ECOGRAPHY

SEASONAL MACRO-DEMOGRAPHY OF NORTH AMERICAN BIRD POPULATIONS REVEALED THROUGH CITIZEN SCIENCE MONITORING

Journal:	*Ecography*
Manuscript ID	ECOG-07349.R3
Wiley - Manuscript type:	Research Article
Keywords:	demography, eBird, productivity, recruitment, survival, population dynamics

Abstract: Avian population sizes fluctuate and change over vast spatial scales, but the mechanistic underpinnings remain poorly understood. A key question is whether spatial and annual variation in avian population dynamics is driven primarily by variation in breeding season recruitment or by variation in overwinter survival. We present a method using large-scale eBird citizen-science data to develop species-specific indices of net population change as proxies for survival and recruitment gain, based on twice-annual, rangewide snapshots of relative abundance in spring and fall. We demonstrate the use of these indices by examining spatially explicit annual variation in survival and recruitment gain in two well-surveyed nonmigratory North American species, Carolina Wren (*Thryothorus ludovicianus*) and Northern Cardinal (*Cardinalis cardinalis*). We show that, while interannual variation in both survival and recruitment gain is slight for Northern Cardinal, eBird abundance data reveal strong and geographically coherent signals of interannual variation in the overwinter survival of Carolina Wren. As predicted, variation in wintertime survival dominates overall interannual population fluctuations of wrens and is correlated with winter temperature and snowfall in the northeastern United States, but not the southern United States. This study demonstrates the potential of volunteer-collected big datasets like eBird for inferring variation in demographic rates and introduces a new complementary approach towards illuminating the macrodemography of North American birds at comprehensive continental extents.

SEASONAL MACRO-DEMOGRAPHY OF NORTH AMERICAN BIRD POPULATIONS REVEALED THROUGH CITIZEN SCIENCE MONITORING

Abstract

Avian population sizes fluctuate and change over vast spatial scales, but the mechanistic underpinnings remain poorly understood. A key question is whether spatial and annual variation in avian population dynamics is driven primarily by variation in breeding season recruitment or by variation in overwinter survival. We present a method using large-scale eBird citizen-science data to develop species-specific indices of net population change as proxies for survival and recruitment, based on twice-annual, rangewide snapshots of relative abundance in spring and fall. We demonstrate the use of these indices by examining spatially explicit annual variation in survival and recruitment in two well-surveyed nonmigratory North American species, Carolina Wren *(Thryothorus ludovicianus)* and Northern Cardinal *(Cardinalis cardinalis)*. We show that, while interannual variation in both survival and recruitment is slight for Northern Cardinal, eBird abundance data reveal strong and geographically coherent signals of interannual variation in the overwinter survival of Carolina Wren. As predicted, variation in wintertime survival dominates overall interannual population fluctuations of wrens and is correlated with winter temperature and snowfall in the northeastern United States, but not the southern United States. This study demonstrates the potential of volunteer-collected big datasets like eBird for inferring variation in demographic rates and introduces a new complementary approach towards illuminating the macrodemography of North American birds at comprehensive continental extents.

Key words: demography, eBird, population dynamics, productivity, recruitment, survival, vital rates, weather

Introduction

Global loss of biodiversity is an alarming trend and escalating crisis (Ceballos et al. 2020, Cafaro et al. 2022, Habibullah et al. 2022, Finn et al. 2023). Despite diverse, interdisciplinary research efforts to understand biodiversity loss, the underlying demographic mechanisms behind species declines are poorly understood (Faaborg et al. 2010b, Knudsen et al. 2011). Whereas a multitude of environmental factors potentially impact populations (ranging from habitat degradation and climate change to invasive species, pollution, and pesticide use), a basic understanding of which specific environmental factors play a role, and which specific periods in the annual cycle drive population change, remains elusive for most species. Exemplifying the broad-scale biodiversity crisis, North America is estimated to support 3 billion fewer birds today than in 1970 (Rosenberg et al. 2019), a 29% decline in the continental avifauna. Species declines are observed across all ecoregions and biomes, suggesting that research conducted at a continental extent is needed to understand the underlying causes. Despite the scale of species declines, to date mostly micro-demographic field studies have been used to measure the underlying recruitment and mortality patterns of birds (e.g., see references therein: Tian and Hua 2023, Maresh Nelson et al. 2024). In addition, large-scale capture-recapture programs have been established to provide demographic information across larger areas through a huge effort in data collection, including the MAPS program (Monitoring Avian Population and Survivorship; Desante et al. 1995), MoSI (Monitoring Overwintering Survival; DeSante et al. 2005) and CES (Constant Effort Sites; Peach et al. 1996). Both local field studies and large-scale capture-recapture programs have important limitations, however, both taxonomically (limited number of species captured in sufficient numbers) and geographically (uneven distribution of field studies and banding stations). Finding broader-scale metrics of avian demography that complement existing programs and can leverage big data from citizen-

science monitoring efforts could therefore greatly increase our understanding of species decline and loss.

Compared to long-term population changes, within-year changes in population size in response to seasonal mortality and reproduction can be orders of magnitude larger. Identifying environmental factors impacting seasonal population changes can provide valuable insights into major limiting factors, and be a steppingstone towards identifying mechanisms relevant for long-term change. A major unanswered question, however, is whether limiting factors occur primarily during the breeding or non-breeding season, a dichotomy termed the Tap vs. Tub hypotheses (Sæther et al. 2004). Using the analogy of a bathtub, its water level (i.e. population size) is determined both by how much water flows in from the tap (i.e. recruitment) and how much water drains out of the tub (i.e. non-breeding mortality). According to the Tub hypothesis (Lack 1954), fluctuations in population size are closely tied to environmental conditions during the non-breeding season that determine the number of birds that survive this critical period. Many factors can affect non-breeding season mortality, which can depend on winter harshness directly, or be mediated through variation in food availability and associated density dependent competition (Newton 1998, Marra et al. 2015). In contrast, the Tap hypothesis (Sæther et al. 2004) considers annual variation in population size to be determined by factors affecting breeding success and the number of new recruits that enter the population. These factors are equally diverse, including seasonal variation in weather (Dunn et al. 2010), food availability (Martin 1987), predation risk (Lima 2009), and phenological variability in the duration available for breeding (Halupka and Halupka 2017). Most bird populations are monitored during the breeding season only, hampering our ability to determine where and when population change is actually occurring. Existing evidence for either hypothesis is mixed (Balogh et al. 2011, Lamanna et al. 2012, Saracco et al. 2012, Brown et al. 2017, Hallworth et al. 2021), and large-scale studies of avian population change have revealed broad geographic patterns but also

74 substantial local variation (DeSante et al. 1999, Rosenberg et al. 2019, Morrison et al. 2022,
75 Saracco et al. 2022, Fink et al. 2023). Because population trends and demographic
76 mechanisms vary spatially, answering broad questions on the causes of population change
77 likely requires a demographic approach that covers large spatial extents and is spatially explicit.
78 Ideally such an approach should be generalizable to many species, without requiring an
79 unrealistic additional burden in terms of data collection. Identifying large-scale drivers of
80 population change, such as weather and climatic conditions, will also benefit from a macro-
81 demographic approach capable of estimating demographic patterns across equally large spatial
82 extents, by integrating over noisy local variation.

83 In this study, we present a novel approach for estimating avian demographic indices
84 from semi-structured citizen science data that is based on repeated sampling of continent-scale
85 relative abundance. Specifically, we explore the potential for leveraging big data from the large
86 and rapidly growing eBird database (Sullivan et al. 2014), exploiting its extensive geographic,
87 temporal, and taxonomic scope that is currently unrivaled by any other monitoring program (La
88 Sorte and Somveille 2020). We apply our methodology to two nonmigratory North American
89 species with well-known and contrasting demographic patterns: Carolina Wren (*Thryothorus*
90 *ludovicianus*), which shows large annual population fluctuations (Ziolkowski et al. 2023) in
91 response to harsh winter conditions (Brooks 1936, Root 1988, Sauer et al. 1996, Mehlman
92 1997, La Sorte and Thompson 2007, Link and Sauer 2007), and Northern Cardinal (*Cardinalis*
93 *cardinalis*) which is expected to be more winter-hardy due to its granivorous diet and urban
94 adaptation (Evans et al. 2015). Both species have undergone northward range expansions
95 (Beddall 1963). We explicitly test whether the new eBird-derived indices detect substantial
96 interannual variation in overwinter survival related to harsh winters at more northerly latitudes
97 (Tub dynamics) in the Carolina Wren but not in Northern Cardinal. In addition, we hypothesize
98 that we would detect additional temperature-associated effects on recruitment in both Carolina

Wren and Northern Cardinal, and that the direction of these effects would vary latitudinally, with warm temperatures detrimental near warm range margins and beneficial near cool range margins (Socolar et al. 2017). While there are many other factors affecting survival and recruitment, we do not intend to explore all these factors and their relationship with demography exhaustively. Instead, we aim to present a proof of concept demonstrating that citizen-science data can provide information on the survival and recruitment of avian populations. Through revealing well-established patterns with our new macrodemographic indices, we lay the groundwork for exploring demography across hundreds of resident and potentially migratory species in North America, contributing new insights into species-specific limiting factors and potential causes of declines.

Methods

General approach

Our approach is based on tracking within-year population change through repeated sampling of relative abundance, as illustrated in Figure 1. We used data from eBird (Sullivan et al. 2014) to derive time series of avian abundance indices across eastern North America with uncertainty. We computed these indices over the cells of a hexagonal grid to obtain spring and fall "snapshots" of bird abundance, corresponding to weeks 13-16 (roughly the month of April) and weeks 40-43 (roughly the month of October) of the calendar year. These population snapshots allow us to define an index of recruitment as the logarithm of the ratio of the autumn population index over that of the preceding spring, and a survival index as the logarithm of the ratio of the spring index over that of the preceding autumn. Since our population indices are confounded by seasonal changes in detectability, they cannot be used to directly compare population sizes between spring and autumn. Therefore, we focus on the annual variation in indices for

recruitment and survival within a given cell. In assuming that this temporal variation is meaningful, we do not have to assume that detection probabilities are similar in April and October or across space. Instead, our key assumption is that the detectability difference between April and October shows up in our population snapshots as an unknown multiplicative factor that is potentially variable across space but constant through time. Under these assumptions, our demographic indices can provide a consistent index of high versus low survival or recruitment years even if it does not provide an absolute measure (Figure 1). Furthermore, given that annual population fluctuations are expected to be approximately log-normal, the relative importance of summer versus winter in governing interannual population fluctuations is directly related to the relative variances in the recruitment and survival indices, and this remains true despite that the detectability difference between seasons is unknown. We should note that these “macro”-demographic indices differ from traditional definitions of recruitment and survival used for example in mark-recapture studies. Our indices represent net gain and net loss in population size during pre-defined breeding and non-breeding periods, respectively. As such, the recruitment (gain) index captures both the effect of population increases due to reproduction and, as an index of net population change, also includes some co-occurring breeding season mortality. The survival (loss) index, on the other hand, captures the effect of non-breeding season mortality only, as no reproduction occurs during non-breeding periods. While seasonal redistribution through immigration and emigration may potentially affect our indices as well, their combined effect is likely small, given high site fidelity and small dispersal distances for most species compared to the spatial scale of our analysis (Haggerty and Morton 2020, Halkin et al. 2021).

To test our Tap and Tub hypotheses, we ask whether temporal variance in the survival index is larger or smaller than the variance in the recruitment index. The index with the larger variance will dominate the overall variance of the time series of the log absolute population size,

so this question amounts to asking whether interannual population fluctuations around the mean trend are more strongly controlled by events during May-September (Tap dynamics) or during November-March (Tub dynamics). Second, we test whether observed variation in recruitment and survival indices is correlated with weather conditions during the relevant period. Third, we ask whether the observed patterns are geographically coherent. When adjacent grid cells are analyzed independently but display geographical coherence in their results, we gain a measure of confidence that the patterns we observe are genuine and not affected by the unstructured sampling inherent in eBird data.

Geographic, temporal, and taxonomic scope

We focused our analysis on the eastern United States and Canada (east of 110° W longitude and south of 50° N latitude) between the years 2006 and 2019. This region contains a high density of eBird checklists going back to the earliest years of our time series and represents a coherent biogeographic unit that contains well-defined populations of multiple resident bird species. We chose 2006 as our start year to maximize the length of our time series while avoiding extremely sparse and potentially unreliable data from earlier years. We chose 2019 as our end year to avoid potentially abrupt changes in the observation process associated with the Covid-19 pandemic (Hochachka et al. 2021).

We considered two species in our analysis: Carolina Wren and Northern Cardinal. These are widespread species that are commonly reported in eBird and are year-round residents within our study area (Haggerty and Morton 2020, Halkin et al. 2021). Year-round residency enabled us to derive regionally specific population snapshots from well-spaced times of year, and to assess the influence of weather on recruitment and survival without the complications of migration. Carolina Wren and Northern Cardinal also provide a useful contrast in that Carolina

Wren displays large annual population fluctuations and the Northern Cardinal does not based on data from the North American Breeding Bird Survey (Sauer and Link 2011).

Bird occurrence and abundance data

We subsetted the eBird data to retain only complete checklists (i.e., observers report all the species they detect and identify) submitted under stationary or traveling protocols of between 0 and 3 km distance, between 5 and 60 minutes duration, and with a checklist calibration index (CCI, a measure of observer efficiency; Johnston et al. 2018) of at least 0 (i.e., above average).

Recruitment and survival indices

To explore spatial variation in population dynamics, we computed indices of recruitment and survival over the cells of an approximately equal-area hexagonal grid with a roughly 285 km spacing between cell centers (Figure S1) (Barnes et al. 2017). Inter-annual fluctuations in North American bird population sizes are approximately multiplicative (Kalyuzhny et al. 2014a), and are only weakly stabilized (Kalyuzhny et al. 2014b), such that populations fluctuate via multiplicative events that are roughly independent of the population size (in contrast, fluctuations in strongly stabilized populations tend to be positive when populations are small and negative when populations are large) (Lande et al. 2003). In this case it is natural to work on the logarithmic scale, where the multiplicative constant becomes an additive constant with no influence on variances or regression slopes. Gradual changes in detectability over multi-year timescales (e.g., due to changes in the eBird user base) are acceptable as long as they impact detection similarly in April and October and as long as they are sufficiently slow so that detection probabilities are always similar in consecutive Aprils and consecutive Octobers.

196 With an eye towards scalability, we eschew the computationally costly machine-learning
197 approaches commonly used to analyze eBird data at a fine landscape level, and instead derive
198 our population indices based on average per-checklist counts with carefully propagated
199 uncertainty. Modeling fine-scale habitat features was not considered a priority since our analysis
200 is focused on macroscale patterns derived from seasonal changes in abundance within the
201 same habitat regions. To support robust uncertainty quantification, we spatially subsampled
202 each grid cell across a finer hexagonal grid with an 18 km spacing between cell centers (Figure
203 S1). We refer to the spatio-temporal units comprising one cell of the fine grid and one four-week
204 period during one year as *micro-cells*, and we refer to the units comprising one cell of the
205 coarse grid and one four-week period during one year as *macro-cells*.

206

207 To compute population indices for the macro-cells, we first computed population indices
208 for each micro-cell as the average per-checklist count of the focal species across all checklists
209 meeting our criteria for inclusion. Let $A_{s,y,c,x,k}$ be the k-th checklist of in total $n_{s,y,c,x}$ checklists
210 belonging to the x^{th} micro-cell, c^{th} macro-cell in year y and season s. Then, the population index
211 $i_{s,y,c,x}$ for the x^{th} micro-cell in the c^{th} macro-cell in the year y and season s is given by:

212

213 $$i_{s,y,c,x} = \sum_{k}^{n_{s,y,c,x}} A_{s,y,c,x,k} / n_{s,y,c,x}$$

214 Some complete checklists in eBird report the presence of a species but not its count; for these
215 checklists we imputed the mean of the remaining checklists in the micro-cell that reported
216 positive counts of the focal species. We excluded micro-cells that contained no complete
217 checklists and micro-cells that contained at least one complete checklist reporting presence-
218 only but no complete count-based checklists reporting nonzero abundance.

We then take our population index for the macro-cell $I_{s,y,c}$ to be the logarithm of the mean index over the total of $n_{s,y,c}$ constituent micro-cells:

$$I_{s,y,c} = \log\left(\sum_{x}^{n_{s,y,c}} i_{s,y,c,x} / n_{s,y,c}\right)$$

We quantified uncertainty in macro-cell level population indices $I_{s,y,c}$ by applying the Bayesian bootstrap (Rubin 1981) over the constituent micro-cells. That is, we generated posterior samples $I^{b}_{s,y,c}$ as:

$$I^{b}_{s,y,c} = \log\left(\sum_{x}^{n_{s,y,c}} i_{s,y,c,x} \cdot w_{b,x} / \sum_{x}^{n_{s,y,c}} w_{b,x}\right)$$

Where *b* indexes the bootstrap replicate and $w_{b,x}$ represents the weight assigned to the x^{th} micro-cell in the b^{th} bootstrap replicate, sampled from a Dirichlet distribution.

We then calculated the demographic index for recruitment $Y^{b}_{recruitment,y,c}$ and survival $Y^{b}_{survival,\,y,c}$ indices by subtracting the spring population index from the subsequent fall index, and the fall index from the next spring index, respectively:

$$Y^{b}_{recruitment,y,c} = I^{b}_{fall,y,c} - I^{b}_{spring,y,c}$$

$$Y^{b}_{survival,y,c} = I^{b}_{spring,\,y+1,c} - I^{b}_{fall,y,c}$$

Finally, we calculate a mean index $Y_{t,y,c}$ and associated standard deviation $\epsilon_{t,y,c}$ for each type of demographic index t $\in$ {recruitment, survival} by calculating the mean and standard deviation across bootstrap replicates.

We fully propagated the posterior uncertainty by performing these subtractions sample-wise through the bootstrapped posteriors. This procedure occasionally produces infinite indices when no individuals are reported in an entire macro-cell ($Y^b_{recruitment,y,c}$ and $Y^b_{survival,y,c}$ can become become ±infinity when all micro-cell counts $i_{s,y,c,x}$ in a macro-cell are zero due to the log-transform in $I^b_{s,y,c}$; in this limit the bootstrap no longer accurately quantifies uncertainty). These infinite indices were rare and primarily occurred in the early years with lower sampling effort and at range boundaries where the species is less common. We removed infinite indices in all downstream analysis, as they are rare and universally reflect sampling variation that our bootstrapping does not capture. In regressions of recruitment and survival indices against weather variables, we simply removed years with infinite indices from analysis. In comparisons of the variance in survival versus the variance in recruitment, we excluded entire cells if they yielded infinite indices in any year.

Comparison with the North American Breeding Bird Survey

To validate the use of eBird for the study of population fluctuations, we used similar methods to create eBird-derived indices of annual June-to-June population change and compared these to indices derived from the North American Breeding Bird Survey (BBS). BBS data are sparse at the level of our macro-cells, and so for the purposes of this analysis we replaced the macro-cell in our method with larger bird conservation regions (BCRs), which are a common unit of aggregation in BBS analyses (Sauer et al. 2003). To ensure that our bootstrapping approach covered a reasonable spatial sample of each BCR, we required that data available for at least 100 micro-cells to compute a population snapshot. To compute yearly population snapshots from the BBS, we fit separate generalized additive models for each BCR with random year effects and Poisson error for route-level counts, see supplementary information (SI) equations

9, 10. This approach is similar to that of Edwards and Smith (2020), except that we use fully independent models for each BCR. See SI for additional details of model fitting.

We then regressed June-to-June log-ratios from our BBS time series against those from our eBird time series via a mixed model incorporating the known measurement error in the predictor and the response, with a random intercept by BCR (SI equation 11, 12). We perform this analysis for Carolina Wren only, as the BBS analysis for Northern Cardinal did not show any interannual variation with sufficient confidence to support a validation based on correlations between time series.

Weather data

We obtained weather data from Daymet (Thornton et al. 2022) via Google Earth Engine (Gorelick et al. 2017) and R package 'rgee' (Aybar et al. 2020), summarized as average values over the spatial macro-cells for which we derived recruitment and survival indices. We focused on three weather variables that we believed *a priori* might influence demography: the average daily high temperature during July and August, reflecting temperatures during the hottest part of the summer; the average daily high temperature during January and February, reflecting temperatures during the coldest part of the winter; and the average snowpack, measured in snow water equivalent, during December-March, reflecting winter weather that might impede foraging.

Analysis of variance and regression

To obtain spatially explicit estimates of the factors governing the temporal dynamics of populations we used a regression-based approach applied to each macro-cell separately

followed by a model-based spatial smoothing of the regression results across cells. To determine whether events during the breeding period versus the non-breeding period exert stronger control on population dynamics, we fit independent cell-specific regressions to the distribution of survival and recruitment indices, with season (survival or recruitment) predicting both the mean and the logarithm of the variance (SI equation 1). The coefficient for the effect of season on the variance provides inference about whether the variance is higher overall for recruitment or survival, and therefore whether overall population fluctuations are primarily under the control of events during May-September or November-March.

To determine if weather conditions drive variation in recruitment or survival, we regressed the indices against mean daily high temperatures during July and August (recruitment indices), mean daily high temperatures during January and February (survival indices), and mean snow-water equivalent during January and February (survival indices). We fit these regressions independently for each cell. To propagate uncertainty in the demographic indices through the regressions we fit both a homoskedastic residual and an additional Gaussian error term whose variance we fixed to the bootstrapped standard error for the measurement (SI equation 3, 4). We fit the regression models only in cells where we retained at least five analyzable years. For the comparisons of variance, we additionally removed all cells where the estimated demographic index was infinite in any year, as ignoring these years could substantially bias variance estimates.

We then spatially smoothed the slope estimates from the regressions of demographic indices against weather variables using both a nonspatial residual and a cell-level spatial random effect with an intrinsic conditional autoregressive (ICAR) prior (Morris et al. 2019) . We again propagated uncertainty by including an extra Gaussian error term whose variance we fixed to the posterior variance in the slope estimate (SI equation 5,6). We present both the conditional autoregressive (CAR) model and the independent cell-specific results. The CAR

model provides spatial smoothing over cells and distinguishes spatially coherent patterns from spatially idiosyncratic patterns that might merely reflect cell-specific noise. The independent cell-specific regressions serve to confirm that the patterns detected in the CAR models reflect genuine geographic coherence across multiple independently analyzed cells and not merely the spatially smoothed influence of single cells that take extreme values. Thus, the independent cell-specific regressions are effectively a robustness check against the possibility that the CAR model produces spurious patterns due to misspecification of an insufficiently heavy-tailed residual term.

Latitudinal pattern

The regression models described above suggested that Carolina Wren populations might be under the control of winter temperatures at high latitudes but not at low latitudes. To further investigate and quantify this pattern, we fit an additional CAR model to predict the slope of Carolina Wren survival regressed on temperature as a function of latitude. We again used a nonspatial Gaussian error term, an ICAR spatial term, and a known error term to propagate uncertainty, but we modified the model described above by additionally including a fixed effect of latitude (SI equations 6, 8).

Model fitting

We fit all regressions in the Stan C++ library for Bayesian modeling and inference (Stan Development Team 2023) using the R package brms (Bürkner 2017) under the default priors, which are generally weakly informative for intercepts and standard deviations and flat for coefficients. We assessed model convergence via lack of divergences and r-hat statistics less than 1.05. Divergences are a sensitive diagnostic unique to Hamiltonian Monte Carlo algorithms

for model fitting for certain forms of non-convergence that frequently manifest in hierarchical modeling situations (Betancourt 2017). We use a contemporary version of r-hat (split, folded, rank-normalized) with improved sensitivity for detecting non-convergence (Vehtari et al. 2021). We also performed posterior predictive checks for each model and visually assessed the model fit for representative subsets of individual macro-cells (not shown) and for aggregated models (SI Figure. S9). As summary checks, we verified for all macro-cells, whether close to 95% of the observations were within the 95% CI of the posterior model predictions, as expected for good model fit (SI Figure S10). We also checked normality assumptions through assessment of the excess kurtosis and skewness of our sampled macrodemographic indices (SI Figure. S11, S12) and the posterior distributions of the regression coefficients (SI Figure. S13, S14).

Results

After applying our stringent data filters, we obtained analyzable time series (series including at least five seasons) in 59 grid cells (macro-cells) for Carolina Wren survival, 60 grid cells for Carolina Wren recruitment, 66 grid cells for Northern Cardinal survival, and 68 grid cells for Northern Cardinal recruitment. Although the longest time-series were concentrated in the northeastern United States, where early eBird uptake was strongest (SI Figure S2), we were able to analyze time-series spanning latitudes from south Florida to southern Canada.

June-to-June population fluctuations at the level of bird conservation regions were strikingly consistent between our eBird-derived indices and BBS-derived indices, with a clearly positive slope statistically indistinguishable from one (Figure 2).

The comparison of recruitment and survival indices reveals higher variance in survival than recruitment for Carolina Wrens in the northern part of their range but not in the southern part (Figure 3a, SI Figure S3a). The sample variances were higher in winter across much of the

358 continent, with especially the northeastern US showing strong evidence for a significant
359 difference. The geographical coherence in the results, which arises from independent cell-scale
360 analyses without spatial smoothing, strongly suggests that results are not due to the semi-
361 structured sampling in eBird but rather are due to a genuine signal of populations controlled
362 primarily by wintertime dynamics in the northern part of the range. No such pattern was
363 apparent for Northern Cardinal, consistent with the idea that this species is hardier and less
364 sensitive to severe winters, and again suggesting that results for Carolina Wren are not artifacts
365 of the sampling process (Figure 3b, SI Figure S3b). Cardinal populations are potentially under
366 stronger control from breeding-season dynamics rather than wintertime dynamics in southern
367 Texas (see below).

368 Weather conditions were strongly predictive of winter survival in Carolina Wren (SI
369 Figure S4a, Figure S4c), which is substantially controlled by winter harshness as measured by
370 temperature and snowfall, especially in the northeastern part of its range (Figure 4a, Figure 4c).
371 Recruitment showed no robust relationship with summer temperatures (Figure 4e, Figure S4e).
372 In contrast, although Northern Cardinals may have experienced similar patterns in survival
373 (Figure 4b, Figure 4d), the degree of certainty was low (SI Figure S4b, Figure S4d). However,
374 their recruitment was positively associated with summer temperatures in the Mississippi Valley
375 and Texas (Figure 4f, SI Figure S4f).

376 The latitudinal ICAR models estimated a clear positive effect of latitude on the slope of
377 the relationship between winter temperature and Carolina Wren survival of roughly 0.006 (95%
378 credible interval -0.003 - 0.014) natural logarithms per degree Celsius per degree latitude
379 (Figure 5). This slope means that a one-degree difference in winter temperatures that has no
380 effect on Carolina Wren survival rates in the south would have roughly a six percent impact on
381 survival rates (and therefore on populations) 1000 km to the north.

382

383 **Discussion**

384 This study presents a novel method for extracting information on seasonal demographic
385 changes from large-scale citizen science data. We show how semi-structured bird counts
386 collected by volunteers can be used to make inferences on processes of seasonal mortality and
387 recruitment across the full range of two resident species. A feature of eBird that enables large-
388 scale demographic analysis is that population abundances are sampled continuously throughout
389 year. This allows us to estimate indices of seasonal change in observed bird counts throughout
390 the season, as abundances change due to mortality and reproduction. We use annual time
391 series of the post- to pre-breeding count change to sample the process of mortality, while the
392 pre- to post-breeding change samples the combined process of recruitment and adult survival
393 during the breeding period. By assigning annual population fluctuations to different parts of the
394 annual cycle in this manner, we provide a way of characterizing population dynamics along the
395 Tap-Tub dichotomy (Sæther et al. 2004).

396 Our results indicate that Carolina Wren population dynamics are primarily Tub-driven at
397 the northern end of the range, as indicated by a higher interannual variance in the survival index
398 as compared to the recruitment index (Figure 3a). In addition, warmer, less snowy winters are
399 associated with higher survival (Figure 4a,c), which aligns with the known susceptibility of
400 Carolina Wrens to harsh winters (Brooks 1936, Sauer et al. 1996, Mehlman 1997, Link and
401 Sauer 2007). Northern Cardinal exhibits strikingly different patterns, with limited evidence for
402 either Tap or Tub dynamics across most of the range (Figure 3b, tap-dynamics was detected in
403 southern Texas only). Only weak associations of demographic indices with weather covariates
404 were found for this species (Figure 4), though recruitment may be elevated when summers are

405 warm in the Mississippi Valley and Texas. These results support our hypothesis that winter
406 harshness is a key factor in Carolina Wren but not for Northern Cardinal.

407 We find a remarkable correspondence between our findings and the outcomes of
408 previous studies, which adds further support that our demographic inferences on Carolina Wren
409 and Northern Cardinal populations are ecologically meaningful and robust. Previous studies,
410 typically conducted at local scales, have consistently demonstrated the influential role of
411 environmental factors, such as overwintering temperature (Brooks 1936, Root 1988, Mehlman
412 1997, Latimer and Zuckerberg 2021), snow cover (Link and Sauer 2007), food availability (Job
413 and Bednekoff 2011), in shaping Carolina Wren populations. These factors are crucial
414 determinants of winter survival and, consequently, may even dictate the northern distribution
415 range of Carolina Wrens (Huang et al. 2016). For instance, researchers using BBS and
416 Christmas Bird Count (CBC) data have quantified a 1.1% decrease of abundance for every day
417 with 4 cm of snow cover in the northern portion of this species winter range (Link and Sauer
418 2007). This phenomenon may be closely tied to their foraging habits, which primarily occur near
419 or at ground level (Haggerty et al. 2001). Conversely, Northern Cardinals exhibit comparatively
420 less interannual variation in their demography (Ziolkowski et al. 2023), owing to their adaptability
421 to urban forest environments (Evans et al. 2015), which offer more consistent food availability
422 during winter months (Leston and Rodewald 2006), including supplemental feeding (Job and
423 Bednekoff 2011). These corresponding findings suggest that our large-scale demographic
424 information offers a valuable complementary insight, extending the geographic scope of these
425 local studies.

426 Our analysis finds similar annual fluctuations in population abundances as those
427 detected by the more structured and standardized survey of the BBS (Figure 2,SI Figure S8). It
428 shows that our analysis framework has sufficient power to detect relatively modest demographic
429 changes in population numbers within seasons and across years. It also indicates that our

stratified sampling design with bootstrap-based uncertainty quantification successfully accommodate the opportunistic nature of citizen science data collection, which inherently introduces spatial and temporal biases in observer effort (Dickinson et al. 2010, Johnston et al. 2021, Fink et al. 2023, Johnston et al. 2023). Spatial biases may emerge due to observers' preferences for particular locations, such as protected areas (Boakes et al. 2010). An illustrative example of temporal bias can be observed in the alterations to data collection practices during the COVID-19 pandemic (Hochachka et al. 2021). Additionally, inclement weather conditions or poor air quality may constrain birder activity, resulting in fewer checklists on specific days in certain areas. We avoided these spatiotemporal biases by aggregating data into relatively large spatial and temporal sampling units (i.e., equal spaced hexagons), with its uncertainty estimates. Resulting demographic indices and their relationships with weather covariates are often similar to those of neighboring hexagons. This spatial consistency in demographic patterns along with spatiotemporal consistency observed in BBS validation provides further support that our approach is robust against spatiotemporal biases in sampling.

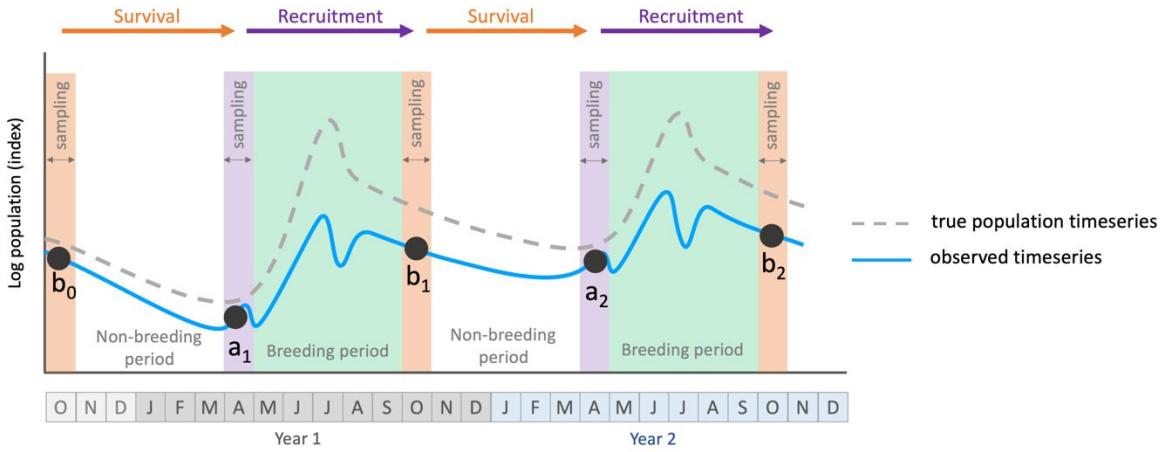
Importantly, detected seasonal changes in counts not only reflects the seasonal change in abundance of species, but also the seasonal change in their detectability. Birds' detectability declines during specific annual cycle events like nesting and molting and increases in spring when birds vocalize more frequently (Wilson and Bart 1985, Riegert 2022). A central assumption in our analysis is that this seasonal detectability change is consistent, meaning that the detectability difference between our population snapshots can be captured by an unknown multiplicative factor that is potentially variable across space, but approximately constant in successive years. Seasonal changes in count can then be interpreted as *relative* indices of demographic population changes that may be compared across years. However, we acknowledge that interannual variation in detectability may affect our estimation. Modeling detectability in eBird either implicitly via the inclusion of detection-related covariates (as in

455 STEM models for eBird data; Fink et al. 2023) or explicitly via occupancy models (Hochachka et
456 al. 2023) or N-mixture models (Goldstein and de Valpine 2022) present interesting avenues for
457 further refinement of our approach.

458 Our macro-demographic approach has potential to expand ecologists' toolkit for
459 obtaining information on demography and overcome several existing challenges. Traditional
460 demographic studies tend to be time consuming and requires banding and resighting individuals
461 in a mark-recapture setting. Detailed demographic studies exist only for a limited number of
462 model species investigated in long-term population research. Extensive banding programs have
463 been initiated to obtain demographic information across large spatial extents and for more
464 species (e.g., MAPS and MoSi in the Americas (Desante et al. 1995, DeSante et al. 2005), and
465 CES schemes in Europe (Peach et al. 1996)). These programs provide invaluable individual-
466 based demographic metrics that remain unrivaled in their specificity and ability to distinguish
467 groups by age and sex. However, they remain limited in their geographic coverage and the
468 number of species that can be sampled in sufficient numbers (Faaborg et al. 2010b). A citizen-
469 science based macro-demographic approach has great potential for complementing existing
470 demographic information. Future work will need to explore how our macro-demographic
471 approach can scale-up to include multiple species over large spatial-temporal extents using
472 freely available and rapidly growing citizen science data. The resulting demographic indices,
473 sampled across large spatial extents instead of at highly local banding sites, are suitable to
474 detect large-scale demographic processes, such as those influenced by large-scale weather
475 and climate, and may prove suitable to detect the influences of other large-scale causal drivers
476 including anthropogenic impacts. Finally, our macro-demographic approach gives us more
477 freedom to temporally subdivide the annual cycle into "snapshots" of interest, allowing us to
478 isolate and study multiple transitions across the annual cycle.

The macro-demographic methodology introduced in this study opens exciting new opportunities in avian ecology research. Expanding this approach to encompass many other resident species, as well as migrants that breed or winter in data rich regions, holds significant potential for uncovering novel insights into population dynamics (Faaborg et al. 2010a, Sullivan et al. 2014, La Sorte et al. 2018). Although, we note that migration poses special challenges of compressed temporal windows for observing population changes pre- and post-breeding, as well as confounding annual variability in migration timing. Still, leveraging the large-scale spatial and temporal coverage of citizen science data allows for a comprehensive examination of annual fluctuations, shedding light on distinct patterns and ecological drivers. Furthermore, its adaptability offers the prospect of exploring finer resolutions within annual cycles, using multiple snapshots to achieve higher time resolution analyses. This opens avenues for researchers to dissect mortality and recruitment dynamics with greater precision, providing a more nuanced understanding of the temporal intricacies within a species' annual life cycle.

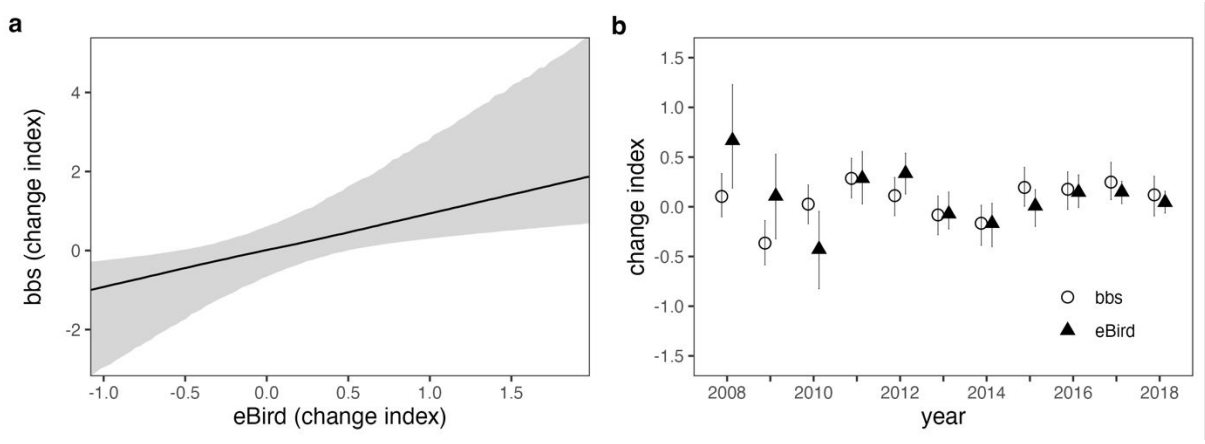
Our case study highlights the potential of citizen-science data in providing demographic information on recruitment, mortality, and its associations with weather and climate. Uncovering these underlying demographic processes will be critical for understanding the causes for demographic boom and bust years, and the mechanisms behind ongoing long-term population declines.



498

499 **Figure 1:** Conceptual overview of our approach for assessing population fluctuations using
500 eBird data. The grey dashed curve represents a hypothetical two-year time series of the
501 logarithm of a population's size, beginning in early fall. The blue curve represents the apparent
502 time series from eBird data, which confounds the population time series with detection effects
503 (e.g., higher detection in spring than fall). We snapshot the eBird time series in fall (circle 'b')
504 and spring (circle 'a'), and we treat the differences between successive snapshots as indices of
505 survival (i.e., $a_1 - b_0$) and recruitment (i.e., $b_1 - a_1$; on the log scale, differences correspond to
506 log-ratios). Because we are interested primarily in the year-to-year variability of these indices
507 and not in their raw values, we can neglect the differences between the apparent log-population
508 and the true log-population provided that these differences are consistent from spring to spring
509 and from fall to fall (a multiplicative detection term becomes an additive term on the log scale).
510 In this example, survival was higher in the second winter than in the first (i.e., $a_2 - b_1 > a_1 - b_0$),
511 and the eBird-derived population snapshots provide an unbiased estimate of the difference
512 between year-1 survival and year-2 survival.

513



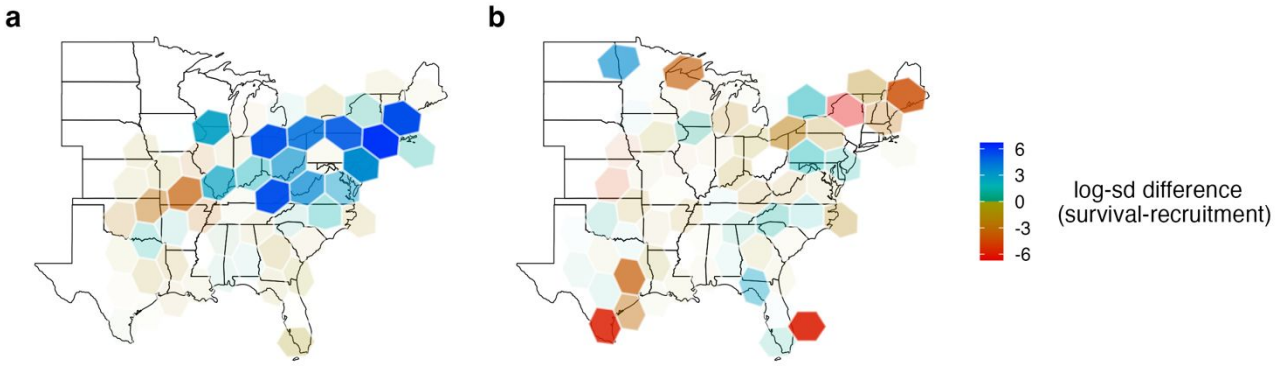
514

515 **Figure 2:** eBird derived indices for June-to-June population fluctuations at the level of bird
516 conservation regions (BCRs) are predictive of fluctuations derived from the United States
517 Breeding Bird Survey (BBS) for the same regions and years. **a**) The slope is estimated to be
518 near unity (0.97, 95% CI 0.34–2.14). **b)** The match in fluctuations through time as visualized for
519 one of the longest and best-aligned time series (BCR 28 includes the Appalachian Mountains
520 from Alabama to southern New York). Data from years prior to 2008 did not meet the inclusion
521 thresholds for eBird analysis.

522

523

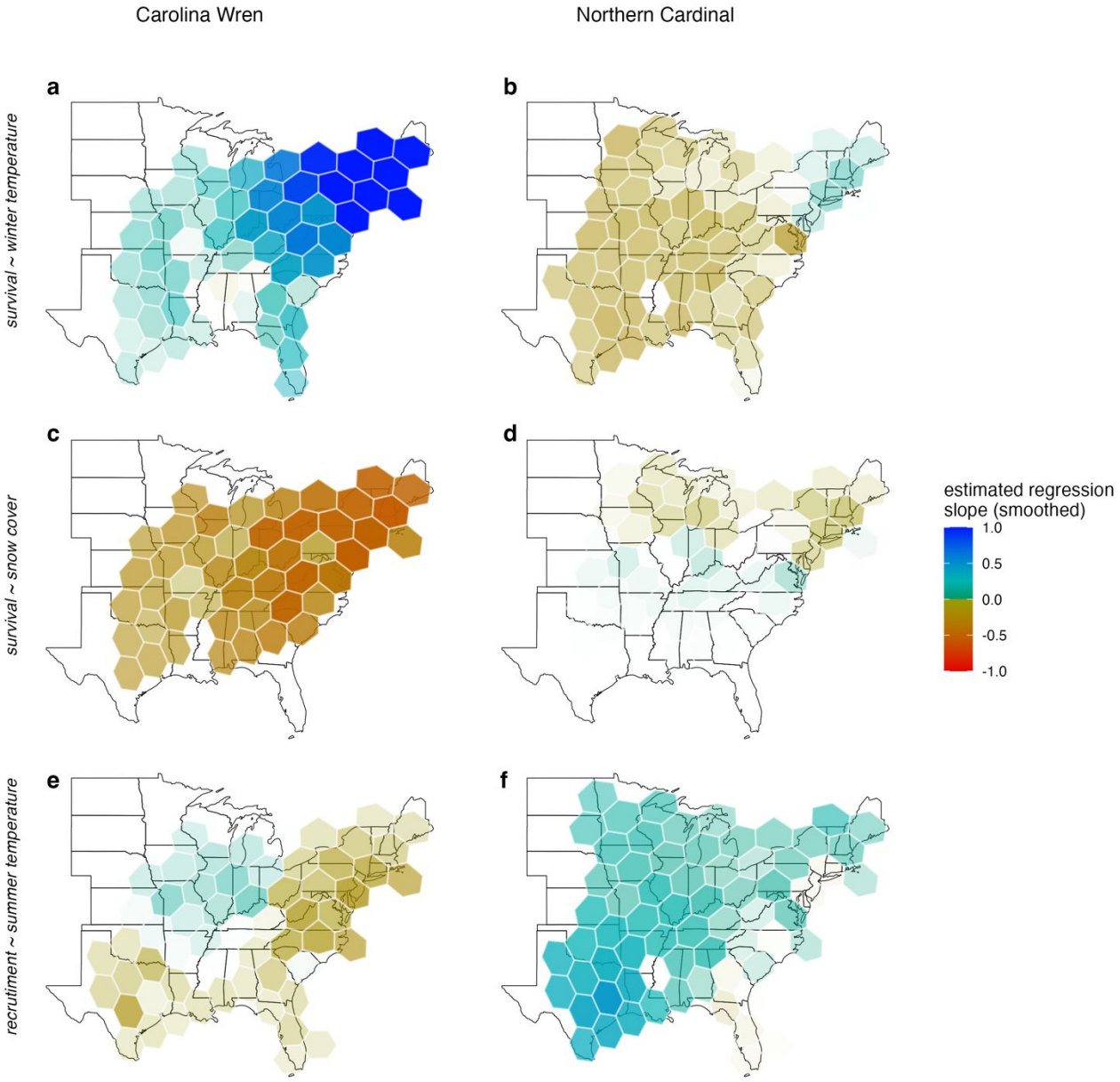
524



525

526 **Figure 3:** For Carolina Wren (a)--but not Northern Cardinal (b)--northern cells tend to show
527 evidence for higher variance in survival, implying that in the north, annual population fluctuations
528 are more strongly controlled by events during winter than by events during summer. The
529 difference in (log) standard deviations estimated from independent cell-level models of survival
530 and recruitment for Carolina Wren (a) and Northern Cardinal (b). The color scale gives the
531 posterior mean effect size for the difference in the logarithm of the standard deviation; the
532 opacity gives the posterior probability that the true effect is in the same direction as mean effect,
533 scaled so that a probability of 0.5 is completely transparent and a probability of 1 is completely
534 opaque. See SI Figure S3 for a color-based representation of these opacity values.

535



536

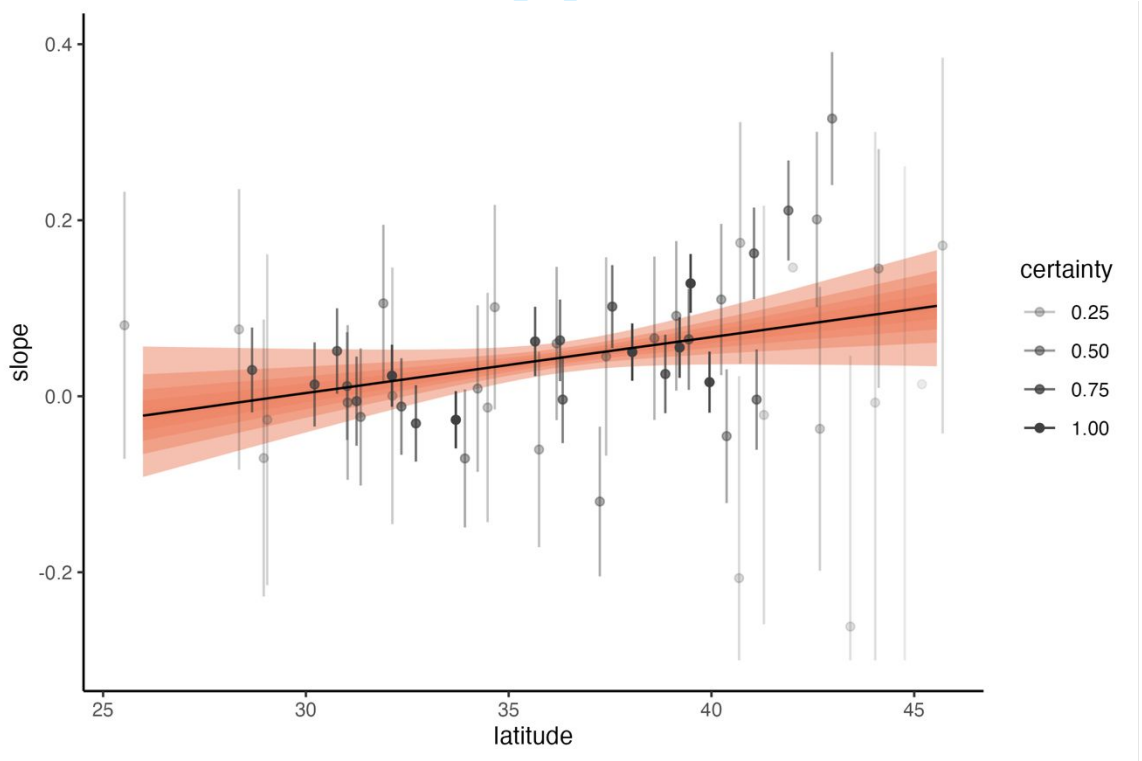
537 **Figure 4**: Survival and recruitment relationships to winter and summer weather in Carolina
538 Wren and Northern Cardinal. Carolina Wren survival is higher in warmer winters (a) and lower in
539 snowier winters (c) in the northeast, whereas recruitment shows no statistically robust
540 relationship to summer temperatures (e). Northern Cardinal shows potentially similar patterns in
541 survival, but with low certainty (b, d), while their recruitment is potentially higher when summer

542 temperatures were warm in the Mississippi Valley and Texas. The color scale gives the
543 posterior mean effect size for the true (smoothed) cell-specific slope for a regression of the
544 demographic index against weather conditions; the opacity gives the posterior probability that
545 the true effect is in the same direction as mean effect, scaled linearly so that a probability of 0.5
546 is completely transparent and a probability of 1 is completely opaque. See SI Figure S4 for a
547 color-based representation of the opacity values, and SI Figures S6 and S7 for unsmoothed
548 versions.

549

550

551



552

553 **Figure 5**: Posterior expectations for the slope of the relationship between winter temperature
554 and survival (natural logarithms per degree C) of Carolina Wren as a function of latitude, based

555 on a conditional autoregressive model. The median expectation is given in black; colored bands
556 delimit credible intervals in steps of 10%, with the widest band giving the 90% credible interval.
557 Points and vertical lines give the posterior mean +/- 1 standard deviation for the cell-specific
558 slopes. Opacity of data points is scaled as the uncertainty of the least uncertain point divided by
559 the uncertainty of the focal point. See SI Figure S5 for an equivalent analysis of Northern
560 Cardinal.

561

562

563

564

565

566 References

567 Aybar, C., Q. Wu, L. Bautista, R. Yali, and A. Barja. 2020. rgee: An R package for interacting
568 with Google Earth Engine. Journal of Open Source Software **5**:2272.
569 Balogh, A. L., T. B. Ryder, and P. P. Marra. 2011. Population demography of Gray Catbirds in
570 the suburban matrix: sources, sinks and domestic cats. Journal of Ornithology **152**:717-
571 726.
572 Barnes, R., K. Sahr, G. Evenden, A. Johnson, and F. Warmerdam. 2017. dggridR: discrete
573 global grids for R. R package version 0.1 **12**:963.
574 Beddall, B. G. 1963. Range expansion of the cardinal and other birds in the northeastern states.
575 The Wilson Bulletin **75**:140-158.
576 Betancourt, M. 2017. A conceptual introduction to Hamiltonian Monte Carlo. arXiv preprint
577 arXiv:1701.02434.
578 Boakes, E. H., P. J. K. McGowan, R. A. Fuller, D. Chang-qing, N. E. Clark, K. O'Connor, and G.
579 M. Mace. 2010. Distorted Views of Biodiversity: Spatial and Temporal Bias in Species
580 Occurrence Data. PLOS Biology **8**:e1000385.
581 Brooks, M. 1936. Winter killing of Carolina wrens. The Auk **53**:449-449.
582 Brown, D. J., C. A. Ribic, D. M. Donner, M. D. Nelson, C. I. Bocetti, and C. M. Deloria-Sheffield.
583 2017. Using a full annual cycle model to evaluate long-term population viability of the
584 conservation-reliant Kirtland's warbler after successful recovery. Journal of Applied
585 Ecology **54**:439-449.
586 Bürkner, P.-C. 2017. brms: An R package for Bayesian multilevel models using Stan. Journal of
587 statistical software **80**:1-28.
588 Cafaro, P., P. Hansson, and F. Götmark. 2022. Overpopulation is a major cause of biodiversity
589 loss and smaller human populations are necessary to preserve what is left. Biological
590 Conservation **272**:109646.
591 Ceballos, G., P. R. Ehrlich, and P. H. Raven. 2020. Vertebrates on the brink as indicators of
592 biological annihilation and the sixth mass extinction. Proceedings of the National
593 Academy of Sciences **117**:13596-13602.
594 Desante, D. F., K. M. Burton, J. F. Saracco, and B. L. Walker. 1995. Productivity indices and
595 survival rate estimates from MAPS, a continent-wide programme of constant-effort mist-
596 netting in North America. Journal of Applied Statistics **22**:935-948.
597 DeSante, D. F., D. R. O'Grady, and P. Pyle. 1999. Measures of productivity and survival derived
598 from standardized mist-netting are consistent with observed population changes. Bird
599 Study **46**:S178-S188.
600 DeSante, D. F., T. S. Sillett, R. B. Siegel, J. F. Saracco, C. A. Romo de Vivar Alvarez, S.
601 Morales, A. Cerezo, D. R. Kaschube, M. Grosselet, and B. Milá. 2005. MoSI (Monitoreo
602 de Sobrevivencia Invernal): assessing habitat-specific overwintering survival of
603 Neotropical migratory landbirds. USDA Forest Service General Technical Report.
604 Dickinson, J. L., B. Zuckerberg, and D. N. Bonter. 2010. Citizen Science as an Ecological
605 Research Tool: Challenges and Benefits. Annual Review of Ecology, Evolution, and
606 Systematics **41**:149-172.
607 Dunn, P. O., D. W. Winkler, A. Møller, W. Fiedler, and P. Berthold. 2010. Effects of climate
608 change on timing of breeding and reproductive success in birds. Effects of climate
609 change on birds **11**:17.
610 Edwards, B. P. M., and A. C. Smith. 2020. bbsBayes: An R Package for Hierarchical Bayesian
611 Analysis of North American Breeding Bird Survey Data.
612 bioRxiv:2020.2005.2027.118901.
613 Evans, B. S., T. B. Ryder, R. Reitsma, A. H. Hurlbert, and P. P. Marra. 2015. Characterizing
614 avian survival along a rural-to-urban land use gradient. Ecology **96**:1631-1640.

615 Faaborg, J., R. T. Holmes, A. D. Anders, K. L. Bildstein, K. M. Dugger, S. A. Gauthreaux Jr, P.
616 Heglund, K. A. Hobson, A. E. Jahn, D. H. Johnson, S. C. Latta, D. J. Levey, P. P. Marra,
617 C. L. Merkord, E. Nol, S. I. Rothstein, T. W. Sherry, T. S. Sillett, F. R. Thompson III, and
618 N. Warnock. 2010a. Recent advances in understanding migration systems of New World
619 land birds. Ecological Monographs **80**:3-48.
620 Faaborg, J., R. T. Holmes, A. D. Anders, K. L. Bildstein, K. M. Dugger, S. A. Gauthreaux Jr., P.
621 Heglund, K. A. Hobson, A. E. Jahn, D. H. Johnson, S. C. Latta, D. J. Levey, P. P. Marra,
622 C. L. Merkord, E. Nol, S. I. Rothstein, T. W. Sherry, T. S. Sillett, F. R. Thompson III, and
623 N. Warnock. 2010b. Conserving migratory land birds in the New World: Do we know
624 enough? Ecological Applications **20**:398-418.
625 Fink, D., A. Johnston, M. Strimas-Mackey, T. Auer, W. M. Hochachka, S. Ligocki, L.
626 Oldham Jaromczyk, O. Robinson, C. Wood, S. Kelling, and A. D. Rodewald. 2023. A
627 Double machine learning trend model for citizen science data. Methods in Ecology and
628 Evolution **14**:2435-2448.
629 Finn, C., F. Grattarola, and D. Pincheira-Donoso. 2023. More losers than winners: investigating
630 Anthropocene defaunation through the diversity of population trends. Biological Reviews
631 **98**:1732-1748.
632 Goldstein, B. R., and P. de Valpine. 2022. Comparing N-mixture models and GLMMs for relative
633 abundance estimation in a citizen science dataset. Scientific Reports **12**:12276.
634 Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore. 2017. Google
635 Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of
636 Environment **202**:18-27.
637 Habibullah, M. S., B. H. Din, S.-H. Tan, and H. Zahid. 2022. Impact of climate change on
638 biodiversity loss: global evidence. Environmental Science and Pollution Research
639 **29**:1073-1086.
640 Haggerty, T. M., and E. S. Morton. 2020. Carolina Wren (Thryothorus ludovicianus), version 1.0.
641 In Birds of the World (A. F. Poole, Editor). Cornell Lab of Ornithology, Ithaca, NY, USA.
642 https://doi.org/10.2173/bow.carwre.01.
643 Haggerty, T. M., E. S. Morton, and R. C. Fleischer. 2001. Genetic monogamy in Carolina wrens
644 (Thryothorus ludovicianus). The Auk **118**:215-219.
645 Halkin, S. L., D. P. Shustack, M. S. DeVries, J. M. Jawor, and S. U. Linville. 2021. Northern
646 Cardinal (Cardinalis cardinalis), version 2.0. In Birds of the World (P. G. Rodewald and
647 B. K. Keeney, Editors). Cornell Lab of Ornithology, Ithaca, NY, USA.
648 https://doi.org/10.2173/bow.norcar.02.
649 Hallworth, M. T., E. Bayne, E. McKinnon, O. Love, J. A. Tremblay, B. Drolet, J. Ibarzabal, S.
650 Van Wilgenburg, and P. P. Marra. 2021. Habitat loss on the breeding grounds is a major
651 contributor to population declines in a long-distance migratory songbird. Proceedings of
652 the Royal Society B: Biological Sciences **288**:20203164.
653 Halupka, L., and K. Halupka. 2017. The effect of climate change on the duration of avian
654 breeding seasons: a meta-analysis. Proceedings of the Royal Society B: Biological
655 Sciences **284**:20171710.
656 Hochachka, W. M., H. Alonso, C. Gutiérrez-Expósito, E. Miller, and A. Johnston. 2021. Regional
657 variation in the impacts of the COVID-19 pandemic on the quantity and quality of data
658 collected by the project eBird. Biological Conservation **254**:108974.
659 Hochachka, W. M., V. Ruiz-Gutierrez, and A. Johnston. 2023. Considerations for fitting
660 occupancy models to data from eBird and similar volunteer-collected data. Ornithology
661 **140**.
662 Huang, Q., J. R. Sauer, A. Swatantran, and R. Dubayah. 2016. A centroid model of species
663 distribution with applications to the Carolina wren Thryothorus ludovicianus and house
664 finch Haemorhous mexicanus in the United States. Ecography **39**:54-66.

Job, J., and P. A. Bednekoff. 2011. Wrens on the edge: feeders predict Carolina wren Thryothorus ludovicianus abundance at the northern edge of their range. Journal of Avian Biology **42**:16-21.
Johnston, A., D. Fink, W. M. Hochachka, and S. Kelling. 2018. Estimates of observer expertise improve species distributions from citizen science data. Methods in Ecology and Evolution **9**:88-97.
Johnston, A., W. M. Hochachka, M. E. Strimas-Mackey, V. Ruiz Gutierrez, O. J. Robinson, E. T. Miller, T. Auer, S. T. Kelling, and D. Fink. 2021. Analytical guidelines to increase the value of community science data: An example using eBird data to estimate species distributions. Diversity and Distributions **27**:1265-1277.
Johnston, A., E. Matechou, and E. B. Dennis. 2023. Outstanding challenges and future directions for biodiversity monitoring using citizen science data. Methods in Ecology and Evolution **14**:103-116.
Kalyuzhny, M., Y. Schreiber, R. Chocron, C. H. Flather, R. Kadmon, D. A. Kessler, and N. M. Shnerb. 2014a. Temporal fluctuation scaling in populations and communities. Ecology **95**:1701-1709.
Kalyuzhny, M., E. Seri, R. Chocron, C. H. Flather, R. Kadmon, and N. M. Shnerb. 2014b. Niche versus neutrality: a dynamical analysis. The American Naturalist **184**:439-446.
Knudsen, E., A. Lindén, C. Both, N. Jonzén, F. Pulido, N. Saino, W. J. Sutherland, L. A. Bach, T. Coppack, T. Ergon, P. Gienapp, J. A. Gill, O. Gordo, A. Hedenström, E. Lehikoinen, P. P. Marra, A. P. Møller, A. L. K. Nilsson, G. Péron, E. Ranta, D. Rubolini, T. H. Sparks, F. Spina, C. E. Studds, S. A. Sæther, P. Tryjanowski, and N. C. Stenseth. 2011. Challenging claims in the study of migratory birds and climate change. Biological Reviews **86**:928-946.
La Sorte, F. A., C. A. Lepczyk, J. L. Burnett, A. H. Hurlbert, M. W. Tingley, and B. Zuckerberg. 2018. Opportunities and challenges for big data ornithology. The Condor **120**:414-426.
La Sorte, F. A., and M. Somveille. 2020. Survey completeness of a global citizen-science database of bird occurrence. Ecography **43**:34-43.
La Sorte, F. A., and F. R. Thompson. 2007. Poleward shifts in winter ranges of North American birds. Ecology **88**:1803-1812.
Lack, D. 1954. The natural regulation of animal numbers. Clarendon. Oxford.
Lamanna, J. A., T. L. George, J. F. Saracco, M. P. Nott, and D. F. DeSante. 2012. El Niño—Southern Oscillation Influences Annual Survival of a Migratory Songbird at a Regional Scale. The Auk **129**:734-743.
Lande, R., S. Engen, and B.-E. Saether. 2003. Stochastic population dynamics in ecology and conservation. Oxford University Press, USA.
Latimer, C. E., and B. Zuckerberg. 2021. Habitat loss and thermal tolerances influence the sensitivity of resident bird populations to winter weather at regional scales. Journal of Animal Ecology **90**:317-329.
Leston, L. F. V., and A. D. Rodewald. 2006. Are urban forests ecological traps for understory birds? An examination using Northern cardinals. Biological Conservation **131**:566-574.
Lima, S. L. 2009. Predators and the breeding bird: behavioral and reproductive flexibility under the risk of predation. Biological Reviews **84**:485-513.
Link, W. A., and J. R. Sauer. 2007. Seasonal components of avian population change: joint analysis of two large-scale monitoring programs. Ecology **88**:49-55.
Maresh Nelson, S. B., C. A. Ribic, N. D. Niemuth, J. Bernath-Plaisted, and B. Zuckerberg. 2024. Sensitivity of North American grassland birds to weather and climate variability. Conservation Biology **38**:e14143.

Marra, P. P., C. E. Studds, S. Wilson, T. S. Sillett, T. W. Sherry, and R. T. Holmes. 2015. Non-breeding season habitat quality mediates the strength of density-dependence for a migratory bird. Proceedings of the Royal Society B: Biological Sciences **282**:20150624.

Martin, T. E. 1987. Food as a limit on breeding birds: a life-history perspective. Annual review of ecology and systematics **18**:453-487.

Mehlman, D. W. 1997. Change in avian abundance across the geographic range in response to environmental change. Ecological Applications **7**:614-624.

Morris, M., K. Wheeler-Martin, D. Simpson, S. J. Mooney, A. Gelman, and C. DiMaggio. 2019. Bayesian hierarchical spatial models: Implementing the Besag York Mollié model in stan. Spatial and Spatio-temporal Epidemiology **31**:100301.

Morrison, C. A., S. J. Butler, J. A. Clark, J. Arizaga, O. Baltà, J. Cepák, A. L. Nebot, M. Piha, K. Thorup, T. Wenninger, R. A. Robinson, and J. A. Gill. 2022. Demographic variation in space and time: implications for conservation targeting. Royal Society Open Science **9**:211671.

Newton, I. 1998. Population limitation in birds. Academic press.

Peach, W., S. Buckland, and S. Baillie. 1996. The use of constant effort mist-netting to measure between-year changes in the abundance and productivity of common passerines. Bird Study **43**:142-156.

Riegert, J. 2022. Detectability of birds under different sampling efforts and during the breeding season: a case study from Central Europe. Journal of Vertebrate Biology **71**:22027.22021-22012.

Root, T. 1988. Energy Constraints on Avian Distributions and Abundances. Ecology **69**:330-339.

Rosenberg, K. V., A. M. Dokter, P. J. Blancher, J. R. Sauer, A. C. Smith, P. A. Smith, J. C. Stanton, A. Panjabi, L. Helft, M. Parr, and P. P. Marra. 2019. Decline of the North American avifauna. Science **366**:120-124.

Rubin, D. B. 1981. The bayesian bootstrap. The annals of statistics:130-134.

Sæther, B.-E., W. J. Sutherland, and S. Engen. 2004. Climate Influences on Avian Population Dynamics. Pages 185-209 Advances in Ecological Research. Academic Press.

Saracco, J. F., R. L. Cormier, D. L. Humple, S. Stock, R. Taylor, and R. B. Siegel. 2022. Demographic responses to climate-driven variation in habitat quality across the annual cycle of a migratory bird species. Ecology and Evolution **12**:e8934.

Saracco, J. F., J. A. Royle, D. F. DeSante, and B. Gardner. 2012. Spatial modeling of survival and residency and application to the Monitoring Avian Productivity and Survivorship program. Journal of Ornithology **152**:469-476.

Sauer, J. R., J. E. Fallon, and R. Johnson. 2003. Use of North American Breeding Bird Survey data to estimate population change for bird conservation regions. The Journal of wildlife management:372-389.

Sauer, J. R., and W. A. Link. 2011. Analysis of the North American Breeding Bird Survey Using Hierarchical Models. The Auk **128**:87-98.

Sauer, J. R., G. W. Pendleton, and B. G. Peterjohn. 1996. Evaluating Causes of Population Change in North American Insectivorous Songbirds. Conservation Biology **10**:465-478.

Socolar, J. B., P. N. Epanchin, S. R. Beissinger, and M. W. Tingley. 2017. Phenological shifts conserve thermal niches in North American birds and reshape expectations for climate-driven range shifts. Proceedings of the National Academy of Sciences **114**:12976-12981.

Stan Development Team. 2023. Stan Modeling Language Users Guide and Reference Manual.

Sullivan, B. L., J. L. Aycrigg, J. H. Barry, R. E. Bonney, N. Bruns, C. B. Cooper, T. Damoulas, A. A. Dhondt, T. Dietterich, A. Farnsworth, D. Fink, J. W. Fitzpatrick, T. Fredericks, J. Gerbracht, C. Gomes, W. M. Hochachka, M. J. Iliff, C. Lagoze, F. A. La Sorte, M.

763 Merrifield, W. Morris, T. B. Phillips, M. Reynolds, A. D. Rodewald, K. V. Rosenberg, N.
764 M. Trautmann, A. Wiggins, D. W. Winkler, W.-K. Wong, C. L. Wood, J. Yu, and S.
765 Kelling. 2014. The eBird enterprise: An integrated approach to development and
766 application of citizen science. Biological Conservation **169**:31-40.
767 Thornton, M. M., R. Shrestha, P. E. Y. Wei, Thornton, S-C. Kao, and B. E. Wilson. 2022.
768 Daymet: Daily Surface Weather Data on a 1-km Grid for North America.
769 Tian, Y., and F. Hua. 2023. Abundance versus vital rates: The extent and predictors of
770 inconsistent conclusions on avian population responses to forest loss and degradation.
771 Biological Conservation **288**:110353.
772 Vehtari, A., A. Gelman, D. Simpson, B. Carpenter, and P.-C. Bürkner. 2021. Rank-
773 normalization, folding, and localization: An improved Rˆfor assessing convergence of
774 MCMC (with discussion). Bayesian analysis **16**:667-718.
775 Wilson, D. M., and J. Bart. 1985. Reliability of singing bird surveys: effects of song phenology
776 during the breeding season. The Condor **87**:69-73.
777 Ziolkowski, D. J., M. Lutmerding, W. B. English, V. I. Aponte, and M.-A. R. Hudson. 2023. North
778 American Breeding Bird Survey Dataset 1966 - 2022: U.S. Geological Survey data
779 release.

780

Supplementary information for: Seasonal macro-demography of North American bird populations revealed through citizen science monitoring

September 27, 2024

1 Model Structures and Equations

Throughout we will index the type of demographic index as $t \in \{\text{recruitment, survival}\}$), year as y and hexagon macrocell as c. We indicate the dependence of each variable to type of demographic index t, year y and macrocell c by index subscripts. In addition to the mathematical description of our models, we also include a brief description of the canonical model formula as used in the R-package BRMS (*1*). For a full description of the model implementation we refer to the full supplemental R code included with this publication.

1.1 Mean and variance by season

In each cell c and for each type of demographic index d, we approximated the bootstrapped posteriors for the fluctuation indices as Normal, and we modeled the demographic indices as

$$\begin{aligned} Y_{t,y,c} &\sim \text{Normal}(z_{t,y,c}, \epsilon_{t,y,c}) \\ z_{t,y,c} &\sim \text{Normal}(\mu_{t,c}, \sigma_{t,c}) \\ \mu_{t,c} &= \alpha_c + \beta_c \times \tau_t \\ \log \sigma_{t,c} &= \gamma_c + \kappa_c \times \tau_t \end{aligned} \tag{1}$$

where $z_{d,y,c}$ is the true index, which yields our noisy estimate $Y_{t,y,c}$ with known standard deviation $\epsilon_{t,c,y}$, α is the intercept, β is the effect of type of demographic index (notated as factor τ), γ is the log-scale intercept for the standard deviation, and κ is the log-scale effect of τ on the standard deviation.

1.1.1 R code implementation in BRMS

In `BRMS` we model each cell c separately, using formula

$$Y_{t,y,c} \mid \text{resp\_se}(\epsilon_{t,y,c}, \text{sigma} = \text{TRUE}) \sim \tau_t, \text{sigma} \sim \tau_t \tag{2}$$

resp_se() specifies the known measurement error of the response. Input data consists of rows values $(Y_{t,y,c}, \epsilon_{t,y,c}, \tau_t)$ that span all years y and indices t. Figure 3 of the main document shows estimates for κ_c

1.2 Weather predictors of recruitment and survival

In each cell c, we fit three regressions (survival index against mean maximum winter temperature, survival index against 'snow water equivalent', and recruitment against mean maximum summer temperature). Each of these regression had a form

$$\begin{aligned} Y_{t,y,c} &\sim \text{Normal}(y_{t,y,c}, \epsilon_{t,y,c}) \\ y_{t,y,c} &\sim \text{Normal}(\mu_{t,y,c}, \sigma_{t,c}) \\ \mu_{t,y,c} &= \alpha_{t,c} + \beta_{t,c} \times w_{t,y,c} \end{aligned} \tag{3}$$

where $y_{t,y,c}$ is the true index, which yields our noisy measurement $Y_{t,y,c}$ with known standard deviation $\epsilon_{t,y,c}$, $\alpha_{t,c}$ is the intercept, $\beta_{t,c}$ is the effect of weather (notated w), and $\sigma_{t,c}$ is the residual standard deviation.

1.2.1 R code implementation in BRMS

In `BRMS` we model each cell c separately, using formula

$$Y_{t,y,c} \mid \text{resp\_se}(\epsilon_{t,y,c}, \text{sigma} = \text{TRUE}) \sim w \tag{4}$$

and input data consisting of rows values $(Y_{t,y,c}, \epsilon_{t,y,c}, w_{t,y,c})$ that span all years y. In each of the three models, we include only data for one index t for the relevant period, i.e. recruitment or survival. Figure S6 shows estimates for the posterior mean of $\beta_{t,c}$

1.3 Spatially smoothed weather relationships

For each of our three weather variables, we approximated our posterior estimates of the regression slopes $\beta_{t,c}$ as Normal and smoothed these estimates across cells using regressions of the form

$$\begin{aligned} \beta_{t,c} &\sim \text{Normal}(b_{t,c}, \epsilon_{t,c}) \\ b_{t,c} &\sim \text{Normal}(\mu_{t,c}, \sigma) \\ \mu_{t,c} &= \alpha + \phi_{t,c} \end{aligned} \tag{5}$$

where $b_{t,c}$ is the true slope in a cell, which yields our noisy estimate $\beta_{t,c}$ with known standard deviation $\epsilon_{t,c}$ (obtained from the posterior mean and standard deviation from the earlier model fit in Eq. 3), α is an intercept, ϕ is a spatial random effect of cell with an intrinsic conditional autoregressive (ICAR) prior, and σ is the residual standard deviation, or in other words the standard deviation of a non-spatial random effect of cell.

In models that additionally contain an effect of latitude L, that effect $(\beta \times L_c)$ is added to $\mu_{t,c}$, as in

$$\mu_{t,c} = \alpha + \phi_{t,c} + \beta \times L_c \tag{6}$$

1.3.1 R code implementation in BRMS

In `BRMS` we model all cells c together, using formula

$$\beta_{t,c} \mid \text{resp\_se}(\epsilon_{t,c}, \text{sigma} = \text{TRUE}) \sim \text{car}(M, \text{gr=cell, type="icar"}) \quad (7)$$

with M the adjacency matrix of locations

In models that additionally contain an effect of latitude L, we used model formula

$$\beta_{t,c} \mid \text{resp\_se}(\epsilon_{t,c}, \text{sigma} = \text{TRUE}) \sim L_c + \text{car}(M, \text{gr=cell, type="icar"}) \quad (8)$$

Input data consists of rows of values for each cell $(\beta_{t,c}, \epsilon_{t,c}, L_c)$, with $\beta_{t,c}$ the slope of the weather regression and associated standard deviation $\epsilon_{t,c}$ and L_c the centroid latitude of the hexagon cell, and the adjacency matrix M.

1.4 Models for BBS counts

We modeled Breeding Bird Survey (BBS) counts within each BCR via

$$\begin{aligned} C_y &\sim \text{Poisson}(p_y) \\ \log(p_y) &= \eta_y + \nu + s(y) \\ \eta &\sim \text{Normal}(0, \sigma) \\ \nu &\sim \text{Normal}(0, \theta) \end{aligned} \quad (9)$$

where p_y is the true count, which yields our noisy estimates C_y of observed counts, y is the year associated with the count, η is a random effect of year with standard deviation σ, ν is an observation-level random effect with standard deviation θ, and $s()$ is a spline constructed by `R` package `mgcv` and cast in its random effects form for fitting via Stan (*2*).

1.4.1 R code implementation in BRMS

In summary, in `BRMS` we model the observed counts C_y using formula:

$$C_y \sim s(y) + (1 \mid i) + (1 \mid y) \quad (10)$$

where i is indexing the count observations.

1.4.2 Details of model fitting for BBS timeseries

For some Bird Conservation Regions (BCRs), numerically simulated trajectories in Hamiltonian Monte Carlo sampling tended to display divergences, which we reduced by increasing the target acceptance rate to induce smaller step sizes when numerically solving for the Hamiltonian trajectories (*3*). Nevertheless, divergences persisted in very low numbers in 5 of the 16 BCRs, with no more than two divergent transitions out of the 4,000 post-warmup transitions obtained for each BCR. Any resulting biases in the posterior (which would pass completely undetected using traditional MCMC fitting engines) are unlikely to seriously impact downstream analysis.

1.5 Comparison of BBS and eBird

We defined log-scale population fluctuations B as the log-scale differences in population index between year y and year y+1 for each BCR, as in

$$B_i = \log(C_{y+1} - C_y) \tag{11}$$

with C_y defined as in Eq. 9 for each BCR of the BBS, and i indexing unique year-BCR combinations. From the posterior samples we obtain an estimate for both the mean of B_i and the associated standard deviation ϵ_i^b.

For eBird we use an index E_i and known standard deviation ϵ_i^e that is equivalent to $Y_{t,y,c}$ and $\epsilon_{t,y,c}$ in Eq. 1, but in this case estimated for a BCR instead of an macro-cell c and using a log-ratio between two spring seasons of two consecutive years, instead of a log-ratio between two consecutive spring and fall seasons.

We approximated the posteriors for these annual log-scale population fluctuations as Normal for both BBS and eBird, and we treated the posterior means as noisy measurements of the unknown true values. We then regressed the BBS values against the eBird values as follows:

$$\begin{aligned} B_i &\sim \text{Normal}(b_i, \epsilon_i^b) \\ b_i &\sim \text{Normal}(\mu_i, \sigma) \\ \mu_i &= \alpha + \beta \times e_i + \eta_i \\ e_i &\sim \text{Normal}(t, u) \\ E_i &\sim \text{Normal}(e_i, \epsilon_i^e) \\ \eta_i &\sim \text{Normal}(0, \theta) \end{aligned} \tag{12}$$

Here, b_i is the true BBS fluctuation, which yields our noisy measurement B_i with known standard deviation ϵ_i^b; e_i is the true eBird fluctuation, which yields our noisy measurement E_i with known standard deviation ϵ_i^e; t and u are the mean and standard deviation (estimated during model fitting) of a regularizing hierarchical prior on the true eBird fluctuations. μ_i is the estimate of the true BBS fluctuation, which is regressed against the eBird fluctuation ϵ_i with α and β the regression intercept and slope, respectively; η is a random intercept for BCR with standard deviation θ;

1.5.1 R code implementation in BRMS

In summary, in `BRMS` we model the observed year-to-year fluctuations B_i using formula:

$$B_i \mid \text{resp\_se}(\epsilon_i^b, \text{sigma} = \text{TRUE}) \sim \text{me}(E_i, \epsilon_i^e) + (1 + \text{me}(E_i, \epsilon_i^e) \mid \text{BCR}) \tag{13}$$

where me() specifies a predictor with measurement error, and resp_se() specifies known measurement error of the response.

2 Supplementary figures

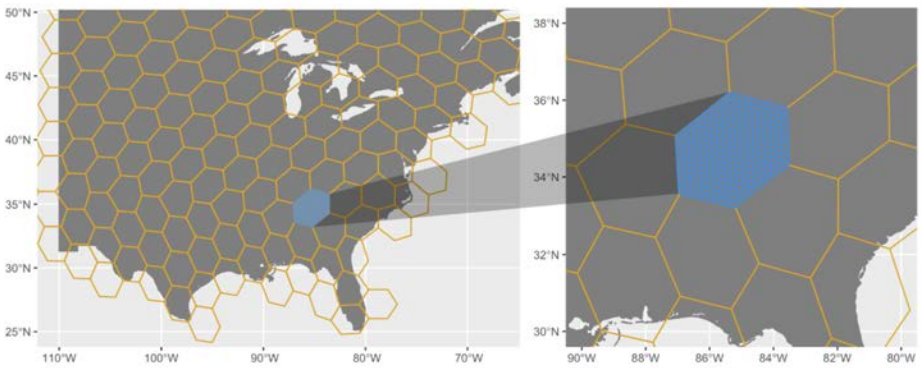


Figure S1: Indices of spring and fall bird populations and winter and summer weather conditions are derived on a hexagonal grid with roughly 285 km spacing between cell centers (left). Within each large hexagon, we construct a fine hexagonal grid with roughly 18 km spacing between cell centers (right). We compute our population index over each of these small cells, and we evaluate the uncertainty in index for the large parent cell by applying the Bayesian bootstrap over the small-cell indices.

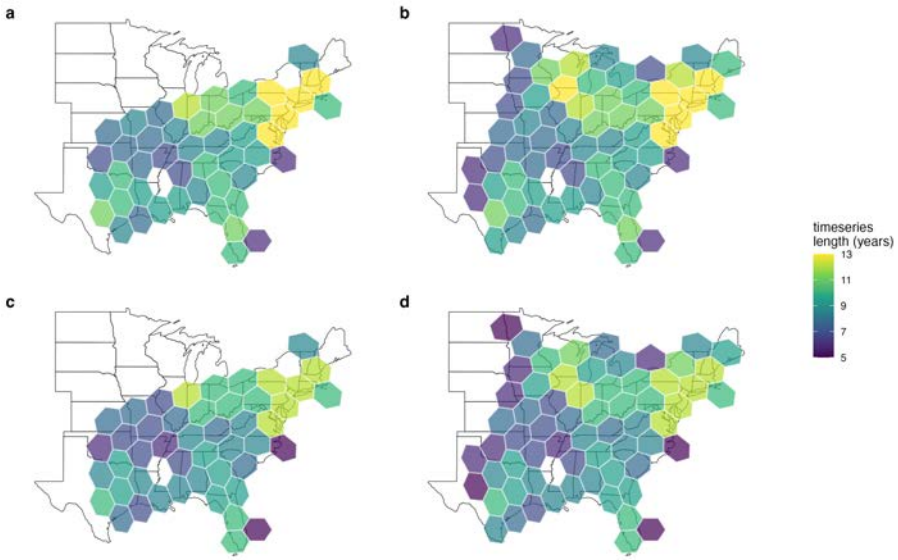


Figure S2: Lengths of timeseries analyzed for recruitment (a, b) and survival (c, d) of Carolina Wren (a, c) and Northern Cardinal (b, d).

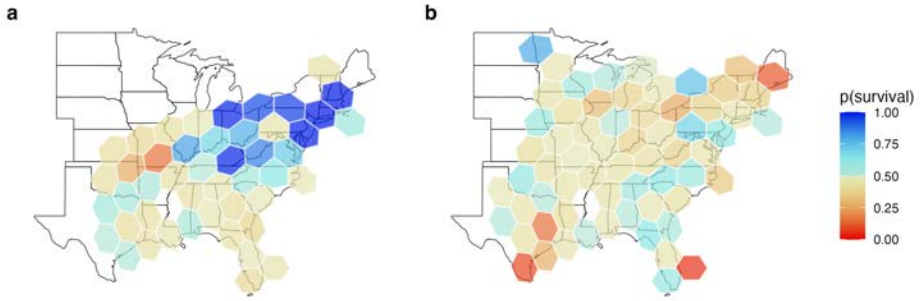


Figure S3: Posterior probabilities that the variance in survival is higher than the variance in recruitment for Carolina Wren (a) and Northern Cardinal (b).

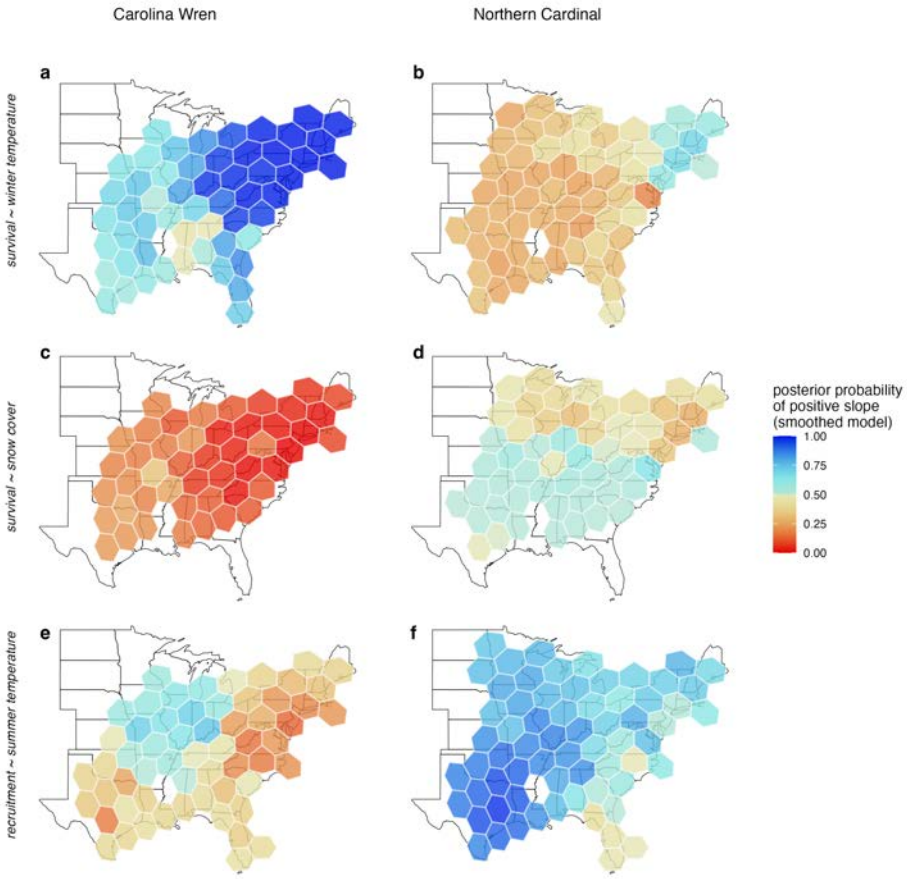


Figure S4: Posterior probabilities of positive slopes in regressions of timeseries against weather variables, smoothed via exact sparse conditional autoregressive models. Results are shown for the effects of winter temperatures on survival (a, b), winter snow cover on survival (c, d), and summer temperatures on recruitment (e, f), with Carolina Wren in the left column and Northern Cardinal in the right.

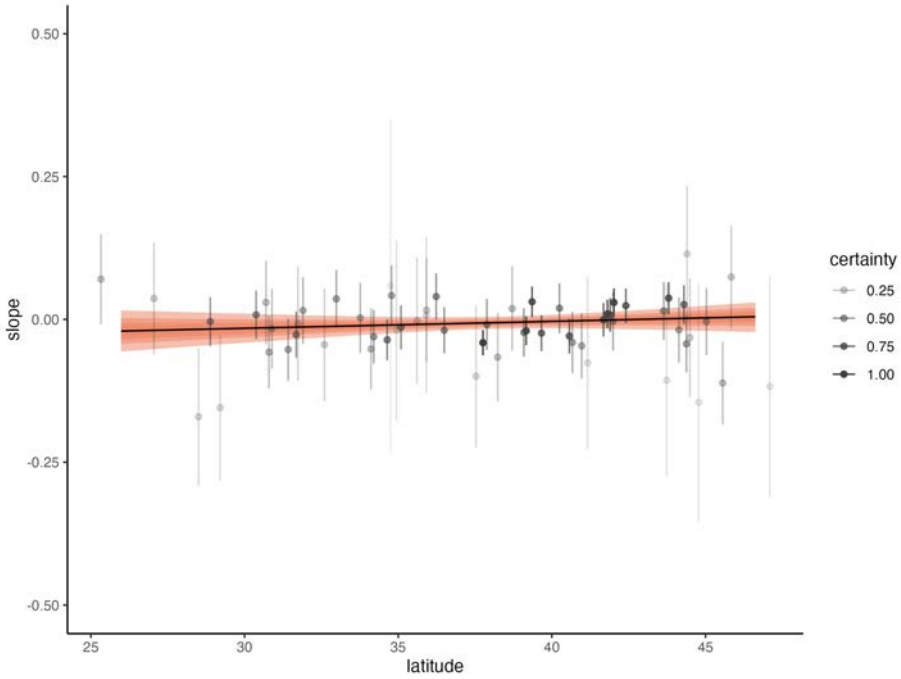


Figure S5: Posterior expectations for the slope of the relationship between winter temperature and survival (natural logarithms per degree C) of Northern Cardinal as a function of latitude, based on a conditional autoregressive model. The median expectation is given in black; colored bands delimit credible intervals in steps of 10%, with the widest band giving the 90% credible interval. Points and vertical lines give the posterior mean +/- 1 standard deviation for the cell-specific slopes. Opacity of data points is scaled as the uncertainty of the least uncertain point divided by the uncertainty of the focal point.

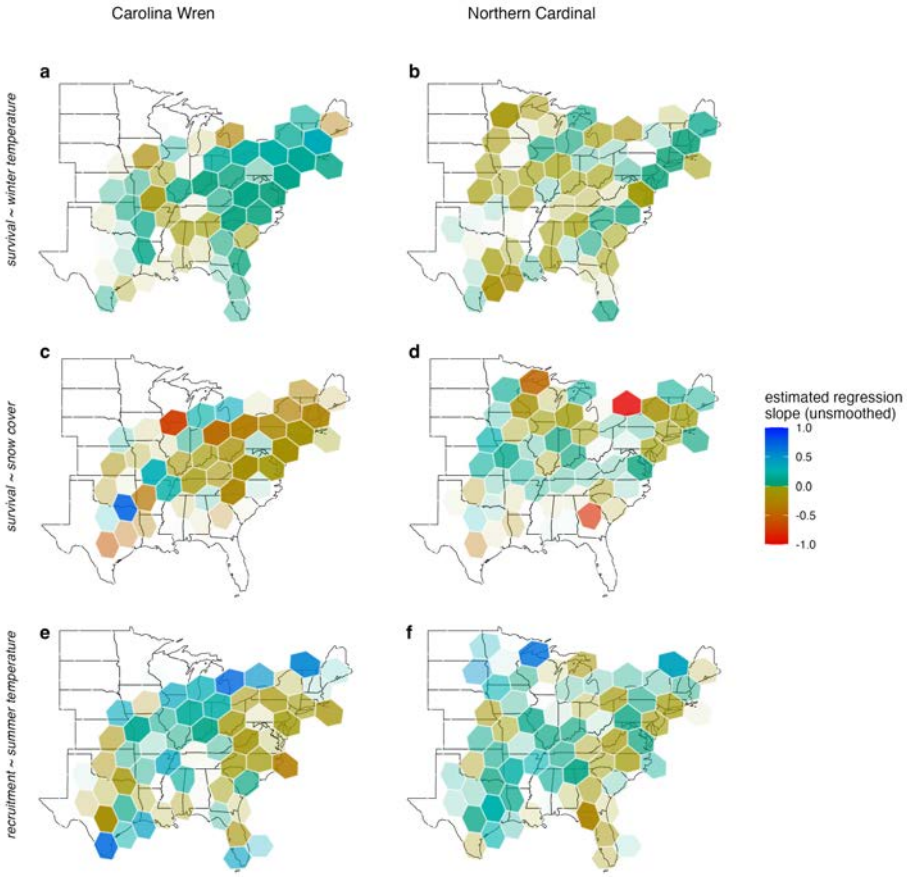


Figure S6: An equivalent of figure 4, but showing the unsmoothed cell-level posteriors.

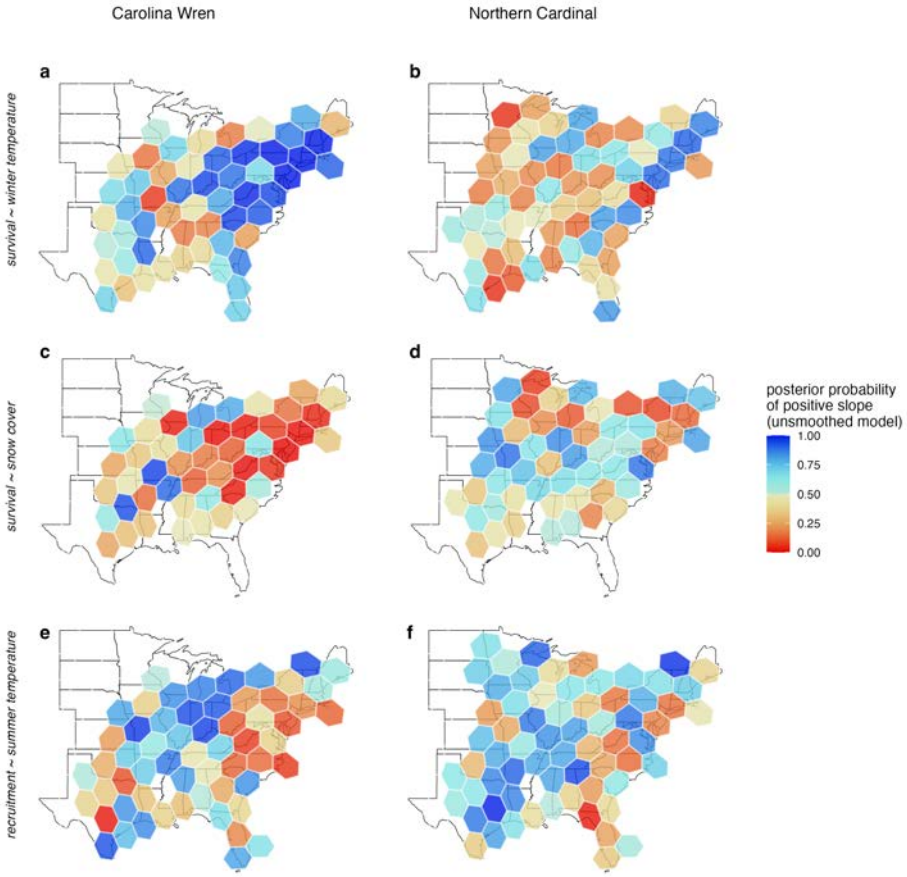


Figure S7: An equivalent of figure S4, showing the posterior probability of effect directionality.

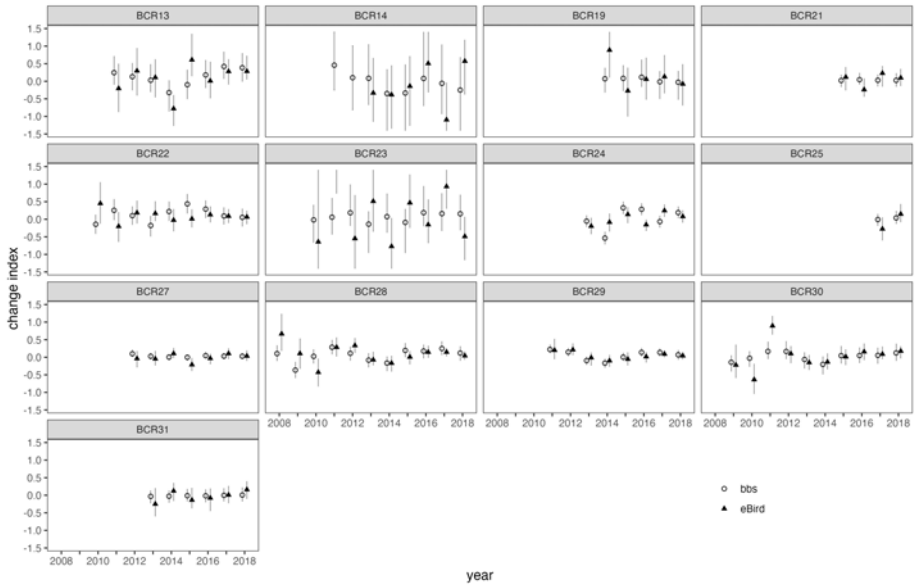


Figure S8: eBird and BBS derived indices for June-to-June population fluctuations at the level of bird conservation regions (BCRs). Data from years prior to 2008 did not meet the inclusion thresholds for eBird analysis.

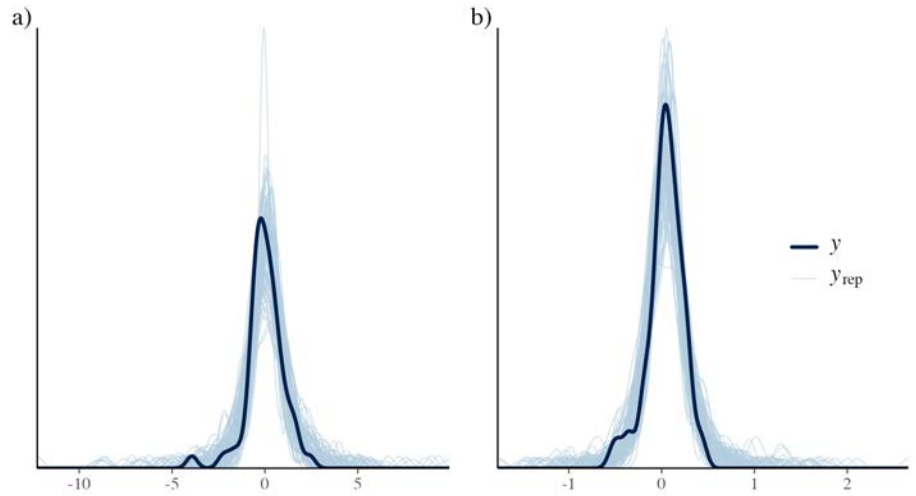


Figure S9: The examples of Posterior Predictive Check (PPC) for the latitudinal weather regression model (panel-a; regression model underlying Figure 5) and BBS validation model (panel-b; regression model underlying Figure 2a). The PPCs compare the empirical distribution of the data y to the distributions of simulated/replicated data yrep from the posterior predictive distribution.

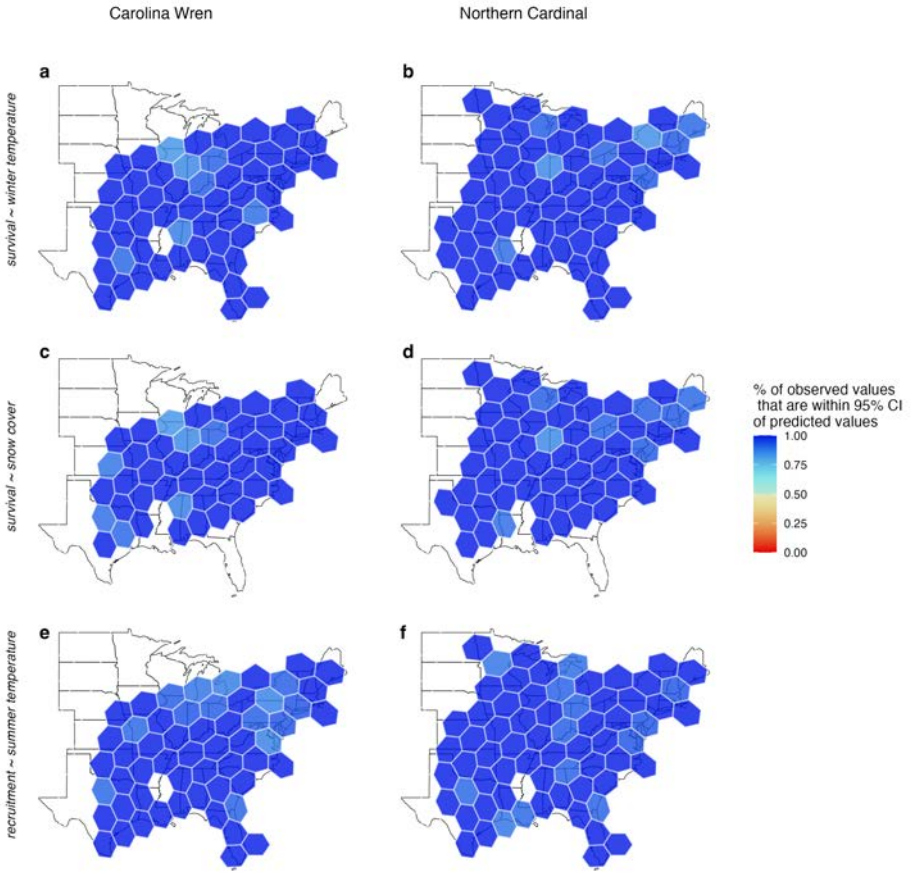


Figure S10: Posterior Predictive Checks of weather regression models. The color scale indicates the percentage of observed values that are within 95% credible interval of predicted values from posterior distribution estimated from independent cell-level weather regressions.

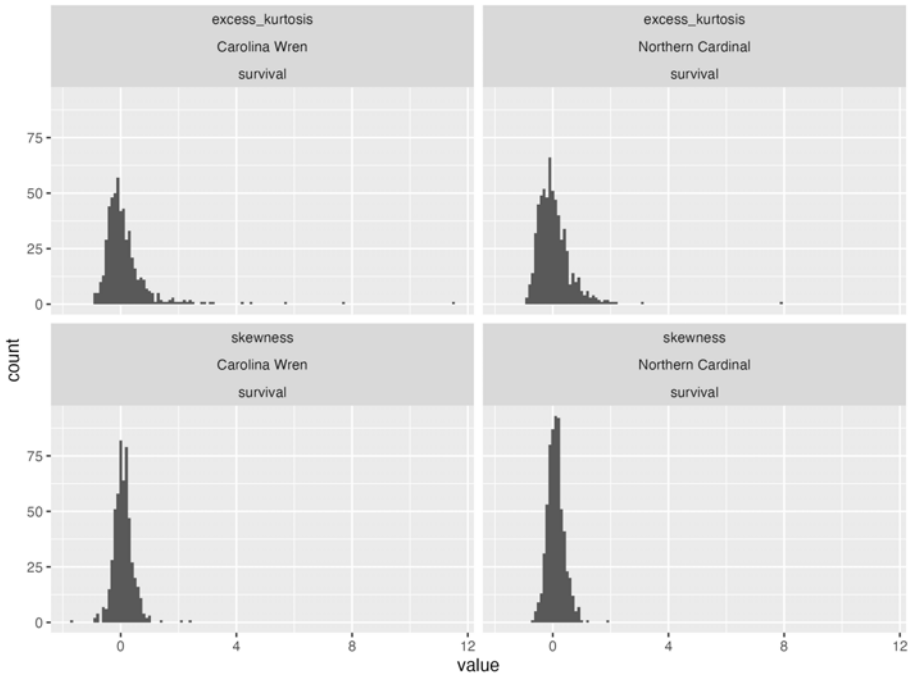


Figure S11: Histograms of the excess kurtosis (top) and skewness (bottom) of survival indices calculated for the bootstrap samples for each year - macrocell combination. Values centered near zero confirm of the adequacy of normal approximations in our regression models.

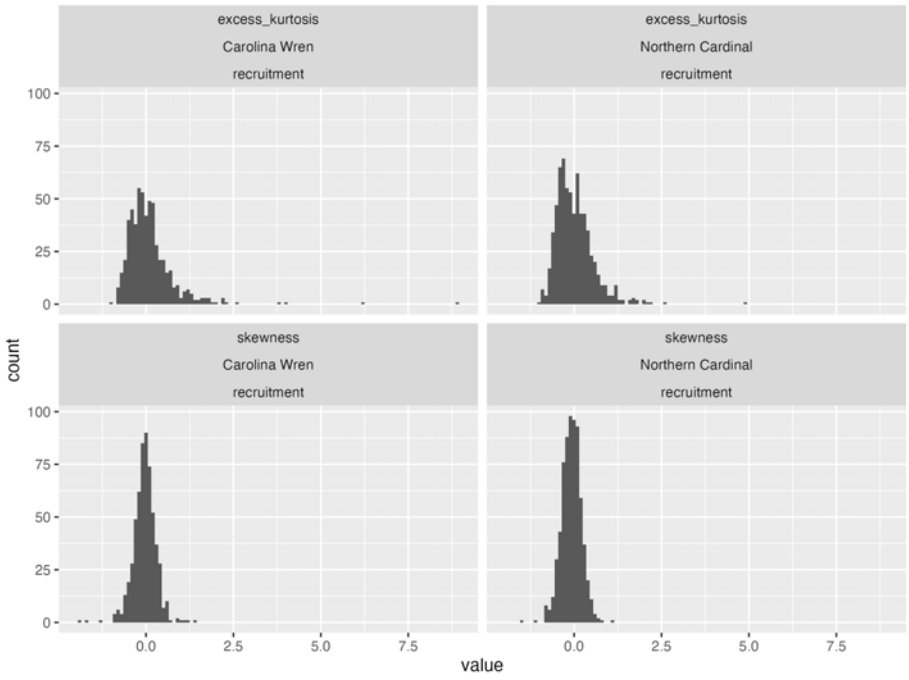


Figure S12: Histogram of the excess kurtosis (top) and skewness (bottom) of recruitment indices calculated for the bootstrap samples for each year - macrocell combination. Values centered near zero confirm of the adequacy of normal approximations in our regression models.

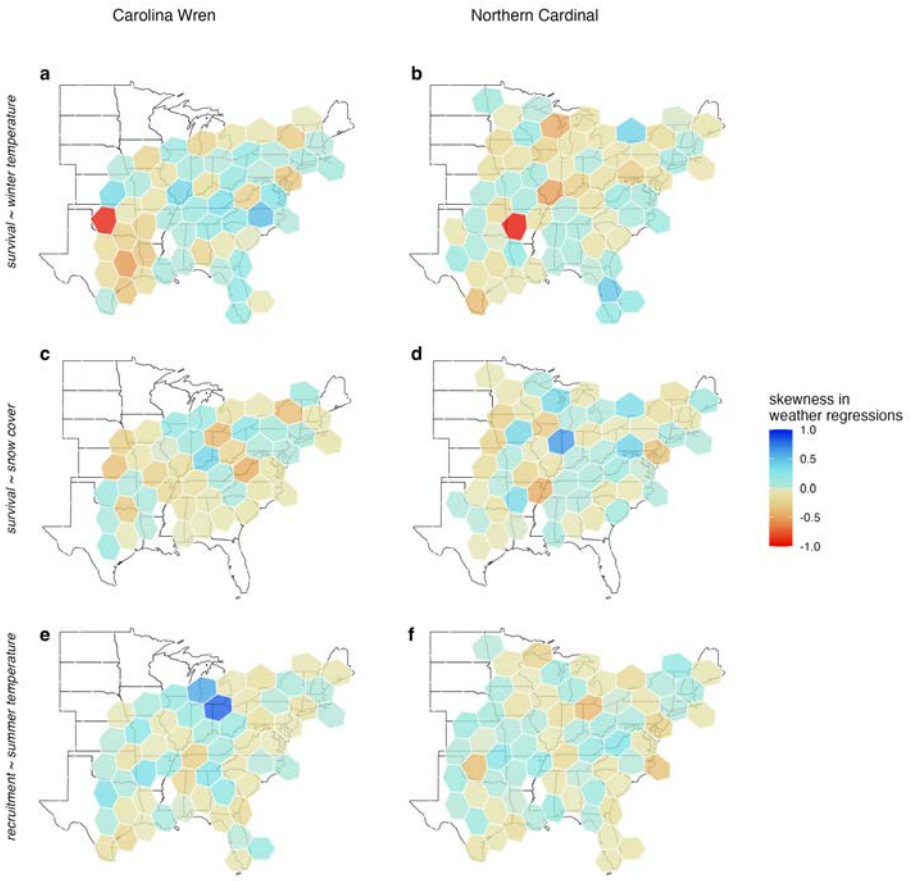


Figure S13: Skewness of the posterior distribution for the regression slope estimated in regressions between demographic indices and weather variables. Skewness was close to zero for most macro-cells, consistent with our normality assumptions.

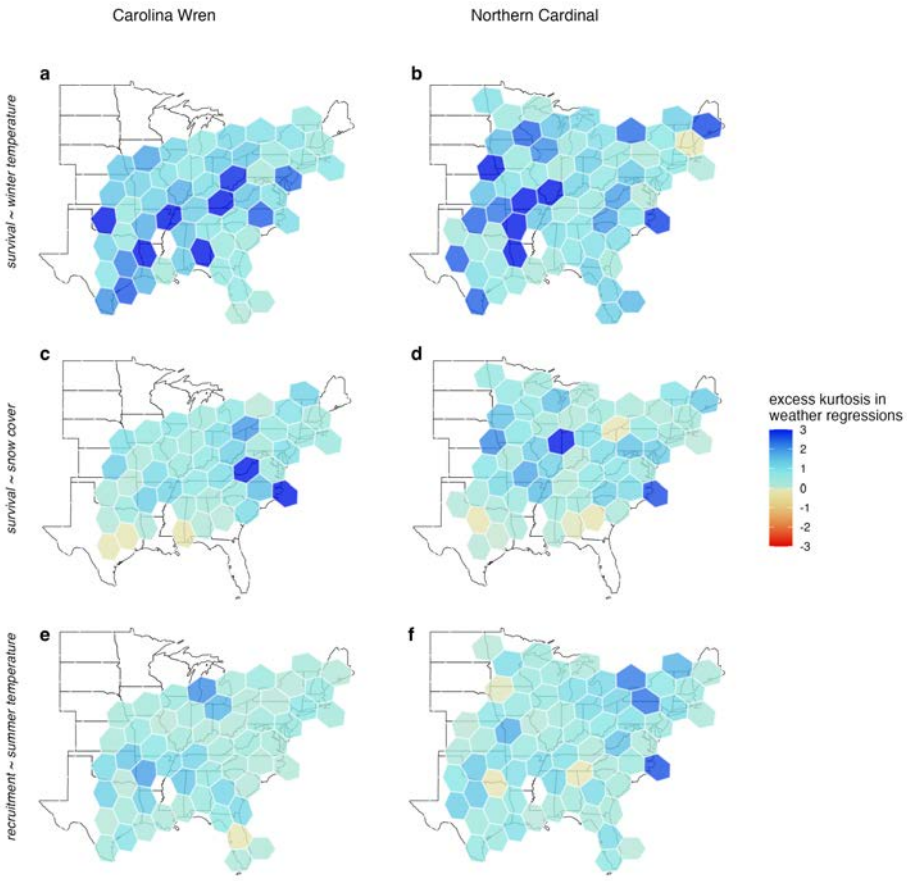


Figure S14: Excess kurtosis of the posterior distribution for the regression slope estimated in regressions between demographic indices and weather variables. Excess kurtosis was close to zero for most macro-cells, consistent with our normality assumptions.

References

(*1*) Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software 80*, 1–28.

(*2*) Stan Development Team Stan Modeling Language Users Guide and Reference Manual, version 2.34, 2023.

(*3*) Betancourt, M. A Conceptual Introduction to Hamiltonian Monte Carlo, 2018.

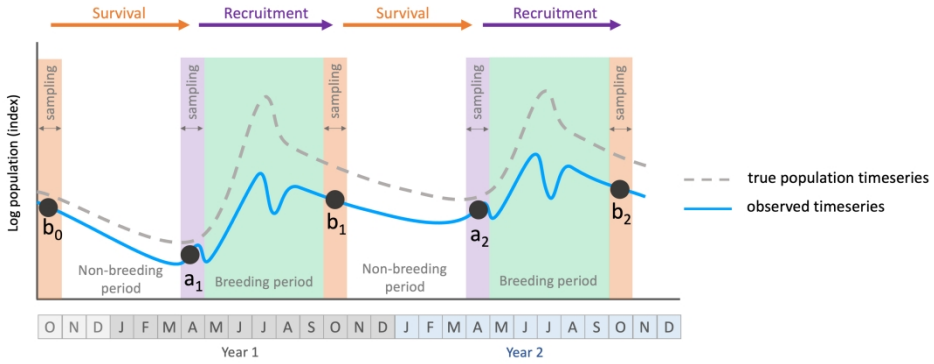


Figure 1: Conceptual overview of our approach for assessing population fluctuations using eBird data. The grey dashed curve represents a hypothetical two-year timeseriestime series of the logarithm of a population's size, beginning in early fall. The blue curve represents the apparent timeseriestime series from eBird data, which confounds the population timeseriestime series with detection effects (e.g., higher detection in spring than fall). We snapshot the eBird timeseriestime series in fall (circle 'b') and spring (circle 'a' and 'c'), and we treat the differences between successive snapshots as indices of survival (i.e., a1c -– b0) and recruitment (i.e., b1 -– a1; on the log scale, differences correspond to log-ratios). Because we are interested primarily in the year-to-year variability of these indices and not in their raw values, we can neglect the differences between the apparent log-population and the true log-population provided that these differences are consistent from spring to spring and from fall to fall (a multiplicative detection term becomes an additive term on the log scale). In this example, survival was higher in the second winter than in the first (i.e., a2 – b1 > a1 – b0), and the eBird-derived population snapshots provide an unbiased estimate of the difference between year-1 survival and year-2 survival.

419x190mm (242 x 242 DPI)

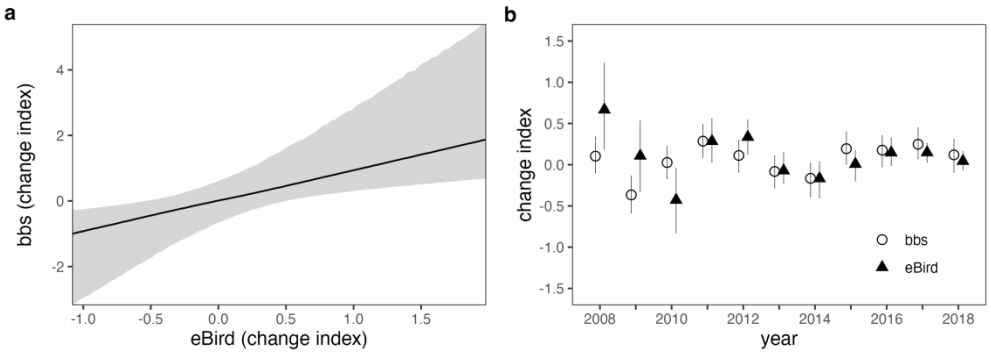
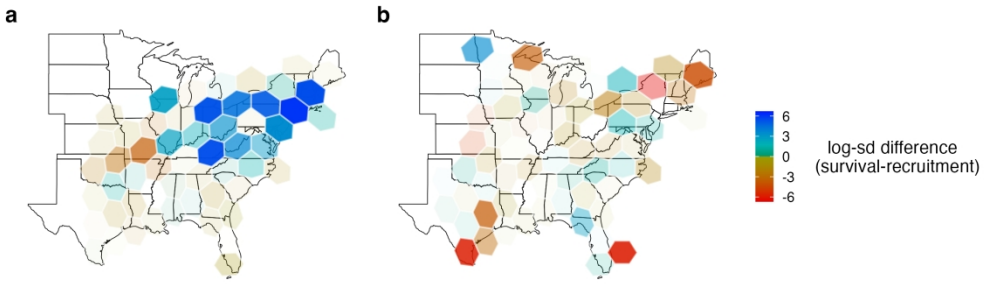


Figure 2: eBird derived indices for June-to-June population fluctuations at the level of bird conservation regions (BCRs) are predictive of fluctuations derived from the United States Breeding Bird Survey (BBS) for the same regions and years. a) The slope is estimated to be near unity (0.97, 95% CI 0.34–2.14). b) The match in fluctuations through time as visualized for one of the longest and best-aligned timeseries (BCR 28 includes the Appalachian Mountains from Alabama to southern New York). Data from years prior to 2008 did not meet the inclusion thresholds for eBird analysis.

559x203mm (118 x 118 DPI)



559x177mm (118 x 118 DPI)

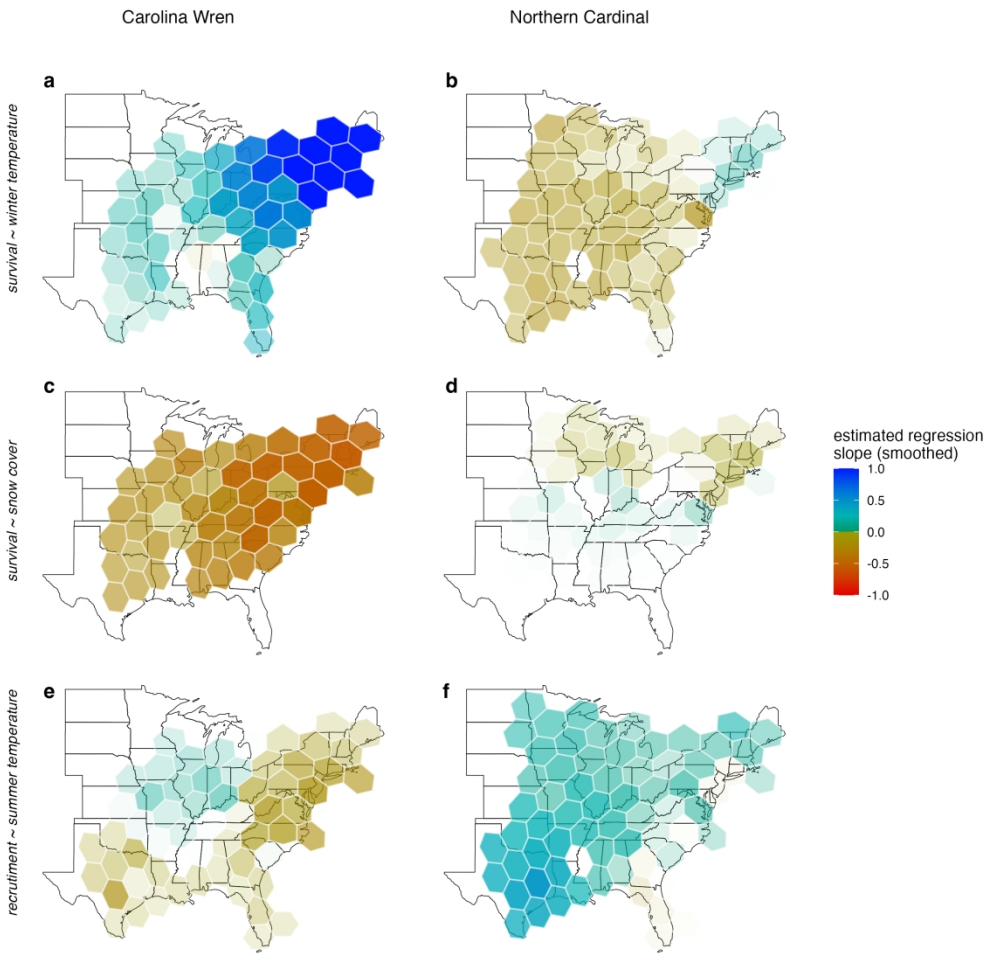


Figure 4: Survival and recruitment relationships to winter and summer weather in Carolina Wren and Northern Cardinal. Carolina Wren survival is higher in warmer winters (a) and lower in snowier winters (c) in the northeast, whereas recruitment shows no statistically robust relationship to summer temperatures (e). Northern Cardinal shows potentially similar patterns in survival, but with low certainty (b, d), while their recruitment is potentially higher when summer temperatures were warm in the Mississippi Valley and Texas. The color scale gives the posterior mean effect size for the true (smoothed) cell-specific slope for a regression of the demographic index against weather conditions; the opacity gives the posterior probability that the true effect is in the same direction as mean effect, scaled linearly so that a probability of 0.5 is completely transparent and a probability of 1 is completely opaque. See figure S4 for a color-based representation of the opacity values, and figures S6 and S7 for unsmoothed versions.

610x635mm (118 x 118 DPI)

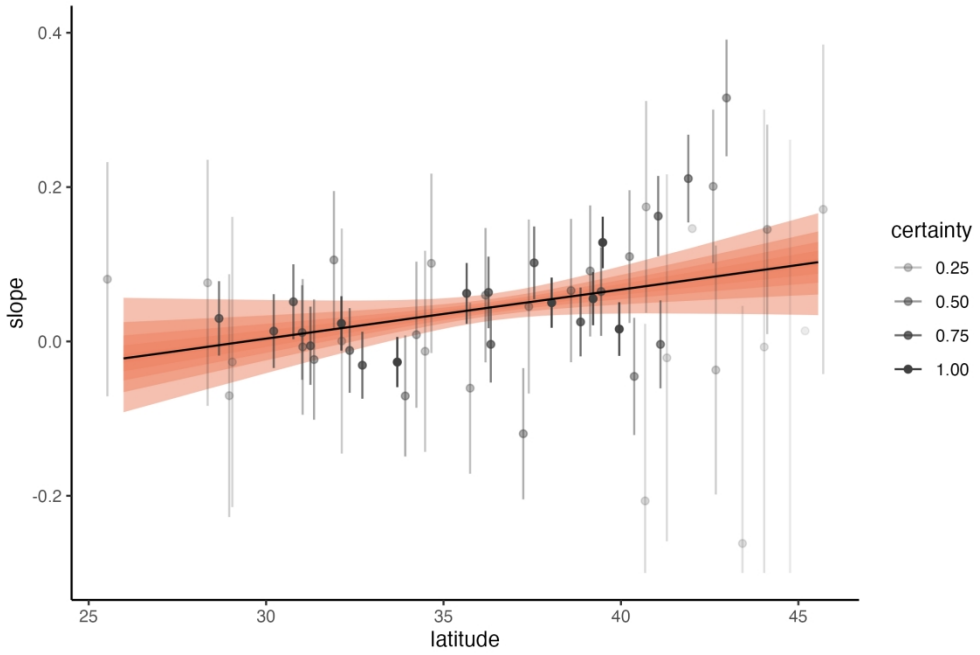


Figure 5: Posterior expectations for the slope of the relationship between winter temperature and survival (natural logarithms per degree C) of Carolina Wren as a function of latitude, based on a conditional autoregressive model. The median expectation is given in black; colored bands delimit credible intervals in steps of 10%, with the widest band giving the 90% credible interval. Points and vertical lines give the posterior mean +/- 1 standard deviation for the cell-specific slopes. Opacity of data points is scaled as the uncertainty of the least uncertain point divided by the uncertainty of the focal point. See figure S5 for an equivalent analysis of Northern Cardinal.

457x305mm (118 x 118 DPI)