January 2022

An Assessment of Geographic and Taxonomic Biases in Research on Climate Change-Related Range Shifts

Evan J. Parker
Yale University

Follow this and additional works at: https://elischolar.library.yale.edu/library_map_prize

Part of the Biodiversity Commons, Climate Commons, and the Research Methods in Life Sciences Commons

Recommended Citation
https://elischolar.library.yale.edu/library_map_prize/14

This Article is brought to you for free and open access by the Library Prizes at EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Library Map Prize by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.
An Assessment of Geographic and Taxonomic Biases in Research on Climate Change-Related Range Shifts

Evan Parker

Department of Environmental Studies
Yale University
Advisor: Walter Jetz
Colloquium Mentor: Kealoha Freidenburg
# Table of Contents

**Abstract**  
3

**Introduction**  
4  
Species distributions in a changing climate  
4  
Variability in range shifts  
5  
Bias in ecological research  
7  
Bias in range shift literature  
8

**Methods**  
9  
Literature review and database creation  
9  
Creation of geographic data  
10  
Geographic coverage  
10  
Taxonomic Coverage  
11  
Within-range coverage  
11

**Results**  
14  
Geographic Coverage  
14  
Taxonomic Coverage  
17  
Within-range coverage  
18

**Discussion**  
21  
Geographic coverage  
21  
Taxonomic coverage  
23  
Within-range coverage  
24  
Limitations  
29

**Conclusion**  
30

**References**  
32

**Appendices**  
41
Abstract

Geographic and taxonomic biases in ecological literature inhibit the ability of research to be generalized or translated to other systems. This is particularly true with research on climate change-related range shifts, where a diverse and representative set of observations is needed to understand the significant heterogeneity of range shifts and improve models. A literature review was conducted and biases were assessed by examining the geographic location of study areas, the taxonomic makeup of study species, and the location of study areas within species ranges. A substantial geographic skew towards terrestrial, mid-latitude North America and Europe was observed. However, taxonomic coverage was even and representative within plants and animals. At the level of individual species ranges, the vast majority of study areas were not representative and study areas failed to effectively include range extremes. Based on these results, scientists must more strategically design studies to include diverse geographic regions and to representatively capture range shifts throughout a species range.
Introduction

Species distributions in a changing climate

Anthropogenic climate change is a clear and worsening phenomenon, with global average surface temperatures having increased by over 1 °C and potential future pathways including increases of over 4 °C (Masson-Delmotte et al. 2021). Such changes in climate are certain to have numerous ecological impacts (Weiskopf et al. 2020). One important response of organisms to changing climate is range shifts. Range shifts occur when the distributions of species are altered in order to maintain suitable climatic niches. The general expectation is that species will shift towards the cool higher latitudes and elevations as global temperatures increase. Range shifts have indeed been documented across a variety of taxa, geographic locations, and dimensions (Chen et al. 2011; Lenoir et al. 2020; Rubenstein et al. in review).

Although species may track changes in climatic niche, shifts also occur in the context of additional factors that can inhibit the tracking ability of a species, such as habitat availability and interspecific competition. An alternative to range shifts is thus range contraction and possible extinction (Morelli et al. 2020). A study of the South African fynbos, for example, found that decreased habitat availability due to a hotter and drier climate could lead to the extinction of 10% of endemic species of the family Proteaceae, with 95% of species losing at least a third of their range (Midgley et al. 2002). Similar studies in other regions of the world produce comparable results, with some predictions including the extinction of entire clades and over 20% of species (Freeman et al. 2018; González-Orozco et al. 2016; Thomas et al. 2004).
Beyond extinction of individual species, range shifts can also lead to ecosystem-level changes. For instance, changing species distributions result in novel community assemblages and species interactions, with predictions of species loss in the community at a given location over time being as high as 33% in Europe and 66% in southern Africa (Bakkenes et al. 2002; Erasmus et al. 2002). Species loss will be at least somewhat counteracted by range expansion of other species, in addition to the increased spread of invasive species (Bellard et al. 2013; Mainka and Howard 2010). These changes in community structure can in turn have impacts on important ecosystem functions and services (Dieleman et al. 2015; Fadrique et al. 2018; Pinsky et al. 2020; Pugnaire et al. 2019).

Variability in range shifts

Given the potential of range shifts to significantly restructure ecosystems, thorough research on when and how they occur is of paramount importance. This is particularly true given the observed heterogeneity of the direction and magnitude of shifts (Rubenstein et al. in review). Although broad hypotheses predict that species will simply move to maintain their climatic niche, observed shifts suggest that variation exists along geographic and taxonomic lines. For instance, terrestrial species are often limited by terrain and human activity, and their ranges consequently track isotherms less closely than those of marine species (Lenoir et al. 2020). Geographic features ranging in prominence from coastlines to local areas of lower temperature can create variability in patterns of range shifts (Burrows et al. 2014). Latitude and elevation can also impact range shifts, as low-elevation and tropical shifts tend to more closely follow temperature than those at high elevations or in
temperate regions (Freeman et al. 2021; Mamantov et al. 2021). Additionally, different areas of a range can respond differently to climate change. In fact, one study suggested that some species may display climate change-induced range shifts in one part of the range, while the broader range shows very little change in relation to climate (Currie and Verne 2017). Alternatively, local shifts might vary throughout a range due to short-term climate variation, lag, competition, or habitat availability (Bates et al. 2014; Li and Park 2020; Renwick et al. 2016).

The importance of these factors also can vary among taxa due to a number of species traits (MacLean and Beissinger 2017; Pacifici et al. 2017; Valladares et al. 2017; Yusefi et al. 2021). For example, habitat specialists generally shift more slowly and are more affected by decreases in habitat availability (Chen et al. 2011; Platts et al. 2019). The exact role of species traits, however, generally remains unclear. For example, one study found that species with traits such as fossoriality and short generation time were less negatively impacted by climate change (Pacifici et al. 2017). However, another study suggested that fecundity is linked to increased shift magnitude as often as it is to slower shifts, while body size is important primarily in birds and marine fish (MacLean and Beissinger 2017). While shifts are evidently linked in some way to species traits, there is significant variability. Overall, it is clear that range shifts occur as the result of complex interactions between climate change exposure, geography, ecological interactions, and species traits. This complexity underlines the need to document range shifts across diverse geographic areas and taxonomic subjects in order to obtain a more complete understanding of how range shifts occur and consequently improve biodiversity management in a changing climate.
Bias in ecological research

Geographic representation is unequal in ecology research. For instance, an analysis of publications in top ecology journals has indicated that protected forests in wealthy countries are significantly overrepresented as study sites. These trends are likely a result of logistics, cultural habits, and the perception of nominally untouched areas as more representative of pure ecological processes (Martin et al. 2012; Mirtl et al. 2018). However, geographic skews in research can bias understanding of ecological processes and create path dependencies that lead to an extended legacy of biased literature. This is particularly problematic given that the overrepresented areas comprise only a small portion of the Earth's area and human impacts continue to grow more widespread, creating a less and less applicable ecological literature. Geographic imbalances in research effort thus inhibit the ability of research to be generalized and effectively applied to conservation or management activities (de los Ríos et al. 2018; Martin et al. 2012; Meyer et al. 2016; Nabout et al. 2012; Oliver et al. 2021).

Taxonomic skews have also been documented in ecological literature. In the fields of climate change vulnerability and invasion ecology, plants are highly represented while invertebrates are often ignored (de los Ríos et al. 2018; Jeschke et al. 2012). A bias can also be seen at finer taxonomic scale, with one study finding skewed coverage of herpetofauna families in climate change research (Winter et al. 2016). Generally, research focuses on taxa that are more charismatic or easily studied. This type of bias also challenges the ability of research to be generalized or translated into management actions, and is particularly concerning given that there has been little improvement in taxonomic coverage in the past 70 years (Troudet et al. 2017).
Bias in range shift literature

The relatively short period over which range shift studies have occurred means that there is a smaller literature on its research biases. Lenoir and Svenning (2014) reviewed 212 studies to examine geographic and taxonomic biases. The authors found that the Northern Hemisphere, particularly Europe, was far more studied than the Southern Hemisphere. Greenland, Antarctica, and the tropics were also all relatively underrepresented. There was also a skew in biome coverage, with a greater focus on temperate and boreal forests. Marine studies were somewhat more representative, but were still concentrated at higher latitudes. The authors also reported a taxonomic bias; animals dominated study species, particularly in marine studies (Lenoir and Svenning 2014). A later follow-up reported similar results (Lenoir et al. 2020). Overall, these studies suggest that range shift research likely follows similar geographic and taxonomic trends as broader ecology literature.

Given the need to account for and understand the significant variation in range shifts across geography and taxonomy, an updated assessment of research bias is imperative. The most recent meta-analysis documented taxonomic and geographic biases, but did not examine within-range bias (Lenoir et al. 2020). It is important to determine how recent research has increased geographic and taxonomic coverage, as well as to expand coverage assessments to consider within-range coverage. Thus, I ask the following question: how well do range shift studies include a diverse and representative set of study areas and species? This paper will expand on previous reviews and examine geographic coverage on a species level to assess how well species ranges are geographically and climatically represented. An examination of these types of coverage will indicate how future research
can better produce a greater fundamental understanding of range shifts and thus better support biodiversity and ecosystem management in the face of a changing climate.

**Methods**

*Literature review and database creation*

A literature review was conducted following the procedure outlined by Rubenstein et al. (2020). The databases Web of Science and Scopus were searched using the following search terms: \( TS = (\text{"climate" OR "global warming" OR "temperature" OR "precipitation"}) \) AND \( (\text{"range" OR "distribution" OR "habitat extent" OR "occupancy"}) \) AND \( (\text{"species"}) \). Retrieved articles were then entered into Colanddr, which uses machine learning to sort articles by relevance. The abstracts of the most relevant articles were reviewed, accepting or rejecting articles based on their use of taxa of interest (animals and plants only), anthropogenic climate change, observed rather than modeled shifts, and species-level shifts. Review of new abstracts ended when 300 were reviewed without finding any new articles. Additionally, 1000 articles were retrieved from Google Scholar using the same search terms, with new articles being reviewed for entry into the database. Other articles were retrieved from the citations of accepted articles and prior literature reviews. At this point, 840 studies had passed an abstract review and underwent a full-text review using the same criteria. After the final review, 313 studies were retained for entry into the database.

Next, data were extracted from each study, several of which were relevant to this paper: location of the study, focal species, whether the observation was of a change of
abundance or occupancy, dimension (elevation, latitude, depth, or longitude), and parameter (leading, trailing, or mean). Leading edges were defined as high-latitude or high-elevation edges, while trailing edges were defined as low-latitude or low-elevation edges. Each unique species shift observation was entered as its own row. This spreadsheet was put into R (v4.1.1; R Core Team 2021) for analysis. A number of R scripts were then run to perform cleaning measures such as taxonomic harmonization, fixing typos, etc (Rubenstein et al. in review; https://github.com/ejparker08/Thesis). This left the final primary database with 33,016 unique range shift observations.

Creation of geographic data

Using the data recorded in the database for each observation, geographic data were created. Each article was scanned to find, if it existed, spatial data for download from the authors (n = 5 studies). If this was not possible, study area shapefiles were drawn in Google Earth Pro (v7.3.4.8248) using article figures (n = 251 studies) or text descriptions (n = 74 studies). This was done at the population, species, or study level, depending on availability. Polygon names and corresponding information about each species were recorded in a separate spreadsheet. Lastly, minimum and maximum elevation of each study area were extracted from each article as available. The shapefiles were then downloaded from Google Earth Pro as a KMZ file. The tool KML to Layer in ArcGIS Pro (v2.8.0) was then used to prepare the shapefiles for analysis.

Geographic coverage

In order to assess geographic coverage, a heat map of research effort was created at the species level. First, the polygons were imported in R. Next, all study area polygons
corresponding to each species were combined using the sf package (v1.0.3; Pebesma 2018) to create species-level polygons representing all of the locations in the database where that species has been studied. Then, that shapefile was exported into ArcGIS Pro. Using a grid of 3025 km$^2$ cells, the number of species polygons intersecting each cell was calculated. The resulting map displays research effort as the number of species assessed within each grid cell. The same process was followed in ArcGIS Pro with the study-level polygons to create a map displaying the number of study areas overlapping with each grid cell.

**Taxonomic Coverage**

Taxonomic coverage was assessed by determining the number of unique species of each taxonomic group represented in the database, regardless of the number of observations or studies. In addition to absolute species counts, relative species counts were also calculated. The total number of described species in each taxa was taken from Catalog of Life (catalogueoflife.org; Bánki *et al.* 2022), and the number of studied species was then divided by this number to obtain research effort relative to taxon size.

**Within-range coverage**

First, range distribution data was obtained for each unique species represented in the database. However, due to taxonomic limitations on range map availability, this section was restricted to species of Amphibia, Aves, Mammalia, and Reptilia (n = 1803 species). Range maps were downloaded from *Map of Life* (mol.org; Jetz *et al.* 2012) in the form of intersected grid tables. Each species had a table with a list of cells that intersected with the range map. This grid had cells with an area of 3025 km$^2$ and each cell had a standardized
name. Each individual species matrix was imported into R and all were combined into a single matrix.

Next, a function was created to calculate all of the cell IDs with which a given study area polygon intersected. After running this function on all species study area polygons, the resulting species-level matrices were combined into a single matrix. Additionally, the latitude of the centroid of each grid cell was calculated and joined to both the range and study area intersection matrices. Lastly, a table with elevational limits of species was downloaded from Map of Life.

Of the species in the selected taxa, species that lacked range data or which only intersected with a single cell and were therefore at too fine of a scale to assess were excluded. For the remaining species (n = 1319 species), a t-test was conducted on the latitude of cells intersected by their study areas or range in order to determine the latitudinal representativeness of study areas. Additionally, the distance between the maximum and minimum latitude and elevation of the study areas and ranges was calculated to assess range edge coverage (Figure 1). Differences of less than 0 were assumed to be the result of multi-species studies and were thus replaced with a value of 0. Due to the relative coarseness of the latitudinal and elevational data, an edge was considered to be assessed if the difference was less than or equal to 1 degree of latitude or 100 m of elevation. The size of the buffers was determined by gradually increasing them until there was little further change in coverage with each unit increase. The fact that coverage increases when using a buffer of these sizes but not larger indicates that the buffer is adjusting for data coarseness, not simply including areas that truly lack coverage.
Example Species Within-Range Metrics

**Western Harvest Mouse** (*Reithrodontomys megalotis*)

**Fire Salamander** (*Salamandra salamandra*)

*Figure 1:* Example Species Within-Range Metrics. On the left are ranges (gray) and study areas (black) overlaid on a 3025 km² grid. Leading edge coverage was determined by calculating the difference of the maximum latitudes of the range and study areas, while minimum latitudes were used for trailing edge coverage. The individual cell latitudinal distributions of range and study areas of each species were turned into density curves, with vertical lines indicating the mean latitudes. A t-test was used to find the difference of these means. The western harvest mouse is an example of a species that is latitudinally representative (*p* > 0.05), but has no latitudinal edge coverage. Instead, it has multiple studies that solely document shifts on the elevational trailing edge. The fire salamander, on the other hand, is not latitudinally representative (*p* < 0.05). It does, however, have a study along the latitudinal trailing edge and coverage of both elevational range edges. These contrasting species illustrate how the two metrics can complement each other.

Within-range distribution of average annual temperature was also considered for two reasons. First, the entire edge of a species range might not necessarily be at its latitudinal or elevational absolute maximum or minimum. Instead, there can be variation in
the latitude or elevation of a range edge due to a number of abiotic and biotic factors. For one, the latitude of the poleward edge of a range can decrease if elevation increases along that edge (Figure 1). However, temperature would remain relatively constant along the edge, and so can provide a better metric for coverage of range edges. Second, average annual temperature is an important climate variable that influences species distribution. Since range shifts likely occur variably across temperature gradients, it is important that range shift assessments are performed not only at temperature extremes, but also representatively across the temperature range of a given species range (Colwell et al. 2008; Coristine and Kerr 2015; Paquette and Hargreaves 2021; Sunday et al. 2012).

Average annual temperature (bio1) data was downloaded from CHELSA (chelsa-climate.org; Karger et al. 2017). Then, the mean values within each grid cell were calculated and put into a table. This table was imported into R and joined with the grid intersection matrices for both study areas and ranges. A t-test was then conducted on the temperature values for each species to determine representativeness of temperature. Additionally, the difference between the maximum and minimum temperature values for study areas and ranges was calculated. For this, a temperature extreme was considered to be covered if the difference was less than or equal to 0.1 °C.

Results

Geographic Coverage

General geographic trends appear to match those of the general ecological literature (Martin et al. 2012; Mirtl et al. 2018). Notably, the vast majority of research effort is
Figure 2: Geographic Distribution of Research Effort. Each grid cell is colored to represent the number of marine species assessed within the cell (A), the number of terrestrial species assessed within the cell (B), the number of marine studies overlapping with the cell (C), and the number of terrestrial studies overlapping with the cell (D). All maps were log-transformed to improve visibility. Plots along the left of each map indicate the relative number of species assessed within (A, B) or the relative number of studies overlapping with (C, D) each 10 degree latitudinal band. See Appendices A-D for larger copies.
concentrated in the mid-latitudes. Specifically, there is an abundance of effort in the United States and Canada, northwestern Europe, South Africa, and parts of Australia (Figure 2). This is in contrast to the dearth of studies in the polar regions, with only 3 polar-focused studies, two on fish and one on Adélie penguins (Landa et al. 2014; Mecklenburg et al. 2007; Taylor and Wilson 1990). Research effort is greater but still minimal in the tropics, with primarily only a scattering of small studies in mountainous regions (Berton et al. 2012; Chen et al. 2009; Forero-Moreno et al. 2011; Freeman and Freeman 2014; Molina-Martínez et al. 2016; Pinault and Hunter 2011; Raxworthy et al. 2008; Seimon et al. 2007). Coverage is much better in the temperate zone, but is still geographically biased within the region. For instance, there is a noticeable lack of research in most of temperate Asia or South America.

There is a broader distribution of marine studies, although there still is a clear concentration of research in mid-latitude United States, northwestern Europe, and Tasmania (Figure 2). A small number of species were studied in open-ocean areas worldwide, as well as several studies off the coast of Angola and subtropical Japan and Australia (Erauskin-Extramiana et al. 2020; Kurihara et al. 2011; Poloczanska et al. 2011; Potts et al. 2014; Worm and Tittensor 2011; Yemane et al. 2014). The contrast of Figures 2A and 2C illustrate that, although the number of assessed species is more equally distributed, the majority of studies are centered in northwestern Europe. While there is some marine research being conducted at many latitudes, the vast majority of research effort is also focused on near-shore areas in or near mid-latitude Western countries.
**Figure 3: Number of Assessed Species per Taxon.** Number of species with at least one range shift observation indicated in orange, as a proportion of each taxon’s total number of described species in Catalog of Life. The values of n indicate the absolute number of assessed species in each taxon. Asterisks * indicates the taxa whose relative research effort is an outlier.

There is a clear skew of research effort among taxonomic groups as reported in previous studies (Figure 3; Lenoir *et al.* 2020). Angiosperms, insects, and birds were particularly well represented, with angiosperms having nearly 100 times as many species assessed as reptiles. However, angiosperms are the only outlier in terms of raw species count (Appendix E). Relative species count gives quite contrasting results. When considered this way, research effort is generally more even. Angiosperms are in the middle of the range of values, while insects actually become the least represented group. However, there are only two taxa that are outliers: birds and gymnosperms, both of which are represented to a greater degree than the other included taxa. It is difficult to determine the number of described species in the “other” group, which includes any species not in the other taxonomic groups. However, the number of non-insect invertebrate species might be as
many as 300,000, which would make that group the least represented taxa (catalogueoflife.org; Bánki et al. 2022). Consequently, it is important to still pay attention to that group of organisms.

**Within-range coverage**

Overall, both representativeness and edge coverage were quite poor. The study areas of less than 7% of species were latitudinally representative of their ranges (Figure 4). While there was some variation among taxa, only 17% of species in the best taxon were representative (Figure 4). Similarly, range edge coverage was also low. Just over 10% of species had no latitudinal or elevational edges assessed, while less than 1% had all four edges assessed (Appendix F). When broken down by dimension, however, several patterns can be seen. First, the elevational range of species is much better covered than the latitudinal range. 617 species had both elevational extremes covered, while only 11 species

![Species Latitudinal Representativeness](image)

**Figure 4:** Species Latitudinal Representativeness. Results of a t-test on the latitudinal distributions of range and study areas for each species, interpreted as representative ($p > 0.05$) and unrepresentative ($p < 0.05$). Absolute number of species that were representative or unrepresentative for each taxa provided as data labels.
**Figure 5**: Species Latitudinal Edge Coverage. Edge coverage was determined by calculating the difference of the maximum and minimum latitudes, respectively, of study areas and ranges for each species. If the difference was less than the 1 degree buffer, the edge was considered to be covered. Absolute number of species with each amount of coverage for each taxon provided as data labels.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>None</th>
<th>Trailing Edge</th>
<th>Leading Edge</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amphibian</td>
<td>68</td>
<td>48</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Bird</td>
<td>760</td>
<td>150</td>
<td>128</td>
<td>3</td>
</tr>
<tr>
<td>Mammal</td>
<td>69</td>
<td>14</td>
<td>10</td>
<td>1</td>
</tr>
<tr>
<td>Reptile</td>
<td>26</td>
<td>11</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>923</td>
<td>223</td>
<td>162</td>
<td>11</td>
</tr>
</tbody>
</table>

**Figure 6**: Species Elevational Edge Coverage. Edge coverage was determined by calculating the difference of the maximum and minimum elevations, respectively, of study areas and ranges for each species. If the difference was less than the 100 m buffer, the edge was considered to be covered. Absolute number of species with each amount of coverage for each taxon provided as data labels.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>None</th>
<th>Trailing Edge</th>
<th>Leading Edge</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amphibian</td>
<td>9</td>
<td>18</td>
<td>16</td>
<td>88</td>
</tr>
<tr>
<td>Bird</td>
<td>127</td>
<td>224</td>
<td>93</td>
<td>504</td>
</tr>
<tr>
<td>Mammal</td>
<td>23</td>
<td>25</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>Reptile</td>
<td>12</td>
<td>4</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>171</td>
<td>274</td>
<td>118</td>
<td>617</td>
</tr>
</tbody>
</table>
had both latitudinal edges covered (Figures 5, 6). For both elevation and latitude, the trailing edge was more frequently studied than the leading edge. Unlike with representativeness, there were taxonomic differences in range edge coverage. Notably, amphibians had better edge coverage than birds and mammals.

**Figure 7:** Species Temperature Representativeness. Results of a t-test on the average annual temperature (bio1) distributions of range and study areas for each species, interpreted as representative (p > 0.05) and unrepresentative (p < 0.05). Absolute number of species that were representative or unrepresentative for each taxa provided as data labels.

**Figure 8:** Species Temperature Extreme Coverage. Temperate extreme coverage was determined by calculating the difference of the maximum and minimum average annual temperatures (bio1), respectively, of study areas and ranges for each species. If the difference was less than the 1 °C buffer, the temperature extreme was considered to be covered. Absolute number of species with each amount of coverage for each taxon provided as data labels.
When compared to latitude and elevation, the results for temperature were broadly similar. While study areas were representative of ranges for three times as many species when considering temperature rather than latitude, the majority still were not (Figure 7). Coverage of temperature extremes, on the other hand, was mixed. The number of species with neither extreme covered was much greater than those with both extremes covered. Temperature coverage was better than latitudinal edge coverage, but worse than elevational edge coverage (Figures 5, 6, 8). Perplexingly, and like with elevation and latitude, the minimum temperature extreme was also covered much more frequently than the maximum (Figure 8).

Discussion

Geographic coverage

This study found a significant geographic skew to the locations of range shift observations similar to that of past meta-analyses (Chen et al. 2011; Lenoir et al. 2020; Lenoir and Svenning 2014). This geographic heterogeneity limits the ability to predict distribution changes for the vast number of species that will inherently be unmonitored, yet nonetheless strongly impacted by climate change. For example, there are differences in the way that tropical and temperate species shift in response to climate change. In the extensive flat regions in parts of the tropics, such as the Amazon and Congo basins, more gradual latitudinal temperature gradients mean that species must shift further in order to maintain the same climate envelope (Şekerçioglu et al. 2012). Additionally, tropical species often have narrower climate niches (Addo-Bekiako et al. 2000; Brown 2014; Ghalambor et
al. 2006; Perez et al. 2016; Vázquez et al. 2004). This might mean that species inhabiting
those areas do shift greater distances, but it also might mean that they fail to maintain their
climate niche. Given the lack of studies assessing latitudinal shifts in either of those basins,
predictions of distribution changes in those regions must be made using models that
cannot be effectively validated against real-world observations, instead relying on those in
completely different, temperate parts of the world (Jarnevich et al. 2015). Even in the
mountainous tropics, species are more likely to shift upwards than they would in temperate
mountains (Colwell et al. 2008). Marine species, which are understudied, are likely to shift
differently than terrestrial species. Many of these differences are related to species traits
and geography. For instance, marine fish are thought to shift further because of increased
dispersal capability, reduced geographic limitations, and differing human pressures (Lenoir
et al. 2020; MacLean and Beissenger 2017). More broadly, heterogeneous influences of
climate, species interactions, and other potentially range-limiting factors across latitudinal
and climatic zones and land-use gradients suggests that range shifts will vary across space,
underlining the importance of improving the geographic distribution of research.

The cause of this biased geographic distribution of research effort can likely be
linked to a number of factors, including geographical access, local funding sources,
participation in data-sharing networks, and language barriers (Amano et al. 2016; Meyer et
al. 2015). These reasons are likely to explain some of the bias towards Europe and North
America. However, there are also scientific considerations. Because of smaller latitudinal
temperature gradients in the tropics, it can simply be harder to develop a methodology that
documents changes in less clearly defined range edges over large areas. Similarly, it is
easier to make observations of near-shore benthic species than it is to study range shifts in
the open-ocean (dos Santos et al. 2015; Earp and Liconti 2020). The availability of existing data sources, such as long-term fisheries data, can also contribute to geographic biases. This is particularly true given that effectively documenting range shifts requires consistent, long-term data collection (Sarukhán et al. 2015). Even though it might be more challenging to conduct research in the currently understudied areas, it is crucial that scientists expand their scope in order to obtain a representative understanding of how range shifts occur. However, political considerations beyond individual scientist choices are likely also important given the clear differences in effort across some national and subnational boundaries, such as California, Sweden, China, and South Africa, and the fact that long-term monitoring can be linked to government policy (Sarukhán et al. 2015).

**Taxonomic coverage**

Overall, taxonomic coverage was the most balanced. Although there was a gradient in the number of assessed species, that gradient appears to be more representative of taxon size than of taxonomic bias. The only exception was birds and gymnosperms, both of which were studied disproportionately more than the other taxa. This is not entirely surprising. Birds are quite charismatic, receive lots of funding, and have an abundance of data from researchers as well as citizen science programs (Clark and May 2002; dos Santos et al. 2015; McClure et al. 2020). Indeed, several of the largest studies on birds did not gather their own data, but used existing data from programs like the North American Breeding Bird Survey (Currie and Venne 2017; La Sorte and Thompson III 2007). Most of the studied gymnosperm species, on the other hand, were temperate conifers with relatively high economic value. Additionally, the advance of a conifer treeline is potentially easier to
observe and document than far-ranging animal species or plants living in continuous, mixed-species communities.

Although birds and gymnosperms are overrepresented, the general spread of research effort among taxa suggests few clear corrective actions. Rather, researchers need to continue to research a representative range of taxa in order to deepen the field’s understanding of range shifts. Additionally, scientists should ensure that the diversity of species assessed within each taxon continues to grow. A broad range of study taxa is necessary to understand not only changes in a single species, but how those changes might translate into shifts in community composition and structure. However, this conclusion is limited in several ways. First, it is dependent upon the taxonomic groupings chosen for this paper. A primary reason why range shifts should be documented across taxa is due to potential differences in shifts related to species traits, ecological roles, etc. However, those traits and their impacts certainly vary within taxa (Chen et al. 2011). Thus, while taxonomic representativeness was assessed at a certain scale here, the results might be different if assessed at a scale finer than class. Therefore, researchers should attempt to document range shifts of new and diverse species whenever possible. Additionally, this study only includes animals and plants, so the extent of research on the other diverse and important kingdoms is unclear. Due to the difficulty of and lesser interest in studying the distributions of those taxa, they are likely highly underrepresented in the range shift literature.

Within-range coverage

The lack of latitudinal representativeness of species ranges by study areas is not unexpected. The majority of species included in this section (n = 866 species) were only
assessed in a single study. While studies vary in size and can cover anywhere from a single edge of the range or the entire range, a single given study is unlikely to be representative of a species latitudinal distribution. Even in species included in multiple studies, there is the potential for repeated coverage of areas that are more accessible or have lots of existing data (Asher et al. 2011; Grewe et al. 2013; Hickling et al. 2005; MacGregor et al. 2019; Mill et al. 2010).

More surprising is the poor coverage of latitudinal and elevation range edges. Given the high number of species assessed in only a single study, it makes sense that there are only 11 species with both latitudinal edges covered. However, since 65% (n = 212 studies) of studies report a shift on a latitudinal edge, certainly more than 30% of species should have at least one latitudinal edge covered. Elevation is much better represented, but a significant portion of species also have neither elevational extreme covered (n = 171 species). A difference between latitude and elevation results exists even though there is an approximately equal proportion of latitudinal and elevational observations (n = 16,241 and 13,609 observations, respectively), suggesting that elevational observations are better at capturing range edges.

However, there are potential explanations for these results. First, there is the issue of local vs. absolute extremes. Many factors, particularly topographic features, can cause variation in the latitude or elevation of a range edge. The big-eared woodrat (Neotoma macrotis), assessed in two studies in the database, is found in the northern Sierra Nevadas of California, but its poleward edge is elsewhere restricted to lower latitudes by the Central Valley and the flatter areas north of the Southern Coast Ranges. A study examining this species’ poleward edge in western California would be missed by assessing absolute
extremes (Moritz et al. 2008; Rowe et al. 2015). Another potential explanation can be found in range shift documentation methodology. A number of studies in the database attempted to document distribution changes not necessarily on a specific range margin, but as an abundance change (n = 132 studies). A study might assess changes in mean latitude or elevation weighted by abundance in order to track distribution changes without needing to specifically include a range margin. Approximately 29% of observations were abundance rather than occupancy, suggesting that this factor might contribute to the lack of range edge coverage. As an example, one study examined changes in the abundance of 75 plant species within Wisconsin, many of whose ranges extended in all directions beyond the study area (Ash et al. 2017). Additionally, there is the issue of defining a range. Different sources through different time periods can easily give different distributions, which are particularly difficult to define for highly mobile species. For instance, two species distribution sources, BirdLife International and Jetz et al. (2012), provide ranges for the red-crowned crane (Grus japonensis) which vary on the scale of 100 km. While neither source is necessarily incorrect, mismatches in what is considered a range margin could lead to the results obtained here. A buffer was used when calculating if an edge was covered or not to compensate for this issue, but might not be able to completely resolve all mismatches.

Taxonomic group was related to both elevational and latitudinal edge coverage. Amphibians were better represented, as were reptiles in terms of latitude and birds in terms of elevation. It is unclear why taxonomic group might interact with edge coverage. One explanation could simply be that the number of assessed species in each taxa is quite small and so it is difficult to compare between taxa. For instance, with so few species and even fewer studies, a large number of amphibians being assessed in a single study in a
montane region might result in greater elevational coverage than if those species were studied in a flatter region (Davidson et al. 2001; Raxworthy et al. 2008). Additionally, calculated range sizes for each species in this database are about equal among taxa. However, the area of all combined study areas for each included species of amphibian was larger than for those of the other taxa. Thus, amphibians might be better represented simply because a larger portion of their range has been studied.

Overall, coverage of elevational and latitudinal range margins is poor. Although there is much better elevational coverage, the relatively fewer elevational range shift observations suggest that much of that is likely due to the accidental inclusion of high- and low-elevation regions rather than intentional assessment of distributional changes at those margins. For instance, studies utilizing large-scale survey data like that from the continent-wide Breeding Bird Survey or Christmas Bird Count would have a large study area with a wide elevational range, even if the observations of the study itself focused on changes at the latitudinal range edges (Currie and Venne 2017; La Sorte and Thompson III 2007). So, several recommendations can be made from these results. First, researchers need to ensure that latitudinal extremes of ranges are better included in range shift observation sites, whether the species is shifting elevationally or latitudinally. Secondly, while researchers are better at including large elevational gradients even in studies assessing latitudinal shifts, there is not a correspondingly high number of observed elevational shifts. Taking advantage of the elevational extremes at which data already exists would be an easy way to improve coverage. Lastly, given that low-latitude and low-elevation range margins are more frequently studied, an emphasis should be put on documenting range shifts at the high-latitude and high-elevation margins. This is especially important
given the traditional biogeographical hypothesis that climate is a greater determinant of species distribution along poleward and upper range margins, while biotic factors play a more important role along the trailing edges (Paquette and Hargreaves 2021; Louthan et al. 2015).

Partially in an effort to avoid the above flaws of comparing absolute extremes of elevation and latitude, temperature was used for an additional comparison. Among a number of climatic variables and other factors, temperature is considered to be a primary factor in determining species distributions (Jarnevich et al. 2015). So, temperature should act as a proxy for elevation and latitude, where areas of minimum temperature tend to represent low-elevation and low-latitude regions, and the inverse for areas of maximum temperature. Additionally, because range shifts can occur variably over temperature gradients, the temperature distribution of study areas should ideally be representative of that of the overall species range (Colwell et al. 2008; Coristine and Kerr 2015; Paquette and Hargreaves 2021; Sunday et al. 2012).

The mean temperature of study areas was not representative of the overall range of most species, even if it is more so than latitude is. This is, again, not surprising given the small number of studies per species. More concerning, coverage of temperature extremes is lower than combined latitude and elevation coverage. 662 species had no covered temperature extreme while only 133 species had not a single elevation or latitude edge covered, suggesting that temperature is not an effective proxy for range edge in this context (Figure 8, Appendix F). Instead, the role of other factors and climatic variables, particularly annual minimum and maximum temperature, might be strongly influential in determining species distributions (Fei et al. 2017; Jarnevich et al. 2015). Additionally, temperature might
have a scale issue. For instance, high-elevation areas that are much smaller than a single grid cell and as cold as any other part of the range, like a mountain peak, might be effectively erased when calculating temperature means in each cell (Berton et al. 2012; Chen et al. 2011; Forero-Medina et al. 2011). Studies with certain methodologies, like the abundance studies discussed earlier, could also contribute to these results.

Nonetheless, it is clear that, for most of the species that have any documented range shifts, the geographic locations of study areas ineffectively capture extremes of elevation, latitude, and temperature. Extremes or edges of a range are particularly important to study. Changes of range margins are tightly linked to significant community-level changes and are crucial to determining the success of a species in a dramatically altered climate (Midgley et al. 2002). Additionally, shifts are potentially more prominent along range margins than inside the range, so an abundance of studies conducted in the middle of the range might result in range shift underestimations (Bates et al. 2014, Currie and Venne 2017). While an ideally studied species would therefore have all range edges covered and have study areas representative of the overall range, only 0.34% (n = 4 species) of species met these conditions. Given the variability in how range shifts occur throughout a species range, it is imperative that researchers improve the distribution of range shift observations within species ranges in order to better capture changes across a variety of conditions.

Limitations

In addition to the specific caveats discussed above, there are some potential limitations to note. For one, papers documenting shifts above the species-level were excluded from the database. Similarly, papers examining taxa besides plants and animals
were also excluded. However, there is no reason to expect that including other taxa or shifts on a different scale would significantly change the findings of this paper. Of potentially greater impact is the decision to include only articles written in English (Amano et al. 2016). While necessary due to the language skills of those compiling the database, including range shifts documented in other languages might improve the geographic distribution of research outside of North America and Europe. In addition, average annual temperature is only one climate variable used to predict species distributions. Indeed, this single variable can hide changes in other variables like temperature seasonality, maximum annual temperature, and minimum annual temperature, some of which might be more important to range shifts. It would be worthwhile to expand this analysis to include other important bioclimatic variables and thus obtain a broader understanding of how research might be biased within the range of the species they study. Lastly, this study ignores differences in the characteristics of the observations themselves. An assessment of methodological differences, change in coverage throughout time, etc., would be a worthwhile endeavor.

**Conclusion**

A review of published range shift observations reveals the existence of significant geographic biases. Research effort remains highly concentrated in terrestrial, mid-latitude Europe and North America. Additionally, research effort is unequally distributed within species ranges, as study areas are unrepresentative of species ranges and fail to include range extremes, particularly latitudinal extremes. Assessed species, on the other hand, are
taxonomically representative among plants and animals. However, because characteristics of range shifts vary geographically, it is essential that researchers improve geographic coverage. By being more strategic about the placement of study areas throughout the world and throughout the range of studied species, researchers can better contribute to representative literature. More representative literature will, in turn, increase the accuracy with which range shifts can be modeled and improve the ability of managers to protect biodiversity and ecosystem services in the face of a significantly changing climate.
References


Amano, T., González-Varo, J.P., Sutherland, W.J., 2016. Languages Are Still a Major Barrier to Global Science. PLOS Biology 14, e2000933.


Conference for Young Marine Researchers in Oldenburg, Germany. Springer International Publishing, Cham, pp. 1–19.


González-Orozco, C.E., Pollock, L.J., Thornhill, A.H., Mishler, B.D., Knerr, N., Laffan, S.W.,


Li, F., Park, Y.-S., 2020. Habitat availability and environmental preference drive species
range shifts in concordance with climate change. Diversity and Distributions 26, 1343–1356.


information basis of biodiversity distributions. Nature Communications 6, 8221.
Species’ traits influenced their response to recent climate change. Nature Climate Change 7, 205–208.


Appendices
Appendix A: Geographic Distribution of Assessed Marine Species. Each cell is colored to represent the log-transformed number of marine species assessed within it. The plot along the left indicates the relative number of species assessed within each 10 degree latitudinal band. This map is a larger version of Figure 2A.
Appendix B: Geographic Distribution of Assessed Terrestrial Species. Each cell is colored to represent the log-transformed number of terrestrial species assessed within it. The plot along the left indicates the relative number of species assessed within each 10 degree latitudinal band. This map is a larger version of Figure 2B.
Appendix C: Geographic Distribution of Marine Studies. Each cell is colored to represent the log-transformed number of marine studies overlapping with the cell. The plot along the left indicates the relative number of studies overlapping with each 10 degree latitudinal band. This map is a larger version of Figure 2C.
Appendix D: Geographic Distribution of Terrestrial Studies. Each cell is colored to represent the log-transformed number of terrestrial studies overlapping with the cell. The plot along the left indicates the relative number of studies overlapping with each 10 degree latitudinal band. This map is a larger version of Figure 2D.
Appendix E: Table of Taxonomic Coverage. Assessed species are those with at least one recorded range shift observation. The total number of described species in each taxon was obtained from Catalog of Life. Relative number of assessed species was calculated by dividing assessed species by described species. Asterisks * indicate outliers. These data were used to create Figure 3.

<table>
<thead>
<tr>
<th></th>
<th>Assessed Species</th>
<th>Described Species</th>
<th>Relative Number of Assessed Species (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-vascular land plants</td>
<td>87</td>
<td>23,486</td>
<td>0.370</td>
</tr>
<tr>
<td>Ferns and lycophytes</td>
<td>143</td>
<td>12,232</td>
<td>1.169</td>
</tr>
<tr>
<td>Gymnosperms</td>
<td>93</td>
<td>1045</td>
<td>8.900*</td>
</tr>
<tr>
<td>Angiosperms</td>
<td>5345*</td>
<td>323,603</td>
<td>1.652</td>
</tr>
<tr>
<td>Insects</td>
<td>3076</td>
<td>1,053,578</td>
<td>0.292</td>
</tr>
<tr>
<td>Fishes</td>
<td>474</td>
<td>33,803</td>
<td>1.402</td>
</tr>
<tr>
<td>Amphibians</td>
<td>155</td>
<td>8054</td>
<td>1.925</td>
</tr>
<tr>
<td>Reptiles</td>
<td>67</td>
<td>6681</td>
<td>1.003</td>
</tr>
<tr>
<td>Birds</td>
<td>1612</td>
<td>10,521</td>
<td>15.322*</td>
</tr>
<tr>
<td>Mammals</td>
<td>111</td>
<td>5940</td>
<td>1.869</td>
</tr>
<tr>
<td>Other</td>
<td>554</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Appendix F: Species Grouped by Number of Covered Edges and Latitudinal Representativeness. Each species was grouped in the x-axis by the sum of all range margins (elevational leading, elevational trailing, latitudinal leading, and latitudinal trailing) that were covered, where a value of 0 indicates that species had no covered edges and a value of 4 indicates all edges were covered. Each species was also grouped in the y-axis by the results of a t-test on the latitudes of its study area(s) and range. If p > 0.05, the study areas were considered representative. If p < 0.05, the study areas were considered unrepresentative.

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representative</td>
<td>15</td>
<td>34</td>
<td>16</td>
<td>10</td>
<td>4</td>
<td>79</td>
</tr>
<tr>
<td>Unrepresentative</td>
<td>118</td>
<td>301</td>
<td>459</td>
<td>219</td>
<td>2</td>
<td>1099</td>
</tr>
<tr>
<td>Total</td>
<td>133</td>
<td>335</td>
<td>475</td>
<td>229</td>
<td>6</td>
<td>1178</td>
</tr>
</tbody>
</table>
Appendix G: Species Latitudinal Representativeness. Each species was grouped by the results of a t-test on the latitudes of its study area(s) and range. If $p > 0.05$, the study areas were considered representative. If $p < 0.05$, the study areas were considered unrepresentative. These data were used to create Figure 4.

<table>
<thead>
<tr>
<th></th>
<th>Amphibian</th>
<th>Bird</th>
<th>Mammal</th>
<th>Reptile</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representative</td>
<td>14</td>
<td>55</td>
<td>9</td>
<td>9</td>
<td>87</td>
</tr>
<tr>
<td>Unrepresentative</td>
<td>116</td>
<td>987</td>
<td>84</td>
<td>43</td>
<td>1230</td>
</tr>
</tbody>
</table>

Appendix H: Species Latitudinal Range Edge Coverage. Each species was grouped by the amount of coverage of its latitudinal range margins by its study area(s). These data were used to make Figure 5.

<table>
<thead>
<tr>
<th></th>
<th>Amphibian</th>
<th>Bird</th>
<th>Mammal</th>
<th>Reptile</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neither</td>
<td>68</td>
<td>760</td>
<td>69</td>
<td>26</td>
<td>923</td>
</tr>
<tr>
<td>Trailing</td>
<td>48</td>
<td>150</td>
<td>14</td>
<td>11</td>
<td>223</td>
</tr>
<tr>
<td>Leading</td>
<td>13</td>
<td>128</td>
<td>10</td>
<td>11</td>
<td>162</td>
</tr>
<tr>
<td>Both</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>4</td>
<td>11</td>
</tr>
</tbody>
</table>

Appendix I: Species Elevational Range Edge Coverage. Each species was grouped by the coverage of its elevation range margins by its study area(s). These data were used to make Figure 6.

<table>
<thead>
<tr>
<th></th>
<th>Amphibian</th>
<th>Bird</th>
<th>Mammal</th>
<th>Reptile</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neither</td>
<td>9</td>
<td>127</td>
<td>23</td>
<td>12</td>
<td>171</td>
</tr>
<tr>
<td>Trailing</td>
<td>18</td>
<td>224</td>
<td>25</td>
<td>7</td>
<td>274</td>
</tr>
<tr>
<td>Leading</td>
<td>16</td>
<td>93</td>
<td>5</td>
<td>4</td>
<td>118</td>
</tr>
<tr>
<td>Both</td>
<td>88</td>
<td>504</td>
<td>8</td>
<td>17</td>
<td>617</td>
</tr>
</tbody>
</table>
Appendix J: Species Temperature Representativeness. Each species was grouped by the results of a t-test on the annual mean temperature (bio1) of its study area(s) and range. If p > 0.05, the study areas were considered representative. If p < 0.05, the study areas were considered unrepresentative. These data were used to create Figure 7.

<table>
<thead>
<tr>
<th></th>
<th>Amphibian</th>
<th>Bird</th>
<th>Mammal</th>
<th>Reptile</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Representative</td>
<td>33</td>
<td>194</td>
<td>22</td>
<td>17</td>
<td>266</td>
</tr>
<tr>
<td>Unrepresentative</td>
<td>99</td>
<td>848</td>
<td>71</td>
<td>35</td>
<td>1053</td>
</tr>
</tbody>
</table>

Appendix K: Species Range Temperature Extreme Coverage. Each species was grouped by the coverage of its range temperature extremes by its study area(s). These data were used to make Figure 8.

<table>
<thead>
<tr>
<th></th>
<th>Amphibian</th>
<th>Bird</th>
<th>Mammal</th>
<th>Reptile</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neither</td>
<td>15</td>
<td>574</td>
<td>56</td>
<td>17</td>
<td>662</td>
</tr>
<tr>
<td>Minimum</td>
<td>44</td>
<td>352</td>
<td>22</td>
<td>13</td>
<td>431</td>
</tr>
<tr>
<td>Maximum</td>
<td>5</td>
<td>32</td>
<td>7</td>
<td>13</td>
<td>57</td>
</tr>
<tr>
<td>Both</td>
<td>68</td>
<td>84</td>
<td>8</td>
<td>9</td>
<td>169</td>
</tr>
</tbody>
</table>