

Stanton, J. C., P. Blancher, K. V. Rosenberg, A. O. Panjabi, and W. E. Thogmartin. 2019. Estimating uncertainty of North American landbird population sizes. *Avian Conservation and Ecology* 14(1):4. <https://doi.org/10.5751/ACE-01331-140104>  
Copyright © 2019 by the author(s). Published here under license by the Resilience Alliance.

Research Paper

## Estimating uncertainty of North American landbird population sizes

Jessica C. Stanton<sup>1</sup>, Peter Blancher<sup>2</sup>, Kenneth V. Rosenberg<sup>3,4</sup>, Arvind O. Panjabi<sup>5</sup> and Wayne E. Thogmartin<sup>1</sup>

<sup>1</sup>U.S. Geological Survey, Upper Midwest Environmental Sciences Center, La Crosse, WI, <sup>2</sup>Environment and Climate Change Canada, Ottawa, ON (emeritus), <sup>3</sup>Cornell Lab of Ornithology, Ithaca, NY, <sup>4</sup>American Bird Conservancy, Washington, D.C., <sup>5</sup>Bird Conservancy of the Rockies, Fort Collins, CO

**ABSTRACT.** An important metric for many aspects of species conservation planning and risk assessment is an estimate of total population size. For landbirds breeding in North America, Partners in Flight (PIF) generates global, continental, and regional population size estimates. These estimates are an important component of the PIF species assessment process, but have also been used by others for a range of applications. The PIF population size estimates are primarily calculated using a formula designed to extrapolate bird counts recorded by the North American Breeding Bird Survey (BBS) to regional population estimates. The extrapolation formula includes multiple assumptions and sources of uncertainty, but there were previously no attempts to quantify this uncertainty in the published population size estimates aside from a categorical data quality score. Using a Monte Carlo approach, we propagated the main sources of uncertainty arising from individual components of the model through to the final estimation of landbird population sizes. This approach results in distributions of population size estimates rather than point estimates. We found the width of uncertainty of population size estimates to be generally narrower than the order-of-magnitude distances between the population size score categories PIF uses in the species assessment process, suggesting confidence in the categorical ranking used by PIF. Our approach provides a means to identify species whose uncertainty bounds span more than one categorical rank, which was not previously possible with the data quality scores. Although there is still room for additional improvements to the estimation of avian population sizes and uncertainty, particularly with respect to replacing categorical model components with empirical estimates, our estimates of population size distributions have broader utility to a range of conservation planning and risk assessment activities relying on avian population size estimates.

## Estimation de l'incertitude quant à la taille des populations d'oiseaux terrestres d'Amérique du Nord

**RÉSUMÉ.** Une mesure importante pour de nombreux aspects de la planification de la conservation des espèces et de l'évaluation des risques est une estimation de la taille totale de la population. Pour les oiseaux terrestres se reproduisant en Amérique du Nord, Partners in Flight (PIF) génère des estimations de la taille des populations mondiales, continentales et régionales. Ces estimations constituent un élément important du processus d'évaluation des espèces du PIF, mais elles sont également utilisées par d'autres pour diverses applications. Les estimations de la taille de la population calculées par PIF le sont à l'aide d'une formule conçue pour extrapoler les dénombrements d'oiseaux enregistré par le Relevé des Oiseaux Nicheurs de l'Amérique du Nord (BBS) aux estimations régionales de population. La formule d'extrapolation comprend de multiples hypothèses et sources d'incertitude, mais il n'y avait auparavant aucune tentative de quantifier cette incertitude dans les estimations publiées de la taille des populations, à l'exception d'un score catégorique de qualité des données. En utilisant une approche de Monte Carlo, nous avons propagé les principales sources d'incertitude découlant des composantes individuelles du modèle jusqu'à l'estimation finale de la taille des populations d'oiseaux terrestres. Cette approche donne lieu à des distributions d'estimations de la taille des populations plutôt qu'à des estimations ponctuelles. Nous avons constaté que l'ampleur de l'incertitude des estimations de la taille des populations était généralement plus étroite que les distances de l'ordre de grandeur entre les catégories de scores de taille des populations utilisées par PIF dans le processus d'évaluation des espèces, ce qui suggère que le classement catégorique utilisé par PIF est fiable. Notre approche offre un moyen d'identifier les espèces dont les limites d'incertitude couvrent plus d'un rang catégorique, ce qui n'était pas possible auparavant avec les scores de qualité des données. Bien qu'il reste encore des améliorations à apporter à l'estimation de la taille et de l'incertitude des populations aviaires, particulièrement en ce qui concerne le remplacement des composantes qualitatives des modèles par des estimations empiriques, nos estimations de la répartition de la taille des populations sont utiles à une gamme d'activités plus large de planification de la conservation et d'évaluation des risques qui reposent sur des estimations des populations aviaires.

**Key Words:** *Breeding Bird Survey; Partners in Flight; population size; species conservation assessment; uncertainty*

## INTRODUCTION

The total population size of animal species is one of the fundamental determining factors in extinction risk (McKinney 1997, Purvis et al. 2000) and, thus, is an important component of species risk assessment programs (Carter et al. 2000, Faber-Langendoen et al. 2012, IUCN 2012). As a general rule, the smaller the population, the greater the risk of extirpation through demographic stochasticity, environmental stochasticity, and random catastrophic events (Lande 1993). In addition, estimates of population size are important for assessing the impact of different sources of anthropogenic mortality in North America (Johnson et al. 2012, Longcore and Smith 2013, Longcore et al. 2013, Loss et al. 2013, 2014, Erickson et al. 2014). Given the importance of population size in risk assessment, knowing the uncertainty of those estimates is equally important. Uncertainty, or the amount of variance inherent in the estimation process, can arise from several sources (Regan et al. 2003), and failing to account for uncertainty when conducting risk assessment can lead to misclassification (Wilson et al. 2011).

As part of the 2004 North American Landbird Conservation Plan (Rich et al. 2004), Partners in Flight (PIF), a consortium of scientific and other organizations focused on avian conservation in North America (Finch and Stangel 1993), introduced an approach for estimating population sizes (Rosenberg and Blancher 2005). This approach resulted in the first published global population size estimates for North American landbirds (Rich et al. 2004). These estimates were used to refine an important component of the PIF species assessment designed to categorize and prioritize species by conservation need across the continent (Hunter et al. 1993, Carter et al. 2000). Recognizing that meeting conservation objectives at continental scales requires regional planning and action, PIF also published a database of population size estimates downscaled to several geographic scales, such as Bird Conservation Regions (BCRs), states in the United States, and provinces/territories in Canadian (Blancher et al. 2007, 2013, PIF 2013).

In the PIF species assessment approach, global population size estimates are expressed as categorical population size (PS) scores (Panjabi et al. 2005, 2012, Rosenberg et al. 2017), with scoring categories separated by orders of magnitude (Table 1). Therefore, for PIF assessment purposes, it is assumed a high degree of precision in population size estimates is not essential to accurately assign a PS score. The current PIF population estimates (Blancher et al. 2007) include a data quality rating to express the relative confidence in the quality or quantity of data on which the components of population estimate are based. However, these data quality ratings do not include a measure of uncertainty around the population size estimates, which would indicate whether an estimate spans more than one PS category.

PIF population size estimates are frequently used for purposes beyond the PIF species assessment. They have aided in establishing continental and regional species prioritization and planning (Potter et al. 2007, Thogmartin et al. 2014, Rosenberg et al. 2016); for those working in avian conservation, transforming abstract concepts such as “index of abundance” into more generally understandable “population size” estimates were found to aid communication with policy makers and the public (Briggs 2006, Bickford et al. 2012). PIF population estimates are also

increasingly being used to assess the impacts of different sources of anthropogenic avian mortality (Johnson et al. 2012, Longcore et al. 2013, Erickson et al. 2014, Loss et al. 2014). Because of the importance of population size estimation for assessing species risk and mortality impacts, PIF has committed to improve and revise the approach and, to the extent possible, address critiques and suggestions that have been raised (Thogmartin et al. 2006, Thogmartin 2010). A key suggestion for improving the usefulness of the population size estimates was to incorporate an explicit measure of uncertainty (Thogmartin et al. 2006). The handbook to PIF 2013 identified the reporting of uncertainty bounds in the population estimates as a “next step” (Blancher et al. 2013).

**Table 1.** Categories of population size (PS) scores used in the Partners in Flight (PIF) species assessments of North American landbirds. Smaller population sizes are associated with greater vulnerability and are assigned larger PS scores (Panjabi et al. 2005, 2012, Rosenberg et al. 2016).

PS Score	Global breeding population size
1	≥ 50,000,000
2	≥ 5,000,000 & < 50,000,000
3	≥ 500,000 & < 5,000,000
4	≥ 50,000 & < 500,000
5	< 50,000

We present a modification to the PIF approach to estimating population sizes for landbirds breeding in the continental United States (U.S.) and Canada (hereafter “North America”). This modification allows for the range of uncertainty underlying the raw data and the other model parameters to be numerically presented in the final estimate. Specifically, we use a Monte Carlo simulation to propagate uncertainty arising from the individual components of the estimation process through to the final estimation of total population size. The result is a distribution of population size estimates for each species in each geographic region, which can be subsequently described by standard descriptive statistics (mean, median, quantiles) rather than as a single point estimate. In addition to incorporating uncertainty, we further refine the PIF population estimation approach to include an update to the time-of-day adjustment factor (Blancher et al. 2013) and the consolidation of a large portion of the process into coded computational scripts written for the R platform (R Core Team 2017), improving reproducibility and transparency.

## METHODS

### The Partners in Flight approach

PIF currently presents population sizes as single point estimates with corresponding categorical data quality ratings. Here we briefly describe the PIF approach to population size estimation (for additional details see Rosenberg and Blancher 2005, Blancher et al. 2007, 2013). The PIF approach relies on a series of logical assumptions for extrapolating survey-level counts to total abundance within defined geographic regions. For landbird species breeding primarily in North America, the primary data source is the North American Breeding Bird Survey (BBS; Pardieck et al. 2016, Rosenberg et al. 2017). The approach is based on a series of adjustment factors applied to mean bird counts

observed along BBS routes. Each adjustment factor is meant to adjust the mean count for the approximate proportion of birds that are missed relative to those that are observed, as well as the amount of area (on average) that is effectively surveyed along a BBS route.

The PIF approach to estimating the population size within each defined geographic area is based on the following equation:

$$\text{Population Size} = \frac{\sum_{j=1}^m \sum_{i=t}^T Y_{ij}/n}{m} * \frac{a}{25.1} * \left(\frac{400}{C_D}\right)^2 * C_P * C_T \quad (1)$$

where  $Y$  is the number of individuals of a species reported along a BBS route  $j$  in year  $i$ ,  $n$  is the number of times a route was surveyed under acceptable conditions during the sequence of years from  $i = t$  to  $T$ , which covers a time period of 10 consecutive years (typically the most recent decade of available BBS data, 2006 to 2015 data were used in the estimates presented below),  $m$  is the number of routes within the geographic area,  $a$  is the land area (km<sup>2</sup>) within the geographic area. The divisor on  $a$ , 25.1, is the presumed area in km<sup>2</sup> that is sampled along a BBS route (50 sample locations each with a 400 m sampling radius). The three constants:  $C_D$ ,  $C_P$ , and  $C_T$ , are species-specific adjustment parameters.  $C_D$  and  $C_P$  are both categorical and derived through a process of literature review, data review, and expert opinion (Blancher et al. 2013).  $C_D$ , the detection-distance adjustment, is used to modify the presumed BBS sampling radius, 400 m, for each species based on habitat, behavior, and song characteristics.  $C_P$ , the pair adjustment, is used to modify the estimate based on the assumption that detection of some species may be biased toward only one member of a breeding pair.  $C_T$ , the time-of-day adjustment, is used to modify the estimate due to variable detection probability throughout the survey period.  $C_T$  is derived through an analysis of stop-level BBS data (available from 1997 to the present) and estimates the ratio of the average count at the peak-detection stop to the mean count across all stops based on polynomial smoothing models (Blancher et al. 2007, 2013).

The smallest geographic units of analysis (here referred to as “physio-political regions”) are regions defined by the intersection of states in the U.S. or provinces/territories in Canada with BCRs (CEC 1999, Bird Studies Canada and NABCI 2014). Land area (km<sup>2</sup>) was derived from a Geographic Information System overlay of BCRs with states or provinces/territories. Large lakes the size of Utah’s Great Salt Lake or larger (e.g., all of the Great Lakes, along with several large lakes between Lake Winnipeg, Manitoba, and Great Bear Lake, Northwest Territories) were excluded from the calculation of land area (Blancher et al. 2007). Digital species range maps (Ridgely et al. 2005) were used to estimate a range adjustment for extrapolating to regions not sufficiently covered by the BBS surveys (Blancher et al. 2013).

The PIF 2013 approach to estimating population sizes does not incorporate variance or uncertainty into the estimate. Rather, PIF 2013 reports categorical data quality ratings based on variation among BBS routes in each region but does not measure or report variation or uncertainty within the BBS route-level counts or in any of the adjustment factors. Our method largely adheres to the PIF approach as described above but incorporates uncertainty in the regional BBS count estimates as well as the individual

adjustment factors, and propagates that uncertainty through to the final population size estimates.

## Mean BBS route count

There are two sources of variance in the number of birds counted along BBS routes that we accounted for in the population size calculation for each species in each region: within-route and between-route variation. To account for within-route variation, we calculated the mean and variance of total counts for each species along each route within a region (for runs meeting acceptable standards for time of day and weather conditions as determined by the BBS; Pardieck et al. 2016) for the most recent 10 years of data (2015 at the time of this analysis). We used the mean and variance to define a discrete distribution for each route to sample from. If the variance was greater than the mean (count data overdispersed), we used a negative binomial distribution (using the method of moment matching; Hobbs and Hooten 2015). If the variance was not greater than the mean we sampled from a Poisson distribution. For routes with only a single run in the last 10 years, we sampled from a Poisson distribution around the observed count.

For each route in each iteration, we drew 10 random values from either a negative binomial or Poisson distribution around the mean count for that route, regardless of how many times a route was actually run in the past 10 years. This was done to avoid weighting routes by the number of runs conducted. For each iteration of the population size calculation, we pooled the simulated observations for all routes within a region and calculated the mean to determine the mean route-level count for the region.

To account for between-route variation, we used a bootstrap approach to select routes to represent a region. For each iteration, we randomly selected, with replacement, a set of routes equal in size to the total number of routes within each region. Because the sampling was done with replacement, a route may have been selected more than once or not at all for a given iteration.

## Time-of-day adjustment

For many species, detectability can vary by time of day with daily cycles in activity level, i.e., territorial calling or foraging (Thompson et al. 2017). The time-of-day adjustment assumes that birds are present at a constant rate throughout the period of a BBS survey (typically from 30 minutes before sunrise until roughly 4 hours after sunrise), but the probability of detection varies with daily bird activity. The proportion of birds present, but not detected, is therefore proportional to the peak count over the mean count across BBS stops. We revised the time-of-day adjustment for each species by adding additional years of data and using generalized additive models rather than polynomials for the smoothed models. We used all available BBS stop-level data for each species (50-stop data are available from the BBS going back to 1997). Our generalized additive models relate the time of day, i.e., stop number, to the number of birds observed at each stop while accounting for route-level variation and observation trends by year. Additional details describing how we fit the smoothed time-of-day models are available in Appendix 1.

We used the fitted curves to calculate a distribution for the time-of-day adjustment statistic by sampling from the distribution of the fitted curve at each stop. The time-of-day adjustment was

estimated as the maximum divided by the mean across stops. The result of this adjustment is that extrapolated population estimates are based on the counts of birds during their peak time of detection, reducing underestimates due to undetected individuals. We acknowledge that there will still be an unmeasured fraction of the population that goes undetected, even during the time of peak detection within the survey, resulting in an underestimate in population sizes, particularly for some cryptic and/or nocturnal/crepuscular species. Alternative approaches to accommodate an empirical estimate of detection probability are feasible for some species, but these approaches remain beyond the scope of the current improvements to the PIF approach employed here.

### Pair adjustment

PIF assigns a pair adjustment multiplier to the PS calculation to account for the probability that for some species, one member of a breeding pair, e.g., singing male, is more likely to be detected during a sight/sound survey than the other partner. For those species, it is assumed that for every bird recorded at the peak time of day, there may be a second bird present that was not recorded. Therefore, the pair adjustment parameter should theoretically not be less than 1 (both sexes equally likely to be detected) nor greater than 2 (only one member of a pair detected). PIF assigned each species to one of five categories (1.0, 1.25, 1.5, 1.75, or 2.0) based on the likelihood that most detections were biased to one sex. To assign these categories, PIF reviewed available evidence related to observational sex-bias. Such evidence included the sex-ratio of the birds observed, the species' breeding phenology, and whether the species was most often detected during the dawn chorus, i.e., a singing male, and/or singly or in groups (Blancher et al. 2013).

We maintained these categorical assignments but incorporated uncertainty in the assignments by replacing them with truncated normal distributions. The means of the distributions were equal to the previously assigned PIF pair adjustment values and truncated at 1.0 and 2.0. We selected a variance term to describe the pair adjustment distributions ( $sd = 0.13$ ) that would allow overlap between the categorical distributions.

### Detection-distance adjustment

The PIF approach to estimating population sizes is particularly sensitive to the detection-distance adjustment (Thogmartin et al. 2006, Thogmartin 2010). Comparisons with field tests of detection distances suggest that the assigned detection-distance categories for the PIF population size calculation may be too large for many species (Confer et al. 2008, Hamel et al. 2009, Matsuoka et al. 2012, Twedt 2015) potentially resulting in overly conservative total population estimates. The PIF approach assigns each species to one of seven detection-distance categories based on a combination of published maximum detection distances (see Rosenberg and Blancher 2005 and references therein), expert opinion, and by comparing relative distances across species in regional point count surveys that included detection distances (Blancher et al. 2013).

We still lack an empirically based approach that can be applied systematically across species to estimate detection distances. To incorporate uncertainty bounds on the PIF approach with respect to the detection-distance adjustment, we randomly sampled from uniform distributions with the lower bound being 80% of the difference between the assigned distance adjustment category for a given species and the next lower category, and the upper bound

being a 10% increase over the assigned category. We assigned the sampling distribution in this manner so that the distribution would (1) include the previously assigned detection-distance adjustment, (2) be broad enough to encompass the uncertainty of this parameter, and (3) partly account for recent empirical estimates of detection distance suggesting that the assigned PIF detection-distance categories may be generally overestimated (see Matsuoka et al. 2012, Twedt 2015).

### Estimation procedure

PIF reports population size estimates for > 450 landbird species regularly breeding within North America by relying on multiple data sources (Rosenberg et al. 2017). For the present study, we focus on the 336 species for which the BBS is the sole or major contributing data source. The remaining species for which PIF reports population size estimates rely on non-BBS data sources, such as species-specific surveys, that were deemed more reliable than BBS for those species (Blancher et al. 2013, Rosenberg et al. 2017).

To incorporate uncertainty into the PIF process, we employed a Monte Carlo approach to estimate population size distributions by randomly sampling from each variable or nonfixed model component prior to calculating the population size estimate. The components of the model that were variable were either species-specific (pair adjustment, detection-distance adjustment, time-of-day adjustment) or species-by-region specific (mean route-level BBS count). Fixed model components were treated as known without error (range adjustment, region area). Population size distributions were derived by making 1000 iterations of the calculation for each species in each region by making independent random draws from each model component. We did not model any correlation between model parameters. Results are presented as the median result of the 1000 iterations with either 80% or 95% bounds. To measure the degree of uncertainty around the population size estimates, we calculated a standardized interquartile distance as the difference between upper and lower 80% estimation bounds divided by the mean population size estimate.

Appendix 2 includes R scripts and for downloading and processing BBS count data, code and accompanying parameter files for running the Monte Carlo population size estimation, and summary output data of population size estimates with uncertainty ranges. Aside from the changes noted here, the PIF population size estimation is as described previously in the PIF database and handbooks (Blancher et al. 2007, 2013).

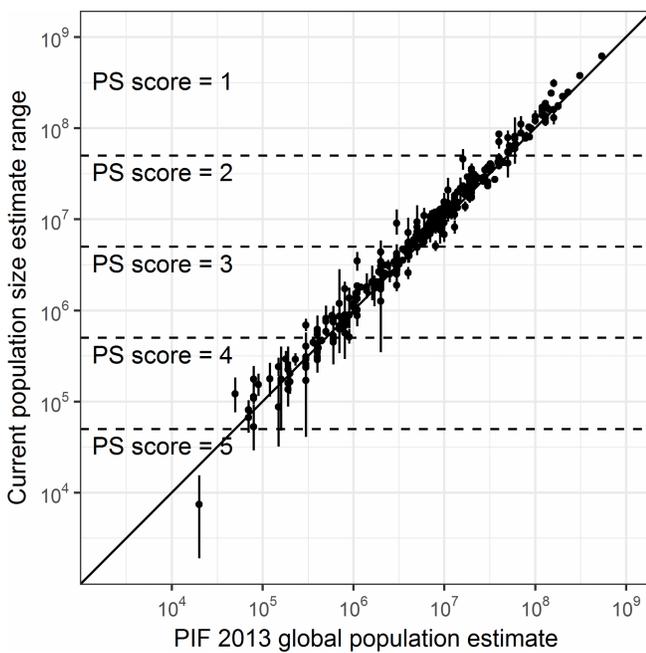
## RESULTS

### Population size estimates

For the 336 landbird species whose estimates rely on the BBS, PIF 2013 estimated 5.85 billion adult birds in North America, and 7.67 billion adult birds globally based on data from 1998 to 2007. Using our refined methods, and BBS data from 2006 to 2015, we estimate 6.34–7.67 billion in North America and 8.25–10.05 billion globally at 80% estimation bounds, which is approximately 8–30% higher than the PIF 2013 estimates. At 95% estimation bounds, we estimate 6.05–8.08 billion adult birds in North America and 7.88–10.63 billion globally for the comparison set of species.

At 80% bounds, 47.6% of species have new uncertainty ranges that encompass the PIF 2013 estimates, this fraction of species increases to 63.5% at the 95% bounds. Of those species for which our new estimate range does not include the PIF 2013 North American estimate, 88.6% and 90.2% have ranges greater than the previous point estimates at 80% and 95% bounds, respectively (Fig. 1).

**Fig. 1.** Comparison between previous global Partners in Flight (PIF) population size estimate (PIF 2013) and estimates from this study. Points indicate the median global population size estimate. Error bars show the 80% bounds on the current estimates. Areas between the dashed lines indicate the corresponding ranges of the PIF population size (PS) score used for species vulnerability assessment.

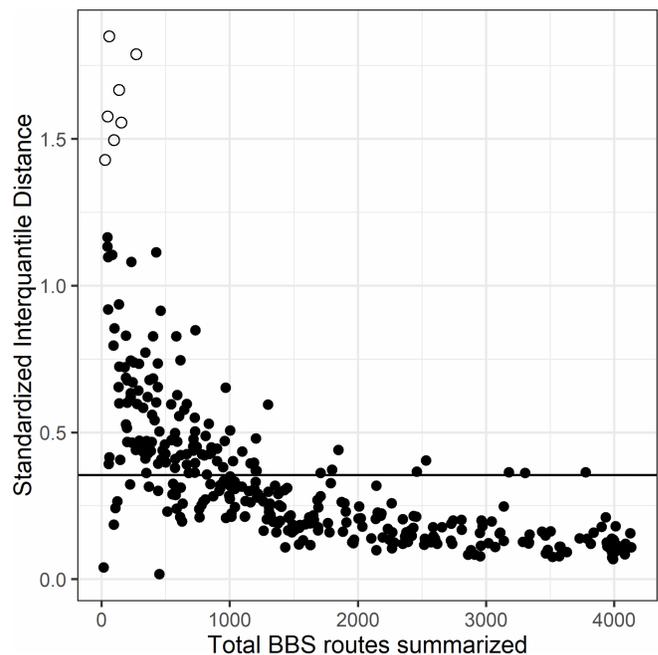


### Population size uncertainty

The mean standardized interquartile distance across species for the combined North American population totals was 0.35 (range: 0.02–1.85). This estimation excludes physio-political regions where population size estimates were based on non-BBS sources. The PIF 2013 population size estimates did not include uncertainty bounds comparable with the new approach; however, PIF's assignment of species to PS scores based on order-of-magnitude categories of abundance (Table 1; Blancher et al. 2007) is robust only if variance in the estimates is generally less than an order of magnitude. Our new uncertainty ranges were generally much tighter than an order of magnitude, with 76% of species having upper 95% estimates less than twice their lower 95% estimates (90% of species for 80% uncertainty estimates). Only seven species, all with small populations or restricted ranges within North America, had upper 95% estimates that were an order-of-magnitude larger or more than the lower 95% estimates (Mexican Whip-poor-will [*Antrostomus arizonae*], Allen's Hummingbird [*Selasphorus sasin*], Flammulated Owl [*Psiloscops*

*flammeolus*], Elf Owl [*Micrathene whitneyi*], Black-whiskered Vireo [*Vireo altiloquus*], Mexican Jay [*Aphelocoma wollweberi*], and Seaside Sparrow [*Ammodramus maritimus*]). The size of the standardized interquartile distance was strongly influenced by the number of BBS routes used in the analysis, with fewer routes analyzed leading to larger uncertainty bounds (Fig. 2).

**Fig. 2.** Standardized interquartile distances for U.S./Canada population size estimates plotted against number of Breeding Bird Survey (BBS) routes summarized. The distance statistic is calculated at 80% bounds and standardized by the mean population size. Note that the number of routes summarized may be greater than the number of routes a species has been observed on in the past 10 years because estimation is made by sampling all BBS routes within a region. Open dots indicate species whose 95% uncertainty range exceeded one order of magnitude. The solid line shows the mean across species.



We found that the PIF 2013 data quality scores (5 color-coded rating categories ranging from good to poor quality; Blancher et al. 2007) were consistent with the widths of our newly calculated uncertainty measures. We quantified the width of uncertainty by calculating a standardized interquartile range (distance between 80% bounds divided by the mean population size estimate). Comparing the uncertainty ranges with the PIF 2013 quality ratings showed increasing uncertainty with increasingly poorer quality ratings. We also found greater variation in the uncertainty widths with increasingly poorer PIF data quality ratings (Fig. 3).

### PIF species assessment scores

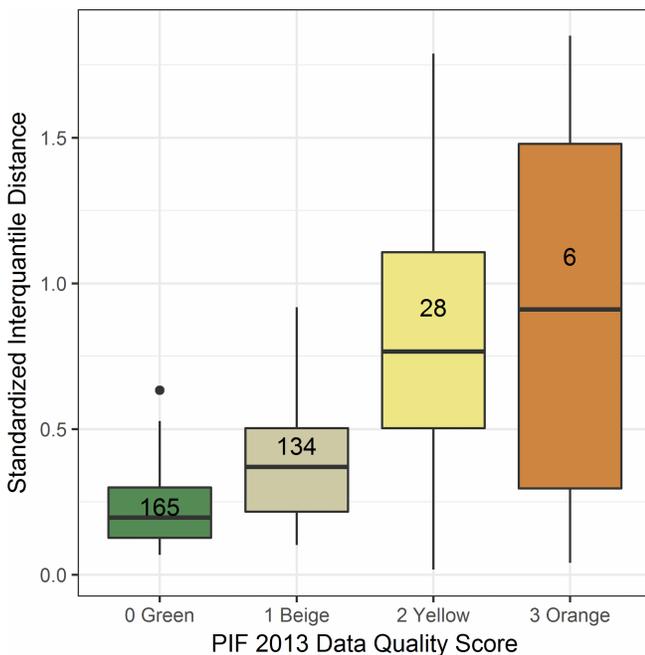
One important consideration of our new methods is whether incorporating uncertainty in PS estimation would lead to uncertainty in assigning PS scores in the PIF species assessment, which could potentially lead to uncertainty in a species' conservation status. We found 87.5% (294) of species fell into single global PS score categories across the entire range of

**Table 2.** Uncertainty in assigning Partners in Flight (PIF) population size (PS) scores, and uncertainty ranges in population size estimates and the four nonfixed model components. The latter uncertainty ranges are reported as standardized interquartile distances, calculated at 80% distributional bounds divided by the mean.

PS score	Number of species	Mean number of routes	% Agreement in PS scoring		Mean standardized interquartile distances for full U.S./Canada population size and individual model components				
			At 80% bounds	At 95% bounds	Population size	Mean BBS count <sup>†</sup>	Time-of-day adjustment	Pair adjustment	Distance adjustment
1	45	2635	93.3	88.9	0.20	0.03	0.08	0.13	0.58
2	152	1730	94.1	86.8	0.26	0.05	0.05	0.15	0.58
3	102	1034	81.4	67.6	0.43	0.14	0.06	0.15	0.58
4 & 5	37	576	70.3	56.8	0.71	0.17	0.07	0.15	0.60
All species	336	1512	87.5	78.0	0.35	0.09	0.06	0.15	0.58

<sup>†</sup>Summarized for each species over all BBS routes with nonzero count averages weighted by the percent of the total population for each region.

**Fig. 3.** Standardized interquartile distance (distance between 80% bounds divided by the mean) grouped by the Partners in Flight (PIF 2013) data quality categories for North American population size estimates. PIF data quality categories are color-coded and correspond to ordered levels of data quality or quantity ranging from green, the highest quality category, through beige, yellow, orange, and red (Blancher et al. 2007). Not shown: Northern Wheatear (*Oenanthe oenanthe*), the sole species with a red PIF data quality score in this analysis, with standardized interquartile range = 0.41. Each box shows the number of species associated with each score. Horizontal lines indicate the median value, boxes span first and third quartiles, whiskers extend 1.5 x the interquartile range (or minimum value for lower whisker if less), remaining value is plotted as a point.



uncertainty at 80% bounds. At 95% bounds of uncertainty, 78.0% (262) of species fell into single PS score categories. We found one species, Mexican Whip-poor-will, in which the population size uncertainty range spanned more than the entire upper and lower bound of the presumed PS score at 80% bounds. Widening the uncertainty range to 95% bounds added an additional two species (Black-whiskered Vireo and Seaside Sparrow). Of these three species, only the Seaside Sparrow's range is entirely within the coverage area of the BBS; Mexican Whip-poor-will and Black-whiskered Vireo global population estimates are based partially on range adjustment extrapolations with only a small portion of the total range within the area of BBS coverage (19% for Mexican Whip-poor-will and 6% for Black-whiskered Vireo).

In general, species with smaller estimated populations within North America were more likely to have larger standardized interquartile distances and were, therefore, also more likely to have uncertainty in the assignment of a PS score (Table 2). Our analysis would result in 20 species being assigned a different PS score relative to scores based on the PIF 2013 population size database, though only nine of those species would not include the PIF 2013 score within the uncertainty bounds of the new analysis.

### Components of uncertainty

Of the four model components serving as sources of uncertainty, interquartile variance was largest for the distance adjustment (mean of 0.58; Table 2), and smallest for the time-of-day adjustment (0.06) and BBS average count (0.09). Variance in the distance adjustment, pair adjustment, and time-of-day adjustment was consistent across all population size categories (Table 2). The variance of the weighted mean BBS count was larger for relatively smaller population sizes (PS score = 3 & 4; Table 2) leading to the pattern of larger uncertainty ranges for smaller population size estimates (Table 2).

### DISCUSSION

The effort to generate standardized population size approximations for several hundred bird species has been an important contribution to bird conservation in North America. These estimates are an important component of the PIF species assessment but have also been used for other conservation planning and assessment purposes. We present new methods for characterizing and expressing the magnitude of uncertainty

around PIF population size approximations and, as such, increase the utility of these estimates in conservation planning and risk assessment. Previous estimates published by PIF 2013 used categorical data quality ratings, which carried an implicit, but unknown, level of uncertainty. Categorical data quality ratings were useful in providing a sense of confidence in the source of data used in estimation, but unless they are calibrated to measurable ranges, they do not provide information about the uncertainty inherent in the estimation process (Akçakaya et al. 2000).

Our improved methodology, replacing categorical ratings with quantitative uncertainty bounds, has important implications for both the PIF species assessment process and for the use of population size estimates based on the PIF approach. These refinements will be particularly important for comparing population size estimates to a regional conservation target or recovery goal, or for assessments of avian mortality impacts that rely on population size estimates to contextualize potential population-level consequences. Our approach provides an explicit uncertainty range on population size estimates, allowing for a quantitative assessment of whether a species meets a given assessment criterion, conservation target, or has the potential to be severely impacted by a novel source of mortality.

### Implications for PIF species assessment

The primary use of population size estimates in the PIF species assessment process is to assign each species a categorical PS score, based on order-of-magnitude ranges of population sizes: the smaller the population size, the higher the PS score and implied level of conservation vulnerability (Panjabi et al. 2005, 2012). Although tremendous precision is not necessary to place most species into the correct PS score category, large uncertainty in estimates could result in species being assigned an incorrect PS score. Previous PIF 2013 data quality ratings did not indicate when the uncertainty in a population size estimate might span more than one PS category. In contrast, our new methodology explicitly provides uncertainty bounds that may span more than one PS category, indicating when there is uncertainty in the score assignment for a particular species that may affect PIF's overall assessment for that species.

Fortunately, we found the width of the uncertainty bounds for the majority of landbird species in our analysis to be much narrower than the order-of-magnitude distance between scoring categories, suggesting a high level of confidence in PIF's PS score assignment for most species. Some species with estimates near a boundary between two categorical scores may nevertheless span more than one category. In addition, we found that confidence in the assignment of a single categorical PS score decreased with increasing vulnerability level (higher PS score), highlighting the importance of understanding uncertainty in estimates for species of high conservation concern. This uncertainty in assigning PS scores need not be viewed as problematic for future PIF species assessments, however, as it allows for calculating how much of the uncertainty distribution lies in one category versus another (Akçakaya et al. 2000). Our approach also opens the possibility to consider a range of PS scores where necessary and may prove beneficial for identifying when an alternative data-source should be sought that might have less uncertainty.

### Patterns and sources of uncertainty

We found a general pattern where the width of the uncertainty bounds for many species roughly approximated the categorical data quality ratings from PIF 2013 (Figure 3). The PIF 2013 data quality ratings were based on three subcomponents: a variance component based on the amount of variation around the route-level counts within a region; a sample-size component based on the number of BBS routes within a region; and a BBS coverage component based on the proportion of a species breeding range that is sampled by the BBS. The overall data-quality rating was assigned based on the lowest scoring subcomponent. Comparing our population size uncertainty ranges with the PIF 2013 categorical data quality ratings showed wider uncertainty bounds associated with lower quality ratings. We also found increasing variation in the width of uncertainty bounds with lower PIF data quality. The pattern of the standardized interquartile distances being generally in concordance with the PIF 2013 data quality rating is not surprising because, aside from the coverage component, the count variance and the number of routes contribute to both metrics of uncertainty. This similarity does not diminish the importance of quantifying variance rather than summarizing it into categories, particularly because the quantified variances presented here incorporate measures of variance in distance and pair adjustments that were not included in PIF's data quality ratings.

At regional scales, wide uncertainty bounds in population size estimates are often a consequence of a species being sampled on few BBS routes within a region. Additionally, whether routes sufficiently sample available land cover or vegetation types within a region can also influence the width of the uncertainty bounds. For rare species and habitat specialists that are not repeatedly sampled at different locations, resampling routes to derive uncertainty bounds can result in some of the less common land covers not being represented in any particular random draw. This occasionally resulted in the lower bound of the uncertainty range being zero in physio-political regions where we know the species to occur. This source of negative bias would be expected for any sort of extrapolation relying upon few data sources (Machtans and Thogmartin 2014), emphasizing the need for as complete and representative a coverage as possible in future surveys. Although PIF continues to provide population size estimates at BCR, state, and province scales, new uncertainty bounds derived from our methodology will allow users to evaluate the reliability of these estimates.

A primary factor influencing species detectability, and subsequent population estimates based on survey data, is variation within and among observers (Sauer et al. 2013). Other attempts at extrapolating BBS data to population size estimates (Runge et al. 2009, Thogmartin et al. 2014) relied on hierarchical models of BBS counts assumed to derive from a Poisson distribution accounting for year and region as well as overdispersion. In these models, observer-level differences are controlled for with an observer random effect and a fixed effect for novice observers. In our approach, the uncertainty caused by observers is subsumed by the time-of-day and distance adjustments because both relate to observation processes that may vary by individual observers.

## Impact of individual model components

With our improved methodology, PIF will be updating population size estimates for all North American landbirds. These new estimates tend to be larger with the refined estimation approach we employ here than the previous PIF 2013 published estimates. Although our new estimates reflect a more recent decade of BBS count data (2006–2015 vs. 1998–2007), we caution against attributing any differences to actual trends in bird populations. The tendency toward higher population size estimates in our analysis likely arises from changes in the adjustment factors used in the calculations (detection-distance, time-of-day). Overall, there is general concordance between the PIF 2013 estimates and our present estimates (Fig. 1) as would be expected because the calculation formula and data sources are similar. Only by applying identical methodology to both decades of BBS data would we be able to directly compare estimates from the two time periods.

The individual model components in the estimation process can have large effects on both population size estimates and the width of the uncertainty bounds around those estimates. A previous review of the general PIF approach to population size estimates pointed out several areas where the method is sensitive to inherent assumptions and adjustment factors used in the estimation (Thogmartin et al. 2006). Our analysis of interquartile variance by component indicated that variance was highest for the distance adjustment factor (Table 2). A sensitivity analysis conducted by Thogmartin (2010) also suggested that relatively small changes in this component have large impacts on the population size estimate. For common birds (PS score = 1 & 2), improvements to the distance factor will have the greatest impact on improving both the precision and accuracy of population estimates. In contrast, for species with small population sizes (PS score = 4 & 5), overall uncertainty was higher than variance in the distance factor, suggesting that for these species finding alternative sources of data for population size estimation may be the best action.

The general shift toward larger population size estimates compared with PIF 2013 was also due to a decision to shift the range of the detection-distance adjustment distribution so that a greater proportion of random draws were lower (as opposed to higher) than the previous fixed detection-distance adjustment values. Several empirical studies of effective detection distances (Confer et al. 2008, Hamel et al. 2009, Matsuoka et al. 2012, Twedt 2015) suggested that the previous detection distances used by PIF were systematically too large, resulting in underestimates of population size, and our decision here was an attempt to address this concern. However, we acknowledge that PIF's categorical distance correction parameter remains imprecise and would be greatly improved with an empirical analysis incorporating a wide range of taxa across a range of land-cover types.

## Future of population size estimation

Here we propose a means to improve an aspect of how PIF estimates and presents population sizes by providing quantitative uncertainty bounds. However, there is still room for additional improvements for estimating the size of North American bird populations. These improvements might proceed down alternative paths. First, PIF may maintain the same fundamental approach while making further refinements to individual model components. Alternately, PIF may pursue entirely new

approaches and data sources beyond what has been used to date. We note that the distributions used for various adjustment factors, as well as the other components of the estimation process, can easily be refined or replaced in the work flow using the R scripts provided with this paper. It will therefore be relatively simple to make further adjustments to individual model components and generate refined population size estimates in the future as more data and new analyses become available. For example, the available analyses to assign effective detection distances do not yet cover the full range of species, are limited in geographic coverage, or are not conducted using road-side point-counts and thus are not directly applicable to estimating population size based on BBS counts. We hope that future work in this area will allow us to improve the detection-distance adjustment factor, because it is an important source of uncertainty.

The main source of data for PIF population estimation has been the BBS, which was designed to estimate avian population trends through time. The BBS methodology does not include techniques to estimate population density for a known area (Link and Sauer 1998, Sauer et al. 2013), which creates challenges for using it as a data source for estimating population size. However, no other systematic bird survey program has the breadth of spatial, temporal, and taxonomic coverage as the BBS (Rosenberg et al. 2017). Future iterations might consider alternative monitoring approaches designed to generate population size estimates without the need to mark and recapture individuals (Royle 2004, Kéry et al. 2005, Dail and Madsen 2011, Pavlacky et al. 2012), as well as current bird survey programs in North America that incorporate field methodology for estimating abundance (Matsuoka et al. 2014, Woiderski et al. 2018). For species and regions where alternative data are available, those approaches may be preferable to extrapolations of data from the BBS. In addition, the rapid growth of eBird (Sullivan et al. 2014), and range-wide, year-round modeling of abundance using eBird data (Fink et al. 2010, Johnston et al. 2015), hold promise for informing population estimates of many species. Integrating available survey data across multiple studies and monitoring programs into a single model may allow for shared inference and calibration of more accurate population estimates (Sólymos et al. 2013). Constructing such a model for all landbirds breeding in North America, as well as extending these approaches to the nonlandbird avifauna, remain areas for future work.

*Responses to this article can be read online at:*  
<http://www.ace-eco.org/issues/responses.php/1331>

---

## Acknowledgments:

*We wish to acknowledge the staff, scientists, and volunteers who have contributed over the years to the invaluable data resource that is the North American Breeding Bird Survey. We also thank the PIF International Science Committee for many fruitful ideas and conversations on this topic. This manuscript benefited from comments provided on earlier drafts from J. Sauer, and four anonymous reviewers. Any use of trade, product, or firm names are for descriptive purposes only and do not imply endorsement by the U.S. Government.*

---

## LITERATURE CITED

- Akçakaya, H. R., S. Ferson, M. A. Burgman, D. A. Keith, G. M. Mace, and C. R. Todd. 2000. Making consistent IUCN classifications under uncertainty. *Conservation Biology* 14:1001-1013. <http://dx.doi.org/10.1046/j.1523-1739.2000.99125.x>
- Bickford, D., M. R. C. Posa, L. Qie, A. Campos-Arceiz, and E. P. Kudavidanage. 2012. Science communication for biodiversity conservation. *Biological Conservation* 151:74-76. <http://dx.doi.org/10.1016/j.biocon.2011.12.016>
- Bird Studies Canada and NABCI. 2014. *NABCI Bird conservation regions*. Bird Studies Canada on behalf of the North American Bird Conservation Initiative, Port Rowan, Ontario, Canada. [online] URL: <http://www.birdscanada.org/research/gislab/index.jsp?targetpg=bcr>
- Blancher, P. J., K. V. Rosenberg, A. O. Panjabi, B. Altman, J. Bart, C. J. Beardmore, G. S. Butcher, D. D. Demarest, R. Dettmers, E. H. Dunn, W. E. Easton, W. C. Hunter, E. E. Iñigo-Elias, D. N. Pashley, C. J. Ralph, T. D. Rich, C. M. Rustay, J. M. Ruth, and T. C. Will. 2007. *Guide to the Partners in Flight population estimates database. North American Landbird Conservation Plan 2004. Partners in Flight Technical Series No 5*.
- Blancher, P. J., K. V. Rosenberg, A. O. Panjabi, B. Altman, A. R. Couturier, and W. E. Thogmartin. 2013. *Handbook to the Partners in Flight population estimates database, Version 2.0. Partners in Flight Technical Series No 6*.
- Briggs, S. V. 2006. Integrating policy and science in natural resources: why so difficult? *Ecological Management and Restoration* 7:37-39. <http://dx.doi.org/10.1111/j.1442-8903.2006.00245.x>
- Carter, M. F., W. C. Hunter, D. N. Pashley, and K. V. Rosenberg. 2000. Setting conservation priorities for landbirds in the United States: the Partners in Flight approach. *Auk* 117:541-548. [http://dx.doi.org/10.1642/0004-8038\(2000\)117\[0541:SCPFLI\]2.0.CO;2](http://dx.doi.org/10.1642/0004-8038(2000)117[0541:SCPFLI]2.0.CO;2)
- Commission for Environmental Cooperation (CEC). 1999. *A proposed framework for delineating ecologically-based planning, implementation, and evaluation units for cooperative bird conservation in the U.S.* U.S. Fish and Wildlife Service, Washington, D.C., USA.
- Confer, J. L., R. E. Serrell, M. Hager, and E. Lahr. 2008. Field tests of the Rosenberg-Blancher method for converting point counts to abundance estimates. *Auk* 125:932-938. <http://dx.doi.org/10.1525/auk.2008.06263>
- Dail, D., and L. Madsen. 2011. Models for estimating abundance from repeated counts of an open metapopulation. *Biometrics* 67:577-587. <http://dx.doi.org/10.1111/j.1541-0420.2010.01465.x>
- Erickson, W. P., M. M. Wolfe, K. J. Bay, D. H. Johnson, and J. L. Gehring. 2014. A comprehensive analysis of small-passerine fatalities from collision with turbines at wind energy facilities. *PLoS ONE* 9:e107491. <https://doi.org/10.1371/journal.pone.0107491>
- Faber-Langendoen, D., J. Nichols, L. Master, K. Snow, A. Tomaino, R. Bittman, G. Hammerson, B. Heidel, L. Ramsay, A. Teucher, and B. Young. 2012. *NatureServe conservation status assessments: methodology for assigning ranks*. NatureServe, Arlington, Virginia, USA.
- Finch, D. M., and P. W. Stangel. 1993. Status and management of neotropical migratory birds: introduction. Pages 1-4 in D. M. Finch and P. W. Stangel, editors. *Status and management of neotropical migratory birds. General Technical Report RM-229*. U. S. Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado, USA. <http://dx.doi.org/10.2737/RM-GTR-229>
- Fink, D., W. M. Hochachka, B. Zuckerberg, D. W. Winkler, B. Shaby, M. A. Munson, G. Hooker, M. Riedewald, D. Sheldon, and S. Kelling. 2010. Spatiotemporal exploratory models for broad-scale survey data. *Ecological Applications* 20:2131-2147. <http://dx.doi.org/10.1890/09-1340.1>
- Hamel, P. B., M. J. Welton, C. G. Smith III, and R. P. Ford. 2009. Test of Partners in Flight effective detection distance for Cerulean Warbler. Pages 328-333 in T. Rich, C. Arizmendi, D. Demarest, and C. Thompson, editors. *Tundra to tropics: connecting birds, habitats and people. Proceedings of the 4th International Partners in Flight Conference*, McAllen, Texas, USA.
- Hobbs, N. T., and M. B. Hooten. 2015. *Bayesian models: a statistical primer for ecologists*. First edition. Princeton University Press, Princeton, New Jersey, USA.
- Hunter, W. C., M. F. Carter, D. N. Pashley, and K. Barker. 1993. The Partners in Flight prioritization scheme. Pages 109-119 in D. M. Finch and P. W. Stangel, editors. *Status and management of Neotropical migratory birds. General Technical Report RM-229*. U.S. Department of Agriculture, Forest Service, Rocky Mountain Forest and Range Experiment Station, Fort Collins, Colorado, USA.
- International Union for Conservation of Nature (IUCN). 2012. *IUCN Red List categories and criteria, version 3.1, second edition*. IUCN Species Survival Commission, Gland, Switzerland.
- Johnson, F. A., M. A. H. Walters, and G. S. Boomer. 2012. Allowable levels of take for the trade in Nearctic songbirds. *Ecological Applications* 22:1114-1130. <http://dx.doi.org/10.1890/11-1164.1>
- Johnston, A., D. Fink, M. D. Reynolds, W. M. Hochachka, B. L. Sullivan, N. E. Bruns, E. Hallstein, M. S. Merrifield, S. Matsumoto, and S. Kelling. 2015. Abundance models improve spatial and temporal prioritization of conservation resources. *Ecological Applications* 25:1749-1756. <http://dx.doi.org/10.1890/14-1826.1>
- Kéry, M., J. A. Royle, and H. Schmid. 2005. Modeling avian abundance from replicated counts using binomial mixture models. *Ecological Applications* 15:1450-1461. <http://dx.doi.org/10.1890/04-1120>
- Lande, R. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. *American Naturalist* 142:911-927. <http://dx.doi.org/10.1086/285580>
- Link, W. A., and J. R. Sauer. 1998. Estimating population change from count data: application to the North American breeding bird survey. *Ecological Applications* 8:258-268. [http://dx.doi.org/10.1890/1051-0761\(1998\)008\[0258:EPCFCD\]2.0.CO;2](http://dx.doi.org/10.1890/1051-0761(1998)008[0258:EPCFCD]2.0.CO;2)
- Longcore, T., C. Rich, P. Mineau, B. MacDonald, D. G. Bert, L. M. Sullivan, E. Mutrie, S. A. Gauthreaux, M. L. Avery, R. L. Crawford, A. M. Manville, E. R. Travis, and D. Drake. 2013. Avian

- mortality at communication towers in the United States and Canada: which species, how many, and where? *Biological Conservation* 158:410-419. <http://dx.doi.org/10.1016/j.biocon.2012.09.019>
- Longcore, T., and P. A. Smith. 2013. On avian mortality associated with human activities. *Avian Conservation and Ecology* 8(2):1. <http://dx.doi.org/10.5751/ACE-00606-080201>
- Loss, S. R., T. Will, and P. P. Marra. 2013. Estimates of bird collision mortality at wind facilities in the contiguous United States. *Biological Conservation* 168:201-209. <http://dx.doi.org/10.1016/j.biocon.2013.10.007>
- Loss, S. R., T. Will, and P. P. Marra. 2014. Refining estimates of bird collision and electrocution mortality at power lines in the United States. *PLoS ONE* 9:e101565. <https://doi.org/10.1371/journal.pone.0101565>
- Machtans, C. S., and W. E. Thogmartin. 2014. Understanding the value of imperfect science from national estimates of bird mortality from window collisions. *Condor* 116:3-8. <http://dx.doi.org/10.1650/CONDOR-13-134.1>
- Matsuoka, S. M., E. M. Bayne, P. Sólymos, P. C. Fontaine, S. G. Cumming, F. K. A. Schmiegelow, and S. J. Song. 2012. Using binomial distance-sampling models to estimate the effective detection radius of point-count surveys across boreal Canada. *Auk* 129:268-282. <http://dx.doi.org/10.1525/auk.2012.11190>
- Matsuoka, S. M., C. L. Mahon, C. M. Handel, P. Sólymos, E. M. Bayne, P. C. Fontaine, and C. J. Ralph. 2014. Reviving common standards in point-count surveys for broad inference across studies. *Condor* 116:599-609. <http://dx.doi.org/10.1650/CONDOR-14-108.1>
- McKinney, M. L. 1997. Extinction vulnerability and selectivity: combining ecological and paleontological views. *Annual Review of Ecology and Systematics* 28:495-516. <http://dx.doi.org/10.1146/annurev.ecolsys.28.1.495>
- Panjabi, A. O., P. J. Blancher, R. Dettmers, and K. V. Rosenberg. 2012. *The Partners in Flight handbook on species assessment. Version 2012. Partners in Flight Technical Series No. 3*. Rocky Mountain Bird Observatory, Brighton, Colorado, USA.
- Panjabi, A. O., E. H. Dunn, P. J. Blancher, W. C. Hunter, B. Altman, J. Bart, C. J. Beardmore, H. Berlanga, G. S. Butcher, S. K. Davis, D. W. Demarest, R. Dettmers, W. Easton, H. Gomez de Silva Garza, E. E. Iñigo-Elias, D. N. Pashley, C. J. Ralph, T. D. Rich, K. V. Rosenberg, C. M. Rustay, J. M. Ruth, J. S. Wendt, and T. C. Will. 2005. *The Partners in Flight handbook on species assessment. Version 2005. Partners in Flight Technical Series No. 3*. Rocky Mountain Bird Observatory, Brighton, Colorado, USA.
- Pardieck, K. L., D. J. Ziolkowski, M.-A. R. Hudson, and K. Campbell. 2016. *North American Breeding Bird Survey Dataset 1966-2015, version 2015.1*. U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland, USA.
- Partners in Flight (PIF). 2013. *PIF population estimates database*. Partners in Flight. [online] URL: <http://pif.birdconservancy.org/PopEstimates>
- Pavlacky, D. C., J. A. Blakesley, G. C. White, D. J. Hanni, and P. M. Lukacs. 2012. Hierarchical multi-scale occupancy estimation for monitoring wildlife populations. *Journal of Wildlife Management* 76:154-162. <http://dx.doi.org/10.1002/jwmg.245>
- Potter, B., G. J. Soulliere, D. N. Ewert, M. G. Knutson, W. E. Thogmartin, J. S. Castrale, and M. Roell. 2007. *Upper Mississippi River and Great Lakes Region Joint Venture Landbird Habitat Conservation Strategy*. U.S. Fish and Wildlife Service, Fort Snelling, Minnesota, USA.
- Purvis, A., J. L. Gittleman, G. Cowlshaw, and G. M. Mace. 2000. Predicting extinction risk in declining species. *Proceedings of the Royal Society B: Biological Sciences* 267:1947-1952. <http://dx.doi.org/10.1098/rspb.2000.1234>
- R Core Team. 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Regan, H. M., H. R. Akçakaya, S. Ferson, K. V. Root, S. Carroll, and L. R. Ginzburg. 2003. Treatments of uncertainty and variability in ecological risk assessment of single-species populations. *Human and Ecological Risk Assessment: An International Journal* 9:889-906. <http://dx.doi.org/10.1080/713610015>
- Rich, T. D., C. J. Beardmore, H. Berlanga, P. J. Blancher, M. S. W. Bradstreet, G. S. Butcher, D. W. Demarest, E. H. Dunn, W. C. Hunter, E. E. Iñigo-Elias, J. A. Kennedy, A. M. Martell, A. O. Panjabi, D. N. Pashley, K. V. Rosenberg, C. M. Rustay, J. S. Wendt, and T. C. Will. 2004. *North American landbird conservation plan*. Cornell Lab of Ornithology, Ithaca, New York, USA.
- Ridgely, R. S., T. F. Allnutt, T. Brooks, D. K. McNicol, D. W. Mehlman, B. E. Young, and J. R. Zook. 2005. *Digital distribution maps of the birds of the Western Hemisphere, version 2.1*. NatureServe, Arlington, Virginia, USA.
- Rosenberg, K. V., and P. J. Blancher. 2005. Setting numerical population objectives for priority landbird species. Pages 57-67 in C. J. Ralph and T. D. Rich, editors. *Bird conservation implementation and integration in the Americas: proceedings of the Third International Partners in Flight Conference. 2002*. U.S. Forest Service, Albany, California, USA.
- 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. *Condor* 119:594-607. <http://dx.doi.org/10.1650/CONDOR-17-57.1>
- Rosenberg, K. V., J. A. Kennedy, R. Dettmers, R. P. Ford, D. Reynolds, J. D. Alexander, C. J. Beardmore, P. J. Blancher, R. E. Bogart, G. S. Butcher, A. F. Camfield, A. Couturier, D. W. Demarest, W. E. Easton, J. J. Giocomo, R. H. Keller, A. E. Mini, A. O. Panjabi, D. N. Pashley, T. D. Rich, J. M. Ruth, H. Stabins, J. Stanton, and T. Will. 2016. *Partners in Flight Landbird Conservation Plan: 2016 revision for Canada and Continental United States*. Partners in Flight Science Committee. [online] URL: <https://www.partnersinflight.org/resources/the-plan/>
- Royle, J. A. 2004. N-mixture models for estimating population size from spatially replicated counts. *Biometrics* 60:108-115. <http://dx.doi.org/10.1111/j.0006-341X.2004.00142.x>
- Runge, M. C., J. R. Sauer, M. L. Avery, B. F. Blackwell, and M. D. Koneff. 2009. Assessing allowable take of migratory birds. *Journal of Wildlife Management* 73:556-566. <http://dx.doi.org/10.2193/2008-090>
- Sauer, J. R., W. A. Link, J. E. Fallon, K. L. Pardieck, and D. J. Ziolkowski. 2013. *The North American Breeding Bird Survey*

1966-2011: summary analysis and species accounts. *North American Fauna* 79:1-32. <http://dx.doi.org/10.3996/nafa.79.0001>

Sólymos, P., S. M. Matsuoka, E. M. Bayne, S. R. Lele, P. Fontaine, S. G. Cumming, D. Stralberg, F. K. A. Schmiegelow, and S. J. Song. 2013. Calibrating indices of avian density from non-standardized survey data: making the most of a messy situation. *Methods in Ecology and Evolution* 4:1047-1058. <http://dx.doi.org/10.1111/2041-210X.12106>

Sullivan, B. L., J. L. Aycrigg, J. H. Barry, R. E. Bonney, N. Bruns, C. B. Cooper, T. Damoulas, A. A. Dhondt, T. Dietrich, A. Farnsworth, D. Fink, J. W. Fitzpatrick, T. Fredricks, J. Gerbracht, C. Gomes, W. M. Hochachka, M. J. Iliff, C. Lagoze, F. A. La Sorte, M. Merrifield, W. Morris, T. B. Phillips, M. Reynolds, A. D. Rodewald, K. V. Rosenberg, N. M. Trautmann, A. Wiggins, D. W. Winkler, W-K. Wong, C. L. Wood, J. Yu, and S. Kelling. 2014. The eBird enterprise: an integrated approach to development and application of citizen science. *Biological Conservation* 169:31-40. <http://dx.doi.org/10.1016/j.biocon.2013.11.003>

Thogmartin, W. E. 2010. Sensitivity analysis of North American bird population estimates. *Ecological Modelling* 221:173-177. <http://dx.doi.org/10.1016/j.ecolmodel.2009.09.013>

Thogmartin, W. E., S. M. Crimmins, and J. Pearce. 2014. Prioritizing bird conservation actions in the Prairie Hardwood transition of the Midwestern United States. *Biological Conservation* 176:212-223. <http://dx.doi.org/10.1016/j.biocon.2014.06.002>

Thogmartin, W. E., F. P. Howe, F. C. James, D. H. Johnson, E. T. Reed, J. R. Sauer, and F. R. Thompson. 2006. A review of the population estimation approach of the North American landbird conservation plan. *Auk* 123:892-905. [http://dx.doi.org/10.1642/0004-8038\(2006\)123\[892:AROTPE\]2.0.CO;2](http://dx.doi.org/10.1642/0004-8038(2006)123[892:AROTPE]2.0.CO;2)

Thompson, S. J., C. M. Handel, and L. B. McNew. 2017. Autonomous acoustic recorders reveal complex patterns in avian detection probability. *Journal of Wildlife Management* 81:1228-1241. <http://dx.doi.org/10.1002/jwmg.21285>

Twedt, D. J. 2015. Estimating regional landbird populations from enhanced North American breeding bird surveys. *Journal of Field Ornithology* 86:352-368. <http://dx.doi.org/10.1111/jfo.12118>

Wilson, H. B., B. E. Kendall, and H. P. Possingham. 2011. Variability in population abundance and the classification of extinction risk. *Conservation Biology* 25:747-757. <http://dx.doi.org/10.1111/j.1523-1739.2011.01671.x>

Woiderski, B. J., N. E. Drilling, J. M. Timmer, M. F. McLaren, C. M. White, N. J. Van Lanen, D. C. Pavlacky Jr., and R. A. Sparks. 2018. *Integrated monitoring in bird conservation regions (IMBCR): 2017 field season report*. Bird Conservancy of the Rockies, Brighton, Colorado, USA.



# APPENDIX 1

## Additional Details on Time-of-Day Adjustment

### Data source and filtering

We downloaded the raw BBS 50-stop data for each species for the period 1997-2014 (Pardieck et al. 2015). During the model testing and development phase we found that limiting the dataset to 500 BBS routes for widespread species (i.e. species that occurred on  $> 500$  routes) greatly increased computational time without significantly altering model results. For these species, we applied two data quality filters to preferentially include route-level data that was representative of species: 1) we excluded BBS routes where the mean number of counts (across years) was an extreme outlier (defined as  $3x$  the interquartile distance) relative to all other BBS routes, and 2) we preferentially included BBS routes if they were either run more than once with peak counts at different stop numbers in different years (unless the peak count occurred at the first stop), or had a large number of stops (defined as  $\geq 75\%$  percentile) where the species was observed. We applied these filters to limit variability due to differences in habitat suitability along the 50 stops of the survey being incorrectly attributed as changes in time-of-day detectability.

### Model

We fit generalized additive models (GAMs) to the raw BBS 50-stop data for each species to relate the time of day (i.e., stop number) to the number of birds observed at each stop while accounting for route-level effects and trends through time. We modeled the number of birds counted at each stop with a negative binomial distribution to allow for overdispersion of counts. The stop number acts as a proxy for the time-of-day since survey volunteers are instructed to begin the survey at a set time relative to dawn (30 minutes before local sunrise). The model finds a smoothed nonlinear function to relate stop number to the number of birds counted using a penalized cubic regression spline. The model included a random

effect for each BBS route and a linear fixed effect for survey year. All models were fit in the R language and environment (R Core Team 2017), using the package ‘mgcv’ (Wood 2011).

The ‘mgcv’ package has a GAM function that is optimized for efficiently fitting models to large datasets (bam; Wood et al. 2015). However, for the negative binomial family, the ‘bam’ function cannot handle fitting the *theta* ( $\theta$ ) parameter, and a fixed value is required. The parameter  $\theta$ , in a negative binomial distribution, specifies the variance such that  $var(y) = \mu + \mu^2/\theta$ , where  $\mu = E(y)$ . To take advantage of this optimized GAM function, we fit the models using a two-step process. First, to estimate  $\theta$  we specified a model using a subset of the data and the ‘gam’ function. For this step, we randomly selected 50 BBS routes from the set of suitable routes (see Data source and filtering above), fit a GAM model as described above, and recorded the  $\theta$  parameter. We repeated this 10 times for each random draw of 50 BBS routes for each species. To make sure the estimation of  $\theta$  was not overly influenced by the number of routes included in the random draw, we tested models with random draws 40, 60, 80, or 100 BBS routes on a subset of species. No pronounced pattern emerged with  $\theta$  changing as a function of the number of routes included in the set. Once the first step was complete, we fit the full model using the mean  $\theta$  value from the previous step as a fixed term in the negative binomial distribution using the ‘bam’ function.

### **Time-of-day adjustment**

We then used the fitted time-of-day curves to calculate a time-of-day adjustment distribution for each species by resampling the fitted curve at each stop and re-calculating the maximum count divided by the mean count for each iteration. We parameterized the time-of-day distributions based on 1000 iterations. For species recorded on < 100 BBS routes, we calculated an average weighted by number of routes from all species in the same taxonomic (genus or family) and/or temporal (diurnal/nocturnal) group with > 100 routes (Table A1). Previous PIF population size estimates used averaging for species with > 50 routes

(Blancher et al. 2007). The decision to increase this to 100 routes was based on an examination of the coefficient of variation of the time of day adjustment statistic as a function of the number of routes.

## Literature Cited

Blancher, P. J., K. V. Rosenberg, A. O. Panjabi, B. Altman, J. Bart, C. J. Beardmore, G. S. Butcher, D. D.

Demarest, R. Dettmers, E. H. Dunn, W. E. Easton, W. C. Hunter, E. E. Iñigo-Elias, D. N. Pashley, C. J.

Ralph, T. D. Rich, C. M. Rustay, J. M. Ruth, and T. C. Will. 2007. *Guide to the Partners in Flight*

*Population Estimates Database. North American Landbird Conservation Plan 2004. Partners in Flight Technical Series No 5.*

Pardieck, K. L., D. J. Ziolkowski, M.-A. R. Hudson, and K. Campbell. 2016. *North American Breeding*

*Bird Survey Dataset 1966 - 2015, version 2015.1.* U.S. Geological Survey, Patuxent Wildlife Research

Center, Laurel, MD, USA. <https://www.pwrc.usgs.gov/BBS/RawData>

R Core Team. 2017. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.

Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)* 73:3–36.

Wood, S. N., Y. Goude, and S. Shaw. 2015. Generalized additive models for large data sets. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 64:139–155.

**Table A1.** Species present on < 100 BBS routes for which a weighted mean by number of BBS routes from the same taxonomic (genus or family) and/or temporal (diurnal/nocturnal) group with > 100 routes was used to estimate the time-of-day adjustment.

<b>Genus mean</b>	
Greater Prairie-Chicken ( <i>Tympanuchus cupido</i> )	Long-billed Thrasher ( <i>Toxostoma longirostre</i> )
Allen's Hummingbird ( <i>Selasphorus sasin</i> )	Bendire's Thrasher ( <i>Toxostoma bendirei</i> )
Rough-legged Hawk ( <i>Buteo lagopus</i> )	California Thrasher ( <i>Toxostoma redivivum</i> )
Western Screech-Owl <sup>a</sup> ( <i>Megascops kennicottii</i> )	Le Conte's Thrasher ( <i>Toxostoma lecontei</i> )
Lewis's Woodpecker ( <i>Melanerpes lewis</i> )	Crissal Thrasher ( <i>Toxostoma crissale</i> )
Gila Woodpecker ( <i>Melanerpes uropygialis</i> )	Bohemian Waxwing ( <i>Bombycilla garrulus</i> )
Golden-fronted Woodpecker ( <i>Melanerpes aurifrons</i> )	American Pipit ( <i>Anthus rubescens</i> )
Nuttall's Woodpecker ( <i>Picoides nuttallii</i> )	Hoary Redpoll ( <i>Acanthis hornemanni</i> )
White-headed Woodpecker ( <i>Picoides albolarvatus</i> )	Lawrence's Goldfinch ( <i>Spinus lawrencei</i> )
Gilded Flicker ( <i>Colaptes chrysoides</i> )	Lapland Longspur ( <i>Calcarius lapponicus</i> )
Gyr Falcon ( <i>Falco rusticolus</i> )	Abert's Towhee ( <i>Melospiza aberti</i> )
Gray Vireo ( <i>Vireo vicinior</i> )	Black-chinned Sparrow ( <i>Spizella atrogularis</i> )
Black-whiskered Vireo ( <i>Vireo altiloquus</i> )	American Tree Sparrow <sup>b</sup> ( <i>Spizelloides arborea</i> )
Mexican Jay ( <i>Aphelocoma wollweberi</i> )	Bell's Sparrow ( <i>Artemisiospiza belli</i> )
Northwestern Crow ( <i>Corvus caurinus</i> )	Seaside Sparrow ( <i>Ammodramus maritimus</i> )
Cave Swallow ( <i>Petrochelidon fulva</i> )	Hooded Oriole ( <i>Icterus cucullatus</i> )
Bridled Titmouse ( <i>Baeolophus wollweberi</i> )	Bronzed Cowbird ( <i>Molothrus aeneus</i> )
Black-crested Titmouse ( <i>Baeolophus atricristatus</i> )	Lucy's Warbler ( <i>Oreothlypis luciae</i> )
Black-tailed Gnatcatcher ( <i>Poliophtila melanura</i> )	Grace's Warbler ( <i>Setophaga graciae</i> )
Gray-cheeked Thrush ( <i>Catharus minimus</i> )	Red-faced Warbler ( <i>Cardellina rubrifrons</i> )
<b>Family mean</b>	
Chukar ( <i>Alectoris chukar</i> )	Northern Hawk Owl <sup>c</sup> ( <i>Surnia ulula</i> )
Willow Ptarmigan ( <i>Lagopus lagopus</i> )	Elf Owl <sup>a</sup> ( <i>Micrathene whitneyi</i> )
Rock Ptarmigan ( <i>Lagopus muta</i> )	Great Gray Owl <sup>c</sup> ( <i>Strix nebulosa</i> )
Mexican Whip-poor-will <sup>a</sup> ( <i>Antrostomus arizonae</i> )	Long-eared Owl <sup>a</sup> ( <i>Asio otus</i> )
Black Swift ( <i>Cypseloides niger</i> )	Flammulated Owl <sup>a</sup> ( <i>Deltarhynchus flammulatus</i> )
Costa's Hummingbird ( <i>Calypte costae</i> )	McCown's Longspur ( <i>Rhynchophanes mccownii</i> )
Harris's Hawk ( <i>Parabuteo unicinctus</i> )	
<b>All Diurnal</b>	
Arctic Warbler ( <i>Phylloscopus borealis</i> )	
Northern Wheatear ( <i>Oenanthe oenanthe</i> )	

<sup>a</sup> Summarizing group is further subset to only nocturnal species.

<sup>b</sup> Summarized by genus *Spizella*.

<sup>c</sup> Summarizing group is further subset to only diurnal species.

---

**Appendix 2.** R scripts and associated parameter files

*Please click here to download file 'appendix2.zip'.*

---