This information product has been peer reviewed and approved for publication as a preprint by the U.S. Geological Survey.

BioShifts v2 and BioShiftR: tools for navigating and analysing species range shift data

Authors

Jake A. Lawlor[1], Brunno F. Oliveira[2,3], Nikki A. Moore[1], Jonathan Lenoir[4], R.M.W.J. Bandara[5], Romain Bertrand[6], I-Ching Chen[7], Sarah Diamond[8], Gaël Grenouillet[6,9], Lesley Lancaster[10], Malin Pinsky[11], Jonathan Rolland[6],Brett Scheffers[12], Fabricio Villalobos[13], Barrett Wolfe[14], Jennifer M. Sunday[1†], Sarah R. Weiskopf[15†], Lise Comte[16†]

[†]Equal contribution last-authorship

Corresponding Author: Jake Lawlor; jake.lawlor@mail.mcgill.ca

Affiliations

1. Department of Biology, McGill University, Montreal, QC, Canada
2. Centre de Synthèse et d'Analyse sur la Biodiversité, Fondation pour la Recherche sur la Biodiversité, Montpellier 34000, France
3. Departamento de Ecologia, Universidade Federal do Rio Grande do Norte, Natal 59072-970, Brazil
4. UMR CNRS 7058 "Ecologie et Dynamique des Systèmes Anthropisés" (EDYSAN), Université de Picardie Jules Verne, Amiens, France
5. Graduate Program in Ecology and Evolution, School of Environmental and Biological Sciences, Rutgers, The State University of New Jersey, New Brunswick, New Jersey, USA
6. Centre de Recherche sur la Biodiversité et l'Environnement (CRBE), Université de Toulouse, CNRS, IRD, Toulouse INP, Toulouse, France
7. Department of Life Sciences, National Cheng Kung University
8. Department of Biology, Case Western Reserve University, Cleveland, Ohio, USA
9. Institut Universitaire de France, Paris, France
10. School of Biological Sciences, University of Aberdeen, Aberdeen, United Kingdom AB24 2TZ
11. Department of Ecology & Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA 95060 USA
12. Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL 32611
13. Red de Biología Evolutiva, Instituto de Ecología A.C., Xalapa 91073, Mexico

14. Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia, 7004
15. U.S. Geological Survey National Climate Adaptation Science Center
16. Conservation Science Partners, Inc., Truckee, California, USA

Acknowledgements

This work is a product of the BIOSHIFTS working group funded by the Centre for the Synthesis and Analysis of Biodiversity (CESAB), a key program supported by the French Foundation for Research on Biodiversity (FRB). We are grateful for feedback from the larger BIOSHIFTS working group and with attendees of the BioShiftR workshop at the Species on the Move Conference. We thank Shannon Bayliss for her help developing the BioShiftR logo. We thank Ciara Johnson, Jessica Krippel, Yi-Wen (Ivory) Chen, Ailey Lindemann-Mitchell and Elly Sponagle for their help checking data from BioShifts and CoRE. JL and LC received additional funding throughout the FRAGSHIFTS project funded by the FRB, the French Office for Biodiversity (OFB) and the French Ministry for Ecological Transition (MTE). JAL received additional support for this work, including a Doctoral Research Grant from Fonds du Recherche du Quebec (FRQNT) (https://doi.org/10.69777/297888) and a Réal-Decoste Excellence Scholarship from FRQNT and Ouranos Inc (https://doi.org/10.69777/371714), and travel support from the Quebec Centre for Biodiversity Science (QCBS) and McGill University. NAM received funding support from the National Sciences and Engineering Research Council of Canada (NSERC) and travel support from the QCBS. JMS was supported by the NSERC Canada Research Chairs program. We also thank the Jules Verne University of Picardy (UPJV) for providing high-performance computational power through the cluster MatriCs. The platform MatriCS is cofinanced by the European Union with the European Regional Development Fund (FEDER) and the Hauts-De-France Regional Council, among others. This research was also enabled in part by support provided by Calcul Québec and the Digital Research Alliance of Canada (alliancecan.ca). Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Author Contributions

Authors from the core group (JAL, BFO, NAM, JL, JMS, SRW and LC) contributed to the project's development and methodology. JAL, NAM, LC, and SRW curated data. JL, LC, and GG acquired funding and administered the project. JAL created scripts to clean and process *BioShifts* data, calculated climate exposure metrics, developed the R package, and created the package webpage and related vignettes. NAM created scripts to collate and create spatial data. BFO contributed to the climate exposure code and completed taxonomic harmonization. JMS, SRW, and LC contributed supervision. JAL, BFO, BW, JR, SRW, and LC validated data or code. JAL, NAM and JMS created visualizations. JL provided resources to access high-performance computational power through the cluster MatriCs at UPJV. JAL, NAM, JL, JMS, SRW, and LC contributed to writing & editing. All authors contributed to the manuscript's conceptualization and review.

Data Availability Statement

All data contained within *BioShifts v2* are available and embedded within the `BioShiftR` package and are stored in the R package repository on GitHub (https://github.com/bioshifts/bioshiftR). Geospatial data included in the database are stored externally on Open Science Framework (https://osf.io/tp4hv/overview), with download functions in the R package prompting a local download in order to use the spatial data. A static version of the *BioShifts v2* database will be archived on FigShare upon acceptance of this article.

Data Sources (where appropriate)

The data included in *BioShifts v2* and `BioShiftR` come from the *BioShifts* dataset (https://doi.org/10.6084/m9.figshare.7413365) and the Contractions and Range Expansions (CoRE) repository (https://doi.org/10.5066/P99VP2TW). Species ranges were obtained from the International Union for Conservation of Nature (IUCN), BirdLife, GARD, BIEN, Butterfly Maps, Aquamaps or estimated using convex hulls based on Global Biodiversity Information Facility occurrences (see Table S3 for DOIs). Climatic data were obtained through CHELSAcruts for terrestrial species (https://www.doi.org/10.16904/envidat.159) and ORAS5 for marine species (https://doi.org/10.24381/cds.67e8eeb7).

Conflicts of Interest

The authors declare no conflicts of interest.

BioShifts v2 and `BioShiftR`: tools for navigating and analysing species range shift data

Abstract

1. Species are redistributing their ranges globally in response to climate change, but the fragmented and complex nature of the documented responses make them challenging to track and analyze. Synthesis of species' range shifts requires careful integration and consideration of varying methods and scales of study.
2. Here, we present an updated database that combines two previous data repositories (*BioShifts v1* and CoRE), complete with methodological information from original studies, spatial datasets of species and study locations, and associated climate exposure metrics calculated at multiple scales, together comprising a clean and interoperable relational database for future studies of range shifts.
3. We make these data accessible through a companion open-source software tool, `BioShiftR`, that simplifies the navigation and analysis of estimated range shift rates in light of varying methods and hypothesized climate drivers.
4. Together, these tools will facilitate tractable and robust downstream studies of range shifts, ultimately improving our understanding of species' biogeographical responses to climate change.

Keywords: Species Redistribution, Range Shifts, Biogeography, Climate Change Ecology, Synthesis, Scientific Software, R Package,

Introduction

Species are redistributing their ranges across terrestrial, freshwater, and marine realms in response to climate change, yet observed responses are highly heterogeneous in magnitude and direction (Lenoir et al., 2020; Rubenstein et al., 2023). As the scientific community grapples to understand the mechanisms underlying highly variable and idiosyncratic biotic responses (Lawlor et al., 2024), a comprehensive set of synthesized data and analytical tools are needed to test hypotheses. Several quantitative and qualitative syntheses compiling scientific studies and empirical reports of species range shifts have been published (Chen et al., 2011; Lenoir & Svenning, 2015; Parmesan et al., 1999; Poloczanska et al., 2013; Taheri et al., 2021; Vitasse et al., 2021), leading to the development of data repositories on biodiversity redistributions (Comte et al., 2020; Rubenstein et al., 2024). However, a centralized database of range shift estimates that integrates information about the estimation methods, potential drivers, and contextual metadata, organized through best practices in database management along with tools to facilitate their comparative analysis do not yet exist.

Integrating and appropriately analyzing range shift estimates across hundreds of studies is challenging due to the complexity of the processes that are being documented, as well as the different procedures through which estimates are made and synthesized. Range shift estimates stem from multiple studies that rely on a wide variety of data types, geographical contexts, and analytical methods, which can greatly impact the magnitude and directions of range shift rates that are detected (Brown et al., 2016; Fredston, 2024). Consequently, systematically accounting for – or at least compiling – metadata variables describing how range shifts were estimated is essential for appropriately synthesizing results across studies.

The *BioShifts* dataset (Comte et al., 2020) and the *Contractions and Range Expansions (CoRE)* repository (Rubenstein et al., 2024) represent two of the most comprehensive syntheses of global range shift estimates to date, but both repositories are difficult to use and integrate due to the complexity of their structures. Although these repositories contain similar information, they each include slightly different sets of range shift estimates from published literature, and are not interoperable due to inconsistencies in metadata collected. For example, although both repositories included information on how studies accounted for unbalanced sampling designs, the number of columns was not the same and the categories differed. In addition, the absence of dedicated analytical tools for organizing and analysing these datasets further limits their usability. Together, these challenges highlight the need for a standardized and well-structured database of range shift estimates including harmonized metadata and user-friendly analytical tools that enable testing hypotheses across heterogeneous data sources.

Although changes in environmental conditions are commonly hypothesized as drivers of range shifts, relevant exposure metrics are not always included in synthesis databases. If range shifts are driven by climate change, then information describing the direction and magnitude of climate shifts during the time period and across the spatial extent of a given study (e.g., climate change velocity *sensu* Loarie et al. 2009), are relevant for predicting and explaining biotic responses. However, quantifying climate exposure at scales relevant to individual species can be challenging, and the ecological relevance of exposure metrics often depends on the spatial resolution at which they are calculated (Da Silva & Diamond, 2024; Heikkinen et al., 2020). Studies conducted across large spatial extents are likely to experience spatially heterogeneous climate exposure, and average values within large study areas might misrepresent the climate pressure experienced by a species whose range extent encompasses only a portion of the study area. Therefore, calculations of climate change exposure across multiple spatial resolutions that are relevant to individual species within a given study area can provide a more meaningful basis for comparing climate-driven expectations to empirical range shift rates (Oliveira et al., 2026).

Here, we present the updated and expanded version of the *BioShifts* database (*BioShifts v2*), enhanced through best practices in data management, along with its companion R package, `BioShiftR`, a software tool to simplify access, navigation, and analysis of the database. *Bioshifts v2* includes over 31,000 range shift rate estimates for over 12,000 species, collated from 244 published scientific articles. Each estimate includes extensive metadata describing study design and estimation procedure (e.g., location, time of observations, method of survey), as well as standardized taxonomy that enables integration with external data sources. The database also includes spatial dataframes that link range shift rate estimates to study polygons and – when available – the intersection between study polygons and focal species' ranges (hereafter, *species-specific study polygons*). Supplementing the data collected from individual articles, the database also contains several derived variables capturing climate change exposure within each study polygon and species-specific study polygon over the relevant time periods (e.g., warming trends, rates of isotherm shifts) calculated from input data at multiple spatial resolutions. `BioShiftR` facilitates easy access to the database, with custom functions to download, filter, and select relevant range shift estimates and associated variables. Together, *Bioshifts v2* and `BioShiftR` provide direct access to the largest standardized repository of species' biogeographic responses to climate change currently available, offering a valuable resource for investigating the causes and consequences of biodiversity redistribution under climate change.

Creation of the *BioShifts v2* database

Here, we provide a brief overview of the steps involved in creating *BioShifts v2* (Fig. 1), but the full methodology is detailed in Appendix 1.

Data from source articles

In order to update the database, we first collated range shift estimates included in *BioShifts v1* and in *CoRE.* To be included in *BioShifts v2*, studies needed to clearly describe their methodology and report quantitative, species-level, latitudinal or elevational range shift rates (see supplement for additional information). Thus, some studies from the original repositories were excluded if they did not meet these new criteria for inclusion. We then checked for inconsistencies between the two existing repositories. When differences arose, we manually rechecked values from original articles to harmonize data extraction between the two repositories. In order to make values interoperable, we standardized range shift units from original studies to km/year in the latitudinal direction or m/year in the elevational direction (codified as "calc_rate"), and we classified estimated shifts into a "direction" value in relation to global hypotheses (expansion or contraction for range edges, latitudinal or elevational shift direction for range centres, no change). In order to aid future meta-analyses and make studies recognizable to original authors, we also retained original values reported by authors, as well as statistical significance from the original studies, when assessed (in “author_reported_*” columns). Some studies reported species’ range shifts as non-significant without providing an exact value for the shift estimate; in order to minimize positive publication bias, we retained unquantified non-significant shifts, assigning their standardized rates as zero (in “calc_rate”), and author reported rates as “null”. Because range shifts were collected from published studies that used a wide range of data types and methods to quantify shift rates, we collected metadata for 16 methodological variables for every estimated shift (Table S1).

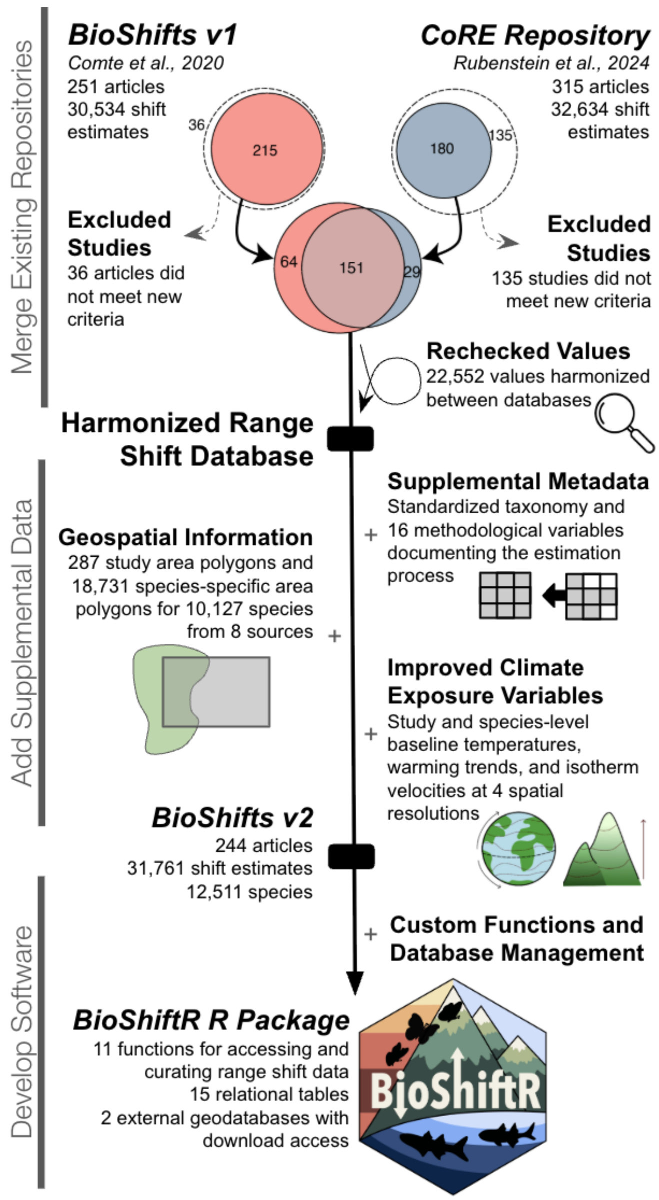

Figure 1. Workflow for the creation of the *BioShifts v2* database and the `BioShiftR` R package. *Bioshifts v2* was created by merging range shift articles from *BioShifts v1* and CoRE that met new criteria (215 and 180, respectively, with 64 and 29 articles unique to each individual repository and 151 present in both), harmonizing

values, and supplementing the range shift dataset with several collected and derived associated metrics (+ symbols). Then, database management and custom joining functions converted *BioShifts v2* into `BioShiftR`.

Next, we updated all species names as published in original articles to current accepted and standardized taxonomy. We retrieved accepted species names (genus and species) and higher taxonomic ranks (family, class, phylum, and kingdom) using backbone taxonomy from the Global Biodiversity Information Facility (GBIF). We retained the species name as originally published ("sp_name_publication"), as well as the checked and updated species name ("sp_name_checked") together with any additional distinctions (e.g., subspecies or variant) specified in the "subspecies" or "sublocality_or_population" columns, and provided standardized numerical identifiers (AphiaIDs) to facilitate connection with external taxonomic databases. Taxonomic harmonization reduced the total number of species with shift estimates from 12,857 unique species (as published in original articles) to 12,511 species (after harmonization).

These steps produced an initial dataset of 31,761 estimated species range shifts measured at three focal range parameters (leading edge, range centre, and trailing edge) for 12,511 species. Estimated range shift rates ranged from -63 to +146 km/year across latitudes, and from -119 to +105 m/year across elevations (where negative values indicate equatorward or downslope shifts, and positive values indicate poleward or upslope shifts; Fig. 2). Shift estimates spanned multiple taxonomic and functional groups, but were heavily dominated by vascular plants, birds, and invertebrates, which together constituted 91% of range shift estimates (Fig. 2).

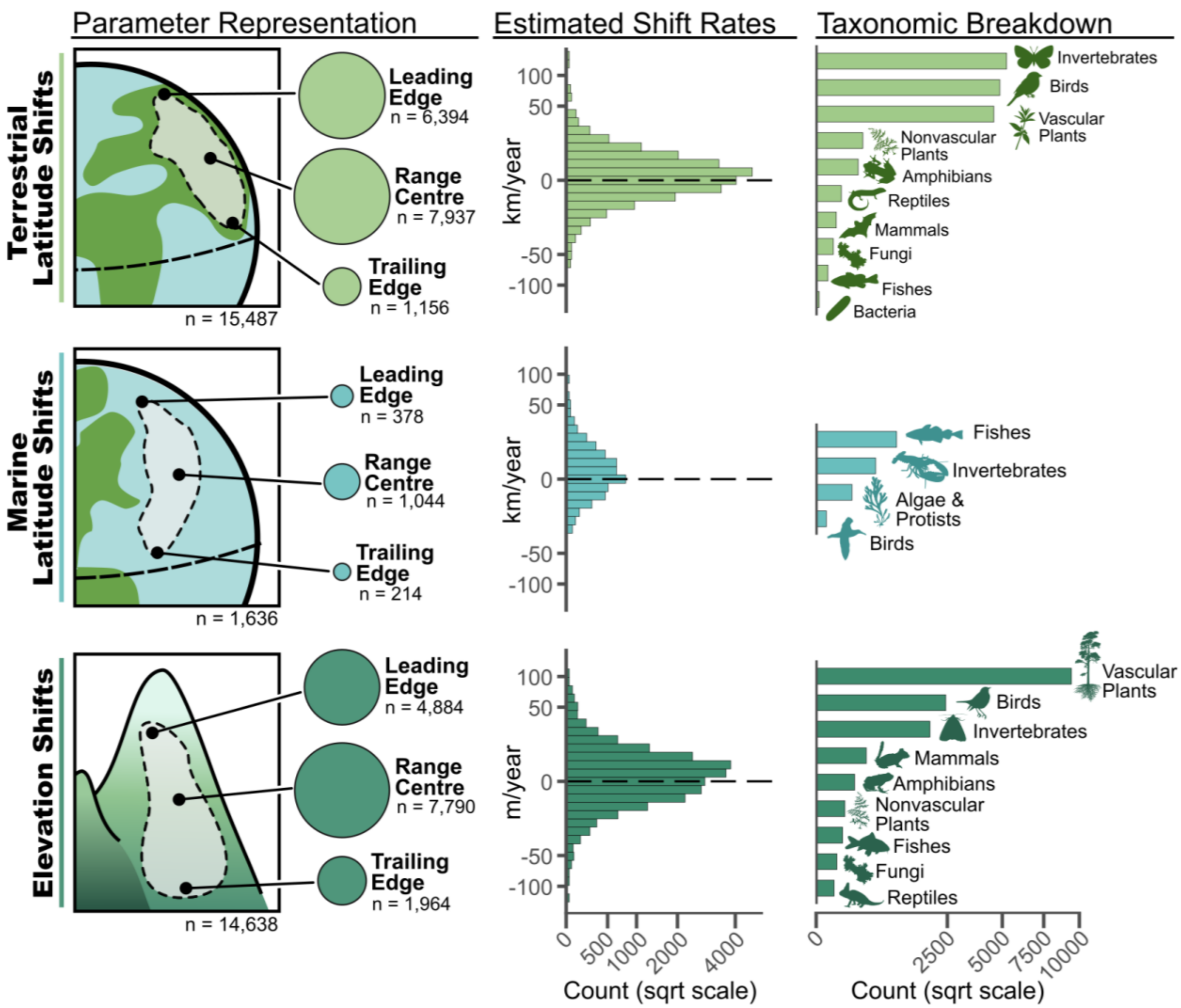

Figure 2. Overview of range shift estimates collated in *BioShifts v2*. Estimates of species' range shift rates across latitudinal gradients in the terrestrial (which includes freshwater; top row) and marine (middle row) realms, and across elevation gradients (bottom row). Left column shows the proportion of estimated shifts across shift types and range parameters; center column shows the distribution of raw estimated shift rates in each group (where positive values are poleward or upslope and negative values are equatorward or downslope shifts); right column shows the taxonomic and functional composition of range shift estimates within each group. Note that axes in columns 2 and 3 are square root-scaled to maximize data visibility.

Derived Data

For every source study, we delineated study polygons based on textual (spatial coordinates, regional descriptors) or graphical (e.g., maps, figures) descriptions from source articles matched to national geographic boundaries or delineated manually. When possible, we further delineated species-specific study polygons by collecting estimates of species' full distribution ranges for each (harmonized) species name from multiple sources, and intersecting species' distribution ranges with study polygons in which the species' range shift was assessed. This process resulted in 18,731 species-specific study polygons that were the same size or smaller than the corresponding full study polygons. We included spatial features representing all full- and species-specific study polygons, and additionally summarized polygons into tabular summary data including center coordinates, latitudinal and elevational extent, and total area.

Within all study- and species-specific polygons, we calculated three metrics of climate exposure covering the temporal duration of each focal study: baseline temperature (the average temperature over the study period), warming trend (the slope coefficient of the linear model linking temperature and time), and isotherm velocity standardized in the direction of the gradient over which the range shift was estimated (latitude or elevation; Fig. 3). Each exposure metric was calculated for every raster cell covering the area of a focal polygon, then cell values were averaged – weighted by area and coverage fraction within the given polygon – to achieve a mean value for the given polygon, as well as metrics of variability between cells (standard deviation and, in some cases, quantiles).

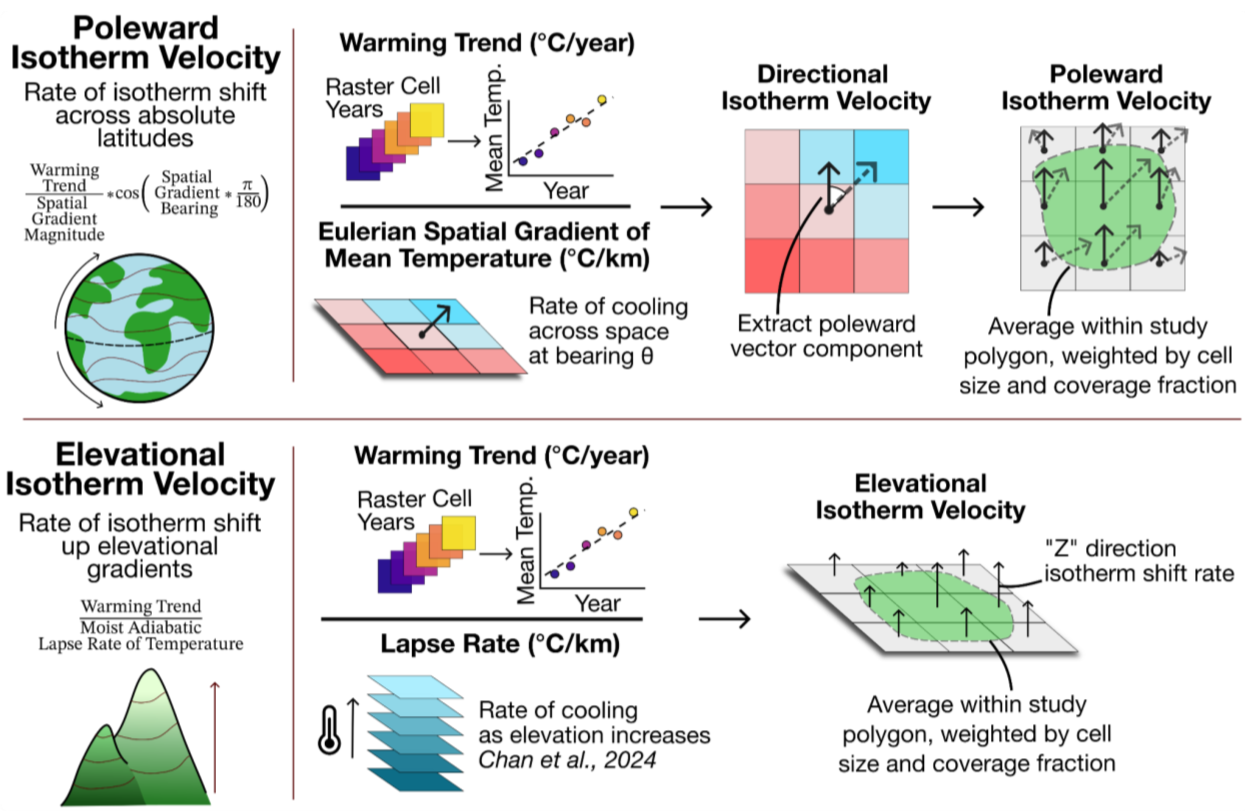

Figure 3: Methods for calculating the velocity of climate change across latitudes and elevations. Isotherm velocities were calculated for each raster cell across latitudinal (top row) and elevational (bottom row) gradients, then cell values were summarized within study polygons or species-specific study polygons for each estimated range shift (Appendix 1).

Because each of these metrics – particularly, isotherm velocity – is sensitive to the spatial resolution of the chosen input data (Fig. 4), we calculated each exposure variable using input temperature data at four spatial resolutions: 1km (terrestrial only), 25km, 50km, and 110km. Elevation shift exposure variables were calculated only with 1km input data because coarser-resolution elevation layers smooth local topographic variation, substantially reducing the accuracy and ecological relevance of derived metrics. We provide exposure variables estimated across both study and species-specific study areas. In the *BioShifts v2* database, 83% of species-specific study areas were smaller than the full study areas of source articles (on average, 74% of the size of the full study area); thus, study-level exposure calculations encompass values from areas that are irrelevant to focal species, and might misrepresent exposure compared to only the portions of study areas in which focal species are likely to live (Fig. 5). Species-level exposure calculations therefore offer a more precise estimate of hypothesized climate drivers for individual species' shifts, but could not be calculated for about 15% of range shift estimates when species' range polygons were unavailable.

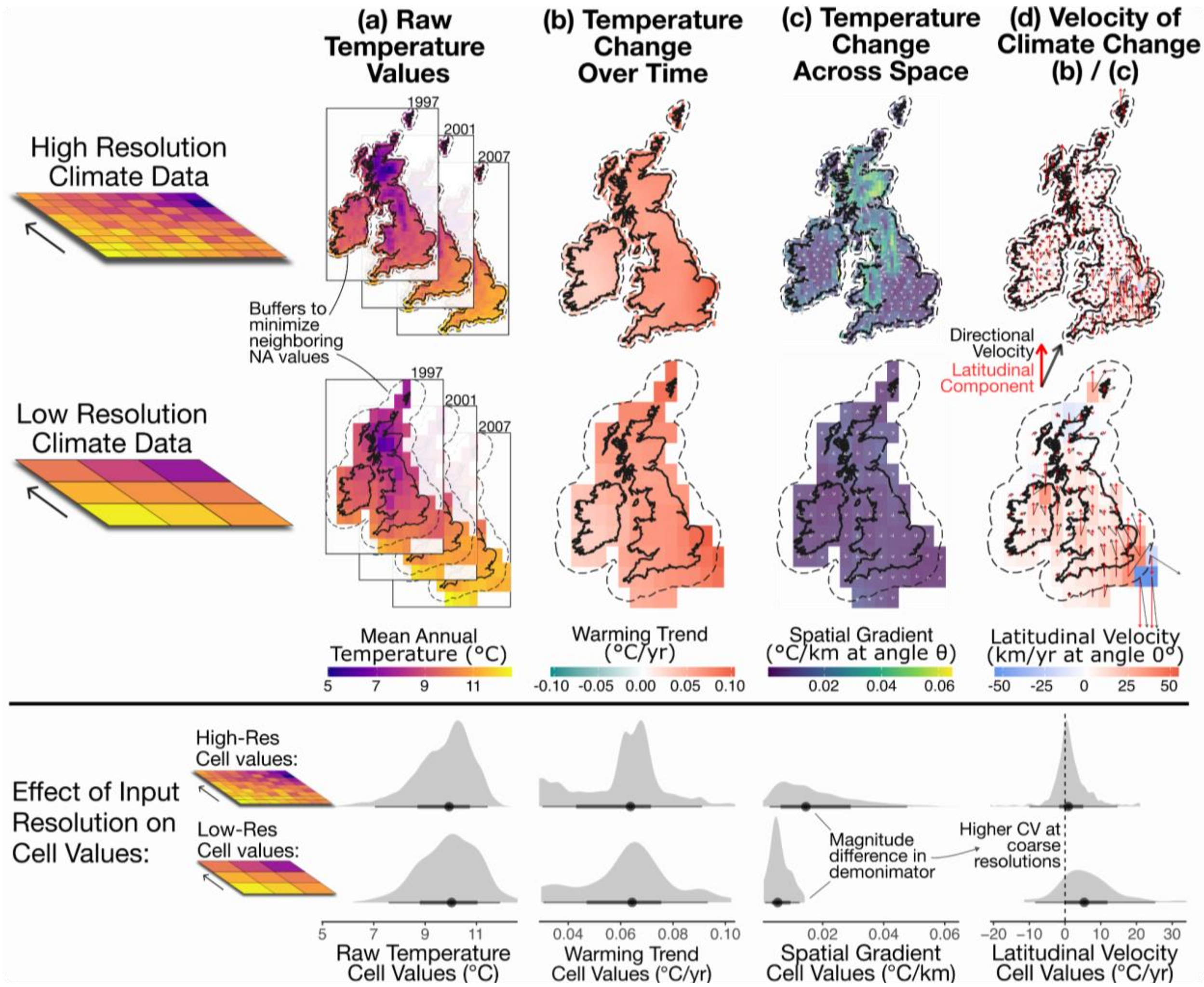

Figure 4. Effect of the spatial resolution of the input temperature raster layer on poleward isotherm velocities. The calculation of poleward isotherm velocity within the same study area (The British Isles) from 1997 to 2007 shown at two spatial resolutions: 25km cells (top row), and 110km cells (middle row). Comparison of cell output values from each step of the calculations (bottom row) shows how raw temperature values (column a) and warming trend values (column b) are largely consistent between coarse and fine spatial resolutions, but spatial gradient values differ greatly (bottom row, column c) with coarser spatial resolutions producing smaller values. These spatial gradient values become the denominator in the climate velocity computations (warming trend / spatial gradient), resulting in absolute velocity values closer to zero when calculated from finer spatial resolution data (bottom row, column d).

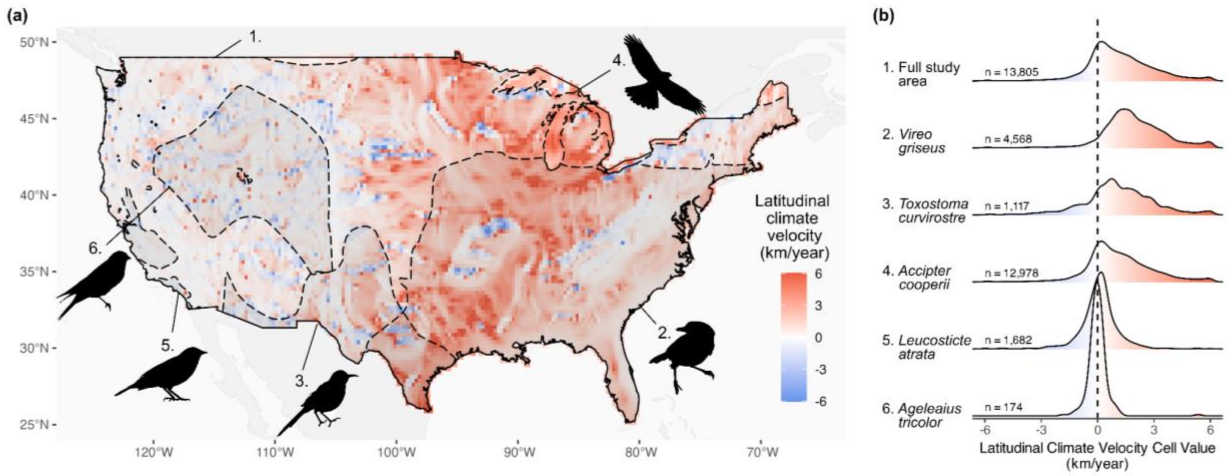

Figure 5. Comparison of poleward isotherm velocities across the full study area (the contiguous USA) and species-specific polygons for five example bird species. (a) Map of poleward isotherm velocities in km/year at a 25-km resolution across the USA (i.e., the study-level polygon outlined by a solid line) along with species-specific polygons (intersection of full study area and species range maps) for five example bird species whose range shifts were documented in *BioShifts v2* (polygons 2-6, dotted lines). (b) Distribution of climate velocity values across cells falling within the full study-level polygon (row 1) and species-specific polygons (2-6).

The `BioShiftR` package

We developed the `BioShiftR` package, implemented in the R software environment, to allow users to conveniently access and organize range shift data to facilitate comparative analysis and hypothesis-testing. The `BioShiftR` package contains 11 functions (Fig. 6, Table S2), which allow users to load a minimal dataset of range shift estimates from *BioShifts v2*, filter to target groups or locations, then add and merge information from associated relational tables including methodological variables, taxonomic information, spatial attributes, and climate exposure metrics for every study polygon or species-specific study polygon for each estimated shift (Appendix 2). In total, the database includes a maximum of 144 associated identifier-, metadata-, or derived data variables that link to each estimated range shift; the R package is therefore designed to streamline the management and navigation of the complex database. The `BioShiftR` package website (https://bioshifts.github.io/BioShiftR/) includes vignettes and documentation for every function.

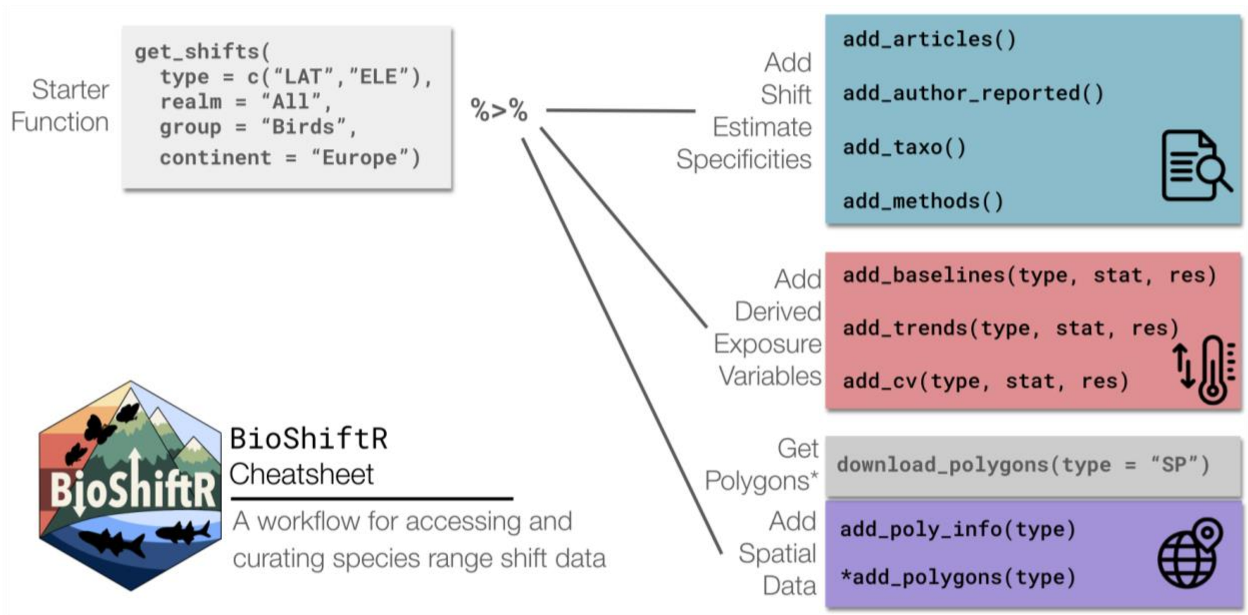

Figure 6. Function guide for `BioShiftR` R package. Workflows begin with the `get_shifts()` function, with helper arguments to filter shift estimates by type, realm, functional group, or location, then pipe to nine additional functions which add metadata, geospatial data, or derived exposure variables. `add_polygons()`, the function to merge in spatial data to the shifts dataset, requires an initial run of `download_polygons()` to store spatial files locally (denoted with an asterisk*).

`BioShiftR` functions are designed to use a "piped" syntax in R to sequentially add associated methodological and derived variables to the full range shifts dataframe, or to a potentially curated subset of it. Workflows begin with the `get_shifts()` function, which uploads a minimum dataframe of estimated shift rates within *BioShifts v2*. Then, additional functions iteratively add information to the minimal dataframe as needed by users (Table S2), linking the additional methodological or derived variables by standardized identifiers (species name, article ID, polygon ID, method ID, or combinations thereof). Because detection methods, taxonomy, climate change exposure, and geographical location might each affect range shift estimates either through influencing the detection process or affecting species' responses to climate change, we provide the ability for the users to account for these different factors in their analytical pipelines. Through piping, different supplemental datasets can be queried and linked in succession, providing information on the range shift assessment method, the species, the study area, or derived climate predictors, which can each be used to filter, visualize, assess biases, or explain detected range shift rates (Fig. 7).

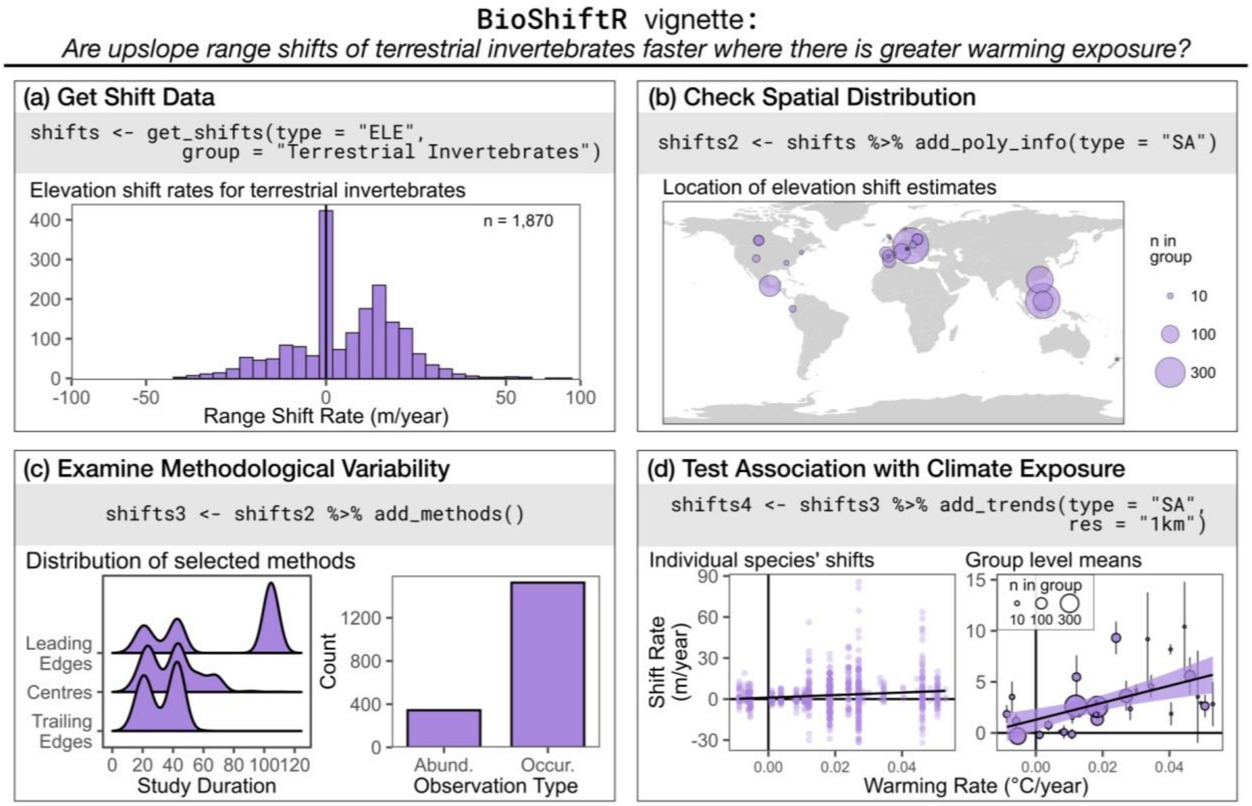

Figure 7. Example workflow to query range shift data for hypothesis testing with the BioShiftR R package. Here, we demonstrate how package functions can be used to (a) access raw range shift estimates for select shift types and species groups and (b-d) iteratively add contextual meta- and derived data in order to check biases, contextualize records, and test hypotheses. Package functions to access and add data are displayed in grey boxes, and plots in each panel are produced from the respective shift datasets (`shifts`-`shifts4`). Here, we only show distributions across space or methods (b, c), but filtering, grouping, or statistically accounting for such differences can be appropriate in some analyses. See the R Package website vignette for code for reproducing the displayed plots.

In addition to internal identifying keys to link across 15 relational tables, *BioShifts v2* includes standardized taxonomy identifiers and geospatial metadata of observations, and thus, range shift estimates can be easily linked to external data sources such as species trait databases and additional geospatial data layers that might influence observed range shift rates. These possible linkages are designed to facilitate expansion of possible hypotheses that *BioShifts v2* data could be used to test by allowing incorporation of data over and above what is provided in the database and package alone.

Conclusions

The *BioShifts v2* database provides the largest collection of documented quantitative species' range shift estimates to date, and its companion R package, `BioShiftR`, offers ease of access for navigating, organizing, and analysing compiled and derived data. Together, these tools will be useful for stimulating novel research on drivers of species redistributions. While our aim was to be comprehensive in collecting the maximum amount of data possible from range shifts and associated modifying variables, we acknowledge that additional variables that we did not collect – such as precipitation shifts – might explain some of the variation in recorded range shift rates.

As such, our tools provide information on species, geographies, and original sources for each documented range shift that will allow users to identify individual shift estimates and supplement the provided data with other hypothesized drivers.

As reports of species range shifts accumulate, dynamic tools will accelerate access to the most up-to-date information. Releasing *BioShifts v2* via `BioShiftR` leaves the door open for future updates to the database and package as new range shift reports are published. Integrating new datasets remains a time-consuming task, but the standards and workflows we have produced to date will provide direction for future data integration. While future updates would benefit most from repeated systematic literature searches of articles published after the timeframe of the searches that created the current version (June 2021), we have provided architecture for authors to submit their own data and articles that we have not yet integrated or that our searches may have missed (Appendix 3). In the future, we hope to update *BioShifts v2* and `BioShiftR` regularly by adding the most recent reports of range shift data, expanding to include non-English language reports, and exploring ways to accelerate and automate this process. Ultimately, we hope that these tools will facilitate further study of species' range shifts to build mechanistic understanding of biogeographical responses to climate change.

Brown, C. J., O'Connor, M. I., Poloczanska, E. S., Schoeman, D. S., Buckley, L. B., Burrows, M. T., Duarte, C. M., Halpern, B. S., Pandolfi, J. M., Parmesan, C., & Richardson, A. J. (2016). Ecological and methodological drivers of species' distribution and phenology responses to climate change. *Global Change Biology*, *22*(4), 1548–1560. https://doi.org/10.1111/gcb.13184

Chen, I.-C., Hill, J. K., Ohlemuller, R., Roy, D. B., & Thomas, C. D. (2011). Rapid Range Shifts of Species Associated with High Levels of Climate Warming. *Science*, *333*(6045), 1024–1026. https://doi.org/10.1126/science.1206432

Comte, L., Grenouillet, G., Bertrand, R., Murienne, J., Bourgeaud, L., Hattab, T., & Lenoir, J. (2020). *BioShifts: A global geodatabase of climate-induced species redistribution over land and sea* (p. 5238495 Bytes) [Dataset]. figshare. https://doi.org/10.6084/M9.FIGSHARE.7413365

Da Silva, C. R. B., & Diamond, S. E. (2024). Local climate change velocities and evolutionary history explain multidirectional range shifts in a North American butterfly assemblage. *Journal of Animal Ecology*, *93*(8), 1160–1171. https://doi.org/10.1111/1365-2656.14132

Fredston, A. (2024). *Measuring the edges of species' geographic ranges*. Ecology and Evolutionary Biology. https://doi.org/10.32942/X2QP69

Heikkinen, R. K., Leikola, N., Aalto, J., Aapala, K., Kuusela, S., Luoto, M., & Virkkala, R. (2020). Fine-grained climate velocities reveal vulnerability of protected areas to climate change. *Scientific Reports*, *10*(1), 1678. https://doi.org/10.1038/s41598-020-58638-8

Lawlor, J. A., Comte, L., Grenouillet, G., Lenoir, J., Baecher, J. A., Bandara, R. M. W. J., Bertrand, R., Chen, I.-C., Diamond, S. E., Lancaster, L. T., Moore, N., Murienne, J., Oliveira, B. F., Pecl, G. T., Pinsky, M. L., Rolland, J., Rubenstein, M., Scheffers, B. R., Thompson, L. M., … Sunday, J. (2024). Mechanisms, detection and impacts of species redistributions under climate change. *Nature Reviews Earth & Environment*. https://doi.org/10.1038/s43017-024-00527-z

Lenoir, J., Bertrand, R., Comte, L., Bourgeaud, L., Hattab, T., Murienne, J., & Grenouillet, G. (2020). Species better track climate warming in the oceans than on land. *Nature Ecology and Evolution*, *4*(8). https://doi.org/10.1038/s41559-020-1198-2

Lenoir, J., & Svenning, J.-C. (2015). Climate-related range shifts—A global multidimensional synthesis and new research directions. *Ecography*, *38*(1), 15–28. https://doi.org/10.1111/ecog.00967

Loarie, S. R., Duffy, P. B., Hamilton, H., Asner, G. P., Field, C. B., & Ackerly, D. D. (2009). The velocity of climate change. *Nature*, *462*(7276), 1052–1055. https://doi.org/10.1038/nature08649

Oliveira, B. F., Bertrand, R., Pinsky, M. L., Casajus, N., Wolfe, B. W., Scheffers, B. R., Villalobos, F., Grenouillet, G., Pecl, G. T., Chen, I.-C., Baecher, J. A., Lawlor, J. A., Sunday, J., Murienne, J., Rolland, J., Thompson, L. M., Lancaster, L. T., Rubenstein, M. A., Moore, N. A., … Comte, L. (2026). Species range shifts often speed ahead of their modeled climatic niches. *Proceedings of the National Academy of Sciences*, *123*(14), e2515903123. https://doi.org/10.1073/pnas.2515903123

Parmesan, C., Ryrholm, N., Stefanescu, C., Hill, J. K., Thomas, C. D., Descimon, H., Huntley, B., Kaila, L., Kullberg, J., Tammaru, T., Tennent, W. J., Thomas, J. A., & Warren, M. (1999). Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, *399*(6736), 579–583. https://doi.org/10.1038/21181

Poloczanska, E. S., Brown, C. J., Sydeman, W. J., Kiessling, W., Schoeman, D. S., Moore, P. J., Brander, K., Bruno, J. F., Buckley, L. B., Burrows, M. T., Duarte, C. M., Halpern, B. S., Holding, J., Kappel, C. V., O’Connor, M. I., Pandolfi, J. M., Parmesan, C., Schwing, F., Thompson, S. A., & Richardson, A. J. (2013). Global imprint of climate change on marine life. *Nature Climate Change*, *3*(10), 919–925. https://doi.org/10.1038/nclimate1958

Rubenstein, M. A., Weiskopf, S. R., Bertrand, R., Carter, S. L., Comte, L., Eaton, M. J., Johnson, C. G., Lenoir, J., Lynch, A. J., Miller, B. W., Morelli, T. L., Rodriguez, M. A., Terando, A., & Thompson, L. M. (2023). Climate change and the global redistribution of biodiversity: Substantial variation in empirical support for expected range shifts. *Environmental Evidence*, *12*(1), 7. https://doi.org/10.1186/s13750-023-00296-0

Rubenstein, M. A., Weiskopf, S. R., Bertrand, R., Carter, S. L., Comte, L., Eaton, M., Johnson, C. G., Lenoir, J., Lynch, A. J., Maslyukova, D. Y., Miller, B. W., Morelli, T. Lyn., Parker, E. J., Poole, C., Rodriguez, M. A., Rudick, T., Terando, A. J., Thompson, L. M., & Turner, K. (2024). *CoRE (Contractions or Range Expansions) Database: Global Database of Species Range Shifts from 1802-2019* [Dataset]. U.S. Geological Survey. https://doi.org/10.5066/P99VP2TW

Taheri, S., García-Callejas, D., & Araújo, M. B. (2021). Discriminating climate, land-cover and random effects on species range dynamics. *Global Change Biology*, *27*(6), 1309–1317. https://doi.org/10.1111/gcb.15483

Vitasse, Y., Ursenbacher, S., Klein, G., Bohnenstengel, T., Chittaro, Y., Delestrade, A., Monnerat, C., Rebetez, M., Rixen, C., Strebel, N., Schmidt, B. R., Wipf, S., Wohlgemuth, T., Yoccoz, N. G., & Lenoir, J. (2021). Phenological and elevational shifts of plants, animals and fungi under climate change in the European Alps. *Biological Reviews*, *96*(5), 1816–1835. https://doi.org/10.1111/brv.12727

Appendix 1. Methods for creating BioShifts v2.

Data sources for species range shift rates

BioShifts v2 was created by integrating and enhancing two existing range shift repositories – the *BioShifts v1* archive (Comte et al., 2020) and the Contractions and Range Expansion (*CoRE*) repository (Rubenstein et al., 2024) – to create a compilation of quantitative estimates of range shifts reported in the scientific literature (i.e., none of the raw range shift values were recalculated and qualitative estimates were excluded) for individual species along latitudinal and elevation gradients, together with their associated significance (if estimated in original sources). We focused on these two repositories because although they were initiated independently and contained unique information, they were both created through systematic literature reviews and compiled >30,000 range shift observations for thousands of species. The *BioShifts v1* initiative compiled quantitative estimates of latitudinal and elevational range shifts (in km/year or m/year) across all taxonomic groups (including terrestrial, freshwater and marine species), while the *CoRE* initiative included both quantitative and qualitative (e.g., direction) estimates of latitudinal, elevational, longitudinal, and depth range shifts of animals and plants, as well as abundance changes across species' ranges. Both also report methodological attributes related to the original study design that are structured in similar, albeit not identical, ways, thereby facilitating data integration. Despite having been initiated through systematic reviews, the set of studies (i.e., articles) included in the two repositories differed, owing to differences in search processes, data availability (e.g., data shared directly by authors) and data curation (e.g., *CoRE* includes more recent records from articles published after *BioShifts v1* publication).

Harmonization of species range shift estimates

Data harmonization between *BioShifts v1* and *CoRE* was completed by systematically comparing range shift estimates that were common to the two repositories and by standardizing estimates that were unique to either repository. For any range shift estimates that did not match between the two repositories, we manually reviewed the study dates, range shift rates, units of measurement, and statistical significance to correct any errors in the original repositories. In some cases, we excluded papers in either of the original repositories from *BioShifts v2* because: (1) the methods in the original study were not described clearly enough to confidently extract the range shift data; (2) the range shift values were extracted from a figure that was deemed not extractable; or (3) the range shift estimates were too coarse or approximative (e.g., range shift angle reported with a precision of 45 degrees). For the range shift estimates unique to either repository, we reviewed the articles to make sure that they met the same criteria as the articles shared between the two repositories, and completed the remaining missing fields using a common database schema. Any range shifts reported at a higher taxonomic level than the species or that reported imprecise start or end years of the period over which the range shift was estimated were excluded. For a given species, when range shifts were reported for more than one geographically distinct survey area or using different temporal resolutions or estimation methods, we considered them as separate range shift estimates. When range shifts were reported for different subspecies, varieties, or subpopulations of the same species within the same survey area, we recorded range shifts separately for each subgroup. Observations were labeled with the same species name and article identifier, but distinguished by methodological identifiers including "subspecies" and "sublocality or population".

Out of the 251 original articles in *BioShifts v1*, 36 were excluded, and out of the 315 articles in *CoRE*, 135 were excluded (most of them because they reported qualitative range shift estimates, abundance shifts, or depth shifts). Of the remaining articles, 151 were common to both *BioShifts v1* and *CoRE* (representing 22,552 range shift estimates), 64 were unique to *BioShifts v1* (representing 7,262 range shift estimates) and 29 were unique to *CoRE* (representing 1,947 range shift estimates). Overall, *BioShifts v2* is composed of a total of 244 articles and 31,761 individual species-level range shift estimates across latitudes or elevations estimated at species' range edges or range centres (Fig. 1, Fig. 2).

Standardizing range shift estimates to a common unit

Original studies reported range shift rates in many different ways, including total shift distances over a period of time (in m, km, or degrees latitude) or directly as rates of change (in m/year, km/year, degrees/year, sometimes measured at a compass angle). In order to standardize values across studies, we converted all range shift estimates to a rate using common units, in km/year poleward for latitudinal studies (with positive values indicating poleward shifts) and in m/year upward for elevational studies (with positive values indicating upslope shifts). For shifts that were originally reported as rates of change in the latitudinal or elevational direction, we kept original shift values, converting units where necessary. For shifts that were reported as total shift distances, we calculated the shift rate by dividing the shift distance in the latitudinal or elevational direction, in standardized units of km or m, by the study duration, in years. To calculate shift distance in the latitudinal direction when shift distances were reported as the distance shifted at a given compass angle across latitude, we multiplied the distance moved at the given compass angle, converted to km, by the cosine of the angle in radians. We calculated the study duration as the difference between the "end year" of the study and the "start year" of the study plus one to get the total number of years elapsed over the entire study period. If the range shift was documented using grouped survey census periods rather than single-sample surveys across years, we calculated the study duration as the difference between the midpoint year of the last study period and the midpoint year of the first study period plus one. For the sake of reproducibility and transparency with the original range shift reports, both standardized range shift rates and the authors' reported values are available in the final database (with "calc_" and "author_reported_" prefixes, respectively), together with information to identify the source articles for each shift estimate ("reference") and the exact location within source articles from which author-reported values originated ("author_source").

Codifying significance status

When available, we retained information on the authors' assessment of significance within the original study that might be useful for future meta-analyses. We systematically codified study-assessment of range shift significance as *yes* if the magnitude of range shift reported was assessed as significant within the original study, as *no* if it was a non-significant shift, or as *uncertain* if significance was not assessed. In some cases, shifts were reported as non-significant, but precise estimates of range shift rates were not reported in original articles; for these records, we recorded the standardized range shift rate as 0 but the author-reported range shift rate with the string, '*NULL'* instead of a magnitude. We retained all standardized range shift rates in the database whether or not the authors assessed the reported range shifts as 'significantly' different from zero to avoid positive publication bias.

Direction of range shift rates

To facilitate interpretation, we classified the overall "direction" of each standardized range shift rate, relative to the expected direction under global climate change for shifts at the range center (i.e., upwards in elevation or polewards in latitude) or expected range dynamics for shifts at the range limits (i.e., range contraction or expansion). For shifts measured at the range centre, we classified negative shifts as towards *lower elevation* or *towards equator* and positive shifts as those towards *higher elevation* or *towards poles*, depending on the gradient considered. For shifts measured at latitudinal or elevational range limits, we classified both negative shifts at the leading edge and positive shifts at the trailing edge as cases of range *contraction*. Likewise, we classified both positive shifts at the leading edge and negative shifts at the trailing edge as cases of range *expansion*. Non-significant shifts or range shift rates reported as 0 were classified as *no change*.

Taxonomic harmonization

We harmonized taxonomy in *BioShifts v2* to standardize the nomenclature reported by original studies. For each species listed in original publications, we retrieved accepted species names (genus and species) and higher taxonomic ranks (family, class, phylum, and kingdom) using the Global Biodiversity Information Facility (GBIF) backbone taxonomy, accessed via the R packages *rgbif* (Chamberlain et al., 2025), *bdc* (Ribeiro et al., 2022), and *taxadb* (Norman et al., 2020). When an accepted name could not be resolved through GBIF, we queried alternative taxonomic authorities, the Integrated Taxonomic Information System (ITIS) and the National Center for Biotechnology Information taxonomy (NCBI) to identify valid synonyms and taxonomic classifications. In cases where no accepted name could be confidently retrieved from any of these sources, we retained the species name and taxonomy as originally reported in the study. This hierarchical procedure allowed us to maximize taxonomic standardization while preserving traceability to the original records and minimizing data loss. Species names as published in original studies ("sp_name_publication"), including any reported subspecies or variety ("subsp"), and harmonized species names ("sp_name_checked") are available in the final database, as well as standardized numerical identifiers (AphiaIDs) to connect *BioShifts v2* to external taxonomic databases.

Methodological variables documenting the estimation process

Species range shifts are estimated using a wide variety of data types (e.g., occurrences from gridded atlases covering distinct time periods, networks of monitoring sites, multi-year abundance surveys), sampling procedures (e.g., true resurveys, bootstrapping to equilibrate surveys across time periods), and statistical definitions of the focal parameter (range edge, range center). All these decisions during the methodological workflow of a given study can result in a high variability across estimates (Brown et al., 2016; Fredston-Hermann et al., 2020; Lai et al., 2024; Lenoir et al., 2020). To account for methodological differences between studies, we included several methodological metadata variables documenting the estimation process for each range shift estimate in the database (Table S1). Methodological variables are usually similar across all range shift estimates of a given range parameter along a given spatial gradient (i.e., latitude or elevation) within a given study, but sometimes can vary between species in one study, or one study can use multiple methods to estimate shifts for each species. Most methodological variables are expected to affect – if anything – the precision of range shift estimates around a hypothetical mean; others do not have a clear influence on range shift detections, but could introduce some uncertainty to account for in statistical models comparing standardized range shift rates collated in the database.

Geospatial information

BioShifts v2 includes two geopackages: one containing spatial polygons delineating the full spatial extent of each study area and another containing polygons delineating the spatial extent that each species occupies within the full study area (i.e., species-specific study areas).

Within each article, we defined a polygon as a geographically distinct survey area across which the range shift(s) were estimated. We provide textual information describing the location of the polygon (e.g., California or French Alps) as reported in the original articles, as well as spatial features delineating these areas. Spatial polygons were delineated based on national geographic boundaries using the maptools package in the R environment (Bivand, 2023)or delineated manually based on any meaningful spatial information from the text (e.g., maps, figures, spatial coordinates) in ArcGIS Pro and Google Earth (Esri, 2024; Gorelick et al., 2017).

While a set of study area polygons were readily available from *BioShifts v1* and were updated to include study areas for new range shift estimates from CoRE, the polygons that delineate the spatial extent that each species occupies within the full study area (species-specific study polygons) were not available and thus represent a novel addition to *BioShifts v*2. The idea to add species-specific study area polygons was born from the realization that for studies covering large geographic areas, the full study area polygon often extends beyond the occupied range of species within the study (Fig. 4), thereby misrepresenting species-specific climate exposure and other landscape-level drivers of species range shifts.

To create species-specific study polygons, we sourced range map polygons from a variety of published databases (including IUCN [IUCN (International Union for Conservation of Nature) 2024], BirdLife [Bird species distribution maps of the world 2019], GARD [Roll *et al.* 2017], BIEN [Maitner 2025], and ButterflyMaps [Daru 2025]) and supplemented them with occurrence-based polygons. Since we wanted to retain a liberal estimate of the area of the full study in which the species could have been observed, we combined polygons representing each species' extant, extinct, breeding, and non-breeding range together into a single polygon that would be used to clip the full study area polygon. For species for which no expert range map was available, we used coarse-resolution regional occurrence data (e.g., species checklists from the GIFT database [Denelle *et al.* 2023]), outputs of species distribution models (e.g., Aquamaps, with a liberal threshold of $p > 0.05$), or custom convex hulls around filtered GBIF occurrences (using default parameters of the `make_range()` function from the *gbif.range* package; mean number of occurrences = 28,158 ± 83,051 SD, range = 30 - 1,575,525) to estimate species' geographic ranges. To retain only one range map polygon per species, we selected range map polygons in the following order of priority when polygons were available from multiple sources: IUCN (n = 1,147), BirdLife (n = 819), GARD (n = 6), BIEN (n = 3,017), Butterfly Maps (n = 202), GIFT (n = 1,107), Aquamaps (n = 111), GBIF convex hull (n = 3,948) (DOIs listed in Table S3).

We then individually clipped the study area polygon by the range map polygon for each species observed within that study, resulting in a total of 18,731 species-specific polygons which were either the same size (in which case no clipping was applied) or smaller than study-level polygons. Of the 18,731 species-specific study area polygons that could be delineated, roughly 17% were identical in shape to the original study area polygons, indicating study areas were fully encompassed inside of the range of a species. However, the 83% of species-specific polygons that do differ from the full study area polygon were, on average, 74% of the size of the full study area, indicating the importance of this step in improving the precision of climate change exposure estimates tailored to each species-by-study-area combination. Since 2,572 species had no range map polygon available from any source, there were 4,692 combinations of species and study areas for which a species-specific polygon is not available.

We further collected spatial metadata for each species-specific and study-level polygon, describing the size, placement, and breadth of polygons along the latitudinal and elevational gradients. Areas and spatial extents were estimated after projecting the polygons to the equal-area Mollweide projection and gridded elevation data were sourced from WorldClim 2.1 at a 1-km resolution (Fick & Hijmans, 2017).

Climate velocities and climate change exposure

We derived three summarized metrics of climate exposure within study-area and (when available) species-specific polygons: (1) the baseline (mean) temperature over the study duration, (2) the warming trend over time, and (3) the velocity of isotherm shifts along latitudinal or elevational gradients. We calculated each of these variables from input raster data of mean annual temperature (air temperature or sea surface temperature) at multiple spatial resolutions (1km for terrestrial study areas only, and 25km, 50km, and 110km for both terrestrial and marine; 1km only for elevation shifts), recognizing that climate change velocity values and mean temperature values are sensitive to spatial resolution (Dobrowski *et al.* 2013; Fig. 3). Mean air temperature values used for climate exposure metrics in terrestrial study areas originated from the CHELSAcruts v1 1km x 1km monthly maximum and minimum global temperature dataset (Karger & Zimmermann, 2018), which we averaged to produce annual mean temperatures for 1901-2016 (Bernhardt et al., 2018), then coarsened to 25, 50, and 110km resolutions by resampling 1km x 1km rasters to averaged values within coarser grids using the `terra` package in R. Mean sea surface temperatures were derived from the ORAS5 global ocean reanalysis monthly dataset (https://data.marine.copernicus.eu), summarized within years for 1958-2019 at 25km x 25km, then coarsened to 50 and 110km resolutions. To represent climatic conditions that may have contributed to observed species range shifts, annual temperature values were lagged by one year. For example, the mean annual temperature value assigned to 1960 corresponds to the average of monthly temperatures observed between January and December

1959. We calculated each exposure metric within each raster cell of a given study area (or species-specific study area) for the time period covering the study's duration, then calculated summary values across cells for each metric (mean and standard deviation for baseline temperature and warming trend; mean, standard deviation, median, 25th and 75th quantile for climate velocity). When studies covered temporal durations that extended beyond the limits of temperature data (1902-2016 for terrestrial, 1959-2019 for marine), we calculated exposure metrics from the maximum amount of climate data possible.

We calculated baseline temperature as the average mean annual temperature in the study area or species-specific study area over the years of the study's duration, and calculated the average value and standard deviation across all raster cells. We calculated the warming trend as the coefficient of the linear model linking mean annual temperature values to years in °C/year for all raster cells, and again summarized cell values to the mean and standard deviation within the focal polygon. We also derived new estimates of the velocity of climate change along latitudinal and elevation gradients using the warming trend and estimates of the spatial gradient of mean temperature – or changes in mean temperature across space. For each gradient (latitudinal and elevational), we standardized isotherm velocities to the component only towards that direction in order to allow direct comparisons to range shift estimates.

Latitudinal velocities were estimated using an adaptation of the gradient-based method from Burrows *et al.* (2011; Fig. 2), building on the `VoCC` package (García Molinos et al., 2019) and the `gVelocity` function from the `climetrics` R package (Taheri et al., 2024). In order to estimate climate velocities along the latitudinal gradient, we used the warming trend of cells within the study area (described above), and the spatial gradient of temperature change across space. Spatial gradient was calculated as the vector sum of the weighted averages of north-to-south and east-to-west slopes of mean temperature change within 9-cell neighborhoods, from which we identified the rate (in °C/km) and direction (compass bearing θ) of temperature cooling across space. In order to reduce the influence of NA values in 9-cell neighborhoods encompassing focal cells (cells within study areas), we maintained all available raster cells within a one-cell-width buffer from the focal study area during this calculation, then retained spatial gradient values for only focal cells. We divided the warming trend of each cell by its spatial gradient to obtain the absolute rate and direction of isotherm shifts (in km/year, towards compass bearing θ under warming and in the opposite direction under cooling). Because division by anomalously flat spatial gradients in the denominator can produce anomalously high velocity values, we truncated cell velocity values to the 5th and 95th percentile cell values within study areas when they contained more than 100 raster cells (similar to `climetrics`). We extracted the poleward component of the directional isotherm velocity vectors by multiplying the vector magnitudes by $\cos(\theta * \pi/180)$ (standardizing vectors towards 0°N) then multiplying by -1 for southern hemisphere cells to represent the rate of isotherm shifts towards the poles. The resulting metric is the latitudinal component of poleward isotherm velocity, with positive values representing poleward isotherm shifts and negative values indicating equatorward shifts. Finally, we calculated the mean, median, 25th and 75th percentile, and standard deviation of poleward isotherm velocities, weighted by cell size and coverage fraction within a polygon for every focal study area polygons, and, when available, species-specific polygons (Fig. 4).

We calculated poleward isotherm velocities from input rasters at four spatial resolutions: 1km (terrestrial only), 25km, 50km, and 110km raster cells. Cell size affects both the variability of temperatures between raster cells (wherein larger cell sizes have fewer extreme values; Fig. 3), and the spatial gradient of temperature between cells (wherein larger cell sizes result in flatter gradients owing to smaller differences between cells and larger distances between them, Fig. 3). Thus, climate velocities calculated from data at coarser spatial resolutions generally have smaller spatial gradient denominators compared to those calculated from data at finer spatial resolutions, resulting in higher climate velocity values (Fig. 3). Given the wide range of body size, mobility, and home range size of species in the dataset, it is also possible that species interact with climate variability and velocity at multiple scales, and thus, we provide all four options for climate velocity and related exposure variables.

We calculated elevational climate velocities as the warming trend of mean annual temperature (°C/year, as above) within a study area divided by the moist adiabatic lapse rate of temperature along elevational gradients (Chan *et al.* 2024; °C/m), resulting in the velocity of isotherm shift up the "Z" axis of elevation and expressed in m/year. Again, values were calculated per raster cell, then summarized to the mean, median, 25th and 75th percentile, and standard deviation within each study area, weighted by cell area and coverage fraction of the cell within the polygon or species-specific polygon. Elevation velocities were calculated only at 1km spatial resolution only (Fig. 2).

Appendix 2. `BioShiftR` Column Metadata

BioShiftR Column Metadata

Reference guide to all columns in the BioShifts v2 database that can be downloaded with the BioShiftR helper functions. Column descriptions are organized by the BioShiftR functions that are used to access them.

1. Core range shifts dataset — get_shifts()

All `BioShiftR` workflows begin with get_shifts(), which returns a minimal dataset containing the calculated range shift rates across latitude or elevation, in km/year or m/year respectively, and the identifier columns needed to connect to other tables of the *BioShifts v2* database.

1.1 Identifier Columns

id

Field	id
Description	Summary identifier value arranged as article_poly_type_param_method. Because the minimal shifts dataframe connects to several other dataframes, each containing data at different levels (e.g., some additional information is at the "source article" level, while some is at the species or shift level), the minimal shifts dataframe contains five separate ID columns (article_id, poly_id, type, param, and method_id), which collectively make up the id column. These identifiers are nested such that each consecutive identifier can—but does not always—contain multiple of the next value. Different combinations of these identifiers merge the minimal shifts dataframe to all other *BioShifts* sub-datasets, and they are designed such that within a single identifier, a species/subspecies will never be represented more than once.
Format	Character
Pattern / Examples	`AXXX_PX_type_param_MXX / A001_P1_LAT_O_M01`

article_id

Field	article_id
Description	Identifier for the source publication from which the range shift estimates were extracted. Multiple shifts can share the same article_id when a single publication documents shifts for multiple species, polygons, range parameters, or methods.
Format	Character
Pattern / Examples	`AXXX / A001, A002, ... A244`

poly_id

Field	poly_id

Description	Identifier for the polygon (study area) within the source publication in which range shifts were documented. Most publications have one polygon (P1), but can have multiple when a publication reported range shifts within multiple study areas (e.g., two separate mountains, three ocean basins).
Format	Character
Pattern / Examples	`PX / P1, P2, ... P9`

type

Field	**type**
Description	Whether the range shift estimate was documented along the latitudinal or elevational gradient.
Format	Character
Values / Notes	LAT (latitudinal), ELE (elevational)

param

Field	**param**
Description	Whether the range shift estimate was documented at the leading edge (LE), defined as the poleward (high-latitude) or upslope (high-elevation) range margin within the study area; range center (O), defined as the centroid or mid-point of the species distribution within the study area; or trailing edge (TE), defined as the equatorward (low-latitude) or downslope (low-elevation) range margin within the study area. We note that we used the definition provided in the original articles; we did not attempt to verify whether the range margins aligned with the cold and warm species range limits, nor did we verify whether the range limits within the study areas aligned with the 'true' range limits of species distributions
Format	Character
Values / Notes	LE (leading edge), TE (trailing edge), O (range center)

method_id

Field	**method_id**
Description	Identifier for the detection method used within the preceding group levels. Usually, studies use only one method for all range shift estimates that they reported (M01), but in some cases, studies reported range shift estimates across multiple timeframes (e.g., 1950–1975, 1975–2000, 1950–2000) or using two different statistical methods, resulting in detections with identical values of all preceding columns. These cases are separated into M01, M02, and so on.
Format	Character
Pattern / Examples	`MXX / M01, M02, ... M24`

1.2 Core Shift Metric Columns

calc_rate

Field	calc_rate
Description	The standardized range shift rate in km/year for latitudinal shifts and m/year for elevational shifts. This is the primary shift metric in the database, recalculated from the original author-reported values to ensure consistency across studies.
Format	Numeric
Values / Notes	Positive values indicate poleward or upslope shifts; negative values indicate equatorward or downslope shifts

calc_unit

Field	calc_unit
Description	The unit of measurement for the calculated range shift rate (calc_rate). Latitudinal shifts are reported in km/year and elevational shifts are reported in m/year.
Format	Character
Values / Notes	km/year (for LAT shifts), m/year (for ELE shifts)

direction

Field	direction
Description	The overall “direction” of each standardized range shift rate, relative to the expected direction under global climate change for shifts at the range center (i.e., upwards in elevation or polewards in latitude) or expected range dynamics for shifts at the range limits (i.e., range contraction or expansion). For shifts measured at the range center, we classified negative shifts as towards *lower elevation* or *towards equator* and positive shifts as those towards *higher elevation* or *towards poles*, depending on the gradient considered. For shifts measured at latitudinal or elevational range limits, we classified both negative shifts at the leading edge and positive shifts at the trailing edge as cases of range *contraction*. Likewise, we classified both positive shifts at the leading edge and negative shifts at the trailing edge as cases of range *expansion*. Non-significant shifts or range shift rates reported as 0 were classified as *no change*.
Format	Character
Values	Lower Elevation (shift towards lower elevation at the range center), Higher Elevation (shift towards lower elevation at the range center), Towards Poles (shift towards higher latitudes at the range center), Towards Equator (shift towards lower latitudes at the range center), Expansion (positive shift at the leading edge or negative shift at the trailing edge), Contraction (negative shift at the leading edge or positive shift at the trailing edge), No Change

781 ### 1.3 Core Shift Taxonomic Columns

782 #### sp_name_publication

Field	sp_name_publication
Description	Species name as reported in the original source publication. May differ from the harmonized name in sp_name_checked due to spelling variations or taxonomic updates.
Format	Character

783 #### sp_name_checked

Field	sp_name_checked
Description	Species name after validation and harmonization against a standard taxonomic backbone. Preferred for analyses requiring consistent species-level identification.
Format	Character

784 #### subsp

Field	subsp
Description	Subspecies or variety as described in the original publication. NA if no subspecies or variety was included.
Format	Character

785 #### eco

Field	eco
Description	Whether the study was performed in the terrestrial or marine realm. We note that the definition of the two realms was based on whether the study occurred within the land masses or the oceans and as such, no distinction was made for freshwater species (considered as terrestrial here).
Format	Character
Values / Notes	Ter (terrestrial or freshwater), Mar (marine)

786

787 ## 2. Source Article Information — add_articles()

788 Adds columns identifying the source publication for each shift.

789 #### article

Field	article
Description	Internal label identifying the publication or other source in which the range shift was originally documented. Labels usually combine author information, year and an abbreviated journal or source type, but are not full bibliographic citations. Use `reference` for the formatted citation where available.

Format	Character

reference

Field	reference
Description	Formatted reference for the source in which the range shift was originally documented, following the pattern: author(s), title, abbreviated source or journal name, volume, pages or article number, and year.
Format	Character

doi_url

Field	doi_url
Description	Digital Object Identifier (DOI) for the source publication, or the url to the web source when a DOI was not available.
Format	Character

id_bioshifts_v1

Field	id_bioshifts_v1
Description	Cross-reference identifier linking the current *BioShifts v2* article to the equivalent article in the first version of the *BioShifts* database (v1). Useful for tracking records across database versions. Can contain NAs if the article did not exist in v1.
Format	Numeric
Values	1-257

id_core

Field	id_core
Description	Cross-reference identifier linking the current *BioShifts v2* article to the equivalent article in the CoRE database. Can contain NAs if the paper was not included in CoRE.
Format	Numeric
Values	1-781

3. Author-Reported Shift Information — add_author_reported()

BioShifts v2 recalculates range shifts as standardized rates across latitude or elevation, which are returned by `get_shifts()` as \`calc_rate\` and \`calc_unit\`. This function provides the original shift values as reported from the source publication, with their original units, significance information, magnitude and angle where relevant, and the location (within the reference publication) from which the reported value was extracted

author_reported

Field	author_reported
Description	The raw range shift value as reported by the original study authors, prior to any recalculation or standardization. Complements `calc_rate` by preserving the original reported shift for reference.
Format	Numeric

802 ## author_reported_unit

Field	author_reported_unit
Description	The unit of measurement for the author-reported shift value. May differ from calc_unit if the original study used non-standard units (e.g., degrees latitude, m/decade).
Format	Character
Values	deg, km, km/year, km/decade, deg/year, m, m/year, m/decade

803 ## author_reported_sig

Field	author_reported_sig
Description	Statistical significance of the author-reported range shift, as stated in the original publication.
Format	Character
Values	Y (yes, shift was found to be significant), N (no, shift was not found to be significant), U (unknown, significance of the shift was not reported)

804 ## author_reported_magnitude

Field	author_reported_magnitude
Description	The magnitude of the author-reported shift if the article provided a shift value with a compass direction (i.e., if the latitudinal component of the range shift was not directly provided). Values are NA if the original shift estimates were provided in a different format.
Format	Numeric

805 ## author_reported_angle

Field	author_reported_angle
Description	The angle of the author-reported shift if the article provided a shift value with a compass direction (i.e., if the latitudinal component of the range shift was not directly provided). Values are NA if the original shift estimates were provided in a different format.
Format	Numeric

806 ## author_source

Field	author_source

Description	Indicates the location or table within the original publication from which the author-reported shift value was extracted (e.g., main text, supplementary table).
Format	Character

4. Methodological Parameters — add_methods()

Adds methodological parameters describing how each individual shift was detected, including temporal coverage and detection methods.

sublocality_or_pop

Field	sublocality_or_pop
Description	When relevant, the distinguishing group, subpopulation, or other identifying difference for multiple shifts recorded for the same species within the same study polygon and focal range parameter and using the same methods (e.g., saplings vs. adults, queen vs. worker bees, fall vs. spring trawl sampling).
Format	Character
Values	Free text descriptions of focal groups within a species and study polygon

start_firstperiod

Field	start_firstperiod
Description	The start year of the first study time period over which the range shift was detected. If detections are made via distinct annual surveys-resurveys (e.g., 1965 versus 2007), start, midpoint, and end years will be the same for each period. For “census” surveys (e.g., occurrence records from 1965-1969 versus 2007-2009), start, midpoint, and end will describe the census period.
Format	Integer

midpoint_firstperiod

Field	midpoint_firstperiod
Description	The midpoint year of the first study time period. Calculated as the mean of start_firstperiod and end_firstperiod. If detections are made via distinct annual surveys-resurveys (e.g., 1965 versus 2007), start, midpoint, and end years will be the same for each period. For “census” surveys (e.g., occurrence records from 1965-1969 versus 2007-2009), start, midpoint, and end will describe the census period.
Format	Integer

end_firstperiod

Field	end_firstperiod
Description	The end year of the first study time period over which the range shift was detected. If detections are made via distinct annual surveys-resurveys (e.g.,

	1965 versus 2007), start, midpoint, and end years will be the same for each period. For “census” surveys (e.g., occurrence records from 1965-1969 versus 2007-2009), start, midpoint, and end will describe the census period.
Format	Integer

815 ### start_lastperiod

Field	start_lastperiod
Description	The start year of the last study time period over which the range shift was detected. If detections are made via distinct single-year surveys-resurveys (e.g., 1965 versus 2007), start, midpoint, and end years will be the same for each period. For “census” surveys (e.g., occurrence records from 1965-1969 versus 2007-2009), start, midpoint, and end will describe the census period.
Format	Integer

816 ### midpoint_lastperiod

Field	midpoint_lastperiod
Description	The midpoint year of the last study time period. Calculated as the mean of start_lastperiod and end_lastperiod. If detections are made via distinct annual surveys-resurveys (e.g., 1965 versus 2007), start, midpoint, and end years will be the same for each period. For “census” surveys (e.g., occurrence records from 1965-1969 versus 2007-2009), start, midpoint, and end will describe the census period.
Format	Numeric

817 ### end_lastperiod

Field	end_lastperiod
Description	The end year of the last study time period over which the range shift was detected. If detections are made via distinct annual surveys-resurveys (e.g., 1965 versus 2007), start, midpoint, and end years will be the same for each period. For “census” surveys (e.g., occurrence records from 1965-1969 versus 2007-2009), start, midpoint, and end will describe the census period.
Format	Integer

818 ### n_periods

Field	n_periods
Description	Total number of distinct temporal units that were used to estimate the range shifts. For example, for range shifts estimated between two census periods (e.g., 1965-1969 and 2007-2009), the number of temporal units is two; for range shifts estimated using annual time series, the number of temporal units would be the number of years in the study period.
Format	Integer

819 ### duration

Field	duration
Description	Total study duration in years, estimated as the difference between the midpoint year of the second study period and the midpoint year of the first study period + 1.
Format	Numeric

grain_size

Field	grain_size
Description	The spatial resolution of the species data used to estimate the shifts: less than 1 km^2, such as a well-delineated site or transect (small): 1–10 km^2, such as a gridded atlas (moderate); 10–100km^2, such as sub-regions or ecoregions (large): or >100km^2, such as states or provinces (very large).
Format	Character
Values	Small (<1km^2); moderate (1-10 km^2); large (10–100 km^2); very_large (>100 km^2)

sampling

Field	sampling
Description	The temporal distribution of sampling used to estimate range shifts.
Format	Character
Values	TWO - two time periods used to estimate range shifts; IRR/AGG - three time periods composed of one or more years or aggregated such as in decadal intervals; MULTIPLE (continuous) - multiple time units that were treated as continuous in the estimation process even if they may be irregular; CONTINUOUS - multiple and regular sampling periods

category

Field	category
Description	Broad methodological category describing how range shifts were detected (e.g., resurvey, time series analysis).
Format	Character
Values	Whether the range shifts were estimated between two distinct years (Survey-Resurvey), between two periods composed of more than one year of species data (CensusPeriods), using (usually annual) time series data (TimeSeries), or using species distribution limit estimates gathered from previous literature, generally spanning multiple years (Synthesis). Cases in which the first and last time period had different compositions are indicated as Survey-Census or Census-Resurvey

obs_type

Field	obs_type

Description	The type of species data used to estimate the range shifts, i.e., occurrences versus abundances. Even if the underlying dataset includes abundances, if the data were transformed to presence-absence to estimate the range shift (e.g., using the northernmost documented presence), the data type is classified as occurrence.
Format	Character
Values	occurrence, abundance

uncertainty_distribution

Field	**uncertainty_distribution**
Description	Whether the data used to estimate the range shifts were pre-processed to account for spatial imbalances across time periods arising from unbalanced sampling designs or varying species detectability.
Format	Character
Values	The same spatial units were resurveyed (Resurveyed), the data were used without any pre-processing (Raw), the data were resampled using statistical techniques such as bootstrapping (Resampling), statistical models such as generalized linear models were used to estimate the range shifts (Model), statistical models accounting for species imperfect detection were used to estimate the range shifts (Detectability), or any combinations of these categories.

position_definition

Field	**position_definition**
Description	The statistical definition of the focal range parameter used to estimate range shifts in the original article, including definitions such as “95th percentile of latitudinal occurrences”, and “Latitude of the provincial boundary of the northernmost province occupied”.
Format	Character
Values	Free text descriptions of methods from original articles

position_definition_category

Field	**position_definition_category**
Description	Broader categorical grouping of the position_definition, allowing comparison across studies that use conceptually similar but operationally distinct definitions. For example, “5% occurrence”, “2.5% occurrence”, “10% occurrence” are grouped as “Quantile”.
Format	Character
Values	Central tendency: range center defined using methods such as mean, median, midpoints, or weighted means. Furthest occurrence: range edge defined by the furthest observed individual. Mean/median of extremes: range edge defined by the average or median position of a defined subset of the furthest range observations, e.g., the mean latitude of the 10 furthest occurrences.

	Model: range positions defined using model output Political boundary: range edges defined using political boundaries such as country borders. Quantile: percentiles or weighted percentiles used to define range edges or centroids Unclear: range edge definition not provided or does not fit into prior categories

5. Taxonomic Information — add_taxo()

Adds full taxonomic classification for each species.

kingdom

Field	kingdom
Description	Taxonomic kingdom to which the species belongs.
Format	Character

phylum

Field	phylum
Description	Taxonomic phylum to which the species belongs.
Format	Character

class

Field	class
Description	Taxonomic class to which the species belongs .
Format	Character

family

Field	family
Description	Taxonomic family to which the species belongs.
Format	Character

taxon_id

Field	taxon_id
Description	Standard taxonomic identifier for the species, drawn from a recognized taxonomic reference (e.g., GBIF, ITIS, or NCBI taxonomy). Enables consistent cross-referencing with external databases.
Format	Character

6. Climate Baseline Variables — add_baselines()

Adds the average temperature within the study area (or species-specific study area) over the study duration. Column names vary by the statistic requested via the type, stat, and res arguments.

baseline_temp_{stat}

Field	baseline_temp_{stat}
Description	Average temperature within the study area (or species-specific study area clipped to the species range, when available) over the duration of the study period. The statistic applied (either mean or standard deviation) is indicated by the {stat} placeholder in the column name, so it will either be baseline_temp_mean or baseline_temp_sd.
Format	Numeric
Values / Notes	Units: degrees C

baseline_res

Field	baseline_res
Description	Spatial resolution (in km) of the raster data used to calculate baseline temperature values.
Format	Character
Values	1km, 25km, 50km, 110km

7. Climate Trend Variables — add_trends()

Adds the average change per year in temperature within the study area over the study duration. Positive values indicate warming trends; negative values indicate cooling. Column names vary by statistic and resolution.

trend_temp_{stat}

Field	trend_temp_{stat}
Description	Average change per year in temperature (temperature trend) within the study area (or species-specific study area) over the study duration. The statistic applied (either mean or standard deviation) is indicated by the {stat} placeholder, so it will either be trend_temp_mean or trend_temp_sd. Positive values indicate warming; negative values indicate cooling.
Format	Numeric
Values / Notes	Units: degrees C/year

trend_res

Field	trend_res
Description	Spatial resolution (in km) of the raster data used to calculate temperature trend values. Terrestrial latitudinal study areas are calculated at 1, 25, 50,

	110km, marine latitudinal studies are calculated at 25, 50, 110km, and elevation studies are calculated only at 1km.
Format	Character
Values	1km, 25km, 50km, 110km

8. Climate Velocity Variables — add_cv()

Adds the velocity of isotherm shifts (climate velocity) in the latitudinal or elevational direction within the study area.

along_gradient

Field	along_gradient
Description	Gradient along which climate change velocity in cv_temp_var is calculated.
Format	Character
Values	latitude, elevation

cv_temp_{stat}

Field	cv_temp_{stat}
Description	Velocity of isotherm (climate velocity) shifts along the latitudinal or elevational direction within the study area or species-specific study area over the study duration. Positive values indicate shifts towards higher latitudes or elevations, while negative velocities indicate shifts towards the equator or lower elevations. Units are km/year for latitudinal shifts and m/year for elevational shifts. Climate velocities for the study area or species-specific study area were calculated using the mean, median, 25th (q25), and 75th (q75) percentiles, and standard deviation (sd). By default, {stat} is mean, i.e., “cv_temp_mean”, and specific metrics can be selected using the stat argument in the add_cv() function.
Format	Numeric

cv_temp_var

Field	cv_temp_var
Description	The variable used for climate velocity calculations (cv_temp_{stat} columns), either "Mean Air Temperature" or “Sea Surface Temperature"
Format	Character

cv_res

Field	cv_res

Description	Spatial resolution of the raster data used to calculate climate velocity values. Note that coarser resolutions will generally produce faster velocities, since climate velocity is calculated as climate trend / spatial gradient.
Format	Character
Values	1km, 25km, 50km, 110km

9. Study Area Polygon Information — add_poly_info()

Adds summary information for the study area or species-specific study area polygon in which the shift was detected.

lat_min

Field	lat_min
Description	Minimum projected y-coordinate of the study area or species-specific study area polygon, in meters, in the Mollweide projection.
Format	Numeric

lat_max

Field	lat_max
Description	Maximum projected y-coordinate of the study area or species-specific study area polygon, in meters, in the Mollweide projection.
Format	Numeric

lat_cent_deg

Field	lat_cent_deg
Description	Latitudinal centroid of the study area polygon, in decimal degrees.
Format	Numeric

lon_cent_deg

Field	lon_cent_deg
Description	Longitudinal centroid of the study area polygon, in decimal degrees.
Format	Numeric

lat_extent_km

Field	lat_extent_km
Description	Latitudinal extent of the study area polygon, in kilometers, calculated as the distance from the maximum to the minimum latitude.
Format	Numeric

ele_mean_m

Field	ele_mean_m
Description	Mean elevation within the study area polygon, in meters above sea level.
Format	Numeric

ele_min_m

Field	ele_min_m
Description	Minimum elevation within the study area polygon, in meters above sea level.
Format	Numeric

ele_max_m

Field	ele_max_m
Description	Maximum elevation within the study area polygon, in meters above sea level.
Format	Numeric

ele_extent_m

Field	ele_extent_m
Description	Elevational range (extent) within the study area polygon, in meters, calculated as ele_max_m minus ele_min_m.
Format	Numeric

area_km2

Field	area_km2
Description	Total area of the study area polygon, in square kilometers.
Format	Numeric

study_area

Field	study_area
Description	Description of the study area location.
Format	Character
Values / Notes	Free text description of the study area location

10. Study Area Polygons — add_polygons()

Adds spatial polygons of study areas or species-specific study areas over which shifts were calculated to the shifts dataframe. Note that this requires polygons to be downloaded locally with the download_polygons() function first. The function converts the input tibble/data frame to an `sf` simple-features tibble/dataframe, allowing it to be used with spatial methods from the `sf` package.

geom

Field	geom
Description	Spatial geometry (polygon) of the study area.
Format	sf geometry column, MULTIPOLYGON geometry in WGS 84 geographic coordinates

Appendix 3. BioShifts Data Submission Protocol

In order to streamline the integration of new range shift reports or old reports that did not appear in our initial searches into future versions of BioShifts, we have created a suite of resources for authors to enter their own range shift estimation data. We created a submission form {link}, in which authors can use our data template to enter their own species-specific range shift data to the best of their ability. Participation in self-submission will ensure that applicable papers reach the BioShifts team for screening, and will speed up data extraction, which is not always externally available to readers of scientific articles. We also created an email address, bioshifts.data@gmail.com, to which other unformatted submissions or inquiries can be sent.

Table S1: Methodological metadata variables included in the *BioShifts v2* database.

Methodological Variable(s)	**Column Name**	**Values**	**Definition**
Ecological Realm	eco	Ter; Mar	Whether the study was performed in the terrestrial or marine realm. We note that the definition of the two realms was based on whether the study occurred within the land masses or the oceans and as such, no distinction was made for freshwater species (considered as terrestrial here).
Range Shift Type	type	ELE; LAT	Whether the range shifts was estimated along the latitudinal or elevational gradient.
Range Parameter	param	LE; O; TE	Whether the range shift estimate was documented at the leading edge (LE), defined as the poleward (high-latitude) or upslope (high-elevation) range margin within the study area; range center (O), defined as the centroid or mid-point of the species distribution within the study area; or trailing edge (TE), defined as the equatorward (low-latitude) or downslope (low-elevation) range margin within the study area. We note that we used the definition provided in the original articles; we did not attempt to verify whether the range margins aligned with the cold and warm species range limits, nor did we verify whether the range limits within the study areas aligned with the 'true' range limits of species distributions (see (Parker et al., 2024) for a discussion).
First Period, Last Period	start_firstperiod; midpoint_firstperiod; end_firstperiod; start_lastperiod; midpoint_lastperiod; end_lastperiod	Integer	The characteristic of the first and last sampling periods used to estimate the range shift. If detections are made via distinct annual surveys-resurveys (e.g., 1965 versus 2007), start, midpoint, and end years will be the same for each period. For "census" surveys (e.g., occurrence records from 1965-1969 versus 2007-2009), start, midpoint, and end will describe the census period.
Study Duration	duration	Integer	Total study duration in years, estimated as the difference between the midpoint year of the last study period and the midpoint year of the first study period + 1.
Number of Sampling Periods	n_periods	Integer	Total number of distinct temporal units that were used to estimate the range shifts. For example, for range shifts estimated between two census periods (e.g., 1965-1969 and 2007-2009), the number of temporal units is two; for range shifts estimated using annual time series, the number of temporal units would be the number of years in the study.
Grain Size	grain_size	small; moderate; large; very_large	The spatial resolution of the species data used to estimate the shifts: less than 1 km^2, such as a well-delineated site or transect (small): 1–10 km^2, such as a gridded atlas (moderate); 10–100km^2, such as sub-regions or ecoregions (large): or >100km^2, such as states or provinces (very large).
Temporal Sampling Type	sampling	TWO; IRR/AGG; MULTIPLE (continuous); CONTINUOUS	The temporal distribution of sampling used to estimate range shifts: estimated between two time units (TWO); three time periods composed of one or more years or aggregated such as in decadal intervals (IRR/AGG); multiple time units that were treated as continuous in the estimation process even if they may be irregular (MULTIPLE (continuous)); or multiple and regular sampling periods (CONTINUOUS).
Observation Data Type	obs_type	occurrence; abundance	The type of species data used to estimate the range shifts, i.e., occurrences versus abundances. We note that even if the underlying dataset includes abundances, if the data were transformed to presence-absence to estimate the range shift (e.g., using the northernmost documented presence), the data type is classified as occurrence.

Data Processing	uncertainty_distribution	Detectability; Model; Raw; Resampling; Resurveyed; and combinations thereof	Whether the data used to estimate the range shifts were pre-processed to account for spatial imbalances across time periods arising from unbalanced sampling designs or varying species detectability: the same spatial units were resurveyed (Resurveyed), the data were used without any pre-processing (Raw), the data were resampled using statistical techniques such as bootstrapping (Resampling), statistical models such as generalized linear models were used to estimate the range shifts (Model), statistical models accounting for species imperfect detection were used to estimate the range shifts (Detectability), or any combinations of these categories.
Sampling Strategy	category	Survey-Resurvey; CensusPeriods; TimeSeries; Synthesis; and combinations thereof	Whether the range shifts were estimated between two distinct years (Survey-Resurvey), between two periods composed of more than one year of species data (CensusPeriods), using (usually annual) time series data (TimeSeries), or using species distribution limit estimates gathered from previous literature, generally spanning multiple years (Synthesis). Cases in which the first and last time period had different compositions are indicated as Survey-Census or Census-Resurvey).
Definition of Parameter Position	position_definition	Free text descriptions of methods from original articles	The statistical definition of the focal range parameter used to estimate range shifts in the original article, including definitions such as “95th percentile of latitudinal occurrences”, and “Latitude of the provincial boundary of the northernmost province occupied”.
Definition of Parameter Position: Categorized	position_definition_category	Categories of position_definition values	Summarized groups of position_definition free text values. For example, “5% occurrence”, “2.5% occurrence”, “10% occurrence” are grouped as “Percentile”.
Focal subpopulation or sublocality within a species/study	sublocality_or_pop	Free text descriptions of focal groups within a species and study polygon	When relevant, the distinguishing group, subpopulation, or other identifying difference for multiple shifts recorded for the same species within the same study polygon and focal range parameter and using the same methods (e.g., saplings vs. adults, queen vs. worker bees, fall vs. spring trawl sampling).

Table S2. Functions of the `BioShiftR` R package

Function	Arguments	Description
Starter Function		
`get_shifts()`	group, realm, continent, type	Acquire minimal range shift data from the *BioShift v2* database; easily filter by generalized taxa, realm, geography, or shift type. **This is the starter function** that provides data to which any other functions can be added.
Add Specificities of Range Shift Estimates		
`add_articles()`	NA	Merge source article identifiers to the range shifts database.
`add_author_reported()`	NA	Merge unstandardized, author-reported range shift estimates to the range shifts database.
`add_taxo()`	NA	Merge taxonomy data to the range shifts database.
`add_methods()`	NA	Merge methodological attributes of range shift estimates to the range shifts database.
Add Spatial Data		
`add_polygons()`	type, directory	Merge spatial polygons to the range shifts database. Choose to merge polygons of study polygons, or species-specific study polygons, when available. Polygon databases are stored externally in Open Science Framework; the function will prompt a download if not already done.
`add_poly_info()`	type	Merge polygon info to the range shifts database, including latitudinal and elevational bounds of focal polygons, total area, and description of the study area.
`download_polygons()`	type, directory, timeout	Download polygons from external source; precursor function to `add_polygons()`. Choose to download study-level (n = 287) or species-specific polygons (n = 18,731), and the directory in which they should be saved.
Add Derived Climate Exposure Variables		
`add_baselines()`	type, stat, res	Merge baseline temperature values to the range shifts database. Choose baseline temperature values summarized within study-level or species-specific

		polygons, the summary statistic of mean temperature calculated across raster cells (mean, sd), and the spatial resolution of input temperature data used to calculate temperature baselines.
`add_trends()`	type, stat, res	Merge average temperature trend values to the range shifts database. Choose trends summarized within study-level or species-specific polygons, the summary statistic of cell-wise temperature trends within the focal polygon (mean, sd), and the spatial resolution of input temperature data used to calculate trends.
`add_cv()`	type, stat, res	Merge poleward or elevational isotherm velocity values to range shifts database. Choose climate velocities summarized over study-level or species-specific polygons, the statistic summarized across raster cells, and the resolution of input temperature data used to calculate velocities.

Table S3. GBIF DOIs for convex hulls

See attached file: gbif_dois.csv

References

Bernhardt, J., Carleton, A. M., & LaMagna, C. (2018). A Comparison of Daily Temperature-Averaging Methods: Spatial Variability and Recent Change for the CONUS. *Journal of Climate*, *31*(3), 979–996. https://doi.org/10.1175/JCLI-D-17-0089.1

Bird species distribution maps of the world. (2019). *Bird species distribution maps of the world* (Version 2019.1) [Dataset]. http://datazone.birdlife.org/species/requestdis

Bivand, R. (2023). *maptools: Tools for Handling Spatial Objects* (Version 1.1-8) [R]. Comprehensive R Archive Network (CRAN). r-project.org

Brown, C. J., O'Connor, M. I., Poloczanska, E. S., Schoeman, D. S., Buckley, L. B., Burrows, M. T., Duarte, C. M., Halpern, B. S., Pandolfi, J. M., Parmesan, C., & Richardson, A. J. (2016). Ecological and methodological drivers of species' distribution and phenology responses to climate change. *Global Change Biology*, *22*(4), 1548–1560. https://doi.org/10.1111/gcb.13184

Burrows, M. T., Schoeman, D. S., Buckley, L. B., Moore, P., Poloczanska, E. S., Brander, K. M., Brown, C., Bruno, J. F., Duarte, C. M., Halpern, B. S., Holding, J., Kappel, C. V., Kiessling, W., O'Connor, M. I., Pandolfi, J. M., Parmesan, C., Schwing, F. B., Sydeman, W. J., & Richardson, A. J. (2011). The Pace of Shifting Climate in Marine and Terrestrial Ecosystems. *Science*, *334*(6056), 652–655. https://doi.org/10.1126/science.1210288

Chamberlain, S., Barve, V., Mcglinn, D., Oldoni, D., Desmet, P., Geffert, L., & Ram, K. (2025). *rgbif: Interface to the Global Biodiversity Information Facility API* (Version 3.8.4) [R]. https://CRAN.R-project.org/package=rgbif

Chan, W.-P., Lenoir, J., Mai, G.-S., Kuo, H.-C., Chen, I.-C., & Shen, S.-F. (2024). Climate velocities and species tracking in global mountain regions. *Nature*, *629*(8010), 114–120. https://doi.org/10.1038/s41586-024-07264-9

Comte, L., Grenouillet, G., Bertrand, R., Murienne, J., Bourgeaud, L., Hattab, T., & Lenoir, J. (2020). *BioShifts: A global geodatabase of climate-induced species redistribution over land and sea* (p. 5238495 Bytes) [Dataset]. figshare. https://doi.org/10.6084/M9.FIGSHARE.7413365

Daru, B. H. (2025). A global database of butterfly species native distributions. *Ecology*, *106*(1), e4462. https://doi.org/10.1002/ecy.4462

Denelle, P., Weigelt, P., & Kreft, H. (2023). GIFT — An R package to access the Global Inventory of Floras and Traits. *Methods in Ecology and Evolution*, *14*(11), 2738–2748. https://doi.org/10.1111/2041-210X.14213

Dobrowski, S. Z., Abatzoglou, J., Swanson, A. K., Greenberg, J. A., Mynsberge, A. R., Holden, Z. A., & Schwartz, M. K. (2013). The climate velocity of the contiguous United States during the 20th century. *Global Change Biology*, *19*(1), 241–251. https://doi.org/10.1111/gcb.12026

Esri. (2024). *ArcGIS Pro* [Computer software]. Esri.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086

Fredston-Hermann, A., Selden, R., Pinsky, M., Gaines, S. D., & Halpern, B. S. (2020). Cold range edges of marine fishes track climate change better than warm edges. *Global Change Biology*, *26*(5), 2908–2922. https://doi.org/10.1111/gcb.15035

García Molinos, J., Schoeman, D. S., Brown, C. J., & Burrows, M. T. (2019). VoCC: An R package for calculating the velocity of climate change and related climatic metrics. *Methods in Ecology and Evolution*, *10*(12), 2195–2202. https://doi.org/10.1111/2041-210X.13295

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

IUCN (International Union for Conservation of Nature). (2024). *Spatial data* [Dataset]. The IUCN Red List of Threatened Species. https://www.iucnredlist.org

Karger, D. N., & Zimmermann, N. E. (2018). *CHELSAcruts—High resolution temperature and precipitation timeseries for the 20th century and beyond* (Version 1.0, p. 3298534883328 bytes, 0, 1146678 bytes) [No Info,application/pdf,application/pdf]. EnviDat. https://doi.org/10.16904/ENVIDAT.159

Lai, Y., Lin, Y. J., & Ko, C. (2024). How would estimation of geographic range shifts of marine fishes be different when using occurrence and abundance data? *Diversity and Distributions*, *30*(10), e13919. https://doi.org/10.1111/ddi.13919

Lenoir, J., Bertrand, R., Comte, L., Bourgeaud, L., Hattab, T., Murienne, J., & Grenouillet, G. (2020). Species better track climate warming in the oceans than on land. *Nature Ecology and Evolution*, *4*(8). https://doi.org/10.1038/s41559-020-1198-2

Maitner, B. (2025). *BIEN: Tools for Accessing the Botanical Information and Ecology Network Database* (Version 1.2.7) [Computer software]. https://CRAN.R-project.org/package=BIEN

Norman, K. E. A., Chamberlain, S., & Boettiger, C. (2020). taxadb: A high-performance local taxonomic database interface. *Methods in Ecology and Evolution*, *11*(9), 1153–1159. https://doi.org/10.1111/2041-210X.13440

Parker, E. J., Weiskopf, S. R., Oliver, R. Y., Rubenstein, M. A., & Jetz, W. (2024). Insufficient and biased representation of species geographic responses to climate change. *Global Change Biology*, *30*(7), e17408. https://doi.org/10.1111/gcb.17408

Ribeiro, B. R., Velazco, S. J. E., Guidoni-Martins, K., Tessarolo, G., Jardim, L., Bachman, S. P., & Loyola, R. (2022). bdc: A toolkit for standardizing, integrating and cleaning biodiversity data. *Methods in Ecology and Evolution*, *13*(7), 1421–1428. https://doi.org/10.1111/2041-210X.13868

Roll, U., Feldman, A., Novosolov, M., Allison, A., Bauer, A. M., Bernard, R., Böhm, M., Castro-Herrera, F., Chirio, L., Collen, B., Colli, G. R., Dabool, L., Das, I., Doan, T. M., Grismer, L. L., Hoogmoed, M., Itescu, Y., Kraus, F., LeBreton, M., … Meiri, S. (2017). The global distribution of tetrapods reveals a need for targeted reptile conservation. *Nature Ecology & Evolution*, *1*(11), 1677–1682. https://doi.org/10.1038/s41559-017-0332-2

Rubenstein, M. A., Weiskopf, S. R., Bertrand, R., Carter, S. L., Comte, L., Eaton, M., Johnson, C. G., Lenoir, J., Lynch, A. J., Maslyukova, D. Y., Miller, B. W., Morelli, T. Lyn., Parker, E. J., Poole, C., Rodriguez, M. A., Rudick, T., Terando, A. J., Thompson, L. M., & Turner, K. (2024). *CoRE (Contractions or Range Expansions) Database: Global Database of Species Range Shifts from 1802-2019* [Dataset]. U.S. Geological Survey. https://doi.org/10.5066/P99VP2TW

Taheri, S., Naimi, B., & Araújo, M. B. (2024). climetrics: An R package to quantify multiple dimensions of climate change. *Ecography*, *2024*(8), e07176. https://doi.org/10.1111/ecog.07176