Supporting Documentation

Data Overview

The data are outputs from work based on ensemble Species Distribution Models (SDMs) that were used to explore how two kinds of socio-political considerations affect bird and mammal biodiversity under climate change. First, we provide national-level data for 189 countries of modelled changes to species richness (projected to 2070), which we used to explore how national level socioeconomic factors, such as per capita GDP, per capita CO_2 emissions, and governance quality are related to measures of biodiversity loss from climate change. This is included in one of our datasets, 'Climate_impacts_by_country.csv'. The remaining datasets consider how species distributions could change in relation to political borders. We estimate the number of species whose ranges currently intersect each political border globally ('transboundary_richness.csv'), and the number of species whose ranges are projected to cross each political border under climate change ('transboundary_range_shifts.csv').

These data are global in spatial scope, and cover >12,700 terrestrial mammals and birds (~80% of them), excluding species with restricted ranges that are most likely determined by factors other than climate.

Data collection/generation methods

These data were generated using ensemble SDMs. Species distribution data were first obtained from the IUCN Red List of Threatened Species and BirdLife International and Handbook of the Birds of the World. Our model ensemble contained four model types (Generalised Additive Models [GAMs], Generalised Linear Models [GLMs], Random Forests [RFs] and Boosted Regression Trees [BRTs]). These models were then trained on five climatic predictor variables (mean annual temperature, temperature seasonality, precipitation in the wettest and driest months, and precipitation seasonality), and were projected to 2070 to produce maps of species future climatic niches under four different climate scenarios used by the Intergovernmental Panel on Climate Change (RCP 2.6, RCP 4.5, RCP 6.0, and RCP 8.5). Three General Circulation Models (GCMs) were used to generate future climate projections: HadGEM2-ES, CCSM4, and MIROC-ESM-CHEM. To account for the spatial dependence in our models we split gridded presence/absence data for each species into ten spatially disaggregated blocks. Non-contiguous portions of the world's terrestrial ecoregions were used as the sampling units to divide the data; these units were then grouped into 10 blocks using the *blockTools* package in R such that the total area and mean bioclimate was approximately equal in each block, and that each block contained the full range of bioclimates. By splitting the data into ten blocks, we were able to use 10-fold cross-validation to assess model performance. Each block was left out in turn to be used as a testing dataset, and models were trained on the remaining 90% of data. Model performance was then assessed using the area under the receiver operator characteristic (ROC) curve (AUC), which tests for discrimination ability.

For each species, we produced an ensemble of 120 projections (4 model types x 10 data blocks x 3 GCMs) for each of the four climate scenarios. We then took the mean presence/absence value across the ensemble, weighted by each model's AUC to give more weight to better performing models. This produced a final projected distribution for each species under each climate scenario.

By stacking projections for each species together, we calculated total species richness, now and in 2070, for each grid cell under each climate scenario. We then aggregated these grid-cell level projections to the national level, by taking the mean across all grid cells in each country. This enabled us to calculate national-level percentage change in bird and mammals species richness for each country, for each climate scenario. These data are presented in 'Climate_impacts_by_country.csv', separately for birds and mammals and also for the two taxonomic groups combined. We also provide a national-level metric of governance quality. This was calculated by taking the mean across all six Governance Indicators provided by the World Bank (http://info.worldbank.org/governance/wgi/). For our analysis data were taken from the year 2018.

To calculate the current number of species whose ranges intersect political borders around the world, we intersected species distribution data with spatial data on the world's political borders obtained via the R package 'rworldmap' (https://cran.r-project.org/web/packages/rworldmap/rworldmap.pdf). This 'transboundary species richness' was calculated using the 'raster' R package (https://cran.r-project.org/web/packages/raster/raster.pdf). Data are presented as the number of species whose ranges intersect each political border in the data file 'Transboundary _richness.csv'.

To calculate the number of transboundary range shifts for each border, we generated species x country matrices for current and future climate conditions, indicating whether or not each species was present or not in each country for both time periods. By comparing the two matrices for each land border we calculated the number of species projected to move into neighbouring countries as their ranges shift. This information is presented as the total number of species projected to cross each border in either direction, in the data file 'transboundary_range_shifts.csv'.

Details of data structure and units of recorded values

This dataset comprises three csv files entitled ‘Climate_impacts_by_country.csv’, ‘transboundary_richness.csv’ and ‘transboundary_range_shifts.csv’.

1. Climate_impacts_by_country.csv
    - This dataset has 189 rows (one for each country) and 31 columns.
    - The first column contains the country code (ISO3), the standardised three-letter code used to represent countries.
    - The last column contains a governance score, which we obtained by calculating the mean across all six World Bank Governance Indicators (possible range -2.5 to 2.5). See http://info.worldbank.org/governance/wgi/
    - The intervening columns contain data on our modelled impacts of climate change on the countries. The column names indicate the nature of the data and consist of three parts separated by an underscore (_):
        1. The first part indicates the climate scenario. This takes one of ‘current’ (present day conditions), ‘rcp26’ (low emissions scenario RCP 2.6 projected to 2070), ‘rcp45’ (moderate emissions scenario RCP 4.5 projected to 2070), ‘rcp60’ (moderate emissions scenario RCP 6.0 projected to 2070) or ‘rcp85’ (high emissions scenario RCP 8.5 projected to 2070).
        2. The second part indicates the taxonomic group, and can be either ‘mammal’, ‘bird’ or ‘both’
        3. The third part indicates the kind of measurement. This can either be ‘SR’, short for species richness (in which case the unit is number of species), or ‘pctChange’, short for percentage change in species richness (i.e. the future richness for this climate scenario minus the current richness, divided by the current richness).

        Example: the column ‘rcp85_bird_SR’ contains data for bird species richness in each country under the high emissions scenario RCP 8.5.

2. Transboundary_richness.csv
    - This dataset has 327 rows (one for each border) and three columns.
    - The first column, ‘border’, is a list of each border in the dataset, indicated by the pairwise combination of ISO3 country codes either side of the border.
    - The second column, ‘transboundary_richness’, indicates the number of bird and mammal species whose distributions currently intersect that border.
    - The third column, ‘transboundary_richness_threatened’, indicates the number of threatened species (listed as either vulnerable [VU], endangered [EN] or critically endangered [CR] by the IUCN Red List of

Threatened Species) whose distributions currently intersect that border.

3. Transboundary_range_shifts.csv
    - This dataset has six columns.
    - The first column, ‘border’, is a list of each border in the dataset with transboundary range shifts, indicated by the pairwise combination of ISO3 country codes either side of the border.
    - The next four columns include data for the number of projected transboundary range shifts across each border, broken down by taxonomic group (mammals or birds) and climate scenario (moderate [RCP 4.5] or high [RCP 8.5])
    - The final column (‘barrier_species’) indicates, for borders that are fortified with a physical barrier such as a wall or fence, the number of non-flying mammals (i.e. mammals that are not bats) whose ranges are projected to cross it under climate change (high emissions scenario, RCP 8.5). The majority of borders do not have such physical barriers; these are indicated with a NA value.

Other information

For context and further explanations of methodology see *Materials and Methods* in the open-access publication:

Titley M.A., Butchart S. H. M., Jones, V.R., Whittingham M.J, Willis, S.G. Global inequities and political borders challenge nature conservation under climate change, *Proceedings of the National Academy of Sciences, USA* (2021, *in press*)