Global Ecology and Biogeography

WILEY

Global Ecology and Biogeography — A Journal of Macroecology

DATA ARTICLE OPEN ACCESS

SPECTRE: Standardised Global Spatial Data on Terrestrial SPecies and ECosystems ThREats

Vasco V. Branco[1,2] | César Capinha[3] | Jorge Rocha[3] | Luís Correia[2] | Pedro Cardoso[1,4]

[1]Laboratory for Integrative Biodiversity Research (LIBRe), Finnish Museum of Natural History Luomus, University of Helsinki, Helsinki, Finland | [2]LASIGE and Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, Lisboa, Portugal | [3]Centre of Geographical Studies, Institute of Geography and Spatial Planning, University of Lisbon, Lisboa, Portugal | [4]Centre for Ecology, Evolution and Environmental Changes (cE3c), Department of Animal Biology & CHANGE—Global Change and Sustainability Institute, University of Lisbon, Lisboa, Portugal

Correspondence: Vasco V. Branco (vasco.branco@helsinki.fi)

Received: 12 December 2023 | **Revised:** 26 August 2024 | **Accepted:** 30 November 2024

Handling Editor: Thiago Sanna Freire Silva

Funding: This work was supported by Fundação para a Ciência e a Tecnologia, Koneen Säätiö.

Keywords: ecosystem services | extinction risk | geographical data | global change | species distribution | threat assessment

ABSTRACT

Motivation: SPECTRE is an open-source database containing standardised spatial data on global environmental and anthropogenic variables that are potential threats to terrestrial species and ecosystems. Its goal is to allow users to swiftly access spatial data on multiple threats at a resolution of 30-arc seconds for all terrestrial areas. Following the standard set by Worldclim, these data allow full comparability and ease of use under common statistical frameworks for global change studies, species distribution modelling, threat assessments, quantification of ecosystem services and disturbance, among multiple other uses. A web user interface, a persistent online repository and an accompanying R package with functions for downloading and manipulating data are provided.

Main Types of Variable Contained: SPECTRE is a GIS product, currently with 21 geoTiff raster layers with an approximate 1×1 km resolution.

Spatial Location and Grain: Global (longitude −180–180, latitude −60–90) terrestrial database with a resolution of 30-arc seconds (approximately 1×1 km at the equator), converted from global sources of different original spatial grain, from $0.03\times0.03$to 10×10 km.

Time Period and Grain: The known time period for all sources present in SPECTRE varies from 1976 to 2020 (all but three after 1990), with a minimum temporal grain of 1 year.

Major Taxa and Level of Measurement: Non-taxa-specific.

Software Format: geoTiff and R.

1 | Background

Biodiversity and ecosystem services loss are ongoing issues that have taken the spotlight in several international agendas, such as the recently approved Kunming–Montreal Global Biodiversity Framework (Convention on Biological Diversity [CBD] 2023). Across all agendas, the need for direct conservation actions and policy according to the best available knowledge is recognised. However, the status of biodiversity and the services it provides to humanity remain unquantified for the most part. This is at least partly due to the lack of data, but often the culprit is the low accessibility and comparability of the many data sources already available.

Standardised data on several variables of interest are already available or being mobilised through different platforms. These include species distributions through GBIF (Global Biodiversity Information Facility 2022), and spatial or temporal trends of communities or populations through PREDICTS (Natural History Museum 2024; Hudson et al. 2016) and BioTIME (Dornelas et al. 2018) respectively. Variables representing potential threats to species are also available from various sources, such as forest loss from ForestWatch (Global Forest Watch 2014), or human density and built area from the Global Human Settlement Layer (Corbane et al. 2018; Schiavina, Freire, and MacManus 2019). Despite their availability, these sources are not directly comparable, that is, they are at different resolutions, projections and formatted in other ways that do not allow for immediate analysis. Due to this, users that intend to simultaneously analyse multiple threats will have to go through the task of standardising data across myriad formats, resolutions and extents.

Here, we present SPECTRE, a global spatial database of threats to terrestrial species and ecosystems, mainly intended towards global and large-scale analyses (see Section 3 for access details). Our goals are fourfold: (1) compiling global spatial layers of several types of threats; (2) standardising them to a common extent and resolution, identical to existing and widely used climate layers; (3) make them easily and openly available to researchers and decision-makers involved in species and ecosystem conservation and (4) provide an online portal together with the tools to easily download and manipulate different layers in the R environment.

With these goals in mind, we have chosen WorldClim (Fick and Hijmans 2017) as our template for SPECTRE. WorldClim is a longstanding, high-quality dataset of global climate layers and has been used extensively in different fields, including species distribution modelling (SDM), understanding the effects of climate change on species and more, with over 1000 citations per year. As such, by adopting WorldClim as our template, we guarantee the easy integration of SPECTRE into myriad studies involving both climate and biodiversity threats and straightforward comparisons between such studies. Our intention is to continuously update SPECTRE (see Section 3, 'Data accessibility & usage notes'), advancing as needed through community suggestions and the best available scientific knowledge on threats to species and ecosystems worldwide.

2 | Methods

2.1 | Description of Data

The first version of SPECTRE comprises a collection of 21 standardised raster layers, each representing a potential threat to species and ecosystems (Figure 1). The layers are classified first according to the list of direct pressures on biodiversity and sustainable use as proposed in the CBD (2020a, 2020b) in their Strategic Goal B (Table 1). Each is then further divided into subsections according to the International Union for the Conservation of Nature (IUCN) threat categories (IUCN 2023). Every layer and its metadata are listed in Table 2.

All layers are based on data collected between 1976 and 2020 (all but three after 1990) but these data do not span the entire time span for all layers. Moreover, we anticipate that threats will exhibit spatial and temporal correlations on a global scale. This means that the regions heavily impacted by a particular threat in 1 year or decade are likely to be affected in other years or decades by the same threat (at a global scale). Therefore, layers covering different time spans may still be directly comparable. Information on layer time grain can be consulted in Table 2. In that table, the time interval of data collection (i.e., first and last years) is recorded under 'Temporal range'. When the temporal range does not match the number of years with data, 'multiple

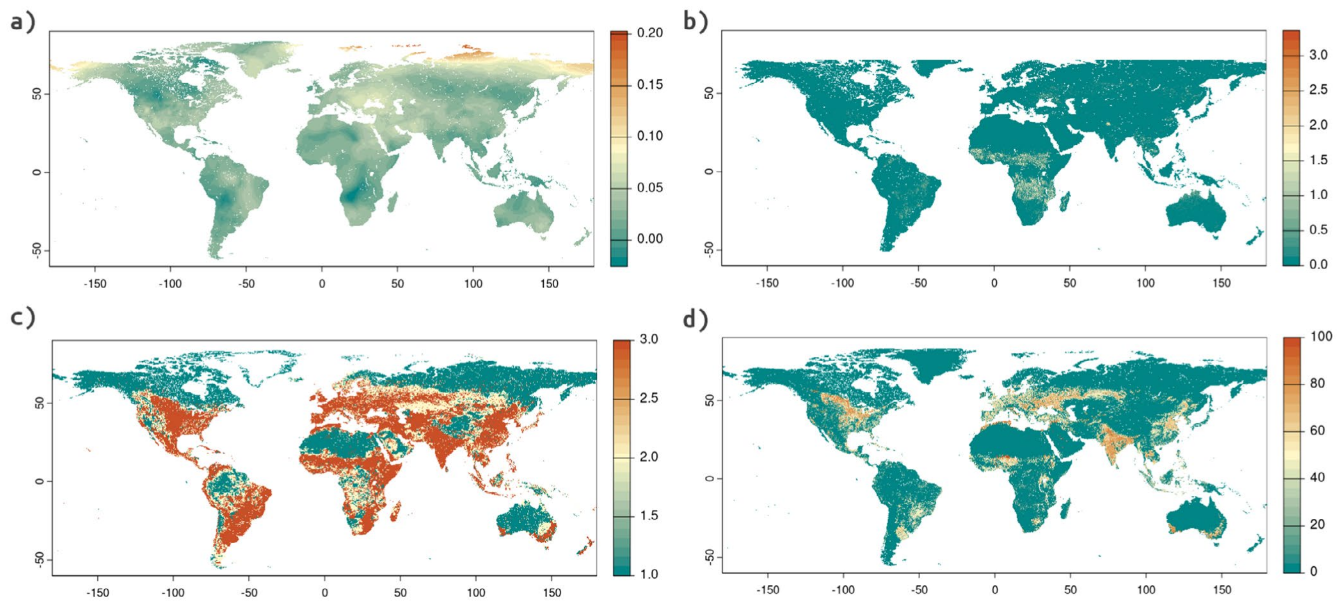

FIGURE 1 | Example global raster layers for (a) temperature trends (*5_1_TEMP_TRENDS*), (b) fire occurrence (*1_10_FIRE_OCCUR*), (c) impact area (*1_7_IMPACT_AREA*) and (d) crop percentage (*2_2_CROP_PERC_IIASA*). Layers in SPECTRE represent common threats to species and ecosystems, both directly as the destruction of species habitats, as well as indirectly, such as projected surface temperature changes for a given area. Note that some areas, such as polar regions, are not represented in all layers, depending on the original source.

TABLE 1 | Threat categories associated with each SPECTRE layer currently included in the database (see Table 2), as defined by the Convention on Biological Diversity and the International Union for the Conservation of Nature (IUCN).

CBD	IUCN	Description (from IUCN)	ID	Layer
Habitat loss	Residential & commercial development	Housing, urban areas, commercial and industrial areas, tourism (the areas themselves) and recreational areas	1.4	BUILT_AREA
			1.6	FOOTPRINT_PERC
			1.7	IMPACT_AREA
			1.8	MODIF_AREA
			1.9	HUMAN_BIOMES
	Agriculture & aquaculture	Physical and ecological (see Pollution for the chemical effects) effects of livestock and crop production, including wood and pulp plantations and non-timber crops, livestock farming and aquaculture both freshwater and marine	1.7	IMPACT_AREA
			1.8	MODIF_AREA
			1.9	HUMAN_BIOMES
			1.11	CROP_PERC_UNI
			1.12	CROP_PERC_IASA
			1.13	LIVESTOCK_MASS
			2.1	FOREST_LOSS_PERC
			2.2	FOREST_TREND
	Energy production & mining	Activities related to renewable and nonrenewable energy and goods such as oil, gas drilling, mining, quarrying, solar farms, windmill turbines and so on, and in some cases their necessities (power lines and pipelines fall both under this threat and Transportation & service corridors)	1.1	MINING_AREA
			1.8	MODIF_AREA
	Transportation & service corridors	Roads, railroads, utility and service lines, shipping lanes and flight paths	1.5	ROAD_DENSITY
			1.6	FOOTPRINT_PERC
			1.8	MODIF_AREA
	Human intrusions & disturbance	Recreational activities and work, as well as war, civil unrest and military exercises	1.3	HUMAN_DENSITY
			1.6	FOOTPRINT_PERC
			1.7	IMPACT_AREA
			1.8	MODIF_AREA
			1.9	HUMAN_BIOMES
	Natural system modifications	Fire and fire suppression, as well as dams and water management and use	1.7	IMPACT_AREA
			1.10	FIRE_OCCUR
	Geological events	The habitat loss effects caused by volcanoes, earthquakes, tsunamis, avalanches and landslides	1.2	HAZARD_ POTENTIAL
Overexploitation	Biological resource use	All biological harvest and nonproduction activities (gathering activities), including hunting, trapping and collecting terrestrial animals, gathering of terrestrial plants, logging, wood harvesting and fishing	1.7	IMPACT_AREA
			1.8	MODIF_AREA
			1.9	HUMAN_BIOMES
			1.13	LIVESTOCK_MASS
			2.1	FOREST_LOSS_PERC
			2.2	FOREST_TREND

(Continues)

TABLE 1 | (Continued)

CBD	IUCN	Description (from IUCN)	ID	Layer
Pollution	Pollution	Pollution includes domestic and urban wastewater, industrial, military, agricultural and forestry effluents, garbage, solid waste, air-borne pollutants and excess energy (aka: thermal pollution)	1.6	FOOTPRINT_PERC
			1.8	MODIF_AREA
			3.1	LIGHT_MCDM2
Invasive species	Invasives & other problematic species, genes & diseases	A broad label that includes problematic species and diseases caused by microorganisms or otherwise (unknown agents, viruses, prions, etc.) of either native, alien or unknown origin, as well as any potentially introduced genetic material	NA	NA
Climate change	Climate change & severe weather	The effects of climate-induced habitat shifts and alterations, as well as severe weather phenomena such as droughts, temperature extremes, storms and flooding	1.2	HAZARD_POTENTIAL
			5.1	TEMP_TRENDS
			5.2	TEMP_SIGNIF
			5.3	CLIM_EXTREME
			5.4	CLIM_VELOCITY
			5.5	ARIDITY_TREND

Note: Categories for which no data meeting our criteria for inclusion was found are denoted with 'NA' in columns 'ID' and 'Layer'.

ranges' are added to 'Temporal range' to denote this. The years in that interval for which the authors provide data are recorded under '#years with data'. For layers where the authors did not provide a year for the data, the number of years with data is classified as 'unknown'. It is important to note that SPECTRE does not include any time-series data. When time series data were available (e.g., Keys, Barnes, and Carter 2021; Mu et al. 2022), we calculated an aggregate measure to reflect the entire period depending on the nature of the data, such as sum, difference, average, harmonisation or slope of the time series: FOREST_TREND, FOREST_LOSS_PERC (see also Table 2). Any other transformations of time series data were done by the authors of each layer. We plan to expand layers in the near future as further compatible ones are made available. In terms of time frames, we consider including layers with time series, extending SPECTRE to support temporal analysis of ecosystem threats.

The criteria for inclusion of layers in SPECTRE are: (1) being publicly available, with a licence that allows open sharing under a Creative Commons CC BY 4.0 licence, or under similar public domain noncommercial licences; (2) being global in its intent and scope (see exceptions under Section 2.2, 'Data transformation' and Section 3, 'Data accessibility & usage notes'); (3) having a minimum spatial resolution of 10×10 km, as it is well established in GIS literature and is used both in WorldClim and in the reference grid of the European Environment Agency, in other large-scale initiatives (e.g., Ford et al. 2020; Lewthwaite and Mooers 2021; Fischer, Walentowitz, and Beierkuhnlein 2022; Silveira et al. 2022; European Environment Agency 2024) and (4) providing the best available products for each threat category according to our own judgement (see a complete list of datasets considered in the process and the rationale behind their exclusion in Table S2). We excluded products that do not present numeric quantitative data or a numeric scale of threat to species or ecosystems. For example, a layer classifying terrain into habitat types such as 'cropland' or 'urban area' cannot be converted into a gradient of threat without a significant interpretative and often subjective effort (e.g., ECOCLIMAP, see Table S2). All layers in SPECTRE were created using R (R Core Team 2023), namely, the package terra (Hijmans et al. 2023) and the software QGIS, version 3.10.11 (QGIS Association 2021). All layers were encoded as FLT4S, reclassified to have -3.4×10^{38} as their NA value and finally masked in the areas considered as NA in WorldClim. Layers requiring either resampling or reprojection were transformed using terra's 'resample' or 'project' functions respectively. The 'resample' function transfers and estimates cell values between rasters with different origins or resolutions, while 'project' also changes the coordinate reference system to match the original layer. In both cases, the resampling method used was the bilinear method for continuous data and the nearest neighbour method for categorical data (Arif and Akbar 2005; ArcGIS Pro 2023). It should be noted that, by using this resampling strategy, we are artificially increasing the data resolution and users should be aware of the original resolution of the different layers. This decision was made to avoid data contamination, which happens when using high-resolution layers to statistically increase the resolution of lower resolution layers, inevitably forcing high correlation between them. Furthermore, the resampling method used is of comparable performance to more advanced model downscaling methods, as found in the systematic review by Li and Heap (2011), with the advantage of being computationally efficient and maintaining spatial continuity of the original data. Nevertheless, the downscaling procedure was necessary for eight layers only. Five of these are the climate change variables which are highly spatially correlated and hence the downscaling can be performed with high confidence. The remaining three layers (HAZARD_POTENTIAL,

TABLE 2 | Characteristics of SPECTRE GIS original layers.

ID	Layer name	Overview	Unit	Temporal range	#years with data	Temporal resolution	Original spatial resolution (km)	Source
1.1	MINING_AREA	Number of mines on a 50-cell radius (sum of multiple years)	[NA]	Undefined range	Unknown	[NA}	1×1	Sonter et al. (2020)
1.2	HAZARD_ POTENTIAL	Number of significant hazards (e.g., earthquakes, floods, volcanoes) potentially affecting cells based on hazard frequency data (sum of multiple years)	[NA]	Multiple ranges (1976–2005), some undefined	Unknown	[NA}	5×5	SEDAC (2005)
1.3	HUMAN_DENSITY	Population density (value of single year)	[persons/km^2]	2015	1	Year	1×1	Schiavina, Freire, and MacManus (2019)
1.4	BUILT_AREA	Proportion of surface area covered by buildings (value of single year)	[%]	2014	1	Year	1×1	Corbane et al. (2018)
1.5	ROAD_DENSITY	Road density (including dirt, gravel, paved, etc.) (harmonisation of multiple years)	[m/km^2]	1997–2015	15	[NA]	8×8	Meijer et al. (2018)
1.6	FOOTPRINT_PERC	Weighted average of anthropogenic impacts on the environment, including population density, electric power infrastructure, crop lands and others (harmonisation of multiple years)	[%]	Multiple ranges (1990–2010)	7	[NA]	1×1	SEDAC (2018)
1.7	IMPACT_AREA	Classification of land into very low-impact areas (1), low impact areas (2) and non-low-impact areas (3) (harmonisation of multiple years)	[NA]	Multiple ranges (2000–2018)	17	[NA]	1×1	Jacobson et al. (2019)
1.8	MODIF_AREA	Proportion of a landscape that has been modified (harmonisation of multiple years)	[NA]	Multiple ranges (1980–2016)	34	[NA]	1×1	Kennedy et al. (2019)
1.9	HUMAN_BIOMES	Ordination of anthropogenic biomes of differing pressure: wildlands (0), forested (1), rangeland (2), croplands (3), villages (4) and dense settlements (5) (harmonisation of multiple years)	[NA]	Multiple ranges (2001–2006)	Unknown	[NA]	10×10	Ellis and Ramankutty (2008)

(Continues)

TABLE 2 | (Continued)

ID	Layer name	Overview	Unit	Temporal range	#years with data	Temporal resolution	Original spatial resolution (km)	Source
1.10	FIRE_OCCUR	Number of fire occurrences between the years 2006 and 2016, calculated by overlaying ignition polygons built from known centre and radius (sum of multiple years)	[NA]	2006–2016	10	Year	Point data, converted to 1 × 1	Andela et al. (2019)
1.11	CROP_PERC_UNI	Proportion of cropland (harmonisation of multiple years)	[%]	Multiple ranges (1990–2014)	14	[NA]	0.25 × 0.25	Waldner et al. (2016)
1.12	CROP_PERC_IASA	Proportion of cropland (average of multiple years)	[%]	Multiple ranges (1986–2009), some unknown	Unknown	Year	1 × 1	Fritz et al. (2015)
1.13	LIVESTOCK_MASS	Estimated total amount of livestock wet biomass based on global livestock head counts (estimate of single year)	[kg]	2006	1	Year	1 × 1	Robinson et al. (2014)
2.1	FOREST_ LOSS_PERC	Proportion of forest tree cover loss between 2007 and 2017 (difference between 2 years)	[%]	2007, 2017	2	Year	1 × 1	Shimada et al. (2014)
2.2	FOREST_TREND	Number of years (0–19) where a loss of forest was detected (sum of multiple years)	[NA]	2001–2019	19	Year	0.03 × 0.03	Global Forest Watch (2014)
3.1	LIGHT_MCDM2	Simulated upward (measured from the zenith) radiance, trained with satellite low-light imaging data and calibrated with radiance data collected by citizen scientists (harmonisation of multiple years)	[mcd/m^2]	2010–2020	Unknown	[NA]	1 × 1	Falchi et al. (2016)
5.1	TEMP_TRENDS	Linear regression coefficients of mean monthly temperature (slope of multiple years)	[NA]	1990–2019	29	Month	5.5 × 5.5	Wan, Hook, and Hulley (2015)
5.2	TEMP_SIGNIF	Temperature trends (5.1) divided by their standard error (using multiple years)	[NA]	1990–2019	29	Month	5.5 × 5.5	Wan, Hook, and Hulley (2015)

(Continues)

TABLE 2 | (Continued)

ID	Layer name	Overview	Unit	Temporal range	#years with data	Temporal resolution	Original spatial resolution (km)	Source
5.3	CLIM_EXTREME	The largest of the absolute of the trend coefficients (5.1) of the months with the lowest or highest mean temperatures (using multiple years)	[NA]	1990–2019	29	Month	5.5×5.5	Harris, Jones, and Osborn (2020)
5.4	CLIM_VELOCITY	Ratio between temperature trends (5.1) and a local spatial gradient in mean temperature calculated as the slope of a plane fitted to the values of a 3×3 cell neighbourhood centred on each cell (using multiple years)	[NA]	2000–2020	21	Month	5.5×5.5	Wan, Hook, and Hulley (2015)
5.5	ARIDITY_TREND	Linear regression coefficients of aridity for the years 1990–2019, that is, MPET/(MPRE+1), with MPET being the monthly potential evapotranspiration of each pixel and MPRE the monthly precipitation (slope of multiple years)	[NA]	1990–2019	29	Month	5.5×5.5	Harris, Jones, and Osborn (2020)

Note: Cells marked as 'Unknown' in '#years with data' represent cases where the authors did not specify which years within the range have available data. As an example, someone interested in crop percentage might choose to read the entry for CROP_PERC_IIASA and be able to quickly verify that it has % as its 'Unit'. After consideration, the user decides to use CROP_PERC_UNI instead as we know which years have data, 14, from its temporal range from 1990 to 2014.

ROAD_DENSITY and HUMAN_BIOMES) are the only for which some caution in their use should be exercised.

We also filled data gaps in cells classified as land by WorldClim that were assigned a null value during the interpolation process. These gaps mainly occurred in cells at the land-ocean interface, where the spatial extent of the interpolated variables was smaller than that of the WorldClim land mask. To address this, we first identified the target cells, defined as those classified as land by WorldClim that showed 'NA' in the interpolated surface and had at least one valid value among their eight nearest neighbours. We then estimated the values for these cells using four different interpolation methods: a simple average of the eight nearest neighbours (i.e., within a 3×3 cell neighbourhood), a simple average of the 24 nearest neighbours (5×5 cell neighbourhood) and their inverse distance-weighted (IDW) equivalents. For each layer, we assessed the accuracy of these methods through cross-validation by comparing the estimated values with those of cells that had known values. The method that produced the lowest mean absolute error was selected to generate the final interpolated values. The most accurate method in all layers was IDW (3×3) except for HUMAN_DENSITY, with the latter using the neighbour average (3×3). The exceptions to this procedure are a few very large areas (e.g., Greenland) or small isolated islands (e.g., South Sandwich Islands) that in some cases do not have neighbouring data for a reasonable extrapolation.

As with any GIS product, SPECTRE has several limitations due to the nature of its source data. For example, layers with an original resolution close to 10 km are usually not suitable for small-scale analyses, such as for small countries or islands. Prospective users should carefully evaluate any limitations and assumptions involved in creating these layers to determine whether the information meets their needs and to understand how these assumptions may affect their analyses. Please consult the metadata for each layer (Table 2) and consider consulting also the source of each layer to assess its appropriateness before use. Additionally, please consult Section 3, 'Data accessibility & usage notes' before using SPECTRE.

2.2 | Data Transformation

We transformed each layer to the same resolution, datum and projection as WorldClim (Fick and Hijmans 2017) so that SPECTRE and WorldClim can be used simultaneously without further adjustments from the user. For all resampling performed, bilinear interpolation was used when dealing with continuous data and nearest neighbour for discrete data (see Section 2.1, 'Description of data'). Additional interpolation was performed for a restricted number of cells receiving NA values but identified as land in WorldClim, mainly located along shorelines. This required different manipulation of data to convert every layer to a 30-arc second resolution (circa 1 km2 at the equator) covering the extent between longitudes −180° and 180° and latitudes between −60° and 90°. The 1×1 km resolution is the most convenient scale compromise for global data representation and analysis in ecology, following the approach set by Ellis and Ramankutty (2008). The temporal scope varied between 1976 and 2020 depending on the availability of information, although *HAZARD_POTENTIAL* uses older historical information. We also provide metadata for all layers, including the original citations (Table 2). Across all layers, larger values represent larger impact, with negative values representing a possibly positive impact (e.g., negative forest loss represents a gain in forest area).

We took measures to reduce incongruence between raster layers, due to missing or out of bounds data associated with water bodies. For this, we used both vector layers of the Global Lakes and Wetlands Database (Lehner and Döll 2004), which contains the shoreline polygons of the world's lakes and reservoirs. We then selected all bodies of water with an area larger than 1 km^2, rasterised the file and masked our WorldClim template with it, assigning all corresponding pixels as NA.

2.2.1 | General Procedures

A vast proportion of layers needed only few processing steps, such as resampling to the 1×1 km standard (*FOREST_TREND*, *FOOTPRINT_PERC*, *LIGHT_MCDM2*, *ROAD_DENSITY*, *CROP_PERC_UNI*, *CROP_PERC_IIASA*, *HUMAN_BIOMES*), reprojection (*HUMAN_DENSITY*, *BUILT_AREA*, *MODIF_AREA*) or masking to terrestrial areas (*HAZARD_POTENTIAL*).

We created *FOREST_TREND* using forest loss data from the Global Forest Watch (2014), available for download in spatial segments. These were collected and individually rescaled. *FOOTPRINT_PERC* (SEDAC 2018) is an index percentage metric indicating anthropogenic impacts on the environment in 2009, created by collating data on human population density, human land use, built-up areas, nighttime lights, land use/land cover and human access (coastlines, roads, railroads, navigable rivers) with a discriminate scoring system. *LIGHT_MCDM2* (Falchi et al. 2016) is a measure of continuous simulated upward (measured from the zenith) radiance, trained with satellite low-light imaging data and calibrated with radiance data collected by citizen scientists. It takes into account all sources contributing to it in a radius of 200 km (representing the impact of medium to long-range effects). *ROAD_DENSITY* (Meijer et al. 2018) is a continuous metric of road density in metres per square kilometre, resulting from the collation of 60 geospatial datasets on road infrastructure with more global coverage than similar products. *CROP_PERC_UNI* and *CROP_PERC_IIASA* (Fritz et al. 2015; Waldner et al. 2016) are both percentage metrics indicating the proportion of cropland in each cell, differing in temporal grain (1990–2014 representing a 2005 baseline year in the absence of complete data for that year) and methodology (adequacy analysis and crowd-sourcing & agreeableness respectively). *HUMAN_BIOMES* (Ellis and Ramankutty 2008) classifies terrain into different anthropogenic biomes of increasing anthropogenic pressure. The original dataset (see Table S1) classified cells into two hierarchical levels with 20 different categories in the lower level. We opted for the broader level with six categories: wildlands (0), forested (1), rangeland (2), croplands (3), villages (4) and dense settlements (5).

HUMAN_DENSITY (Schiavina, Freire, and MacManus 2019) and *BUILT_AREA* (Corbane et al. 2018) are continuous metrics of human density for the year 2015, expressed in people per km^2 and percentage of built-up presence respectively. Lastly, *MODIF_AREA* (Kennedy et al. 2019) is a continuous metric between 0 and 1 that reflects the proportion of a landscape that has been modified by collating remote sensing imagery and ground inventory data from 1980 to 2016 (with the mode year of 2016, as this was the intended snapshot year, past data being used only when no 2016 data were available) on a group of five anthropogenic stressors: (1) human settlement (population density & built-up areas), (2) agriculture (cropland & livestock), (3) transportation (major roads, minor roads, two tracks (a kind of two direction dirt road created through pedestrian and vehicle passage) & railroads), (4) mining and energy production (mining, oil wells & wind turbines) and (5) electrical infrastructure (powerlines, nighttime lights).

Additionally, two layers also needed to be masked to terrestrial areas: *HAZARD_POTENTIAL* and *MINING_AREA*. *HAZARD_POTENTIAL* (SEDAC 2005) is a discrete metric of the number of significant hazards (aka: earthquakes, volcanoes, landslides, floods, droughts and cyclones) potentially affecting cells based on hazard frequency data, based on threat analyses of high temporal grain. *MINING_AREA* is a discrete metric of the number of mines (preoperational, operational, closed) in a 50-cell radius, as mining threats were found to extend to this range (Sonter et al. 2017), considering both exploration of materials critical to renewable energy technology and infrastructures, and other materials.

2.2.2 | Further Transformations

Some of the layers present in SPECTRE required extensive processing of the source data: *IMPACT_AREA*, *FOREST_LOSS_PERC*, *LIVESTOCK_MASS*, *FIRE_OCCUR*, *TEMP_TRENDS*, *TEMP_SIGNIF*, *CLIM_EXTREME*, *CLIM_VELOCITY* and *ARIDITY_TREND*.

In *IMPACT_AREA* (Jacobson et al. 2019), the original data were composed of two binary rasters which were combined in a single layer representing very low (1), low (2) and high (3) impact. The final result was reprojected to our standard. Similarly, *FOREST_LOSS_PERC* (Shimada et al. 2014) is the result of two rasters of tree cover proportion (0–1), for the years 2007 and 2017, the latter being subtracted from the former. The end result is a raster with positive percentages being tree cover losses and negative percentages being tree cover gains.

We created *LIVESTOCK_MASS* using a similar methodology to that described in Bar-On, Phillips, and Milo (2018). The average wet biomass values present in Dong et al. (2006) were used to infer average wet biomass values for developed, transitioning and developing economies and the inferred values were combined with a country polygon vector map (Natural Earth 2018) and a classification of world economies (United Nations 2006) and then applied to 2006 livestock count rasters from the Gridded Livestock of the World (GLW2, Robinson et al. 2014) for cattle, chickens, ducks, goats, pigs and sheep, to estimate total livestock biomass.

FIRE_OCCUR is a fire occurrence layer created by taking point ignition data (Andela et al. 2019) for the years 2006–2016, approximately 1 million points per year, and for each point drawing a polygon using the ignition's radius. The value of each pixel is then determined by counting the number of polygons overlapping that pixel.

We constructed five layers for climate change: *TEMP_TRENDS*, *TEMP_SIGNIF*, *CLIM_EXTREME*, *CLIM_VELOCITY* and *ARIDITY_TREND*, based on methodology described in Bowler et al. (2020), using two datasets, Wan, Hook, and Hulley (2015) and Harris, Jones, and Osborn (2020). All layers used monthly data collected between 1990 and 2019. *TEMP_TRENDS* is a continuous metric of temperature trends, based on the linear regression coefficients of mean monthly temperature. Similarly, *TEMP_SIGNIF* is a continuous metric of temperature trend significance, the temperature trends divided by its standard error. *CLIM_EXTREME* is a continuous metric calculated as the largest of the absolutes of the coefficients of two linear regressions, one of the months with the lowest mean temperatures, and another of the month with the highest mean temperatures.

CLIM_VELOCITY is a continuous 0–1 metric of the velocity of climate change calculated as the ratio between two normalised features, *TEMP_TRENDS* and a local spatial gradient in mean temperature. This spatial gradient uses MODIS/Terra Land Surface Temperature (Wan, Hook, and Hulley 2015) and is calculated through the average maximum technique, defined as the slope of a plane fitted to the values of a 3×3 cell neighbourhood centred on a grid point (ArcGIS Resources 2021). This reflects the speed at which an organism would have to move to track its current climatic range. To our knowledge, two studies have calculated climate velocity at a global scale. Loarie et al. 2009 used the WorldClim Version 1.4 Annual Mean Temperature and Total Annual Precipitation bioclimatic variables to estimate climate velocity for 2050–2100, based on the emissions scenario temporal gradient present in the Special Report on Emissions Scenarios (SRES) A1B (IPCC 2000). It has several differences to our layer, which is expected given both the different temporal ranges as well as the diluted effects of temperature trends over three estimates (upper, mean and lower), accentuating patterns characteristic of the local spatial gradient used. More recently, Asamoah et al. (2022) have also produced climate velocity layers for both historic data and projected climate scenarios. The Spearman's Rho between Asamoah's and our layers is very close to 1.

The fifth layer, *ARIDITY_TREND*, is a continuous metric based on the linear regression coefficients of aridity, which is given by the expression:

$$\mathrm{ARIDITY\_TREND} = \mathrm{MPET} / (\mathrm{MPRE} + 1),$$

with MPET being the monthly potential evapotranspiration of each grid point and MPRE being the monthly precipitation.

3 | Data Accessibility and Usage Notes

SPECTRE is publicly accessible through several services. The Finnish IT Centre for Scientific Computing (CSC 2023) spatial data download service, Paituli (https://paituli.csc.fi/) allows for easy download of all layers in a shared zipped folder. The CSC is hosting SPECTRE for the foreseeable future. Layer updates will be made subsequently to their publication.

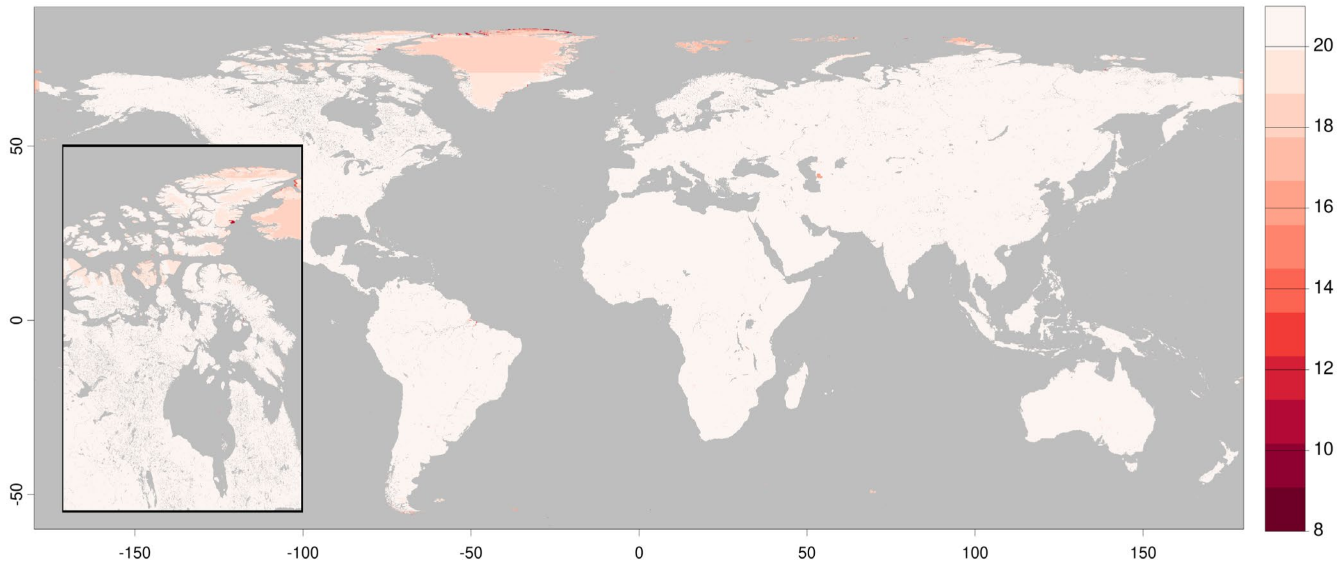

FIGURE 2 | Global map representing the number of layers with information for a given pixel in SPECTRE. On the bottom left corner, a close up of the Hudson bay and surrounding areas, highlighting the occasional lack of data in higher latitudes and coastal regions.

To consult all layers considered for inclusion in SPECTRE, their metadata and any additional minor usage notes, please consult Table S2. All data can also be accessed through the *gecko* R package (Branco, Correia, and Cardoso 2023). *gecko* is a collection of geographical analysis functions aimed primarily at ecology and conservation science studies, focusing on processing of both point and raster data. It also allows WMS and WCS request capabilities, making it a seamless transition from SPECTRE data access to processing. The entirety of our methodology, including the R scripts needed to recreate SPECTRE and the raster used in Figure 2 are found both on Github (https://github.com/VascoBranco/spectre.methodology) and Zenodo (https://zenodo.org/doi/10.5281/zenodo.10256190). Lastly, a dashboard allowing quick visualisation and download is also available on the project page (https://biodiversityresearch.org/spectre/). We have tested it over 10 development iterations with the help of 20 users and we are continuously improving it.

SPECTRE inherits current data availability and research biases, with entries being overrepresentative of those threats potentially categorised under habitat loss and fragmentation. Additionally, SPECTRE currently has layers representing threats fitting all categories of the classification system proposed (Table 1) except for 'Invasives & other problematic species, genes & diseases'. To the extent of our knowledge, there is no geospatial information representing the impact of invasive alien species on local ecosystems that meets our criteria for inclusion in either range or spatial resolution. Furthermore, we are also missing data on some aspects of categories considered. For example, under the IUCN threat 'Natural system modifications' we have data on fire occurrence which fits the current definition of 'Fire and fire suppression [..]' but we do not have data on dams or water management and use, also aspects of this category.

The layers also have specific limitations, namely, the error incurred through downscaling of the source data and the temporal and spatial grain of the sources used, limiting potential usage of the end product. The base 10 × 10 km resolution of some layers might make analyses on some microregions such as small islands unadvisable. As with all GIS products, keep in mind that full accuracy cannot be assured when running analyses for areas not much larger than the original resolution. It is, however, to the full extent of our knowledge, the best available information.

Whenever possible consider consulting the original metadata for each dataset for further details and be fully aware of how they were first built, the inherent limitations and if they serve your intended purpose. SPECTRE layers are thematically related (e.g., percentage of building surface area, road density) and several are composite and make use of the same source data in one way or another. This means that layers can be highly correlated. Be judicious in selecting your data and consider performing a correlation analysis in your study area.

Acknowledgements

We would like to thank Victor Cazalis and Luca Santini for their feedback on the SPECTRE dashboard. V.V.B. and P.C were supported by the Kone Foundation, Finland. L.C. was supported by Portuguese Foundation for Science and Technology (FCT) through LASIGE Research Unit, refs. UIDB/00408/2020 and UIDP/00408/2020. C.C. and J.R. acknowledge support from FCT awarded to the CEG/IGOT Research Unit (UIDB/00295/2020 and UIDP/00295/2020).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are openly available in Zenodo at https://doi.org/10.5281/zenodo.10256191, as well as other platforms including Github (https://github.com/VascoBranco/spectre.methodology) and Paituli (https://paituli.csc.fi/).

References

Andela, N., D. C. Morton, L. Giglio, et al. 2019. "The Global Fire Atlas of Individual Fire Size, Duration, Speed and Direction." *Earth System Science Data* 11, no. 2: 529–552. https://doi.org/10.5194/essd-11-529-2019.

ArcGIS Pro. 2023. "Resample function." Esri. Accessed November 24, 2023. https://pro.arcgis.com/en/pro-app/3.0/help/analysis/raster-functions/resample-function.htm.

Arif, F., and M. Akbar. 2005. "Resampling Air Borne Sensed Data Using Bilinear Interpolation Algorithm." *IEEE International Conference on Mechatronics* 2005: 62–65. https://doi.org/10.1109/ICMECH.2005.1529228.

Asamoah, E. F., M. Di Marco, J. E. M. Watson, L. J. Beaumont, O. Venter, and J. M. Maina. 2022. "Land-Use and Climate Risk Assessment for Earth's Remaining Wilderness." *Current Biology* 32, no. 22: 4890–4899. https://doi.org/10.1016/j.cub.2022.10.016.

Bar-On, Y., R. Phillips, and R. Milo. 2018. "The Biomass Distribution on Earth." *Proceedings of the National Academy of Sciences* 115, no. 25: 6506–6511. https://doi.org/10.1073/pnas.1711842115.

Bowler, D., A. Bjorkman, M. Dornelas, et al. 2020. "Mapping Human Pressures on Biodiversity Across the Planet Uncovers Anthropogenic Threat Complexes." *People and Nature* 2, no. 2: 380–394. https://doi.org/10.1002/pan3.10071.

Branco, V. V., L. Correia, and P. Cardoso. 2023. "Gecko: Geographical Ecology and Conservation Knowledge Online." R Package Version 1.0.0. https://cran.r-project.org/package=gecko.

Convention on Biological Diversity. 2020a. "Aichi Biodiversity Targets." CBD. Accessed November 12, 2020. https://www.cbd.int/sp/targets/.

Convention on Biological Diversity. 2020b. "Global Biodiversity Outlook 5." CBD. Accessed November 12, 2020. https://www.cbd.int/gbo/gbo5/publication/gbo-5-en.pdf.

Convention on Biological Diversity. 2023. "Kunming-Montreal Global Biodiversity Framework." CBD. Accessed November 15, 2023. https://www.cbd.int/gbf/.

Corbane, C., A. Florczyk, M. Pesaresi, P. Politis, and V. Syrris. 2018. "GHS Built-Up Grid, Derived From Landsat, Multitemporal (1975-1990-2000-2014)." *European Commission, Joint Research Centre (JRC)*. https://doi.org/10.2905/jrc-ghsl-10007.

CSC. 2023. "IT Center for Science." Accessed November 13, 2023. https://www.csc.fi/.

Dong, H., J. Mangino, T. Mcallister, and D. Have. 2006. "Emissions From Livestock and Manure Management (Intergovernmental Panel on Climate Change)." https://www.ipcc-nggip.iges.or.jp/public/2019rf/vol4.html.

Dornelas, M., L. H. Antão, F. Moyes, et al. 2018. "BioTIME: A Database of Biodiversity Time Series for the Anthropocene." *Global Ecology and Biogeography* 27, no. 7: 760–786. https://doi.org/10.1111/geb.12729.

Ellis, E. C., and N. Ramankutty. 2008. "Putting People in the Map: Anthropogenic Biomes of the World." *Frontiers in Ecology and the Environment* 6, no. 8: 439–447. https://doi.org/10.1890/070062.

European Environment Agency. 2024. "EEA Reference Grid." Accessed November 14, 2024. https://www.eea.europa.eu/en/datahub/datahubitem-view/3c362237-daa4-45e2-8c16-aaadfb1a003b.

Falchi, F., P. Cinzano, D. Duriscoe, et al. 2016. "The New World Atlas of Artificial Night Sky Brightness." *Science Advances* 2, no. 6. https://doi.org/10.1126/sciadv.1600377.

Fick, S. E., and R. J. Hijmans. 2017. "WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37, no. 12: 4302–4315. https://doi.org/10.1002/joc.5086.

Fischer, J., A. Walentowitz, and C. Beierkuhnlein. 2022. "The Biome Inventory – Standardizing Global Biogeographical Land Units." *Global Ecology and Biogeography* 31, no. 11: 2172–2183. https://doi.org/10.1111/geb.13574.

Ford, S. A., M. R. Jepsen, N. Kingston, et al. 2020. "Deforestation Leakage Undermines Conservation Value of Tropical and Subtropical Forest Protected Areas." *Global Ecology and Biogeography* 29, no. 11: 2014–2024. https://doi.org/10.1111/geb.13172.

Fritz, S., L. See, I. McCallum, et al. 2015. "Mapping Global Cropland and Field Size." *Global Change Biology* 21, no. 5: 1980–1992. https://doi.org/10.1111/gcb.12838.

Global Biodiversity Information Facility. 2022. "Global Biodiversity Information Facility." GBIF. Accessed May 26, 2022. https://www.gbif.org/.

Global Forest Watch. 2014. "World Resources Institute." Accessed 24 November 2023. http://www.globalforestwatch.org/.

Harris, I. C., P. D. Jones, and T. Osborn. 2020. "CRU TS4.04: Climatic Research Unit (CRU) Time-Series (TS) Version 4.04 of High-Resolution Gridded Data of Month-By-Month Variation in Climate (Jan. 1901–Dec. 2019)." Centre for Environmental Data Analysis. https://catalogue.ceda.ac.uk/uuid/89e1e34ec3554dc98594a5732622bce9.

Hijmans, R. J., E. Bivand, E. Pebesma, and M. D. Sumner. 2023. "Terra: Spatial Data Analysis." R Package Version 1.7-29. https://CRAN.R-project.org/package=terra.

Hudson, L. N., T. Newbold, S. Contu, et al. 2016. "The Database of the PREDICTS (Projecting Responses of Ecological Diversity in Changing Terrestrial Systems) Project." *Ecology and Evolution* 7, no. 1: 145–188. https://doi.org/10.1002/ece3.2579.

IPCC. 2000. "Summary for Policymakers. Emissions Scenarios: A Special Report of IPCC Working Group II. Intergovernmental Panel on Climate Change."

IUCN. 2023. "Threats Classification Scheme (Version 3.3)." IUCN Redlist. https://www.iucnredlist.org/about/barometer-of-life.

Jacobson, A. P., J. Riggio, A. M. Tait, and J. E. M. Baillie. 2019. "Global Areas of Low Human Impact ('Low Impact Areas') and Fragmentation of the Natural World." *Scientific Reports*. https://doi.org/10.5061/dryad.z612jm67g.

Kennedy, C. M., J. R. Oakleaf, D. M. Theobald, S. Baruch-Mordo, and J. Kiesecker. 2019. "Managing the Middle: A Shift in Conservation Priorities Based on the Global Human Modification Gradient." *Global Change Biology* 25, no. 3: 811–826. https://doi.org/10.1111/gcb.14549.

Keys, P. W., E. A. Barnes, and N. H. Carter. 2021. "A Machine-Learning Approach to Human Footprint Index Estimation With Applications to Sustainable Development." *Environmental Research Letters* 16, no. 4: 44061. https://doi.org/10.1088/1748-9326/abe00a.

Lehner, B., and P. Döll. 2004. "Development and Validation of a Global Database of Lakes, Reservoirs and Wetlands." *Journal of Hydrology* 296, no. 1–4: 1–22. https://doi.org/10.1016/j.jhydrol.2004.03.028.

Lewthwaite, J. M. M., and A. Ø. Mooers. 2021. "Geographical Homogenization but Little Net Change in the Local Richness of Canadian Butterflies." *Global Ecology and Biogeography* 31, no. 2: 266–279. https://doi.org/10.1111/geb.13426.

Li, J., and A. D. Heap. 2011. "A Review of Comparative Studies of Spatial Interpolation Methods in Environmental Sciences: Performance and Impact Factors." *Ecological Informatics* 6, no. 3–4: 228–241. https://doi.org/10.1016/j.ecoinf.2010.12.003.

Loarie, S. R., P. B. Duffy, H. Hamilton, G. P. Asner, C. B. Field, and D. D. Ackerly. 2009. "The Velocity of Climate Change." *Nature* 462: 1052–1055. https://doi.org/10.1038/nature08649.

Meijer, J. R., M. A. J. Huijbregts, K. C. G. J. Schotten, and A. M. Schipper. 2018. "Global Patterns of Current and Future Road Infrastructure."

Environmental Research Letters 13, no. 6: 64006. https://doi.org/10.1088/1748-9326/aabd42.

Mu, H., X. Li, Y. Wen, et al. 2022. "A Global Record of Annual Terrestrial Human Footprint Dataset From 2000 to 2018." *Scientific Data* 9: 176. https://doi.org/10.1038/s41597-022-01284-8.

Natural Earth. 2018. "Natural Earth Data (vector data v. 4.1.1)." Accessed November 24, 2023. https://www.naturalearthdata.com/.

Natural History Museum. 2024. "PREDICTS database." Accessed December 9, 2024. https://www.nhm.ac.uk/our-science/research/projects/predicts/science.html.

QGIS Association. 2021. "QGIS Geographic Information System." QGIS. http://www.qgis.org.

R Core Team. 2023. "R: A Language and Environment for Statistical Computing." R Foundation for Statistical Computing. https://www.R-project.org/.

ArcGIS Resources. 2021. "How Slope Works." Esri. Accessed November 24, 2023. https://resources.arcgis.com/en/help/main/10.1/index.html#//009z000000vz000000.

Robinson, T. P., G. R. W. Wint, G. Conchedda, et al. 2014. "Mapping the Global Distribution of Livestock." *PLoS One* 9, no. 5: e96084. https://doi.org/10.1371/journal.pone.0096084.

Schiavina, M., S. Freire, and K. MacManus. 2019. "GHS-Pop R2019A—GHS Population Grid Multitemporal (1975-1990-2000-2015)." *European Commission, Joint Research Centre (JRC)*. https://doi.org/10.2905/42E8BE89-54FF-464E-BE7B-BF9E64DA5218.

SEDAC. 2005. "Global Multihazard Frequency and Distribution, v1: Natural Disaster Hotspots." *NASA Socioeconomic Data and Applications Center*. https://doi.org/10.7927/H45718Z5.

SEDAC. 2018. "Global Human Footprint (Geographic), v3 (1995–2004): Last of the Wild." *NASA Socioeconomic Data and Applications Center*. https://doi.org/10.7927/H46T0JQ4.

Shimada, M., T. Itoh, T. Motooka, et al. 2014. "New Global Forest/Non-forest Maps From Alos Palsar Data (2007–2010)." *Remote Sensing of Environment* 155: 13–31. https://doi.org/10.1016/j.rse.2014.04.014.

Silveira, M. V. F., C. H. L. Silva-Junior, L. O. Anderson, and L. E. O. C. Aragão. 2022. "Amazon Fires in the 21st Century: The Year of 2020 in Evidence." *Global Ecology and Biogeography* 31, no. 10: 2026–2040. https://doi.org/10.1111/geb.13577.

Sonter, L. J., M. C. Dade, J. E. M. Watson, and R. K. Valenta. 2020. "Renewable Energy Production Will Exacerbate Mining Threats to Biodiversity." *Nature Communications* 11: 4174. https://doi.org/10.1038/s41467-020-17928-5.

Sonter, L. J., D. Herrera, D. J. Barrett, G. L. Galford, C. J. Moran, and B. S. Soares-Filho. 2017. "Mining Drives Extensive Deforestation in the Brazilian Amazon." *Nature Communications* 8, no. 1. https://doi.org/10.1038/s41467-017-00557-w.

United Nations. 2006. "World Economic Situation and Prospects 2006." United Nations New York. Accessed April 16, 2021. https://www.un.org/en/development/desa/policy/wesp/wesp_archive/2006wesp.pdf.

Waldner, F., S. Fritz, A. Di Gregorio, et al. 2016. "A Unified Cropland Layer at 250 m for Global Agriculture Monitoring." *Datamation* 1, no. 1: 3. https://doi.org/10.3390/data1010003.

Wan, Z., S. Hook, and G. Hulley. 2015. "MOD11C3 MODIS/Terra Land Surface Temperature/Emissivity Monthly L3 Global 0.05Deg CMG V006." NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD11C3.006.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.