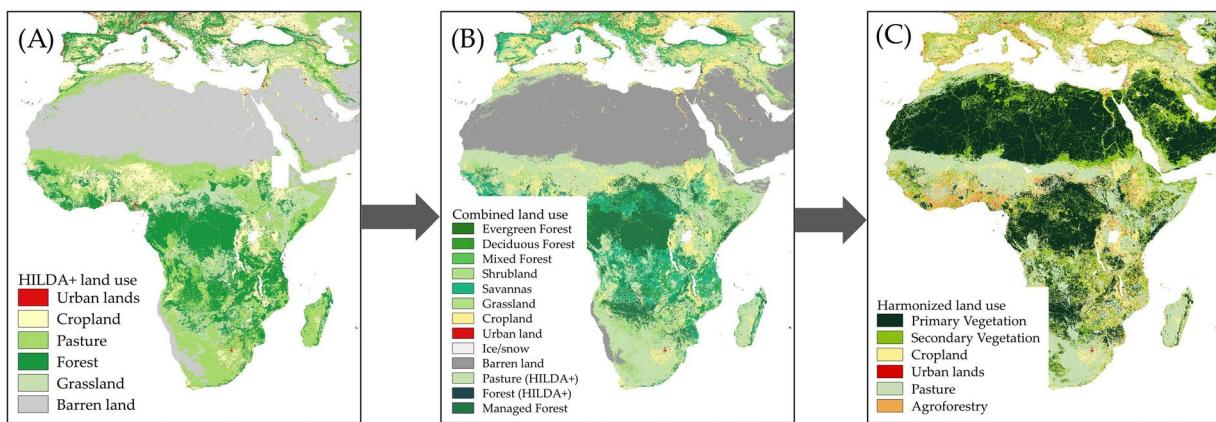


Fig. 2 Example of land use harmonization processes in Africa. (A) original HILDA+ land use includes six broad classes, (B) combined land use/cover with MCD12Q1 and other land use datasets to refine vegetation types, (C) output of harmonized land use map with habitat states for biodiversity integrity assessment.

Data on human-induced impacts. Beyond land-use fraction derived from land-use maps, a range of other anthropogenic factors also play a significant role in regulating biodiversity intactness. Human population density is widely considered one of the primary agents that significantly correlates with biodiversity loss, as population growth and human interventions lead to resource exploitation, habitat disturbance and fragmentation, and pollution⁵³. We acquired population density data from WorldPop (~1 km resolution) for each year from 2000 to 2020, which applies a random forest algorithm based on census and land cover data to estimate and distribute population on grid cells⁵⁴.

Transportation and traffic projects fragment habitats and ecological systems⁵⁵. Obtaining a balanced global road network for sustained time period is challenging as some countries have detailed data availability while others have relatively limited data sources⁵⁶. A dense road network can cause overestimation and vice versa, while currently available data frequently lacks hierarchical information to eliminate local roads and pathways. Esri World Roads provides relatively homogeneous data, providing details of highways and major roads suitable for human impact analysis. The vector map was transformed into raster data by estimating the proximity to roads using Euclidean distance function in ArcGIS Pro.

Another factor is accessibility, represented by travel time to nearby urban centers using various modes of transportation while considering land cover and topography⁵⁷. Natural characteristics and topographical patterns have to some extent been proven to control ecological dynamics⁵⁵. For instance, flat plain ecosystems are often more disturbed due to high population density and agricultural and industrial activities. Therefore, accessibility data is relevant when considering both natural and anthropogenic drivers of biodiversity intactness. Since there is currently no annual data on roads and accessibility, this data was treated as time-invariant driver.

Crop and livestock distribution datasets. Spatial distribution datasets for crops and livestock provide essential information for attributing biodiversity loss to specific crop types and livestock categories. BII loss was attributed to crop and livestock groups based on their proportion of total crop/livestock at a given location (or grid cell). This study employed SPAM (Spatial Production Allocation Model) to represent crop distributions and GLW (Gridded Livestock of the World) to represent livestock distribution.

SPAM is a global crop distribution dataset used to attribute the impact of individual crops on biodiversity loss. SPAM applies both spatial and non-spatial inputs to estimate the spatial distribution of the physical area based on land suitability criteria for each specific crop⁵⁸. Currently, there are three available SPAM datasets from 2005, 2010, and 2020 with relatively uniform crop systems (i.e., 42 crop types in 2005/2010 and 46 crop types in 2020) and spatial resolution (~10 km)^{59–61}. However, our annual estimates of biodiversity loss footprint require continuous time series data rather than discrete snapshots to enable a robust temporal analysis. Therefore, we extended the SPAM framework by incorporating SPAM as reference data and FAOSTAT statistics to generate annual global SPAM crop distribution maps for 2000–2020, using suitability probability while ensuring national-level consistency and pixel-scale constraints (see Supplementary Information 1 for full method).

GLW is a spatial dataset of the global distribution of eight major livestock categories (i.e., buffalo, cattle, chicken, duck, horse, goat, pig, and sheep)⁶². The current versions cover 2010, 2015, and 2020 at a spatial resolution of approximately ~10 km. Our study considered five of the eight livestock types estimated from the dasymetric allocation (DA) method, including buffalo, cattle, horses, goats, and sheep. These specific animals were chosen because their production systems are directly associated with pasture and grazing land use. Other livestock types are predominantly raised in confined or mixed systems, and although their production may significantly impact biodiversity through feed production and related pollution (e.g., nutrient runoff, ammonia emission, and waste-related contaminants), they do not directly contribute to pasture conversion. Therefore, excluding them ensures that our analysis focuses solely on the direct impacts of livestock production on land use and biodiversity rather than indirect effects. Annual datasets for each livestock type were also generated to support time series analysis. We applied a spatial allocation framework that corresponds to the SPAM crops

Land use category	Definition and Description
Primary vegetation	All vegetation and other landscapes with no or low human impact and disturbance or clearing, including natural and natural disturbance vegetation (e.g., natural forests, shrublands, grasslands, savannas, scrubs, and tundra), desert, barren, rock and ice.
Secondary vegetation	Regrowing/regenerating vegetation landscapes and managed forests, implying medium human intervention following disturbance or clearing (e.g., logging and shifting cultivation), such as naturally regenerated forests and vegetation landscapes, production forests, planted forests, plantation forests, urban green spaces, and urban forests.
Pasture	Managed pasture, grassland, and meadows (covered mainly by herbaceous plants) used for livestock grazing and hay production, which includes steppes, savannah, and mosaic landscapes.
Agroforestry	Woody crops including tree/shrub crops that do not belong to typical herbaceous croplands. It includes fruit/orchards and plantations (e.g., oil palm, coconut, and coffee plantations).
Cropland	Herbaceous crops, mosaic crops, and paddy fields, such as annual grain crops (e.g., wheat, rice, maize), and vegetables.
Urban areas	Artificial surfaces, including urban and built-up areas, roads, industrial sites, and urban infrastructures and utilities.

Table 2. Definition and description of harmonized land use.

allocation model based on data already available from GLW and FAOSTAT to produce each yearly GLW raster layer. This approach assumes that areas with higher livestock density indicate more suitable environments for their presence, allowing for a spatially explicit estimation of livestock distribution over time. In livestock production, the production of meat and dairy production may differently affect the environment and biodiversity. Some multi-regional input/output datasets (e.g., EXIOBASE) distinguish between these two categories in their estimations. To align with this distinction and enhance the granularity of our analysis, we further disaggregated each livestock type (except horses) into meat and dairy livestock distribution maps (see Supplementary Information 2 for full method).

For each year, a shared proportion of each crop or livestock type at a given location was calculated as its share within the total area or total number of livestock heads at that pixel (Eq. 1). By this definition, the sum of shared proportions across all crops or livestock types at any given pixel equals one (1).

$$Prop_{i,j} = \frac{a_{i,j}}{A_j} \quad (1)$$

where $Prop_{i,j}$ is the shared proportion of crop or livestock i at pixel j , $a_{i,j}$ is the area for crops or head count for livestock i at pixel j , and A_j is the total crop area or livestock head count at pixel j .

Land use harmonization. Land use is a critical driver of environmental changes and biodiversity loss, reflecting human interventions that modifies and disrupts natural ecosystems for production, agricultural cultivation, and infrastructure purposes^{5,63}. However, biodiversity integrity is typically assessed using a classification system that emphasizes habitat states and management regimes rather than conventional land use and land cover categories. This conceptual framework was conceptualized by Land-Use Harmonization (LUH), which integrates historical and future projections of land-use to generate land-use fraction maps from 850 to 2100³⁸. However, a relatively coarse resolution (~0.25 arc-degree), which is challenging for finer and small-scale biodiversity assessments attempting to capture local trends adequately⁴³. Therefore, we combined multiple available land use, land cover and auxiliary datasets to generate high-resolution harmonized land use datasets (HHLU), with a classification system aligned to LUH categories, at a global scale for biodiversity integrity assessment (Table 2). An overview of the methodological workflow is provided in Figure S3 with detailed explanations presented below.

A data quality assessment comparing HILDA+ data with other available land use datasets revealed notable inconsistencies and misclassifications within the HILDA+ data. For example, forested areas in Taiwan and natural grassland areas in the northern latitudes were misclassified as grassland and pasture areas in HILDA+, respectively (Figure S2). These discrepancies highlight the need to integrate additional datasets to reduce such uncertainties and improve the classification accuracy of HILDA+. To address this, we integrated HILDA+ with land cover layer from the University of Maryland's classification scheme from the MCD12Q1 product (land cover layer 2, ~500 m) to adjust and specify forest class from HILDA+ at a finer resolution (Fig. 2-B).

Prior to the adjustment and harmonization processes, all data layers were reprojected and resampled to align with MCD12Q1 data (WGS84 lat/lon geographic coordinates, ~500 meters), ensuring consistency and preserving spatial detail from finer-resolution layers such as MCD12Q1 land cover, forest management, and global pasture watch datasets. For continuous value data, the average value will be taken, and for discrete classes, the most frequent value (mode) or majority-pixel will be used in the upscaling process. During the downscaling process, resampling was conducted to ensure the retention of the values of finer component pixels within each coarser pixel. Conversely, during the upscaling process, the average value was considered for continuous data, while the mode (i.e., that most frequently occurring within a targeted coarse pixel) was applied for categorical (discrete) data to preserve the most dominant class.

Following these procedures, cultivated grassland was extracted from the GPW dataset and resampled to match the projection and resolution of HILDA+ data. These annual cultivated grassland layers and the human pasture layer were then adopted to adjust the pasture class in HILDA+ for each year by removing areas classified as pasture but lying outside the limits of cultivated grassland and human pasture layers⁴⁵. Specifically, pasture areas in HILDA+ that fell outside these boundaries were reclassified as natural grassland rather than pasture used for grazing purposes. The adjusted HILDA+ layers were resampled to 500 meters to preserve the spatial

detail after harmonization with the MODIS land use layers. While most land use classes and information were directly derived from HILDA+ and retained their original classes (e.g., cropland, urban, and barren land), the forest class was further integrated with MCD12Q1 to correct possible misclassification within the forest category and specify forest types (Fig. 2-B). Both HILDA+ and MCD12Q1 datasets only capture forest classes but do not distinguish between different forest management regimes. However, different forest management types (e.g., wild forest, production forest, agroforestry, and orchards) have divergent impacts on biodiversity. To capture these differences, the refined forest classes were further subdivided into various forest classes and managed forests (agroforestry) by integrating global management forest data⁴⁷. The combined land use is relatively detailed, reflecting critical land use and land cover categories (Fig. 2-B).

Subsequently, a second procedure was applied to the combined land use layers to delineate vegetation and habitat states such as primary and secondary vegetation. From a biodiversity intactness perspective, land use and management regimes are more critical than land use categories alone, as they indicate land use intensity rather than just land use purposes. Therefore, the combined land use data was reclassified at a second level, separating forest and grassland into primary vegetation (naturally regenerating vegetation without human impacts) and secondary vegetation (naturally regenerating vegetation with human activities, e.g., logging and clear cuts). Two datasets were utilized to distinguish these habitat states, including intact forest landscapes (<https://intact-forests.org>) and annual terrestrial human footprint^{48,49}. Areas identified as intact forest landscapes were directly assigned to the primary vegetation class. Forests and other vegetation in wilderness and intact states in the terrestrial human footprint were also assigned as primary vegetation, while the remaining vegetation was classified as secondary. Ultimately, we produced annual harmonized land use (level 2, HHLU) at 500-meter resolution encompassing six land use types (i.e., primary vegetation, secondary vegetation, cropland, pasture, agroforestry, and urban areas) (Table 2 and Fig. 2-C).

The level 2 HHLU data was then aggregated to generate land use fraction maps at a coarser ~2.5 km resolution. This was done by applying a 5×5 pixel moving window, where for each window, the number of pixels belonging to each land use category was counted and then divided by the total number of pixels within this window (25 pixels). The resulting land-use fraction values represent the proportional coverage of each land-use type within a larger pixel. This process was repeated across the entire dataset to produce continuous raster maps where the fraction of all land use categories sums to 1, which can capture mixed land uses at a coarser scale for biodiversity modeling.

Biodiversity modeling. Linear mixed-effect (LMM) models were employed to characterize species diversity by comparing different sampling methods between distinguishing studies and sampling blocks—with studies and sampling blocks as random effects^{26,27,39}. Two LMMs were constructed to represent biodiversity components of BII through diversity abundance and compositional similarity. These models were constructed using estimated parameters from the PREDICTS database, along with land use and human impact variables drawn from the data pool. Prior to biodiversity modeling, land use fractions for each land use category and human impact variables were extracted at sampling block locations provided in the PREDICTS database (Figure S1). Rescaled abundance, compositional similarity (Bray-Curtis index), land use fractions, and human impact variables were first transformed to approximate normality prior to applying linear-mixed models. Rescaled abundance and human impact variables were transformed using the natural logarithm transformation with added constant [$\ln(x + 1)$] to allow the inclusion of zero in the analysis. Compositional similarity was transformed using the logit transformation function from the *car* R-library, with a small adjustment of 0.01 to constrain values within the open interval [0, 1]⁶⁴.

The transformed abundance and compositional similarity were used as dependent variables in two separate linear-mixed effect models, with land use fractions and all other human impact variables included as independent predictors. Initially, a broader set of predictors was considered, including nighttime light (NTL) and proximity to cities. However, only the human impact variables listed in the data used section made significant contributions to at least one of the LMM models. The LMM models were fitted using a maximum likelihood estimator, implemented via the *lme4* R-library⁶⁵. The land use fraction from all land use categories and human impact variables were considered as fixed effects. Random effects were specified across sampling blocks to account for spatial and study-level variability. Variable selection was done by stepwise elimination using the *lmerTest* R-library⁶⁶. The final models were evaluated based on two indicators of Variance Inflation Factor (VIF) and the coefficient of determination (R^2). VIF assists in controlling free multicollinearity, with all predictors not exceeding $VIF < 4.0$, while R^2 reveals the power of fitted models. The fitted models of Abundance and Compositional Similarity achieved R^2 of 0.521 and 0.774, respectively. The selected predictors and their coefficients ($p < 0.001$) for the transformed abundance and compositional similarity models are presented in Table 3.

Spatial prediction of BII maps. The fitted global models were applied to generate annual spatial BII maps from 2000 to 2020. These models were used to predict spatially the two component maps of transformed abundance and compositional similarity. This required the raster layers of the predictors (except for land use fractions) to be transformed using a natural logarithm, consistent with the transformations applied to point-based data. Human impact raster layers were upscaled to match the land use fraction layers using average aggregation. Land use fractions and population layers were treated as dynamic variables for individual years, while proximity to road and travel time were considered static variables due to the unavailability of annual data.

The intermediate maps of transformed abundance and compositional similarity were subsequently back-transformed to original components using the inverse of natural logarithm and logit functions, respectively. Resulting maps of abundance (Ab) and compositional similarity (Cs) were normalized to reduce model uncertainty and rescale to a standard 0–1 range, ensuring that maximum values correspond to baseline intact habitats with minimal human disturbance. Since the observation period (2000–2020) covers only the last two decades—well after the Industrial Revolution, which marked the start of rapid habitat destruction—and because

Predictors	Coefficient	Standard Error	Sig. level	VIF
Abundance model				
(Intercept)	3.252E-01	1.164E-02	< 2E-16***	
Primary vegetation	3.622E-02	5.493E-03	4.41E-11 ***	1.033
Croplands	-2.534E-02	4.928E-03	2.77E-07***	1.041
Urban lands	-6.276E-02	7.838E-03	1.25E-15***	1.074
Log-Proximity to roads	-3.859E-03	1.199E-03	0.00129**	1.068
Compositional Similarity model				
(Intercept)	5.020E-01	1.729E-01	0.00379**	
Log-Geographical distance	-7.383E-02	1.018E-03	< 2E-16***	1.001
Primary vegetation	5.336E-01	6.378E-02	< 2E-16***	1.189
Urban lands	-1.086E+00	1.583E-01	7.37E-12***	1.499
Log-Population density	-8.599E-02	1.541E-02	2.51E-08***	1.610
Log-Travel time	-1.188E-01	2.514E-02	2.35E-06***	1.851

Table 3. Predictors and corresponding coefficients and VIF values yielded from linear mixed-effect models for transformed abundance and compositional similarity. Significance level: *** is $p < 0.001$; ** $p < 0.01$; * $p < 0.05$. Log is natural logarithm transformation.

pre-industrial reference values are unavailable, relative reference values were statistically derived using the 90th percentile of Ab and Cs from the 2000 maps, rather than relying on any pristine assumption or site-level data from PREDICTS database. The 90th was chosen over the maximum to minimize the influence of extreme outliers and sampling bias, thereby providing a more robust and representative estimate of high biodiversity conditions across global landscapes. These reference values were then used to rescale annual Ab and Cs maps by dividing each pixel in the annual maps by the corresponding reference values of Ab and Cs, respectively. The final BII maps were obtained by multiplying the normalized annual Ab and Cs layers, with any pixel values exceeding one (1) capped at 1, representing fully intact habitat.

Biodiversity loss allocation to agricultural activities. Biodiversity loss allocation aimed to quantify biodiversity loss footprints, represented by BII loss, caused by agricultural production in crop and livestock systems⁶⁷. We integrated spatial data on BII loss with spatial distributions of agricultural production (i.e., time series data of SPAM and GLW). The original BII reflects the remaining biodiversity intactness under land use changes and anthropogenic pressures. Therefore, biodiversity loss (BII_{loss}) was inversely defined as the proportion of biodiversity that has been degraded due to any anthropogenic impacts, calculated as the complement of BII (Eq. 2).

At any given location, BII loss resulted from multiple pressures as presented in described models such as land use, infrastructure development, and other human interventions. This allows a complex picture to emerge. For example, BII loss observed in a cropland pixel may result not only from crop cultivation itself but also from adjacent influences such as human settlement agglomeration and infrastructure development. Fully attributing this BII loss solely to cropland may therefore lead to oversimplification and an overestimation of cropland's contribution. However, the interpretation of overall biodiverse loss based on factor contributions in the original multiplicative formulation of BII, where BII is derived by multiplying Ab and Cs, is not straightforward. Land use and human pressures can have nonlinear effects on BII and may influence both components. Therefore, an additional step was implemented to decompose and quantify the contribution of individual land use categories to the overall BII loss (Supplementary Information 3). The decomposition revealed that croplands, agroforestry systems (including fruits and orchards), and pasturelands (associated with livestock grazing) contributed approximately 21.02%, 15.85%, and 16.19% of the total BII loss, respectively. These values are referred to as contribution factors (CF) for each land use category in relation to total BII loss. More specifically, BII loss attributed to a particular land use category (lu) was estimated as a function of total BII loss, the contribution factor of that land use, and its corresponding land use fraction derived from the HHLU dataset ($BII_{loss,lu}$, Eq. 3).

Subsequently, the relative biodiversity loss was quantified as the BII loss footprint, representing the approximate area that experienced BII loss (km^2 , Eq. 4). In other words, for example, if a pixel i with a total area of a_i experienced a BII loss of 20%, the corresponding BII loss footprint would be approximately 20% of that pixel's total area.

In essence, a single pixel in the BII loss layer (approximately 625 hectares) may encompass a mix of different crop and livestock production systems. To allocate biodiversity loss more precisely, the BII loss footprint associated with each land use type was further disaggregated to estimate the specific footprints attributable to individual crops and livestock items. It should be noted that the BII loss footprint attributed to cropland and agroforestry corresponds to crop production, whereas the loss associated with pastureland is allocated to livestock systems. The land use-level BII loss was further apportioned to specific product items within the crop, agroforestry, and livestock sectors by multiplying it with their respective proportional shares in SPAM and GLW data (Eq. 5)^{67,68}. This allows us to estimate the individual contributions of each product type to the total BII loss footprint (Figure S4).

$$BII_{loss} = (1 - BII) \quad (2)$$

$$BII_{loss,lu} = CF_{lu} \times F_{lu} \times BII_{loss} \quad (3)$$

$$FP_i^{BII_{loss,lu}} = BII_{loss,lu} \times a_i \quad (4)$$

$$FPx_i^{BII_{loss}} = FP_i^{BII_{loss,lu}} \times P_{x,lu(i)} \quad (5)$$

where BII_{loss} is the total biodiversity intactness loss; $BII_{loss,lu}$ is the portion of biodiversity intactness loss attributable to land use lu ; CF_{lu} is the contribution factor of land use lu to total BII loss; F_{lu} is the land use fraction of land use lu [0:1]; $FP_i^{BII_{loss,lu}}$ (km^2) is the BII loss footprint induced by lu at pixel i ; a_i is the area of pixel i ($\sim 6.25 \text{ km}^2$); $FPx_i^{BII_{loss}}$ (km^2) is the BII loss footprint attributable to crop, agroforestry, or livestock x at pixel i ; and $P_{x,lu(i)}$ is the shared proportion of crop, agroforestry, or livestock x within land use lu at pixel i .

The spatial output maps of BII loss footprints represent the area of BII loss attributed to individual crops, agroforestry, and livestock commodity groups aligned with SPAM and GLW broader categories. Global inventories and multi-regional input/output (MRIO) databases typically capture economic activities and environmental pressures at the national level. To ensure consistency and further enhance the footprint analysis, we quantified total BII loss footprints for each agricultural item group at the national level in an annual time series. Additionally, within each country, these total footprints were specified for each corresponding ecoregion. Livestock systems had been disaggregated into dairy and meat production systems at an earlier stage of GLW time series generation. The approximately forty crop and agroforestry item groups in the SPAM dataset are still relatively too broad to capture the granularity of crop-level footprints, and we aimed to provide data at a more detailed and crop-specific level. To enhance granularity and ensure compatibility with other datasets such as FAOSTAT, we further disaggregated the national footprints at the item group level into specific FAOSTAT commodity items based on their relative production shares. In particular, production data of all crop items corresponding to SPAM groups (Table S3) was acquired from the FAOSTAT database and used to calculate each item's share within its respective SPAM group. If a group contains only one crop item, it is assigned to the full proportion. The BII loss footprint attributed to each specific crop item was then obtained by multiplying the total footprint of the general SPAM group by the item's share. The final footprint dataset provides ecoregional and national-level footprints for 154 crop items and 9 livestock categories in dairy and meat production systems.

Data Records

The dataset is organized into four main folders on the Figshare repository (<https://doi.org/10.6084/m9.figshare.28303442>⁶⁹, Fig. 3), each containing annual global spatial data and tabular information of the BII loss footprint attributable to each crop and livestock commodity. The spatial datasets comprise (1) high-resolution harmonized land use (HHLU), (2) land use fractions, and (3) BII maps from 2000 to 2020, all provided in GeoTIFF raster format (*.tif) using the “EPSG:4326” (WGS84) coordinate reference system (CRS). The spatial resolution of HHLU is ~ 0.004 arc-degrees (approximately 500 meters at the Equator), while both land use fractions and BII maps have a resolution of ~ 0.02 arc-degrees ($\sim 2.5 \text{ km}$ at the Equator). Each map folder also includes a “readme” file detailing file name structure and value range. In addition, a map style (*.qml) file is provided in each folder to supply map styling for visualization in QGIS.

The HHLU folder includes one GeoTIFF file (*.tif) per year, with discrete values from 1 to 7 representing distinct land use and habitat states. Values 1 and 2 correspond to primary minimal-used vegetation and primary vegetation, respectively, which were combined into a single primary vegetation class for BII modeling (Fig. 4). Secondary vegetation, cropland, urban lands, pasture, and agroforestry are coded as 3, 4, 5, 6, and 7, respectively.

The land-use fractions are organized into six separate folders, each corresponding to one land-use category. Each folder contains one GeoTIFF file (*.tif) per year representing the fractional coverage of one land use category as described in the section in the text on land-use harmonization. Pixel values range from 0 to 1, indicating the coverage proportion of each land use within that pixel (Fig. 5), with the sum of all six fractions at any location totaling 1.0.

The BII dataset provided in a separate folder contains one GeoTIFF file (*.tif) per year, representing the remaining biodiversity intactness globally, which reflects the remaining biodiversity intactness across the world. BII values range from 0 to 1. A higher BII value indicates a higher level of biodiversity integrity (Fig. 6).

The dataset of BII loss footprints for crops and livestock categories is given in tabular form (Comma-separated values, CSV format, *.csv), which is stored in a separate folder, detailing quantitative statistics on the spatial imprint of each crop and livestock commodity across biomes and countries/territories annually from 2000 to 2020. The footprint dataset includes fourteen columns (Table 4). Each row represents the area of BII loss footprint (km^2) associated with a specific product in a given year, country/territory, UN region, continent, and biome. This data captures the BII loss footprint of 154 crops aligned with FAO categories, and five livestock categories covering both meat and dairy production systems, resulting in a total of nine livestock commodities (with horses assumed only in meat production).

Technical Validation

Technical validation comprises four main tasks: (1) assessing data consistency of HHLU maps by comparison with currently available land use maps with the same classification system; (2) verifying the BII map outputs against multiple independent datasets; (3) validating spatial agreement between the extended SPAM and GLW maps and reference data; and (4) conducting uncertainty analyses of land use and land use fractions alongside sensitivity analysis of the biodiversity models.

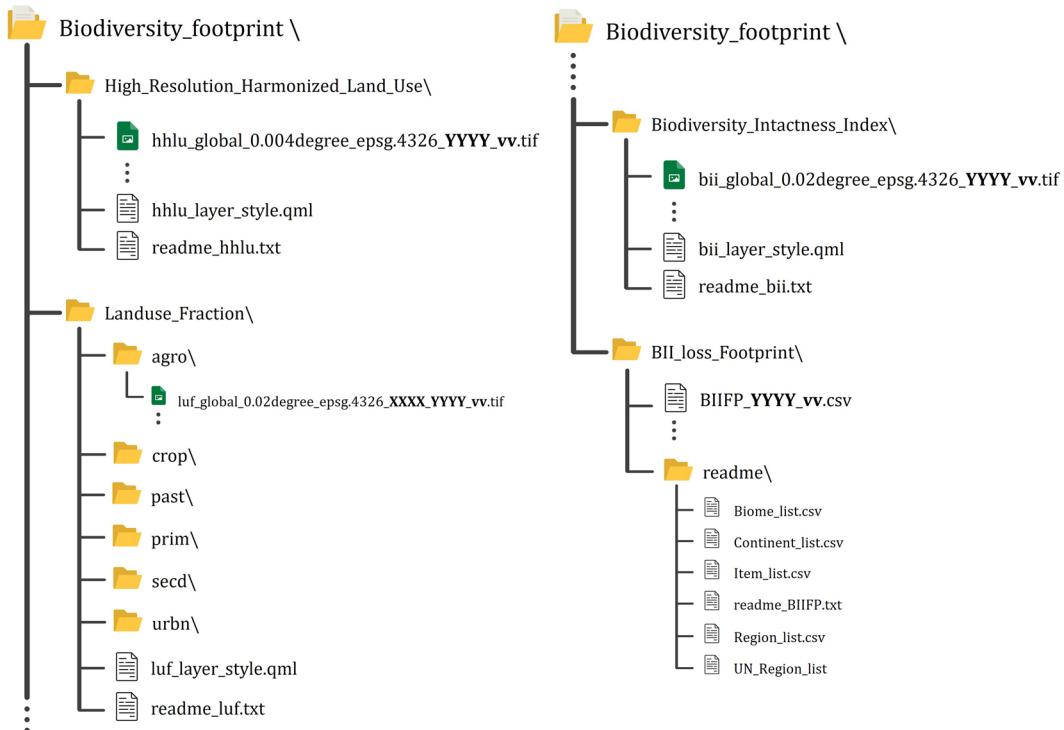


Fig. 3 Folder structure of the data repository. XXXX is the four letters of land use type (e.g., prim), YYYY is the observation year (e.g., 2020), and vv is the data version.

Validation of land use and land-use fractions. Land use data for biodiversity integrity assessments are defined using the specific classification system of the LUH dataset that describes habitat states and potential human disturbance through land use types such as primary and secondary vegetation habitats. Acquiring comparable data for cross-comparison poses a challenge. We addressed this by validating the generated datasets by using multiple independent datasets. First, although the land use types in the HHLU are closely aligned with the LUH system, the cropland and pastureland categories are best aligned with the FAO land use system. Therefore, we used national-level cropland and pastureland area (i.e., permanent meadows and pastures) from FAOSTAT over the observed period as a reference for comparing with the corresponding categories in the HHLU dataset. It should be noted that cropland in the HHLU dataset for the validations includes both cropland and agroforestry in order to maintain consistency with other datasets. The results indicate that the HHLU dataset reliably captures cropland and pastureland areas in most countries and territories, with correlation coefficients of 0.962 (95% CI: 0.950–0.971) for cropland and 0.949 (95% CI 0.932–0.962) for pastureland (Figure S5).

Subsequently, all land use classes were evaluated using accuracy assessment procedures based on the down-scaled LUH data for 2005⁷⁰ (Supplementary Information 4). The accuracy assessment employed a bootstrap method to estimate accuracy uncertainty and confidence intervals. The HHLU map, generated in accordance with the LUH framework, demonstrated relatively high reliability, with an overall accuracy of approximately 78.7% (95% CI: 78.3–79.1%) and a kappa coefficient of 0.724 (95% CI: 0.718–0.729). The classification performance for each land use class was further assessed through user's accuracy. Primary vegetation and urban land exhibited the highest user's accuracy among the land use categories, with confidence intervals of approximately 87.2–88.5% and 81.7–84.9%, respectively. Other land use types also achieved relatively high classification performance: secondary vegetation (73.7–75.5%), cropland (71.3–73.2%), and pasture (77.3–79.4%).

In addition to assessing the accuracy of land use classes, this process further validated the land use fraction maps derived from HHLU to ensure their spatial alignment with the downscaled LUH data for 2005⁷⁰. To make this comparison, we obtained five layers from the downscaled LUH data including primary and secondary vegetation, cropland, pasture, and urban land. Ice and desert regions were excluded from all comparative analyses. First, the downscaled LUH layers were resampled to match the spatial resolution of the HHLU land use fractions. These layers were then compared at both the pixel and country levels using Pearson correlation to verify their consistency at different levels. Specifically, all pixels were extracted and compared for the pixel-level analysis, while national averages of land use fractions were used for the country-level evaluation. Since the HHLU integrates multiple data sources with varying spatial resolutions, spatial uncertainty may be propagated from the input layers into the resulting land use fractions. To account for this, we also tested correlations across different spatial aggregation levels. At the original spatial resolution (~2.5 km), the overall consistency between HHLU and the downscaled LUH was moderate, with a correlation coefficient of approximately 0.729 at the pixel-based level for all classes (Fig. 7). This consistency improved significantly at a coarser spatial scale of ~10 km ($r=0.784$), but no further increase was observed at ~20 km resolution. Among individual classes at the pixel level, cropland showed the highest spatial agreement ($r=0.828$), while other classes also exhibited

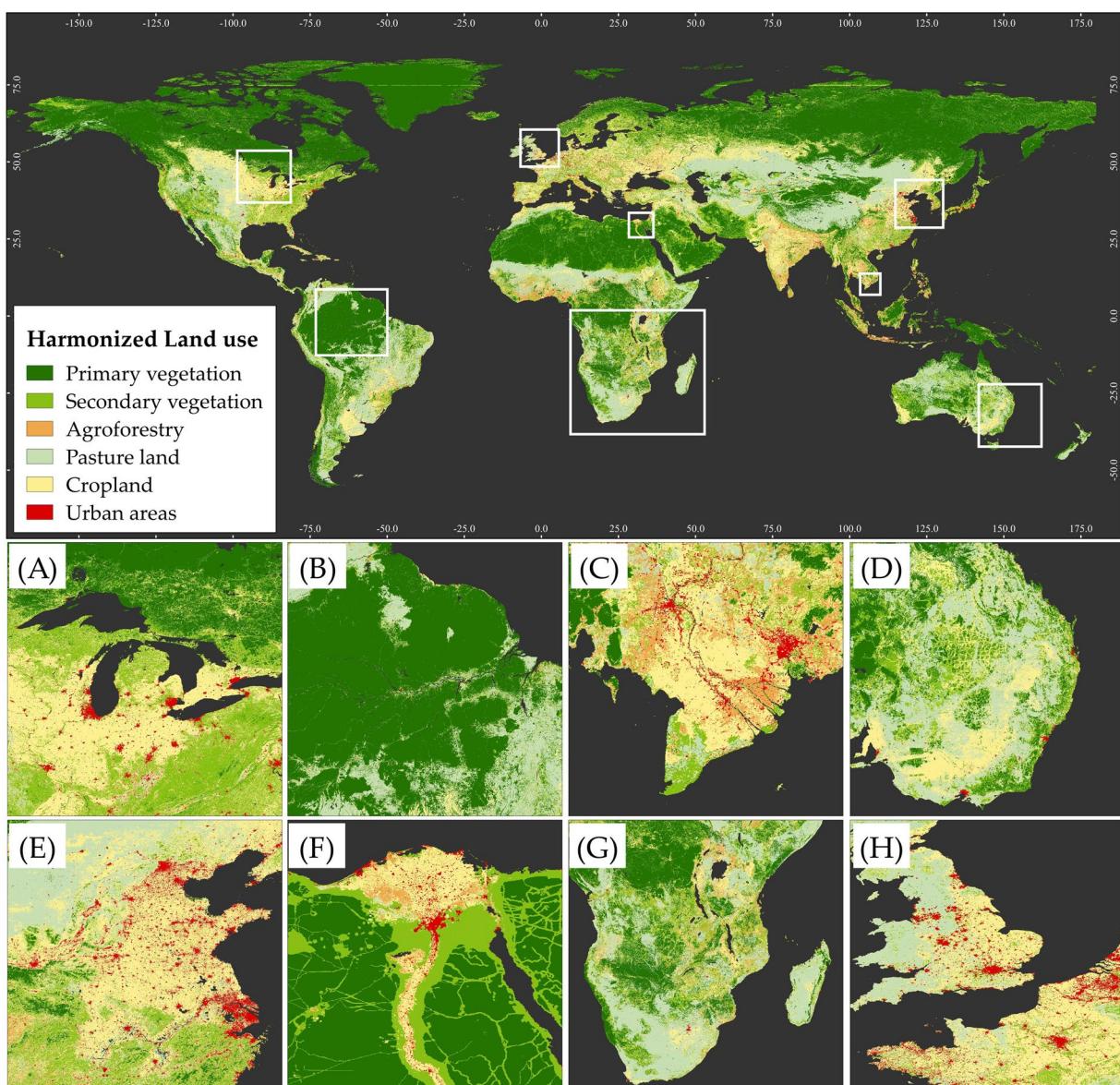


Fig. 4 Harmonized land use map in 2020 and fragmented zoom-in examples in (A) Northeast America, (B) Amazon, (C) Mekong Delta, (D) Eastern Australia, (E) Northeast China, (F) Nile Delta in Egypt, (G) Southern Africa, and (H) Western Europe, depict dominant land use categories globally.

relatively high agreements (approximately $r = 0.8$) under the ~ 10 km aggregation. At the country-level, where aggregation effects were minimized, the correlation peaked, with an overall coefficient of approximately 0.751 (95% CI: 0.720–0.780) (Figure S6). Urban land, cropland, and pasture were the most consistent classes, with correlation coefficients of 0.878 (95% CI: 0.837–0.909), 0.871 (95% CI: 0.828–0.903), and 0.845 (95% CI: 0.795–0.884), respectively.

Secondary vegetation showed with relatively low spatial agreement in both scale analyses, with correlation coefficients ranging 0.331–0.420. This observed inconsistency is likely due to differences in data sources and data generation approaches between LUH and HHLU. The original LUH modeled and generated soft land use fractions, while the maps in this study derived from historical datasets, leading to inconsistency within even the same land use class (Figure S7). For example, the lowest consistency was observed in African countries, where forest management and intact forest landscape indicated large areas of primary forest, while the LUH dataset tended to classify these regions as secondary vegetation (Figure S7). Pixel values in the HHLU-derived dataset often reach the maximum value within aggregation windows where land use is homogeneous, a scenario that rarely occurs in the LUH dataset. This difference arises because our land use fractions were generated from land use-based estimates rather than model-derived fractions, which may amplify apparent inconsistencies. Although our land use fractions showed relatively low consistency for secondary vegetation with the original LUH, they generally offer a more historically grounded and observation-based representation of land use

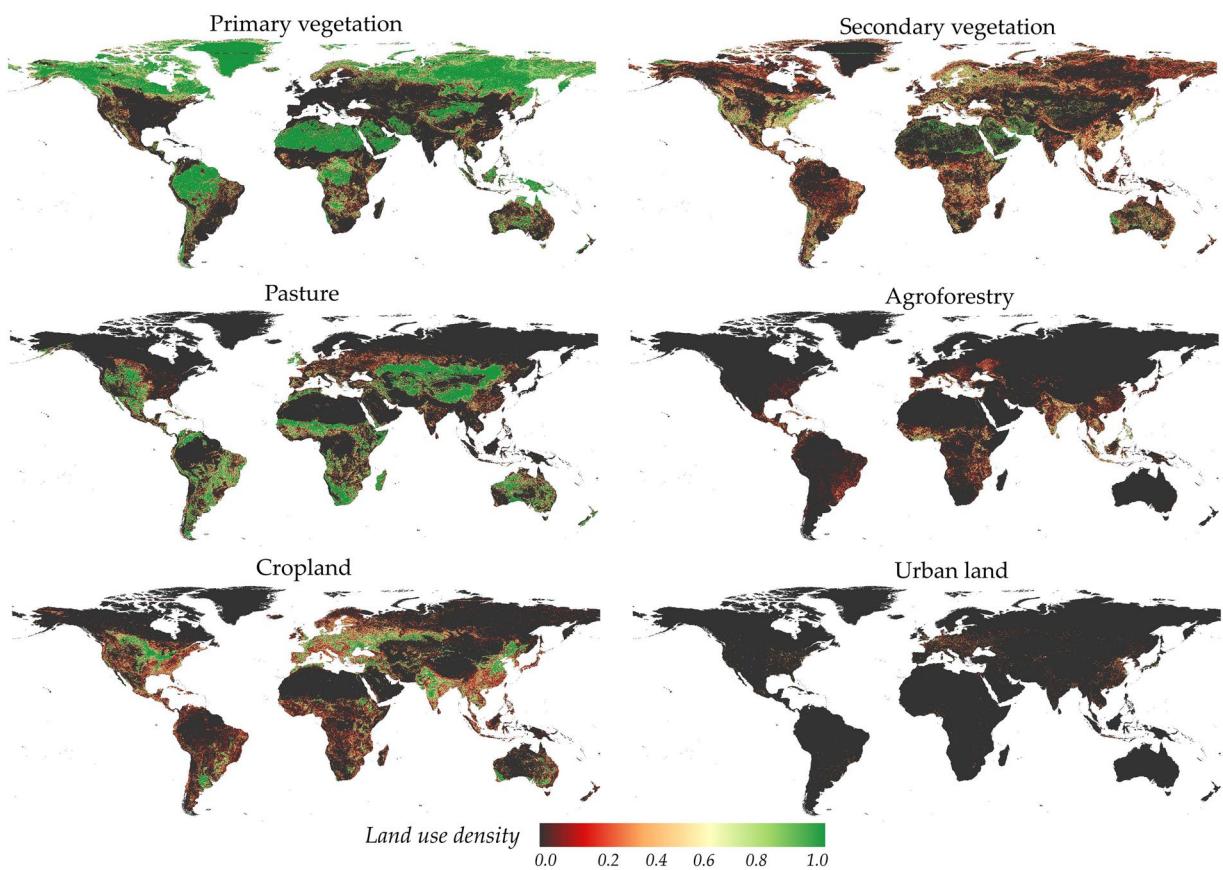


Fig. 5 Land use fractions of six harmonized land use categories ranging from 0 to 1 (0–100%). An example of a land use fraction was calculated from a harmonized land use map in 2020.

patterns for biodiversity integrity assessments that require accurate historical baselines rather than relying solely on model-derived estimates.

Given the classification performance of the HHLU and aggregation of land use in generating land use fraction maps, it is essential to estimate the associated uncertainty to inform users of potential limitations and confidence in these datasets. Therefore, an uncertainty assessment for land use fractions to evaluate the robustness of HHLU dataset across classes for biodiversity modeling as also conducted. The overall uncertainty of land use fractions arises from two sources: uncertainty in classification performance and uncertainty in aggregation to estimate fractions resulting from the mixing or heterogeneity of land use within an aggregation window. The first was derived from an accuracy assessment, while the latter was estimated by the bootstrap method. More explicitly, we implemented a block bootstrap approach (5×5 pixels) over each land use raster layer of the HHLU. For each aggregation window, the pixel values were repeatedly resampled with replacement and the fraction of the target land use type was calculated across $n = 50$ bootstrap iterations. The mean proportion provided the estimated land use fraction, while the standard deviation across iterations served as uncertainty estimates for each aggregated block. To estimate overall uncertainty, we combined the classification uncertainty and the aggregation uncertainty arising from land use heterogeneity using the following equation:

$$\text{Uncertainty} = \sqrt[3]{se_{lui}^2 + se^2} \quad (6)$$

where se_{lui} is uncertainty from classification, and se is uncertainty from the aggregation process.

The estimated uncertainty of land use fractions varied across land use types (Fig. 8). Primary vegetation and urban land exhibited relatively high uncertainties, approximately 8% (95% CI: 4.1–12.0% for primary vegetation; 1.9–17.0% for urban land), indicating moderate spatial variability. Secondary vegetation showed a comparable mean uncertainty of 6.8%, but its confidence interval was much wider (95% CI: 1.8–21.5%). In contrast, cropland and pasture demonstrated lower average uncertainties of 4.5% and 5.8%, respectively, though both had wide confidence intervals, indicating spatial variability in data reliability across regions. These uncertainties imply that cropland and pasture were generally well-represented in the HHLU datasets, while secondary vegetation and urban areas should be interpreted more cautiously due to their high spatial uncertainty.

Sensitivity analysis of BII models and validation of BII maps. *Sensitivity analysis of BII models.* A sensitivity analysis of BII models to land use and anthropogenic pressure is essential to ensure the robustness of biodiversity modeling and assessment. Biodiversity responses to environmental changes are complex, and

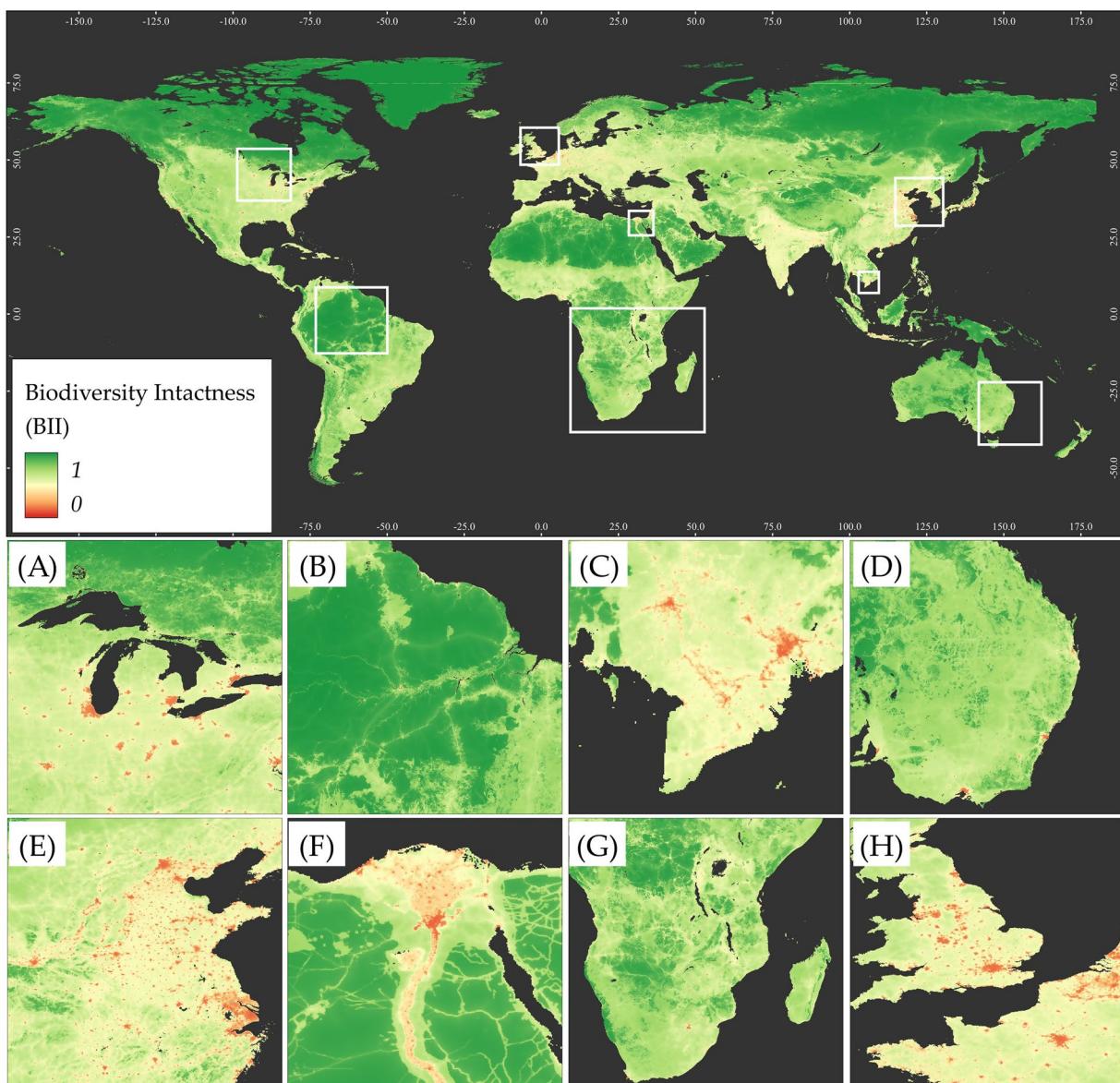
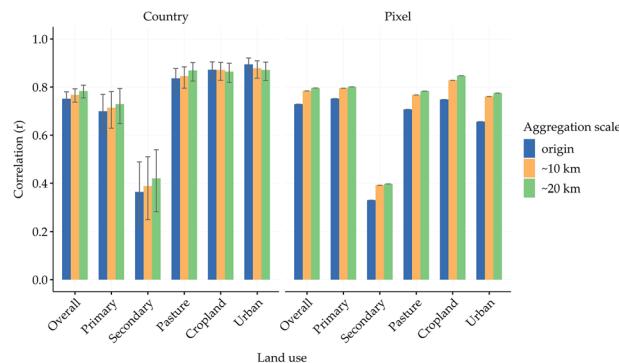
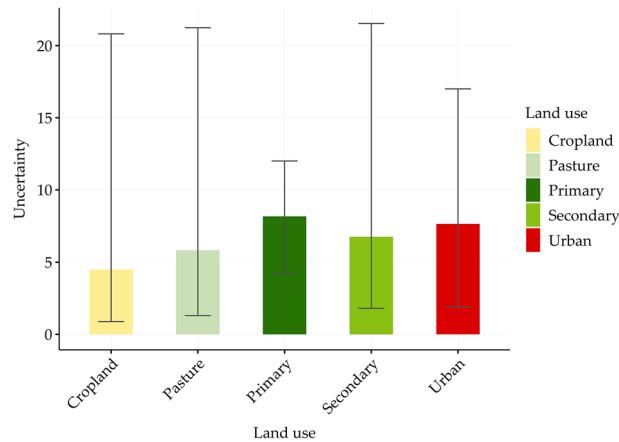


Fig. 6 Biodiversity intactness index (BII) in 2020 and examples in (A) Northeast America, (B) Amazon, (C) Mekong Delta, (D) Eastern Australia, (E) Northeast China, (F) Nile Delta in Egypt, (G) Southern Africa, and (H) Western Europe. Green represents intact ecosystems and red indicates ecosystems with high human intervention.

different drivers can affect various aspects of biodiversity in distinct ways^{53,71}. By explicitly evaluating how sensitive abundance and compositional similarity are to land use against human pressures, we can identify which key inputs meaningfully contribute to each biodiversity component. Separate models for Ab and Cs were fitted using only land use predictors, only human pressures, and compared against the baseline models (or the fitted models) that combined both sets of predictors. This revealed that abundance is more sensitive to certain types of land use such as primary vegetation, cropland, and urban areas, whereas compositional similarity is influenced more by human pressures including population density, proximity to roads, and travel time. Correlation tests of simulated BII indicated that BII from the fitted models correlates better with the land use model ($r=0.95$) than the human pressure model ($r=0.92$). Land use variables are better at capturing the overall biodiversity intactness patterns reflected in the fitted model, particularly for secondary vegetation, cropland, pasture, and urban lands. This sensitivity analysis emphasized the critical role of land use in biodiversity modeling, while also demonstrating that incorporating human pressures provides valuable complementary information for efficiently capturing biodiversity patterns³. Uncertainty in land use fractions can propagate through the BII model: the highest uncertainty in BII estimates occurred in urban landscapes, followed by secondary vegetation and cropland. In contrast, propagated uncertainty for pasture and primary vegetation remained comparatively low, reflecting more stable BII estimates for these landscapes.

Column	Description
FAOID	FAO numeric country code
ISO3	ISO 3166-1 alpha-3 country code
Biome	Name of terrestrial biome based on the Worldwide Fund (WWF) classification system
Region	Full name of country or territory
Region_SHO	Short country or territory name
UN_Region	United Nations regional classification (e.g., Eastern Europe)
Continent	Continent Name (e.g., Africa)
Year	Year of assessment (2000 – 2020)
Category	Product category (i.e., Crop or Livestock)
Item_FAO	FAO numeric code for product
Item_BIIFP	Internal numeric code aligned with BII footprint to separate dairy and meat production (e.g., 1 is dairy and 2 is meat, Table S4)
Item_Name	Common product name (e.g., Wheat, Barley)
Unit	Unit of measurement (km^2) characterizes spatial footprint area
Value	The estimated area (km^2) of BII loss footprint attributed to a specific product in a specific biome, country, and year.

Table 4. Data columns of BII loss footprint data and descriptions.**Fig. 7** Correlation of land use fractions between HHLU and downscaled LHU across different levels and spatial aggregations. Error bars show 95% CI.**Fig. 8** Average uncertainty in land use fractions across land use types. Error bars illustrate 95% confidence intervals.

Validation of BII maps. The spatial BII maps were generated using a combination of abundance and compositional similarity models and validated against three independent reference datasets: the updated global biodiversity intactness in 2005 [BII(S)]⁷², high-resolution BII maps from 2017–2020 [BII(G)]³⁹, and Biodiversity Habitat Index (BHI) for 2020⁷³. The earlier BII datasets also used the PREDICTS database to model BII in discrete years, while BHI estimated the retention of terrestrial species diversity under habitat loss, degradation,

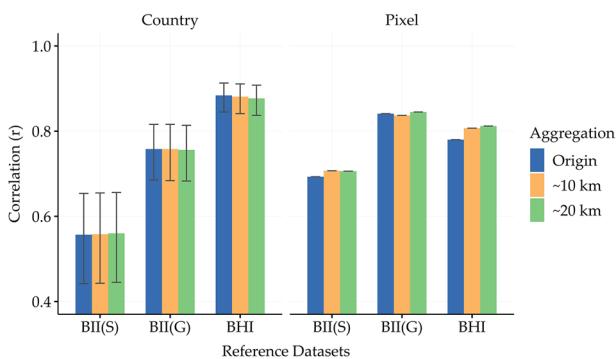


Fig. 9 Correlation of simulated BII and reference datasets at pixel and country levels across different spatial aggregation scales.

and fragmentation pressures, providing a complementary perspective on biodiversity integrity. Similar to land use fractions, validation for BII also examined consistency across pixel- and country-levels, examining how spatial aggregation influences the correlation strength and reliability of BII estimates. BII maps were compared with reference datasets from corresponding years after resampling to match the modeled BII spatial resolution (~2.5 km) and all tested aggregation scales. At the pixel level, correlations for BII(G) were highest, ranging from 0.837 to 0.845, followed by BHI (0.780–0.812) and BII(S) (0.639–0.707). At the country level, BHI showed the highest correlations (0.877–0.884), followed by BII(G) (0.756–0.758) and BII(S) (0.557–0.560), with confidence intervals reflecting moderate reliability. At both the pixel and country levels, the correlations remained relatively stable as aggregation increased, indicating consistent spatial agreement regardless of pixel aggregation (Fig. 9). The biodiversity models emphasized the prominence of land use in BII estimation. As BII(S) was generated using the LUH data, inconsistencies between LUH-based land use and our land use fractions may propagate into BII maps, leading to lower consistency with BII(S). However, our simulated BII maps still exhibited high consistency with BII(G) and BHI, indicating the strong alignment of the simulated BII with these reference datasets. Additionally, correlations at the country level were generally lower than those at the pixel level, underscoring spatial heterogeneity and the influence of data aggregation, which can dilute biodiversity signals when synthesizing raster-based data to the country-level summaries.

Moreover, a comparison at biome/ecoregion level was also provided to furnish additional perspectives and to assist users in discerning regions where the dataset may be confidently applied and regions where more cautious interpretation may be necessary (Figure S8⁷⁴ and Table S5). The comparison revealed varying consistency across biomes. Forested biomes, such as Tropical and subtropical moist broadleaf forests ($r = 0.805$), Temperate conifer forests ($r = 0.662$), and Boreal forests/taiga ($r = 0.662$), Mangroves ($r = 0.641$), Temperate broadleaf and mixed forests ($r = 0.634$), Deserts and xeric shrublands ($r = 0.632$), Tropical and subtropical dry broadleaf forests ($r = 0.613$), had high averaged correlations among the references, indicating strong agreement in these biomes. In contrast, open and less vegetated biomes, e.g., tundra ($r = 0.323$) and temperate grasslands, savannas, and shrublands ($r = 0.428$) exhibited lower correlations. In general, the BII data from Gassert *et al.*³⁹ tend to report higher values for individual and global biomes, which may lead to overestimation or more optimistic assessments of biodiversity integrity in contrast to the BHI which offers a more conservative evaluation. The BII maps generated in this study show moderate to high levels of consistency and reliability with at least one reference dataset, which is expected to balance the limitations present in existing datasets. Overall, these validations demonstrated that our BII maps are both consistent and reliable, offering a potential improvement on current datasets for further biodiversity integrity assessments.

Validation of spatial crop and livestock distribution maps. Spatial maps of SPAM and GLW generated by the extended methods were underwent rigorous validation to ensure both quantitative accuracy and spatial agreement. These maps redistributed FAO-reported data to specific locations and regions based on crop and livestock probability distributions. Since ordinary correlation assessments at the national level cannot capture spatial homogeneity, we employed the spatial correlation test in *terra* R-package to assess spatial agreement between the reference and generated maps, explicitly accounting for spatial heterogeneities between map pairs⁷⁵. Additionally, the results were evaluated using Pearson correlation coefficients. Validation was conducted by comparing SPAM data for three reference years (2005, 2010, and 2020) and GLW data for 2010, 2015, and 2020. The validation results, summarized in Fig. 10, illustrate spatial correlation coefficients (r) as two-dimensional metrics: average values across all years and minimum values for specific crop and livestock categories. Both SPAM and GLW maps produced through the redistribution methods showed strong spatial agreement with reference data across all years, with mean correlation coefficients of 0.914 ± 0.056 for SPAM and $r = 0.960 \pm 0.045$ for GLW. Although spatial heterogeneity varies among individual categories, all demonstrate moderate to highly reliability. Notably, tropical fruits (TROF, 2020) and plantain (PLNT, 2005) exhibited the lowest correlation coefficients at 0.734 and 0.760, respectively, while horses (HORS, 2015) showed the lowest spatial agreement among livestock categories ($r = 0.860$) (Fig. 10).

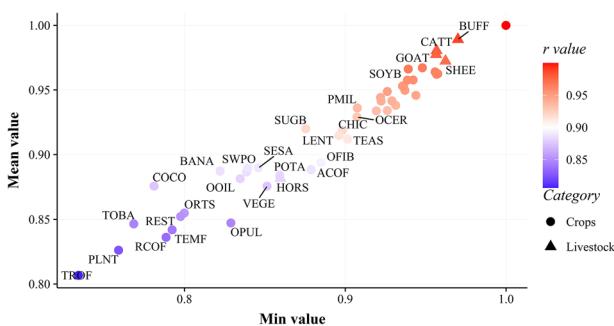


Fig. 10 Spatial correlation between the reference and generated maps of SPAM and GLW. Circles and triangles represent crops and livestock items, respectively. Abbreviations of SPAM are shown in Table S3. GLW livestock categories: BUFF = Buffalo, SHEE = Sheep, GOAT = Goat, HORS = Horses, CATT = Cattle.

Usage Notes

The validated annual HHLU dataset represents one of the first efforts to produce a consistent and high-resolution global land-use dataset. It addresses uncertainties present in earlier datasets and is dedicated supporting robust biodiversity assessments. The land-use definitions align seamlessly with the established LUH classification system, ensuring broad applicability across biodiversity and ecological integrity studies. Furthermore, land-use fractions serve as an intermediate resource, facilitating a wide range of ecological analyses.

The annual BII maps demonstrate strong concordance with existing datasets and have the potential to resolve inconsistencies in biodiversity intactness observations across certain biomes. These maps offer valuable support for a wide range of applications, including environmental research, policymaking, and conservation initiatives. By analyzing temporal trends, researchers can identify regions experiencing biodiversity decline or recovery, evaluate the effectiveness of conservation strategies, and investigate the complex interactions between biodiversity and environmental drivers, such as urbanization, deforestation, mining extraction, climate change, and land use dynamics.

The BII footprint on specific crops and livestock items has been meticulously disaggregated across 154 crop types, 09 livestock categories, 14 biomes, and 193 countries and territories over the past two decades. This dataset allocates BII loss footprint directly attributable to agricultural production, integrating the most current available data while filling critical gaps—such as buffalo populations in North America and the distinction between dairy and meat livestock sectors—that are absent from the FAO dataset. This detailed data empowers researchers and policymakers to aggregate and analyze diverse factors underlying terrestrial biodiversity integrity loss driven by agricultural production systems. It not only reflects trends in biodiversity integrity but also illuminates the dynamic of changes occurring within biomes, countries, and production sectors, enabling identification of key drivers of biodiversity decline and informing targeted conservation and sustainable consumption strategies. Some key trends can be observed in the summary in Fig. 11 below.

By providing essential biodiversity footprint data, this resource enables the precise identification of local impacts from agricultural production on terrestrial ecosystems, highlights biodiversity hotspots at greatest risk, helps prioritize conservation efforts, and enables projections of future trends under varying land use and consumption scenarios. Moreover, the dataset facilitates comprehensive footprint analyses by quantifying biodiversity loss impacts on human activities, industries, and international trade through integration with multi-regional input/output models. This approach reveals the ecological costs and responsibilities embedded within global supply chains, uncovers hidden ecological inequalities among countries and regions, and supports the development of demand-side policies to promote more sustainable production and consumption patterns. Incorporating this footprint data into international trade frameworks can further drive sustainable trade policies by offering a standardized metric to assess the biodiversity impacts of imports and exports, thereby encouraging trade agreements that balance biodiversity conservation with economic growth.

It is important to note that certain datasets (e.g., global road network and travel time) were treated as static drivers due to the difficulty of obtaining consistent temporal datasets. As the expansion of road and transportation infrastructures can substantially contribute to a loss of biodiversity intactness, this static assumption may lead to an underestimation of overall biodiversity decline. However, because these factors are assumed to be globally consistent, relative comparison between regions and countries remains valid and unaffected. Moreover, it should be noted that the dataset represents consolidated regional units rather than strictly adhering to current geopolitical boundaries. This consolidation resulted from data harmonization over the observed period, facilitating the alignment of FAO datasets and other data sources during the disaggregation processes.

Data availability

Supporting data used to harmonize land-use dataset and generate BII maps can be accessed directly from the respective data sources. All datasets produced in this work are openly available for education and non-commercial purposes via the project data archive on Figshare (<https://doi.org/10.6084/m9.figshare.28303442>)⁶⁹.

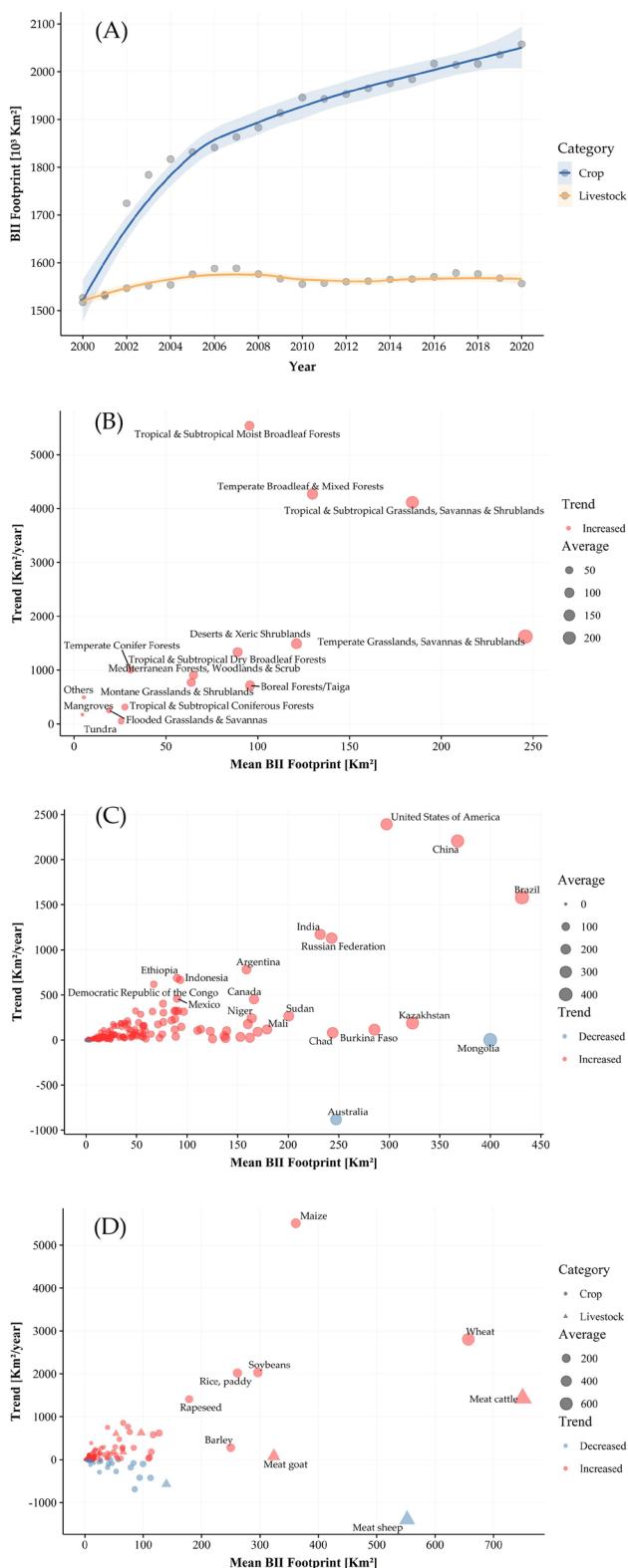


Fig. 11 Summary of BII loss footprints: **(A)** overall trend of total BII loss footprint on crops and livestock from 2000 to 2020, and comparisons between average footprint and overall trend on **(B)** biomes, **(C)** countries/territories, and **(D)** commodities. The footprint from the livestock sector has gradually declined, whereas the footprint associated with crop production has increased persistently over time. Increases in biodiversity footprints were observed across all biomes, with particularly notable growth in tropical and subtropical moist broadleaf forests, temperate broadleaf and mixed forests, and tropical and subtropical grasslands, savannas, and shrublands. At the country level, footprints increased in nearly all countries, except for Australia and Mongolia, where substantial declines were observed. Among commodities, the most significant increases in footprints were associated with maize, wheat, and meat cattle, while a notable decrease was seen for meat sheep.

Code availability

All datasets and statistical models in this work were generated using R (version 4.4.1) and Python (version 3.11.5). No custom code or programs were developed during the generation and processing of these datasets.

Received: 30 January 2025; Accepted: 26 August 2025;

Published online: 02 October 2025

References

1. Beery, T. *et al.* Disconnection from nature: Expanding our understanding of human–nature relations. *People Nat.* **5**, 470–488, <https://doi.org/10.1002/pan3.10451> (2023).
2. IPBES. *Summary for Policymakers of the Global Assessment Report on Biodiversity and Ecosystem Services of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*, <https://doi.org/10.5281/zenodo.3553579> (Bonn, Germany, 2019).
3. Butchart, S. H. M. *et al.* Global Biodiversity: Indicators of Recent Declines. *Science (80-)* **328**, 5982, 1164–1168, <https://doi.org/10.1126/science.1187512> (2010).
4. Dove, S., Böhm, M., Freeman, R., McRae, L. & Murrell, D. J. Quantifying reliability and data deficiency in global vertebrate population trends using the Living Planet Index. *Glob. Chang. Biol.* **29**, 4966–4982, <https://doi.org/10.1111/gcb.16841> (2023).
5. Hill, S. L. L. *et al.* Worldwide impacts of past and projected future land-use change on local species richness and the Biodiversity Integrity Index. *bioRxiv* 311787, <https://doi.org/10.1101/311787> (2018).
6. Raimondo, D. *et al.* Using Red List Indices to monitor extinction risk at national scales. *Conserv. Sci. Pract.* **5**, 1–12, <https://doi.org/10.1111/csp2.12854> (2023).
7. Leclère, D. *et al.* Bending the curve of terrestrial biodiversity needs an integrated strategy. *Nature* **585**, 551–556, <https://doi.org/10.1038/s41586-020-2705-y> (2020).
8. Pereira, H. M. *et al.* Global trends and scenarios for terrestrial biodiversity and ecosystem services from 1900 to 2050. *Science (80-)* **384**, 458–465, <https://doi.org/10.1126/science.adn3441> (2024).
9. Richardson, K. *et al.* Earth beyond six of nine planetary boundaries. *Sci. Adv.* **9**, 1–16, <https://doi.org/10.1126/sciadv.adh2458> (2023).
10. Mace, G. M. *et al.* Approaches to defining a planetary boundary for biodiversity. *Glob. Environ. Chang.* **28**, 289–297, <https://doi.org/10.1016/j.gloenvcha.2014.07.009> (2014).
11. Dudley, N. & Alexander, S. Agriculture and biodiversity: a review. *Biodiversity* **18**, 45–49, <https://doi.org/10.1080/14888386.2017.1351892> (2017).
12. Chaudhary, A., Pfister, S. & Hellweg, S. Spatially Explicit Analysis of Biodiversity Loss Due to Global Agriculture, Pasture and Forest Land Use from a Producer and Consumer Perspective. *Environ. Sci. Technol.* **50**, 3928–3936, <https://doi.org/10.1021/acs.est.5b06153> (2016).
13. Vačkářová, D., Medková, H., Krpec, P. & Weinzettel, J. Ecosystem services footprint of international trade: Economic value of ecosystem services lost due to crop production. *Ecosyst. Serv.* **64**, <https://doi.org/10.1016/j.ecoser.2023.101560> (2023).
14. Lanz, B., Dietz, S. & Swanson, T. The Expansion of Modern Agriculture and Global Biodiversity Decline: An Integrated Assessment. *Ecol. Econ.* **144**, 260–277, <https://doi.org/10.1016/j.ecolecon.2017.07.018> (2018).
15. Marques, A. *et al.* Increasing impacts of land use on biodiversity and carbon sequestration driven by population and economic growth. *Nat. Ecol. Evol.* **3**, 628–637, <https://doi.org/10.1038/s41559-019-0824-3> (2019).
16. Weinzettel, J., Hertwich, E. G., Peters, G. P., Steen-Olsen, K. & Galli, A. Affluence drives the global displacement of land use. *Glob. Environ. Chang.* **23**, 433–438, <https://doi.org/10.1016/j.gloenvcha.2012.12.010> (2013).
17. Balvanera, P. *et al.* Chapter 2.1. Status and Trends - Drivers of Change, in *Global assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services* 54–200, <https://doi.org/10.5281/zenodo.3831673> (Bonn, Germany, 2019).
18. Carrasco, L. R., Chan, J., McGrath, F. L. & Nghiêm, L. T. P. Biodiversity conservation in a telecoupled world. *Ecol. Soc.* **22**, <https://doi.org/10.5751/ES-09448-220324> (2017).
19. Crenna, E., Marques, A., Notte, A., La & Sala, S. Biodiversity Assessment of Value Chains: State of the Art and Emerging Challenges. *Environ. Sci. Technol.* **54**, 9715–9728, <https://doi.org/10.1021/acs.est.9b05153> (2020).
20. Marques, A., Verones, F., Kok, M. T., Huijbregts, M. A. & Pereira, H. M. How to quantify biodiversity footprints of consumption? A review of multi-regional input–output analysis and life cycle assessment. *Curr. Opin. Environ. Sustain.* **29**, 75–81, <https://doi.org/10.1016/j.cosust.2018.01.005> (2017).
21. Marquardt, S. G. *et al.* Consumption-based biodiversity footprints – Do different indicators yield different results? *Ecol. Indic.* **103**, 461–470, <https://doi.org/10.1016/j.ecolind.2019.04.022> (2019).
22. Damiani, M. *et al.* Critical review of methods and models for biodiversity impact assessment and their applicability in the LCA context. *Environ. Impact Assess. Rev.* **101**, 107134, <https://doi.org/10.1016/j.eiar.2023.107134> (2023).
23. Lenzen, M. *et al.* International trade drives biodiversity threats in developing nations. *Nature* **486**, 109–112, <https://doi.org/10.1038/nature11145> (2012).
24. Wilting, H. C., Schipper, A. M., Bakkenes, M., Meijer, J. R. & Huijbregts, M. A. J. Quantifying Biodiversity Losses Due to Human Consumption: A Global-Scale Footprint Analysis. *Environ. Sci. Technol.* **51**, 3298–3306, <https://doi.org/10.1021/acs.est.6b05296> (2017).
25. Boakes, E. H., Dalin, C., Etard, A. & Newbold, T. Impacts of the global food system on terrestrial biodiversity from land use and climate change. *Nat. Commun.* **15**, <https://doi.org/10.1038/s41467-024-49999-z> (2024).
26. Newbold, T. *et al.* Has land use pushed terrestrial biodiversity beyond the planetary boundary? A global assessment. *Science (80-)* **353**, 288–291, <https://doi.org/10.1126/science.aaf2201> (2016).
27. De Palma, A. *et al.* Annual changes in the Biodiversity Integrity Index in tropical and subtropical forest biomes, 2001–2012. *Sci. Rep.* **11**, 1–13, <https://doi.org/10.1038/s41598-021-98811-1> (2021).
28. Martin, P. A., Green, R. E. & Balmford, A. The biodiversity intactness index may underestimate losses. *Nat. Ecol. Evol.* **3**, 862–863, <https://doi.org/10.1038/s41559-019-0895-1> (2019).
29. Faith, D. P., Ferrier, S. & Williams, K. J. Getting biodiversity intactness indices right: Ensuring that ‘biodiversity’ reflects ‘diversity’. *Glob. Chang. Biol.* **14**, 207–217, <https://doi.org/10.1111/j.1365-2486.2007.01500.x> (2008).
30. Little, I. T., Hockey, P. A. R. & Jansen, R. Assessing biodiversity integrity for the conservation of grazed and burnt grassland systems: avian field metabolic rates as a rapid assessment tool. *Biodivers. Conserv.* **24**, 1443–1471, <https://doi.org/10.1007/s10531-015-0868-x> (2015).
31. Koch, J., Schaldach, R. & Göpel, J. Can agricultural intensification help to conserve biodiversity? A scenario study for the African continent. *J. Environ. Manage.* **247**, 29–37, <https://doi.org/10.1016/j.jenvman.2019.06.015> (2019).
32. Gopel, J., Schungel, J., Stuch, B. & Schaldach, R. Assessing the effects of agricultural intensification on natural habitats and biodiversity in Southern Amazonia. *PLoS One* **15**, 1–21, <https://doi.org/10.1371/journal.pone.0225914> (2020).
33. Shumba, T. *et al.* Effectiveness of private land conservation areas in maintaining natural land cover and biodiversity intactness. *Glob. Ecol. Conserv.* **22**, e00935, <https://doi.org/10.1016/j.gecco.2020.e00935> (2020).

34. Hill, S. L. L. *et al.* Measuring Forest Biodiversity Status and Changes Globally. *Front. For. Glob. Chang.* **2**, 1–11, <https://doi.org/10.3389/fgfc.2019.00070> (2019).
35. Burton, V. J. *et al.* Effects of land use and soil properties on taxon richness and abundance of soil assemblages. *Eur. J. Soil Sci.* **74**, 1–15, <https://doi.org/10.1111/ejss.13430> (2023).
36. Marquardt, S. G. *et al.* Identifying regional drivers of future land-based biodiversity footprints. *Glob. Environ. Chang.* **69**, 102304, <https://doi.org/10.1016/j.gloenvcha.2021.102304> (2021).
37. Mahlich, L., Jung, C. & Schaldach, R. The Biodiversity Footprint of German Soy-Imports in Brazil. *Sustainability* **14**, <https://doi.org/10.3390/su142316272> (2022).
38. Hurt, G. C. *et al.* Harmonization of global land use change and management for the period 850–2100 (LUH2) for CMIP6. *Geosci. Model Dev.* **13**, 5425–5464, <https://doi.org/10.5194/gmd-13-5425-2020> (2020).
39. Gassert, F., Mazzarello, J. & Hyde, S. *Global 100m Projections of Biodiversity Intactness for the Years 2017–2020 [White Paper]*. https://ai4edatasetspublicassets.blob.core.windows.net/assets/pdfs/io-biodiversity/Biodiversity_Intactness_whitepaper.pdf (2022).
40. Matej, S. *et al.* Options for reducing a city's global biodiversity footprint – The case of food consumption in Vienna. *J. Clean. Prod.* **437**, 140712, <https://doi.org/10.1016/j.jclepro.2024.140712> (2024).
41. Sanyé-Mengual, E., Biganzoli, F., Valente, A., Pfister, S. & Sala, S. What are the main environmental impacts and products contributing to the biodiversity footprint of EU consumption? A comparison of life cycle impact assessment methods and models. *Int. J. Life Cycle Assess.* **28**, 1194–1210, <https://doi.org/10.1007/s11367-023-02169-7> (2023).
42. Winkler, K., Fuchs, R., Rounsevell, M. & Herold, M. Global land use changes are four times greater than previously estimated. *Nat. Commun.* **12**, 1–10, <https://doi.org/10.1038/s41467-021-22702-2> (2021).
43. Ganzenmüller, R. *et al.* Land-use change emissions based on high-resolution activity data substantially lower than previously estimated. *Environ. Res. Lett.* **17**, <https://doi.org/10.1088/1748-9326/ac70d8> (2022).
44. Friedl, M. & Sulla-Menache, D. *MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006 [Data Set]*, <https://doi.org/10.5067/MODIS/MCD12Q1.006> (2019).
45. Ramankutty, N., Evan, A. T., Monfreda, C. & Foley, J. A. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Global Biogeochem. Cycles* **22**, 1–19, <https://doi.org/10.1029/2007GB002952> (2008).
46. Parente, L. *et al.* Annual 30-m maps of global grassland class and extent (2000–2022) based on spatiotemporal Machine Learning. *Sci. Data* **11**, 1–22, <https://doi.org/10.1038/s41597-024-04139-6> (2024).
47. Lesiv, M. *et al.* Global forest management data for 2015 at a 100 m resolution. *Sci. Data* **9**, 1–14, <https://doi.org/10.1038/s41597-022-01332-3> (2022).
48. Potapov, P. *et al.* The last frontiers of wilderness: Tracking loss of intact forest landscapes from 2000 to 2013. *Sci. Adv.* **3**, 1–13, <https://doi.org/10.1126/sciadv.1600821> (2017).
49. Mu, H. *et al.* A global record of annual terrestrial Human Footprint dataset from 2000 to 2018. *Sci. Data* **9**, 1–9, <https://doi.org/10.1038/s41597-022-01284-8> (2022).
50. Venter, O. *et al.* Global terrestrial Human Footprint maps for 1993 and 2009. *Sci. Data* **3**, 1–10, <https://doi.org/10.1038/sdata.2016.67> (2016).
51. Hudson, L. *et al.* *The 2016 Release of the PREDICTS Database V1.1 [Data Set]*. <https://doi.org/10.5519/j4sh7e0w> (2023).
52. Contu, S. *et al.* *Release of Data Added to the PREDICTS Database (November 2022) [Data Set]*. <https://doi.org/10.5519/jg7i52dg> (2022).
53. Luck, G. W. A review of the relationships between human population density and biodiversity. *Biol. Rev.* **82**, 607–645, <https://doi.org/10.1111/j.1469-185X.2007.00028.x> (2007).
54. WorldPop. *Global High Resolution Population Denominators Project*. www.worldpop.org <https://doi.org/10.5258/SOTON/WP00660> (2018).
55. Nguyen, C. T. *et al.* A regional assessment of ecological environment quality in Thailand special economic zone: Spatial heterogeneous influences and future prediction. *L. Degrad. Dev.* **34**, 5770–5787, <https://doi.org/10.1002/ldr.4876> (2023).
56. Meijer, J. R., Huijbregts, M. A. J., Schotten, K. C. G. J. & Schipper, A. M. Global patterns of current and future road infrastructure. *Environ. Res. Lett.* **13**, <https://doi.org/10.1088/1748-9326/aabd42> (2018).
57. Weiss, D. J. *et al.* A global map of travel time to cities to assess inequalities in accessibility in 2015. *Nature* **553**, 333–336, <https://doi.org/10.1038/nature25181> (2018).
58. You, L., Wood, S., Wood-Sichra, U. & Wu, W. Generating global crop distribution maps: From census to grid. *Agric. Syst.* **127**, 53–60, <https://doi.org/10.1016/j.agsy.2014.01.002> (2014).
59. International Food Policy Research Institute & International Institute for Applied Systems Analysis. *Global Spatially-Disaggregated Crop Production Statistics Data for 2005 Version 3.2*, <https://doi.org/10.7910/DVN/DHXBIX> (2016).
60. International Food Policy Research Institute. *Global Spatially-Disaggregated Crop Production Statistics Data for 2010 Version 2.0*, <https://doi.org/10.7910/DVN/PRFF8V> (2019).
61. International Food Policy Research Institute. *Global Spatially-Disaggregated Crop Production Statistics Data for 2020 Version 1.0.0*, <https://doi.org/10.7910/DVN/SWPENT> (2024).
62. Gilbert, M. *et al.* Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Sci. Data* **5**, 1–11, <https://doi.org/10.1038/sdata.2018.227> (2018).
63. Li, F., Wu, S., Liu, H. & Yan, D. Biodiversity loss through cropland displacement for urban expansion in China. *Sci. Total Environ.* **907**, 167988, <https://doi.org/10.1016/j.scitotenv.2023.167988> (2024).
64. Fox, J. & Weisberg, S. An R companion to applied regression (Third Edition). *Sage Publ.* (2018).
65. Bates, D., Mächler, M., Bolker, B. M. & Walker, S. C. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, <https://doi.org/10.18637/jss.v067.i01> (2015).
66. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* **82**, 1–26 (2017).
67. Medková, H., Vačkář, D. & Weinzel, J. Appropriation of potential net primary production by cropland in terrestrial ecoregions. *J. Clean. Prod.* **150**, 294–300, <https://doi.org/10.1016/j.jclepro.2017.03.002> (2017).
68. Weinzel, J., Vačkář, D. & Medková, H. Human footprint in biodiversity hotspots. *Front. Ecol. Environ.* **16**, 447–452, <https://doi.org/10.1002/fee.1825> (2018).
69. Nguyen, C. T., Vačkářová, D. & Weinzel, J. Consistent Global Datasets on Land Use, Biodiversity Intactness Index, and Biodiversity Intactness Footprint of Agricultural Production from 2000 to 2020. *Figshare* <https://doi.org/10.6084/m9.figshare.28303442> (2025).
70. Hoskins, A. J. *et al.* Downscaling land-use data to provide global 30'' estimates of five land-use classes. *Ecol. Evol.* **6**, 3040–3055, <https://doi.org/10.1002/ece3.2104> (2016).
71. Oliver, T. H. & Morecroft, M. D. Interactions between climate change and land use change on biodiversity: Attribution problems, risks, and opportunities. *Wiley Interdiscip. Rev. Clim. Chang.* **5**, 317–335, <https://doi.org/10.1002/wcc.271> (2014).
72. Sanchez-Ortiz, K., Newbold, T., Purvis, A. & De Palma, A. Global Maps of Biodiversity Intactness Index (Sanchez-Ortiz *et al.*, 2019 - BioRxiv). *Dataset* <https://doi.org/10.6084/m9.figshare.7951415.v1> (2019).
73. Harwood, T. *et al.* *BHI v2: Biodiversity Habitat Index: 30s Global Time Series. V1*. <https://doi.org/10.25919/3j75-f539> (2022).
74. Olson, D. M. *et al.* Terrestrial ecoregions of the world: A new map of life on Earth. *Bioscience* **51**, 933–938 (2001).

75. Hijmans, R. J. *terra*: Spatial Data Analysis. R package version 1.8-3. at <https://rspatial.github.io/terra/>, <https://rspatial.org/> (2024).
76. Hudson, L. N. *et al.* The database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) project. *Ecol. Evol.* 7, 145–188, <https://doi.org/10.1002/ece3.2579> (2017).

Acknowledgements

The research is supported by the Czech Science Foundation (grant no. 23-07984X: “Pathways towards Environmental Sustainability”). The authors also thank Laura Henderson and Nigel K. Downes for their meticulous proofreading.

Author contributions

J.W. and D.V. conceived the research through the initial project proposal. C.T.N., J.W. and D.V. developed the methodology. C.T.N. acquired and processed the data, conducted the formal analysis, and prepared the visualizations. All authors contributed to the discussion and interpretation of the results. C.T.N. prepared the manuscript. C.T.N., D.V. and J.W. edited the manuscript. All authors reviewed and approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41597-025-05901-0>.

Correspondence and requests for materials should be addressed to C.T.N.

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

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



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025