***Target Journal:***Global Change Biology

***Recommended Reviewers:***

* Derek Sikes, [dssikes@alaska.edu](mailto:dssikes@alaska.edu)
* Toke Høye, [tth@ecos.au.dk](mailto:tth@ecos.au.dk)
* Maria Hällfors, maria.hallfors@syke.fi

**RISING MINIMUM TEMPERATURES CONTRIBUTE TO 50 YEARS OF SHIFTING ARCTIC AND BOREAL BUTTERFLY COMMUNITIES IN NORTH AMERICA**

Vaughn Shirey1\*, Naresh Neupane1, Robert Guralnick2,andLeslie Ries1

1 Department of Biology, Georgetown University; Washington, DC; United States of America

2 Florida Museum of Natural History – University of Florida; Gainesville, FL; United States of America

\* Corresponding author; +1 (484)-945-3100, vms55@georgetown.edu

**Abstract:** Global climate change has been identified as a major driver of observed insect declines, yet in many regions there are critical knowledge gaps for how communities are responding to climate. Poleward regions are of particular interest because warming is most rapid while biodiversity data are most sparse. Building on recent advances in occupancy modeling of presence-only data, we reconstructed 50 years (1970-2019) of butterfly population trends in response to rising minimum temperatures in one of the most under sampled regions of the continent. Among 90 modeled species, we found that cold-adapted species are far more often in decline compared to their warm-adapted, more southerly distributed counterparts. Further, in a post-hoc analysis using species’ traits, we find that species’ range-wide average annual temperature and wingspan are a consistent predictor of occupancy changes. Species with warmer ranges and larger wingspans were most likely to be increasing in occupancy. Our results provide the first look at macroscale butterfly biodiversity shifts in a critically under sampled region of North America. Further, these results highlight the potential of leveraging the wealth of presence only data, the most abundant source of historical insect biodiversity. New approaches to the modeling of presence only data will match recent increases in community science participation with sparse historical records to reconstruct trends even in poorly sampled regions.

**Keywords:** arctic, biodiversity, boreal, butterflies, climate change, insect decline, Lepidoptera

**Authorship Contributions:** VS came up with the study idea and design. NN obtained and processed climate data. VS conducted the analysis and original interpretation. RG, NN, and LR provided feedback on the analysis. VS wrote the original manuscript, and all authors edited the manuscript together.

**Funding:** VS was funded by the Georgetown University Department of Biology, Georgetown Graduate Student Government STEM for the Public Good Award, a National Science Foundation Graduate Research Fellowship (#1937959). Additional funding came from a Global Biodiversity Information Facility Young Researchers Award.

**Acknowledgments:** Firstly, we want to thank all of the museum staff, volunteers, and community scientists who aided in the mobilization of biodiversity data used in this study. We would also like to thank Laura Melissa Guzman, Rassim Khelifa, Leithen M’Gonigle, Sarah Johnson, Hanna Jackson, and Elijah Reyes for insightful conversations and collaboration on the use of occupancy models with presence-only data. Martha Weiss, Gina Wimp, Greg Breed and other members of the Ries Lab provided feedback on the original manuscript. Finally, model development, testing, and implementation were performed on the Georgetown High-performance Computing Cluster (aided by Woonki Chung), Georgetown Massive Data Institute (Lisa Singh), and the University of Florida HiperGator High-performance Computing System.

**1. INTRODUCTION**

Multiple lines of scientific evidence have pointed to the potential for an alarming recent decline in insect biodiversity across the planet (Cardoso *et al.* 2020; Wagner *et al.* 2021a, b). Major changes in the abundance and composition of insect populations increase the risk of losing vital insect-mediated ecosystem services, including pollination, decomposition, and pest control (Kremen *et al.* 2007, Zhou *et al.* 2023), and even the potential for collapse of some ecological networks (Dunne *et al.* 2002; Memmott *et al.* 2004; Grames *et al.* 2023). Although the list of potential agents driving insect decline is multifaceted, climate change has been put forth as a major direct contributor but also as a force that interacts with other global drivers [(Wagner *et al.* 2021b)](https://www.zotero.org/google-docs/?8BW3yT). As ectotherms, insects are especially sensitive to local scale and macroscale climatic conditions in ways that may integrate to signals of change at macro-scales (Deutsch *et al.* 2008). Among insects, butterflies have by far the densest distribution and natural history data due to their high public profile and apparency in nature. Numerous foundational studies in global change biology have used butterflies to better understand how life history, ecological traits, and species interactions may drive responses to novel climatic conditions [(Parmesan *et al.* 1999; Pöyry *et al.* 2009; Breed *et al.* 2013)](https://www.zotero.org/google-docs/?vZFL5N) .

Surface-level temperatures have risen more drastically in high-latitude regions than anywhere else (Holland & Bitz 2003), a trend that is expected to continue under nearly all climate change forecasts (Ono *et al.* 2022). Under current, “business-as-usual” carbon-dioxide emission scenarios, surface temperatures are projected to increase in the Arctic by an average of 10° Celsius through 2100, roughly four times greater than other global regions (You *et al.* 2021). Further, this will likely result in rainier rather than snowy winters across polar region (Bintanja & Andry 2017; McCrystall *et al.* 2021). For butterflies in high-latitude regions to persist under such drastically shifting climatic conditions, they must be able to either move to where conditions are favorable, adapt to novel climates *in situ* or face significant risk of population declines (Corlett & Westcott 2013; Kellermann & van Heerwaarden 2019). The rapid climatic changes in this region and the potential for significant barriers to movement (e.g., closed-canopy habitats) may contribute to a high degree of “climate debt” for butterflies (Lewthwaite *et al.* 2018), in other words, a lag between the pace of climate change and observed species responses.

Despite the magnitude of change happening in high-latitude regions, North American boreal and Arctic regions remain critically under sampled for butterflies, and are even more biased that in many other regions due to high concentration of human populations in the south [(Shirey *et al.* 2021)](https://www.zotero.org/google-docs/?9tXWmg). These sampling gaps are major impediments to conservation, especially for invertebrates [(Cardoso *et al.* 2011)](https://www.zotero.org/google-docs/?CUS0Ah) . Historically, data that do exist in this region come largely from natural history museum collections while more contemporary data largely arise from community science platforms, such as records from online platforms like iNaturalist and are almost all opportunistic or “presence-only” data (hereafter, presence-only). Presence-only data rarely capture all species in the broader community (Kelling *et al.* 2019). For example, collectors/observers may more often record species of particular interest or those that are more detectable. Cryptic, small-bodied species may go under-reported (Meyer *et al.* 2016; Adamo *et al.* 2021). Further, there is no guarantee of resampling at the same site over time. Finally, accounting for data biases is particularly challenging because absences or non-detections go unrecorded, and effort is unknown. Thus, these data have long been unsuitable for typical statistical models, which are largely rooted in generalized linear modeling paradigms that have demanded abundance or presence-absence data. To confront the conundrum that the most highly available distribution data could not be used in typical statistical models, a new class of species distribution models, including popular software like MaxENT, which generate “pseudoabsence” data based on complex algorithms (), have been widely used and have helped expand our knowledge about large-scale distributions (Elith & Leathwick 2009); however, some studies have found these methods to be unreliable (Yackulic *et al.* 2013).

The issue of using presence-only data in large scale ecological analysis has resulted in many conversations about the most appropriate models and also best practices for the use of these data. These debates span multiple metrics of biodiversity including when or how presence-only data is sufficient to estimate abundances (Ries *et al.* 2019; Wepprich 2019), phenological patterns (Larsen & Shirey 2021), and range dynamics (Yackulic *et al.* 2013; Ascher *et al.* 2020; Guzman *et al.* 2021). For example, Ries *et al.* 2019 and Wepprich 2019 found that approaches used to estimate abundance declines in the Monarch butterfly, *Danaus plexippus* (Lepidoptera: Nymphalidae), were severely biased in an effort to track abundances over 100 years (Ries *et al.* 2019; Wepprich 2019). In addition, Larsen & Shirey (2021) found that lack of curation and inappropriate modeling of presence-only data to in a study to infer phenometrics resulted in inferences that clearly did not make ecological sense; specifically the study found that butterfly populations in high latitudes having similar emergence timing as populations at lower latitudes (Larsen & Shirey 2021). Finally, the strategy of imputing non-detections, censoring inference to species ranges or to regions where they could plausibly occur, and not accounting for heterogeneous detection probability within an occupancy modeling framework can produce biased estimates of occupancy trends with a particular tendency to find declines that are likely artefacts (Ascher *et al.* 2020; Guzman *et al.* 2021). Underscoring all of these examples is the pressing need for approaches to produce reliable trend estimates from presence-only data, by far the most abundant data source, both in the past and likely into the future.

New advances in statistical modeling to more robustly address these biases have the potential to unlock the wealth of presence-only for inferring population trends, even at continental extents. For example, occupancy-detection models, (MacKenzie *et al.* 2002; Kéry & Royle 2015) have long been used to reconstruct ecological signals over space and time from presence-absence, more accurately (and hereafter, detection/non-detection), where sites are repeatedly visited and the patterns in a series of detections and non-detections across all visits are used to more robustly model occupancy patterns by disentangling the observation process from true ecological patterns. Yet occupancy models demand non-detection data (specifically, zeros for all species in the community when not observed) and so their use with presence-only data has traditionally been considered inappropriate. Recently, proposals for the use of occupancy models with presence-only data center on appropriate inferences and so explicit additions of zeros in the data set by leveraging records of other species has been used as a proxy for effort (Guzman *et al.* 2021; Jackson *et al.* 2022) but this too has been controversial (). A recent simulation study has confirmed that, when multiple species records in the same location are used as a proxy for community sampling, the inference of zeros in an occupancy modeling framework provides robust ecological signal from presence-only data (Shirey *et al.* 2022). This approach is increasing in popularity and has already been used to reconstruct sensible ecological trends from presence-only data using presence-only and other non-standardized datasets (van Strien *et al.* 2013; Jönsson *et al.* 2021; Engelhardt *et al.* 2022; Jackson *et al.* 2022).

Here, we used an occupancy-detection approach to account for a complex and undocumented detection process to reconstruct patterns of change in North America above 45°N latitude over a 50-year period (1970-2019). Specifically, we focused on the influence of changes in minimum temperature and, separately, precipitation (a highly correlated variable) on species-specific occupancy patterns through space and time. A rich body of literature from Europe (Parmesan *et al.* 1999; Hill *et al.* 2002; Pöyry *et al.* 2009) where structured surveys, where effort is known and all individuals observed are recorded, are available served as a guidepost for our *a priori* hypotheses of general expectations in these polar and sub-polar climates. Following observed trends, we hypothesized that warm-associated species are likely faring, on average, better than their cold-associated counterparts as warmer winters support proliferation of species at their northern range-edges but are less likely to show declines in the south because the intensity of warming is much lower (). In contrast, cold-associated species may be pushed to the northernmost limits of their range and exhibit declines in occupancy probability, either through maladaptation to warmer climates or being pushed out by an influx of better adapted species, although specific mechanisms are difficult to establish with observational data (Parmesan *et al.* 1999; Hill *et al.* 2002; Pöyry *et al.* 2009; Heikkinen *et al.* 2010; Hällfors *et al.* 2021, 2023). In addition to the importance of understanding trends in poleward latitudes, as the first empirical implementation of this approach, there is the benefit that we are estimating spatial patterns which have been not only been demonstrated in other regions, but also make ecological sense. This is useful because in regions with presence-only data, there is rarely an opportunity to validate results with independent data sets. In this work, we specifically aimed to use sparse, presence-only data on butterflies in North America above 45°N to:

1. Reconstruct historical occupancy trends for butterflies over 50 years including species-specific responses to climate change.
2. Split these trends into three zones: southern, mid-latitude, and northern components of each species range above 45°N latitude to assess subrange trends.
3. Relate overall trends in species occupancy to species-specific ecological traits including range size, range-wide temperature, wingspan, overwintering stage, and host plant family breadth in order to make identify species most likely to experience declines/increases in occupancy using a model selection paradigm (Candidate models are shown in Table 1)

**2. METHODOLOGY**

Occupancy-detection models are designed to disentangle the process of observation from the underlying ecology (MacKenzie *et al.* 2002; van Strien *et al.* 2013; Kéry & Royle 2015). Further, multi-species occupancy models can be used to leverage community patterns to strengthen the inferences for multiple species at the same time (Dorazio & Royle 2005; Zipkin *et al.* 2010). In a recent simulation study, these models have demonstrated the ability to accurately reconstruct ecological trends from presence-only datasets that meet certain pre-qualifying conditions indicative of sampling multiple species at one location (Shirey *et al.* 2022). A graphical representation of our workflow with details about this implementation of our model, including methods for inferring zeros is illustrated in Supplemental Figure 1.

To implement occupancy-detection models for this study, we established grids of 100×100 and 200×200-kilometer square cells across our study region to determine if the spatial grain of our study impacted our results. We set our analysis to the 50 years between 1970 and 2019 and divided this period into ten, five-year-long intervals. Previous simulation results found that five temporal bins were found to provide sufficient granularity to detect trends, but ten were best able to reconstruct simulated trends and also provide more nuanced patterns through time. In contrast, using only two time periods, a common method in these studies, often produced poor results (Shirey et al., 2022). In each of these periods we estimated the probability that a given species is an occupant of a specific grid cell. These “occupancy intervals” were further broken down into five one-year-long “visit intervals” to provide a basis for separate estimation of detection and occupancy, which requires tracking multiple visits as repeated trials within the larger occupancy interval (MacKenzie *et al.* 2002).

***2.2 Climate Data***

We used minimum temperatures and precipitation data to inform the ecological process of our occupancy-detection model. Land use is also a strong contributor to the occupancy of butterflies across our study region, but to our knowledge, no comprehensive dataset of land cover exists for high-latitude regions of North America extending back into the 1970s or from even more recent time periods that would allow for our 5-year temporal grain. We also did not include pesticide data in our analysis as these data are difficult to come by and to aggregate at the spatial grain of our analysis. Thus, we opted to inform our models solely with climatic predictors. Monthly minimum temperatures from 1970 to 2015 were extracted from the National Oceanic and Atmospheric Administration's (NOAA) 20th-century reanalysis, 20CRv3 (Slivinski *et al.* 2019). These data are available at 0.1-degree spatial and monthly temporal resolution from 1836 to 2015; reanalysis products (where observation data are synthesized with mechanistic climate models are not yet available for 2016 onward, variables were obtained from the Climate Prediction Center (CPC, which are available at 0.5-degree spatial and monthly temporal resolutions) (Climate Prediction Center 2022). Minimum average temperatures increased over the vast majority of our study range, up to 4° Celsius over the 50-year study period (Figure 1a). For precipitation, monthly precipitation data were extracted from the Climate Research Unit (CRU) TS4.04 global precipitation dataset which are also available at 0.5-degree spatial resolution, but only back to 1900 (Harris *et al.* 2020a, b). Trends in precipitation were generally more heterogenous but reflect, on average, increasing rainfall across the region.

Climate raster data reflecting monthly average minimum temperature and monthly precipitation for the years 1970 through 2019 were read into R using the package “raster” (Hijmans *et al.* 2015). All raster data were reprojected using this same package to the project coordinate reference system (North America Albers Equal Area Conic) and this was also used as the base projection for all distribution data in this analysis. The raster data were summarized within each 100×100-kilometer (or 200×200-kilometer) grid cell using the mean value across for each five-year occupancy interval (e.g., 1970-1974) weighted by their coverage of the cell. Finally, exploration of our climate data revealed that average minimum temperature and average precipitation were highly correlated (Pearson’s r = 0.67 for the 100×100-kilometer scale and Pearson’s r = 0.73 for the 200×200-kilometer scale), thus we opted to run our models separately for both climate covariates (results for temperature are shown in the main text and precipitation in the Supplementary material). We opted to focus on temperature over precipitation as prior research has demonstrated temperature to be more important for range dynamics in butterflies (Keret *et al.* 2020).

***2.1 Species Range, Trait, Phylogeny, and Occurrence Data***

We used species range boundaries to constrain model inference to historical regions applying a conservative 100-kilometer buffer to allow for range shifts because past shifts in poleward butterflies fell into distances from X-X (ref). If any portion of the buffer intersected with a grid cell, that cell was assumed to be potentially occupied by that species. Prior work has shown that failure to censor locations where a species could not occur can produce biased and/or misleading estimates of occupancy (Guzman *et al.* 2021; Shirey *et al.* 2022).

We used the same range maps to calculate range size and range-wide climatic metrics which served as a basis for species range-related traits, including total range size through central America and average, range-wide temperature and precipitation across all of North America. Species range map data were derived from published field guides on North American butterfly species, including The Kaufman Field Guide to Butterflies of North America (Brock & Kaufman 2006) and A Swift Guide to Butterflies of Mexico and Central America (Glassberg 2018). The range maps were digitized as part of work by Earl et al. (2021) and were re-used here. Range-wide average annual temperature and precipitation traits were extracted from each species’ range averaging conditions over each polygon using the WorldClim2/BioClim dataset (representing average climatic conditions from 1970-2000) [(Fick & Hijmans, 2017)](https://www.zotero.org/google-docs/?o92OvI). Along with this derived climatic trait data, we aggregated other ecological traits including wingspan, host plant family breadth, and disturbance tolerance from the global LepTraits database (Shirey et al., 2022) to test how these traits are associated with observed population trajectories. Thus, the final dataset of butterfly traits consisted of average annual range-wide temperatures and precipitation, geographic range size, disturbance tolerance, host plant family breadth, average wingspan, and overwintering stage. Finally, we obtained a recently produced phylogeny (Earl *et al.* 2021) of North American butterfly species as a final data product to be used in our *post hoc* analysis of occupancy trends. We used this phylogeny to test for phylogenetic signal in our modeled occupancy results.

Species occurrence data for all butterflies above 45°N from 1970 to 2019 were obtained from the Global Biodiversity Information Facility (GBIF)(GBIF.org 2022), Integrated Digitized Biocollections (iDigBio)(see Supplemental File S1 for a list of accessed collections), and the Symbiota Collections of Arthropods Network (SCAN) (Heinrich *et al.* 2015)(see Figure 1b, c for a summary of these data). Duplicate records for the same specimen were filtered out. We removed records with exceptionally low coordinate precision (greater than 25 kilometers of a reported uncertainty radius). We reconciled the taxonomic names using the R package “taxize” (Scott Chamberlain & Eduard Szocs 2013) and, for non-matches, we resolved the taxonomy by hand where tractable (e.g. by matching unresolved synonymies to currently accepted names). We used the Lamas taxonomy as a backbone (Lamas 2015); because of the taxonomic *Celestrina* complex (Wright & Pavulaan 1999; Pavulaan & Wright 2000, 2005), we excluded this group from our analysis since we could not make redetermination for all occurrence records. We then limited our analysis to include only species with at least 500 reported occurrences across the spatiotemporal scope of our study. This resulted in the retention of 33.7% of the species in our dataset (from 315 unique species names to 106 names). Finally, we removed all species from the analysis that are known migrants (*i.e.,* are not historically known to overwinter anywhere north of 45° in North America).In the iterative process of modeling, we noted several species for which the MCMC chains experienced notable convergence issues and removed these species from the analysis as well.

***2.3 Inferring Detection from Presence-only Observations***

Occupancy-detection models use detection/non-detection data to reconstruct occupancy trends (MacKenzie *et al.* 2002; Kéry & Royle 2015). Our simulation study showed that if at least 50% of all spatiotemporal sampling bins contain community data, then the data set is sufficient to go forward with model implementation. If not, our simulation results suggest data are too sparse or not collected/observed under an appropriate methodology for this approach (Shirey et al., 2022) To proxy this probability that field observations in our dataset were community-focused in our dataset, we spatiotemporally aggregated all of our occurrence data at the point level. The occurrence records must have been collected during the same year and must also have been collected/observed within 5-kilometers of one another to be included in an aggregate cluster. We then used the number of species in these spatiotemporal aggregations as a means to classify clusters. “Community clusters” were point aggregate clusters where more than one distinct species was reported; “singleton clusters” were point clusters or single points where only one species was reported. We took the percentage of points that fell within community clusters compared to the filtered occurrence dataset as the proxy of community-focused visitation probability. This percentage was 69.64% and which was above a 50% found to be a sufficient cutoff in simulation studies (Shirey et al., 2021).

Our approach for imputing non-detection is detailed in Supplemental Figure S1 from Shirey et al. (2022) and we provide a more brief description here. , For the purposes of this study, we assume that the detection of at least two species within a single, one-year visit interval in each study cell (e.g., two different species recorded in 1970, the first visit interval of five in the occupancy interval 1970-1974) is sufficient to impute non-detection data for all other species that could potentially occur at that grid cell (*i.e.,* their range plus a 100-kilometer buffer that intersects the cell). Although a relatively low bar for imputing zeros into a 5-year occupancy interval cell, simulations have shown that such an approach generates results that match simulated trends and so can be thought of as a minimum requirement, useful when considering regions with particularly sparse data (Shirey *et al.* 2022).

***2.4 Occupancy-detection Modeling***

For the occupancy subcomponent of our model, we included two environmental predictors of occupancy, average minimum temperature (or average precipitation) across a 5-year-long period (e.g., 1970-1974, as species-specific slopes), and the terrestrial surface area of the grid cell (to account for cells along coastlines or with other large bodies of water). We modeled the butterflies as a single community, and we examined species-specific responses by including species-specific intercept in our model to account for differences in baseline occupancy among species. Thus, the occupancy component of our model is:

where is the probability that a given species, , is an occupant of cell in a 5-year occupancy interval [](https://www.codecogs.com/eqnedit.php?latex=t#0); is the mean occupancy probability for butterflies (on a linear scale); is a species-specific intercept; is the effect of terrestrial surface area on occupancy probability; and is the species-specific effect (slope) of minimum temperature or precipitation on occupancy. The parameter is only estimated in the temperature model to estimate the potential quadratic effect of minimum temperature on occupancy probability. We used uninformative priors for all parameters, thus assuming values were to be drawn from normal distributions with a mean of zero and a variance that was estimated at the parameter level from a uniform distribution.

The detection component of our model was informed by a random-effects structure which included an average intercept value estimated across all grid cell, occupancy interval, and species combinations, a fixed-effect of occupancy interval on detection (to account for potentially increasing detectability due to modern survey methods and digital platforms such as iNaturalist), and two random intercepts that varies by (a) species and (b) cell by occupancy interval. Mathematically, our detection component is defined as:

where is the detection probability of a given species, , during a given occupancy interval, , at a particular cell, ; the parameter is the mean detection probability for all butterflies (on a linear scale); is a random species-specific intercept; is a cell by occupancy interval random intercept; and is the effect of occupancy interval on detection probability. For all parameters, we assumed the values were to be drawn from normal distributions with a mean of zero and a variance that was estimated at the parameter level from a uniform distribution. Finally, the likelihood in our model was defined by the product of the latent occupancy state (either present or absent per a Bernoulli draw of the calculated occupancy probability ) and the probability of detection, .

We ran our occupancy models using JAGS (Plummer 2003) on each detection/non-detection dataset for 150,000 iterations, 50,000 of which were discarded as “burn-in,” retaining the samples of every 100 iterations across four chains for a total of 4,000 samples from the posterior distribution. We assessed convergence across these chains by examining both Gelman-Rubin diagnostic values (Gelman & Rubin 1992) (using 1.1 as an upper threshold), and by visually inspecting the trace plots for all parameters (provided in the Supplementary Material). We used the R packages “jagsUI” (Kellner 2021) and “MCMCvis” (Youngflesh 2018) to complete the majority of this work. Visualizations of model performance metrics are included in the Supplemental Material.

***2.5 Post-hoc Trait Analysis***

Species’ traits have been demonstrated to be associated with changes in species’ ranges over time and in the relative risk of decline (Keinath *et al.* 2017) or extinction (Fagan *et al.* 2001; Chichorro *et al.* 2019, 2022). We used our chosen traits (range size, range-wide mean temperature, wingspan, host plant family breadth, overwintering life-stage, and disturbance tolerance to examine if any were associated with mean expected mean occupancy shifts from the 1970s to 2010s. Using a model-selection approach, we specified 14 models based on *a priori* hypotheses about which traits might be most important when predicting occupancy trends across our study. Our first model, Model A, or our ecological “null” model, is an intercept-only model that assumes occupancy trend is best predicted by a singular mean value across all species. In other words, this model assumes that no trait is better able to account for differences in species-specific occupancy trends. To this model, we added species-specific intercept terms that are correlated by a variance-covariance matrix from our phylogeny (Model B, the “null + phylogenetic intercept” model). The additional models (Models C-N) follow a similar structure but in these cases, we selected trait(s) that we believe may be especially informative for predicting occupancy trend and implemented each combination both with or without phylogenetic intercept terms. Table 1 describes each of these models and provides relevant background for the selection of trait information.

After running our models, we then compared them to each other using leave-one-out (LOO) cross-validation in order to select the top candidate model (Vehtari *et al.* 2017). From our LOO scores, we calculated the expected log pointwise predictive density (ELPD-LOO) or a measure of predictive capacity(Vehtari *et al.* 2017). Higher values of this metric indicate better predictive capacity of the model. Phylogenetic signal among occupancy estimates was measured using Pagel’s [](https://www.codecogs.com/eqnedit.php?latex=%5Clambda#0) where values close to zero indicate no phylogenetic signal in the response and values close to one indicate strong correlation with phylogeny (Pagel 1999; Freckleton *et al.* 2002). This work was performed using the R package “brms” which interfaces with Stan (Bürkner 2017). We ran all of our post-hoc models for 200,000 iterations, discarding 100,000 as “burn-in” and thinning by 50 across four chains for a total of 4,000 samples from the posterior. We again assessed convergence by using the Gelman-Rubin diagnostic values and by inspecting trace plots. We provide all model code and complete data files used in our entire analyses via GitHub (<https://github.com/vmshirey/HighLatitudeNorthAmericanButterflyOccupancy>) and DataDryad (is there a DOI you can refer to here when posted?).

**3. RESULTS**

We modeled the response of 90 species of butterfly to changes to temperature from 1970-2019. On average, occupancy probability increased with increasing minimum temperatures for the majority of warm-adapted species (Figure 2, 3). Intuitively, warm-associated species showed the greatest positive response to increases in minimum temperature and cold-associated species showed acute declines in their response (Figure 2); however, modeled results also showed and unexpected reversal in response for cold-associated species when temperatures increase (Figure 2), although confidence was much lower for that response. Notably, several cold-adapted species exhibit average occupancy declines across their modeled ranges including *Boloria freija* (BOLFRE, -7.5% in mid-latitude sites), *B. chariclea* (BOLCHA, -6.8% in mid-latitude sites), *Boloria eunomia* (BOLEUN, -5.9% in mid-latitude sites), and *Agriades glandon* (AGRGLA, -3.1% in mid-latitude sites). In contrast, warm-adapted/southern species have, in general, exhibited occupancy stability or average increases alongside rising minimum temperatures (Figure 3). For example, the species, *Pieris rapae* (PIERAP), *Pterourus rutulus* (PTERUT) and *Cercyonis pegala* (CERPEG) have all exhibited average occupancy increases of roughly 4% or greater across all geographic components of their modeled range. Several species for which occupancy trends appear to marginally increase/decrease or that have remained relatively stable compared to the 1970s include *Coenonympha tullia* (COETUL) and *Glaucopsyche lygdamus* (GLALYG).

In sum, we find trends at the northern third of each species’ model range are generally more positive than trends at the southern third (average northern occupancy shifts are greater than southern shifts for 60 species compared to 29 species where average southern occupancy shifts are greater than those in the northern subcomponent of their range). Average trend estimates for cold-associated species are typically negative across all geographic subcomponents of their ranges (Figure 3a, b, c). For warm-associate species, the trends in each geographic subcomponent are stable or increasing (bearing in mind that “southern” and “mid-latitude” sites do not include sites below 45°N). Due to the very small number of sites modeled for *P. glaucus*, a comparison could not be made between northern and southern sites. A full account of average occupancy shifts as well as maps reproducing the shifts from the 1970s for each species are included in the supplementary material for all models.

The top model from our post-hoc analysis revealed that a model including range-wide temperature (Models C/D) and a model including average wingspan (Model G/H) were the best models in terms of predicting occupancy shift at the 100-kilometer scale (Supplemental Table S2). Generally, species with warmer North American ranges and species that have larger wingspans exhibited larger gains/smaller losses in occupancy probability over the period of this research (Figure 4a, b). Across other models not presented in the main text, the Model C/D group were consistently the top predictive models of occupancy trend for temperature-based occupancy trends (Supplemental Table S2-S3). An analysis of the phylogenetic signal in occupancy trend revealed weak evidence for a strong phylogenetic signal with Pagel’s being estimated at 0.0 (0.0 – 0.01 95% Bayesian credible interval) for Model D and Model H across all models. This indicates that phylogenetic signal among occupancy trends is low.

**4. DISCUSSION**

Our models demonstrate that minimum temperature (and in corollary, precipitation) predicts the overall 50-year occupancy trajectory of butterfly species in our study region. On average, minimum temperatures have increased by an average of 0.86 degrees Celsius between the 1970s and 2010s (Figure 1). Such increases may be related to occupancy via two distinct effects on butterflies. First, elevated minimum temperatures may be detrimental to cold-associated species as lack of snow cover and/or false phenological cues of winter termination may increase the risk of exposure of diapausing butterflies to harsh winter temperatures and early spring frosts. The impact of winter heatwaves has been well quantified in experimental studies and may be a core driver of population decline at the trailing southern edge of species’ ranges as has been found in the Baltimore Checkerspot, *Euphydryas phaeton* (Lepidoptera: Nymphalidae) (Abarca *et al.* 2019). Second, elevated minimum temperatures are likely beneficial to more southern-distributed species as previously harsh winters become milder, allowing for northward range expansion as has been found in some species (Crozier 2003, 2004). Indeed, poleward shifts are a commonly noted phenomenon among species elsewhere as they aim to track their thermal optima with a changing climate (Parmesan *et al.* 1999; Pöyry *et al.* 2009; Breed *et al.* 2013). Spatially, our results confirm that stronger positive shifts in species occupancy are occurring for a majority of species (n = 60 or roughly 66%) in the northern geographic context of their modeled range (Figure 3). The estimated effect of minimum temperature on species-specific occupancy from our model supports these aforementioned hypotheses (Figure 2), and as evidenced by cold-associated species generally declining in occupancy probability as minimum temperatures rise (Figure 3). In contrast, many of the warm-associated species in our study benefit from the same increasing minimum temperatures (Figure 2). Further, species with intermediate range-wide temperatures exhibit a positive response to rising minimum temperatures (Figure 2).

Our post-hoc analysis of species traits confirmed that a species’ range-wide temperature was the top predictor of 50-year occupancy shift across our study region (Supplemental Table S2; Model C/D). Indeed, species’ thermal niches are nearly ubiquitous predictors of range shifts across the tree of life (McMahon & Hays 2006; Scridel *et al.* 2017; Braschler *et al.* 2020). Further, given that we informed our estimates of occupancy probability by climatic variables, it is a positive confirmation that range-wide average annual temperature was a strong predictor of trends and that our methodology was excelled at reconstructing ecologically sensible trends from sparse, presence-only data. Taken together, these results suggest a more nuanced perspective of insect biodiversity decline and suggest that a scenario of ecological “winners” and “losers” (Jackson *et al.* 2022) is likely occurring as opposed to ubiquitous declines. The ability to discern historical and predict future *species-specific* trends will be important with respect to conservation assessments and intervention strategies.

Recent work in the western United States has revealed parallel sensitivity of butterfly species to global climate change. In particular, warmer autumn months has been identified as a potential driver for fewer butterflies being seen by community scientists in the western United States (below our study region) (Forister *et al.* 2021). In comparison to work from the Forister team, we find that similar species in our study also exhibit declines including, the Large Marble, *Euchloe ausonides* (Lepidoptera: Pieridae) and the Anicia Checkerspot, *Euphydryas anicia* (Lepidoptera: Nymphalidae) are in declines. Our trends are an important addition to this work and provide context for some of the same species in the northernmost reaches of their ranges.

Since prediction is an important part of the scientific enterprise, we aimed to also test which traits, aside from range-wide temperature (if any), might also predict occupancy declines in our study region. Among our models, Model G/H, using average wingspan as a predictor, emerged as a near equivalents to Model C/D candidate for predicting occupancy trend (Supplemental Table S2). In these models, butterflies with larger wingspans are more likely to be increasing in occupancy probability which smaller butterflies are more likely to be in decline (Figure 4b). Wingspan is often used as a proxy for mobility where species with larger wingspans are typically considered to have greater flying ability and, consequently, improved odds of navigating to new, suitable habitats (Sekar 2012). We discuss the important of mobility in this region later in the discussion. Further, we note that wingspan as a proxy for mobility appears to operate at both intra- and interspecific scales of variation. For example, in the Monarch butterfly (*Danaus plexippus*, not modeled here due to its exceptional migratory status), long-distance migrants exhibit larger wingspans than generations that do not make the long flight to the overwintering grounds in Mexico (Dockx 2007; Freedman & Dingle 2018; Freedman *et al.* 2020). Emerging work on the impacts of climate on morphology have shown that butterflies fluctuate in size (Bowden *et al.* 2015; Daly 2018) and thus future work should consider how intraspecific variation in morphology, driven by global change, may predict, exacerbate, or mitigate occupancy trends at various spatial scales. For example, if butterflies are becoming smaller due to climate change, this may, in turn, impact their mobility, and consequently, the ability for them to track their thermal niche. The degree to which these effects are not captured by models that don’t account for intraspecific variation in these factors across a species range remains to be tests.

While we examined the impacts of climate on species-specific occupancy trends, the importance of changing land-cover/use on boreal and Arctic butterflies cannot be ignored. Butterflies are especially sensitive to habitat type and disturbance as well as the availability of corridors to navigate to new, suitable habitats. In boreal Canada, butterfly abundance and species richness increases along human-made cutlines (Riva *et al.* 2018a). Additionally, increasing frequency and intensity of forest fire in the region may also contribute to both opportunities and challenges for high-latitude butterflies (Girardin & Mudelsee 2008; Hanes *et al.* 2019). Forest fire (of varying severity) may contribute to initial negative butterfly abundance patterns, but over time, butterflies may benefit from the presence of early successional and open canopy habitats (Johansson *et al.* 2020; Mason Jr *et al.* 2021; Ulyshen *et al.* 2022). Further, open canopy areas (including roadsides, smaller cutlines, and trails) may act as corridors for butterfly movement (Haddad 1999; Haddad & Tewksbury 2005; Riva *et al.* 2018b). Low connectivity of habitat/mobility of species can make it more challenging for more southern distributed butterflies to navigate and colonize new habitats as they track their thermal tolerances (Hodgson *et al.* 2012). Our model results point to a scenario where larger butterflies typically fare better with respect to occupancy declines, which, in the absence of connectivity data, suggest by proxy that species’ mobility is an important factor for persistence under rapidly changing climates in the region. Limited ability to colonize newly available habitat in this region is the leading hypothesis surrounding why so much climate debt has accrued for butterflies here (Lewthwaite *et al.* 2017, 2018).

Finally, to contextualize our results back into the world of modeling practice, we note several key findings from researchers on the forefront of statistical modeling with presence-only data. First, we note that several statistical advances have emerged from research teams to process community sourced survey data to estimate various biodiversity metrics (Dennis *et al.* 2021; Belitz *et al.* 2022). These advancements are exciting and with the integration of emerging technologies, including those from machine-learning (Joseph 2020) we may be approaching a technological singularity in biodiversity science, especially as it pertains to reconstructing historical ecologies. Finally, we note that there are the robust relationships between occupancy and abundance (Gaston *et al.* 2000; Zuckerberg *et al.* 2009; but see Dennis *et al.* 2019). Estimating abundance has been the statistical “Holy Grail” for ecologists working to reconstruct historical trends, especially since we can never go back in time to conduct structured surveys of historical populations. We are optimistic about the use of hierarchical and integrated modeling approaches in this space (Davis *et al.* 2023) and, to the extent that occupancy can be considered a proxy for underlying abundance patterns, suggest that well-crafted occupancy-detection approaches are a huge leap forward in the use of presence-only data to assess insect declines.

**5. CONCLUSION**

Global climate change is impacting high-latitude butterfly communities across North America; however, it is important to recognize that not all species are being impacted equally. Our research, in tandem with recent occupancy analyses of other groups demonstrates the importance of understanding how climatic shifts will impact insects on a species-specific level. Further, we show that occupancy-detection models can be used with sparse, presence-only data to extract clear ecological signals over large spatiotemporal scales. As such, occupancy-detection models will be an important tool for ecologists and conservation biologists, especially for hyper-diverse groups like insects, where structured monitoring data are often unavailable. While an overall perspective shows that rising minimum temperatures may benefit the majority of butterflies; cold-adapted species typically do not benefit. Species’ range-wide temperature and wingspan are the best predictors of overall occupancy trend, pointing to both the ability of the occupancy-detection approach for reconstructing historical trends from presence-only data but also the importance of mobility for butterflies in the region as they aim to track their thermal tolerances. Further research on the interplay of traits including the role of intraspecific variation in thermal tolerance, phenological and morphological adaptation, and genetic variation/effective population size are needed to contextualize these broad scale patterns further and support predictive frameworks for insect biodiversity decline.

**Conflict of Interest:** The authors declare no conflict of interest in completing this work.

**Data Availability Statement:** The code utilized in this analysis are freely available via GitHub at <https://github.com/vmshirey/HighLatitudeNorthAmericanButterflyOccupancy> and via DataDryad at the following DOI: xxx. Data related to model inferences can be found on GitHub and also at DataDryad. Range maps used in this analysis can be made available upon request.

**REFERENCES**

Abarca, M., Larsen, E.A. & Ries, L. (2019). Heatwaves and novel host consumption increase overwinter mortality of an imperiled wetland butterfly. *Frontiers in Ecology and Evolution*, 7, 193.

Adamo, M., Chialva, M., Calevo, J., Bertoni, F., Dixon, K. & Mammola, S. (2021). Plant scientists’ research attention is skewed towards colourful, conspicuous and broadly distributed flowers. *Nature plants*, 7, 574–578.

Ascher, J.S., Marshall, L., Meiners, J., Yanega, D. & Vereecken, N.J. (2020). Heterogeneity in large-scale databases and the role of climate change as a driver of bumble bee decline. *Science*, 8, 685.

Belitz, M.W., Larsen, E.A., Shirey, V., Li, D. & Guralnick, R.P. (2022). Phenological research based on natural history collections: practical guidelines and a Lepidopteran case study. *Functional Ecology*, Early View.

Bintanja, R. & Andry, O. (2017). Towards a rain-dominated Arctic. *Nature Climate Change*, 7, 263–267.

Bowden, J.J., Eskildsen, A., Hansen, R.R., Olsen, K., Kurle, C.M. & Høye, T.T. (2015). High-Arctic butterflies become smaller with rising temperatures. *Biology Letters*, 11, 20150574.

Brackley, A. (2021). *No Longer Forgotten: Pupation as a Critical Link in the Lepidopteran Life Cycle*. Georgetown University.

Brackley, A., Lill, J.T. & Weiss, M.R. (2021). Adaptive ontogenetic shifts in larval responses to environmental cues. *Entomologia Experimentalis et Applicata*, 169, 1147–1156.

Braschler, B., Duffy, G.A., Nortje, E., Kritzinger-Klopper, S., du Plessis, D., Karenyi, N., *et al.* (2020). Realised rather than fundamental thermal niches predict site occupancy: Implications for climate change forecasting. *Journal of Animal Ecology*, 89, 2863–2875.

Breed, G., Stichter, S. & Crone, E. (2013). Climate-driven changes in northeastern US butterfly communities. *Nature Climate Change*, 3, 142–145.

Brock, J.P. & Kaufman, K. (2006). *Kaufman field guide to butterflies of North America*. Houghton Mifflin Harcourt.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of statistical software*, 80, 1–28.

Cardoso, P., Barton, P., Birkhofer, K., Chichorro, F., Deacon, C., Fartmann, T., *et al.* (2020). Scientists’ warning to humanity on insect extinctions. *Biological Conservation*, 242.

Cardoso, P., Erwin, T.L. & Borges, P.A.V. (2011). The seven impediments in invertebrate conservation and how to overcome them. *Biological Conservation*, 144, 2647–2655.

Chichorro, F., Juslén, A. & Cardoso, P. (2019). A review of the relation between species traits and extinction risk. *Biological Conservation*, 237, 220–229.

Chichorro, F., Urbano, F., Teixeira, D., Väre, H., Pinto, T., Brummitt, N., *et al.* (2022). Trait-based prediction of extinction risk across terrestrial taxa. *Biological Conservation*, 274, 109738.

Climate Prediction Center. (2022). Global Temperature data provided by the NOAA/OAR/ESRL PSL.

Corlett, R.T. & Westcott, D.A. (2013). Will plant movements keep up with climate change? *Trends in ecology & evolution*, 28, 482–488.

Crozier, L. (2003). Winter warming facilitates range expansion: cold tolerance of the butterfly Atalopedes campestris. *Oecologia*, 135, 648–656.

Crozier, L. (2004). Warmer winters drive butterfly range expansion by increasing survivorship. *Ecology*, 85, 231–241.

Daly, K.M. (2018). Morphological and phenological responses of butterflies to seasonal temperature increase in Alaska. PhD Thesis. .

Davis, C.L., Guralnick, R.P. & Zipkin, E.F. (2023). Challenges and opportunities for using natural history collections to estimate insect population trends. *Journal of Animal Ecology*, 92, 237–249.

Dennis, E.B., Kéry, M., Morgan, B.J., Coray, A., Schaub, M. & Baur, B. (2021). Integrated modelling of insect population dynamics at two temporal scales. *Ecological Modelling*, 441, 109408.

Dennis, E.B., Morgan, B.J., Fox, R., Roy, D.B. & Brereton, T.M. (2019). Functional data analysis of multi-species abundance and occupancy data sets. *Ecological Indicators*, 104, 156–165.

Deutsch, C.A., Tewksbury, J.J., Huey, R.B., Sheldon, K.S., Ghalambor, C.K., Haak, D.C., *et al.* (2008). Impacts of climate warming on terrestrial ectotherms across latitude. *Proceedings of the National Academy of Sciences*, 105, 6668–6672.

Dockx, C. (2007). Directional and stabilizing selection on wing size and shape in migrant and resident monarch butterflies, Danaus plexippus (L.), in Cuba. *Biological Journal of the Linnean Society*, 92, 605–616.

Dorazio, R.M. & Royle, J.A. (2005). Estimating size and composition of biological communities by modeling the occurrence of species. *Journal of the American Statistical Association*, 100, 389–398.

Downes, J. (1965). Adaptations of insects in the Arctic. *Annual Review of Entomology*, 10, 257–274.

Dunne, J.A., Williams, R.J. & Martinez, N.D. (2002). Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology letters*, 5, 558–567.

Earl, C., Belitz, M.W., Laffan, S.W., Barve, V., Barve, N., Soltis, D.E., *et al.* (2021). Spatial phylogenetics of butterflies in relation to environmental drivers and angiosperm diversity across North America. *iScience*, 24, 102239.

Elith, J. & Leathwick, J.R. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual review of ecology, evolution, and systematics*, 40, 677–697.

Engelhardt, E.K., Biber, M.F., Dolek, M., Fartmann, T., Hochkirch, A., Leidinger, J., *et al.* (2022). Consistent signals of a warming climate in occupancy changes of three insect taxa over 40 years in central Europe. *Global Change Biology*, 28, 3998–4012.

Fagan, W., Meir, E., Prendergast, J., Folarin, A. & Karieva, P. (2001). Characterizing population vulnerability for 758 species. *Ecology Letters*, 4, 132–138.

Forister, M., Halsch, C., Nice, C., Fordyce, J., Dilts, T., Oliver, J., *et al.* (2021). Fewer butterflies seen by community scientists across the warming and drying landscapes of the American West. *Science*, 371, 1042–1045.

Freckleton, R.P., Harvey, P.H. & Pagel, M. (2002). Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist*, 160.

Freedman, M.G. & Dingle, H. (2018). Wing morphology in migratory North American monarchs: characterizing sources of variation and understanding changes through time. *Animal Migration*, 5, 61–73.

Freedman, M.G., Dingle, H., Strauss, S.Y. & Ramírez, S.R. (2020). Two centuries of monarch butterfly collections reveal contrasting effects of range expansion and migration loss on wing traits. *Proceedings of the National Academy of Sciences*, 117, 28887–28893.

Gaston, K.J., Blackburn, T.M., Greenwood, J.J., Gregory, R.D., Quinn, R.M. & Lawton, J.H. (2000). Abundance–occupancy relationships. *Journal of Applied Ecology*, 37, 39–59.

GBIF.org. (2022). Occurrence Download for High Latitude, North American Butterflies.

Gelman, A. & Rubin, D.B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 457–472.

Girardin, M.P. & Mudelsee, M. (2008). Past and future changes in Canadian boreal wildfire activity. *Ecological Applications*, 18, 391–406.

Glassberg, J. (2018). *A swift guide to butterflies of Mexico and Central America*. Princeton University Press.

Grames, E.M., Montgomery, G.A., Youngflesh, C., Tingley, M.W. & Elphick, C.S. (2023). The effect of insect food availability on songbird reproductive success and chick body condition: Evidence from a systematic review and meta-analysis. *Ecology Letters*.

Guzman, L.M., Johnson, S.A., Mooers, A.O. & M’Gonigle, L.K. (2021). Using historical data to estimate bumble bee occurrence: Variable trends across species provide little support for community-level declines. *Biological Conservation*, 257, 109141.

Haddad, N.M. (1999). Corridor and distance effects on interpatch movements: a landscape experiment with butterflies. *Ecological Applications*, 9, 612–622.

Haddad, N.M. & Tewksbury, J.J. (2005). Low-quality habitat corridors as movement conduits for two butterfly species. *Ecological Applications*, 15, 250–257.

Hällfors, M., Poyry, J., Heliola, J., Kohonen, I., Kuussaari, M., Leinonen, R., *et al.* (2021). Combining range and phenology shifts offers a winning strategy for boreal Lepidoptera. *Ecology Letters*, 24, 1619–1632.

Hällfors, M.H., Heikkinen, R.K., Kuussaari, M., Lehikoinen, A., Luoto, M., Pöyry, J., *et al.* (2023). Recent range shifts of moths, butterflies, and birds are driven by the breadth of their climatic niche. *Evolution Letters*, qrad004.

Hanes, C.C., Wang, X., Jain, P., Parisien, M.-A., Little, J.M. & Flannigan, M.D. (2019). Fire-regime changes in Canada over the last half century. *Canadian Journal of Forest Research*, 49, 256–269.

Harris, I., Osborn, T.J., Jones, P. & Lister, D. (2020a). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific data*, 7, 1–18.

Harris, I.C., Jones, P.D. & Osborn, T. (2020b). CRU TS4.04: Climatic Research Unit (CRU) Time-Series (TS) version 4.04 of high-resolution gridded data of month-by-month variation in climate (January 1901 – December 2019).

Heikkinen, R., Luoto, M., Leikola, N., Poyry, J., Settele, J., Kudrna, O., *et al.* (2010). Assessing the vulnerability of European butterflies to climate change using multiple criteria. *Biodiversity and Conservation*, 19, 695–723.

Heinrich, P.L., Gilbert, E., Cobb, N.S. & Franz, N. (2015). Symbiota collections of arthropods network (SCAN): A data portal built to visualize, manipulate, and export species occurrences.

Hijmans, R.J., Van Etten, J., Cheng, J., Mattiuzzi, M., Sumner, M., Greenberg, J.A., *et al.* (2015). Package ‘raster.’ *R package*, 734, 473.

Hill, J., Thomas, C., Fox, R., Telfer, M., Willis, S., Asher, J., *et al.* (2002). Responses of butterflies to twentieth century climate warming: implications for future ranges. *Proceedings of the Royal Society B-Biological Sciences*, 269, 2163–2171.

Hodgson, J.A., Thomas, C.D., Dytham, C., Travis, J.M. & Cornell, S.J. (2012). The speed of range shifts in fragmented landscapes.

Holland, M.M. & Bitz, C.M. (2003). Polar amplification of climate change in coupled models. *Climate dynamics*, 21, 221–232.

Jackson, H.M., Johnson, S.A., Morandin, L.A., Richardson, L.L., Guzman, L.M. & M’Gonigle, L.K. (2022). Climate change winners and losers among North American bumblebees. *Biology Letters*, 18, 20210551.

Johansson, V., Gustafsson, L., Andersson, P. & Hylander, K. (2020). Fewer butterflies and a different composition of bees, wasps and hoverflies on recently burned compared to unburned clear-cuts, regardless of burn severity. *Forest Ecology and Management*, 463, 118033.

Jönsson, G.M., Broad, G.R., Sumner, S. & Isaac, N.J. (2021). A century of social wasp occupancy trends from natural history collections: spatiotemporal resolutions have little effect on model performance. *Insect Conservation and Diversity*, 14, 543–555.

Joseph, M.B. (2020). Neural hierarchical models of ecological populations. *Ecology Letters*, 23, 734–747.

Keinath, D.A., Doak, D.F., Hodges, K.E., Prugh, L.R., Fagan, W., Sekercioglu, C.H., *et al.* (2017). A global analysis of traits predicting species sensitivity to habitat fragmentation. *Global Ecology and Biogeography*, 26, 115–127.

Kellermann, V. & van Heerwaarden, B. (2019). Terrestrial insects and climate change: adaptive responses in key traits. *Physiological Entomology*, 44, 99–115.

Kelling, S., Johnston, A., Bonn, A., Fink, D., Ruiz-Gutierrez, V., Bonney, R., *et al.* (2019). Using semistructured surveys to improve citizen science data for monitoring biodiversity. *BioScience*, 69, 170–179.

Kellner, K. (2021). *jagsUI: A Wrapper Around “rjags” to Streamline “JAGS” Analyses*.

Keret, N., Mutanen, M., Orell, M., Itamies, J. & Valimaki, P. (2020). Climate change-driven elevational changes among boreal nocturnal moths. *Oecologia*, 192, 1085–1098.

Kéry, M. & Royle, J.A. (2015). Applied Hierarchical Modeling in Ecology: Analysis of Distribution, Abundance and Species Richness in R and BUGS.

Kremen, C., Chaplin-Kramer, R., & others. (2007). Insects as providers of ecosystem services: crop pollination and pest control. In: *Insect conservation biology: proceedings of the royal entomological society’s 23rd symposium*. CABI Publishing, pp. 349–382.

Lamas, G. (2015). *Catalog of the butterflies (Papilionoidea)*. Available from author.

Larsen, E.A. & Shirey, V. (2021). Method matters: pitfalls in analysing phenology from occurrence records. *Ecology Letters*, 24, 1287–1289.

Layne Jr, J.R. & Kuharsky, D.K. (2000). Triggering of cryoprotectant synthesis in the woolly bear caterpillar (Pyrrharctia isabella Lepidoptera: Arctiidae). *Journal of Experimental Zoology*, 286, 367–371.

Lewthwaite, J.M., Debinski, D.M. & Kerr, J.T. (2017). High community turnover and dispersal limitation relative to rapid climate change. *Global Ecology and Biogeography*, 26, 459–471.

Lewthwaite, J.M.M., Angert, A.L., Kembel, S.W., Goring, S.J., Davies, T.J., Mooers, A.Ø., *et al.* (2018). Canadian butterfly climate debt is significant and correlated with range size. *Ecography*, 41, 2005–2015.

MacKenzie, D.I., Nichols, J.D., Lachman, G.B., Droege, S., Andrew Royle, J. & Langtimm, C.A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83, 2248–2255.

Mason Jr, S.C., Shirey, V., Ponisio, L.C. & Gelhaus, J.K. (2021). Responses from bees, butterflies, and ground beetles to different fire and site characteristics: a global meta-analysis. *Biological Conservation*, 261, 109265.

McCrystall, M.R., Stroeve, J., Serreze, M., Forbes, B.C. & Screen, J.A. (2021). New climate models reveal faster and larger increases in Arctic precipitation than previously projected. *Nature communications*, 12, 1–12.

McMahon, C.R. & Hays, G.C. (2006). Thermal niche, large-scale movements and implications of climate change for a critically endangered marine vertebrate. *Global Change Biology*, 12, 1330–1338.

Memmott, J., Waser, N.M. & Price, M.V. (2004). Tolerance of pollination networks to species extinctions. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 271, 2605–2611.

Meyer, C., Jetz, W., Guralnick, R.P., Fritz, S.A. & Kreft, H. (2016). Range geometry and socio-economics dominate species-level biases in occurrence information. *Global Ecology and Biogeography*, 25, 1181–1193.

Ono, J., Watanabe, M., Komuro, Y., Tatebe, H. & Abe, M. (2022). Enhanced Arctic warming amplification revealed in a low-emission scenario. *Communications Earth & Environment*, 3, 27.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature*, 401, 877–884.

Parmesan, C., Ryrholm, N., Stefanescu, C., Hill, J., Thomas, C., Descimon, H., *et al.* (1999). Poleward shifts in geographical ranges of butterfly species associated with regional warming. *Nature*, 399, 579–583.

Pavulaan, H. & Wright, D.M. (2000). The Biology, Life History, and Taxonomy of Celastrina neglectamajor (Lycaenidae: Polyommatinae).

Pavulaan, H. & Wright, D.M. (2005). Celastrina serotina (Lycaenidae: Polyommatinae): a new butterfly species from the northeastern United States and eastern Canada.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In: *Proceedings of the 3rd international workshop on distributed statistical computing*. Vienna, Austria., pp. 1–10.

Pöyry, J., Luoto, M., Heikkinen, R., Kuussaari, M. & Saarinen, K. (2009). Species traits explain recent range shifts of Finnish butterflies. *Global Change Biology*, 15, 732–743.

Ries, L., Zipkin, E.F. & Guralnick, R.P. (2019). Tracking trends in monarch abundance over the 20th century is currently impossible using museum records. *Proceedings of the National Academy of Sciences*, 116, 13745–13748.

Riva, F., Acorn, J.H. & Nielsen, S.E. (2018a). Localized disturbances from oil sands developments increase butterfly diversity and abundance in Alberta’s boreal forests. *Biological Conservation*, 217, 173–180.

Riva, F., Acorn, J.H. & Nielsen, S.E. (2018b). Narrow anthropogenic corridors direct the movement of a generalist boreal butterfly. *Biology Letters*, 14, 20170770.

Scott Chamberlain & Eduard Szocs. (2013). taxize - taxonomic search and retrieval in R. *F1000Research*.

Scridel, D., Bogliani, G., Pedrini, P., Iemma, A., von Hardenberg, A. & Brambilla, M. (2017). Thermal niche predicts recent changes in range size for bird species. *Climate Research*, 73, 207–216.

Sekar, S. (2012). A meta-analysis of the traits affecting dispersal ability in butterflies: can wingspan be used as a proxy? *Journal of Animal Ecology*, 81, 174–184.

Shirey, V., Belitz, M.W., Barve, V. & Guralnick, R. (2021). A complete inventory of North American butterfly occurrence data: narrowing data gaps, but increasing bias. *Ecography*, 44.

Shirey, V., Khelifa, R., M’Gonigle, L.K. & Guzman, L.M. (2022). Occupancy-detection models for natural history museum data: promise and pitfalls. *Methods in Ecology and Evolution*, Early View.

Slivinski, L.C., Compo, G.P., Whitaker, J.S., Sardeshmukh, P.D., Giese, B.S., McColl, C., *et al.* (2019). Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, 145, 2876–2908.

van Strien, A.J., van Swaay, C.A.M. & Termaat, T. (2013). Opportunistic citizen science data of animal species produce reliable estimates of distribution trends if analysed with occupancy models. *Journal of Applied Ecology*, 50, 1450–1458.

Ulyshen, M.D., Hiers, J.K., Pokswinksi, S.M. & Fair, C. (2022). Pyrodiversity promotes pollinator diversity in a fire-adapted landscape. *Frontiers in Ecology and the Environment*, 20, 78–83.

Vehtari, A., Gelman, A. & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and computing*, 27, 1413–1432.

Wagner, D., Fox, R., Salcido, D. & Dyer, L. (2021a). A window to the world of global insect declines: Moth biodiversity trends are complex and heterogeneous. *Proceedings of the National Academy of Sciences of the United States of America*, 118.

Wagner, D.L., Grames, E.M., Forister, M.L., Berenbaum, M.R. & Stopak, D. (2021b). Insect decline in the Anthropocene: Death by a thousand cuts. *Proceedings of the National Academy of Sciences*, 118, e2023989118.

Wepprich, T. (2019). Monarch butterfly trends are sensitive to unexamined changes in museum collections over time. *Proceedings of the National Academy of Sciences*, 116, 13742–13744.

Wright, D.M. & Pavulaan, H. (1999). Celastrina idella (Lycaenidae: Polyommatinae): a new butterfly species from the Atlantic coastal plain.

Yackulic, C.B., Chandler, R., Zipkin, E.F., Royle, J.A., Nichols, J.D., Campbell Grant, E.H., *et al.* (2013). Presence-only modelling using MAXENT: when can we trust the inferences? *Methods in Ecology and Evolution*, 4, 236–243.

You, Q., Cai, Z., Pepin, N., Chen, D., Ahrens, B., Jiang, Z., *et al.* (2021). Warming amplification over the Arctic Pole and Third Pole: Trends, mechanisms and consequences. *Earth-Science Reviews*, 217, 103625.

Youngflesh, C. (2018). MCMCvis: Tools to visualize, manipulate, and summarize MCMC output. *Journal of Open Source Software*, 3, 640.

Zhou, Y., Zhang, H., Liu, D., Khashaveh, A., Li, Q., Wyckhuys, K.A., *et al.* (2023). Long-term insect censuses capture progressive loss of ecosystem functioning in East Asia. *Science Advances*, 9, eade9341.

Zipkin, E.F., Royle, J.A., Dawson, D.K. & Bates, S. (2010). Multi-species occurrence models to evaluate the effects of conservation and management actions. *Biological Conservation*, 143, 479–484.

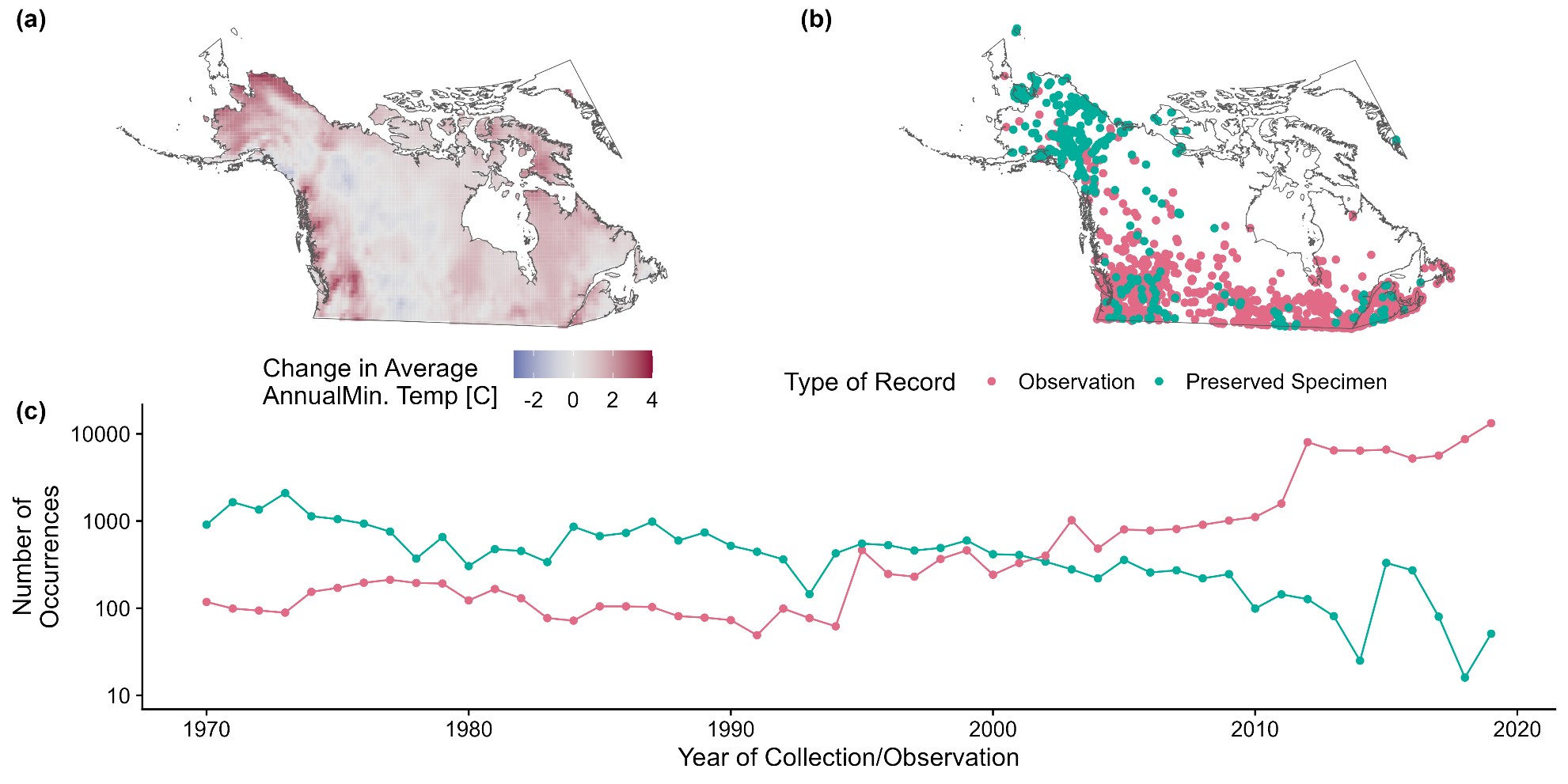
Zuckerberg, B., Porter, W.F. & Corwin, K. (2009). The consistency and stability of abundance–occupancy relationships in large-scale population dynamics. *Journal of Animal Ecology*, 78, 172–181.

**TABLES**

**Table 1.** Descriptions of the trait-based models of occupancy trend used in our post-hoc analysis. Model codes are provided as well as mathematical definitions of the model. Supporting literature for each model is also provided.

|  |  |  |
| --- | --- | --- |
| **Model ID** | **Model Specification** | **Rationale** |
| A  “Null” |  | Ecological “null” model, *i.e.,* all occupancy trends are predicted by an average intercept. |
| B  “Null + Phylo” |  | Model A with a phylogenetic random intercept term, *i.e.,* species occupancy trends are predicted by an average intercept and species-specific, phylogenetic intercept. |
| C  “Temp” |  | Occupancy trends are best predicted by species’ range-wide average annual temperature. Warmer, southern species will exhibit increases in occupancy probability. |
| D  “Temp + Phylo” |  | Occupancy trends are best predicted by species’ range-wide average annual temperature and a species-specific, phylogenetic intercept. Warmer, southern species will exhibit increases in occupancy probability. |
| E  “Range” |  | Occupancy trends are best predicted by species’ range-size. Widely distributed species will exhibit increases in occupancy probability due to wider environmental tolerance. |
| F  “Range + Phylo” |  | Occupancy trends are best predicted by species’ range-size and a species-specific phylogenetic intercept. Widely distributed species will exhibit increases in occupancy probability due to wider environmental tolerance. |
| G  “Size” |  | Occupancy trends are best predicted by species’ average wingspan. Larger species will exhibit increases in occupancy probability due to greater mobility (Sekar 2012). |
| H  “Size + Phylo” |  | Occupancy trends are best predicted by species’ average wingspan and a species-specific phylogenetic intercept. Larger species will exhibit increases in occupancy probability due to greater mobility (Sekar 2012). |
| I  “Resource” |  | Occupancy trends are best predicted by species’ family-level host plant breadth. Species with greater host plant breadth will exhibit increases in occupancy probability due to broader resource availability. |
| J  “Resource + Phylo” |  | Occupancy trends are best predicted by species’ family-level host plant breadth and a species-specific phylogenetic intercept. Species with greater host plant breadth will exhibit increases in occupancy probability due to broader resource availability. |
| K  “Overwinter” |  | Occupancy trends are best predicted by range-wide temperature, overwintering life stage, and the interaction between those two factors. We included an interaction term here since physiological or behavioral overwintering strategies could vary by a species’ thermal adaptability. For example, some species of Lepidoptera possess the ability to synthesize antifreeze compounds while others may burrow underground or find other sheltering mechanisms to survive winter conditions (Downes 1965; Layne Jr & Kuharsky 2000; Brackley 2021). Further, responses to environmental cues can vary ontogenetically (Brackley *et al.* 2021). |
| L  “Overwinter + Phylo” |  | Occupancy trends are best predicted by range-wide temperature, overwintering life stage, the interaction between those two factors and a species-specific, phylogenetic intercept. Following the same logic as Model K. |
| M  “Complex” |  | Occupancy trends are best predicted by a model that includes all available trait information. Many traits predict occupancy trend. |
| N  “Complex + Phylo” |  | Occupancy trends are best predicted by a model that includes all available trait information and a species-specific, phylogenetic intercept. Many traits predict occupancy trend. |

**FIGURES**

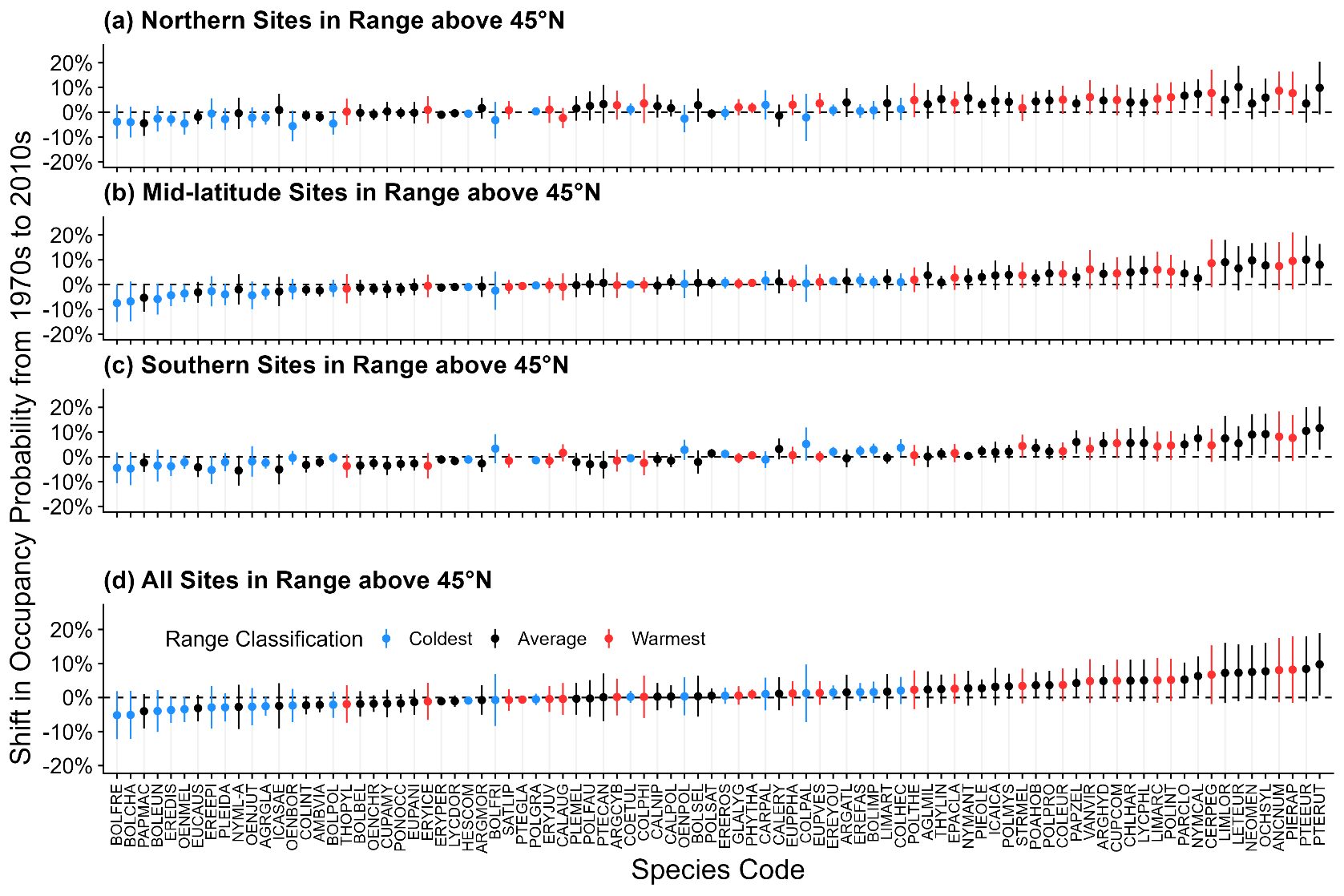


**Figure 1.** An overview of the study region encompassed by our analysis including (a) the change in average annual minimum temperature from the 1970s to 2010s, (b) the spatial distribution of occurrence records from natural history collections and community science platforms, and (c) the number of occurrence records from natural history museum collections and community science platforms by year over the same 50-year timeframe. A random sample of 5,000 of the occurrence records (c) are shown to indicate spatial bias in available records while avoiding overplotting.

Chart

Description automatically generated

**Figure 2.** Species-specific responses to average minimum temperature (faded, grey lines) as well as the response of the butterflies with the coldest quartile of North American ranges when all species’ entire ranges are compared (blue, *n = 23*), warmest quartile of North American ranges (red, *n = 23*), and middle half of species or those species with average range-wide temperatures (grey, *n = 44*). Species-specific responses are shown in grey and truncated to show modeled results only for the range of temperatures that each species experienced within its range above 45°N between 1970-2019. All lines are derived from model estimated parameters (linear and quadratic) of species to the average minimum temperature covariate and species-specific responses are shown in Figure 3.



**Figure 3.** Species-specific occupancy shifts from the 1970s to 2010s where points indicate the mean occupancy shift and lines indicate one standard deviation of the variation in occupancy shift among (a) the northernmost third of cells within that species’ range in the study region, (b) the mid-latitude third of sites within that species’ range, (c) the southernmost third of sites within that species’ range, and (d) all of the sites within that species’ range (all ranges truncated at 45°N). Species are colored by quartiles of the average annual temperature in their North American ranges, so it is directly comparable to Figure 2.

Chart, scatter chart

Description automatically generated

**Figure 4.** The predictive relationship between species’ (a) range-wide average annual temperature (Model C) and (b) average wingspan (Model G) for the mean occupancy probability shift from the 1970s to 2010s. Shaded regions indicate Bayesian credible intervals while points represent the individual species in our post-hoc analysis.

**SUPPLEMENTAL MATERIAL**

***Table of Contents***

**Supplemental Table S1.** Average occupancy trend estimates across all models/species

**Supplemental Table S2.** Model selection for post-hoc 100-km temperature analysis

**Supplemental Table S3.** Model selection for post-hoc 200-km temperature analysis

**Supplemental Table S4.** Model selection for post-hoc 100-km precipitation analysis

**Supplemental Table S5.** Model selection for post-hoc 200-km precipitation analysis

**Supplemental Figure S1.** Graphical illustration of the approach used in this study

**Supplemental Figure S2.** Occupancy trends from the 200km temperature model

**Supplemental Figure S3.** Occupancy trends from the 100km precipitation model

**Supplemental Figure S4.** Occupancy trends from the 200km precipitation model

**Supplemental Figure S5.** Parameter estimates for the 200km temperature post-hoc Model C.

**Supplemental Figure S7.** Parameter estimates for the 100km precipitation post-hoc Model C.

**Supplemental Figure S8.** Parameter estimates for the 200km precipitation post-hoc Model C.

**ADDITIONAL SUPPLEMENTARY MATERIAL FILES**

**Supplementary File S1.** Mean occupancy shift maps for species modeled using across all models (denoted by specific .pdf files names for each, default is the 100-kilometer temperature model).

**Supplemental File S2:** List of accessed collections through iDigBio.

**Supplemental File S3:** Trace plots and other model assessment metrics for our 100-kilometer temperature occupancy-detection model.

**Supplemental File S4:** Trace plots and other model assessment metrics for our 200-kilometer temperature occupancy-detection model.

**Supplemental File S5:** Trace plots and other model assessment metrics for our 100-kilometer precipitation occupancy-detection model.

**Supplemental File S6:** Trace plots and other model assessment metrics for our 200-kilometer precipitation occupancy-detection model.

**Supplemental Table S1.** Mean occupancy trend for modeled butterfly species across all models in our analysis. Occupancy trends reflect the overall average trend and variance (+/- one standard deviation) across all cells modeled for each species. Inferences from each model type are denoted by column headers (100T = 100km temperature, 200T = 200km temperature, 100P = 100km precipitation, and 200P = 200km precipitation. Cells are conditionally formatted by the direction and magnitude of decline/increase where dark red indicates strong declines in occupancy probability and dark blue indicates strong increases in occupancy probability. White shading indicates marginal decline/increase.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **CODE** | **SCIENTIFIC NAME** | **100T MODEL** | **200T MODEL** | **100P MODEL** | **200P MODEL** |
| AGLMIL | *Aglais milberti* | 2.4% (+/-5.4%) | 2.8% (+/-5.8%) | 0.0% (+/-0.1%) | 0.1% (+/-0.3%) |
| AGRGLA | *Agriades glandon* | -2.5% (+/-2.7%) | -2.2% (+/-2.3%) | -0.2% (+/-0.6%) | -0.1% (+/-0.4%) |
| AMBVIA | *Amblyscirtes vialis* | -2.2% (+/-2.1%) | -0.7% (+/-0.8%) | -0.7% (+/-1.6%) | -0.2% (+/-0.7%) |
| ANCNUM | *Ancyloxypha numitor* | 8.1% (+/-9.5%) | 6.3% (+/-7.7%) | 0.9% (+/-0.8%) | 0.9% (+/-0.9%) |
| ARGATL | *Argynnis atlantis* | 1.5% (+/-5.2%) | 1.8% (+/-6.0%) | -0.1% (+/-0.3%) | 0.1% (+/-0.3%) |
| ARGCYB | *Argynnis cybele* | 0.2% (+/-5.4%) | 0.5% (+/-4.6%) | -0.4% (+/-0.9%) | -0.2% (+/-0.8%) |
| ARGHYD | *Argynnis hydaspe* | 4.9% (+/-4.7%) | 4.2% (+/-4.6%) | 2.3% (+/-1.8%) | 2.5% (+/-2.1%) |
| ARGMOR | *Argynnis mormonia* | -0.7% (+/-4.3%) | 0.0% (+/-3.7%) | -0.6% (+/-0.7%) | -0.4% (+/-0.5%) |
| BORBEL | *Boloria bellona* | -1.9% (+/-3.7%) | -1.2% (+/-3.0%) | -0.9% (+/-3.5%) | -0.3% (+/-2.3%) |
| BOLCHA | *Boloria chariclea* | -5.1% (+/-7.1%) | -3.6% (+/-6.1%) | -0.9% (+/-2.4%) | -0.5% (+/-1.9%) |
| BOLEUN | *Boloria eunomia* | -3.9% (+/-6.1%) | -3.2% (+/-4.8%) | -0.7% (+/-1.9%) | -0.2% (+/-1.3%) |
| BOLFRE | *Boloria freija* | -5.2% (+/-7.1%) | -4.4% (+/-6.3%) | -1.7% (+/-3.7%) | -1.1% (+/-3.4%) |
| BOLFRI | *Boloria frigga* | -0.7% (+/-7.6%) | -1.4% (+/-7.6%) | -1.4% (+/-2.9%) | -0.9% (+/-3.3%) |
| BOLIMP | *Boloria improba* | 1.6% (+/-3.1%) | -1.8% (+/-3.7%) | -2.2% (+/-2.3%) | -1.9% (+/-2.7%) |
| BOLPOL | *Boloria polaris* | -2.1% (+/-3.8%) | -4.4% (+/-5.8%) | -3.0% (+/-3.5%) | -2.0% (+/-4.8%) |
| BOLSEL | *Boloria selene* | 0.4% (+/-6.0%) | 2.1% (+/-6.6%) | 1.1% (+/-2.2%) | 0.5% (+/-1.4%) |
| CALAUG | *Callophrys augustinus* | -0.4% (+/-4.7%) | 0.1% (+/-4.1%) | 0.6% (+/-1.3%) | 0.6% (+/-1.6%) |
| CALERY | *Callophrys eryphon* | 1.1% (+/-4.9%) | 1.3% (+/-4.4%) | 0.3% (+/-0.4%) | 0.2% (+/-0.6%) |
| CALNIP | *Callophrys niphon* | 0.3% (+/-4.0%) | 1.8% (+/-4.0%) | 0.3% (+/-0.3%) | 0.8% (+/-1.2%) |
| CALPOL | *Callophrys polios* | 0.3% (+/-3.4%) | 0.9% (+/-3.9%) | -0.2% (+/-0.5%) | 0.0% (+/-0.2%) |
| CARPAL | *Carterocephalus palaemon* | 1.0% (+/-4.8%) | 1.7% (+/-4.9%) | 0.9% (+/-2.0%) | 0.7% (+/-1.5%) |
| CERPEG | *Cercyonis pegala* | 6.7% (+/-8.6%) | 5.7% (+/-7.2%) | -0.1% (+/-0.3%) | 0.0% (+/-0.2%) |
| CHLHAR | *Chlosyne harrisii* | 4.9% (+/-6.2%) | 4.5% (+/-5.3%) | 2.5% (+/-2.2%) | 2.3% (+/-2.2%) |
| COETUL | *Coenonympha tullia* | 0.2% (+/-1.8%) | 0.9% (+/-3.4%) | 0.0% (+/-0.2%) | 0.0% (+/-0.1%) |
| COLEUR | *Colias eurytheme* | 3.7% (+/-4.8%) | 3.4% (+/-5.1%) | -0.3% (+/-0.9%) | -0.1% (+/-0.8%) |
| COLHEC | *Colias hecla* | 2.1% (+/-4.0%) | -2.2% (+/-5.8%) | -2.8% (+/-3.0%) | -2.4% (+/-3.9%) |
| COLINT | *Colias interior* | -2.3% (+/-2.7%) | -0.8% (+/-4.2%) | -0.5% (+/-1.8%) | -0.1% (+/-0.9%) |
| COLPAL | *Colias palaeno* | 1.3% (+/-8.4%) | 0.0% (+/-6.6%) | -3.0% (+/-5.1%) | -1.2% (+/-3.8%) |
| COLPHI | *Colias philodice* | 0.2% (+/-6.2%) | 1.1% (+/-8.6%) | -0.1% (+/-1.0%) | 0.0% (+/-0.9%) |
| CUPAMY | *Cupido amyntula* | -1.7% (+/-4.0%) | -0.4% (+/-3.5%) | -0.3% (+/-1.1%) | -0.1% (+/-0.5%) |
| CUPCOM | *Cupido comyntas* | 4.9% (+/-6.2%) | 6.3% (+/-8.0%) | 1.5% (+/-1.3%) | 1.0% (+/-0.9%) |
| EPACLA | *Epargyreus clarus* | 2.6% (+/-4.4%) | 1.4% (+/-2.5%) | 0.0% (+/-0.0%) | -0.4% (+/-0.4%) |
| EREDIS | *Erebia discoidalis* | -3.7% (+/-3.8%) | -3.2% (+/-2.9%) | -1.4% (+/-2.5%) | -0.7% (+/-2.0%) |
| EREEPI | *Erebia epipsodea* | -2.9% (+/-6.3%) | -1.8% (+/-6.0%) | -0.6% (+/-1.0%) | -0.3% (+/-0.6%) |
| EREFAS | *Erebia fasciata* | 1.6% (+/-2.7%) | -2.1% (+/-4.8%) | -2.4% (+/-2.5%) | -2.9% (+/-3.2%) |
| EREROS | *Erebia rossii* | 0.6% (+/-2.4%) | 0.1% (+/-4.7%) | -2.6% (+/-2.9%) | -2.5% (+/-3.1%) |
| EREYOU | *Erebia youngi* | 1.5% (+/-2.1%) | -0.1% (+/-2.4%) | -2.4% (+/-2.1%) | -2.5% (+/-2.3%) |
| ERYICE | *Erynnis icelus* | -1.1% (+/-5.4%) | -0.1% (+/-6.0%) | -0.1% (+/-1.6%) | 0.0% (+/-0.6%) |
| ERYJUV | *Erynnis juvenalis* | -0.4% (+/-3.8%) | 0.5% (+/-3.1%) | 0.8% (+/-0.6%) | 0.7% (+/-0.6%) |
| ERYPER | *Erynnis persius* | -1.1% (+/-0.9%) | -0.2% (+/-0.4%) | -0.2% (+/-0.5%) | 0.1% (+/-0.3%) |
| EUCAUS | *Euchloe ausonides* | -3.1% (+/-3.9%) | -2.2% (+/-4.4%) | -1.0% (+/-2.5%) | -0.3% (+/-1.9%) |
| EUPANI | *Euphydryas anicia* | -1.3% (+/-3.7%) | 0.1% (+/-2.8%) | -0.3% (+/-0.2%) | 0.2% (+/-0.2%) |
| EUPPHA | *Euphydryas phaeton* | 1.2% (+/-3.6%) | 2.8% (+/-4.5%) | 2.8% (+/-2.1%) | 3.2% (+/-2.7%) |
| EUPVES | *Euphyes vestris* | 1.4% (+/-3.5%) | 0.4% (+/-2.0%) | 2.5% (+/-3.2%) | 2.4% (+/-3.2%) |
| GLALYG | *Glaucopsyche lygdamus* | 0.6% (+/-2.7%) | 1.2% (+/-3.6%) | 0.1% (+/-0.1%) | 0.1% (+/-0.3%) |
| HESCOM | *Hesperia comma* | -0.9% (+/-0.9%) | -0.1% (+/-0.3%) | 0.1% (+/-0.2%) | 0.2% (+/-1.0%) |
| ICAICA | *Icaricia icarioides* | 3.2% (+/-5.5%) | 2.8% (+/-3.3%) | -0.6% (+/-0.4%) | 0.2% (+/-0.2%) |
| ICASAE | *Icaricia saepiolus* | -2.4% (+/-6.7%) | -1.2% (+/-5.6%) | -0.8% (+/-2.1%) | -0.3% (+/-1.1%) |
| LETEUR | *Lethe eurydice* | 7.3% (+/-8.3%) | 6.7% (+/-7.5%) | 1.6% (+/-1.5%) | 1.4% (+/-1.7%) |
| LIMARC | *Limenitis archippus* | 5.1% (+/-6.6%) | 5.0% (+/-6.3%) | 1.2% (+/-2.2%) | 0.8% (+/-1.9%) |
| LIMART | *Limenitis arthemis* | 1.7% (+/-5.2%) | 2.3% (+/-5.0%) | 0.7% (+/-1.9%) | 0.4% (+/-1.7%) |
| LIMLOR | *Limenitis lorquini* | 7.2% (+/-8.9%) | 7.1% (+/-8.8%) | 3.8% (+/-3.3%) | 3.6% (+/-3.3%) |
| LYCDOR | *Lycaena dorcas* | -1.0% (+/-1.7%) | -0.8% (+/-2.4%) | -0.2% (+/-0.5%) | 0.0% (+/-0.1%) |
| LYCPHL | *Lycaena phlaeas* | 5.1% (+/-6.2%) | 4.2% (+/-5.5%) | 0.9% (+/-1.7%) | 0.6% (+/-2.0%) |
| NEOMEN | *Neophasia menapia* | 7.5% (+/-7.8%) | 6.6% (+/-8.5%) | 3.7% (+/-3.1%) | 3.3% (+/-3.0%) |
| NYMANT | *Nymphalis antiopa* | 2.7% (+/-5.1%) | 2.5% (+/-4.5%) | 1.1% (+/-2.4%) | 0.7% (+/-2.4%) |
| NYMCAL | *Nymphalis californica* | 6.3% (+/-5.7%) | 6.4% (+/-5.6%) | 0.8% (+/-0.5%) | 1.1% (+/-0.8%) |
| NYML-A | *Nymphalis l-album* | -2.7% (+/-6.5%) | -2.5% (+/-6.8%) | -1.0% (+/-2.6%) | -0.4% (+/-1.9%) |
| OCHSYL | *Ochlodes sylvanoides* | 7.7% (+/-8.4%) | 6.4% (+/-7.7%) | 2.5% (+/-2.2%) | 2.6% (+/-2.3%) |
| OENBOR | *Oeneis bore* | -2.3% (+/-4.9%) | -6.4% (+/-7.1%) | -2.7% (+/-3.7%) | -1.2% (+/-4.0%) |
| OENCHR | *Oeneis chryxus* | -1.8% (+/-2.5%) | -0.5% (+/-2.8%) | -0.4% (+/-1.2%) | 0.0% (+/-0.1%) |
| OENJUT | *Oeneis jutta* | -2.7% (+/-5.5%) | -1.9% (+/-3.8%) | -0.5% (+/-1.7%) | -0.1% (+/-0.6%) |
| OENMEL | *Oeneis melissa* | -3.4% (+/-3.7%) | -4.0% (+/-4.3%) | -1.3% (+/-2.1%) | -0.6% (+/-2.1%) |
| OENPOL | *Oeneis polixenes* | 0.3% (+/-5.6%) | -2.8% (+/-5.2%) | -1.3% (+/-2.2%) | -0.5% (+/-1.9%) |
| PAPMAC | *Papilio machaon* | -4.0% (+/-5.1%) | -3.6% (+/-4.1%) | -1.4% (+/-2.9%) | -0.8% (+/-2.9%) |
| PAPZEL | *Papilio zelicaon* | 4.3% (+/-4.3%) | 4.3% (+/-4.1%) | 1.7% (+/-1.6%) | 1.5% (+/-1.3%) |
| PARCLO | *Parnassius clodius* | 5.3% (+/-5.0%) | 6.4% (+/-6.1%) | 2.6% (+/-1.8%) | 1.7% (+/-1.4%) |
| PHYTHA | *Phyciodes tharos* | 0.9% (+/-1.4%) | 0.7% (+/-1.2%) | 0.1% (+/-0.1%) | 0.0% (+/-0.2%) |
| PIEOLE | *Pieris oleracea* | 2.8% (+/-2.6%) | 1.8% (+/-2.0%) | 1.1% (+/-2.2%) | 0.5% (+/-2.0%) |
| PIERAP | *Pieris rapae* | 8.2% (+/-9.7%) | 8.2% (+/-9.1%) | 0.6% (+/-1.2%) | 0.8% (+/-1.9%) |
| PLEIDA | *Plebejus idas* | -2.9% (+/-4.2%) | -1.5% (+/-3.9%) | -0.7% (+/-1.6%) | -0.2% (+/-0.9%) |
| PLEMEL | *Plebejus melissa* | -0.3% (+/-4.7%) | -2.0% (+/-3.4%) | -3.0% (+/-3.1%) | -2.5% (+/-2.4%) |
| POAHOB | *Poanes hobomok* | 3.6% (+/-3.8%) | 3.9% (+/-4.3%) | 2.6% (+/-2.8%) | 2.0% (+/-2.8%) |
| POLMYS | *Polites mystic* | 3.3% (+/-3.7%) | 2.1% (+/-2.8%) | 1.1% (+/-1.8%) | 0.9% (+/-1.8%) |
| POLTHE | *Polites themistocles* | 2.3% (+/-5.7%) | 1.6% (+/-3.8%) | 0.0% (+/-0.3%) | 0.2% (+/-0.3%) |
| POLFAU | *Polygonia faunus* | -0.3% (+/-5.4%) | 0.6% (+/-5.9%) | 0.2% (+/-0.6%) | 0.2% (+/-0.7%) |
| POLGRA | *Polygonia gracilis* | -0.5% (+/-1.7%) | -0.2% (+/-2.2%) | 0.0% (+/-0.1%) | 0.1% (+/-0.3%) |
| POLINT | *Polygonia interrogationis* | 5.2% (+/-6.2%) | 3.4% (+/-4.5%) | 3.1% (+/-3.3%) | 2.4% (+/-2.7%) |
| POLPRO | *Polygonia progne* | 3.6% (+/-4.2%) | 4.2% (+/-5.3%) | 0.4% (+/-0.7%) | 0.3% (+/-0.9%) |
| POLSAT | *Polygonia satyrus* | 0.5% (+/-2.2%) | 1.3% (+/-2.2%) | 0.0% (+/-0.1%) | 0.1% (+/-0.6%) |
| PONOCC | *Pontia occidentalis* | -1.7% (+/-2.8%) | -1.1% (+/-2.5%) | -0.6% (+/-1.7%) | -0.2% (+/-1.2%) |
| PTECAN | *Pterourus canadensis* | 0.1% (+/-7.0%) | 0.8% (+/-7.4%) | 0.5% (+/-1.2%) | 0.2% (+/-0.8%) |
| PTEEUR | *Pterourus eurymedon* | 8.4% (+/-9.6%) | 6.9% (+/-8.5%) | 3.1% (+/-2.5%) | 3.0% (+/-2.6%) |
| PTEGLA | *Pterourus glaucus* | -0.6% (+/-0.6%) | 0.1% (+/-1.2%) | -3.6% (+/-0.8%) | -0.2% (+/-0.4%) |
| PTERUT | *Pterourus rutulus* | 9.7% (+/-9.2%) | 8.3% (+/-9.4%) | 3.4% (+/-3.2%) | 2.8% (+/-2.9%) |
| SATLIP | *Satyrium liparops* | -0.7% (+/-3.2%) | -0.6% (+/-2.8%) | 1.6% (+/-1.5%) | 1.8% (+/-2.0%) |
| STRMEL | *Strymon melinus* | 3.4% (+/-5.1%) | 5.9% (+/-6.1%) | 0.1% (+/-0.1%) | 0.2% (+/-0.3%) |
| THOPYL | *Thorybes pylades* | -1.9% (+/-5.5%) | -2.0% (+/-5.3%) | -0.6% (+/-1.4%) | -0.3% (+/-1.6%) |
| THYLIN | *Thymelicus lineola* | 2.5% (+/-4.3%) | 3.7% (+/-6.5%) | 0.9% (+/-1.0%) | 0.6% (+/-0.9%) |
| VANVIR | *Vanessa virginiensis* | 4.8% (+/-6.4%) | 2.8% (+/-3.9%) | 3.0% (+/-3.3%) | 2.7% (+/-3.2%) |

**Supplemental Table S2.** Model comparison metrics for models predicting average occupancy shift in our study region since the 1970s. Results are shown for the 100-kilometer scale, temperature analysis. Top candidate model is underlined and in bold. We considered models equivalent if the difference in ELPD +/- the standard error of the ELPD estimate overlapped zero and favored models which used fewer predictors if this was the case.

|  |  |  |  |
| --- | --- | --- | --- |
| **MODEL** | **ELDP-LOO** | **DELTA ELPD** | **SE DELTA ELPD** |
| A | 145.6 | -5.1 | 2.5 |
| B | 147.0 | -3.7 | 3.2 |
| **C** | **150.7** | **0.0** | **0.0** |
| **D** | **150.1** | **-0.6** | **1.5** |
| E | 144.7 | -6.1 | 2.9 |
| F | 145.9 | -4.9 | 3.6 |
| **G** | **147.5** | **-3.3** | **3.8** |
| **H** | **147.1** | **-3.7** | **4.0** |
| I | 144.1 | -6.6 | 2.6 |
| J | 145.6 | -5.1 | 3.3 |
| K | 144.3 | -6.4 | 2.9 |
| L | 145.0 | -5.8 | 2.6 |
| M | 146.0 | -4.7 | 4.3 |
| N | 145.1 | -5.7 | 4.4 |

**Supplemental Table S3** Model comparison metrics for models predicting average occupancy shift since the 1970s. Results are shown for the 200-kilometer scale, temperature analysis. Top candidate model is underlinedand bold. We considered models equivalent if the difference in ELPD +/- the standard error of the ELPD estimate overlapped zero and favored models which used fewer predictors if this was the case.

|  |  |  |  |
| --- | --- | --- | --- |
| **MODEL** | **ELDP-LOO** | **DELTA ELPD** | **SE DELTA ELPD** |
| A | 147.5 | -8.3 | 3.3 |
| B | 150.9 | -5.0 | 4.1 |
| **C** | **155.8** | **0.0** | **0.0** |
| **D** | **155.0** | **-0.8** | **1.6** |
| E | 146.5 | -9.3 | 3.8 |
| F | 149.4 | -6.5 | 4.2 |
| G | 148.3 | -7.5 | 3.9 |
| H | 149.9 | -5.9 | 4.3 |
| I | 146.1 | -9.7 | 3.4 |
| J | 149.0 | -6.8 | 4.2 |
| K | 148.9 | -6.9 | 3.2 |
| L | 150.4 | -5.4 | 2.7 |
| M | 149.9 | -6.0 | 4.2 |
| N | 149.3 | -6.5 | 4.3 |

**Supplemental Table S4.** Model comparison metrics for models predicting average occupancy shift since the 1970s. Results are shown for the 100-kilometer scale, precipitation analysis. Top candidate model is underlined and bold. We considered models equivalent if the difference in ELPD +/- the standard error of the ELPD estimate overlapped zero and favored models which used fewer predictors if this was the case.

|  |  |  |  |
| --- | --- | --- | --- |
| **MODEL** | **ELDP-LOO** | **DELTA ELPD** | **SE DELTA ELPD** |
| A | -408.3 | -5.4 | 3.6 |
| B | -408.0 | -5.6 | 2.7 |
| **C** | **-419.1** | **0.0** | **0.0** |
| **D** | **-416.9** | **-1.1** | **1.7** |
| E | -406.5 | -6.3 | 3.5 |
| F | -405.9 | -6.6 | 2.7 |
| G | -404.8 | -7.1 | 3.0 |
| H | -403.9 | -7.6 | 3.7 |
| I | -406.5 | -6.3 | 3.4 |
| J | -405.8 | -6.7 | 3.0 |
| K | -408.5 | -5.3 | 3.2 |
| L | -406.5 | -6.3 | 3.4 |
| M | -406.9 | -6.1 | 4.2 |
| N | -404.7 | -7.2 | 4.4 |

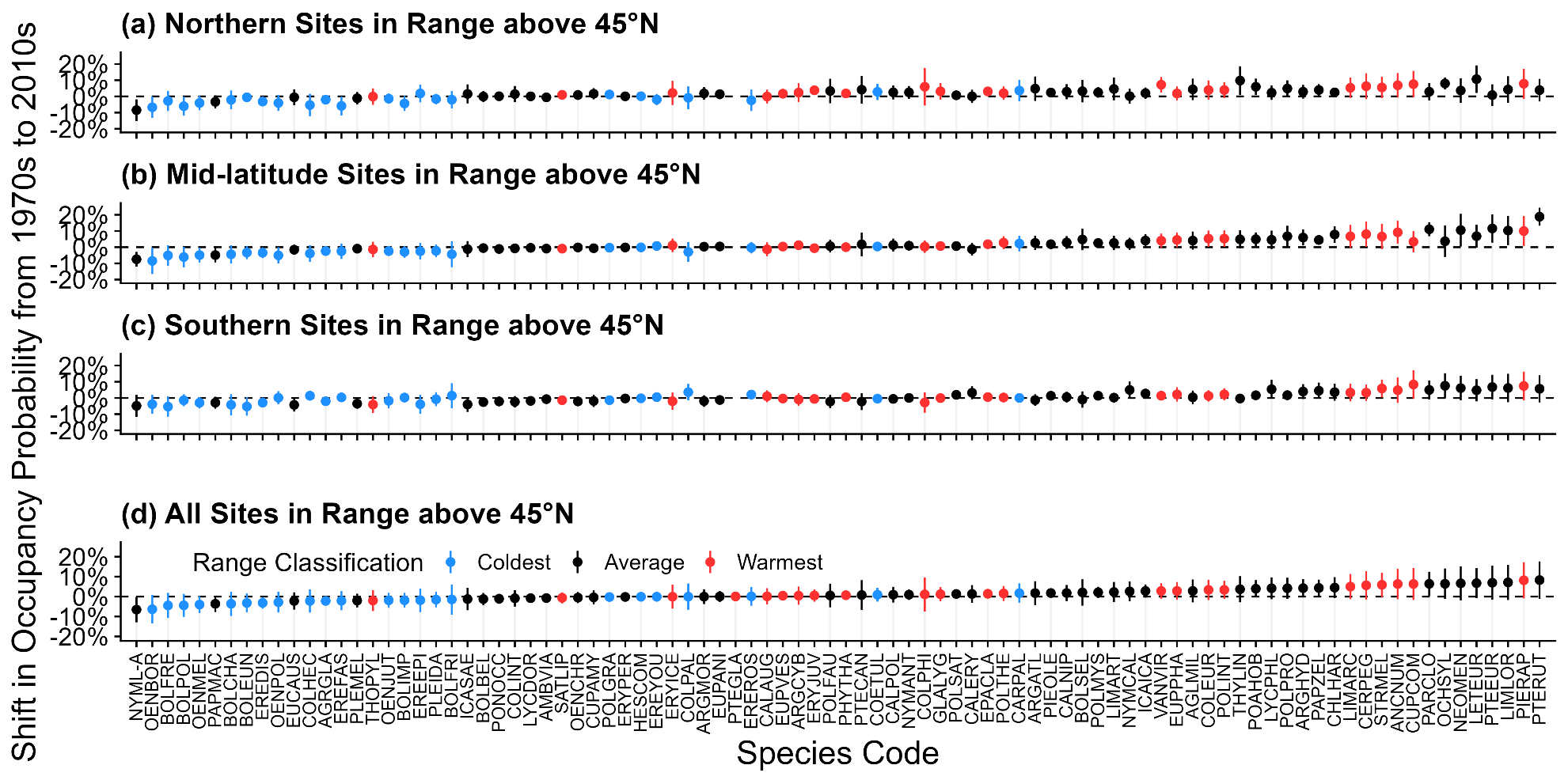
**Supplemental Table S5.** Model comparison metrics for models predicting average occupancy shift since the 1970s. Results are shown for the 200-kilometer scale, precipitation analysis. Top candidate model is underlined and bold. We considered models equivalent if the difference in ELPD +/- the standard error of the ELPD estimate overlapped zero and favored models which used fewer predictors if this was the case.

|  |  |  |  |
| --- | --- | --- | --- |
| **MODEL** | **ELDP-LOO** | **DELTA ELPD** | **SE DELTA ELPD** |
| A | 225.0 | -6.9 | 3.5 |
| B | 226.1 | -5.8 | 3.6 |
| **C** | **231.9** | **0.0** | **0.0** |
| **D** | **231.6** | **-0.3** | **1.3** |
| E | 225.0 | -6.9 | 3.7 |
| F | 225.4 | -6.5 | 3.5 |
| G | 226.1 | -5.8 | 3.5 |
| H | 226.1 | -5.8 | 3.7 |
| I | 224.4 | -7.5 | 3.7 |
| J | 225.3 | -6.6 | 4.0 |
| K | 225.1 | -6.9 | 3.1 |
| L | 224.8 | -7.1 | 3.7 |
| **M** | **228.2** | **-3.7** | **4.7** |
| **N** | **227.2** | **-4.7** | **4.7** |

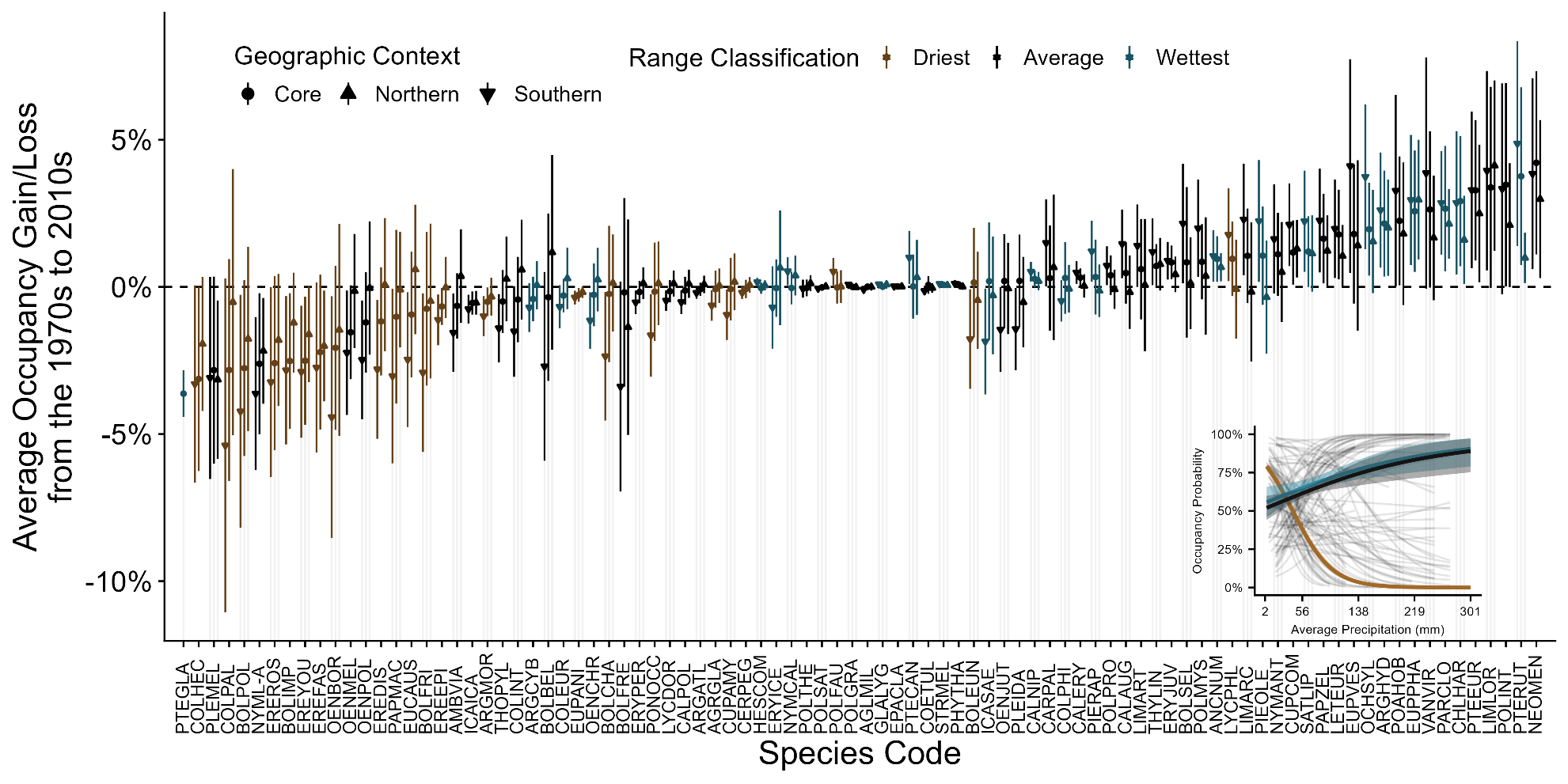
**Graphical user interface, application

Description automatically generated**

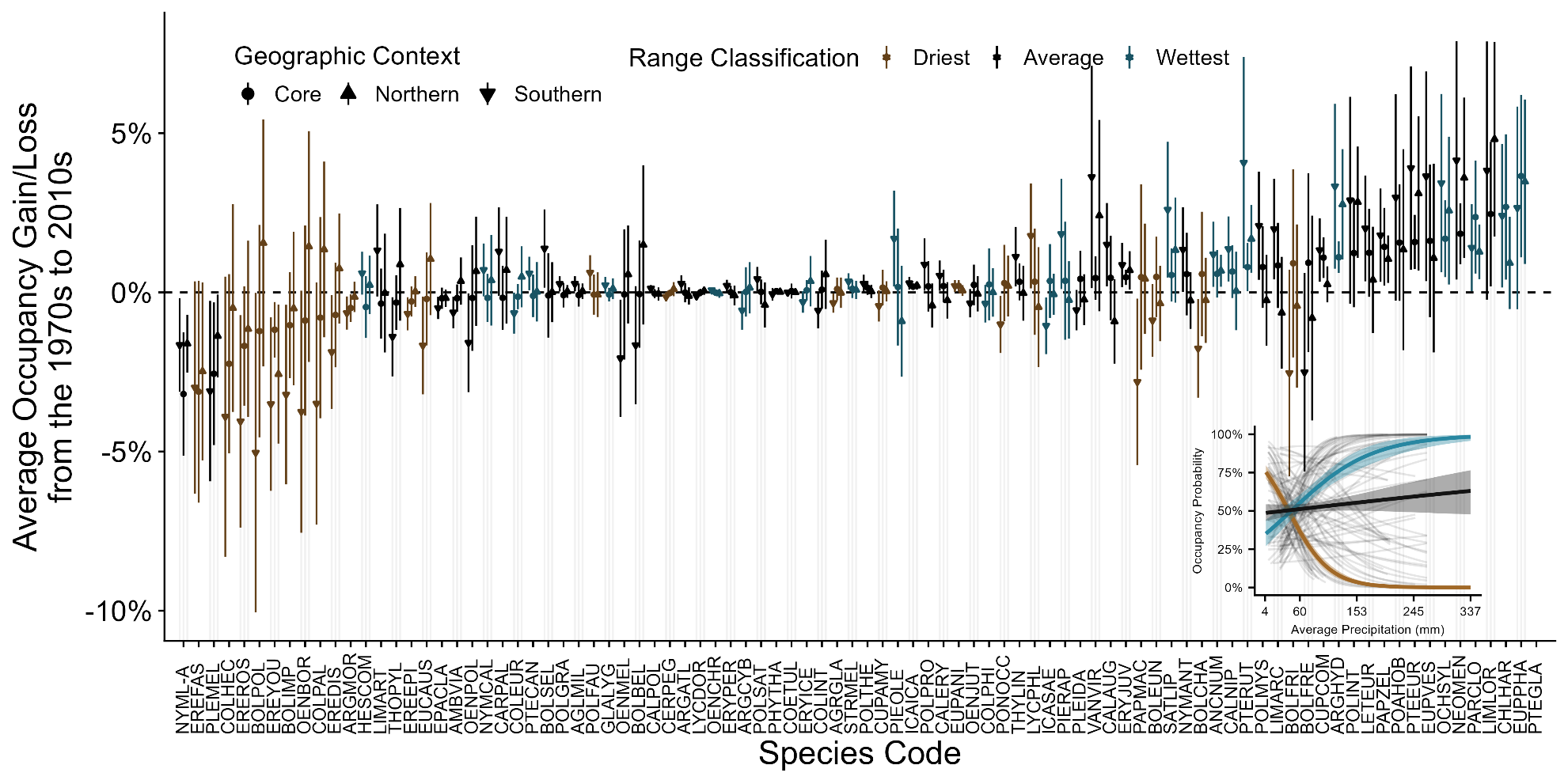
**Supplemental Figure S1.** A summary of the methodological workflow used in this study from data compilation to non-detection imputation and occupancy-detection modeling. The post-hoc analysis using species traits and phylogeny is also visualized.

****

**Supplemental Figure S2.** Species-specific occupancy shifts from the 1970s to 2010s where points indicate the mean occupancy shift and lines indicate one standard deviation of the variation in occupancy shift among (a) the northernmost third of cells within that species’ range in the study region, (b) the mid-latitude third of sites within that species’ range, (c) the southernmost third of sites within that species’ range, and (d) all of the sites within that species’ range (all ranges truncated at 45°N). Species are colored by quartiles of the average annual temperature in their North American ranges, so it is directly comparable to Figure 2.Results shown here are from the 200-kilometer temperature model. Note that there were too few sites to model *Pterourus glaucus* at this scale*.*

****

**Supplemental Figure S3.** Species-specific occupancy shifts from the 1970s where points indicate the average occupancy shift for a given geographic context (core/mid-latitude, southern, and northern grid cells) and lines indicate one standard deviation of variation among relevant sites. The inset panel illustrates the relationship between precipitation and occupancy probability for each species (thin, grey lines), the middle 50% of butterflies (black line), the species with the wettest quarter of ranges (blue line), and the species with the driest quarter of ranges (brown line). Only precipitation that each species has experienced in its range are shown by the species-specific lines. Results shown here are from the 100-kilometer precipitation model.

****

**Supplemental Figure S4.** Species-specific occupancy shifts from the 1970s where points indicate the average occupancy shift for a given geographic context (core/mid-latitude, southern, and northern grid cells) and lines indicate one standard deviation of variation among relevant sites. The inset panel illustrates the relationship between precipitation and occupancy probability for each species (thin, grey lines), the middle 50% of butterflies (black line), the species with the wettest quarter of ranges (blue line), and the species with the driest quarter of ranges (red line). Only precipitation that each species has experienced in its range are shown by the species-specific lines. Results shown here are from the 200-kilometer precipitation model

Chart

Description automatically generated

**Supplemental Figure S5.** Model predicted relationship between species range-wide average annual precipitation and occupancy trend (Model C). Inference is based on trends from the 200-kilometer temperature occupancy analysis.

Chart

Description automatically generated

**Supplemental Figure S6.** Model predicted relationship between species range-wide average annual precipitation and occupancy trend (Model C). Inference is based on trends from the 100-kilometer precipitation occupancy analysis.

Chart

Description automatically generated

**Supplemental Figure S7.** Model predicted relationship between species range-wide average annual precipitation and occupancy trend (Model C). Inference is based on trends from the 200-kilometer precipitation occupancy analysis.