Global 100m Projections of Biodiversity Intactness for the years 2017-2020

Technical White Paper

August 2022

Authors

Francis Gassert[1] , Joe Mazzarello[2†], Sam Hyde[2]

[1] Vizzuality, 123 Calle de Fuencarral, Madrid, Spain.
[2] Impact Observatory, Impact Observatory, 3640 Brandywine Street, Washington, DC 20008, USA.
[†] Corresponding author <joe@impactobservatory.com>

Key Words

Biodiversity, Ecology, Human pressures, Gridded data, Remote sensing

Abstract

This paper describes a global spatial model estimating human impacts on terrestrial Biodiversity Intactness and the resulting global 100-meter gridded maps for the years 2017-2020. This paper builds on past studies that map Biodiversity Intactness using the PREDICTS database of spatially referenced observations of biodiversity across 32,000 sites from over 750 studies. Our approach differs from previous studies by modeling the relationship between observed biodiversity metrics and contemporary, global, geospatial layers of human pressures, with the intention of providing a high resolution monitoring product into the future.

Introduction

Indicators of biodiversity are essential tools for planning and monitoring measures to conserve the world's remaining biodiversity. The nations of the world, through the United Nations Convention on Biological Diversity (CBD), set decadal strategic plans and targets for preserving the world's precious biodiversity, most recently in Aichi, Japan in 2010. These global targets are due for an upgrade, and so are the data products used to assess progress towards those targets. In particular, measures of biodiversity intactness are key indicators for monitoring global progress toward Aichi Target 12 (OECD 2019). These indicators are also most useful if they have the ability to "detect changes in systems within timeframes and on scales relevant to decision-making" (CBD 2010). However, many existing indicators of biodiversity are out of date and update infrequently, leaving policymakers with an information gap for understanding recent change.

Given advancements in data collection, machine learning, and cloud computing, we have established the software and methods to produce a high-resolution global dataset of biodiversity intactness that can be run on a continued basis into the future enabling regular updates to this critical indicator.

Materials and methods

We follow the prior work of Hill et al. (Hill et al. 2018) and Newbold et al. (Newbold et al. 2016) in constructing our global maps of biodiversity intactness. The Biodiversity Intactness Index (BII), described by Scholes and Biggs (2005), summarizes the change in ecological communities in response to human pressures. We estimate intactness as a combination of two metrics: *Abundance*, the quantity of individuals, and *Compositional Similarity*, how similar the composition of species is to an intact baseline. We fit linear mixed effects models to estimate the predictive capacity of spatial datasets of human pressures on each of these metrics, and project results spatially across the globe.

Using the fitted model coefficients, we generate maps of projected Biodiversity Intactness for 2017-2020. Maps are computed at a 0.00090º resolution (approximately 100 m at the equator) in the WGS84 projection using open source Python libraries on the Microsoft Azure Planetary Computer infrastructure.

PREDICTS database

We estimate model parameters by fitting to observed Abundance and Compositional Similarity derived from the PREDICTS database (Hudson et al. 2017; 2014). The PREDICTS database includes millions of measurements of species across over 26,000 sampling locations in 94 countries, assembled from approximately 500 studies of geographic variation in biodiversity. The PREDICTS database was specifically assembled to provide a representative dataset investigating the effect of human impacts related to land use. Observations are hierarchically organized by study site, study block (groups of sites observed together), and studies.

We prepare PREDICTS data for model fitting following Hill et al. (Hill et al. 2018). Abundance is calculated as the sampling-effort-corrected number of individuals at each site relative to the maximum abundance observed within the study. Compositional Similarity is measured using the Bray-Curtis dissimilarity index between each site and each site with a “primary minimal” (i.e. intact) land use classification in the same study. The Bray-Curtis accounts for the dissimilarity of both the occurrence and quantity of individuals of each species identified in one or both sites (Bray and Curtis 1957).

Modeling framework

We develop a model that predicts how much the abundance and composition of species is likely to have changed from a theoretical intact baseline, across the world. As the species of interest and composition of species will differ across studies and regions, we use a mixed effects modeling approach to account for study level and site group level differences and to isolate the fixed effects of human pressures on abundance and composition. In contrast with previous global Biodiversity Intactness studies that model the effects of human pressures based on observed land uses reported in PREDICTS studies themselves, we specifically use only independent global spatially explicit datasets as predictors in our modeling framework.

Predictors

In order to select the predictor variables for our model, we first sought to identify the best globally mapped proxies for known causal mechanisms of biodiversity degradation (Table 1). We then eliminated sources for which data was deemed to be unavailable or insufficient. For example, dasymetric gridded datasets that are derived from national level statistics such as crop yields, livestock density, and GDP, are likely to have negligible predictive utility as studies are always within individual countries. As we seek to develop a high-resolution monitoring product, we prefer data sources with higher spatial resolution that are likely to be continued to be measured into the future.

Table 1. Potential spatial covariates for mechanisms of biodiversity degradation.

Mechanism	Description	Potential covariates
Land use change	Deforestation, land degradation, fragmentation, intensive grazing, and other forms of land use change	Land cover - Agriculture - Urban - Bare ground - Fragmentation Forest use - Plantations - Managed forest Agricultural intensity - Greenness - Crop yield Livestock grazing - Livestock density - Pasture
Direct exploitation & indirect pressures	Resource extraction, hunting, tourism	Distance to roads Travel time to cities Population density Nighttime lights GDP Other Infrastructure
Climate change	Shifting climate regimes	*Not applicable to analytical spatio-temporal scale*
Pollution	Nutrient loading, pesticides, etc.	*No known spatial covariates*
Invasive and non-native species	Non-native species may outcompete or alter the composition of ecosystems	*No known spatial covariates*
"Natural" disruption	Fires, disease, drought, climatic variation etc. May be related to human activity, but can be treated as stochastic	Climatic regimes Topography Fire scars

The final set of candidate predictors (Table 2) include land cover, forest management, and other human pressures proxied by population, accessibility to cities, and nighttime lights. As it can take several years for land deforested for development to appear as another land use type, we also include recent tree cover loss as a predictor.

For each predictor, we extract average values within a buffer of each PREDICTS site to reduce the effect of noise in high resolution datasets and to reduce potential error introduced by inaccuracies in the reported coordinates of PREDICTS sites. We use a square kernel to compute buffers for computational efficiency in projections. For land cover data, we include both 100m and 1km buffers to account for edge effects. For land use data, we use only 100m buffers. For diffuse human pressure proxies, we use 1km buffers. Where multiple years are available, we select the year of data closest to the median year of observation in the PREDICTS data, 2006. To limit the effect of outliers, we cap distance-to-roads values greater than 10km to 10km and travel time to cites values greater than 24hrs to 24hrs. Human pressure proxy variables were log-transformed to account for diminishing marginal effects.

We opt to not include interactions between variables to simplify model interpretation. The overlapping nature of the predictor variables and log-transformations should be able to capture many nonlinear effects.

Table 2. Selected candidate variables for model fitting and selection.

Variable	**Description**	**Source**
Independent variables		
ln(TotalAbundance + 1)	Log-transformed total abundance of all observed species at site	PREDICTS (Hudson et al. 2016)
logit(Bray)	Logit-transformed compositional similarity of species at site compared to intact sites within study	PREDICTS (Hudson et al. 2016)
Dependent variables		
lcCrops_100m	Portion of land classified as crops within a 100 m buffer square kernel	10m Annual Land Use Land Cover: 9-class (K. Karra et al. 2021)
lcCrops_1000m	Portion of land classified as crops within a 1 km buffer	10m Annual Land Use Land Cover: 9-class (K. Karra et al. 2021)
lcBuiltArea_100m	Portion of land classified as built area within a 100 m buffer	10m Annual Land Use Land Cover: 9-class (K. Karra et al. 2021)
lcBulitArea_1000m	Portion of land classified as built area within a 1 km buffer	10m Annual Land Use Land Cover: 9-class (K. Karra et al. 2021)

ln(distRoads + 1)	Log-transformed distance to the nearest road in meters.	Open Street Map (OpenStreetMap contributors 2021)
ln(accessibility + 1)	Log-transformed travel time to the nearest city in minutes.	Oxford MAP Travel Time to Cities (Weiss et al. 2018)
ln(pD2006_1000m + 1)	Log-transformed average population density within a 1 km buffer square kernel.	WorldPop 100m unconstrained population counts (WorldPop et al. 2018)
ln(nL2012_1000m + 1)	Log-transformed average nighttime light intensity (nW/cm^2/sr) within a 1 km buffer square kernel.	Viirs Nighttime Lights Version 2 median annual composites (Elvidge et al. 2021)
managedForest_100m	Portion of land classified as plantations, woodlots, or agroforestry within a 100m buffer	Global forest management (Lesiv et al. 2022)
forestLoss2006_100m	Portion of land classified as tree cover loss within a 100m buffer	Hansen Global Forest Change v1.9 (Hansen et al. 2013)

Model selection

We fit mixed effects models of Abundance and Compositional Similarity on all predictors using the maximum likelihood estimator of the `lme4` R-package. For the Abundance model, we include a random effect on each study block. For the Compositional Similarity model, we allow for comparisons across study blocks, including a random effect on each study, and also include a fixed effect on the log-transformed distance between each pair of sites.

Final model selection is performed to minimize the Akaike Information Criterion (AIC) using backwards stepwise elimination of insignificant predictors with the `lmerTest` R-package. Model fitting results are reported in Tables 3-4.

Table 3. Abundance model coefficients and statistics

Variable	Estimate	Std. Error
(Intercept)	3.70E+00	5.21E-02***
ln(distRoads + 1)	-9.82E-03	3.55E-03**
ln(pD2006_1000m + 1)	-9.63E-02	2.89E-02***
lcCrops_100m	-7.27E-02	3.44E-02*
lcBuiltArea_1000m	2.54E-01	1.12E-01*
lcBuiltArea_100m	-1.77E-01	5.23E-02***
forestLoss2006_100m	-1.56E-01	3.87E-02***

Significance levels: * p<0.05, ** p<0.01, *** p<0.001

Table 4. Community similarity model coefficients and statistics

Variable	Estimate	Std. Error
(Intercept)	2.18E-01	1.69E-01
ln(distRoads + 1)	1.81E-02	1.99E-03***
ln(accessibility + 1)	1.37E-01	7.67E-03***
ln(pD2006_1000m + 1)	2.53E-01	2.05E-02***
ln(nL2012_1000m + 1)	-3.14E-01	2.20E-02***
lcCrops_1000m	-9.35E-01	3.38E-02***
lcCrops_100m	-2.48E-01	1.64E-02***
lcBuiltArea_1000m	-7.46E-01	7.26E-02***
forestManagement_100m	-7.70E-01	1.99E-02***
forestLoss2006_100m	3.39E-01	2.21E-02***
ln(geog_dist + 1)	-1.28E-01	1.73E-03***

Significance levels: * p<0.05, ** p<0.01, *** p<0.001

Projected Biodiversity Intactness 2017-2020

We project the effects of human pressures on biodiversity globally for the period of 2017-2020 as the sum of the product of each layer for the given year and its modeled coefficient. The land cover, population, nighttime lights, and tree cover loss datasets each cover the entire period. The forest management, distance to roads, and accessibility to cities datasets are treated as static layers across the time period.

Maps for Abundance and Compositional Similarity are created separately, and normalized to a 0-1 scale by dividing by the modeled value for “no human influence”, i.e. the projected value at the maximum modeled distance of 10km from a road, 24hrs from a city, and with no other human land uses or pressures. Projected Biodiversity Intactness is then computed as *Abundance * Compositional Similarity*. Mapped results for 2020 are shown in Figure 1.

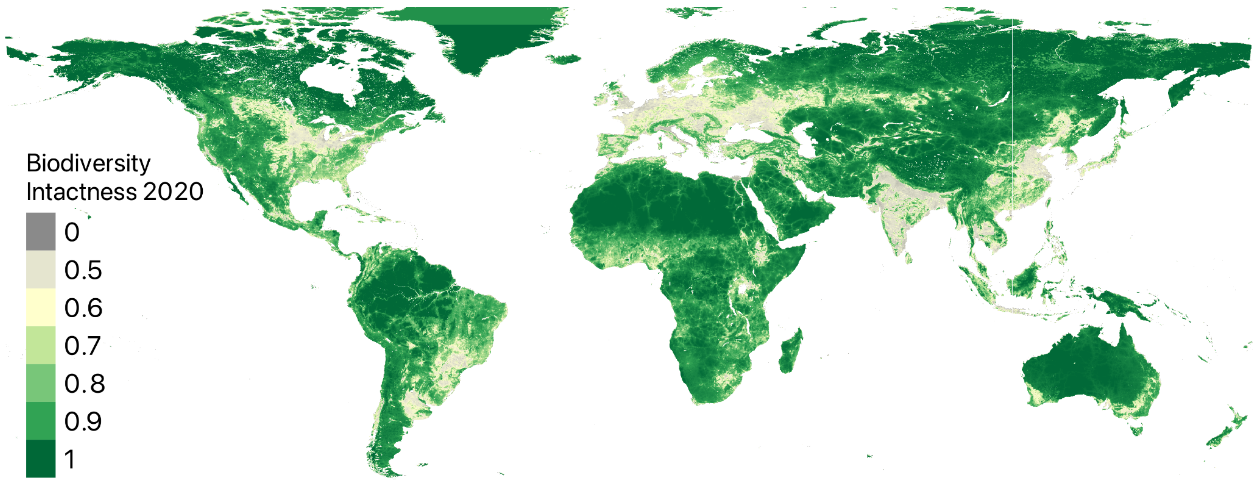

Figure 1. Projected Biodiversity Intactness for the year 2020

Comparison with analogous products

We compare our results to two external datasets: the 2015 Biodiversity Intactness Index (BII) by Newbold et al. (2016) and the 2020 Biodiversity Habitat Index (BHI) by Harwood et al. (2022). Both of these datasets use the PREDICTs dataset to spatially model biodiversity intactness, but differ in their approach. For each dataset we extract values for 10,000 randomly-sampled points across the world’s land area. We use our 2017 values to compare with the 2015 BII, (the closest temporal match) and 2020 values to compare with the 2020 BHI. All three datasets provide index scores between 0 and 1. We calculate both the Pearson’s correlation coefficient, an indicator of relative alignment in scores, as well as the Mean Absolute Difference in scores (Table 6).

Table 6. Statistics for comparison with analogous datasets for 10000 randomly sampled points.

Comparison	Mean Absolute Difference	Pearson's Correlation Coefficient
2020 vs BHI 2020	0.09	0.65
2017 vs BII 2015	0.23	0.73

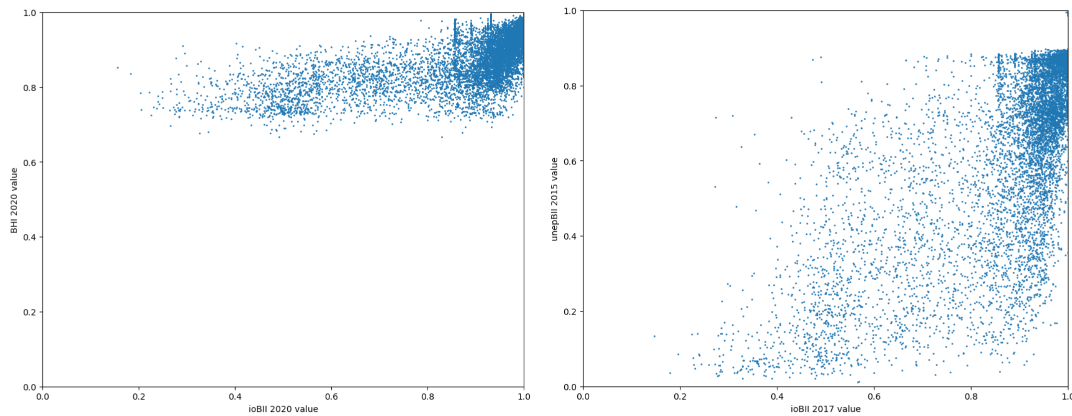

Figure 2. Comparison of scores from this study with scores from the Biodiversity Habitat Index (A) and the Biodiversity Intactness Index (B) for 10,000 randomly sampled points around the world.

Discussion

The projected biodiversity intactness data presented provide a globally high-resolution map of potential degradation of biodiversity due to human influence worldwide. These maps are generated such that they can be continued to be operationally produced as new monitoring data is made available, and are intended to be used as a tool for monitoring large scale influence on biodiversity.

Nonetheless, caution should be taken in interpreting scores at the local scale. Although these maps are based on thousands of local scale observations of species, they only capture a small portion of the factors that affect the richness and abundance of biodiversity in local ecosystems. We identify three key limitations in our approach that should be kept in mind when interpreting these maps.

First, when comparing study sites to determine levels of intactness, the level of biodiversity at "intact" sites may reflect an already substantially degraded state. Few areas of the world remain unaffected by human influence, and maps may be interpreted as an upper bound on biodiversity intactness.

Second, pressures that cannot be readily observed from remote sensing, such as pollution, poaching, disease, livestock grazing, habitat fragmentation, and the competitive pressure of introduced species will not be well reflected in these maps. Although these pressures may be broadly captured by proxy variables such as proximity to roads and population centers, they may lead to incorrect local-scale patterns in the map. For example, since plantation forests are explicitly measured, but pasturing is not, pasture lands adjacent to managed forest may receive higher intactness scores despite greater actual degradation.

Third, we do not attempt to account for the sensitivity of different ecosystems to human pressures. As such, isolated, stressed, or otherwise vulnerable communities may be less intact than the scores show.

Noting these limitations, we believe that the ability for this methodology and data to provide large scale consistent monitoring of biodiversity intactness is highly relevant for timely global policy decision making. Moreover, the approach described here is intended to be able to be run continuously into the future, and incorporate new and better input data as they become available. We hope that this open, automated approach will spread to other key environmental indicators and continue to bridge the knowledge gaps between science and policy making.

Bibliography

Bray, J. Roger, and J. T. Curtis. 1957. "An Ordination of the Upland Forest Communities of Southern Wisconsin." *Ecological Monographs* 27 (4): 325–49. https://doi.org/10.2307/1942268.

CBD. 2010. "Examination of the Outcome-Oriented Goals and Targets (and Associated Indicators) and Consideration of Their Possible Adjustment for the Period beyond 2010." Convention on Biological Diversity. https://www.cbd.int/kb/record/decision/12273.

Elvidge, Christopher D., Mikhail Zhizhin, Tilottama Ghosh, Feng-Chi Hsu, and Jay Taneja. 2021. "Annual Time Series of Global VIIRS Nighttime Lights Derived from Monthly Averages: 2012 to 2019." *Remote Sensing* 13 (5). https://doi.org/10.3390/rs13050922.

Hansen, M. C., P. V. Potapov, R. Moore, M. Hancher, S. A. Turubanova, A. Tyukavina, D. Thau, et al. 2013. "High-Resolution Global Maps of 21st-Century Forest Cover Change." *Science* 342 (6160): 850–53. https://doi.org/10.1126/science.1244693.

Harwood, Tom, Chris Ware, Andrew Hoskins, Simon Ferrier, Alex Bush, Maciej Golebiewski, Samantha Hill, et al. 2022. "BHI v2: Biodiversity Habitat Index: 30s Global Time Series." https://data.csiro.au/collection/csiro%3A54237v1.

Hill, Samantha L.L., Ricardo Gonzalez, Katia Sanchez-Ortiz, Emma Caton, Felipe Espinoza, Tim Newbold, Jason Tylianakis, Jörn P. W. Scharlemann, Adriana De Palma, and Andy Purvis. 2018. "Worldwide Impacts of Past and Projected Future Land-Use Change on Local Species Richness and the Biodiversity Intactness Index." *BioRxiv*, January, 311787. https://doi.org/10.1101/311787.

Hudson, Lawrence N, Tim Newbold, Sara Contu, Samantha L. L. Hill, and et al. 2016. "The 2016

Release of the PREDICTS Database." Natural History Museum. https://doi.org/10.5519/0066354.

Hudson, Lawrence N., Tim Newbold, Sara Contu, Samantha L. L. Hill, Igor Lysenko, Adriana De Palma, Helen R. P. Phillips, et al. 2014. "The PREDICTS Database: A Global Database of How Local Terrestrial Biodiversity Responds to Human Impacts." *Ecology and Evolution* 4 (24): 4701–35. https://doi.org/10.1002/ece3.1303.

———. 2017. "The Database of the PREDICTS (Projecting Responses of Ecological Diversity In Changing Terrestrial Systems) Project." *Ecology and Evolution* 7 (1): 145–88. https://doi.org/10.1002/ece3.2579.

K. Karra, C. Kontgis, Z. Statman-Weil, J. C. Mazzariello, M. Mathis, and S. P. Brumby. 2021. "Global Land Use / Land Cover with Sentinel 2 and Deep Learning." In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 4704–7. https://doi.org/10.1109/IGARSS47720.2021.9553499.

Lesiv, Myroslava, Dmitry Schepaschenko, Marcel Buchhorn, Linda See, Martina Dürauer, Ivelina Georgieva, Martin Jung, et al. 2022. "Global Forest Management Data for 2015 at a 100 m Resolution." *Scientific Data* 9 (1): 199. https://doi.org/10.1038/s41597-022-01332-3.

Newbold, Tim, Lawrence N. Hudson, Andrew P. Arnell, Sara Contu, Adriana De Palma, Simon Ferrier, Samantha L. L. Hill, et al. 2016. "Has Land Use Pushed Terrestrial Biodiversity beyond the Planetary Boundary? A Global Assessment." *Science* 353 (6296): 288–91. https://doi.org/10.1126/science.aaf2201.

OECD. 2019. "The Post-2020 Biodiversity Framework: Targets, Indicators and Measurability Implications at Global and National Level."

OpenStreetMap contributors. 2021. *Planet Dump Retrieved from Https://Planet.Osm.Org*.

Scholes, R. J., and R. Biggs. 2005. "A Biodiversity Intactness Index." *Nature* 434 (7029): 45–49. https://doi.org/10.1038/nature03289.

Weiss, D. J., A. Nelson, H. S. Gibson, W. Temperley, S. Peedell, A. Lieber, M. Hancher, et al. 2018. "A Global Map of Travel Time to Cities to Assess Inequalities in Accessibility in 2015." *Nature* 553 (7688): 333–36. https://doi.org/10.1038/nature25181.

WorldPop, Center for International Earth Science Information Network (CIESIN), School of Geography and Environmental Science, University of Southampton, Department of Geography and Geosciences, University of Louisville, and Departement de Geographie, Universite de Namur. 2018. *Global High Resolution Population Denominators Project - Funded by The Bill and Melinda Gates Foundation (OPP1134076)*. worldpop.org.

Acknowledgements

Funding for this work was provided by the Microsoft Corporation. We thank Neil Burgess, Sam Hill, Michelle Harrison, and Calum Maney at UNEP-WCMC for their expertise and support.