Methods

## Strata-free estimates

Partners in Flight (PIF) estimate bird populations within geo-political regions based on summaries of counts on all North American Breeding Bird Survey (BBS) routes within each region. Summarising observed counts on BBS routes across these regions assumes that the BBS routes are a representative sample of the region. This assumption is reasonable for many species and many regions, but not all. With the development of new spatial information on the paths of BBS routes and range-wide relative abundance surfaces for many species from eBird ((Fink et al., n.d.)), this assumption is no longer necessary. Instead of summarising BBS observations within broad scale geo-political strata (e.g., Bird Conservation Regions by political jurisdictions), we can use the relationship between observed counts on each route and the relative abundance surface within in the area immediately surrounding each BBS route to account for the sampling bias in the BBS roadside surveys. It is relatively trivial to assume that the BBS observations are a representative sample of the birds, habitats, and land area within a short buffer (e.g.,400 m radius) of the route paths.

## eBird relative abundance

We calculated the area weighted mean of the eBird relative abundance cell values within the buffer zone of route-j () to use as a predictor on the observed BBS counts at that route.

We used the mean weekly relative abundance values for the weeks that overlap the BBS field season. We explored alternative summaries of the eBird relative abundance surfaces including the seasonal medians, and maximum relative abundance within the BBS field season (Stillman et al. 2023). We chose to use means of weeks within the BBS field season because these showed the most consistent and linear relationships to the observed BBS counts when compared across many species.

We included all routes on which the species was observed during BBS surveys in the last 10 years of observations (2013 - 2023, excluding 2020 when the BBS was cancelled) and that overlapped positive values of the eBird relative abundance surface. We also excluded the BBS routes where the mean relative abundance values fell in the lowest 10% of the relative abundance values because for some species and routes these very low value predictors generated some extreme predictions and convergence failures.

The relationship between the counts on BBS routes and the eBird relative abundances within these route path buffers along with the PIF population estimates approach to scaling BBS counts to total population size, provide the logical stepping stones to calibrate the eBird relative abundance values to represent a density of birds.

## Proportional relative abundance

Both the BBS counts and the values of the eBird weekly abundance surfaces are designed to represent an index of abundance; i.e., relative abundance, abundance at one location, relative to other locations. For the BBS, the methodological controls in the survey protocol ensure that the observed counts represent relative abundance of the birds in the areas adjacent to the road (), and the analytical adjustments in the eBird analyses generate estimates of weekly or seasonal relative abundance across a species’ range (). These indices of abundance are by definition, proportional to true abundance (), and so each index has some associated proportionality constant ( or ).

The PIF population estimation equation (Stanton et al. 2019) represents one estimate of the BBS proportionality constant , i.e., a factor that scales the observed counts by BBS observers to true abundance. Combining the equations above, it becomes clear that because each index is proportional to true abundance, there is also a proportional relationship between the two indices and their proportionality constants.

Further re-arranging suggests a relationship for predicting the BBS counts based on the ebird abundance indices and the quotient of their two proportionality constants.

This quotient of two constants, is itself a constant, which here we refer to as ().

So finally, the right side of this equation gives the basis for estimating using the relationship between the ebird relative abundance surface and the observed BBS counts. Then once we have an estimate of , we can then combine it with to estimate the true abundance of birds across the landscape. This term is just a re-expression of the PIF population estimates equation that translates counts on BBS routes to true abundance. Therefore, multiplying these two terms together generates a calibration factor to scale from the eBird estimated relative abundance to true total abundance in a grid cell.

## Estimating the relationship between BBS counts and eBird relative abundance

We fit a Bayesian hierarchical model that estimates the term in the above equation. We modeled the expected count during each BBS as a function of the log-transformed mean eBird relative abundance values within a 400 m buffer surrounding the BBS route-path. The model also accounts for: 1) the variation among BBS routes and observers, 2) the variation among years; 3) the seasonal variation during the period of acceptable BBS survey conditions and how that seasonal pattern varies among BBS strata.

We log transformed the area weighted mean relative abundance values at each BBS route and used them as an offset in a log-link model that predicts each observed BBS count () on route-j, in year-t. Treating these as an offset respects the multiplicative relationship derived above.

This assumes a linear relationship with a slope of 1 between the log of the eBird relative abundance surface and the log of the observed BBS counts. This assumption is explicit if each relative abundance value is proportional to true abundance, but also necessary so that the final population estimates respect the spatial distribution in the eBird relative abundance surface. Of course, either or both set of estimates may include errors and therefore fail to meet the assumption of relative abundance in important ways beyond what we have accounted for here. Exploring those possible failures is beyond the scope of this work, but it will include assessing the slope for each species and for different summaries of the eBird relative abundance surface (e.g., the mean, median, or maximum seasonal abundance, and the weeks included in the season to match the BBS observations).

We modeled the observed BBS counts in a similar way to many of the common status and trend models for the BBS (Smith et al. 2024a) as realizations of a negative binomial distribution with mean and inverse dispersion parameter .

In this model formulation, the intercept () represents a scaling factor between the observed BBS counts on route-r, and the eBird relative abundance for route-r, after accounting for the variation due to the year-t (), and day of the BBS season-d in stratum-s (). We used a hierarchical structure to model the variation in the scaling factor among routes , while accounting for the repeated observations on each route by a given observer . The hyperparameter represents the mean log-scale eBird-BBS scaling factor across all routes. We used a t-distribution to allow the variation among routes to have heavier tails than the normal and to provide a robust estimator that is less sensitive to extreme outliers (e.g., routes with unusually high or low counts).

The year-term was fit as a random-walk time-series, centered on the year of the relative abundance surface (i.e., for the eBird status and trends version 2023).

This specific centering ensures that the mean calibration () matches the year of the eBird relative abundance surface, and that the model can adjust for non-linear population trends of each species while using information from many years of BBS observations. We then allowed this random-walk time-series to vary in space (for all years except 2022, when it remained fixed at 0), by grouping BBS routes by strata-s, defined by the spatial intersections of Province, State, and Territory boundaries with the North American Bird Conservation Regions ((Smith et al. 2024b)). A spatially explicit hierarchical structure was used to estimate strata-specific using the same intrinsic conditional autoregressive (iCAR) parameterisation used to model strata-level variation in the spatial variant of the first-difference model of the r-package *bbsBayes2* ((Edwards et al. 2023)), following Smith et al. (2024a). For example, in the year 2021, the difference from the centered year 2022 in stratum-s was,

The seasonal variation was modeled as a hierarchical GAM smooth term that adjusts for the variation in the day each BBS survey was conducted. We used the hierarchical GAM to smooth the variation in counts across the BBS season while allowing the shape of that seasonal smooth to vary across the species’ range and still shrink towards the mean seasonal pattern of neighbouring strata when data were relatively sparse ([Pedersen et al. (2019)](Smith and Edwards 2021)(Smith et al. 2024b)). To model that spatial variation in the seasonal component, the hierarchical GAM smooths were fit using the same iCAR parameterisation used to model annual smooth terms in spatial variants of the GAM and GAMYE models in the r-package *bbsBayes2* ((Edwards et al. 2023)), following Smith et al. (2024).

To account for the asymmetries of the log-scale retransformation (i.e., ) and some extreme tails in the distribution of the route-level terms, we used a modification of a retransformation approach that is unbiased if the distribution of the estimated route-effects was not log-normal, following (Duan 1983). This approach involves retransforming the route-level calibration values and then averaging across the retransformed values. We modified this approach by first excluded the 5% of routes in the extreme tails of the distribution of route-level values (we retained routes with estimates in the middle 95% of the distribution). We chose to use this 95% trimmed mean calculation so the mean value would be less sensitive to some extreme tails of these route-level distributions. These extreme values often reflected routes with few years of observations on the route and/or by the relevant observer and a mismatch between the estimated relative abundance surrounding the route. We then retransformed the calibration value for each of the remaining R-routes and averaged across these routes to estimate the calibration factor ().

We can then combine the posterior draws of the scaling factor , with Monte Carlo processes similar to those used in Stanton et al. (2019) to propagate the uncertainty of the scaling factor with the estimates of adjustment factors that match BBS observation conditions. The re-scaling to represent density accounts for: 1) the estimated availability of birds for detection (and its uncertainty) during a 3-minute BBS point count conducted on an average BBS survey day; 2) the estimated detectability distance for roadside counts and the uncertainty in its estimate; and 3) a correction factor used in previous PIF population estimates to account for the bias in the detectability between male and female birds of each species (Stanton et al. 2019).

In this calibration, the area in square kilometers surveyed by a 50-stop BBS route, based on the estimated Effective Detection Radius (EDR) for a roadside count from the NA-POPS project (Edwards et al. 2023). We incorporated the uncertainty of the EDR estimate into each sample of the posterior distribution, by making a random draw from a normal distribution with the mean and standard deviation based on the estimated mean and standard error of the mean from the NA-POPS analyses. We selected the EDR estimates from the better supported of either the roadside model or the null model in the most recent estimates from NA-POPS.

The time of day correction in this model was based on the estimated availability from the best supported removal model in the NA-POPS analyses. Availability represents the probability that a bird is available for detection during a 3-minute BBS count . The inverse of that probability represents an estimate of the multiplicative correction factor necessary to account for the proportion of birds that were present, but not available for detection during that time . We drew independent random samples from a normal distribution centered on for each posterior sample, and constrained the normal distribution to values between 0.001 and 0.999 to avoid impossible probabilities.

Finally, the pair correction factor was incorporated based on the correction for the ratio of the observed proportion of birds that are female to the expected 0.5 (assuming an even sex ratio in the population). These values can be calculated based as either , or following the values used in previous analyses (Stanton et al.. 2019, truncated normal distribution with a mean equal to the species’ assigned pair correction value, standard deviation equal to 0.13, and truncated at 1.0 and 2.0).

This calibration factor mutliplied by an eBird relative abundance value in a given cell, represents an estimated number of birds per .

For each regional estimated population, we calculated the summed relative abundance values within the region, then multiplied that summed relative abundance by each posterior draw of the calibration factor, and acountted for the area of each cell (e.g., for the 3km cells in the highest resolution seasonal abundance surfaces 9 ). Variation across all posterior draws provides an estimate of the uncertainty of the population size estimate.

## Comparisons with the existing estimation approach

We compared the estimated populations from this model to those derived from the existing modeling process (following Stanton et al. 2019), for a selection of species to highlight some of the key benefits and challenges of this approach. We chose a selection of species classified as resident and migratory for the purposes of their estimated eBird relative abundance surfaces.

We compared the estimates on the finest spatial resolution available for the existing modeling process: the analytical strata represented by the intersection of political jurisdictions (provinces, states, and territories) with the Bird Conservation Regions (BCRs). At that fine scale, we predicted that the differences in population estimates would be largely based on two factors: the relative-abundance sampling bias of the BBS routes relative to the stratum, and the interaction between the relative abundance in each stratum and the temporal trend in abundance over the 10-year span of BBS observations included in the calculations.

We estimated the sampling bias of the BBS by calculating a ratio of the mean relative abundance values that overlap the BBS route-paths over the mean relative abundance values across the entire stratum. Strata with positive bias by the BBS (where the BBS routes sample high abundance areas relative to the stratum), our model should generate lower estimates of population size than the existing analytical process. Strata with negative bias by the BBS (where the BBS routes sample relatively low abundance areas), we expect higher estimates of population size.

We also estimated the ratio of the mean predicted abundance in each stratum and each year, with the predicted abundance in the year of the eBird relative abundance surface. The existing esimation process estimates a ten-year average population (e.g., the average over the 2013-2021 period). This ratio of the predicted counts in 2022 over the mean of the predicted counts in all years, reflects the average expected change in the population estimates due to the local population trend. For example, in a stratum where the time-series of observations from the BBS suggest a decline over the period 2013-2023, estimated populations from the new model will be lower to adjust for that ongoing trend.

Edwards, Brandon P. M., Adam C. Smith, Teegan D. S. Docherty, Marcel A. Gahbauer, Caitlyn R. Gillespie, Alexis R. Grinde, Taylor Harmer, et al. 2023. “Point Count Offsets for Estimating Population Sizes of North American Landbirds.” *Ibis* 165 (2): 482–503. <https://doi.org/10.1111/ibi.13169>.

Fink, D, T Auer, A Johnston, M Strimas-Mackey, S Ligocki, O Robinson, W Hochachka, et al. n.d. “eBird Status and Trends.” <https://doi.org/10.2173/ebirdst.2022>.

Pedersen, Eric J., David L. Miller, Gavin L. Simpson, and Noam Ross. 2019. “Hierarchical Generalized Additive Models in Ecology: An Introduction with Mgcv.” *PeerJ* 7 (May): e6876. <https://doi.org/10.7717/peerj.6876>.

Smith, Adam C, Allison D. Binley, Lindsay Daly, Brandon P M Edwards, Danielle Ethier, Barbara Frei, David Iles, Timothy D Meehan, Nicole L Michel, and Paul A Smith. 2024a. “Spatially Explicit Bayesian Hierarchical Models Improve Estimates of Avian Population Status and Trends.” *Ornithological Applications* 126 (1): duad056. <https://doi.org/10.1093/ornithapp/duad056>.

———. 2024b. “Spatially Explicit Bayesian Hierarchical Models Improve Estimates of Avian Population Status and Trends.” *Ornithological Applications* 126 (1): duad056. <https://doi.org/10.1093/ornithapp/duad056>.

Smith, Adam C, and Brandon P M Edwards. 2021. “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* 123 (1): duaa065. <https://doi.org/10.1093/ornithapp/duaa065>.

Stanton, Jessica, Peter Blancher, Kenneth Rosenberg, Arvind Panjabi, and Wayne Thogmartin. 2019. “Estimating Uncertainty of North American Landbird Population Sizes.” *Avian Conservation and Ecology* 14 (1). <https://doi.org/10.5751/ACE-01331-140104>.

Stillman, Andrew N., Paige E. Howell, Guthrie S. Zimmerman, Emily R. Bjerre, Brian A. Millsap, Orin J. Robinson, Daniel Fink, Erica F. Stuber, and Viviana Ruiz-Gutierrez. 2023. “Leveraging the Strengths of Citizen Science and Structured Surveys to Achieve Scalable Inference on Population Size.” *Journal of Applied Ecology* 60 (11): 2389–99. <https://doi.org/10.1111/1365-2664.14512>.