Global Ecology and Biogeography

WILEY

Global Ecology and Biogeography — A Journal of Macroecology

DATA ARTICLE OPEN ACCESS

LandFrag: A Dataset to Investigate the Effects of Forest Loss and Fragmentation on Biodiversity

Thiago Gonçalves-Souza[1,2,3] | Maurício Humberto Vancine[4] | Nathan J. Sanders[2] | Nick M. Haddad[5] | Lucas Cortinhas[6] | Anne Lene T. O. Aase[7] | Willian Moura de Aguiar[8] | Marcelo Adrian Aizen[9] | Víctor Arroyo-Rodríguez[10,11] | Arturo Baz[12] | Maíra Benchimol[13] | Enrico Bernard[14] | Tássia Juliana Bertotto[15] | Arthur Angelo Bispo[16] | Juliano A. Bogoni[17] | Gabriel X. Boldorini[3] | Cibele Bragagnolo[18] | Berry Brosi[19] | Aníbal Silva Cantalice[3] | Rodrigo Felipe Rodrigues do Carmo[3,20] | Eliana Cazeta[13] | Adriano G. Chiarello[21] | Noé U. de la Sancha[22,23] | Raphael K. Didham[24,25] | Deborah Faria[13] | Bruno Filgueiras[3] | José Eugênio Côrtes Figueira[26] | Gabriela Albuquerque Galvão[3] | Michel Varajão Garey[27] | Heloise Gibb[28] | Carmelo Gómez-Martínez[29] | Ezequiel González[30] | Reginaldo Augusto Farias de Gusmão[3] | Mickaël Henry[31] | Shayana de Jesus[16] | Thiago Gechel Kloss[32] | Amparo Lázaro[29] | Victor Leandro-Silva[3] | Marcelo G. de Lima[33] | Ingrid da Silva Lima[3] | Ana Carolina B. Lins-e-Silva[3,20] | Ralph Mac Nally[34] | Arthur Ramalho Magalhães[3] | Luiz Fernando Silva Magnago[35] | Shiiwua Manu[36] | Eduardo Mariano-Neto[37] | David Nyaga Mugo Mbora[38,39] | Felipe P. L. Melo[40,41] | Morris Nzioka Mutua[42] | Selvino Neckel-Oliveira[43] | André Nemésio[44] | André Amaral Nogueira[45] | Patricia Marques Do A. Oliveira[46] | Diego G. Pádua[47] | Luan Paes[48] | Aparecida Barbosa de Paiva[49] | Marcelo Passamani[50] | João Carlos Pena[51] | Carlos A. Peres[52] | Bruno X. Pinho[53] | Jean-Marc Pons[54] | Victor Mateus Prasniewski[55] | Jenny Reiniö[56] | Magda dos Santos Rocha[57] | Larissa Rocha-Santos[13] | Maria J. Rodal[20] | Rodolpho Credo Rodrigues[58] | Nathalia V. H. Safar[59] | Renato P. Salomão[60] | Bráulio A. Santos[61] | Mirela N. Santos[62] | Jessie Pereira dos Santos[63] | Sini Savilaakso[64] | Carlos Ernesto Gonçalves Reynaud Schaefer[65] | Maria Amanda Menezes Silva[66] | Fernando R. da Silva[67] | Ricardo J. Silva[68] | Marcelo Simonelli[69] | Alejandra Soto-Werschitz[70] | John O. Stireman III[71] | Danielle Storck-Tonon[72] | Neucir Szinwelski[48] | Marcelo Tabarelli[41] | Camila Palhares Teixeira[73] | Ørjan Totland[74] | Marcio Uehara-Prado[75] | Fernando Zagury Vaz-de-Mello[76] | Heraldo L. Vasconcelos[43] | Simone A. Vieira[77] | Jonathan M. Chase[78,79]

Correspondence: Thiago Gonçalves-Souza (tgoncalv@umich.edu)

Received: 19 July 2024 | **Revised:** 6 February 2025 | **Accepted:** 16 February 2025

Handling Editor: Thiago Sanna Freire Silva

Funding: This work was supported by Institute for Global Change Biology.

Keywords: biodiversity conservation | biodiversity loss | functional traits | habitat fragmentation | habitat loss | landscape metrics

ABSTRACT

Motivation: The accelerated and widespread conversion of once continuous ecosystems into fragmented landscapes has driven ecological research to understand the response of biodiversity to local (fragment size) and landscape (forest cover and fragmentation) changes. This information has important theoretical and applied implications, but is still far from complete. We compiled the most comprehensive and updated database to investigate how these local and landscape changes determine species composition, abundance and trait diversity of multiple taxonomic groups in forest fragments across the globe.

Main Types of Variables Contained: We gathered data for 1472 forest fragments, providing information on the abundance and composition of 9154 species belonging to vertebrates, invertebrates, and plants. For 2703 of these species, we obtained more

Thiago Gonçalves-Souza and Maurício Humberto Vancine should be considered joint first author.

Nathan J. Sanders and Jonathan M. Chase should be considered joint senior author.

For affiliations refer to page 8.

than 20 functional traits. We provided the spatial location and size of each fragment and metrics of landscape composition and configuration.

Spatial Location and Grain: The dataset includes 1472 forest fragments sampled in 121 studies from all continents except Antarctica. Most datasets (77%) are from tropical regions, 17% are from temperate regions, and 6% are from subtropical regions. Species abundance and composition were collected at the plot or fragment scale, whereas the landscape metrics were extracted with buffer size ranging from a radius of 200–2000 m.

Time Period and Grain: Data on the abundance of species and community composition were collected between 1994 and 2022, and the landscape metrics were extracted from the same year that a given study collected the abundance and composition data.

Major Taxa and Level of Measurement: The studied organisms included invertebrates (Arachnida, Insecta and Gastropoda; 41% of the datasets), vertebrates (Amphibia, Squamata, Aves and Mammalia; 44%), and vascular plants (19%), and the lowest level of identification was species or morphospecies.

Software Format: The dataset and code can be downloaded on Zenodo or GitHub.

1 | Introduction

The growing expansion of agriculture and infrastructure is causing the annual loss of millions of hectares of forest worldwide, especially in the tropics (Global Forest Watch 2024). These massive land-use changes have caught the attention of ecologists and conservation biologists because they are shaping ecological patterns and processes across different spatial scales (Lôbo et al. 2011; Tscharntke et al. 2012; Haddad et al. 2015; Arroyo-Rodríguez et al. 2020; Hansen et al. 2020). In fact, ecological theory postulates that in human-modified landscapes, the structure of species assemblages can depend on both local—(e.g., fragment size) and landscape-scale factors (e.g., forest cover and fragmentation), but their relative role remains debated (e.g., Fletcher et al. 2018; Fahrig et al. 2019). For instance, the habitat amount hypothesis predicts that species density in a given site is more strongly and positively related to the amount of available habitat in the surrounding landscape than to the size of the habitat fragment within which the site is located. Although some studies support this prediction (Watling et al. 2020), others suggest that species density is weakly related to habitat amount (e.g., Martínez-Ruiz et al. 2024). Similarly, the response of species to fragmentation seems to be generally weak, but positive and negative responses have also been frequently reported (Fahrig 2017). Therefore, further research is required to ascertain the primary drivers of biodiversity loss in human-modified landscapes, the species and/or group of species that are most susceptible to forest spatial changes, and the landscape and/or regional contexts under which biodiversity responses become predominantly negative (Fletcher et al. 2018; Fahrig et al. 2019).

Global reviews and meta-analyses can be valuable approaches to obtain a more comprehensive understanding of the response of biodiversity to spatial changes of the habitat. To this end, an increasing number of global datasets have been compiled in the last decade (e.g., BIOFRAG: Pfeifer et al. 2014; PREDICT: Hudson et al. 2014; FragSAD: Chase et al. 2019). These datasets have been useful to test important ecological hypotheses, such as the 'ecosystem decay hypothesis' (Chase et al. 2020), and the 'extinction filter hypothesis' (Betts et al. 2019; Weeks et al. 2023). BIOFRAG primarily contains data on the presence/absence of populations (single species) or communities (multiple species) in forest fragments. Although it contains the spatial locations of the fragments, it does not provide species abundances to calculate diversity metrics within fragments, estimates of fragment size for all studies or information about the structure of the surrounding landscape (Pfeifer et al. 2014). Also, BIOFRAG is not currently open access. In turn, FragSAD is open access and provides information on species abundances and fragment size for all fragments (Chase et al. 2019, 2020), which allows an explicit evaluation of the effects of sampling effort on diversity (Chase et al. 2020). However, it does not include the exact spatial location of the fragments or plots, which prevents the assessment or control of the effects of spatial gradients, as well as the more precise description of the surrounding matrix. To date, most landscape studies using multiple taxonomic groups have relied on complex metrics of landscape structure, such as the '*edge influence*' index (Betts et al. 2019; Weeks et al. 2023), indirect measurements of matrix quality (e.g., matrix age and matrix type; Chase et al. 2020), or a few configuration metrics (e.g., mean fragment size; Riva and Fahrig 2023).

Here, we gathered a global database of the composition and abundance of 9154 species sampled in 1472 forest fragments as part of 121 studies of different taxonomic groups in tropical, subtropical, and temperate regions (Figure 1). For 2703 of these species, we were also able to compile information on morphological, trophic, habitat, and reproductive traits. We recorded the spatial location of each study fragment, its size (in hectares), and the spatial structure of the landscape surrounding each fragment. In particular, we calculated 10 landscape variables (Table 1) that have been at the core of important ecological debates (e.g., Fahrig et al. 2022). Therefore, this dataset has broad applicability in ecological research, and can potentially be used to address many research questions in landscapes with fragmented forests at all spatial scales, from local to global. Examples of potential research include, but are not limited to, testing the relative importance of fragment- and landscape-related variables affecting biodiversity changes, and determining the traits of species which are the 'winners' and 'losers' following forest loss and fragmentation (see, e.g., Riva and Fahrig 2023; Pinho et al. 2024; Zhang et al. 2024). Importantly, since the effect of landscape structure on biodiversity can go undetected if the structure is assessed at the wrong scale (Jackson and Fahrig 2015), we calculated landscape metrics in circular landscapes of different sizes (from 200 to 2000 m radius). This multiscale information can be highly valuable to identify the so-called 'scale of effect' of each landscape metric on each response (Jackson and Fahrig 2015)—an emerging topic in landscape ecology that can

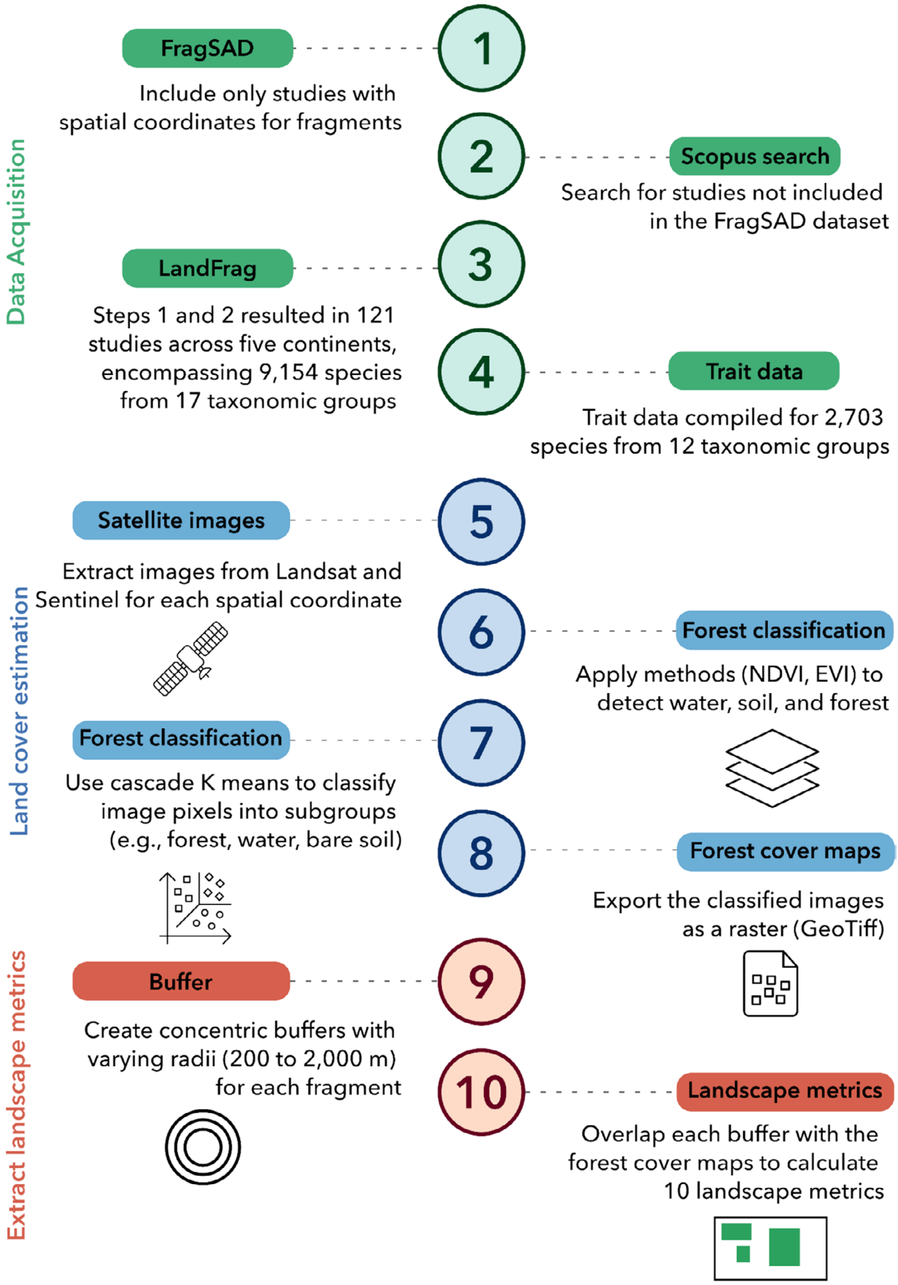

FIGURE 1 | Conceptual figure summarising the 10 most important steps to collect species and trait data (data acquisition, green), obtain satellite image and estimate forest cover (land cover estimation, blue), and to extract the landscape metrics (red).

be used for assessing important hypotheses on spatial scaling issues (Miguet et al. 2016).

2 | Methods

2.1 | Data Acquisition

We used FragSAD (Chase et al. 2019) as the starting point of this new dataset. While FragSAD includes species abundance and composition in forest fragments, it does not provide the exact spatial location of each surveyed forest fragment. Thus, we first revisited all studies in the FragSAD and selected only those that collected species in at least four forest fragments; we also excluded those from non-forest habitats such as grasslands. We retrieved the spatial location (latitude and longitude) of all fragments in a study by: (i) extracting the information from the article (tables) or maps with the coordinate system or (ii) contacting the first or corresponding author(s) of the original papers to request the spatial locations of the fragments. When we extracted data from maps, we used Google Earth to obtain the coordinates from the centroid of the studied fragment. In addition, we included data published since the original 2019 FragSAD publication by performing a new search on Scopus using the same keywords from Chase et al. (2019): ('habitat fragment*' OR 'habitat loss' OR 'forest fragment*' OR 'forest loss' OR 'fragment area' OR 'fragment size' OR 'island area') AND ('diversity' OR 'species diversity' OR 'species richness' OR 'abundance') from 2019 to 2022. For this search, we included those studies with open-access data, and contacted authors listed in the search to ask for published or unpublished datasets

TABLE 1 | Forest landscape metrics describing composition and configuration categories, including area, edge effects, aggregation and shape complexity.

Metric category	Metric group	Landscape metric	Description	Units
Composition	Area and edge	Mean area of fragments	Mean of all fragment areas of forest in the buffer	Hectares
Composition	Area and edge	Percentage of forest in the landscape	Total area of forest in the buffer, divided by the total area of the buffer multiplied by 100	Percent
Configuration	Aggregation	Number of fragments	Number of forest fragments and continuous forests in the landscape	Number
Configuration	Aggregation	Fragment density	Number of fragments in the buffer, divided by the total area of the buffer multiplied by 10,000 and 100	Number/100 ha
Configuration	Aggregation	Inter-fragment isolation distance	Distance from each fragment to its nearest adjacent fragment. The mean distance is then calculated using the distance for all fragments pairs in the landscape	Meters
Configuration	Aggregation	Aggregation index	It equals the number of like adjacencies divided by the theoretical maximum possible number of like adjacencies for that class	None
Configuration	Area and edge	Edge density	Total length of edge (m) in the buffer, divided by the total area of the buffer (m^2) multiplied by 10,000	Meters per ha
Configuration	Area and edge	Largest patch index	Area (m^2) of the largest fragment in the buffer, divided by the total area of the buffer multiplied by 100	Percent
Configuration	Shape	Perimeter-area ratio	Perimeters (m) of the fragment divided by the area (ha) of the fragment	Meters per ha
Configuration	Shape	Contiguity index	Contiguity value of a pixel in a given fragment divided by the area of that fragment	None

fitting our criteria. We called this dataset 'LandFrag', which stands for 'LANDscape and FRAGment data'. The LandFrag dataset gathered information of 1472 forest fragments derived from 121 studies. These studies gathered data from 679,928 individuals, representing 9154 species, documented across five continents and 32 countries (Figures 1 and 2; Supporting Information 1 and 2). These studies were primarily conducted in tropical regions, with 93 out of 121 studies focusing on these areas, and 56% of the studies being in South America. The prominence of fragmentation research in tropical areas has been demonstrated in other datasets (e.g., Pfeifer et al. 2014; Chase et al. 2019). Therefore, conclusions about global patterns must be approached with caution.

We combined a list of all species from these studies into the following taxonomic groups:

- Invertebrates (51 datasets):
  - Arthropoda: Arachnida—Araneae and Opiliones, Coleoptera, Diptera, Hemiptera, Hymenoptera—Formicidae, bees and wasps, Isoptera, Lepidoptera, Orthoptera; and Mollusca: Gastropoda.
- Vertebrates (55 datasets):
  - Amphibia, Aves, Mammalia, Squamata
- Vascular plants (19 datasets)

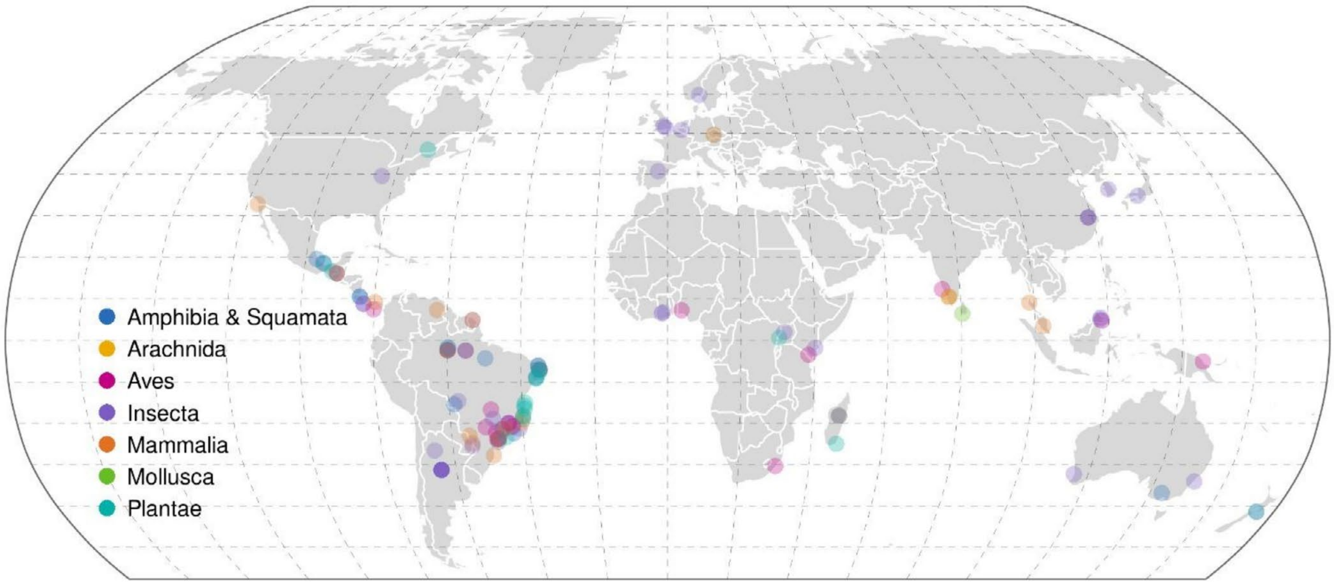

FIGURE 2 | Global distribution of the 121 habitat loss and fragmentation studies organised by taxonomic groups.

2.2 | Aggregating Studies by Sampling Design

Because the sampling design can affect our ability to compare species diversity across fragments and studies (Chase et al. 2020), we classified the 121 studies into three major categories depending on their sampling methods: (i) studies with standardised samples, where the same sampling effort was established in different fragments regardless of their size; (ii) studies using different efforts, usually proportional to fragment area, but which could be standardised by sample size (i.e., number of individuals) because they provide information on the number of individuals per fragment; and (iii) studies with a pooled design, where the total number of individuals of each species is reported for each fragment, but there is no information on the effort (e.g., number of samples) per fragment (see also Chase et al. 2020).

2.3 | Trait Data

We compiled functional traits for 2703 species (29% of all collected species; Table 2; Supporting Information 1). We were not able to provide trait data for 71% of the species because they were (1) unstudied organisms in our collected database, and (2) unidentified species (e.g., morphospecies identified at the class or family level). We used published datasets (Wilman et al. 2014; Oliveira et al. 2017; Kattge et al. 2020; Pinho et al. 2021; Shirey et al. 2022; Tobias et al. 2022) or the original paper describing a given species or genus to obtain trait data. Specifically, for those organisms with organised databases (Gastropoda, Hymenoptera, Lepidoptera, Amphibia, Aves, Mammalia and Plants), we matched the list of species found in the LandFrag dataset with species included in trait databases. For all other taxonomic groups, we searched for trait information in the manuscript (or monograph) that originally described each species (Table 2). For animals, most of the selected traits are associated with morphological features, trophic levels, habitat type, foraging time, and reproductive mode. For plants, the most common traits were dispersal syndrome, leaf Nitrogen, wood density, seed mass and leaf dry mass (Table 2).

2.4 | Estimating Land Cover

We used Google Earth Engine to classify images and estimate land cover for each plot or fragment. We applied cloud mask functions to generate a median image based on the sampling year (or study publication year when the sampling date was not available) for each plot/fragment. The collections used to extract these images were Landsat 4, 5, 7, 8 and Sentinel 2. The datasets consist of atmospherically corrected surface reflectance data produced by the sensors TM (Landsat 4 and 5), ETM+ (Landsat 7), OLI/TIRS (Landsat 8), and the Sentinel MSI sensor. For automatization, we excluded Landsat 7 when possible to ensure extracting information with no interference caused by specific limitations on this satellite (Wijedasa et al. 2012). We used different algorithms depending on the year in which each sampling was carried out. For studies performed before 2013, we selected image collections from Landsat 4 and 5. When there was no image available for a given location, we selected images from Landsat 7. We selected Landsat 8 for studies performed between 2013 and 2015, and Sentinel 2 for studies performed after 2015. Because the studies were conducted in different years, we were unable to use images from the same satellite.

After extracting the images for all plots or fragments, we applied normalised bands of MNDWI (Modified Normalised Difference of Water Index), NDVI (Normalised Difference Vegetation Index) and EVI (Enhanced Vegetation Index) (Wang et al. 2019) to detect water and live green vegetation that will further be used to estimate forest cover. We also used satellite bands of Short Wave Infrared (SWIR) to detect bare soil, near infra-red (NIR) for vegetation and Blue bands for water and shadows. In addition, a mask cloud function was also implemented to prevent interference caused by cloud and shadow in the final forest classification (Anzalone et al. 2024). We selected the classifier Cascade K-means that is based on an unsupervised method of cluster analysis where it splits the image pixels according to their datasets given random trained samples and separates them into subgroups (Caliński and Harabasz 1974). The robustness

TABLE 2 | Trait information and availability for all taxonomic groups in the LandFrag database.

Major taxa	Minor taxa	Trait information	Trait names	Number of species with trait information	Source
Arachnida	Araneae	Not available	None	None	None
Arachnida	Opiliones	Available	Body length and width, femur IV length, habitat	13	Manuscripts containing the original species description
Coleoptera	—	Not available	None	None	None
Diptera	—	Available	Body size, trophic level, reproductive mode, foraging strat, larvae diet, larvae trophic level	46	Manuscripts containing the original species description
Gastropoda	—	Available	Age at maturity, longevity, clutch size, maximal shell size, survival of dry period, inundation tolerance	27	Ellers et al. 2018
Hemiptera	—	Not available	None	None	None
Hymenoptera	Aculeata	Not available	None	None	None
Hymenoptera	Formicidae	Available	Morphological traits (e.g., head width and length, clypeus length, mandible length) and dominant colour	26	Parr et al. 2017
Hymenoptera	Parasitic wasps	Available	Body size, forewing length, diet, reproductive mode and foraging strata	100	Manuscripts containing the original species description
Isoptera	—	Not available	None	None	None
Lepidoptera	—	Available	Wingspan, forewing length, voltinism.	279	Shirey et al. 2022
Orthoptera	—	Available	Pronotum and femur size	75	Manuscripts containing the original species description
Amphibia	—	Available	Body size and mass, foraging habitat and period and diet	227	Oliveira et al. 2017 and manuscripts containing the original species description
Aves	—	Available	Morphological variables (e.g., beak length, body mass, tarsus length), habitat, trophic level	234	Tobias et al. 2022
Mammalia	—	Available	Body mass, diet, foraging strata and activity period, lifestyle	262	Wilman et al. 2014
Squamata	—	Available	Maximum snout to vent length, diet, substrate, reproductive mode, activity period	94	Manuscripts containing the original species description
Plants	—	Available	Dispersal syndrome, leaf nitrogen, wood density, seed dry mass, leaf dry mass	1320	(Kattge et al. 2020; Pinho et al. 2021)

and efficiency of this method in handling large datasets are better when compared to the traditional *K*-means classifier (Deka and Saha 2023). To train the models, we generated 400,000 random pixels in the landscape of interest using all normalised bands. We used this number of random pixels because it is a nice balance of the trade-offs among sample size, computational cost, and classification accuracy. We exported all files (i.e., forest cover maps) as GeoTiff, and we used them to calculate the landscape metrics.

2.5 | Extracting Landscape Variables at Different Scales

One way to investigate the effect of forest loss and fragmentation (Figure 3A) is to extract the landscape metrics surrounding the fragments. To do that, we established concentric buffers (Figure 3B) of 200–2000 m, in 100 m increments. These buffers were created from the centre of each sampled plot/fragment for each study site using the GeoTiff files obtained from the image classification algorithm. This multiscale approach is relevant to select the most appropriate scale affecting different species or taxonomic groups. We fit the previously created binary raster to each buffer and projected these rasters to the WGS84 UTM zone corresponding to each plot/fragment. We used these buffers and the forest cover maps to extract the following landscape metrics: forest cover, fragment density, number of fragments, forest edge density, largest patch index, contiguity index, euclidean nearest-neighbour distance, proximity index and perimeter. We describe each metric along with its category, group, scale, units and an ecological meaning in Table 1. We provided a non-exhaustive list of landscape metrics, but users can extract other metrics available in the *landscapemetrics* package (Hesselbarth et al. 2019) by adapting the code provided (Supporting Information).

Studies spanned the full gradient from intact (e.g., a buffer hosting a single forest fragment within a landscape with 100% forest cover) to intensively fragmented landscapes (e.g., landscapes with low forest cover and a large number of very small forest fragments). This whole gradient was likewise found for all taxonomic groups, allowing for comparison both within and across them (Figure S1). Users can also use landscape metrics at different scales to investigate how the scale of effect can change within and between taxonomic groups or ecosystems (Jackson and Fahrig 2015). We summarised all steps to create the LandFrag dataset in Figure 1; Supporting Information 2.

3 | Value and Potential Uses of the LANDFRAG Dataset

The functioning and ecological integrity of natural ecosystems largely depend on biodiversity and the ecological traits of the species that inhabit them (Andresen et al. 2018). Therefore,

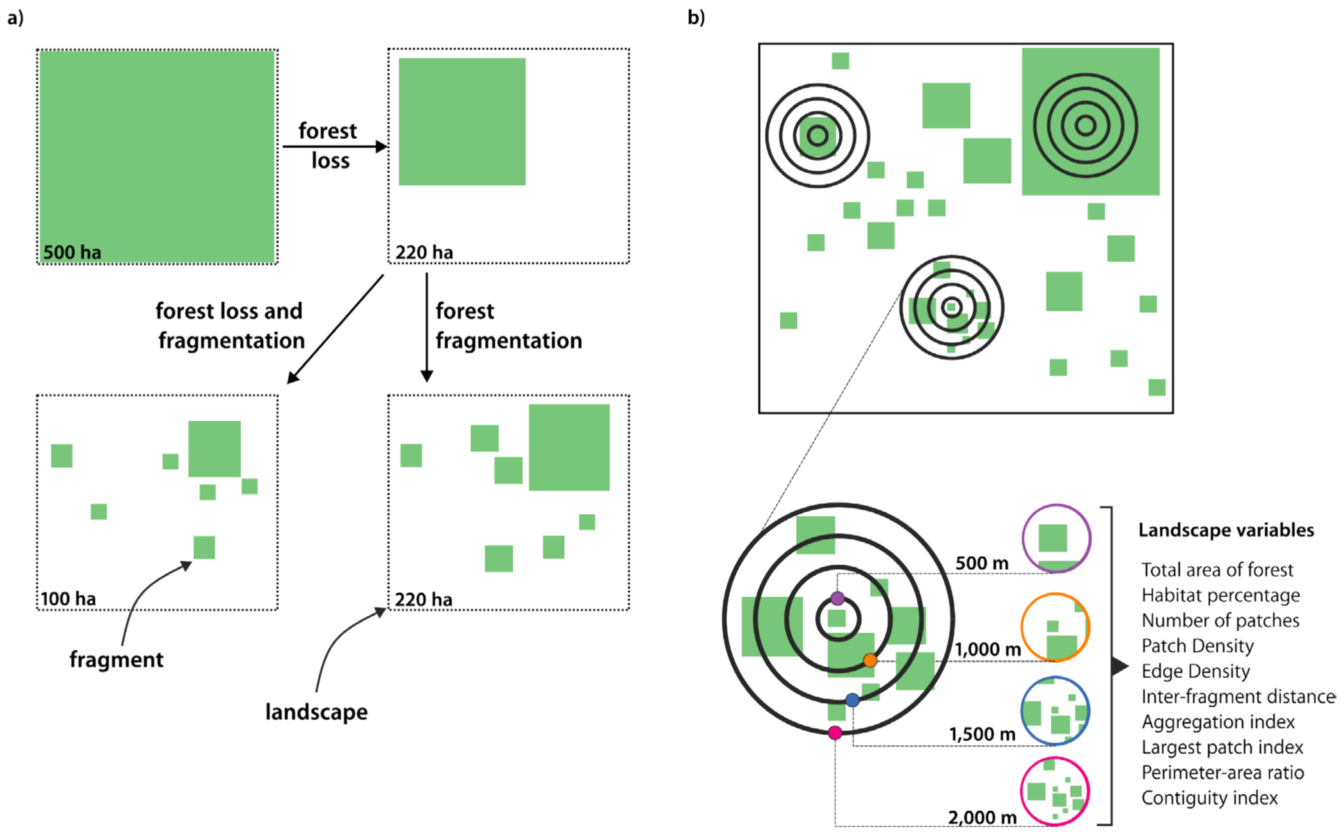

FIGURE 3 | (A) Schematic representation of forest loss and fragmentation. The green squares illustrate either continuous or fragmented forest within a given landscape. The area, measured in hectares (ha), represents the total forest cover present in that landscape. (B) Illustrative description of the buffers used to evaluate the landscapes surrounding the focal fragments. Buffer size varies from 200 to 2000 m of radius (with 100 m intervals), and variables were estimated based on each corresponding buffer size (without scale). We illustrated only four buffers with a radius of 500, 1000, 1500 and 2000 m.

the current biodiversity crisis in human-dominated landscapes threatens the sustainability of natural ecosystems and our own survival (Cardinale et al. 2012). To predict, prevent and if possible, reverse such anthropogenic damage, ecological research has been evaluating the relative importance of different threats to biodiversity in these landscapes. Although the theoretical advances in the matter have been very fruitful (Ewers and Didham 2006; Tscharntke et al. 2012; Fahrig 2017), empirical evidence is still limited and often debated. The LandFrag dataset is a global and up-to-date resource that can be used to carry out comprehensive assessments of these and other interesting theoretical models, which are of great applied value. Other potential uses of this database include the assessment of (1) the taxonomic and functional structure of communities in human-modified landscapes, (2) the functional signal of the responses of species to forest disturbances across scales and forest types, (3) the relationship between taxonomic and functional diversity metrics and (4) the taxonomic and functional predictors of the scale of landscape effects.

This dataset has a wide range of applications in biodiversity research and can be used to address a multitude of intriguing ecological and conservation questions at all spatial scales, from local to global. For example, for decades, ecologists and conservation biologists have noted that species diversity is usually lower in smaller and edge-affected forest fragments (Haddad et al. 2015; Fletcher et al. 2018). However, recent evidence suggests that species density in a given site can be more strongly related to habitat amount (e.g., forest cover) in the surrounding landscape than to the size of the fragment where the site is located (Watling et al. 2020). The LandFrag dataset can be used to untangle in what situations fragment- and landscape-scale variables determine biodiversity change. For instance, one potential source of confusion when comparing studies performed in different regions is that the effect of forest loss can vary depending on the species (Davies et al. 2004; Newbold et al. 2013; Pfeifer et al. 2017; Watling et al. 2020; Saldívar-Burrola et al. 2022) and the spatial extent (landscape size) at which forest cover is measured (Jackson and Fahrig 2015). This is also one approach that could be tested with LandFrag. Because this new dataset provides traits for many species, it is also possible to investigate which traits influence whether a given species will persist in fragmented landscapes, a topic that remains poorly investigated (Davies et al. 2004; Pinho et al. 2024).

Lastly, there is a relevant debate in the literature about the relative effects of forest loss and fragmentation on species and trait diversity. In fact, the effect of forest fragmentation on biodiversity remains contentious, since positive, negative and weak effects have been reported in the literature, and in many cases, fragmentation effects are confounded with the effect of forest loss (Chase et al. 2020; Riva et al. 2024; Zhang et al. 2024). Therefore, the LandFrag dataset should stimulate future research on forest loss and fragmentation, as well as their effects on biodiversity, providing new directions and potential solutions for biodiversity conservation (Gonçalves-Souza et al. 2025).

Affiliations

[1]Institute for Global Change Biology, School for Environment and Sustainability, University of Michigan, Ann Arbor, Michigan, USA | [2]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan, USA | [3]Programa de Pós-Graduação em Etnobiologia e Conservação da Natureza, Departamento de Biologia, Universidade Federal Rural de Pernambuco, Universidade de Pernambuco, Campus Mata Norte, Recife, PE, Brazil | [4]Laboratório de Ecologia Espacial e Conservação, Departamento de Biodiversidade, Instituto de Biociências, Universidade Estadual Paulista (Unesp), Rio Claro, SP, Brazil | [5]Kellogg Biological Station and Department of Integrative Biology, Michigan State University, East Lansing, Michigan, USA | [6]Independent Researcher, Belém, PA, Brazil | [7]Department of Ecology and Natural Resource Management (INA), Norwegian University of Life Sciences, Ås, Norway | [8]Department of Biological Sciences, State University of Feira de Santana, Feira de Santana, Bahia, Brazil | [9]Instituto de Investigaciones en Biodiversidad y Medioambiente (INIBIOMA), Universidad Nacional del ComahueCONICET, San Carlos de Bariloche, Río Negro, Argentina | [10]Instituto de Investigaciones en Ecosistemas y Sustentabilidad, Universidad Nacional Autónoma de México, Morelia, Michoacán, Mexico | [11]Escuela Nacional de Estudios Superiores, Universidad Nacional Autónoma de México, Mérida, Yucatán, Mexico | [12]Department of Life Sciences, Faculty of Science, University of Alcalá, Alclá de Henares, Spain | [13]Applied Ecology and Conservation Lab, Departamento de Ciências Biológicas, Universidade Estadual de Santa Cruz, Ilhéus, Bahia, Brazil | [14]Laboratório de Ciência Aplicada à Conservação da Biodiversidade, Departamento de Zoologia, Universidade Federal de Pernambuco, Recife, Pernambuco, Brazil | [15]Programa de Pós-graduação em Geociências, Departamento de Geologia, Universidade Federal de Pernambuco, Recife, Pernambuco, Brazil | [16]Laboratório de Etnobiologia e Biodiversidade, Núcleo Takinahaky, Universidade Federal de Goiás, Goiânia, Goiás, Brazil | [17]Programa de Pós-Graduação em Ciências Ambientais, Laboratório de Mastozoologia, Universidade do Estado de Mato Grosso, Campos de Júlio, Brazil | [18]Department of Ecology and Evolutionary Biology, Universidade Federal de São Paulo, Diadema, Brazil | [19]Department of Biology, University of Washington, Seattle, Washington, USA | [20]Departamento de Biologia, Universidade Federal Rural de Pernambuco, Recife, Pernambuco, Brazil | [21]Departamento de Biologia, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Ribeirão Preto, SP, Brazil | [22]Department of Environmental Science and Studies, DePaul University, Chicago, Illinois, USA | [23]The Field Museum, Chicago, Illinois, USA | [24]School of Biological Sciences, The University of Western Australia, Western Australia, Perth, Western Australia, Australia | [25]CSIRO Health and Biosecurity, Centre for Environment and Life Sciences, Floreat, Western Australia, Australia | [26]Departamento de Genética, Ecologia e Evolução, Instituto de Ciências Biológicas, UFMG, Belo Horizonte, Brazil | [27]Laboratório de Ecologia de Metacomunidades, Instituto Latino-Americano de Ciências da Vida e da Natureza. Universidade Federal da Integração Latino-Americana, Paraná, Brazil | [28]School of Environment and Life Sciences, Deakin University, Waurn Ponds, Victoria, Australia | [29]Instituto Mediterráneo de Estudios Avanzados (IMEDEA CSIC-UIB), Esporles, Spain | [30]Instituto Multidisciplinario de Biología Vegetal (IMBIV), Universidad Nacional de Córdoba (UNC), CONICET, Córdoba, Argentina | [31]INRAE, UR406 Abeilles et Environnement, Avignon, France | [32]Department of General Biology, Federal University of Viçosa, Viçosa, Minas Gerais, Brazil | [33]Center for Large Landscape Conservation and IUCN/WCPA—CCSG, Bozeman, Montana, USA | [34]School of BioSciences, University of Melbourne, Parkville, Victoria, Australia | [35]Centro de Formação em Ciências Agroflorestais, Universidade Federal do Sul da Bahia, Porto Seguro, BA, Brazil | [36]A.P. Leventis Ornithological Research Institute (APLORI), University of Jos, Plateau, Nigeria | [37]Instituto de Biologia, Universidade Federal da Bahia, Salvador, BA, Brazil | [38]Department of Biology, and the Program in Environmental Science, Whittier College, Whittier, California, USA | [39]Tana River Primate National Reserve, Hola, Kenya | [40]School of Animal, Rural and Environmental Sciences, Nottingham Trent University, Nottinghamshire, UK | [41]Centro de Biociências, Universidade Federal de Pernambuco, Recife, PE, Brazil | [42]Entomology Section, National Museums of Kenya, Nairobi, Kenya | [43]Departamento de Ecologia e Zoologia, Universidade Federal de Santa Catarina,

Florianópolis, SC, Brazil | [44]Instituto de Biologia, Universidade Federal de Uberlândia (UFU), Uberlândia, MG, Brazil | [45]Independent Researcher, São Paulo, SP, Brazil | [46]Programa de Pós-Graduação em Biologia Animal, Departamento de Zoologia, Universidade Federal de Pernambuco, Recife, PE, Brazil | [47]Laboratorio de Entomologia General y Aplicada, Centro de Investigación de Estudios Avanzados del Maule, Universidad Católica del Maule, Talca, Maule, Chile | [48]Centro de Ciências Biológicas e da Saúde, Universidade Estadual do Oeste do Paraná, Cascavel, PR, Brazil | [49]Programa de Pós-Graduação em Ecologia e Recursos Naturais, Universidade Federal do Ceará, CE, Brazil | [50]Programa de Pós-Graduação em Ecologia Aplicada, Instituto de Ciências Naturais, Departamento de Ecologia e Conservação, Laboratório de Ecologia e Conservação de Mamí feros, Universidade Federal de Lavras, Lavras, MG, Brazil | [51]Araucária Innovation and Sustainability Lab (LASI), Environmental Studies Center (CEA), São Paulo State University, Rio Claro, São Paulo, Brazil | [52]Center for Ecology, Evolution and Conservation, School of Environmental Sciences, University of East Anglia, Norwich, UK | [53]Institute of Plant Sciences, University of Bern, Bern, Switzerland | [54]UMR 7205 CNRS MNHN SU EPHE UA, Institut de Systématique, Evolution, Biodiversité, Muséum National d'Histoire Naturelle, Paris, France | [55]Universidade Federal da Fronteira Sul, Campus Realeza, Realeza, PR, Brazil | [56]Independent Researcher, Vygruppen AS, Oslo, Norway | [57]Instituto de Pesquisa Waita, Belo Horizonte, Brazil | [58]Independent Researcher, Rua Marciano Santos, Uberlândia, MG, Brazil | [59]Programa de Pós-Graduação em Botânica, Departamento de Botânica, Universidade Federal de Viçosa, Viçosa, MG, Brazil | [60]Facultad de Estudios Superiores Iztacala, Universidad Nacional Autónoma de México, Tlalnepantla de Baz, Mexico | [61]Departamento de Sistemática e Ecologia, Centro de Ciências Exatas e da Natureza, Universidade Federal da Paraíba, Cidade Universitária, João Pessoa, PB, Brazil | [62]Programa de Pós-Graduação em Biologia Vegetal, Departamento de Botânica, Universidade Federal de Pernambuco, PE, Brazil | [63]Departamento de Biologia Animal, Instituto de Biologia, Universidade Estadual de Campinas, São Paulo, Brazil | [64]Center for International Forestry Research, Jawa Barat, Indonesia | [65]Departamento de Solos, Universidade Federal de Viçosa, Viçosa, MG, Brazil | [66]Departamento de Ensino, Instituto Federal de Educação, Ciência e Tecnologia do Ceará—IFCE, Acopiara, CE, Brazil | [67]Laboratório de Ecologia Teórica: Integrando Tempo, Biologia e Espaço (LET.IT.BE), Departamento de Ciências Ambientais, Universidade Federal de São Carlos—UFSCAR, Sorocaba, SP, Brazil | [68]Coleção Entomológica de Tangará da Serra (CEnTg), Universidade do Estado de Mato Grosso (UNEMAT), Tangará da Serra, MT, Brazil | [69]Instituto Federal do Espírito Santo, Vitória, ES, Brazil | [70]Cuerpo Académico de Ecología y Diversidad Faunística, Facultad de Ciencias Naturales, Universidad Autónoma de Querétaro, Santiago de, Querétaro, Mexico | [71]Department of Biological Sciences, Wright State University, Dayton, Ohio, USA | [72]Programa de Pós-Graduação em Ambiente e Sistemas de Produção Agrícola, Universidade do Estado de Mato Grosso, Tangará da Serra, MT, Brazil | [73]Programa de Pós-Graduação em Análise e Modelagem de Sistemas Ambientais, Instituto de Geociências, Universidade do Estado de Minas Gerais, Belo Horizonte, MG, Brazil | [74]Department of Biological Sciences, University of Bergen, Bergen, Norway | [75]Independent Researcher, Brasilia, DF, Brazil | [76]Laboratório de Zoologia, Universidade Federal de Mato Grosso, Cuiabá, MT, Brazil | [77]Universidade Estadual de Campinas, Campinas, SP, Brazil | [78]German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany | [79]Institute of Computer Science, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

Acknowledgements

This study was supported by the Institute for Global Change Biology in the School for Environment and Sustainability and the Department of Ecology and Evolutionary Biology at the University of Michigan. M.H.V. (#2022/01899-6, São Paulo Research Foundation—FAPESP), J.C.P. (#2018/00107-3), and F.R.d.S. (#2013/50714-0 and #2022/04012-2) were supported by grants from FAPESP. A.L. was supported by the project PRPPID2020-117863RB-I00 (MCIN/AEI/10.13039/501100011033). F.P.L.M. (470574/2013-5), M.B. (304189/2022-7), and L.F.S.M. (307984/2022-2) received support from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico). M.V.G. was funded by UNILA (PRPPG 205/2021 and 77/2022). N.S. received support from Fundação Araucária (PBA2022011000042).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The dataset and code can be downloaded on Zenodo https://zenodo.org/records/12206838 or GitHub https://github.com/mauriciovancine/landfrag.

References

Andresen, E., V. Arroyo-Rodríguez, and F. Escobar. 2018. "Ecological Networks in the Tropics: An Integrative Overview of Species Interactions from Some of the Most Species-Rich Habitats on Earth." In *Tropical Biodiversity: The Importance of Biotic Interactions for Its Origin, Maintenance, Function, and Conservation*, edited by W. Dáttilo and V. Rico-Gray, 1–13. Springer International Publishing.

Anzalone, A., A. Pagliaro, and A. Tutone. 2024. "An Introduction to Machine and Deep Learning Methods for Cloud Masking Applications." *Applied Sciences* 14: 2887.

Arroyo-Rodríguez, V., L. Fahrig, M. Tabarelli, et al. 2020. "Designing Optimal Human-Modified Landscapes for Forest Biodiversity Conservation." *Ecology Letters* 23: 1404–1420.

Betts, M. G., C. Wolf, M. Pfeifer, et al. 2019. "Extinction Filters Mediate the Global Effects of Habitat Fragmentation on Animals." *Science* 366: 1236–1239.

Caliński, T., and J. Harabasz. 1974. "A Dendrite Method for Cluster Analysis." *Communications in Statistics* 3: 1–27.

Cardinale, B. J., J. E. Duffy, A. Gonzalez, et al. 2012. "Biodiversity Loss and Its Impact on Humanity." *Nature* 486: 59–67.

Chase, J. M., S. A. Blowes, T. M. Knight, K. Gerstner, and F. May. 2020. "Ecosystem Decay Exacerbates Biodiversity Loss With Habitat Loss." *Nature* 584: 238–243.

Chase, J. M., M. Liebergesell, A. Sagouis, et al. 2019. "FragSAD: A Database of Diversity and Species Abundance Distributions From Habitat Fragments." *Ecology* 100: e02861.

Davies, K. F., C. R. Margules, and J. F. Lawrence. 2004. "A Synergistic Effect Puts Rare, Specialized Species at Greater Risk of Extinction." *Ecology* 85: 265–271.

Deka, P., and U. Saha. 2023. "Introduction of k-Means Clustering Into Random Cascade Model for Disaggregation of Rainfall From Daily to 1-Hour Resolution With Improved Preservation of Extreme Rainfall." *Journal of Hydrology* 620: 129478.

Ewers, R. M., and R. K. Didham. 2006. "Confounding Factors in the Detection of Species Responses to Habitat Fragmentation." *Biological Reviews* 81: 117–142.

Fahrig, L. 2017. "Ecological Responses to Habitat Fragmentation per se." *Annual Review of Ecology, Evolution, and Systematics* 48: 1–23.

Fahrig, L., V. Arroyo-Rodríguez, J. R. Bennett, et al. 2019. "Is Habitat Fragmentation Bad for Biodiversity?" *Biological Conservation* 230: 179–186.

Fahrig, L., J. I. Watling, C. A. Arnillas, et al. 2022. "Resolving the SLOSS Dilemma for Biodiversity Conservation: A Research Agenda." *Biological Reviews* 97: 99–114.

Fletcher, R. J., R. K. Didham, C. Banks-Leite, et al. 2018. "Is Habitat Fragmentation Good for Biodiversity?" *Biological Conservation* 226: 9–15.

Global Forest Watch. 2024. "World Resources Institute."

Gonçalves-Souza, T., J. M. Chase, N. M. Haddad, et al. 2025. "Increasing Species Turnover Does Not Alleviate Biodiversity Loss in Fragmented Landscapes." *Nature*. https://doi.org/10.1038/s41586-025-08688-7.

Haddad, N. M., L. A. Brudvig, J. Clobert, et al. 2015. "Habitat Fragmentation and Its Lasting Impact on Earth's Ecosystems." *Science Advances* 1: e1500052.

Hansen, M. C., L. Wang, X.-P. Song, et al. 2020. "The Fate of Tropical Forest Fragments." *Science Advances* 6: eaax8574.

Hesselbarth, M. H. K., M. Sciaini, K. A. With, K. Wiegand, and J. Nowosad. 2019. "Landscapemetrics: An Open-Source R Tool to Calculate Landscape Metrics." *Ecography* 42: 1648–1657.

Hudson, L. N., T. Newbold, S. Contu, et al. 2014. "The PREDICTS Database: A Global Database of How Local Terrestrial Biodiversity Responds to Human Impacts." *Ecology and Evolution* 4: 4701–4735. https://doi.org/10.1002/ece3.1303.

Jackson, H. B., and L. Fahrig. 2015. "Are Ecologists Conducting Research at the Optimal Scale?" *Global Ecology and Biogeography* 24: 52–63.

Kattge, J., G. Bönisch, S. Díaz, et al. 2020. "TRY Plant Trait Database-Enhanced Coverage and Open Access." *Global Change Biology* 26: 119–188.

Lôbo, D., T. Leão, F. P. L. Melo, A. M. M. Santos, and M. Tabarelli. 2011. "Forest Fragmentation Drives Atlantic Forest of Northeastern Brazil to Biotic Homogenization." *Diversity and Distributions* 17: 287–296.

Martínez-Ruiz, M., V. Arroyo-Rodríguez, R. Arasa-Gisbert, M. A. Hernández-Ruedas, and M. San-José. 2024. "Maintenance of Different Life Stages of Old-Growth Forest Trees in Deforested Tropical Landscapes." *Ecology* 105: e4273.

Miguet, P., H. B. Jackson, N. D. Jackson, A. E. Martin, and L. Fahrig. 2016. "What Determines the Spatial Extent of Landscape Effects on Species?" *Landscape Ecology* 31: 1177–1194.

Newbold, T., J. P. W. Scharlemann, S. H. M. Butchart, et al. 2013. "Ecological Traits Affect the Response of Tropical Forest Bird Species to Land-Use Intensity." *Proceedings of the Royal Society B: Biological Sciences* 280: 20122131.

Oliveira, B. F., V. A. São-Pedro, G. Santos-Barrera, C. Penone, and G. C. Costa. 2017. "AmphiBIO, a Global Database for Amphibian Ecological Traits." *Scientific Data* 4: 170123.

Pfeifer, M., V. Lefebvre, T. A. Gardner, et al. 2014. "BIOFRAG – A New Database for Analyzing BIOdiversity Responses to Forest FRAGmentation." *Ecology and Evolution* 4: 1524–1537.

Pfeifer, M., V. Lefebvre, C. A. Peres, et al. 2017. "Creation of Forest Edges has a Global Impact on Forest Vertebrates." *Nature* 551: 187–191.

Pinho, B. X., F. P. L. Melo, C. J. F. ter Braak, et al. 2024. "Winner–Loser Plant Trait Replacements in Human-Modified Tropical Forests." *Nature Ecology & Evolution* 9, no. 2: 1–15. https://doi.org/10.1038/s41559-024-02592-5.

Pinho, B. X., M. Tabarelli, C. J. F. ter Braak, et al. 2021. "Functional Biogeography of Neotropical Moist Forests: Trait–Climate Relationships and Assembly Patterns of Tree Communities." *Global Ecology and Biogeography* 30: 1430–1446.

Riva, F., and L. Fahrig. 2023. "Landscape-Scale Habitat Fragmentation Is Positively Related to Biodiversity, Despite Patch-Scale Ecosystem Decay." *Ecology Letters* 26: 268–277.

Riva, F., N. Koper, and L. Fahrig. 2024. "Overcoming Confusion and Stigma in Habitat Fragmentation Research." *Biological Reviews* 99: 1411–1424.

Saldívar-Burrola, L. L., M. Martínez-Ruíz, V. Arroyo-Rodríguez, et al. 2022. "Can Secondary Forests Mitigate the Negative Effect of Old-Growth Forest Loss on Biodiversity? A Landscape-Scale Assessment of Two Endangered Primates." *Landscape Ecology* 37: 3223–3238.

Shirey, V., E. Larsen, A. Doherty, C. A. Kim, and L. Ries. 2022. "LepTraits 1.0 A Globally Comprehensive Dataset of Butterfly Traits." Scientific Data, 9, 382.

Tobias, J. A., C. Sheard, A. L. Pigot, et al. 2022. "AVONET: Morphological, Ecological and Geographical Data for all Birds." *Ecology Letters* 25: 581–597.

Tscharntke, T., J. M. Tylianakis, T. A. Rand, et al. 2012. "Landscape Moderation of Biodiversity Patterns and Processes-Eight Hypotheses." *Biological Reviews* 87: 661–685.

Wang, Y., J. Ma, X. Xiao, X. Wang, S. Dai, and B. Zhao. 2019. "Long-Term Dynamic of Poyang Lake Surface Water: A Mapping Work Based on the Google Earth Engine Cloud Platform." *Remote Sensing* 11: 313.

Watling, J. I., V. Arroyo-Rodríguez, M. Pfeifer, et al. 2020. "Support for the Habitat Amount Hypothesis From a Global Synthesis of Species Density Studies." *Ecology Letters* 23: 674–681.

Weeks, T. L., M. G. Betts, M. Pfeifer, et al. 2023. "Climate-Driven Variation in Dispersal Ability Predicts Responses to Forest Fragmentation in Birds." *Nature Ecology & Evolution* 7: 1079–1091.

Wijedasa, L. S., S. Sloan, D. G. Michelakis, and G. R. Clements. 2012. "Overcoming Limitations With Landsat Imagery for Mapping of Peat Swamp Forests in Sundaland." *Remote Sensing* 4: 2595–2618.

Wilman, H., J. Belmaker, J. Simpson, C. de la Rosa, M. M. Rivadeneira, and W. Jetz. 2014. "EltonTraits 1.0: Species-Level Foraging Attributes of the World's Birds and Mammals." *Ecology* 95: 2027.

Zhang, H., J. M. Chase, and J. Liao. 2024. "Habitat Amount Modulates Biodiversity Responses to Fragmentation." *Nature Ecology & Evolution* 8, no. 8: 1–11. https://doi.org/10.1038/s41559-024-02445-1.

Supporting Information

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