Supplemental_analyses

Model structure

The model is a relatively simple, hierarchical log-link negative binomial regression, like other models commonly applied to the BBS [Sauer and Link (2011)](Adam C. Smith et al. 2014), except applied at the route-level instead of pooling routes into geographic strata e.g., (Betts et al. 2022). In this model, each route has a separate slope and intercept and 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 a log-linear slope parameter.

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

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

We modeled the observed counts ($C_{r,j,t}$) of Rufous Hummingbirds on route-r, in year-t, by observer-j as 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), and 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).

We estimated the the first-year observer-effect η, as an independent parameter with weakly informative prior (below). All other parameters were estimated as hierarchical-effects, sharing information among routes or among observers.

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 remaining variation in relative abundance (α'''_r) and slope (β'''_r) that were not explained by habitat.

$$\alpha_r = \quad \alpha' + \alpha''_r + \alpha'''_r$$

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 the average of the annual habitat suitabilities in a buffer surrounding each route-path ($\alpha_r'' = \rho_{\alpha_r} * \mu_{habitatsuitability_r}$). We estimated the effect of the rate of change in habitat suitability over time on the route-level slope as a product of a route-specific coefficient (ρ_{β_r}) and an estimate of the average rate of change in habitat suitability on each route ($\delta_{habitatsuitability_r}$), calculated using a simple linear regression through the annual estimates of habitat suitability in a buffer surrounding each route-path ($\beta_r'' = \rho_{\beta_r} * \delta_{habitatsuitability}$). The habitat suitability predictors were standardized (both centered and scaled) to improve convergence. The route-specific coefficients for the effects of habitat suitablility on the intercept and slope were allowed to vary among routes, but were centered on a hyperparameter mean effects across routes $\rho_{\alpha_r} \sim Normal\left(P_\alpha, \sigma_{\rho_\alpha}\right)$ 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 suitablility on slope (P_β), represent a clear species-level estimate of the overall effects of habitat on abundance and trend, after adjusting for the species mean abundance and trend, as well as the residual spatially dependent variation in abundance and trend.

We estimated the spatially varying component of the intercepts and slopes using an intrinsic 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 neighbouring routes, with an estimated standard deviation that is proportional to the inverse of the number of neighbours for that route (Morris et al. 2019). Specifically, the component of the intercept that represents the residual spatially dependent relative abundance (α_r''') was drawn from a normal distribution centered on the mean of the intercepts for all neighbouring routes.

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

The spatially varying component of the slope (β_r) were estimated similarly as random route-level terms from a normal distribution centered on the mean of the slopes for all neighbouring routes using the same iCAR structure.

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

Spatial components

In other work, we have fit models with two different approaches to modeling the spatially explicit relationships among routes: a Gaussian process (GP) model that uses a matrix of Euclidian distances separating the start-locations of each BBS route, treating distance between routes as a continuous measure; and 2) an intrinsic Conditional Autoregressive (iCAR) structure that uses a sparse matrix of adjacencies between pairs of routes based on a tessellation of

the intervening space, treating spatial relationships as a series of discrete neighbours. Because the observations from a given BBS route are collected along a transect that is approximately 40km in length, it is not obvious which treatment of the spatial relationships better reflects reality. Spatial models are well developed for data collected at points and for data collected within discrete areas (Pebesma and Bivand 2023), but the BBS transects are neither points nor areas. Our prior work has shown that for route-level models of BBS data, the GP and iCAR approaches result in almost identical predictions, but the iCAR approach is much less computationally demanding (models fit in minutes rather than hours or days).

Both of these approaches are simplifications of the true spatial relationships among the BBS routes. The GP approach simplifies the spatial relationships by assuming each route represents a point in space and ensures that the covariance between pairs of routes declines with distance, but that measure of intervening distance only applies to the distances between the start points of the routes, not to the full transect. The iCAR approach simplifies the spatial structure by assuming each route represents a discrete area of space (a polygon surrounding the route), but the neighbouring routes may be separated by a wide range of distances depending on the spatial distribution and spatial density of those routes. For example, the GP could consider two distant routes as effectively independent, irresepective of how many routes were located in the intervening space. By contrast, the iCAR structure could consider these same two routes as having a very close connection if there were no intervening routes. In some cases, treating two relatively distant routes as close neighbours may be useful if their relative proximity provides useful information to inform the parameter estimates but may also introduce error into the estimate of spatial variance by considering the relatively distant neighbours as similar to relatively close neighbours.

The iCAR spatial structures require a discrete representation of spatial neighbourhood relationships (Ver Hoef et al. 2018), we used a Voronoi tessellation to generate these discrete neighbourhood relationships (Pebesma and Bivand 2023). iCAR models are often applied to contiguous areal stratifications, such as regular grids, census regions, or political jurisdictions, which have natural neighbourhood relationships defined by their adjacencies [Ver Hoef et al. (2018)](Meehan, Michel, and Rue 2019). To generate contiguous discrete spatial units without imposing a regular grid structure, we created a Voronoi tessellation centered on the start points of each BBS route ((Pebesma 2018)). We further limited the adjaceny matrix to the approximate boundaries of the species' range, by clipping the tessellated surface using the standard BBS analytical strata where the species occurs (province/states intersected with the Bird Conservation Regions; (Link and Sauer 2002) and a concave polygon surrounding the route start locations (Gombin, Vaidyanathan, and Agafonkin 2020). 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. Within the species' range boundaries, routes were considered neighbours if their Voronoi polygon shared a linear segment along a separating boundary. Our approach to generating these adjacency relationships may introduce variance, because some neighbouring 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 the relative spatial adjacencies

(each route is considered adjacent to its nearest neighbours) if not always the true Euclidean space.

Fitting the model

The model assumes that population trends at a given route can be reasonably described using a continuous log-linear slope parameter. This assumption of *trend* as an average rate of change across the entire time-series is probably most reasonable for relatively short periods of time (e.g., 10-20 years). For longer periods of time it is likely that natural populations will follow some more complex, non-linear path and therefore that the assumption of a continuous rate of change in the population over the full time-series is less likely (Adam C. Smith and Edwards 2020). Indeed, for Rufous Hummingbird populations in North America appear to show a break point in the population trend in approximately 2005. Between 1985 and 2005, the continental population was relatively stable, and after 2005 it declined steeply. To account for this nonlinearity in the population trajectory, we fit the model separately for the 20-year period from 1985-2005 and for the 15-year period from 2006-2021.

Before fitting the model, we prepared the BBS counts, the neighbourhood structures necessary to estimate the iCAR residual spatial component, and joined them to the habitat suitablity predictors. The full code and data necessary to replicate the data-preparation is available in the online supplement. In brief, we selected all routes on which the species had been observed during the time-period (1985-2005 or 2006-2021), and for which we had GIS route-path information that would allow us to estimate the route-specific annual habitat suitability values.

We fit the model using the probablistic programming language Stan (Stan Development Team 2022), accessed through the R-package `cmdstanr` (Gabry and Češnovar 2022). We used a warm-up of 2000 iterations, and `cmdstanr` default settings for other arguments, followed by a draw of 2000 samples from which we estimated the posterior distributions. All parameters in all models converged based on Rhat < 1.01 and bulk effective sample sizes > 500.

Results

1985-2005

During the first 20 years from 1985-2005, the species overall population was generally stable. The model estimated an overall change in the population of approximately -11% [-27-5.9]. Trends varied among routes and regions (Figure 1). The effect of habitat on mean relative abundance was strong and positive ($P_{\alpha_r} = 22$ [9.3-51]), if the residual trend was modeled without the spatial-dependency. However, when the spatial dependency was included, most of the variation in the relative abundance was allocated to the spatial component. There was a positive effect of change in the habitat suitability on trends, such that routes with habitat-loss had more negative population trends. However, this effect was somewhat uncertain $P_{\beta_r} = 5$ [-0.91-11], and 0 was included within the 95% posterior credible interval.

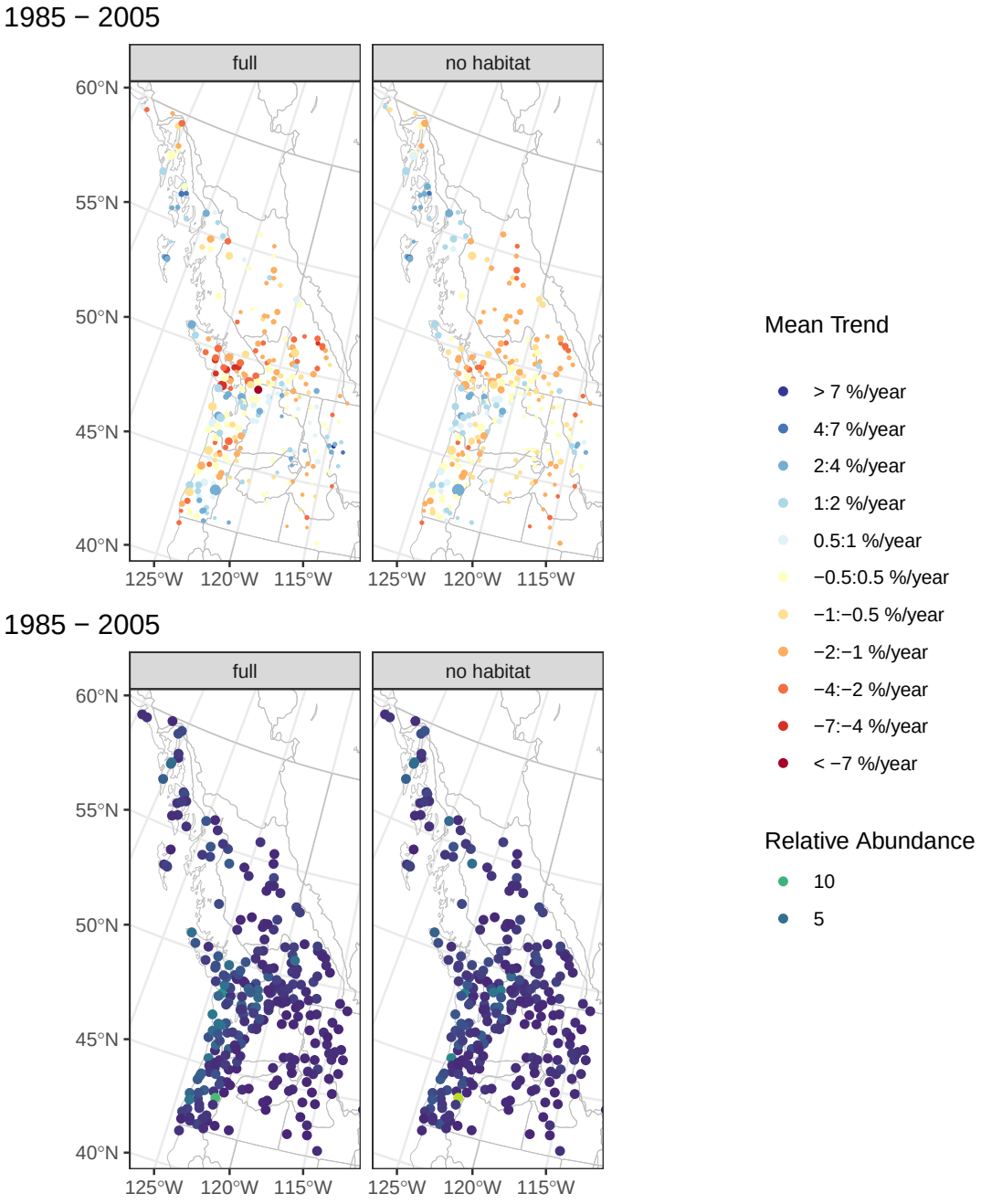

Figure 1: Map of the trends for Rufous Hummingbird from 1985-2005 The colours represent the trends in the uppper panel and the relative abundance in the lower panel. The left panel represents the full estimated trends and abundance on each route, including both the effect of habitat-suitability and the residual component not related to habitat. The right panel represents the trends and relative abundances after removing the effect of habitat-suitability. In the top-left panel, the greater declines in central-coastal regions are evident from the darker red points compared to the top-right panel. In the bottom-left panel, the higher abundance near the coast is evident from the lighter colours. The bottom-right panel shows much more even relative abundance across the species' range, showing that habitat suitability accounts for much of the variation in abundance

2006-2021

During the later 15 years from 2006-2021, the species overall population declined steeply. The model estimated an overall change in the population of approximately -43% [-52–33]. Trends were negative across the species' range, but most negative in the coastal regions where the species is also most abundant (Figure 2). The effect of habitat suitability on mean relative abundance was strong and positive ($P_{\alpha_r} = 20$ [8.7-46]), and this effect was robust, whether the residual abundance component was spatially autocorrellated or random. There was a clear positive effect of change in the habitat suitability on trends, such that routes with habitat-loss had more negative population trends $P_{\beta_r} = 2.6$ [0.28-4.9]. The greater loss of habitat in the coastal region accounts for most of the increased rates of decline in that region (Figure 2), the residual trend component alone (Figure 2, right panel) does not show the same coastal-decline pattern.

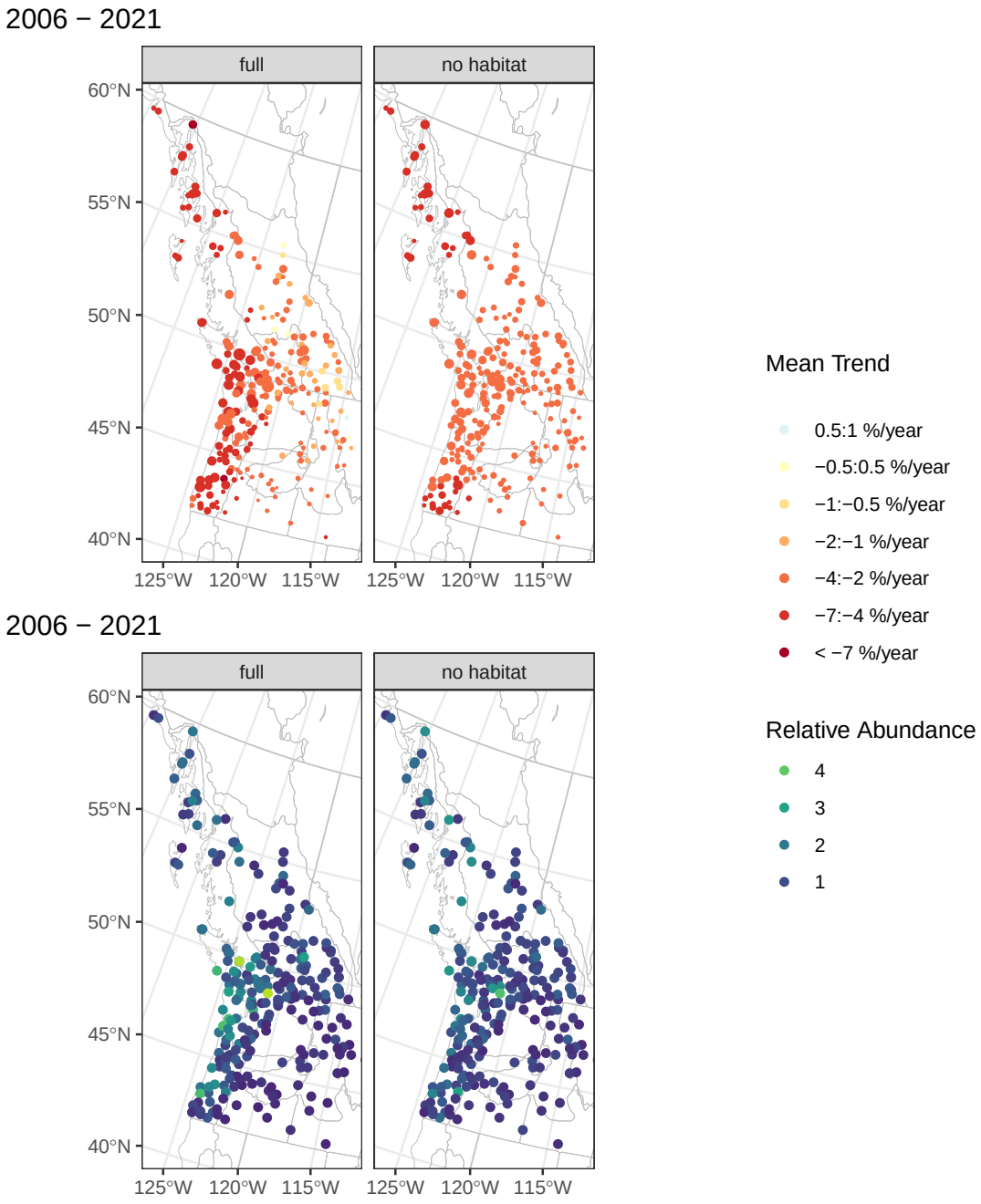

Figure 2: Map of the trends for Rufous Hummingbird from 2006-2021 The colours represent the trends in the uppper panel and the relative abundance in the lower panel. The left panel represents the full estimated trends and abundance on each route, including both the effect of habitat-suitability and the residual component not related to habitat. The right panel represents the trends and relative abundances after removing the effect of habitat-suitability. In the top-left panel, the greater declines in coastal regions are evident from the darker red points compared to the top-right panel. In the bottom-left panel, the higher abundance near the coast is evident from the lighter colours. The bottom-right panel shows much more even relative abundance across the species' range, showing that habitat suitability accounts for much of the variation in abundance

References

Betts, Matthew G., Zhiqiang Yang, Adam S. Hadley, Adam C. Smith, Josée S. Rousseau, Joseph M. Northrup, Joseph J. Nocera, Noel Gorelick, and Brian D. Gerber. 2022. "Forest Degradation Drives Widespread Avian Habitat and Population Declines." *Nature Ecology & Evolution* 6 (6): 709–19. https://doi.org/10.1038/s41559-022-01737-8.

Gabry, Jonah, and Rok Cešnovar. 2022. *Cmdstanr: R Interface to 'CmdStan'*. https://mc-stan.org/cmdstanr.

Gombin, Joël, Ramnath Vaidyanathan, and Vladimir Agafonkin. 2020. *Concaveman: A Very Fast 2D Concave Hull Algorithm*. https://CRAN.R-project.org/package=concaveman.

Link, William A., and John R. Sauer. 2002. "A Hierarchical Analysis of Population Change with Application to Cerulean Warblers." *Ecology* 83 (10): 2832–40. https://doi.org/10.1890/0012-9658(2002)083%5B2832:AHAOPC%5D2.0.CO;2.

Meehan, Timothy D., Nicole L. Michel, and Håvard Rue. 2019. "Spatial Modeling of Audubon Christmas Bird Counts Reveals Fine-Scale Patterns and Drivers of Relative Abundance Trends." *Ecosphere* 10 (4): e02707. https://doi.org/https://doi.org/10.1002/ecs2.2707.

Morris, Mitzi, Katherine Wheeler-Martin, Dan Simpson, Stephen J. Mooney, Andrew Gelman, and Charles DiMaggio. 2019. "Bayesian Hierarchical Spatial Models: Implementing the Besag York Mollié Model in Stan." *Spatial and Spatio-Temporal Epidemiology* 31 (November): 100301. https://doi.org/10.1016/j.sste.2019.100301.

Pebesma, Edzer. 2018. "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 10 (1): 439–46. https://journal.r-project.org/archive/2018/RJ-2018-009/index.html.

Pebesma, Edzer, and Roger Bivand. 2023. *Spatial Data Science: With Applications in R*. 1st ed. Boca Raton: Chapman; Hall/CRC. https://doi.org/10.1201/9780429459016.

Sauer, John R., and William A. Link. 2011. "Analysis of the North American Breeding Bird Survey Using Hierarchical ModelsAnálisis de Los Censos de Aves Reproductivas de Norte América Mediante Modelos Jerárquicos." *The Auk* 128 (1): 87–98. https://doi.org/10.1525/auk.2010.09220.

Smith, Adam C, and Brandon 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*, no. duaa065 (December). https://doi.org/10.1093/ornithapp/duaa065.

Smith, Adam C., Marie-Anne R. Hudson, Constance Downes, and Charles 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 (2): 119–34. https://doi.org/10.22621/cfn.v128i2.1565.

Stan Development Team, 2.29. 2022. *Stan Modeling Language Users Guide and Reference Manual*. https://mc-stan.org.

Ver Hoef, Jay M., Erin E. Peterson, Mevin B. Hooten, Ephraim M. Hanks, and Marie-Josèe Fortin. 2018. "Spatial Autoregressive Models for Statistical Inference from Ecological Data." *Ecological Monographs* 88 (1): 36–59. https://doi.org/10.1002/ecm.1283.