Patterns and drivers of population trends on individual Breeding Bird Survey routes using spatially explicit models and route-level covariates.

Adam C. Smith, Veronica Aponte, Allison D. Binley, Amelia R. Cox, Lindsay Daly, Courtney Donkersteeg, Brandon P.M. Edwards, Willow B. English, Marie-Anne R. Hudson, David Iles, Kendall Jefferys, Barry Robinson, Christian Roy

Adam C. Smith – adam.smith@ec.gc.ca, Canadian Wildlife Service, Environment and Climate Change Canada, Ottawa, ORCID: 0000-0002-2829-4843

Veronica Aponte - Canadian Wildlife Service, Environment and Climate Change Canada, Ottawa,

Allison D. Binley - Department of Biology, Carleton University, Ottawa, Canada ORCID: 0000-0001-8790-9935

Amelia R. Cox - Canadian Wildlife Service, Environment and Climate Change Canada, Yellowknife, Canada

Lindsay Daly - Canadian Wildlife Service, Environment and Climate Change Canada, Ottawa, ORCID: 0000-0002-0892-5505

Courtney Donkersteeg - Department of Biology, Carleton University, Ottawa, Canada

Brandon P.M. Edwards, Department of Biology, Carleton University, Ottawa, Canada & Canadian Wildlife Service, Environment and Climate Change Canada, Ottawa, ORCID: 0000-0003-0865-3076

Willow B English - Canadian Wildlife Service, Environment and Climate Change Canada, Ottawa, ORCID: 0000-0002-0863-8581

Marie-Anne R. Hudson - Canadian Wildlife Service, Environment and Climate Change Canada, Ottawa, ORCID: 0000-0002-9599-0697

Kendall M Jefferys - Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, UK ORCID: 0000-0003-4439-9394

Barry Robinson - Canadian Wildlife Service, Environment and Climate Change Canada, Edmonton Canada ORCID: 0000-0002-2646-2508

Christian Roy - Canadian Wildlife Service, Environment and Climate Change Canada, Gatineau, Canada ORCID: 0000-0002-5599-6234

Acknowledgements:

We sincerely thank the thousands of U.S. and Canadian participants and the regional and national coordinators who have conducted and coordinated the North American Breeding Bird Survey for almost 60 years.

Author Contributions:

All authors contributed to the conceptual development of the ideas and concepts, study design, editing and drafting of the text. ACS, ARC, KMJ, and CR conducted the analyses.

Data Availability

Analyses reported in this article can be reproduced using the data and code provided at https://github.com/AdamCSmithCWS/Route-level_BBS_trends. (this will be permanently archived on publication)

Keywords

Ecological Monitoring, Gaussian Process, iCAR, population abundance

Abstract

Spatial patterns in population trends, particularly those at finer geographic scales, can help us better understand the factors driving population change in North American birds. The standard trend models for the North American Breeding Bird Survey (BBS) were designed to estimate trends within broad geographic strata, such as countries, Bird Conservation Regions, U.S. states, and Canadian territories or provinces. Calculating trend estimates at the level of the BBS's individual survey transects ("routes") allows us to explore finer spatial patterns and simultaneously estimate the effects of covariates, such as habitat loss or annual weather, on both relative abundance and trend (changes in relative abundance through time). Here, we describe four related hierarchical Bayesian models that estimate trends for individual BBS routes, implemented in the probabilistic programming language Stan. All four models estimate route-level trends and relative abundances using a hierarchical structure that shares information among routes, and three of the models share information in a spatially explicit way. The spatial models use either an intrinsic Conditional Autoregressive (iCAR) structure or a distance-based Gaussian Process (GP) to estimate the spatial components. We fit all four models to data for 71 species and then, because of the intensive computations required, fit two of the models (one spatial and one non-spatial) for an additional 216 species. In a leave-future-out cross-validation, the spatial models outperformed the non-spatial models for 284 out of 287 species. The best approach to modeling the spatial components depends on the species being modeled; the Gaussian Process had the highest predictive accuracy for 69% of the species tested here and the iCAR was better for the remaining 31%. We also present two examples of route-level covariate analyses focused on spatial and temporal variation in habitat for Rufous Hummingbird (*Selasphorus rufus*) and Horned Grebe (*Podiceps auritus*). In both examples, the inclusion of covariates improved our understanding of the patterns in the rate of population change for both species. Route-level models for BBS data are useful for visualizing spatial patterns of population change, generating hypotheses on the causes of change, comparing patterns of change among regions and species, and testing hypotheses on causes of change with relevant covariates.

Introduction

The North American Breeding Bird Survey (BBS) is the main source of bird population change information in North America. The BBS provides data at geographic scales ranging from national to regional across much of Canada and the United States for up to 500 species of birds (Hudson et al. 2017, Sauer et al. 2017). The BBS exemplifies the power of citizen science, given this standardized monitoring program has been running since 1965. BBS data are collected annually over 5400 routes by observers conducting 50, 3-minute point-counts along a roughly 40-km long section of roadside (Hudson et al. 2017). BBS data are often used to estimate the change in a species' population over time (i.e. trend) across regions such as Bird Conservation Regions (BCRs) or the intersection of states/provinces with BCRs (Sauer et al. 2003, Soykan et al. 2016, Link et al. 2020, Smith and Edwards 2020). These regional-scale summaries have been critical for identifying and prioritizing species in peril (Government of Canada 2010, IUCN 2012, Rosenberg et al. 2017) and understanding broad-scale patterns of change in North American birds (North American Bird Conservation Initiative Canada 2019, Rosenberg et al. 2019, North American Bird Conservation Initiative 2022).

The BBS dataset can also be analyzed at a finer spatial resolution to complement the regional estimates, and to address different ecological questions and conservation efforts. Incorporating the explicit spatial relationships among individual survey sites (BBS routes) provides the information necessary to estimate abundance and trends at a fine resolution (Smith et al. 2023). Many factors influence the relative abundance and trends in bird populations, and they act and interact to induce spatial patterns across a range of spatial scales (Morrison et al. 2010). Factors such as habitat change (Stanton et al. 2018, Betts et al. 2022), biotic factors like prey availability (Drever et al. 2018), or broad-scale patterns in abiotic factors like precipitation, temperature, and phenology (Renfrew et al. 2013, Wilson et al. 2018) can induce spatial patterns in trends or abundance and can act across different periods in the species' annual cycles (Morrison et al. 2010, Wilson et al. 2011). Likewise, conservation actions occur at many scales, from the broad scales of international conventions to the fine scales of an individual wetland (Prairie Habitat Joint

Venture 2020). Fine-scale estimates of population trends may provide a more useful unit for local conservation efforts and a better scale to model covariates with fine-scale effects such as species interactions, local land cover, and agricultural practices (Thogmartin et al. 2004, Paton et al. 2019, Mirochnitchenko et al. 2021).

The factors affecting population trends may differ from those affecting relative abundance, so it makes sense to model abundance and trend separately. Earlier fine-scale models designed for application across the full BBS dataset did not explicitly model the rate of population change as a parameter in the model (Bled et al. 2013). However, more recent work has demonstrated the utility of modeling both abundance and trends, such as a recent analysis of the effects of forest change on species' trends on BBS routes (Betts et al. 2022). A spatially-explicit hierarchical regression can model both spatial patterns in mean relative abundance and the rate of change in relative abundance (Ver Hoef et al. 2018, Wright et al. 2021). Separating these parameters in the model also allows for the inclusion of covariates (Meehan et al. 2019) to better understand the processes affecting local abundance (e.g., mean habitat amount or edge vs core of a species' range) and trends (local habitat change through time, or differences in climate change effects at Northern or Southern range limits).

Spatially explicit models in ecology treat individual sample units as either points within continuous space (Golding and Purse 2016), or discrete areas with neighborhood relationships (Ver Hoef et al. 2018). Intrinsic Conditional Autoregressive (iCAR) structures are areal and model spatial relationships using the adjacency between a discrete spatial area and its neighbors (Besag and Kooperberg 1995). These structures have been used to model the relatively fine-scale population trends in Christmas Bird Count data (Meehan et al. 2019) and the annual relative abundance of birds using BBS data (Bled et al. 2013). Gaussian Process (GP) models use the Euclidean distance between points and can model fine-scale spatial patterns in a species' relative abundance, treating spatial distances among survey sites to estimate the covariance of parameters in space (Golding and Purse 2016).

Here we describe and demonstrate four models that share two goals: 1) to estimate bird population trends and relative abundance for each BBS route; and 2) to visualize spatial patterns in both trend and relative abundance across a species' monitored range. Three of the models share information on relative abundance and trend in a spatially explicit way, while the fourth model lacks any spatial information. We describe two models that rely on an iCAR structure to model the spatial relationships: the first is the iCAR model, which uses only the iCAR structure to model variation in abundance and trends; and the second is a version of the BYM model, named for Besag, York, and Mollié (Besag et al. 1991), which is identical to the iCAR model but includes an additional random effect on the route-level trends to allow extra non-spatial variation in trends. The third is an isotropic Gaussian Process (GP) model that models covariance among routes in the abundance and trends using the Euclidean distances among routes. Finally the fourth model is a non-spatial version that estimates route-level variation in trends and abundances as a simple, log-normally distributed random effect. We fit all four models to 71 species and fit the non-spatial model and the iCAR model to another 216 species, selected based on sufficient data and computational requirements (details below). We compare the predictive accuracy of models for a given species in a leave-future-out cross-validation to assess the benefits of including spatial information to predict observations in the next year based on the estimated trend. Finally, we provide two examples of route-level covariate analyses to demonstrate elaborations of these models that directly estimate the effects of covariates on trends and the utility of modeling BBS data at a relatively fine, route-level scale.

Methods

Data

We used data for the Baird's Sparrow (*Centronyx bairdii*) as an example species to demonstrate the spatial structures, model fit, and convergence. We chose Baird's Sparrow as it has sufficient data to produce robust estimates but has a very restricted distribution, confined to the northern Great Plains region (Figure 1), which reduces model run-time. We used data for an additional 70 species (Table S1) to fit all four of the models and compared the predictions and predictive accuracy among the models. We chose these 70 species (71 including Baird's Sparrow) because they have small ranges with relatively few BBS routes, which minimizes the size of distance matrices and/or adjacency matrices for computational efficiency, and yet are also commonly observed during surveys and so provide high-quality data on any given route. Species with large ranges that appear on many routes will increase the computational power required to run the models, increasing the model run-time. Specifically, from 2006 to 2021, these small-range species were observed on 125-400 BBS routes, with at least 600 total observations of each species (positive counts), and at least an average of four observations per route. We are only able to compare the fit and predictive accuracy of all four models using these small-range species because the computational time required to fit the GP model was prohibitive for large inter-route distance matrices, given that days or even weeks are required for convergence for a single species. For the additional 216 species that were observed on 400 or more BBS routes during 2006-2021, we compare the predictions and predictive accuracy of the non-spatial model to one of the spatial models (iCAR) to assess the benefits of including spatial information when estimating trends for a large number of species.

We limited all but one of our analyses to a 15-year period, which we considered short enough that a log-linear slope can be a meaningful summary of the population change (Buckland et al. 2004, Thompson and La Sorte 2008). In effect, 15 years is likely long enough to estimate a meaningful rate of change on each route, but also short enough to reduce the likelihood of complex non-linear population patterns. The only exception is the Horned Grebe covariate example, where we used a 43-year period because the covariate

was designed to adjust for annual fluctuations and non-linear patterns in regional moisture/drought cycles (details below). This 15-year period that we demonstrate here is somewhat arbitrary and for many species or ecological questions, it may be very informative to fit these models (or modifications of these models) to longer or shorter periods of time.

Model structure

The four models are relatively simple, hierarchical log-link negative binomial regressions similar to other models commonly applied to BBS data (Sauer and Link 2011, Smith et al. 2014), but modeling trend and abundance as spatially-varying coefficients (Barnett et al. 2021, Thorson et al. 2023). In all four models, each route has a separate slope and intercept but there are no annual intercepts to model annual or non-linear temporal patterns in population change. Therefore, the interpretation of "trend" in these models is limited to this log-linear slope parameter (i.e., a single mean rate of change over the entire modeled time-series).

$$C_{r,j,t} = Negative\ Binomial\left(\lambda_{r,j,t}, \phi\right)$$

$$log\left(\lambda_{r,j,t}\right) = \alpha_r + \beta_r * (t - t_m) + \eta I[j,t] + \omega_j$$

We modeled the observed counts ($C_{r,j,t}$) of a given species on route-r, in year-t, by observer-j as realizations of a negative binomial distribution, with mean $\lambda_{r,j,t}$ and inverse dispersion parameter ϕ. The log of the mean ($\lambda_{r,j,t}$) of the negative binomial distribution was modeled as an additive combination of route-level intercepts (α_r), observer-effects (ω_j), a first-year observer-effect ($\eta I[j,t]$), and route-level slope parameters (β_r) for the continuous effect of year (t) centered on the mid-year of the time-series (t_m).

For the parameters that were common to all models, we estimated observer effects drawn from a normal distribution with estimated variances ($\omega_j \sim Normal(0, \sigma_\omega)$), the inverse dispersion parameter as the inverse of a half, standard t-distribution with 3 degrees of freedom ($\phi \sim |t(3,0,1)|^{-1}$), and the first-year

observer-effect η, as an independent parameter with a weakly informative prior ($\eta \sim Normal(0, 1)$). All other parameters were estimated as hierarchical effects, sharing information among routes or among observers. To encourage convergence, we constrained each of the random effects in the model, including the spatial route-level parameters, to sum to zero. These constraints often improved model sampling efficiency, but they do not affect the interpretation of the final route-level slopes or intercepts. The models here varied only in the estimation of the route-level intercepts and slope terms. Three of the models used spatial information to estimate the intercepts and slopes (i.e., effectively shrinking towards a local mean of neighboring routes), while the fourth model estimated the intercepts and slopes as simple exchangeable random effects (i.e., shrinking towards a global mean of all routes).

To estimate route-level abundance, while accounting for variation among observers, we modeled separate intercepts for routes and observers. Using separate observer and route effects has not been commonly included in hierarchical Bayesian models for the BBS (Sauer and Link 2011, Smith et al. 2014, Link et al. 2020, Edwards and Smith 2021), until recently (Betts et al., 2022, Smith et al., 2023). In general, observers and routes are correlated in the BBS dataset, by design as an experimental control for variation among observers (Kendall et al. 1996). However, observers and routes vary in the number of surveys conducted and the database still contains a lot of information on variation among routes and among observers: considering only the years modeled here (2006-2021), more than 69% of surveys were conducted on routes that have had more than one observer during those 15 years, and 55% of surveys were conducted by observers who have surveyed more than one route. Separating observer from route effects is also possible due to the added spatial information included in the route-level intercept estimate, the sum to zero constraints in the model parameterization, the weakly informative priors that constrain parameters to plausible values given the log-link model, and the improved efficiency of the Hamiltonian Monte Carlo (HMC) samplers in Stan (Betancourt 2018, Stan Development Team 2022) over the Markov Chain Monte Carlo (MCMC) samplers in earlier probabilistic programming languages such as JAGS (Plummer 2003). Finally, we also used an informative prior on the standard deviation of the observer

effects (σ_ω), and we ensured that all parameters had converged when fitting the models (details below). We used a half-normal prior on the standard deviation among observers, scaled to imply that variation among observers is unlikely to result in variation in mean counts greater than a factor of approximately six (i.e., it is very unlikely that a change in observer on a route will result in a six-fold increase, or reduction, in a given species abundance; $\sigma_\omega \sim |Normal(0,0.3)|$), and that variation among observers is less than variation among routes. We suggest this prior is reasonable given that BBS observers are highly skilled and familiar with the local bird community (Link and Sauer 1997).

Spatial structures

We fit models with two different approaches to account for spatially explicit relationships among routes: 1) an intrinsic Conditional Autoregressive (iCAR) structure that uses a sparse matrix of adjacencies between pairs of routes, treating spatial relationships as a series of discrete neighbors; and 2) an isotropic Gaussian process (GP) model that uses a matrix of Euclidean distances separating the start locations of each BBS route, treating distance between routes as a continuous measure of separation.

We used these two different approaches because the spatial locations of BBS observations are not perfectly represented by either discrete areas or points in space. It is not obvious whether the iCAR or the GP better reflects reality (Pebesma and Bivand 2023), because the observations from a given BBS route are collected along a transect that is approximately 40 km long. Both approaches are necessary simplifications of the true spatial processes underlying variation in abundance and trends among BBS routes. The iCAR approach (also used for the spatial relationships in the BYM model) simplifies the spatial structure by assuming each route represents a discrete area of space (i.e. a polygon surrounding the route), but the neighboring routes may be separated by a wide range of distances depending on the spatial distribution and spatial density of those routes. The GP approach simplifies spatial relationships by assuming each route represents a point in space, but the measure of intervening distance only applies to the distance between the start points of the routes, not to the full transect. To illustrate the differences between the approaches, the GP considers the abundance or trends of two distant routes as effectively

independent if the distance is large enough relative to the estimated distance decay function. In contrast, the iCAR structure considers these same two routes as having a very close connection if there are no intervening routes. In some cases, treating two relatively distant routes as close neighbors may be useful if their relative proximity could inform the parameter estimates, but may also introduce error into the estimate of spatial variance (Pebesma and Bivand 2023).

We used a Voronoi tessellation to generate the discrete neighborhood relationships required to support the iCAR model (Ver Hoef et al. 2018, Pebesma and Bivand 2023). iCAR models are often applied to contiguous area-based stratifications, such as regular grids, census regions, or political jurisdictions, which have natural neighborhood relationships defined by their adjacencies (Ver Hoef et al. 2018, Meehan et al. 2019). To generate contiguous discrete spatial units without imposing a regular grid structure, we used a Voronoi tessellation to create contiguous polygons, centered on the start point of each BBS route (Pebesma 2018). We further limited the adjacency matrix to the approximate boundaries of the species' range by clipping the tessellated surface using the standard BBS analytical strata where the species occurs (Sauer and Link 2011) and a concave polygon surrounding start locations of all routes with data for that species (Gombin 2023). This clipping ensured that adjacency relationships did not extend beyond the borders of the species' range and allowed the adjacency matrix to respect large-scale, complex range boundaries (e.g., gaps in forest bird ranges created by the great plains). Within the species' range boundaries, routes were considered neighbors if their Voronoi polygons shared a linear segment along a separating boundary (an edge; Figure 1). This approach to generating adjacency relationships distorts Euclidian space to create a matrix of relative spatial relationships because some neighboring routes may be much further apart than others. However, it is sufficiently flexible to ensure a comprehensive and contiguous network of among-route links, and accurately represents those relative spatial adjacencies (i.e. each route is considered adjacent to its nearest neighbors). We separately modeled the spatial dependence of intercept parameters (route mean relative abundance) and slope parameters (route trends), under the assumption that each parameter may be influenced by different ecological processes acting at different

spatial scales. To fit the GP model, we used a squared exponential kernel to model the isotropic distance-based decline in covariance, assuming that the covariance declines exponentially in all directions with the squared distance between each BBS route's start point. We adapted functions and code in the Stan probabilistic programming language from the "rethinking" R-package for inclusion in our GP model (McElreath 2023). Similar to the iCAR approach, we used independent GPs to model the covariance of the intercept parameters and the slope parameters. We estimated the full matrix for between-route distances using functions in the "sf" package for R (Pebesma 2018).

Intrinsic Conditional Autoregressive model - iCAR

We estimated the route-level intercepts and slopes using an iCAR structure, where the parameter for route-r is drawn from a normal distribution, centered on the mean of that parameter's values in all neighboring routes, with an estimated standard deviation that is proportional to the inverse of the number of neighbors for that route (Morris et al. 2019). Specifically, the intercept term that represents the mean relative abundance on each route (α_r) is estimated as an additive combination of a species-mean (α_r') and a random route-level term (α_r'') drawn from a normal distribution centered on the mean of the intercepts for all neighboring routes ($\alpha_r = \alpha' + \alpha_r''$).

$$\alpha_r'' \sim Normal\left(\frac{\sum_{n \in N_r} \alpha_n''}{N_r}, \frac{\sigma_{\alpha''}}{N_r}\right)$$

The slopes representing the trend on each route (β_r) were estimated similarly as a species-level mean trend plus random route-level terms from a normal distribution centered on the mean of the slopes for all neighboring routes ($\beta_r = \beta' + \beta_r''$).

$$\beta_r'' \sim Normal\left(\frac{\sum_{n \in N_r} \beta_n''}{N_r}, \frac{\sigma_{\beta''}}{N_r}\right)$$

Besag York Mollié iCAR model - BYM

We used an implementation of the Besag, York, Mollié (BYM) spatial iCAR model (Besag et al. 1991) to estimate route-level slopes. This model is an elaboration on the iCAR model where we estimated the slopes as additive combinations of a spatial random effect and a non-spatial random effect (Besag et al. 1991).

$$\beta_r = \beta' + \beta''_{space_r} + \beta''_{non-space_r}$$

We estimated the spatial component using the same structure as for the iCAR model.

$$\beta''_{space_r} \sim Normal\left(\frac{\sum_{n \in N_r} \beta''_{space_n}}{N_r}, \frac{\sigma_{\beta_{space}}}{N_r}\right)$$

We estimated the additional non-spatial component as a simple random effect drawn from a normal distribution with an estimated standard deviation:

$$\beta''_{non-space_r} \sim Normal\left(0, \sigma_{\beta_{non-space}}\right)$$

The additional random effect included in the BYM model allows the route-level trend estimates to vary more among neighboring routes, if supported by the data (Besag et al. 1991).

Gaussian Process model - GP

In the Gaussian Process (GP) model, the intercepts and slopes were also estimated as the sum of a route-level random term and a species-level mean ($\beta_r = \beta' + \beta''_r$). The slope and intercept random terms for each route (β''_r and α''_r) are estimated as zero-mean, multivariate normal distributions, $\beta'' \sim MultivariateNormal(0, K_\beta)$ and $\alpha'' \sim MultivariateNormal(0, K_\alpha)$, with covariance matrices (K_β and K_α) estimated using a squared exponential kernel function (Gelman et al. 2013, pg 501). The

covariance of the slope parameters for two routes ($k_\beta(\beta_1'', \beta_2'')$) is a function of the distance between them ($d_{1,2}$) plus the two parameters that control the magnitude of the covariance when distance is zero (θ_β) and the scale of the spatial dependency (ρ_β).

$$k_\beta(\beta_1'', \beta_2'') = \theta_\beta^2 * e^{\left(-\rho_\beta^2 * d_{1,2}^2\right)}$$

We estimated the intercept parameters using the same squared exponential kernel function with separate parameters for the magnitude and scale of the spatial dependency.

$$k_\alpha(\alpha_1'', \alpha_2'') = \theta_\alpha^2 * e^{\left(-\rho_\alpha^2 * d_{1,2}^2\right)}$$

The parameters of GP models can be quite sensitive to prior distributions (McElreath 2020). We scaled the distance matrix in units of 1000 km and set a half-standard t-distribution prior on θ_α^2 and θ_β^2 with 5 degrees of freedom (Gelman et al. 2013). The half-t prior on θ^2 places most prior density at relatively small values and includes a relatively long tail that allows for larger values, if supported by the data. For most species, we used a weakly informative inverse gamma distribution prior with scale and shape = 5 for ρ_α^2 and ρ_β^2. For some species, the values of ρ^2 failed to converge with this prior, so we set an alternative and more informative prior using a gamma distribution with scale and shape = 2. The gamma and inverse gamma priors on ρ^2 both avoid 0, ensuring that spatial dependency decreases with distance. The weakly informative inverse gamma includes a long right tail that allows the model to estimate spatial dependency that declines steeply with distance (e.g., $\rho_\alpha^2 > 500$ and therefore covariance values near 0 for routes separated by the approximate 40-km length of a BBS route), but for some species, this long tail created convergence difficulties. For these species, we used the gamma prior with a shorter right tail and effectively constrained the estimates of ρ^2 to values < 20. This places most of the prior density at values that imply there is some spatial dependency that may extend out to larger distances (500 km – 3000 km).

Non-spatial model

To assess the benefits of assuming spatial dependence among BBS routes, we compared the predictions and predictive accuracy of the spatial models to an otherwise identical model that lacked spatial information. This non-spatial model had all the same parameters as the spatial models, except that the route-level intercepts and slopes were estimated as simple random effects.

$$\beta_r'' \sim N\left(0, \sigma^2_{\beta_{non-space}}\right)$$

$$\alpha_r'' \sim N\left(0, \sigma^2_{\alpha_{non-space}}\right)$$

Remaining priors

We used weakly informative (Gelman 2006, Lemoine 2019) standard normal priors for the mean species-level intercept and the first-year effect parameter. The mean species-level slope parameter was given a weakly informative normal prior ($\beta' \sim Normal(0, 0.1)$). We consider this prior weakly informative as it reflects our belief that extreme rates of change are unlikely (it places approximately 95% of the prior density for the survey-wide population trends between -20 and +20%/year).

For the iCAR, BYM, and non-spatial models, the priors for the standard deviations of the spatial variation and non-spatial variation of the route-level slopes ($\sigma_{\beta_{spatial}}$ and $\sigma_{\beta_{non-space}}$) had gamma priors with shape = 3 and scale = 30. These gamma priors were weakly informative such that the standard deviation of trends was constrained to more probable scales based on the log-link of the model and to avoid estimates of zero (Chung et al. 2013). Specifically, this gamma prior places the mean of the prior density at approximately 10% per year, and 99% of the prior density on the standard deviation of route-level trends at values less than 28% per year, while also including a long tail so that the model can estimate more extreme variation, if supported by the data (Chung et al. 2013). The standard deviation of the intercept terms in these models ($\sigma_{\alpha_{spatial}}$ and $\sigma_{\alpha_{non-space}}$) were given a half-normal prior with standard deviation = 2. This weakly informative prior placed most prior density at values < 5, and reflects our belief that across

347 a species' range, mean relative abundance for a fixed survey effort can vary a great deal but is unlikely to
348 vary by more than a few orders of magnitude (Fink et al. 2023a). For some species, this relatively wide
349 prior created convergence issues, so for these species we re-fit the models with a prior that considered the
350 observed variation in mean counts among routes for a given species. Specifically, we used a half-normal
351 prior with the standard deviation equal to the observed standard deviation of mean log-transformed
352 observed counts among routes. We are confident that this prior is only weakly informative and likely
353 over-estimates the among-route variance because the observed data includes variation among routes, as
354 well as variation among observers.

355 ### Convergence

356 We fit all models using 1000-2000 warmup iterations and an equal number of sampling iterations for each
357 of the four independent chains (or three independent chains for each iteration of cross-validation). We
358 assessed convergence by monitoring for divergent transitions and estimating split-Rhat values and bulk
359 effective sample sizes for all parameters. We considered convergence to have failed if any Rhat was >
360 1.03 or if any parameter's effective sample size is < 100 (although the vast majority of parameters had
361 effective sample sizes > 1000 and Rhat < 1.01). If any models failed to converge, we re-fit the models
362 with the alternative priors described above.

363 ### Model assessment

364 To assess the benefits of adding spatial information into the model, we compared the 1-step-ahead, leave-
365 future-out (LFO) predictive success of the four models for the 71 species with relatively small ranges
366 (Roberts et al. 2017, Bürkner et al. 2020). We also compared the iCAR spatial model with the non-spatial
367 version of the model using a LFO assessment for an additional 216 species (Table 1). We used the LFO
368 approach to directly test the temporal predictions of the models (i.e. test the accuracy of predictions of
369 next year's observations). In this application of LFO, we fit the model to the first eight years of data
370 (2006-2013; the minimum length of time we considered sufficient for prediction), and used the parameter
371 estimates from this model to predict the counts in the following year (2014). Then we iterated this

approach making predictions for the remaining years (2015-2019, and 2021), predicting the observed data in year n using data for all years up to year (n-1) to fit the model. We could not assess predictive accuracy for the year 2020 because the BBS survey season was canceled due to concerns over COVID-19. The cross-validation process generated predictions for every count in the dataset and an estimate of the log pointwise predictive density (lppd) of the observed count, given the model and the data in all previous years (Gelman et al. 2014). For interpretation and visualization, we calculated pairwise differences in lppd between pairs of models for each count and transformed summaries of these lppd differences across many counts into approximate z-scores (mean divided by the standard error of the point-wise differences in lppd). These z-scores provide a way to summarise the support in the data for each model, accounting for the variation across all observations and providing an interpretable and consistent scale to summarize pair-wise model comparisons across species with different numbers of observations (Link and Sauer 2016). They are an approximation of the test statistic in a paired t-test; e.g., absolute values greater than approximately 2 could be interpreted as a "significant difference" in predictive success, although we put little emphasis on this kind of arbitrary threshold in our interpretation here.

Route-level covariate examples

Modeling covariates of finer-scale trends and relative abundances is a major benefit of modeling BBS trends and abundances at the route level. To demonstrate this, we present two examples, each including route-level predictors to inform estimates of abundance and trend. The first example uses data on the Rufous Hummingbird (*Selasphorus rufus*) and models the effect of habitat suitability on relative abundance and trend. The second example looks at the effects of annual variation in available habitat—the number of ponds surrounding a BBS route in a given year in the Prairie Pothole region (PPR)—on the expected counts of a water bird, the Horned Grebe (*Podiceps auratus*).

Rufous Hummingbird covariate example

This example application is an elaboration of the iCAR route-level trend model, where the route-level intercepts and slopes are additive combinations of two components: 1) one that is a function of a route-

level predictor, and 2) one that is a residual component, estimated using the iCAR structure (Ver Hoef et al. 2018). The route-level predictors are derived from a previous study on Rufous Hummingbirds that modeled variation in habitat suitability over space and time. Habitat suitability was estimated using an independent dataset of bird observations and annual remotely sensed data on weather, climate, landcover, and elevation. More detail on the methods used to estimate habitat suitability is and the model used here is available in the supplemental methods. Mean habitat suitability across all years in a 200m buffer surrounding each BBS route was used as a predictor on the intercept (i.e., mean habitat suitability as a predictor on the mean relative abundance on a given route). The rate of change in habitat suitability over time within the same buffer was used as a predictor on the slope (i.e., change in habitat as a predictor on the trend in the species' abundance). This model structure relies on relatively simple assumptions that the amount of habitat around a BBS route should predict the mean number of birds observed, and that the change in habitat amount should predict the change in the number of birds.

We estimated the route-level intercepts and slopes as an additive combination of a mean species-level intercept or slope (α' or β'), a varying intercept or slope that was a function of the mean habitat suitability on the route (α_r''') or rate of change in habitat suitability on the slope (β_r'''), and spatially varying effects for the residual variation in relative abundance (α_r'') and slope (β_r'') that were not explained by habitat.

$$\alpha_r = \alpha' + \alpha_r'' + \alpha_r'''$$

$$\beta_r = \beta' + \beta_r'' + \beta_r'''$$

This partitioning of the intercept and slope parameters allows the model to generate two alternative estimates of the mean abundance and trend on each route. The full trend ($\beta' + \beta_r'' + \beta_r'''$) represents the estimated trend on a given route, including the effects of habitat change. The residual trend ($\beta' + \beta_r''$) represents an alternate trend if habitat suitability stayed constant on a given route. Similarly, the full relative abundance ($\alpha' + \alpha_r'' + \alpha_r'''$) represents the estimated relative abundance on a given route,

including the effects of habitat change. The residual relative abundance ($\alpha' + \alpha_r''$) represents an alternate abundance that we would expect if habitat suitability stayed constant across all routes.

We estimated the effect of mean habitat suitability on the route-level intercept as a simple product of a route-specific coefficient (ρ_{α_r}) and mean habitat suitability (across all years; 2006-2021) in a 200m buffer around each route-path ($\alpha_r''' = \rho_{\alpha_r} * MeanSuitability_r$). To model the effects of habitat change on population trend, we estimated the effect of the rate of change in habitat suitability on each route ($ChangeSuitability_r$) with a route-specific coefficient (ρ_{β_r}). The route-specific coefficients for the effects of habitat suitability on the intercept and slope were allowed to vary among routes, but were centered on hyperparameter mean effects across routes $\rho_{\alpha_r} \sim Normal(P_\alpha, \sigma_{\rho_\alpha})$ and $\rho_{\beta_r} \sim Normal\left(P_\beta, \sigma_{\rho_\beta}\right)$. As such, the hyperparameters for the effect of mean habitat suitability on the intercept (P_α) and the effect of change in habitat suitability on slope (P_β) represent a clear species-level estimate of the effects of habitat change on abundance and trend.

Horned Grebe covariate example

This example application was an elaboration of the iCAR route-level trend model, where trends and relative abundances are estimated while accounting for the annual variation in climatically dependent habitat. The route-level predictors are derived from a study of the effects of moisture/drought patterns on Horned Grebe (more detail in the supplemental methods), a waterbird species that breeds in small to moderately sized shallow, freshwater ponds (Stedman 2020). To represent annual variation in available habitat for the Horned Grebe in the Canadian Prairie Pothole Region (PPR), we used data collected by the U.S. Fish and Wildlife Service (USFWS) and the Canadian Wildlife Service (CWS) on the number of ponds (natural or artificial ponds that are flooded seasonally, semi-permanently, and permanently) during the Waterfowl Breeding Population and Habitat Survey (Smith 1995). Annual fluctuations in moisture affect the number of wetlands available, which in turn has a strong influence on waterbird populations

that are highly dependent on wetlands abundance (Sorenson et al. 1998, Johnson et al. 2005, Roy 2015, Steen et al. 2016). The model was based on the iCAR model and added an additional iCAR component to create a varying-coefficient model on the effects of available habitat on the observed counts during a given survey on a given route.

We estimated the effect of the number of ponds in a buffer surrounding BBS routes as a spatially-varying coefficient representing the route-specific effect of local ponds ($\rho_r * ponds_{r,t}$). Local ponds are the number of ponds surrounding a BBS route each year, where $ponds_{r,t}$ represents the log(1 + number of ponds) surrounding BBS route r in year t, centered on the mean number of ponds across years for each route. This route-specific centering ensured we could separately estimate the route-level intercepts and the effects of the annual variations in ponds and ensured that it only represented the temporal variation in ponds and not the spatial variation. The effects of ponds per route were centered on a mean hyperparameter (ρ_r') and allowed to vary among routes using the same iCAR spatial structure as for the slopes and intercepts (ρ_r'').

$$\rho_r = \rho_r' + \rho_r''$$

$$\rho_r'' \sim Normal\left(\frac{\sum_{n \epsilon N_r} p_n''}{N_r}, \frac{\sigma_{p''}}{N_r}\right)$$

Finally, we also fit the same data to the simple iCAR model (i.e., an identical model with no covariates) to compare the difference in estimated trends with and without accounting for the annual variations in available habitat.

Results

In general, there are clear spatial patterns in the estimated trends and relative abundances from the spatial models, with similar patterns among the three types of spatial models. Those patterns are obscured or completely lacking from the non-spatial version of the model (e.g., the results for Baird's Sparrow in

Figures 2 and 3). The GP model tended to smooth the spatial pattern in trends more than the iCAR model, which in turn smoothed more than the BYM model (Figure 2). The spatial smoothing in relative abundance was stronger in both the iCAR and BYM models than the GP model for Baird's Sparrow (Figure 2). The covariance in relative abundance of Baird's Sparrow among routes was effectively 0 at distances of only 100 km (posterior mean of ρ_{α}^{2} = 650), whereas the covariance in trend was relatively strong even at distances > 1000 km (posterior mean of ρ_{β}^{2} = 1.5, Figure S1). Predictions of route-level trends had smaller standard errors when including spatial information, and trend precision generally increased with the degree of spatial smoothing (Figure S2). For Baird's Sparrow, all three spatial models had better predictive accuracy than the non-spatial model, with z-scores of pairwise differences between one of the spatial models and the non-spatial model ranging from 2.7 – 3.3 (Figure S4). The iCAR model had better predictive accuracy than the BYM model (z-score of the difference = 3.8; Figure 4), and there was little difference in predictive accuracy between the iCAR and GP models (z-score difference = -0.51; Figure 4).

The leave future out (LFO) cross-validation shows that the iCAR and GP models out-perform (i.e. more accurately predicted next-year's data) the non-spatial model for almost all the 71 small-range species (Figure 4 and Figure S4). The BYM model had lower predictive accuracy than the other spatial models. It had lower accuracy than the iCAR model for all species and was the only spatial model that had clearly lower predictive accuracy than the non-spatial model (i.e., four species for which the z-score difference is < -2, Figure 4 and Figure S4). The iCAR model and the GP model had similar predictive accuracy for many species; 69% (49 of 71 species) of the species were better predicted by the GP model and the remaining species were better predicted by the iCAR model (Figure 4). When including the additional 216 species for which fitting the GP model was prohibitively time-consuming (days or even weeks are required for convergence for a given species), the iCAR model had higher predictive accuracy than the non-spatial model for 283 of 287 species, and predictive accuracy was very similar for the remaining four (Figure 5).

The iCAR model generated trend prediction maps with clear spatial patterns that likely relate to spatially dependent variation in processes affecting populations (Figure 6). These patterns are not evident in predictions from an identical model without spatial information (Figure 6). The spatial patterns in route-level trends vary widely among species (Figures S3 and S5), suggesting varied drivers of population change across the continent and among species.

In general, the iCAR and GP models were comparable in predictive accuracy for the 71 small-range species we analyzed (Figure S9). In addition, the spatial patterns in predicted trends were very similar between these two models, even for species where the predictive accuracy differed between the models (Figure 7). For example, the GP model had higher predictive accuracy than the iCAR model (z-score difference = -4.3, Figure S4) for Canyon Towhee (*Melozone fusca*), but the opposite was true for Western Bluebird (*Sialia mexicana*; z-score difference = 2.3, Figure S4). Regardless, the spatial pattern in predicted trends between the two models is quite similar for both species (Figure 7 and Figure S3). For both species, and in general, the GP model trend estimates had narrower credible intervals (higher estimated precision) than the iCAR model (Figure S6). Precision of the iCAR trend estimates also showed a clear relationship to the number of neighbors for any given route, in that routes with few neighbors (on the edges of the species' range) were much less precise than estimates in the core of the species' range (Figure S6).

Including habitat suitability in the Rufous Hummingbird population model had an effect on estimates of route-level abundance and improved estimates of the spatial pattern in long-term trends (Figure 8). However, much of the overall decline was not related to covariates describing route-level habitat change, as the negative population trends across the species' range remained after removing the effects of local habitat change covariates (right panel, Figure 8). The effect of habitat suitability on mean relative abundance was strong and positive (P_α = 3 [95% CI 2.2:3.8]), such that routes with higher overall habitat suitability had higher mean counts. From 2006-2021, the Rufous Hummingbird's overall population declined steeply, decreasing by approximately -43% (95% CI -52:-33). There was an effect of change in

habitat suitability on trends, such that routes with habitat loss had more negative population trends P_β = 0.025 (95% CI 0.003:0.047). Trends were negative across the species' range, but most negative in the coastal regions where the habitat has changed the most and where the species is also most abundant (left panel, Figure 8, and Figure S7). The change in habitat suitability affected the spatial patterns in trend (Figure 8), with the greater loss of habitat in the coastal regions (Figure S7) accounting for most of the increased rates of decline in the core of the species' range. The residual trend component alone does not show the same coastal-decline pattern (right panel, Figure 8).

Annual variation in the number of ponds around BBS routes affected the overall rate of population change in Horned Grebes and showed a spatial relationship (Figure 9). In a model including the annual pond variation, the Horned Grebe population declined overall at a rate of -1.9 %/year from 1975-2017. After removing the effect of annual pond variation, the long-term rate of decline was -2.2 %/year. The effect of annual fluctuations in the number of ponds was positive across the region: the mean value of P = 0.42 (95% CI 0.29:0.55), but there was also a spatial gradient in intensity. The effect of the number of ponds per year was strongest in the northwest part of the Prairies (Figure 9) and declined to the south and east.

Discussion

Spatially explicit, route-level models are useful for visualizing fine spatial patterns at scales more relevant to local conservation, understanding the drivers of population change, and estimating the effects of covariates on relative abundance and trends (e.g., Betts et al. 2022). At this fine spatial scale, incorporating spatial information improves the models' predictions of future data. This improvement is particularly clear for both the iCAR and the GP models, where the spatial models had higher accuracy for out-of-sample predictions than the non-spatial model for almost every species we compared. Fine spatial patterns in trend estimates across a species' range are useful for generating hypotheses on the ecological drivers of population change. Route-level models also allow for the incorporation of local habitat covariates on abundance and trend at fine scales, which is important as some covariates affect bird populations at scales much smaller than the strata often used for broad-scale analyses, such as Bird Conservation Regions (BCRs) or states/provinces/territories (Thogmartin et al. 2004, Paton et al. 2019, Monroe et al. 2022). Route-level patterns are also useful in guiding conservation and/or further monitoring efforts, such as identifying small areas for conservation purposes or diverging population trends within management areas (i.e., strata or BCR).

These route-level, spatial models generate smoothed patterns of variation in population trends across a species' range, which will greatly facilitate hypothesis generation and direct investigation to better understand the drivers of population change similar to (Fink et al. 2023b). For example, the spatial models show relatively smooth patterns in Baird's Sparrow trends across the species' range (Figure 2), which are not evident in the simpler, non-spatial model. In the spatial models, Baird's Sparrow has increased in the west and decreased in the eastern portion of its range. This latitudinal pattern may suggest hypotheses related to spatial variation in factors such as climate, or habitat amount, which could then be directly tested by incorporating covariates representing these factors into a subsequent model. Similarly, the complex spatial patterns in the trends of Hairy Woodpecker (*Dryobates villosus*, Figure 6) show some latitudinal variation in trends in the west that is not as clear in the east, suggesting that there may be

560 distinct processes driving trends in these two regions. Comparisons of these patterns among species may
561 be particularly informative. For example, the similar southeast to northwest gradients in trends for
562 Canyon Towhee and Western Bluebird may suggest some similarity in the underlying drivers of
563 population change (Figure 7). These observations are meant to illustrate the types of hypothesis-
564 generating that these fine-scale, spatially explicit models can help generate.

565 All three of the spatial models (iCAR, GP, and BYM) generated broadly similar spatial patterns in route-
566 level trends for the subset of species we compared (Figure 4 and Figure S3). The best spatial structure to
567 use will depend on the species and the goals of a given study. For the species in this study, there is little
568 support for the extra variation in route-level trends in the BYM model, given it had lower predictive
569 accuracy than the simpler iCAR model in all cases. The iCAR structure outperformed the GP models for
570 31% of the species, and is more computationally efficient. Overall, the GP model outperformed the iCAR
571 model for most (69%) of the species we compared. The GP model also produces smoother spatial patterns
572 in population trends than the other spatial models and for some, the difference is striking (e.g. Black-
573 throated Gray Warbler *Setophaga nigrescens*, California Quail, *Callipepla californica,* and the Golden-
574 winged Warbler *Vermivora chrysoptera* in Figure S3). For the first two species, the GP outperformed the
575 iCAR for accuracy, while for the third species, the iCAR was better (Figure S4). Until GP models become
576 more efficient to implement (Hoffmann and Onnela 2023), the iCAR structure may be preferable for
577 larger datasets (e.g., broad-ranging species and or longer time-series). The iCAR structure may also
578 provide more direct control to model discontinuities in the spatial relationships, such as complex range
579 boundaries (Ver Hoef et al. 2018, Pebesma and Bivand 2023), since there are many ways to define
580 neighborhood relationships (Freni-Sterrantino et al. 2018). A species with limited dispersal may be
581 particularly sensitive to the Euclidean distance between points and therefore better modeled with the GP,
582 but the simplification of space using the iCAR structure may be sufficient for most wide-ranging
583 migratory birds. For example, for some species, there are routes on the periphery of the BBS sampling
584 distribution or the periphery of a species' range that are separated from most other routes by relatively

large distances. These “isolated” routes are treated very differently by the iCAR and GP models: they are considered close neighbors in the iCAR model irrespective of the intervening distance, whereas in the GP model, the large separation from other routes reduces their correlation with their nearest neighbors. Interestingly, when we compared the predictive accuracy between GP and iCAR models for routes that are more isolated than most (nearest neighboring route where the species was detected > 200km away), the simplified relative-spatial relationships of the iCAR tend to outperform the continuous spatial treatment of the GP for these isolated routes (Figure S8).Therefore, the more accurate representation of the long distances separating these isolated routes in the GP model does not necessarily result in more accurate predictions, and in some cases it may be more effective to treat space as a series of relative spatial relationships.

These route-level BBS models provide many opportunities for further comparisons, applications, and elaborations. Fine-scale estimates could be summarized across species and within regions, such as summaries of the spatial patterns in grassland bird trends or summaries for a given species within BCRs or states/provinces/territories and compared to estimates from models fit at those broader spatial scales. The spatial patterns in trend estimates also allow for comparison of BBS data to other fine-grained maps of species trend and relative abundance, such as eBird (Sullivan et al. 2014, Fink et al. 2023a) or the Integrated Monitoring in Bird Conservation Regions (IMBCR) program (Pavlacky et al. 2017). Comparison of trend estimates between the two programs for the same species and periods of time could provide useful validation of and or help understand differences between the two sources of information. Similarly, there are many possible avenues to integrate information across programs for a given period (e.g., recent trends) or through time (e.g., long-term information from the BBS with more recent information from eBird and/or IMBCR). We see an almost limitless potential for customizing route-level BBS models to include covariates testing hypotheses of drivers of population abundance and trends (e.g., Betts et al. 2023). The examples of covariate models in addition to our application of LFO cross-

validation will hopefully provide useful tools to better understand the causes of population change in North American birds.

Separating the route-level intercepts from the observer-level intercepts allows us to better model patterns in relative abundance. It should also allow for improved modeling of among-observer variation. Although many previous BBS analyses have treated each observer-route combination as an independent sampling unit (Link et al. 2020, Smith and Edwards 2020), doing so necessarily allocates some of the biological variation in abundance in space (i.e., among-route variation in abundance) to an effect that is treated as sampling noise (among-observer variation). The model will struggle to separately estimate intercepts for observers and routes in situations where there are few data to inform the estimates (e.g., intercepts for observers who only contribute data to a route that has never been surveyed by another observer). However, we suggest that a model that includes a few of these weakly estimable parameters is likely preferable to a model that fails to attempt to separate the biological variation among routes from the sampling noise of observer variation, at least in the situations where there are data to support their separation. In a practical sense, this separation of the observer from route effects is improved by the hierarchical structure of the models, spatial information, weakly informative priors, and the improved efficiency of HMC algorithms over the Gibbs sampling algorithms of earlier Bayesian BBS models. Although initially motivated by our desire to directly model route-level abundance, this approach is equally applicable to other BBS analyses (Smith et al. 2023), and is included in the models in the R-package bbsBayes2 (Edwards et al. 2023).

In both covariate examples, incorporating spatial covariates into the trend analyses tested hypotheses related to the drivers of population change and helped identify specific areas for further research and conservation action. For the Rufous Hummingbird, the model shows higher mean abundance on routes with more habitat suitability and positive effects of the change in habitat suitability on the species' trend (more negative trends on routes where habitat has decreased). Interestingly, it also shows that during this period, the variation among routes in habitat change does not account for all of the decline in the species'

634 population (Figure 8, and Figure S7), suggesting that factors other than local habitat or factors acting
635 outside of the breeding range may be driving the overall decline. However, covariates other than habitat
636 suitability could represent local habitat better for the Rufous Hummingbird and by using other covariates
637 we may have had a different relationship between local habitat and abundance. For the Horned Grebe, the
638 effect of annual fluctuations in available wetland habitat (the number of ponds) is positive overall and
639 also varies in magnitude across the species' range. The effect is strongest in the western prairies where the
640 effects of drought are often strongest (Johnson et al. 2005, Millett et al. 2009, Roy 2015). These results
641 highlight the importance of continued investment in wetland conservation programs for waterbird
642 populations breeding in the Prairie Potholes Region, and the vulnerability of these species to climate
643 change since their breeding habitat is highly sensitive to climatic conditions.

644 Finer-scale modeling is useful for hypothesis generation and testing, and finer-scale estimates can
645 inform conservation at scales relevant to communities and the volunteers that collect the data. Our fine-
646 scale modeling is made possible by the structured, longitudinal data from the BBS. Complicated
647 questions can be asked due to the quality of the BBS data, such as our rufous hummingbird example,
648 where we estimated the component of trend that was a function of local, breeding-season habitat amount
649 and the remainder that was presumably a function of other factors including those operating outside the
650 breeding range, and which serves as a counter-factual trend we would expect if habitat amount had
651 remained constant. Using finer scales allows for the consideration of a different suite of potential
652 covariates and mechanisms, when compared to larger regional or national scales. Finer scale models can
653 also be used to inform different scales of decisions and communities. Decisions on land use for industries
654 such as agriculture, forestry, and housing are often made at fine scales (Sodhi et al. 2011, Malek et al.
655 2019). Likewise, habitat protection and restoration by community organizations, municipal governments,
656 and non-governmental organizations occur at fine scales (Sheppard 2005, Aronson et al. 2017). For
657 example, the Horned Grebe covariate analysis confirms the vulnerability of waterbird species in the
658 northwestern Prairie Potholes Region and supports a current initiative to protect critical shallow wetlands

in the region (Prairie Habitat Joint Venture 2020). Community support is important for the success of conservation initiatives (Berkes 2004, Bennett and Dearden 2014), and so providing estimates at scales relevant to communities may increase community support for conservation and encourage a feeling of stewardship. Further, routes are a relevant scale for the volunteer observers dedicated to the BBS, with the average BBS volunteer participating for 12 years. Producing estimates at a route-level allows volunteers' to see the direct results of their efforts over the years, a large motivator for many citizen science volunteers (Phillips et al. 2019). These fine-scale models can investigate a different set of questions than regional models and provide estimates to inform local-scale decisions and inform the ever-important BBS volunteers that enable all BBS research.

References Cited

Aronson, M. F., C. A. Lepczyk, K. L. Evans, M. A. Goddard, S. B. Lerman, J. S. MacIvor, C. H. Nilon, and T. Vargo (2017). Biodiversity in the city: key challenges for urban green space management. Frontiers in Ecology and the Environment 15:189–196.

Barnett, L. A. K., E. J. Ward, and S. C. Anderson (2021). Improving estimates of species distribution change by incorporating local trends. Ecography 44:427–439.

Bennett, N. J., and P. Dearden (2014). Why local people do not support conservation: Community perceptions of marine protected area livelihood impacts, governance and management in Thailand. Marine Policy 44:107–116.

Berkes, F. (2004). Rethinking Community-Based Conservation. Conservation Biology 18:621–630.

Besag, J., and C. Kooperberg (1995). On conditional and intrinsic autoregressions. Biometrika 82:733–746.

Besag, J., J. York, and A. Mollié (1991). Bayesian image restoration, with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics 43:1–20.

Betancourt, M. (2018). A Conceptual Introduction to Hamiltonian Monte Carlo. arXiv:1701.02434 [stat].

Betts, M. G., Z. Yang, A. S. Hadley, A. C. Smith, J. S. Rousseau, J. M. Northrup, J. J. Nocera, N. Gorelick, and B. D. Gerber (2022). Forest degradation drives widespread avian habitat and population declines. Nature Ecology & Evolution 6:709–719.

Bled, F., J. Sauer, K. Pardieck, P. Doherty, and J. A. Royle (2013). Modeling Trends from North American Breeding Bird Survey Data: A Spatially Explicit Approach. PLoS ONE 8:e81867.

Buckland, S. T., K. B. Newman, L. Thomas, and N. B. Koesters (2004). State-space models for the dynamics of wild animal populations. Ecological Modelling 171:157–175.

Bürkner, P.-C., J. Gabry, and A. Vehtari (2020). Approximate leave-future-out cross-validation for Bayesian time series models. Journal of Statistical Computation and Simulation 90:2499–2523.

Chung, Y., S. Rabe-Hesketh, V. Dorie, A. Gelman, and J. Liu (2013). A nondegenerate penalized likelihood estimator for variance parameters in multilevel models. PSYCHOMETRIKA 78:685–709.

Drever, M. C., A. C. Smith, L. A. Venier, D. J. H. Sleep, and D. A. MacLean (2018). Cross-scale effects of spruce budworm outbreaks on boreal warblers in eastern Canada. Ecology and Evolution 8:7334–7345.

Edwards, B. P. M., and A. C. Smith (2021). bbsBayes: An R Package for Hierarchical Bayesian Analysis of North American Breeding Bird Survey Data. Journal of Open Research Software 9:19.

Edwards, M. B. P., A. C. Smith, and S. LaZerte (2023). bbsBayes2: Hierarchical Bayesian Analysis of North American BBS Data. [Online.] Available at https://github.com/bbsBayes/bbsBayes2.

Fink, D., T. Auer, A. Johnson, M. Strimas-Mackey, S. Ligocki, O. Robinson, W. Hochachka, L. Jaromczyk, C. Crowley, K. Dunham, A. Stillman, et al. (2023a). eBird Status and Trends, Data Version: 2022; Released: 2023. *Cornell Lab of Ornithology, Ithaca, New York*. [Online.] Available at https://science.ebird.org/status-and-trends/species/lessca/abundance-map.

Fink, D., A. Johnston, M. Strimas-Mackey, T. Auer, W. M. Hochachka, S. Ligocki, L. Oldham Jaromczyk, O. Robinson, C. Wood, S. Kelling, and A. D. Rodewald (2023b). A Double machine learning trend model for citizen science data. Methods in Ecology and Evolution 14:2435–2448.

Freni-Sterrantino, A., M. Ventrucci, and H. Rue (2018). A note on intrinsic conditional autoregressive models for disconnected graphs. Spatial and Spatio-temporal Epidemiology 26:25–34.

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian Analysis 1:515–534.

Gelman, A., J. Carlin B., H. S. Stern, D. Dunson B., A. Vehtari, and D. Rubin B. (2013). Bayesian Data Analysis. 3rd edition. Chapman and Hall/CRC.

Gelman, A., J. Hwang, and A. Vehtari (2014). Understanding predictive information criteria for Bayesian models. Statistics and Computing 24:997–1016.

Golding, N., and B. V. Purse (2016). Fast and flexible Bayesian species distribution modelling using Gaussian processes. Methods in Ecology and Evolution 7:598–608.

Gombin, J. (2023). Concaveman package. [Online.] Available at https://github.com/joelgombin/concaveman.

Government of Canada (2010). Cosewic / Cosepac - Definitions associated with quantitative criteria. [Online.] Available at https://www.cosewic.ca/index.php/en-ca/assessment-process/wildlife-species-assessment-process-categories-guidelines/quantitative-criteria-definitions.html.

Hoffmann, T., and J.-P. Onnela (2023). Scalable Gaussian Process Inference with Stan. [Online.] Available at http://arxiv.org/abs/2301.08836.

Hudson, M.-A. R., C. M. Francis, K. J. Campbell, C. M. Downes, A. C. Smith, and K. L. Pardieck (2017). The role of the North American Breeding Bird Survey in conservation. The Condor 119:526–545.

IUCN (2012). IUCN Red List Categories and Criteria: Version 3.1. 2nd edition. IUCN, Gland, Switzerland and Cambridge, UK.

Johnson, W. C., B. V. Millett, T. Gilmanov, R. A. Voldseth, G. R. Guntenspergen, and D. E. Naugle (2005). Vulnerability of Northern Prairie Wetlands to Climate Change. BioScience 55:863–872.

Kendall, W. L., B. G. Peterjohn, and J. R. Sauer (1996). First-Time Observer Effects in the North American Breeding Bird Survey. The Auk 113:823–829.

Lemoine, N. P. (2019). Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. Oikos 128:912–928.

Link, W. A., and J. R. Sauer (1997). New Approaches to the Analysis of Population Trends in Land Birds: Comment. Ecology 78:2632–2634.

Link, W. A., and J. R. Sauer (2016). Bayesian cross-validation for model evaluation and selection, with application to the North American Breeding Bird Survey. Ecology 97:1746–1758.

Link, W. A., J. R. Sauer, and D. K. Niven (2020). Model selection for the North American Breeding Bird Survey. Ecological Applications 30:e02137.

Malek, Ž., B. Douw, J. V. Vliet, E. H. V. D. Zanden, and P. H. Verburg (2019). Local land-use decision-making in a global context. Environmental Research Letters 14:083006.

McElreath, R. (2020). Statistical Rethinking: A Bayesian Course with Examples in R and STAN. 2nd edition. Chapman and Hall/CRC, New York.

McElreath, R. (2023). rethinking. [Online.] Available at https://github.com/rmcelreath/rethinking.

Meehan, T. D., N. L. Michel, and H. Rue (2019). Spatial modeling of Audubon Christmas Bird Counts reveals fine-scale patterns and drivers of relative abundance trends. Ecosphere 10:e02707.

Millett, B., W. C. Johnson, and G. Guntenspergen (2009). Climate trends of the North American prairie pothole region 1906–2000. Climatic Change 93:243–267.

Mirochnitchenko, N. A., E. F. Stuber, and J. J. Fontaine (2021). Biodiversity scale-dependence and opposing multi-level correlations underlie differences among taxonomic, phylogenetic and functional diversity. Journal of Biogeography 48:2989–3003.

Monroe, A. P., J. A. Heinrichs, A. L. Whipple, M. S. O'Donnell, D. R. Edmunds, and C. L. Aldridge (2022). Spatial scale selection for informing species conservation in a changing landscape. Ecosphere 13:e4320.

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., R. A. Robinson, J. A. Clark, and J. A. Gill (2010). Spatial and temporal variation in population trends in a long-distance migratory bird. Diversity and Distributions 16:620–627.

North American Bird Conservation Initiative (2022). The State of the Birds, United States of America.

North American Bird Conservation Initiative Canada (2019). The State of Canada's Birds, 2019. Environment and Climate Change Canada.

Paton, G. D., A. V. Shoffner, A. M. Wilson, and S. A. Gagné (2019). The traits that predict the magnitude and spatial scale of forest bird responses to urbanization intensity. PLOS ONE 14:e0220120.

Pavlacky, D. C., P. M. Lukacs, J. A. Blakesley, R. C. Skorkowsky, D. S. Klute, B. A. Hahn, V. J. Dreitz, T. L. George, and D. J. Hanni (2017). A statistically rigorous sampling design to integrate avian monitoring and management within Bird Conservation Regions. PLOS ONE 12:e0185924.

Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal 10:439–446.

Pebesma, E., and R. Bivand (2023). Spatial Data Science: With Applications in R. 1st edition. Chapman and Hall/CRC, Boca Raton.

Phillips, T. B., H. L. Ballard, B. V. Lewenstein, and R. Bonney (2019). Engagement in science through citizen science: Moving beyond data collection. Science Education 103:665–690.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling.

Prairie Habitat Joint Venture (2020). Prairie Habitat Joint Venture: The Prairie Parklands Implemenetation Plan 2013-2020. [Online.] Available at https://www.phjv.ca/wp-content/uploads/2020/12/PHJV-Implemenetation-Plan-PRAIRIE-PARKLAND-2013-2020-Final.pdf#page=33.

Renfrew, R. B., D. Kim, N. Perlut, J. Smith, J. Fox, and P. P. Marra (2013). Phenological matching across hemispheres in a long-distance migratory bird. Diversity and Distributions 19:1008–1019.

Roberts, D. R., V. Bahn, S. Ciuti, M. S. Boyce, J. Elith, G. Guillera-Arroita, S. Hauenstein, J. J. Lahoz-Monfort, B. Schröder, W. Thuiller, D. I. Warton, et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 40:913–929.

Rosenberg, K. V., P. J. Blancher, J. C. Stanton, and A. O. Panjabi (2017). Use of North American Breeding Bird Survey data in avian conservation assessments. The Condor 119:594–606.

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.

Roy, C. (2015). Quantifying Geographic Variation in the Climatic Drivers of Midcontinent Wetlands with a Spatially Varying Coefficient Model. PLOS ONE 10:e0126961.

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 67:372–389.

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., K. L. Pardieck, D. J. Ziolkowski, A. C. Smith, M.-A. R. Hudson, V. Rodriguez, H. Berlanga, D. K. Niven, and W. A. Link (2017). The first 50 years of the North American Breeding Bird Survey. The Condor 119:576–593.

Sheppard, S. R. (2005). Participatory decision support for sustainable forest management: a framework for planning with local communities at the landscape level in Canada. Canadian Journal of Forest Research 35:1515–1526.

Smith, A. C., A. D. Binley, L. Daly, B. P. M. Edwards, D. Ethier, B. Frei, D. Iles, T. D. Meehan, N. L. Michel, and P. A. Smith (2023). Spatially explicit Bayesian hierarchical models improve estimates of avian population status and trends. Ornithological Applications:duad056.

Smith, A. C., and B. P. M. Edwards (2020). North American Breeding Bird Survey status and trend estimates to inform a wide range of conservation needs, using a flexible Bayesian hierarchical generalized additive model. The Condor. https://doi.org/10.1093/ornithapp/duaa065

Smith, A. C., M.-A. R. Hudson, C. Downes, and C. M. Francis (2014). Estimating breeding bird survey trends and annual indices for Canada: how do the new hierarchical Bayesian estimates differ from previous estimates? The Canadian Field-Naturalist 128:119–134.

Smith, G. W. (1995). A Critical Review of the Aerial and Ground Surveys of Breeding Waterfowl in North America. U.S. Department of the Interior, National Biological Service.

Sodhi, N. S., R. Butler, W. F. Laurance, and L. Gibson (2011). Conservation successes at micro-, meso- and macroscales. Trends in Ecology & Evolution 26:585–594.

Sorenson, L. G., R. Goldberg, T. L. Root, and M. G. Anderson (1998). Potential Effects of Global Warming on Waterfowl Populations Breeding in the Northern Great Plains. Climatic Change 40:343–369.

Soykan, C. U., J. Sauer, J. G. Schuetz, G. S. LeBaron, K. Dale, and G. M. Langham (2016). Population trends for North American winter birds based on hierarchical models. Ecosphere 7:e01351.

Stan Development Team (2022). Stan. *stan-dev.github.io*. [Online.] Available at //mc-stan.org/about/team/.

Stanton, R. L., C. A. Morrissey, and R. G. Clark (2018). Analysis of trends and agricultural drivers of farmland bird declines in North America: A review. Agriculture, Ecosystems & Environment 254:244–254.

Stedman, S. J. (2020). Horned Grebe (Podiceps auritus). In Birds of the World (S. M. Billerman, B. K. Keeney, P. G. Rodewald and T. S. Schulenberg, Editors). Cornell Lab of Ornithology.

Steen, V. A., S. K. Skagen, and C. P. Melcher (2016). Implications of Climate Change for Wetland-Dependent Birds in the Prairie Pothole Region. Wetlands 36:445–459.

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, et al. (2014). The eBird enterprise: An integrated approach to development and application of citizen science. Elsevier 169:31–40.

Thogmartin, W. E., J. R. Sauer, and M. G. Knutson (2004). A Hierarchical Spatial Model of Avian Abundance with Application to Cerulean Warblers. Ecological Applications 14:1766–1779.

Thompson, F. R., and F. A. La Sorte (2008). Comparison of Methods for Estimating Bird Abundance and Trends from Historical Count Data. The Journal of Wildlife Management 72:1674–1682.

Thorson, J. T., C. L. Barnes, S. T. Friedman, J. L. Morano, and M. C. Siple (2023). Spatially varying coefficients can improve parsimony and descriptive power for species distribution models. Ecography 2023:e06510.

Ver Hoef, J. M., E. E. Peterson, M. B. Hooten, E. M. Hanks, and M.-J. Fortin (2018). Spatial autoregressive models for statistical inference from ecological data. Ecological Monographs 88:36–59.

Wilson, S., S. L. LaDeau, A. P. Tøttrup, and P. P. Marra (2011). Range-wide effects of breeding- and nonbreeding-season climate on the abundance of a Neotropical migrant songbird. Ecology 92:1789–1798.

Wilson, S., A. C. Smith, and I. Naujokaitis-Lewis (2018). Opposing responses to drought shape spatial population dynamics of declining grassland birds. Diversity and Distributions 24:1687–1698.

Wright, W. J., K. M. Irvine, T. J. Rodhouse, and A. R. Litt (2021). Spatial Gaussian processes improve multi-species occupancy models when range boundaries are uncertain and nonoverlapping. Ecology and Evolution 11:8516–8527.

Figures and captions

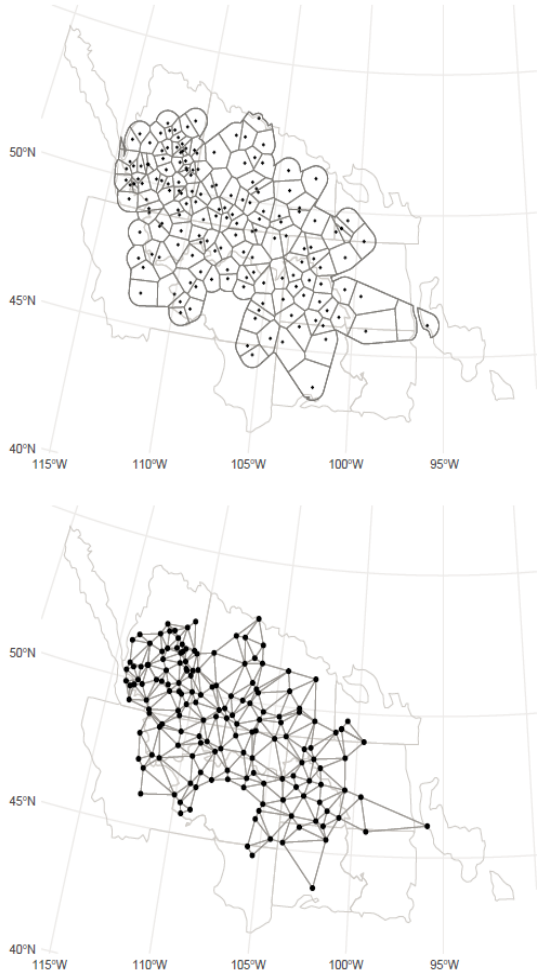

Figure 1. North American Breeding Bird Survey (BBS) route start locations (points) for routes on which Baird's Sparrow was observed (2006-2021), demonstrating the process used to identify the discrete neighbor relationships for the iCAR and BYM spatial models. The top panel shows the Voronoi tessellated surface used to assign the intervening space to the nearest BBS route start location, which is intersected with a concave polygon and the standard BBS strata (state/provinces/territories by Bird Conservation Regions). The lower panel shows routes considered neighbors using lines linking points that share an edge separating their associated Voronoi polygons.

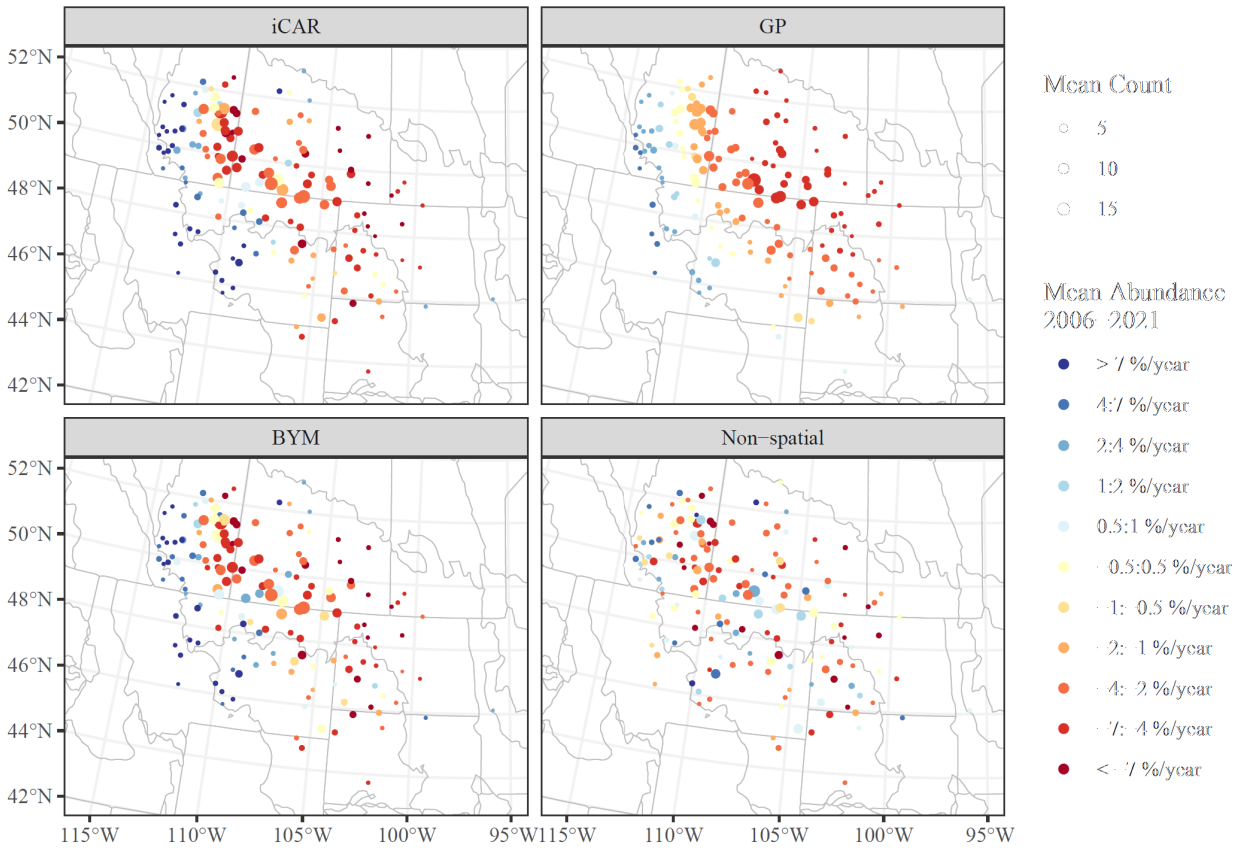

Figure 2. Estimates of trend (colors) and mean relative abundance (size of the points) for Baird's Sparrow populations on BBS routes from 2006-2021, from three spatially explicit models (iCAR, GP, and BYM) and one non-spatial model. Points with warm colors (reds) represent routes with decreasing counts through time, points with cool colors (blues) represent routes with increasing counts through time. The three spatially explicit models suggest very similar spatial patterns in trends, although the GP model suggests smoother spatial variation in trend than either the iCAR or BYM models. Grey lines within the maps represent boundaries of state/provinces/territories and Bird Conservation Regions.

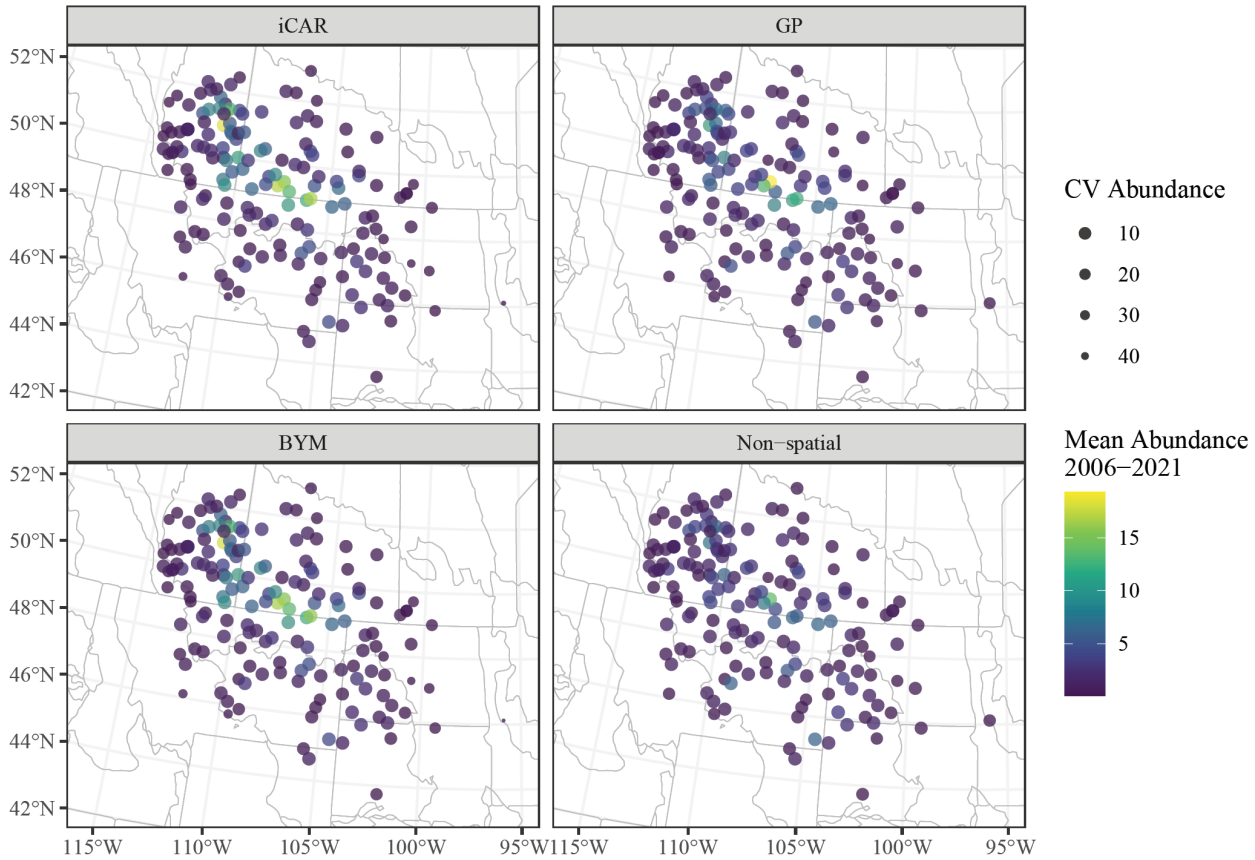

Figure 3. Estimates of mean relative abundance (colors) and the Coefficient of Variation for the estimates (CV = size) for Baird's Sparrow populations on BBS routes from 2006-2021, from three spatially explicit models and one non-spatial model. Points with brighter colors (greens and yellows) represent routes with higher estimated mean counts, and points with more precise estimates of abundance (smaller CV) are larger. The iCAR and BYM models estimate almost identical spatial patterns in abundance with a relatively clear peak in the center of the species' range, and relatively smoother spatial variation than either the GP or the non-spatial model.

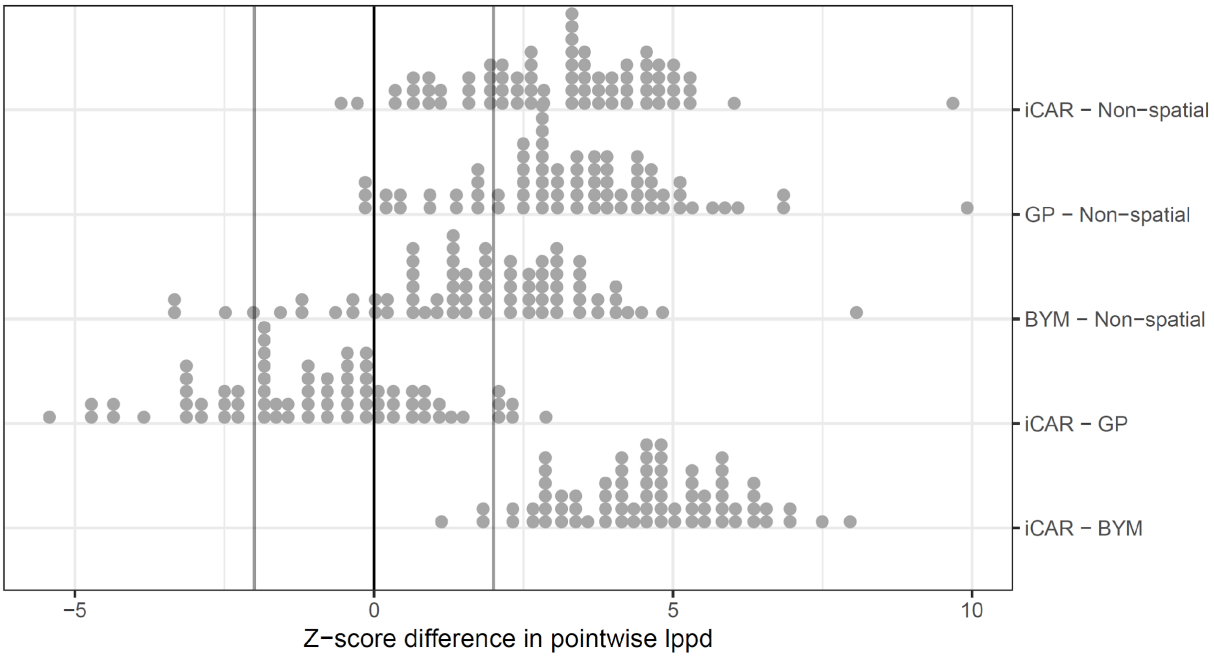

Figure 4. Leave Future Out (LFO) cross-validation results for 71 small-range species from the North American Breeding Bird Survey (BBS) database, comparing among the four different models. The stacked dot-plots represent species-level summaries of the differences in log posterior predictive density (lppd) between pairs of models. Each point represents one species for a given model comparison. Z-score values on the x-axis represent the difference between the lppd for the two models indicated on the y-axis. Z-scores > 0 (points that fall to the right of the black vertical line) represent species for which the predictive accuracy of the first model is higher than that of the second model (e.g., all but two species in the iCAR vs non-spatial comparison), and vice versa. Z-scores > 2 or < -2 (points that fall to the right or left of the vertical dark gray lines, respectively) represent species for which the mean of the differences between the two models are clear and could be considered "significant" in some statistical frameworks. The top three dot-plots show the comparisons between each of the three spatial models and the non-spatial model. The lower two plots compare the predictive accuracy among the three spatial models and show that the iCAR model out-performs the BYM model for all species, and that the GP model out-performs the iCAR model for some species but not for others. See Figure S4 for species-level comparisons.

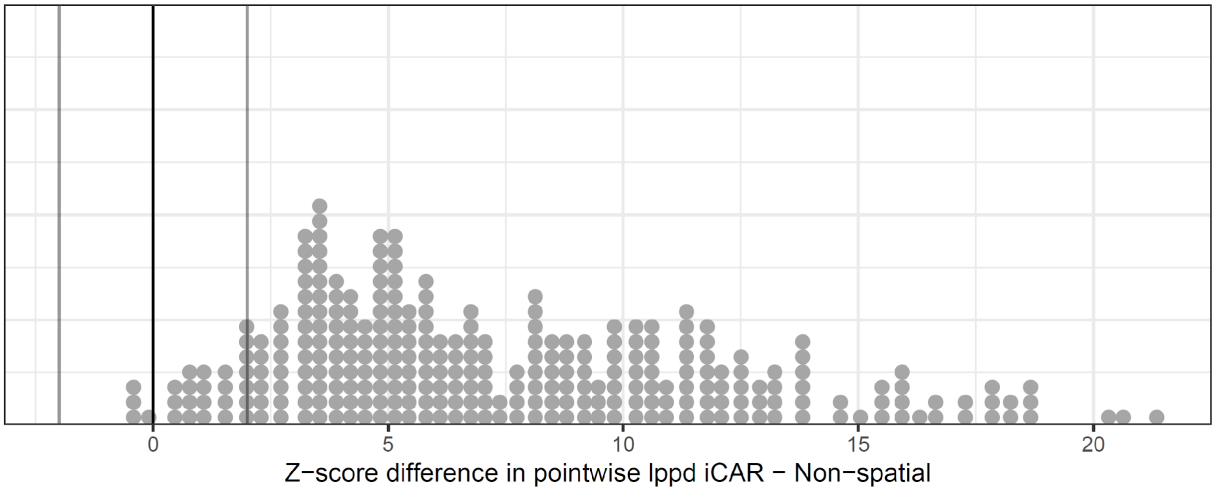

Figure 5. Leave Future Out (LFO) cross-validation results for all 287 species (including the 71 species results in Figure 4) from the North American Breeding Bird Survey (BBS) database, comparing the iCAR spatial model and the non-spatial model. The stacked dot-plots represent species-level summaries of the differences in log posterior predictive density (lppd) between the two models. Each point represents one species. Z-score values represent the difference between the lppd for the two models accounting for the variation across all counts, and the stacked dots form a histogram. Points that fall to the right of the black vertical line represent species for which the predictive accuracy of the spatial model is higher than that of the non-spatial model. The iCAR spatial model outperforms the non-spatial model for all but four species. For those four species, the predictive accuracy of the two models is very similar and does not approach -2, which would support a clear difference between the two models in favor of the non-spatial model.

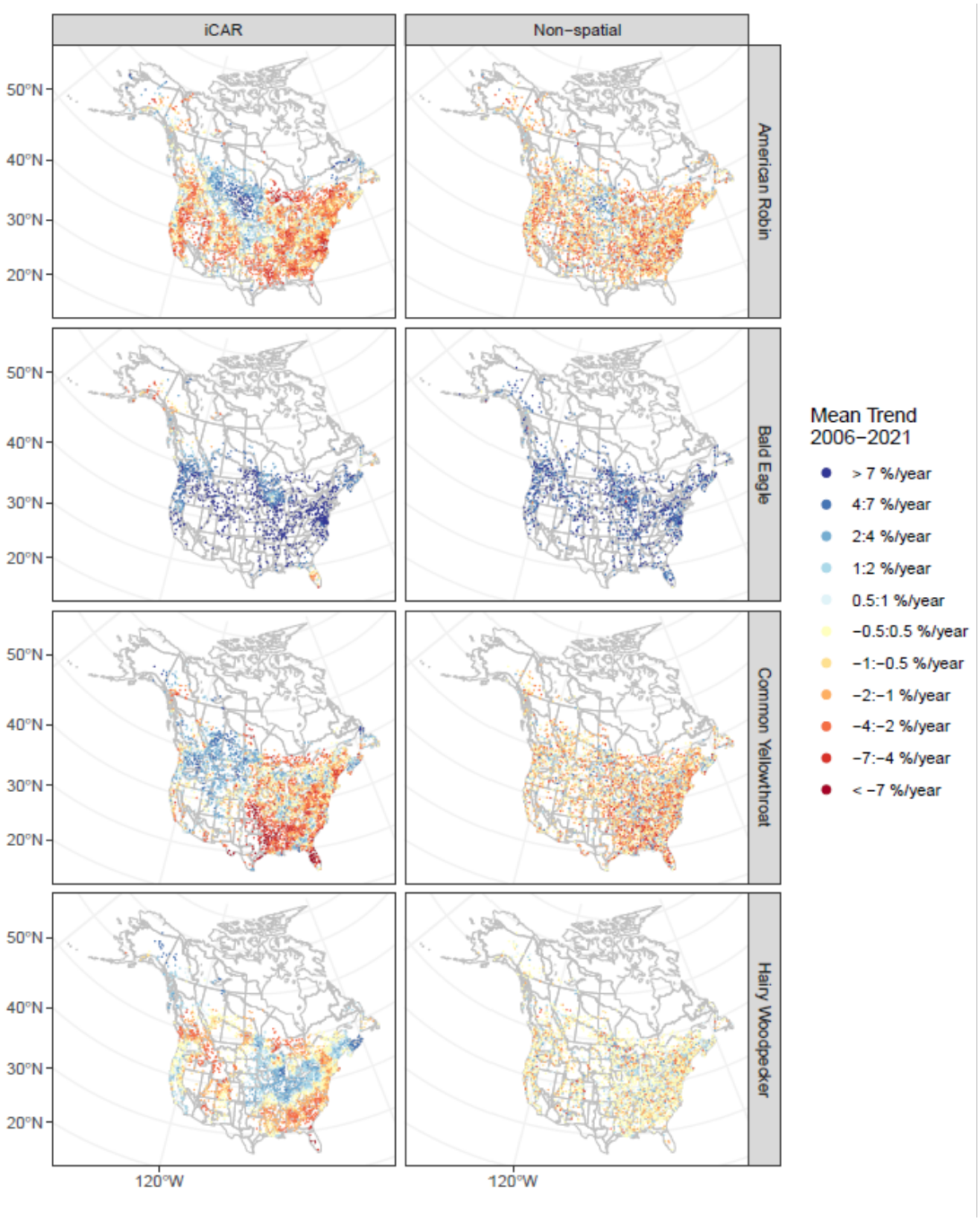

Figure 6. Examples of the spatial patterns in estimated route-level trends for four broad-range species from an iCAR spatial model (left column) compared to trends estimated from an otherwise identical, non-spatial version of the model (right column). All points are the same size in this plot because the mean abundances vary too much among species to display meaningful variation in this plot.

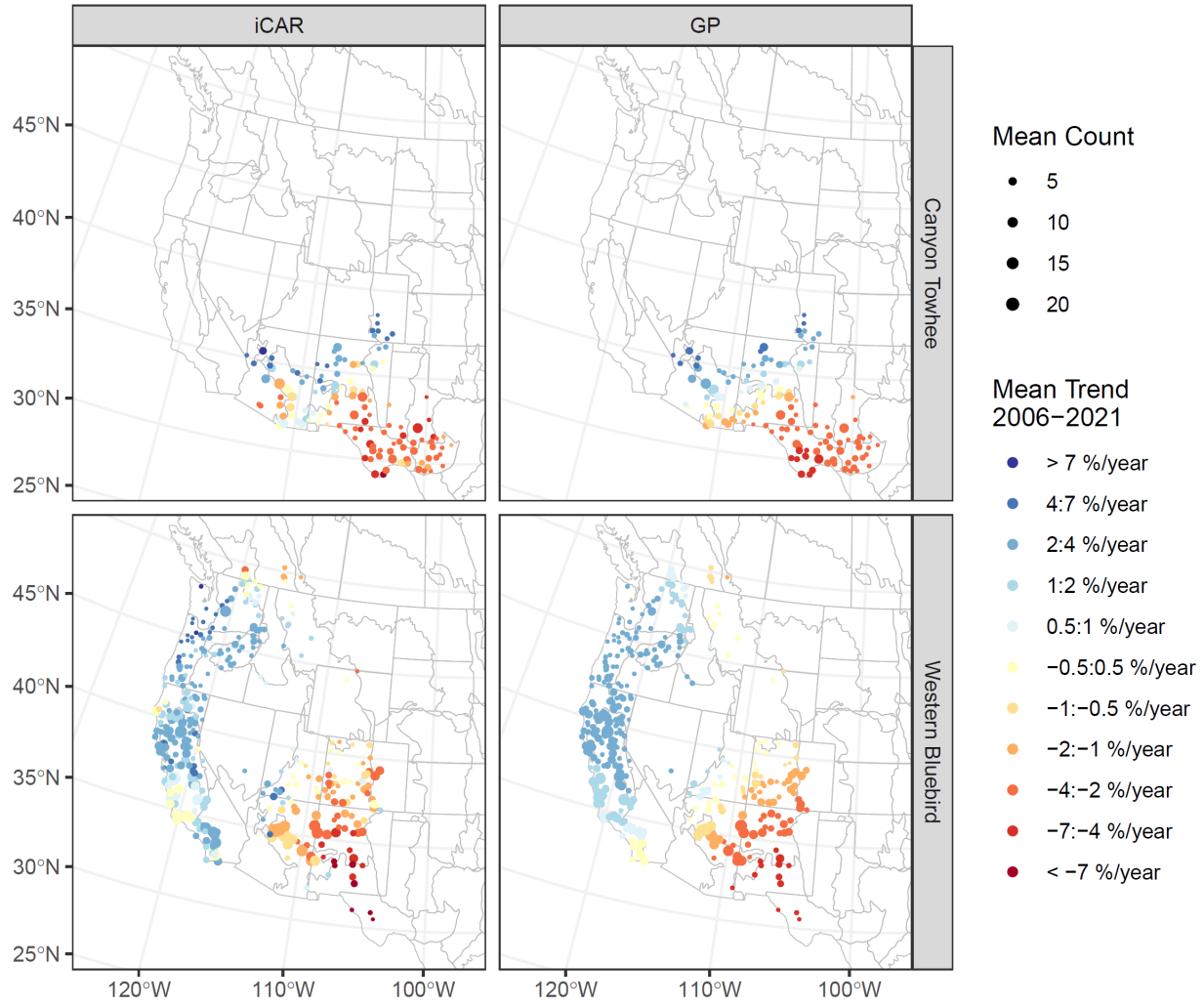

Figure 7. An example illustrates that the spatial patterns in estimated trends for iCAR and GP models are quite similar, even when one of the models strongly out-performs the other in a cross-validation analysis. For the Canyon Towhee (*Melozone fusca*), the GP model clearly out-performs the iCAR model in predictive accuracy (z-score comparison iCAR – GP = -4.3, Figure 4). For the Western Bluebird (*Sialia mexicana*), the iCAR model out-performs the GP model in predictive accuracy (z-score comparison iCAR – GP = 2.9, Figure 4). Despite the relatively strong difference in predictive accuracy, the spatial patterns are quite similar for both models.

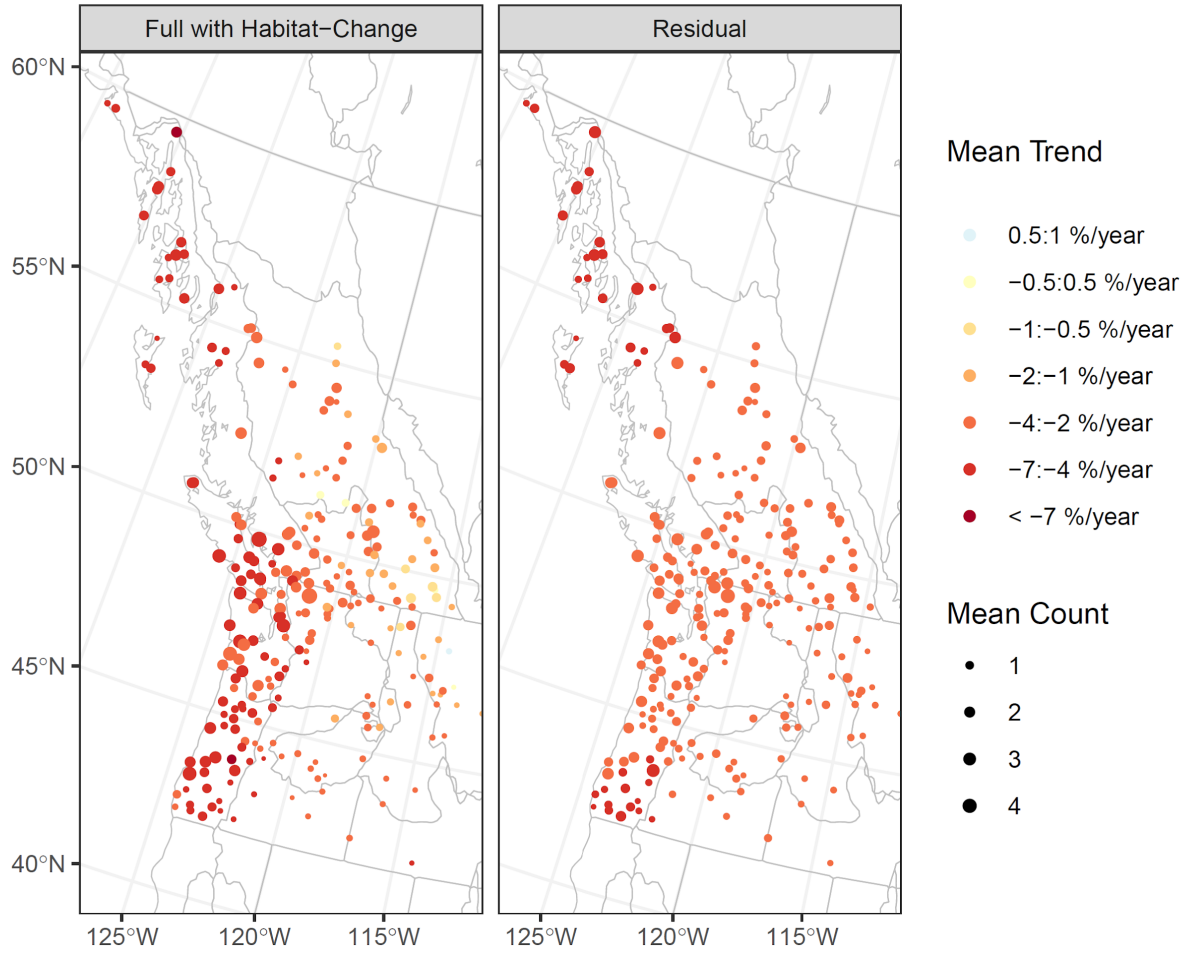

Figure 8. Map of route-level trend estimates for Rufous Hummingbird (*Selasphorus rufus*) from 2006-2021. The colors represent two sets of trends estimated from the model: "Full with Habitat-Change" represent trends that include the spatially explicit random effects and the effects of local habitat change (left panel) and the "Residual" represent only the residual spatially explicit estimate of trend, after removing the effects of habitat change (right panel). Habitat change appears to be driving most of the variation in trends within the core latitudes of the species range (45°-55° N). The faster rates of decline (darker red) in the western regions and the slower rates of decline (lighter yellow) in the east are evident in the left panel that includes the effects of habitat and are missing from the panel on the right.

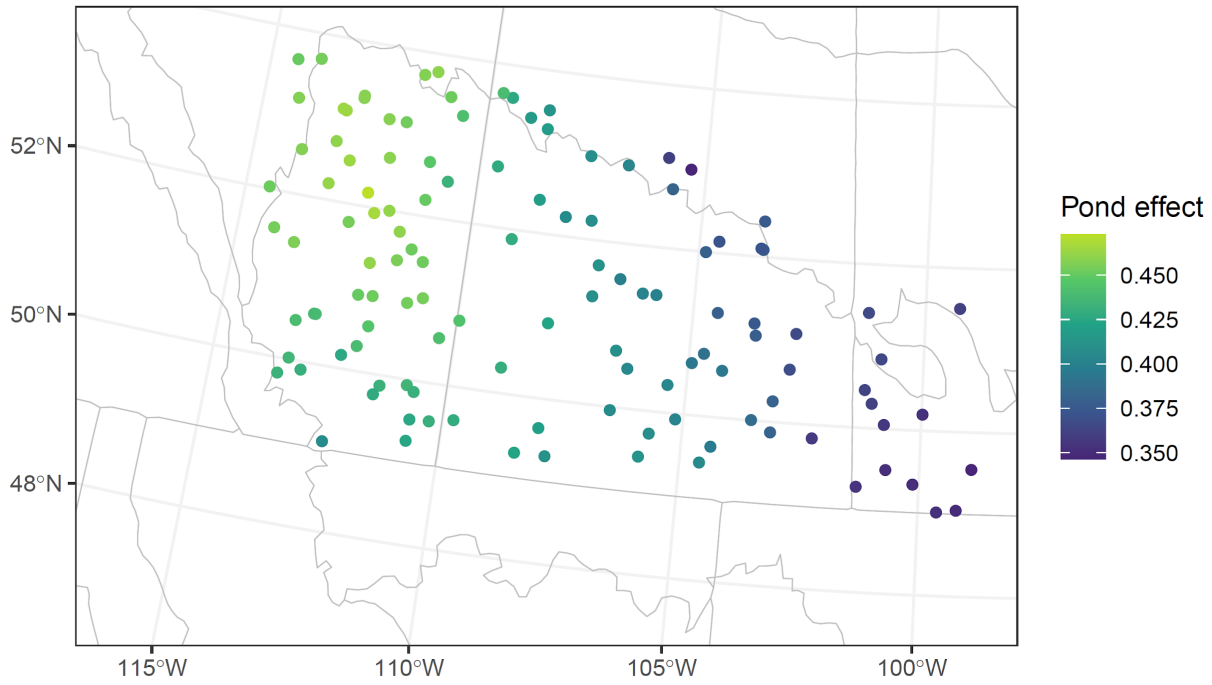

Figure 9. A map of the spatial variation in the effects of annual fluctuations in available habitat (ponds) on the expected counts of Horned Grebe on the North American Breeding Bird Survey (BBS) routes (1975-2017). The pond effect is estimated as a spatially varying coefficient using the iCAR structure among routes and was strongest in the western prairies. Pond effect represents the log-scale effect of annual variation in the number of ponds surrounding a BBS route in a given year on the annual expected count after adjusting for long-term trends, observer-effects, and the other parameters included in all of the models we used.