Global 1 km land cover for ecological modelling from very high resolution imagery

Elia Lo Parrino[1,*,°], Andrea Simoncini[1,°], Gentile Francesco Ficetola[1,2], Mattia Falaschi[1]

[1] Department of Environmental Science and Policy, Università degli Studi di Milano, via Celoria 10, 20133, Milan, Italy

[2] University Grenoble Alpes, Laboratoire d'Écologie Alpine (LECA), F-38000, Grenoble, France

*Corresponding author: Elia Lo Parrino, Department of Environmental Science and Policy, Università degli Studi di Milano, via Celoria 10, 20133, Milan, Italy; elia.loparrino@unimi.it

°These authors contributed equally.

ABSTRACT

Land cover strongly influences species distributions and ecological processes, yet global datasets often have insufficient spatial resolution to capture fine-scale heterogeneity. This limitation can reduce the accuracy of biodiversity and environmental modelling.

We developed a new global land cover dataset at ~1 km resolution by aggregating the 10 m ESA WorldCover 2021 product to a 30-arcsecond grid. The original categorical map was processed to generate continuous layers representing the percentage cover of 11 land cover categories—trees, shrublands, grasslands, croplands, built-up, bare/sparse vegetation, snow/ice, permanent water bodies, herbaceous wetlands, mangroves, and mosses/lichens—plus a layer representing the proportion of emerged land. OpenStreetMap landmass boundaries were used to mask marine cells. The dataset covers the global land surface from 60°S to 84°N and 180°W to 180°E. It provides improved representation of fine-scale habitat heterogeneity compared to widely used coarser products. We demonstrate its utility by comparing species distribution models built with our dataset against models developed using an established global land cover dataset, showing higher predictive performance with the new data.

By combining global extent with enhanced spatial detail, this dataset enables more accurate assessments of species–environment relationships, biodiversity patterns, and land-use impacts. It is intended for integration into ecological, macroecological, and conservation models. The dataset is openly available as GeoTIFF rasters for the year 2021.

Background

Land cover changes are reshaping global biodiversity: they can modify key habitats for threatened species (Cordier et al. 2021; Riva et al. 2023), favour the spread of invasive species (Ficetola et al. 2010), concur in causing range shifts (Poniatowski et al. 2020), modify landscape connectivity (Leonard et al. 2017), and more broadly drive environmental suitability for organisms (Falaschi et al. 2025). To reconstruct and/or anticipate such effects, models assessing species probability of occurrence as a function of land cover gained momentum in the last decade, potentially providing powerful tools for macroecology and conservation science (Torres et al. 2018; Liu et al. 2020; Lo Parrino et al. 2023).

Thus, global standardized maps are essential to develop Species Distribution Models (SDMs) and capture the relationships between environmental features and species' occurrence at large scales. While climatic maps are globally harmonized and constantly updated (e.g., CHELSA, WorldClim; Fick & Hijmans 2017; Karger et al. 2017), land cover maps are often more local. For instance, the CORINE Land Cover is a 100 m resolution map with several temporal updates, but it covers only European countries. Remote sensing has been widely used to map and characterize various landscape features, such as aquatic habitats (Wu et al. 2024), forests (Aziz et al. 2024), agricultural areas (Owusu et al. 2024), and urban environments (Chen et al. 2024). Despite recent progress, many global land cover datasets suffer from outdated temporal baselines, inconsistent classification schemes across regions, or low performance in heterogeneous landscapes such as tropical forests or montane areas. Moreover, data are collected by different satellite sensor systems at various spatial, temporal, and spectral resolutions and maps are generated with different classification approaches (Chen et al. 2017). Furthermore, land use maps often provide categorical outputs, masking intra-pixel variability, hence failing to represent spatially complex areas and offering limited insights into habitat mosaics that are ecologically crucial, particularly for edge species or generalists (Hansen et al. 2002; Blanco et al. 2013; Tuanmu & Jetz 2014). This causes the overrepresentation of dominant land cover types and false absences for less frequent land cover categories. Until recently, most global land cover maps were derived from satellite data at a coarse spatial resolution (300-1000 m; Ban et al. 2015). However, recent satellite missions provided data at a much finer resolution, such as 10 m (Xu et al. 2024). Incorporating accurate and detailed data enables better assessments of landscape heterogeneity, especially for landscape elements that are not mapped at coarse resolutions (e.g., small water bodies; Céréghino et al. 2008) and for species affected by small habitat patches (e.g., arthropods; Norhisham et al. 2024). Additionally, land cover often drives the distribution of species at a finer grain compared to climate and it is thus essential that it captures well the fine-scale heterogeneity (Nieto-Lugilde et al. 2015).

Most widely used global datasets for bioclimatic variables, hydrography, and land cover have ~1 km resolution (Lehner et al. 2008; Tuanmu & Jetz 2014; Fick & Hijmans 2017; Karger et al. 2017). Here, we provide a ~1 km grid resolution global land cover dataset for 11 categories, derived from the very high-resolution (~10 m) ESA WorldCover 2021, a global land cover map based on Sentinel-1 and Sentinel-2 data (Zanaga et al. 2022). The dataset we propose bridges a key resolution gap between detailed but local 10–30 m products and coarse global maps (starting resolution >300 m), offering a harmonized, high-quality representation of global land cover. The present maps can be easily integrated with other widely used environmental layers in global-scale modeling, requiring minimal computational effort. The 1 km resolution is also compatible with the positional accuracy of most occurrence records in biodiversity repositories such as iNaturalist and the Global Biodiversity Information Facility (GBIF). By calculating land cover from very high-resolution satellite data, we retain crucial information about habitat heterogeneity, reducing biases in downstream analyses such as SDMs or ecological niche estimates.

Data acquisition and processing

The original ESA WorldCover 2021 included the following categories, that we retained: "Tree cover", "Shrubland", "Grassland", "Cropland", "Built-up", "Bare/sparse vegetation", "Snow and Ice", "Permanent water bodies", "Herbaceous wetland", "Mangrove", and "Moss and lichen". All

WorldCover maps were provided in the EPSG:4326 WGS84 coordinate reference system, that we also employed. The source dataset is divided into 18 macrotiles (60×60°), each composed of multiple 3×3° tiles at 0.3 arcseconds (~ 10 m) resolution. Each WorldCover pixel indicated the dominant land cover. The overall accuracy of the ESA WorldCover 2021, as detailed from the Product Validation Report (available at: https://worldcover2021.esa.int/), was 76.7%.

Our goal was to generate global maps at 30 arcseconds (~1 km) for each category, where each pixel represents the percentage surface of the target land cover class (Fig. 1). To achieve this, we cropped each original 3×3° tile to the terrestrial areas using the landmass shapefile from OpenStreetMap (https://osmdata.openstreetmap.de/data/land-polygons.html) to avoid considering marine cells and to provide maps tailored for analyses on terrestrial ecosystems. Then, we aggregated the result by a factor of 100, creating a 30 arcseconds map where pixel values described the count of 10 m cells for each class. Lastly, we divided values by 100 to obtain the percentage cover per pixel. For the "Permanent waterbodies" class, marine cells were excluded using a negative internal buffer (~20 m, i.e., 2 times the resolution of the original raster) from the landmass shapefile. All tiles were then merged into a single global raster at 30 arcseconds resolution (~1 km) for each land cover class. Thus, the values of the pixels in the final maps represent the percentage of each land cover relative to the total surface of the pixel. Additionally, we created a global raster representing the percentage of landmass per pixel, enabling, when needed, the estimation of the percentage of each land cover relative to the surface of landmass per pixel. The analysis was developed using the 'terra' R package (Hijmans, 2023) and the workflow is summarized in Fig. 1.

Comparison with an established land cover dataset

We compared the new set of produced variables with the widely used EarthEnv Global 1 km consensus land cover (hereafter "EE") of Tuanmu & Jetz 2014. We compared the performance of alternative species distribution models (SDMs) including climate from CHELSA (Karger et al. 2017), while land cover was either from EE or the land cover presented here (termed "ESA"). The

EE is one of the most used land cover datasets, with 384 citations on Scopus (accessed on 04/07/2025). It is a consensus map derived from two different global land cover products, GlobCover (300 m grid resolution) and MODIS2005 (500 m). It provides global maps at ~1 km resolution for 12 land cover categories. This consensus is based on products that are now outdated and at a coarse resolution, limiting the ability to capture fine-scale heterogeneity. In contrast, the dataset we present is derived from a source with a higher resolution (10 m), more recent imagery (2021) and dual-sensor integration (Sentinel-1 and Sentinel-2). By aggregating to 1 km resolution, we ensure compatibility with widely used bioclimatic and hydrological layers, while retaining more accurate information on the fine-scale variation of land cover.

We modelled habitat suitability for four species: *Nepa cinerea* (class Esapoda, phylum Arthropoda), *Bufo bufo* (Amphibia, Chordata), *Dryas octopetala* (Magnoliopsida, Magnioliophyta), and *Sylvilagus floridanus* (Mammalia, Chordata). These species belong to heterogeneous taxonomic groups and have variable geographic distributions and degrees of ecological specialization. We retrieved observations of the species from the GBIF website (https://www.gbif.org). Besides land cover variables either from EE or ESA, all models considered four fundamental climatic descriptors from CHELSA (Karger et al. 2017): annual mean temperature, temperature seasonality, annual total precipitation, and precipitation seasonality. SDMs were run using the Maxent algorithm (Phillips et al. 2004). A four-fold cross-validation consisting of spatially independent blocks was used to test model performance (Muscarella et al. 2014) and the performance on the withheld datasets was estimated using the Continuous Boyce Index (CBI, Hirzel et al. 2006) and the Area Under the receiver operating characteristic Curve (AUC, Fielding & Bell 1997). Each of the four blocks was used once for testing, while the remaining three blocks were used for model training (i.e., k-fold cross-validation). For each species, we performed two alternative SDMs, with the same model hyperparameters but different land cover sets (climate + EE, climate + ESA). Then, we assessed the performance metrics to assess the differences between the two datasets of land cover used. Lastly, the permutation importance (i.e., the drop in AUC after the random permutation of a given variable; Smith & Santos 2020) was evaluated for each of the tested species and land cover dataset.

We found a superior predictive performance of SDMs developed with the ESA land cover compared to EE for all species and both according to the CBI (Fig. 2) and the AUC (Fig. S1), with marked differences in three out of four species (Fig. 2; Fig. S1; Supplementary methods). Additionally, based on a priori expert-based knowledge on the environmental requirements of the selected species, permutation importance revealed a higher ecological realism of SDMs developed with the ESA land cover (Tables S1-S4). For instance, the permutation importance of water increased from 12.1% to 27.8% for *N. cinerea* (freshwater insect) and from 0% to 12.8% for *B. bufo* (freshwater-breeding amphibian). The importance of grassland for *D. octopetala* (a high-elevation flowering plant) increased from 0.7% to 29.7%, indicating a better ability of the ESA dataset to capture the spatial distribution of key proximal variables.

Usage notes

The final set of variables has an overall size of 2.22 GB. The downloaded rasters can be incorporated into coding frameworks in R by loading them with the 'terra' R package (Hijmans, 2023) using the 'rast' function. Maps are provided in the WGS84 (EPSG:4326) coordinate reference system. When using these variables for SDMs, we recommend including only uncorrelated variables that are deemed relevant for the species ecology and to integrate them with further predictors (e.g., climatic) that capture other nuances of the modelled species' niches (Dormann et al. 2013; Guisan et al. 2017; Fourcade et al. 2018).

Discussion

Our dataset contributes to bridging the gap between high-resolution remote sensing products and the standard resolution used in global biodiversity and environmental modelling, offering a more precise yet computationally tractable alternative to existing coarse resolution products. This product fills a critical gap for large-scale studies, especially those requiring integration with other 1 km environmental layers, and we showed that it can increase the predictive ability and ecological realism of SDMs.

We acknowledge that the thematic resolution of the dataset should be expanded in future developments, for instance, discriminating the different forest and/or vegetation types. However, its capacity to represent sub-pixel heterogeneity offers an important advancement over currently available maps, enabling finer ecological inference in fragmented and transitional landscapes. In particular, the main novelties and advantages of our maps include:

- Improved representation of small and heterogeneous features, such as riparian zones, wetlands, and urban mosaics. For instance, accurate representations of potential wildlife habitats within cities are needed to better assess urban biodiversity patterns (Gelmi-Candusso et al. 2024). Additionally, until now, the lack of fine resolution global maps of aquatic habitats hindered the reliability of niche modelling for freshwater species (Lo Parrino et al. 2023);
- Availability of fractional cover information per class per pixel, avoiding the oversimplification of single-class assignments. Coarse characterizations may limit our understanding of biodiversity patterns across spatial scales (Gelmi-Candusso et al. 2024). Maps representing the proportion of the focal land use class within each cell likely provide a better representation of habitat availability in the context of niche modelling. Moreover, the basic maps here provided represent the proportion of each land class on the total cell surface (~1 km^2). This approach may be relevant in certain contexts, making this dataset flexible and suitable for multiple applications;
- Inclusion of a landmass mask and filtered water body layer to improve accuracy in coastal and aquatic contexts. This is particularly relevant for organisms exploiting inland waters close to the coast, as some maps classify seawater and freshwater under the same category (Xu et al. 2020), leading to inaccurate approximations of habitat availability for freshwater species.

This resource thus has the potential to support a wide range of macroecological and conservation applications, including habitat prioritization, protected area planning, and global biodiversity monitoring. As biodiversity faces increasing pressures from land use changes, global high-resolution tools, such as the maps presented here, are essential for informed decision-making and effective conservation planning.

Figures

Figure 1. Schematic representation of the workflow used to produce a set of global land cover maps at ~ 1 km grid resolution for macroecological and biogeographic applications, starting from the global ESA WorldCover 2021 10 m satellite-based land cover.

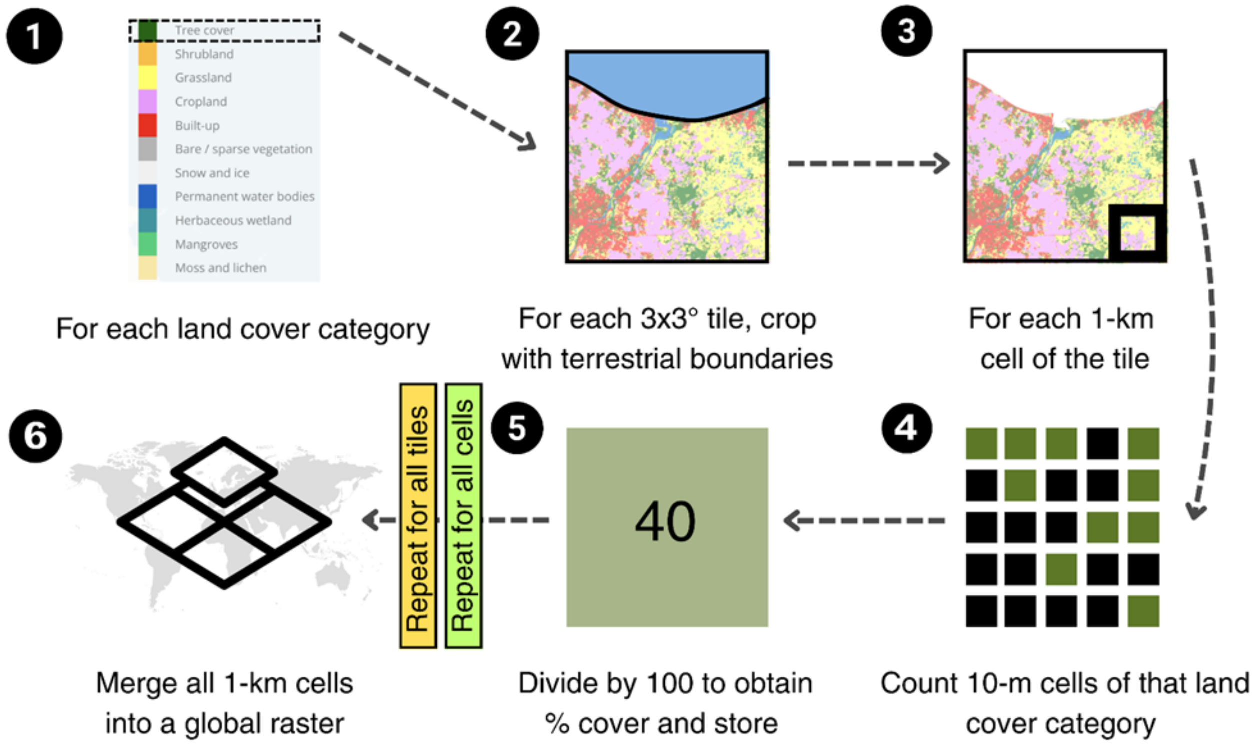
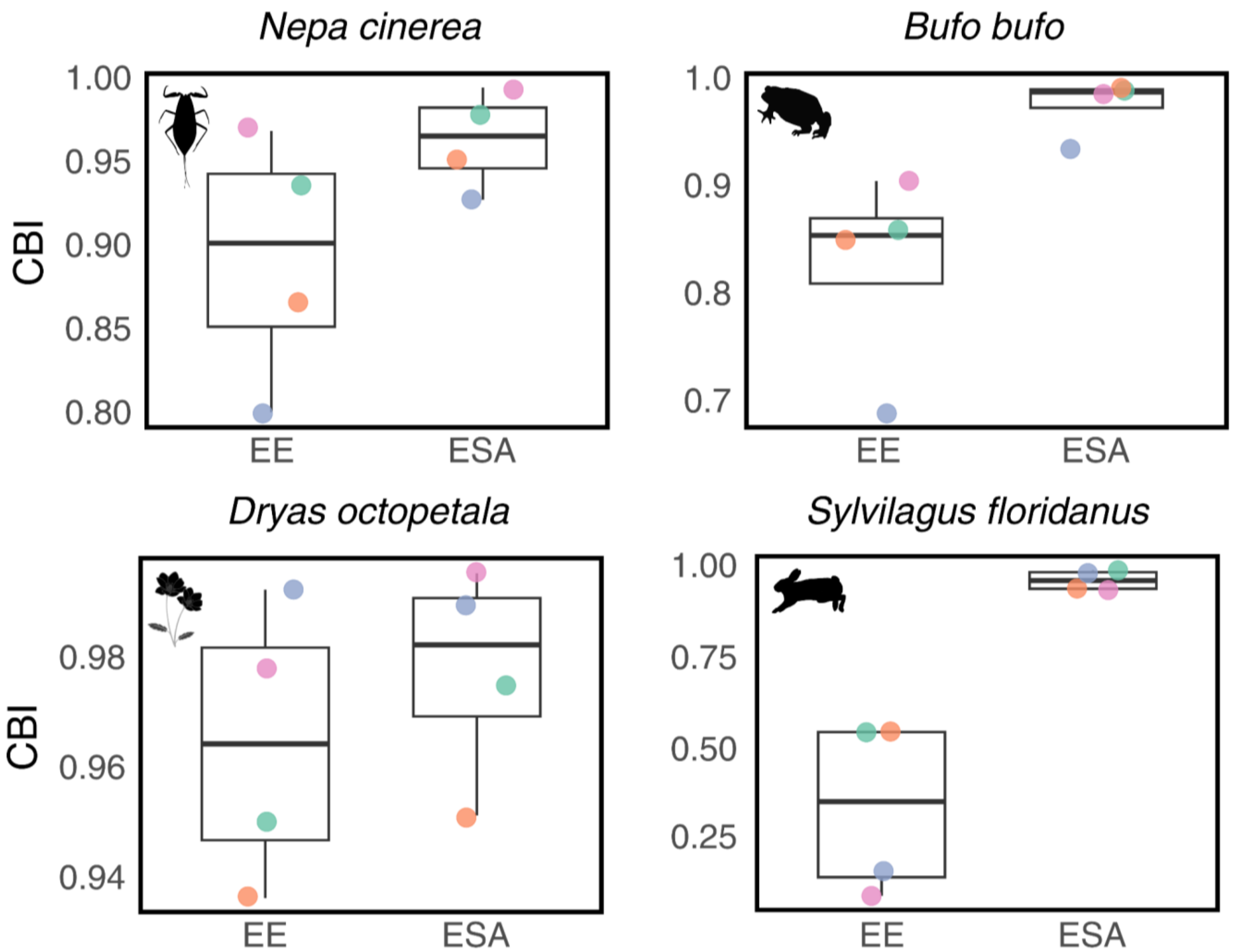


Figure 2. Comparison of the predictive performance of species distribution models developed with climatic and land cover variables, the latter either from the Consensus Land Cover (EE) of Tuanmu & Jetz 2014, or from the set presented in this study and derived from the ESA WorldCover 2021. The metric used is the Continuous Boyce Index (CBI, ranging from -1 to 1). Points represent the performance on withheld test datasets across four spatially independent cross-validation replicates (indicated with different colours). See the text and the supplementary methods for further details and results according to another metric. Species silhouettes from PhyloPic (https://www.phylopic.org).



References

Aziz G, Minallah N, Saeed A, Frnda J, Khan W. 2024. Remote sensing based forest cover classification using machine learning. Scientific Reports 14:69.

Ban Y, Gong P, Giri C. 2015. Global land cover mapping using Earth observation satellite data: Recent progresses and challenges. ISPRS Journal of Photogrammetry and Remote Sensing 103:1–6.

Blanco PD et al. 2013. A land cover map of Latin America and the Caribbean in the framework of the SERENA project. Remote Sensing of Environment 132:13–31.

Céréghino R, Biggs J, Oertli B, Declerck S. 2008. The ecology of European ponds: Defining the characteristics of a neglected freshwater habitat. Hydrobiologia 597:1–6.

Chen B, Huang B, Xu B. 2017. Multi-source remotely sensed data fusion for improving land cover classification. ISPRS Journal of Photogrammetry and Remote Sensing 124:27–39.

Chen G, Zhou Y, Voogt JA, Stokes EC. 2024. Remote sensing of diverse urban environments: From the single city to multiple cities. Remote Sensing of Environment 305:114108.

Cordier JM, Aguilar R, Lescano JN, Leynaud GC, Bonino A, Miloch D, Loyola R, Nori J. 2021. A global assessment of amphibian and reptile responses to land-use changes. Biological Conservation 253:108863.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., …, Lautenbach, S. 2013. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography 36:27–46.

Falaschi M, Simoncini A, Ancillotto L, Viviano A, Menchetti M, Mazza G, Mori E. 2025. Crossing borders: Connectivity analyses reveal potential patterns of range expansion of the Northern raccoon in Europe. NeoBiota 99:71–92.

Ficetola GF, Maiorano L, Falcucci A, Dendoncker N, Boitani L, Padoa-Schioppa E, Miaud C, Thuiller W. 2010. Knowing the past to predict the future: Land-use change and the distribution of invasive bullfrogs. Global Change Biology 16:528–537.

Fick SE, Hijmans RJ. 2017. WorldClim 2: new 1 km spatial resolution climate surfaces for global land areas. International Journal of Climatology 37:4302–4315.

Fielding, A. H., & Bell, J. F. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environmental conservation 24:38-49.

Fourcade, Y., Besnard, A. G., & Secondi, J. 2018. Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. Global Ecology and Biogeography 27:245-256.

Gelmi-Candusso TA, Rodriguez P, Fidino M, Rivera K, Lehrer EW, Magle S, Fortin MJ. 2024. Leveraging open-source geographic databases to enhance the representation of landscape heterogeneity in ecological models. Ecology and Evolution 14:e70402.

Guisan, A., Thuiller, W., & Zimmermann, N. E. 2017. Habitat suitability and distribution models: with applications in R. Cambridge University Press.

Hansen MC, DeFries RS, Townshend JRG, Sohlberg R, Dimiceli C, Carroll M. 2002. Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data. Remote Sensing of Environment 83:303–319.

Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C., & Guisan, A. 2006. Evaluating the ability of habitat suitability models to predict species presences. Ecological modelling 199:142-152.

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., ... & Kessler, M. 2017. Climatologies at high-resolution for the earth’s land surface areas. Scientific data 4:170122.

Lehner B, Verdin K, Jarvis A. 2008. New global hydrography derived from spaceborne elevation data. Eos 89:93–94.

Leonard PB, Sutherland RW, Baldwin RF, Fedak DA, Carnes RG, Montgomery AP. 2017. Landscape connectivity losses due to sea level rise and land use change. Animal Conservation 20:80–90.

Liu C, Wolter C, Xian W, Jeschke JM. 2020. Most invasive species largely conserve their climatic niche. Proceedings of the National Academy of Sciences of the United States of America 117:23643–23651.

Lo Parrino E, Falaschi M, Manenti R, Ficetola GF. 2023. All that changes is not shift: methodological choices influence niche shift detection in freshwater invasive species. Ecography 2023:e06432.

Muscarella, R., Galante, P. J., Soley-Guardia, M., Boria, R. A., Kass, J. M., Uriarte, M., & Anderson, R. P. 2014. ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. Methods in Ecology and Evolution 5:1198–1205.

Nieto-Lugilde, D., Lenoir, J., Abdulhak, S., Aeschimann, D., Dullinger, S., Gégout, J. C., ... & Svenning, J. C. 2015. Tree cover at fine and coarse spatial grains interacts with shade tolerance to shape plant species distributions across the Alps. Ecography 38:578–589.

Norhisham AR, Yahya MS, Nur Atikah S, Jamian S, Bach O, McCord M, Howes J, Azhar B. 2024. Non-crop plant beds can improve arthropod diversity including beneficial insects in chemical-free oil palm agroecosystems. Cogent Food and Agriculture 10:2367383.

Owusu A et al. 2024. A framework for disaggregating remote-sensing cropland into rainfed and irrigated classes at continental scale. International Journal of Applied Earth Observation and Geoinformation 126:103607.

Phillips, S. J., Dudík, M., & Schapire, R. E. 2004. A maximum entropy approach to species distribution modeling. In Proceedings of the twenty-first international conference on Machine learning: 83.

Poniatowski D, Beckmann C, Löffler F, Münsch T, Helbing F, Samways MJ, Fartmann T. 2020. Relative impacts of land-use and climate change on grasshopper range shifts have changed over time. Global Ecology and Biogeography 29:2190–2202.

Riva F, Barbero F, Balletto E, Bonelli S. 2023. Combining environmental niche models, multi-grain analyses, and species traits identifies pervasive effects of land use on butterfly biodiversity across Italy. Global Change Biology 29:1715–1728.

Smith, A. B., & Santos, M. J. 2020. Testing the ability of species distribution models to infer variable importance. Ecography 43:1801-1813.

Torres U, Godsoe W, Buckley HL, Parry M, Lustig A, Worner SP. 2018. Using niche conservatism information to prioritize hotspots of invasion by non-native freshwater invertebrates in New Zealand. Diversity and Distributions 24:1802–1815.

Tuanmu MN, Jetz W. 2014. A global 1 km consensus land-cover product for biodiversity and ecosystem modelling. Global Ecology and Biogeography 23:1031–1045.

Wu Y, Knudby A, Pahlevan N, Lapen D, Zeng C. 2024. Sensor-generic adjacency-effect correction for remote sensing of coastal and inland waters. Remote Sensing of Environment 315:114433.

Xu P, Herold M, Tsendbazar NE, Clevers JGPW. 2020. Towards a comprehensive and consistent global aquatic land cover characterization framework addressing multiple user needs. Remote Sensing of Environment 250:112034.

Xu P et al. 2024. Comparative validation of recent 10 m-resolution global land cover maps. Remote Sensing of Environment 311:114316.

Zanaga D, Van De Kerchove R, Daems D, De Keersmaecker W, Brockmann C, Kirches G, Wevers J, Cartus O, Santoro M, Fritz S. 2022. ESA WorldCover 10 m 2021 v200. Zenodo. doi: https://doi.org/10.5281/zenodo.7254221