

## **Evaluating time-removal models for estimating availability of boreal birds during point count surveys: Sample size requirements and model complexity**

Authors: Sólymos, Péter, Matsuoka, Steven M., Cumming, Steven G., Stralberg, Diana, Fontaine, Patricia, et al.

Source: The Condor, 120(4) : 765-786

Published By: American Ornithological Society

URL: <https://doi.org/10.1650/CONDOR-18-32.1>

---

BioOne Complete ([complete.BioOne.org](https://complete.BioOne.org)) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at [www.bioone.org/terms-of-use](https://www.bioone.org/terms-of-use).

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

---

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.



RESEARCH ARTICLE

# Evaluating time-removal models for estimating availability of boreal birds during point count surveys: Sample size requirements and model complexity

Péter Sólymos,<sup>1,3\*</sup> Steven M. Matsuoka,<sup>2,3</sup> Steven G. Cumming,<sup>4</sup> Diana Stralberg,<sup>3</sup> Patricia Fontaine,<sup>3</sup> Fiona K. A. Schmiegelow,<sup>5</sup> Samantha J. Song,<sup>6</sup> and Erin M. Bayne<sup>1</sup>

<sup>1</sup> University of Alberta, Department of Biological Sciences, Edmonton, Alberta, Canada

<sup>2</sup> U.S. Geological Survey, Alaska Science Center, Anchorage, Alaska, USA

<sup>3</sup> Boreal Avian Modelling Project, University of Alberta, Edmonton, Alberta, Canada

<sup>4</sup> Université Laval, Département des sciences du bois et de la forêt, Québec City, Québec, Canada

<sup>5</sup> Northern Environmental and Conservation Sciences Program, Department of Renewable Resources, University of Alberta, c/o Yukon College, Whitehorse, Yukon, Canada

<sup>6</sup> Environment and Climate Change Canada, Canadian Wildlife Service, Edmonton, Alberta, Canada

\* Corresponding author: [solyms@ualberta.ca](mailto:solyms@ualberta.ca)

Submitted February 12, 2018; Accepted June 19, 2018; Published August 29, 2018

## ABSTRACT

We used conventional and finite mixture removal models with and without time-varying covariates to evaluate availability given presence for 152 bird species using data from point counts in boreal North America. We found that the choice of model had an impact on the estimability of unknown model parameters and affected the bias and variance of corrected counts. Finite mixture models provided better fit than conventional removal models and better adjusted for count duration. However, reliably estimating parameters and minimizing variance using mixture models required at least 200–1,000 detections. Mixture models with time-varying proportions of infrequent singers were best supported across species, indicating that accounting for date- and time-related heterogeneity is important when combining data across studies over large spatial scales, multiple sampling time frames, or variable survey protocols. Our flexible and continuous time-removal modeling framework can be used to account for such heterogeneity through the incorporation of easily obtainable covariates, such as methods, date, time, and location. Accounting for availability bias in bird surveys allows for better integration of disparate studies at large spatial scales and better adjustment of local, regional, and continental population size estimates.

**Keywords:** boreal birds, population size, abundance, detectability, bias–variance tradeoff

## Évaluation des modèles d'élimination du temps pour estimer la disponibilité des oiseaux boréaux lors des inventaires par points d'écoute : exigences relatives à la taille de l'échantillon et complexité du modèle

## RÉSUMÉ

Nous avons évalué des modèles conventionnels et de retrait des mélanges finis avec et sans covariables variant dans le temps pour 152 espèces d'oiseaux, en utilisant des données provenant de points d'écoute en Amérique du Nord boréale. Nous avons trouvé que le choix des modèles avait un impact sur l'estimabilité des paramètres inconnus du modèle et qu'il affectait le biais et la variance dans les décomptes corrigés. Les modèles de mélanges finis offraient un meilleur ajustement que les modèles de retrait conventionnels et ils étaient mieux ajustés pour tenir compte de la durée du décompte. Toutefois, l'estimation fiable des paramètres et la minimisation de la variance pour les modèles de mélanges demandaient au moins 200–1000 détections. Les modèles de mélanges avec des proportions variant dans le temps des chanteurs non fréquents s'appliquaient le mieux pour toutes les espèces, ce qui indique que le fait de tenir compte de l'hétérogénéité reliée à la date et au temps est important lorsque l'on combine des données provenant de plusieurs études sur de grandes échelles spatiales, des périodes d'échantillonnage multiples ou des protocoles d'inventaire variables. Notre cadre de modélisation flexible et éliminant continuellement le temps peut être utilisé pour tenir compte d'une telle hétérogénéité en incorporant des covariables facilement disponibles, telles que les méthodes, la date, le temps et l'emplacement. La prise en compte du biais de disponibilité dans les inventaires d'oiseaux permet de mieux intégrer les études disparates à des échelles spatiales plus larges et un meilleur ajustement des estimations de population à l'échelle locale, régionale et continentale.

**Mots-clés :** oiseaux boréaux, taille de la population, abondance, détectabilité, compromis biais-variance

## INTRODUCTION

It has long been recognized that nearly all avian field surveys underestimate abundances (Leopold 1933, Ken-deigh 1944), unless the estimates are adjusted for the proportion of birds present but undetected at the times and locations surveyed (Hayne 1949, Anderson and Pospahala 1970). Such adjustments require an estimate of the probability of detecting birds present during surveys (detectability). Detectability is the product of the probability that birds make themselves available for detection by emitting detectable cues (availability) and the probability that the available birds will be perceived by a bird surveyor (perceptibility; Diefenbach et al. 2007).

All abundance estimators developed to account for imperfect detectability impose conditions on survey design, requiring either some form of replication, for example, of visits or observers, or ancillary data collection, for example, records of distance or time for each observation (Nichols et al. 2009, Matsuoka et al. 2014, but see Sólymos and Lele 2016). The added costs and expertise necessary to collect these ancillary data have prevented most of the standardized continental surveys, such as the North American Breeding Bird Survey (BBS), from adopting these methods (O'Connor et al. 2000). Some researchers (e.g., Thogmartin et al. 2006, 2010, Matsuoka et al. 2012, Twedt 2015) contend that this limits the rigor of spatial models of avian abundance and population size at large geographic scales, reducing the utility of these models for setting and evaluating progress toward meeting population objectives (e.g., Partners in Flight; Rosenberg and Blancher 2005). Unstandardized counts may also lead to improper inferences about the influence of covariates, thus potentially biasing the identification of suitable habitats for sensitive species (Sólymos et al. 2013).

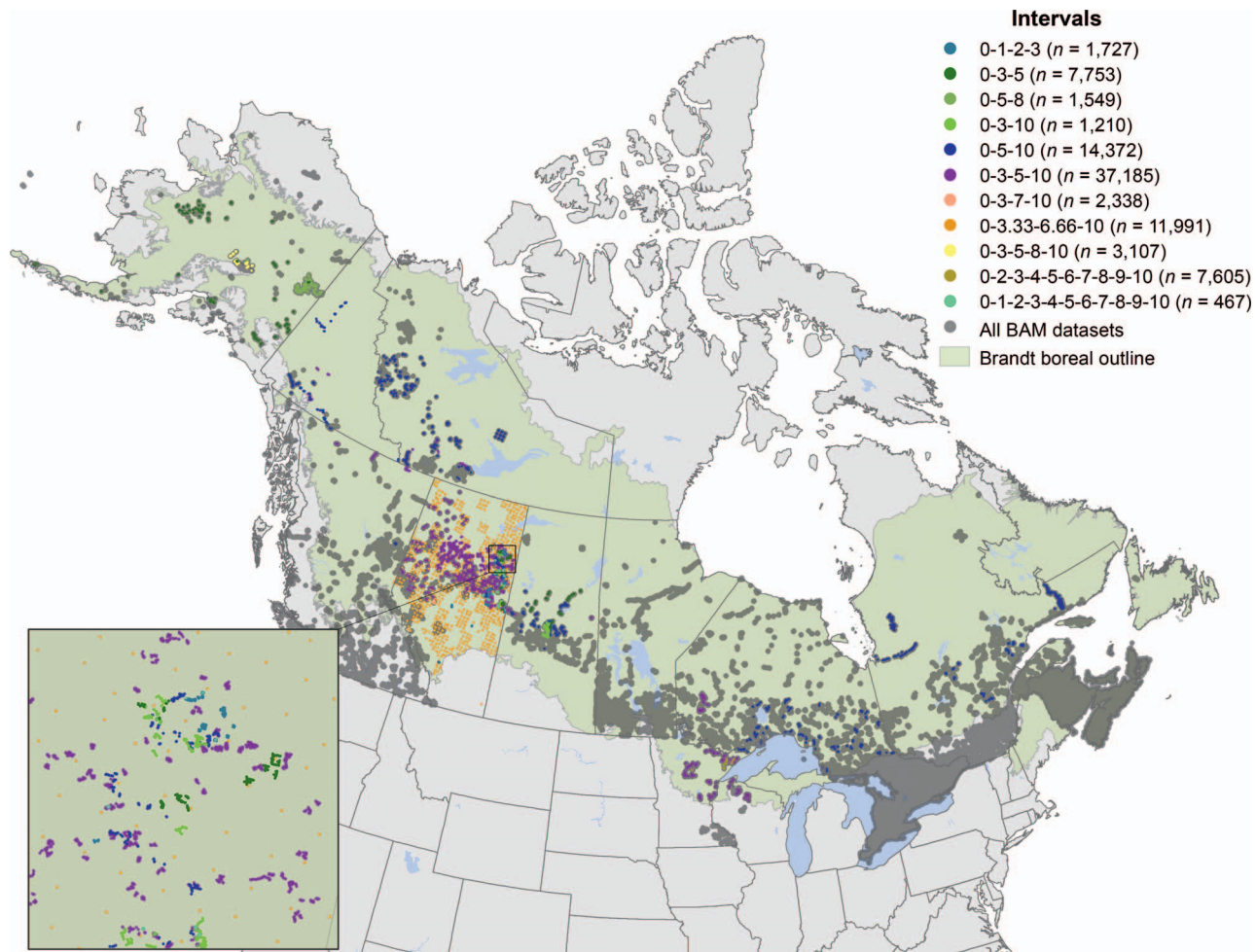
Detection probabilities are strongly influenced by environmental conditions, bird behaviors, and survey protocols (Johnson 2008, Matsuoka et al. 2014), and are expected to show greater variation between studies collating datasets than within individual studies. For point count surveys, duration varies greatly among studies (Barker et al. 2015), but rarely changes within studies. As more and more large-scale studies rely on compiled point count data from disparate local studies (e.g., Bart 2005, Kelling et al. 2009, Dickinson et al. 2010, Hampton et al. 2013), standardization becomes critical because count duration greatly affects observations. The number of individual birds counted over a 10-min period is roughly 60% higher than during a 3-min period (Farnsworth et al. 2002, Etterson et al. 2009, Matsuoka et al. 2014), and both durations are routinely used for avian surveys.

The time-removal model (Moran 1951, Zippin 1956, 1958) was originally developed for estimating wildlife and

fish abundances from capture–mark–recapture studies (Seber 1982). It was later reformulated for avian surveys with the goal of improving estimates of bird abundance, population size, and population trend by accounting for the availability bias inherent in point count data (Farnsworth et al. 2002, Sólymos et al. 2013). The model has now been widely applied to point count studies because the required ancillary data, that is, tallies of the first detections of individual birds in sequential time intervals, are easy to collect (Johnson 2008, Matsuoka et al. 2014). The removal model applied to point count surveys estimates the probability that a bird is available for detection as a function of the average number of detectable cues that an individual bird gives per min (singing rate), and the known count duration (Seber 1982, Barker and Sauer 1995, Alldredge et al. 2007a, Sólymos et al. 2013). However, the removal model has not been rigorously assessed to determine whether availability bias varies with duration, and whether fitting more complex parameterizations of the removal model can minimize such bias.

Singing rates of birds vary with time of day, time of year, breeding status, and stage of the nesting cycle (Wilson and Bart 1985, Rosenberg and Blancher 2005, Stacier et al. 2006). Thus, removal model estimates of availability may be improved by accounting for variation in singing rates using covariates for time of day and day of year (Reidy et al. 2011, Sólymos et al. 2013, Amundson et al. 2014). The removal model can also accommodate behavioral heterogeneity in singing by subdividing the sampled population of a species at a given point into a finite mixture of birds with low and high singing rates, which requires the additional estimation of the proportion of birds in the sampled population with low singing rates (Farnsworth et al. 2002, Alldredge et al. 2007a, 2007b, Reidy et al. 2011).

A key concern when fitting removal models of increasing complexity is the tradeoff between bias and variance, and the moderating role of sample size in this tradeoff. Several studies have recommended finite mixture models over conventional removal models to reduce bias in abundance estimates (Farnsworth et al. 2002, Etterson et al. 2009, Reidy et al. 2011), but these studies analyzed large sample sizes to support the more complex models. In contrast, efforts to fit finite mixture models to smaller datasets have resulted in models with unidentified parameters or point estimates of availability that were similar, but far less precise, than those from the conventional removal model (Amundson et al. 2014). The ability to fit more complex removal models that both reduce bias and minimize variance in estimates of availability may be, in part, a function of the number of detections, but the sample size requirements and bias–variance tradeoffs have not been systematically assessed for removal models.



**FIGURE 1.** Locations of point count surveys from 1991 to 2014 compiled across boreal and hemiboreal regions (according to Brandt 2009) of North America by the Boreal Avian Modelling Project (BAM). We analyzed the point counts for which surveyors recorded when each bird was first detected relative to multiple time intervals (colored dots). Gray dots represent point counts with a single time interval. Numbers of point count surveys ( $n$ ) in each of 11 different combinations of the count duration (min) and the number and length of the time intervals are given in the legend. The inset shows an expanded view of a region in northeastern Alberta, Canada, where many projects with different counting methodologies overlapped.

We formulated and evaluated a continuous time-removal model for avian point counts by estimating the availability probabilities of 152 terrestrial bird species using a large database of point count surveys conducted across boreal and hemiboreal regions of North America (Figure 1, Appendix Tables 2 and 3; Cumming et al. 2010, Barker et al. 2015). We compared different parameterizations of removal models to determine the relationships among model complexity, bias, variance, sample size, and duration. To address general estimation of availability, we fit all models to 152 species that met minimum sampling requirements, but varied widely in singing rates and sample sizes of detections. To address protocol variability, we compared the reliability of the different model classes for adjusting the raw survey counts for variation in count duration. We also examined what factors influenced bias

and variance in the adjustments. Finally, we compared our removal model adjustments of availability to those applied to BBS data across North America (Blancher et al. 2013) to evaluate the extent to which removal models may improve continental estimates of landbird population sizes. We finish by summarizing our recommendations for selecting among the various removal model types given study design and data characteristics.

## METHODS

### Point Count Data

The Boreal Avian Modelling Project (BAM) was initiated to facilitate the conservation of boreal birds. As part of these efforts, BAM has collated data from georeferenced point count surveys conducted in boreal, hemiboreal, and



**TABLE 1.** Relative support for different model classes and candidate models of temporal covariate effects to explain the probability of availability during point count surveys of terrestrial landbird species in northern North America. We fit each of the models to 147 species and summarized relative support for each model as the number of species for which the model had the lowest Akaike's information criterion corrected for small sample sizes (AIC<sub>c</sub>). Model classes included:  $M_0$  = time-invariant conventional;  $M_f$  = time-invariant finite mixture;  $M_0^\varphi$  = conventional with time-varying rate parameter;  $M_f^\varphi$  = finite mixture with time-varying rate parameter; and  $M_f^c$  = finite mixture with time-varying proportion of infrequent singers. The total is the sum of model types  $M_0^\varphi$ ,  $M_f^\varphi$ , and  $M_f^c$  by covariate model ID; frequencies for winter resident species are also shown (frequency for migrant species = Total – Residents).

Model ID and covariate effects <sup>b</sup>	$M_0^\varphi$	$M_f^\varphi$	$M_f^c$	Total	Residents
0: Null (time-invariant: $M_0$ and $M_f$ )	1	0	0	1	0
1: DY	1	6	2	9	2
2: SR	1	5	4	10	2
3: DY + DY <sup>2</sup>	2	3	4	9	0
4: SR + SR <sup>2</sup>	0	7	17	24	4
5: DY + SR	0	5	3	8	1
6: DY + DY <sup>2</sup> + SR	2	0	1	3	0
7: DY + SR + SR <sup>2</sup>	2	2	19	23	3
8: DY + DY <sup>2</sup> + SR + SR <sup>2</sup>	1	1	4	6	1
9: LS	0	3	3	6	0
10: LS + LS <sup>2</sup>	0	2	10	12	4
11: LS + SR	1	0	6	7	0
12: LS + LS <sup>2</sup> + SR	0	1	4	5	0
13: LS + SR + SR <sup>2</sup>	1	3	18	22	6
14: LS + LS <sup>2</sup> + SR + SR <sup>2</sup>	0	0	2	2	0
Total	12	38	97	147	23

<sup>b</sup> Covariate effects included linear or quadratic effects of time since sunrise (SR), ordinal day of the year (DY), and the ordinal day relative to local spring (LS).

sub-Arctic biomes by independent inventory, monitoring, and research studies (Cumming et al. 2010, Barker et al. 2015). The BAM database (version 5) contains 250,822 off-road point count surveys conducted at 145,289 point count locations by 131 separate studies between 1991 and 2014. Of these surveys, 3% were 3 min, 60% were 5 min, and 37% were 10 min in total point count duration. We applied removal models to a subset of 89,304 off-road point counts at 25,754 locations conducted by 71 projects in boreal and hemiboreal regions (hereafter, 'boreal') in which observers recorded when they first detected each bird relative to  $\geq 2$  time intervals (hereafter referred to as the 'full dataset'; Figure 1). We excluded roadside surveys because of possible differences in singing behavior between birds breeding along roadsides vs. in off-road areas (Kociolek et al. 2011). The number of time intervals varied from 2 to 10, and there were 11 distinct combinations of count duration and the number and length of the time intervals (Figure 1). We restricted our analyses to

detections of vocalizing gallinaceous gamebirds, shorebirds, woodpeckers, and songbirds, which included species that are diurnally active and are first detected primarily by their auditory displays during breeding surveys. We analyzed species detected aurally during a minimum of 75 point count surveys in the full dataset (median = 1,240 point count surveys, maximum = 31,334). We present results for the 152 species for which we could successfully estimate the constant singing rate parameter of the conventional removal model (Appendix Tables 2 and 3).

Count duration varied from 3 min to 20 min, but we omitted time intervals beyond 10 min to minimize the inclusion of movements by birds during counts, which could violate the closed population assumption of the removal model. For example, individuals that move toward the observer are more likely to be detected than ones that move away, and would inflate counts (Glennie et al. 2015). Point count radii in our dataset were at least 50 m, with 100-m radius (24%) and unlimited distance (72%) counts making up the majority. The removal model captures availability, while perceptibility is often estimated via distance sampling (Diefenbach et al. 2007). The conditional model formulation uses the cumulative counts over the time intervals relative to the total count, which makes the estimation of availability independent of the observer and distance-related perceptibility component and of the true underlying population abundance. By conditioning on the total counts, the abundance and other nuisance variables (such as perceptibility) are canceled out of the model likelihood (Sólymos et al. 2013).

As a result of this conditional modeling approach, we treated within- and between-year repeat visits to the same locations as independent observations. We checked that our assumption of conditional independence was valid, and we found neither systematic bias in our parameter estimates nor shrinking of standard errors as a result of treating data from revisits as independent (Appendix Figure 6). Using data from within- and between-year revisits to the same locations allowed us to better estimate date- and time-related patterns in our dataset, because revisits increased our sampling coverage of different dates and times, and therefore strengthened our inference regarding longitudinal variation. The diurnal and seasonal timing of the point counts varied widely, so we standardized the timing of each point count relative to local sunrise and spring, respectively, using the National Oceanic and Atmospheric Administration's Sunrise/Sunset Calculator (<https://www.esrl.noaa.gov/gmd/grad/solcalc/sunrise.html>, Meeus 1999), implemented in the R package *maptools* (Bivand and Lewin-Koh 2016). We quantified the ordinal day of each point count, and also a standardized version of the ordinal day of each point count relative to the 30-yr average of the local start of the growing season (McKenney et al. 2006). Temporal covariates were standardized by

dividing the observed values by the maximum possible values: time since local sunrise (SR; mean = 2.2 hr, range = -5.7 to 12.5 hr, maximum possible = 24 hr); ordinal day of year (DY; mean = 166, range = 128 to 200, maximum possible = 365); and days since local spring (LS; mean = 48.5, range = -8 to 94, maximum possible = 365).

### Continuous Time-removal Models

Time-removal models are based on a removal experiment wherein animals are trapped and thereby removed from the closed population of animals being sampled (Zippin 1958). When applying removal models to avian point count surveys, the counts of singing birds ( $Y_{ij}, \dots, Y_{ij}$ ) within a given point count  $i$  ( $i = 1, \dots, n$ ) are tallied relative to when each bird is first detected in multiple and consecutive time intervals, with the survey start time  $t_{i0} = 0$ , the end times of the time intervals  $t_{ij}$  ( $j = 1, 2, \dots, J$ ), and the total count duration of the survey  $t_{ij}$ . Each individual bird is counted once, so that individuals are 'mentally removed' from a closed population of undetected birds by the surveyor. In the continuous time-removal model (hereafter, 'removal model'), singing events by individual birds are assumed to follow a Poisson process. We can use the rate parameter of the Poisson process ( $\phi$ ) to estimate the singing rates of birds during a point count (Allredge et al. 2007a).

In the conventional removal model ( $M_0$ ), all individuals of a single species at a given location and time are assumed to be homogeneous in their singing rates. The time to first detection follows the exponential distribution  $f(t_{ij}) = \phi \exp(-t_{ij} \phi)$ , and the cumulative density function of times to first detection in time interval  $(0, t_{ij})$  gives us the probability that a bird sings at least once during the point count as  $p(t_{ij}) = 1 - \exp(-t_{ij} \phi)$  (Barker and Sauer 1995, Allredge et al. 2007a, 2007b, Sólymos et al. 2013). Although the term 'detection' implies both the perceptibility and detectability processes, the perceptibility component is not part of our models due to the conditional likelihood approach. The perceptibility component should be addressed in a distance sampling framework independent of the removal models (Handel et al. 2009, Sólymos et al. 2013).

In the continuous time formulation of the finite mixture (or 2-point mixture) removal model, the cumulative density function during a point count is given by  $p(t_{ij}) = (1 - c) 1 + c [1 - \exp(-t_{ij} \phi)] = 1 - c \exp(-t_{ij} \phi)$ , where  $\phi$  is the singing rate for the group of infrequently singing birds, and  $c$  is the proportion of birds during the point count that are infrequent singers (see Jewell 1982 for a discussion of mixtures of exponential distributions). The remaining proportions  $(1 - c; \text{the intercept of the cumulative density function})$  of the frequent singers are assumed to be detected instantaneously at the start of the first time interval (Allredge et al. 2007b). In the simplest form of

the finite mixture model, the proportion and singing rate of birds that sing infrequently is homogeneous across all times and locations (model  $M_f$ ).

The singing rate in the conventional removal model ( $\phi_i$ ) may also vary with respect to survey-specific covariates. We hereafter refer to these models collectively as time-varying conventional removal models ( $M_0^\phi$ ). Previously, researchers have applied covariate effects to the parameter  $\phi_i$  of the finite mixture model, assuming that the parameter  $c$  is constant irrespective of time and location (i.e. only the infrequent singer group changes its singing behavior; Farnsworth et al. 2002, Aldredge et al. 2007a, Etterson et al. 2009, Reidy et al. 2011). An alternative parameterization is for  $c_i$  rather than  $\phi$  to be the time-varying parameter, allowing individuals to switch between the frequent and infrequent singer group depending on covariates. We defined 2 time-varying finite mixture model classes: (1)  $M_f^\phi$ , with covariate dependent rate parameter  $p(t_{ij}) = 1 - c \exp(-t_{ij} \phi_i)$ ; and (2)  $M_f^c$ , with the proportion of infrequent singers dependent on covariates:  $p(t_{ij}) = 1 - c_i \exp(-t_{ij} \phi)$ .

In all 3 classes of time-varying models ( $M_0^\phi$ ,  $M_f^\phi$ , and  $M_f^c$ ), we modeled covariate effects as linear or quadratic effects of time since local sunrise (SR), ordinal day of year (DY), and days since local spring (LS) in 14 different candidate models (Table 1). We estimated model parameters using the conditional maximum-likelihood procedure described by Sólymos et al. (2013) and implemented in the R package detect (R Core Team 2017, Sólymos et al. 2018). This included using the `cmulti` function, with type argument 'rem' for conventional ( $M_0$ ,  $M_0^\phi$ ) models and type argument 'fmix' and 'mix' for the 2 mixture model types ( $M_f^\phi$  and  $M_f^c$ , respectively; both options give identical results for  $M_f$ ). Covariate effects for  $\phi_i$  were modeled using the logarithmic link function to ensure that  $\phi_i > 0$ , whereas covariate effects for  $c_i$  were modeled using the logistic link function to ensure that  $c_i$  was in the interval  $(0, 1)$ .

### Sample Size Requirements for Removal Models

We first examined the minimum sample size required to reliably estimate the parameters of the time-invariant  $M_0$  and  $M_f$  models. We stratified the full point count dataset by project, and then estimated the model parameters for 3,002 combinations of the 65 projects and 152 species that had a minimum of 2 nonzero total counts (hereafter referred to as the 'project-level dataset'). We classified an estimation as a success if the Nelder-Mead optimization algorithm for the conditional maximum-likelihood estimator (MLE) was able to converge on a point estimate of  $\phi$  for  $M_0$  or on a joint estimate of  $(\phi, c)$  for  $M_f$  and if the algorithm could estimate the corresponding variances,  $\text{Var}[\log(\phi)]$  for  $M_0$ , and jointly  $\text{Var}[\log(\phi)]$  and  $\text{Var}[\text{logit}(c)]$  for  $M_f$ . We fit logistic regression models separately

to the outcomes of estimates and variances for  $M_0$  and  $M_f$  models to predict the probability of successfully estimating the model parameters as a function of log number of point counts with nonzero totals for the species in a given project, and calibrated the fitted models to predict the minimum number of nonzero counts required to have a 90% probability of successfully estimating removal model parameters and variances. The number of nonzero counts was highly correlated with the mean count (Spearman's  $r = 0.82$ ,  $P < 0.001$ ) in the full dataset; therefore, we did not separately assess relationships with mean counts.

### Adjusting Point Counts for Variation in Count Duration

We used a subset of the full dataset for 133 species from 38,540 point count surveys with 0–3, 3–5, and 5–10 min intervals (hereafter referred to as the ‘validation dataset’) to assess the bias and variance based on the time-invariant models ( $M_0$  and  $M_f$  estimates based on the validation data) when predicting the number of birds counted in the full 10-min point count, using the data from the 0–3 min and 0–5 min portions of the same point counts as inputs. The expected total number of individuals of a species counted across point counts of duration  $t$  can be written as  $E[Y_t] = N p_t$ , where  $p_t$  is the estimated availability probability for duration  $t$  from the removal model ( $M_0$  or  $M_f$ ), and  $N$  is the unobserved total number of individuals of that species present across the point counts. It follows that  $N = E[Y_3]/p_3 = E[Y_5]/p_5 = E[Y_{10}]/p_{10}$  for intervals of the same point count, in which case  $N$  is assumed to be constant due to the closed population. We computed the bias of the removal model when adjusting the 3-min and 5-min point counts to a 10-min standard, respectively, as  $(\Sigma Y_3/p_3) / (\Sigma Y_{10}/p_{10})$  and  $(\Sigma Y_5/p_5) / (\Sigma Y_{10}/p_{10})$ , with a value of 1 resulting when the adjustment was unbiased (hereafter referred to as the ‘corrected relative count’). For each species, we fit  $M_0$  or  $M_f$  models and drew  $B = 1,000$  replications from the asymptotic distribution of the parameter estimates to represent uncertainty, and we calculated bias (mean signed difference from the expected value of 1) and variance (mean squared deviation of the corrected relative counts from their sample mean across the  $B$  replications) based on corrected relative counts from the  $B$  replicates. We graphically displayed the distribution of bias and variance across species as a function of sample size, count duration (3 or 5 min), and model type ( $M_0$  or  $M_f$ ). Bias and variance are presented as proportions of the 10-min corrected count. We used the median and 5% and 95% quantiles to represent the distribution of bias and the 5–95% interval to represent the distribution of variance of the species whose sample sizes were equal to or greater than the sample size selected for comparison.

### Relative Support for Models and Plotting Time-varying Effects

For each of the 152 species, we fit 44 candidate models that included the time-invariant ( $M_0$  and  $M_f$ ) and time-varying models with temporal covariate effects ( $M_0^o$ ,  $M_f^o$ , and  $M_f^c$ ; 14 models for each) based on the full dataset (Table 1, Appendix Table 3). Although our dataset was composed of many point counts, the sample size for conditional likelihood estimation is defined by the number of point counts with  $>0$  total counts. As a result, sample sizes varied considerably across species. Therefore, we compared the relative support for each model within a species using Akaike's information criterion corrected for small sample sizes ( $AIC_c$ ). We assessed model support across species by counting the number of species for which each model was best-supported (smallest  $AIC_c$  value). We only compared models for the 147 species for which all model parameters were estimated. We compiled information on each species' migratory strategy (winter resident vs. migrant) from Poole (2012), and used a chi-square test to see whether migratory strategy affected how often ordinal day (DY) and days since local spring (LS) were selected in the best-fitting models (Appendix Table 3).

We examined temporal patterns in availability probabilities across the species for which the best-fitting removal model included temporal covariates (SR, DY, or LS; Table 1, Appendix Table 3). We did so by predicting availability probabilities based on a 3-min point count duration for each species as a function of the observed covariate values from the full dataset (response curves). We conditioned the predictions on the mean of the other covariates. We overlaid the response curves of all species that shared the covariate effect, and described the general patterns in availability probabilities across species using a 2-dimensional kernel density estimate with a bivariate normal kernel (Wand and Jones 1995) using the R package KernSmooth (Wand 2015). We identified the contour for the 75% bivariate quantile to identify the highest density region, that is, where most of the species' response curves overlapped. We also identified 5%, 25%, 50%, 75%, and 95% contours based on the cumulative density conditional on the covariate values, that is, percentage of the species with the same or higher availability given the date and time. We show average responses for all 3 time-varying model classes ( $M_0^o$ ,  $M_f^o$ , and  $M_f^c$ ).

### Removal Models vs. PIF Approach to Adjusting for Incomplete Availability

We compared adjustments for incomplete availability of individual species based on removal models with those used in the Partners in Flight (PIF) Population Estimates Database (Blancher et al. 2013) to determine how PIF estimates of continental population sizes might be improved by using removal model-based estimates of



availability. PIF uses a time adjustment (here referred to as  $T_{adj}$ ; Blancher et al. 2013) to account for diurnal variation in availability, which is calculated as the ratio of the maximum count estimated from a polynomial regression to the mean count based on 50 stops in the BBS data (Rosenberg and Blancher 2005, Blancher et al. 2013), and assumes that the probability of availability = 1 at the time of the maximum count.

We defined our species adjustment factor as the reciprocal of average availability ( $U_{inv} = 1 / p_{avg}$ ). We calculated the correlation between  $T_{adj}$  and  $U_{inv}$  and compared the distributions across 129 species for which PIF adjustments were available (Blancher et al. 2013; Appendix Table 3). We also assessed how maximum availabilities ( $p_{max}$ ) compared with the PIF assumption of the maximum being 1. We calculated average ( $p_{avg}$ ) and maximum ( $p_{max}$ ) availability probabilities by using the best-supported model for each species (Table 1, Appendix Table 3) for predicting the probability of availability for each point count in the entire BAM dataset (Figure 1) given the covariate values, and then calculating the average and maximum of the predicted values for each individual species. We set count duration to be 3 min, to be consistent with the BBS methodology that  $T_{adj}$  was based on.

## RESULTS

### Sample Size Requirements for Time-invariant Removal Models

We were able to estimate the singing rate parameter ( $\phi$ ) for all 152 species using the conventional removal model ( $M_0$ ) fit to the full dataset from the boreal study area (Appendix Table 2). The mean probability of availability increased from  $0.51 \pm 0.15$  SD for 3-min point count duration to  $0.68 \pm 0.15$  SD for 5-min and  $0.88 \pm 0.12$  SD for 10-min count duration (Figure 2). A logistic regression model ( $\beta_0 = -1.16 \pm 0.10$  SE,  $\beta_{\log(\# \text{ nonzero counts})} = 0.92 \pm 0.04$  SE) predicted that a species needed to be detected during  $45 \pm 6$  SE point counts for the  $M_0$  model to have a 90% probability of successfully converging on a project-specific estimate of  $\phi$ . Sample size requirements for  $\text{Var}[\log(\phi)]$  were identical to that of the point estimates.

We jointly estimated the parameters for the  $M_f$  models ( $\phi$ ,  $c$ ) for 147 species using the full dataset. Mean probabilities of availability based on the  $M_f$  models across the 147 species were 9–16% lower than for the  $M_0$  models, and increased from  $0.46 \pm 0.17$  SD for 3-min, to  $0.56 \pm 0.17$  SD for 5-min, to  $0.74 \pm 0.17$  SD for 10-min count duration (Figure 2). A logistic regression model ( $\beta_0 = -1.16 \pm 0.09$  SE,  $\beta_{\log(\# \text{ nonzero counts})} = 0.66 \pm 0.03$  SE) predicted that a species needed to be detected during  $158 \pm 24$  SE point counts for the  $M_f$  model to have a 90% probability of successfully converging on project-level

estimates of  $\phi$  and  $c$ . The sample size requirement for a 90% probability of successfully estimating the associated variance was substantially larger,  $963 \pm 177$  SE ( $\beta_0 = -1.93 \pm 0.09$  SE,  $\beta_{\log(\# \text{ nonzero counts})} = 0.60 \pm 0.03$ ). The number of nonzero counts for the 5 species for which we were unable to estimate  $M_f$  model parameters varied between 88 and 419 (Appendix Table 2).

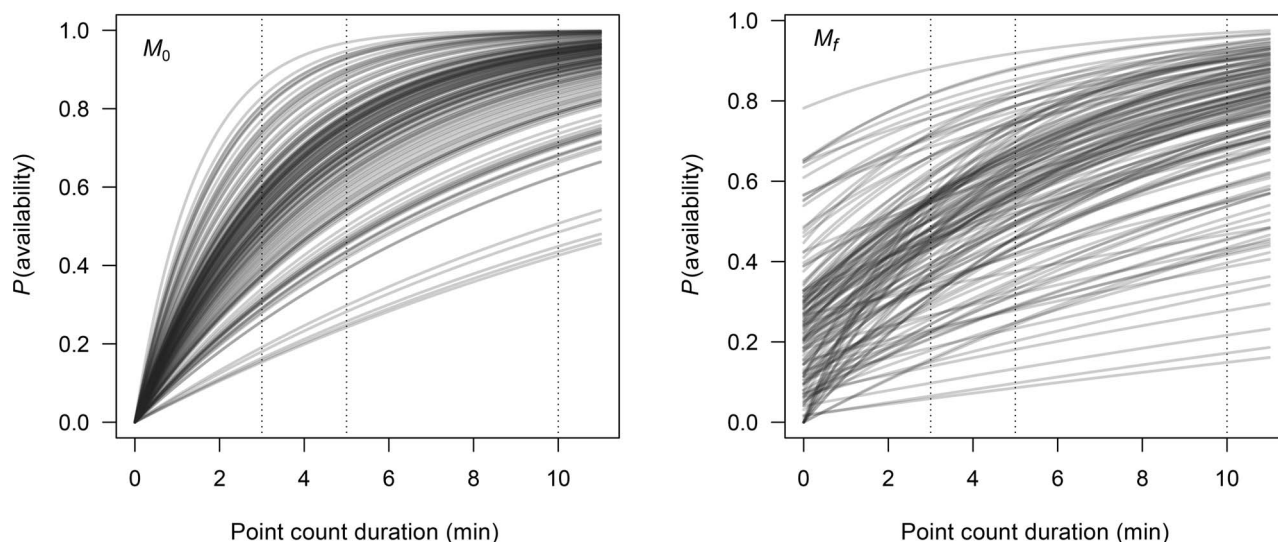
### Adjusting Point Counts for Variation in Count Duration

The bias when correcting 3- and 5-min duration counts to 10-min counts was affected by count duration. A 5-min count duration decreased median bias and tightened the 90% quantile range of bias closer to 0 (no bias) across species for both the  $M_0$  and  $M_f$  models, when compared with the distribution of 3-min based biases. The median bias with respect to 10-min counts for the  $M_0$  model was 0.05 for 3-min and  $-0.03$  for 5-min counts, and did not improve further with increasing sample size (Figure 3). Although the 90% quantile range for  $M_0$  models shortened with increasing sample size, it stayed centered around the biased median. Median bias with respect to 10-min counts for the  $M_f$  model was small (absolute value  $<0.01$ ) regardless of sample size; the 90% quantile range shortened with sample size. The high end of the 90% quantile range was close to 0 (no bias), whereas the lower limit spread wider into the negative direction, changing from  $-0.27$  at a sample size of 20 to  $-0.12$  at a sample size of 200 (Figure 3).

Variance was also greatly affected by count duration. Increasing the count duration from 3 to 5 min resulted in reducing the 90% limit for the  $M_f$  models from 0.06 to 0.03. Variance of the corrected relative counts under the  $M_0$  model was  $<0.01$  irrespective of duration and sample size, but nevertheless showing similar relative patterns. The 90% limit for variance of corrected relative counts under the  $M_f$  models decreased to 0.03 and 0.01 with 200 nonzero counts for the 3- and 5-min counts, respectively (Figure 3).

Probability of availability and corrected relative counts based on the time-invariant removal models ( $M_0$  and  $M_f$ ) varied among species with different sample sizes (Figure 4). The Rusty Blackbird (*Euphagus carolinus*), with 254 nonzero counts, showed high uncertainty in both probabilities and corrected relative counts, but bias with respect to 10-min counts was small for both models because the finite mixture model intercept (probability at time 0,  $1 - c$ ) was close to 0 ( $c = 0.96 \pm 0.18$  SE), and thus the singing rates were similar ( $M_0$ :  $\phi = 0.16 \pm 0.00$  SE;  $M_f$ :  $\phi = 0.14 \pm 0.05$  SE). The high uncertainty for the  $M_f$  model was a result of imprecise  $c$  estimates and high (0.6) correlation between  $c$  and  $\phi$  (high rate with low intercept ( $1 - c$ ), and vice versa). The Western Wood-Pewee (*Contopus sordidulus*) had 493 nonzero counts in the validation set, and, as a





**FIGURE 2.** Estimates of availability probabilities for landbird species in northern North America from conventional ( $M_0$ , 152 species) and finite mixture ( $M_f$ , 147 species) removal models relative to point count duration (min). Dotted vertical lines indicate common point count durations (3, 5, and 10 min).

result, its  $M_f$  estimates, especially  $c$ , were more precise, with low correlation ( $\phi = 0.11 \pm 0.05$  SE,  $c = 0.73 \pm 0.03$  SE, correlation = 0.15). As the intercept ( $1 - c$ ) was higher, the singing rates differed more between models ( $M_0$ :  $\phi = 0.27 \pm 0.00$  SE). This resulted in larger differences between the 2 models in corrected relative counts, with  $M_f$  being unbiased, although uncertainties (ranges of confidence intervals) were comparable. The Connecticut Warbler (*Oporornis agilis*) had 1,162 nonzero counts, and had the most precise parameter estimates of the 3 example species and the highest intercept ( $M_0$ :  $\phi = 0.40 \pm 0.00$  SE;  $M_f$ :  $\phi = 0.17 \pm 0.04$  SE,  $c = 0.54 \pm 0.03$  SE, correlation = 0.56), resulting in low bias and variance compared with the other 2 species.

### Relative Support for Models and Plotting Time-varying Effects

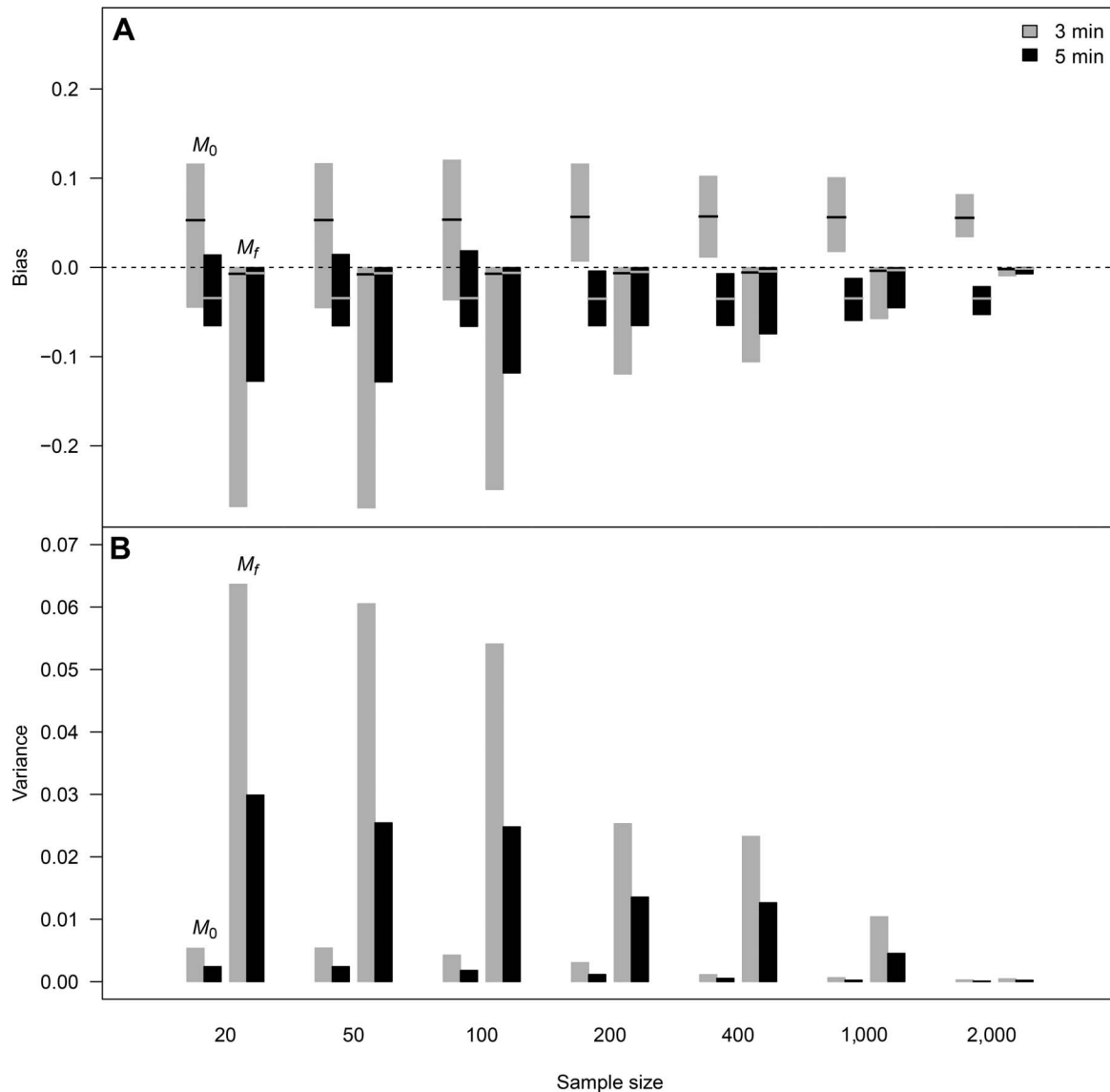
We fit 44 conventional and finite mixture removal models to 147 species detected during at least 75 point counts in the full dataset (fitting finite mixture models was unsuccessful for 5 species; Appendix Tables 2 and 3). Finite mixture models were best-supported for 92% of the 147 species ( $M_f^0$  best-supported for 26%,  $M_f^c$  best-supported for 66%; Table 1). The only species for which the time-invariant ( $M_0$ ) model was best-supported based on  $AIC_c$  was the Yellow-billed Cuckoo (*Coccyzus americanus*;  $n = 196$ ). The best-supported models included time since local sunrise (SR) for 75%, ordinal day (DY) for 39%, and days since local spring (LS) for 37% of the species. Best models included both time since sunrise and either ordinal day or days since spring for 52% of the species (Table 1).

Availability probabilities generally declined across the observed ranges of times since sunrise, and were consistently high between 0 and 4 hr after sunrise. Responses were more variable before and after this 4-hr period. Species' responses to time since sunrise were predominantly nonlinear (77 species with a quadratic and 33 with a linear relationship). Availability declined slowly with respect to date, and was highest across species between ordinal days 155 and 175 and 30–60 days after local spring. Species' responses to dates were predominantly linear for both ordinal day (40 species with a linear relationship and 18 with a quadratic relationship) and days since spring (35 species with a linear and 19 with a quadratic relationship). The temporal patterns of predicted availabilities across all species were very similar between the different classes of time-varying removal models ( $M_0^0$ ,  $M_f^0$ , and  $M_f^c$ ; Figure 5).

The selection frequency for ordinal day (DY) was higher for migratory species (41%, 51 out of 124) than for winter residents (30%, 7 out of 23), but the difference was not significant ( $\chi^2 = 0.54$ ,  $P = 0.47$ ). The selection frequency for days since local spring (LS) was somewhat lower for migrants (35%, 44 out of 124) than for resident species (43%, 10 out of 23), also with a nonsignificant ( $\chi^2 = 0.93$ ,  $P = 0.34$ ) difference (Table 1).

### Removal Models vs. PIF Approach to Adjusting for Incomplete Availability

Across 129 species of landbird, the PIF adjustment factor ( $T_{adj}$ ) for BBS counts had median value of 1.27 and ranged from 1.03 to 25.88. The mean of the average availability ( $p_{avg}$ ) across these species was  $0.43 \pm 0.17$  SD; the mean of



**FIGURE 3.** Sample size effects on (A) bias and (B) variance in removal model predictions of the number of individuals of a species counted during a 10-min point count when using the counts of individuals from 3-min (gray) and 5-min (black) point counts. Data from 133 landbird species with varying numbers of nonzero counts (horizontal axis) were used. The distributions represented by each box are from species whose sample size was equal to or greater than the sample size noted on the horizontal axis. Model types (conventional [ $M_0$ ] and finite mixture [ $M_f$ ]) are shown side by side; boxes show the 90% quantile range, and lines (for bias) represent the median.

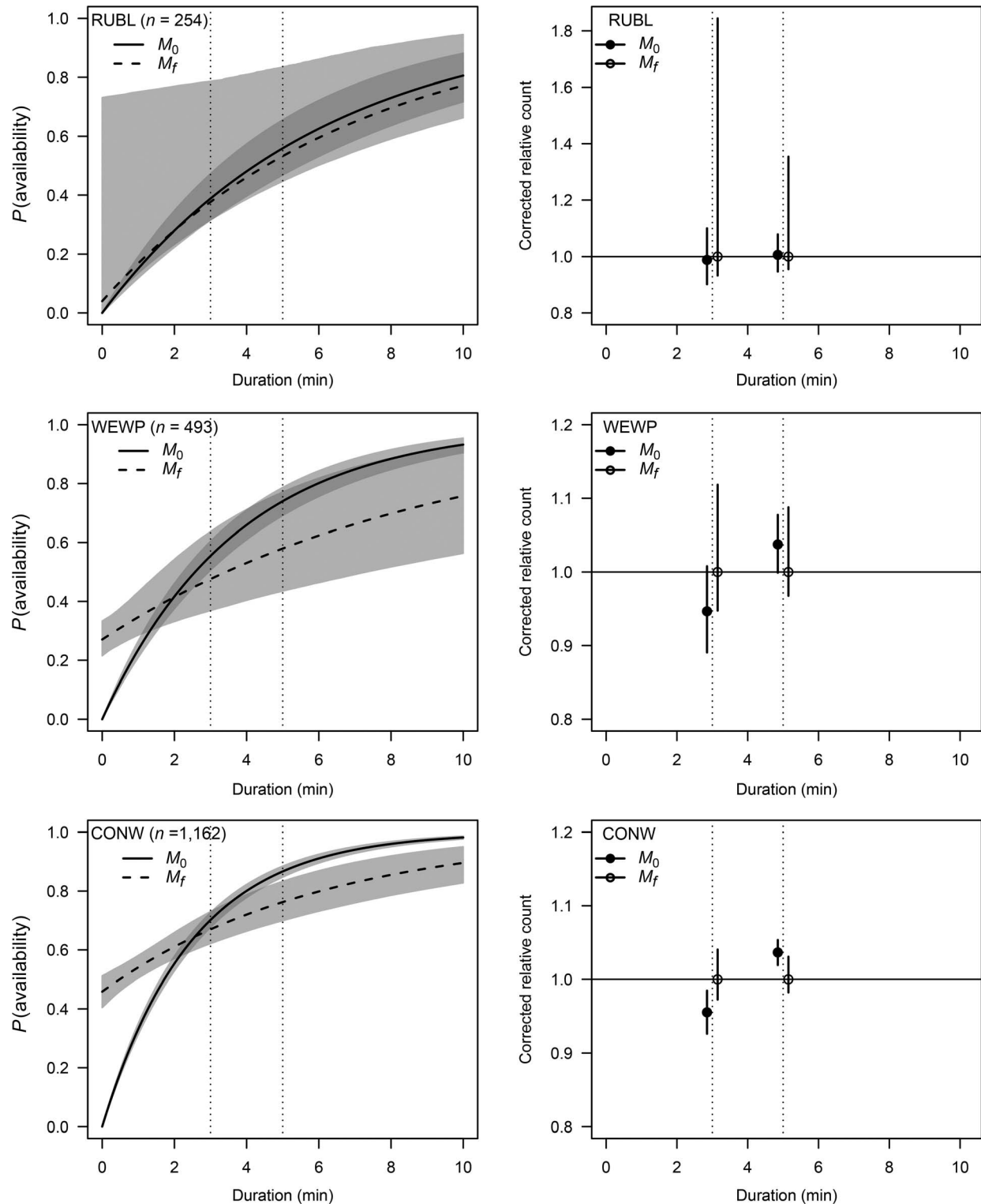
the maximum availability ( $p_{max}$ ) was  $0.66 \pm 0.23$  SD. Our inverse availability based adjustment factor ( $U_{inv}$ ) derived from fitted removal models was 1.7 times higher than  $T_{adj}$  ( $U_{inv}$  median = 2.57, range = 1.25–89.69);  $U_{inv}$  and  $T_{adj}$  were not correlated (Spearman's  $r = 0.005$ ,  $P = 0.95$ ). Average availability was very low, and, as a result,  $U_{inv}$  was very high for the Eastern Bluebird ( $n = 99$ ,  $U_{inv} = 6,788,866.26$ ) and Rusty Blackbird ( $n = 420$ ,  $U_{inv} = 10,022.05$ ) based on their best-supported finite mixture models (Appendix Tables 2 and 3). We show these results to demonstrate the problems that arise when numbers of

detections are too small for the mixture model. These 2 species were excluded from the comparisons.

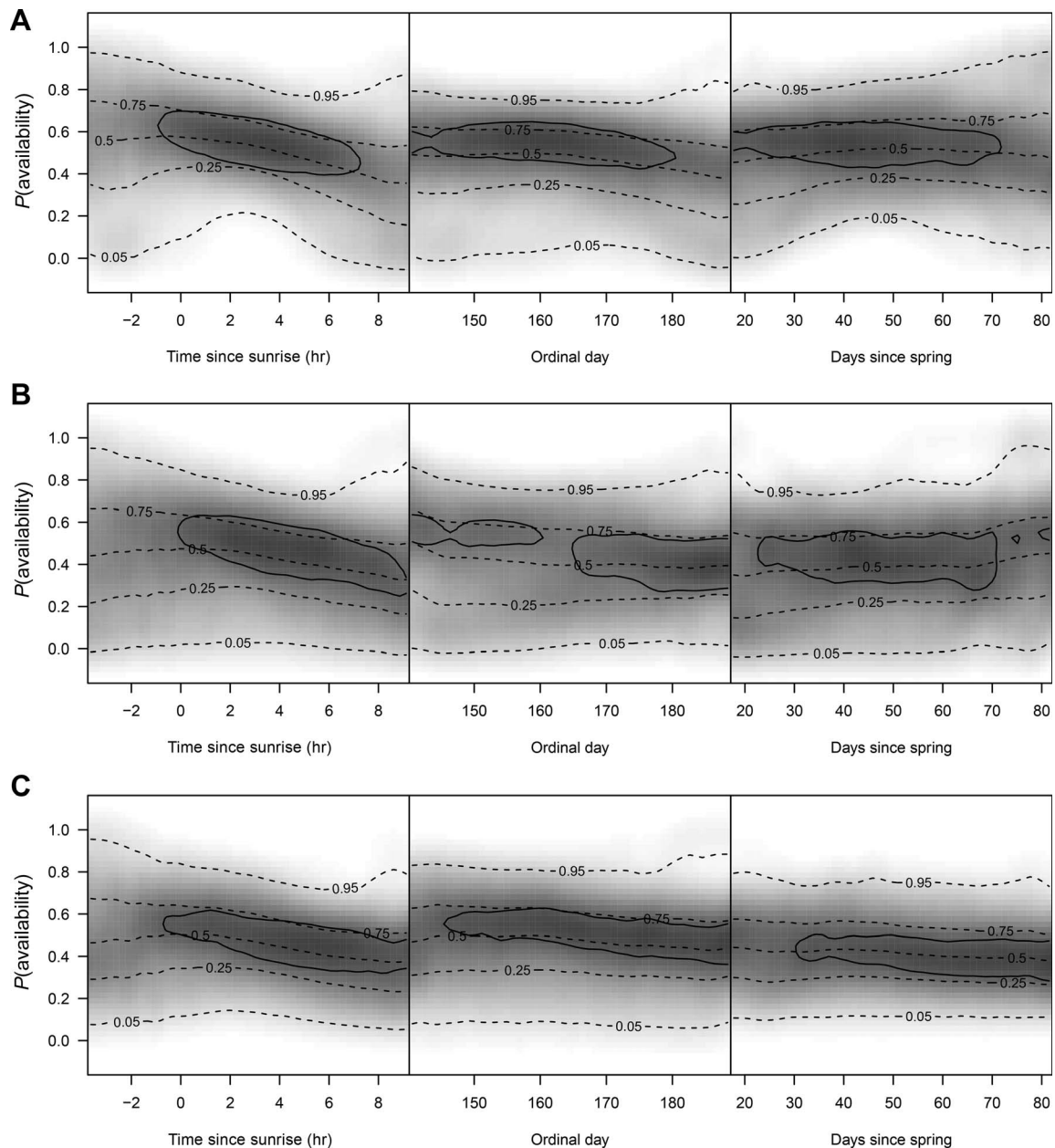
## DISCUSSION

### Conventional vs. Finite Mixture Models

We evaluated alternate formulations of the time-removal model (a conventional removal model and a finite mixture removal model), with and without covariates, for 152 bird species, using a compilation of point count data spanning northern North America. We found that predicted



**FIGURE 4.** Availability probabilities (left) estimated from time-invariant conventional ( $M_0$ ) and finite mixture ( $M_f$ ) removal models for 3 example species, the Rusty Blackbird (*Euphagus carolinus*; RUBL), Western Wood-Pewee (*Contopus sordidulus*; WEWP), and Connecticut Warbler (*Oporornis agilis*; CONW), with 95% confidence intervals (gray shading). Corrected relative counts (right) were calculated as the corrected count divided by the 10-min-based corrected count. Ratio = 1 indicates unbiasedness; continuous vertical lines are 95% confidence intervals. Dotted vertical lines indicate common point count durations (3 and 5 min).



**FIGURE 5.** Availability probabilities estimated from 3 classes of time-varying removal models—**(A)**  $M_0^p$  (conventional with time-varying rate parameter), **(B)**  $M_f^p$  (finite mixture with time-varying rate parameter), and **(C)**  $M_f^c$  (finite mixture with time-varying proportion of infrequent singers)—with covariate effects of hour since local sunrise, ordinal day, and day relative to local spring overlaid for all terrestrial landbird species in northern North America for which the respective variable was part of the best-fit model (see Table 1; other variables were conditioned at their means). Availability probabilities were based on 3-min point count duration. Darker colors in the 2-dimensional kernel density plots indicate regions in which time-varying response curves of multiple species overlap; the continuous contour line indicates the highest density region (75% quantile of the bivariate density). Dashed contours indicate cumulative proportions of species with availability probabilities less than or equal to the value in the contour label conditional on the value of the focal covariate on the horizontal axis.

availability probabilities under conventional and finite mixture models were very similar in their ranges of probability values and shapes of their response curves to predictor variables. However, finite mixture models were supported for 92% of species with sufficient nonzero

counts, in accordance with previous studies that found that mixture models provided better fit (Farnsworth et al. 2002, Etterson et al. 2009, Reidy et al. 2011). By incorporating variation in singing behavior within a survey event, the finite mixture model relaxes the homogeneous singing rate



assumption of the conventional removal models, thus reducing bias when using availability probabilities to adjust survey counts for detectability.

Count duration had a profound effect on average counts. In our data, the first 3-min and first 5-min counts contained, on average, 59% and 73% of the full 10-min counts from the same surveys. Both conventional and finite mixture models reduced this bias considerably. Finite mixture models adjusted the shorter-duration counts with low bias ( $<1\%$ ), whereas conventional models tended to underpredict counts by 5% for 3-min counts and overpredict counts by 3% for 5-min counts when compared with the 10-min counts from the same surveys. The corrected counts using the conventional models were characterized by much smaller variances ( $<1\%$ ) than those from finite mixture models ( $<6\%$ ). However, the variance of finite mixture model-based corrected counts reached acceptable levels for species with at least 200 nonzero counts, the sample size for which bias reduction was greatest.

We evaluated the number of nonzero counts needed to reliably estimate the parameters of the removal models, and found that conventional models required a minimum of 50 nonzero counts, whereas finite mixture models required at least 160 nonzero counts. This corroborates the substantial reduction in variance of the adjusted counts both within and among species at 200 nonzero counts. The low precision of the finite mixture model parameters can considerably inflate the variance of corrected relative counts at low sample sizes (Figures 3 and 4), and at least 1,000 nonzero counts were required to reliably estimate the variance-covariance matrix for finite mixture models. Because the number of nonzero counts was correlated with mean counts (relative abundances) of the study species, it follows that less abundant species will require larger samples for reasonable removal model estimates.

The 1,000 nonzero counts per species required for mixture models (Amundson et al. 2014) is somewhat greater than what is required by other models of detection probability, such as distance sampling (Buckland et al. 2001:240). However, when sample sizes were sufficient, the mixture models were nearly unbiased for most species when adjusting point counts for protocol variation. Because associated variances were often high, even with much larger sample sizes, we recommend that researchers compare results from conventional and finite mixture models, preferring finite mixture models when the estimates and their variances are not outlying values. Otherwise, the benefits of potential bias reduction might be obscured, especially at low detection rates.

### Implications for Survey Design and Modeling

Counts of longer duration increased the numbers of detections and were therefore effective for reducing bias

and variance in the estimates of availability and corrected relative counts. We found that increasing the count duration from 3 to 5 min reduced the range of variation in bias (with respect to 10-min counts) across species, and reduced variance by at least 50% for most species, using both conventional and finite mixture models. Thus, when designing field surveys, lengthening the count duration to 5 or 10 min is an important consideration to increase the accuracy and precision of population estimates by sampling design (Barker et al. 1993, Ralph et al. 1995, Matsuoka et al. 2014). We recognize that shorter count durations (3 min) are favored for logistical reasons, but the choice of count duration must be based on a careful consideration of the tradeoff between survey costs and study objectives.

Longer count durations increase the chances that individuals may move in or out of the surveyed area or will be double counted, which violates assumptions of the removal model and most other models of detectability applied to point counts. Double counting was found to comprise 0.9–3.4% of detections during experimental point counts (Simons et al. 2009), and may also have been low in our study. We excluded count intervals beyond 10 min to minimize the effect of movement-related biases on our estimates. Although longer-duration counts might have been inflated compared with counts in nested subintervals, this would not have changed our conclusions in terms of relative bias among different models.

We found that, among the 135 species for which finite mixture models were found to fit better than conventional models, for 72% there was more support for the models with a time-varying proportion of infrequent singers, when compared with the models with time-varying singing rates. This indicates that our proposed mechanism of individuals switching membership between the groups of frequent and infrequent singers is more realistic than the alternative of fixed proportions as used by previous studies (Farnsworth et al. 2002, Aldredge et al. 2007a, Etterson et al. 2009, Reidy et al. 2011). The overall patterns regarding responses of singing rates to dates and times were very similar, regardless of support across the model classes. The periods when we found detectability to be high (0–4 hr after sunrise and June 4–24 [ordinal day 155–175]; see Results) agreed well with long-standing protocol recommendations (Ralph et al. 1995). Conducting point counts during these times of day and year will make surveys much more efficient for accumulating detections and will thereby improve the resulting accuracy and precision of abundance estimates.

Besides these general guidelines, local and regional context and study goals should also shape survey timing. Although sunrise time already reflects geographic differences, weather and resulting local spring date can vary considerably across years, thus influencing the timing of

migration for nonresidents, and the timing of high vocal activity for most species. Regional differences can also introduce heterogeneity into bird availability through geographic advancement of breeding through the growing season. We captured this breeding progression with the days since spring variable (vs. ordinal days). The fact that the 2 predictors related to time of year were supported by different species suggests differences among species' plasticity of response to interannual variations in weather conditions. We found no significant differences between winter resident and migratory species with respect to selection frequencies of time-of-year covariates. This indicates that, although ordinal day and multidecadal averages (our days since spring variable) are important predictors of availability, time-of-year covariates are inadequate for establishing linkages between breeding phenology and weather events for migratory species (Both and Visser 2001, Both et al. 2010). Estimating year-specific effects with respect to days since spring vs. ordinal day might be indicative of changes in breeding phenology with a changing climate, which will be important to account for in programs monitoring long-term changes in bird population sizes. Accounting for year-to-year variation in covariates is important given the large interannual variation in spring green-up (15–20 days in the boreal region; Delbart et al. 2006) and the strong indications of advancing phenology in boreal and hemiboreal regions with climate change (Linkosalo et al. 2009, Beaubien and Hamann 2011).

## Conclusions

Our study demonstrates the importance of accounting for temporal heterogeneity in bird availability when analyzing avian point count data. This is especially true when combining data across studies with different locations, sampling timeframes, or protocols. Location influences the expected time of the breeding season, and the timing of surveys relative to the local breeding season affects availability, for example, through time-varying behaviors. It is difficult to retrospectively control for these sources of heterogeneity introduced by sampling design, and assuming constant availability in space and time is unrealistic, especially when analyzing heterogeneous and/or spatially extensive datasets. Thus, a model-based approach to analysis is required. We argue that the continuous time-removal modeling framework outlined in this manuscript is sufficient. Easily obtainable covariates, such as the ones used in this study, can help to better correct availability bias of counts.

Comparing our average availability-based corrections with the time adjustment factors developed by Partners in Flight (PIF; Rosenberg and Blancher 2005, Blancher et al. 2013), we found our removal model-based adjustment factors to be 1.6 times higher. This was largely due to the

PIF assumption of 100% availability during the times of day with the maximum counts, whereas our maximum availability estimates for a 3-min point count averaged 66%. As a result, PIF population size estimates are most likely to be underestimates for most species. This availability-related underestimation is relatively small compared with the 5-fold difference in population sizes when adjusting point counts for perceptibility probability using distance sampling versus PIF's expert-derived adjustment for maximum detection distances (Matsuoka et al. 2012). Similarly, both field tests (Confer et al. 2008) and sensitivity analyses (Thogmartin 2010) have indicated that PIF population size estimates are more strongly influenced by the choice of adjustment for perceptibility (maximum detection distance) than availability (time adjustment). We feel that availability probabilities from removal models could replace PIF's time adjustment (Rosenberg and Blancher 2005, Blancher et al. 2013) to modestly improve population size estimates, particularly if covariates were included in the removal models to account for date- and time-related heterogeneity in singing rates (Thogmartin et al. 2006). We did not find a correlation between correction factors based on removal models and PIF's time adjustments, even after adjusting for maximum availability. This is most likely due to geographic differences (mostly nonboreal vs. boreal regions, respectively) and the on- vs. off-road nature of the BBS and BAM datasets, which are complexities to be tackled by future studies.

The probability of availability, given the presence of an individual, can be efficiently estimated by conventional or finite mixture type removal models, with or without covariates for bird behavior. In this paper, we have outlined a flexible and continuous-time parameterization of these model types implemented in R package *detect* (Sólymos et al. 2018). We have extended removal model methodology to flexibly model time-varying finite mixture models using covariates for either the singing rate ( $\phi$ ) or the proportion of infrequent singers ( $c$ ) parameter. We found that the choice of model had an impact on the estimability of unknown model parameters, and that this choice also affected the bias and variance of the corrected relative counts. In general, a conventional removal model is recommended for smaller sample sizes. Model selection approaches can be used to compare support for different parameterizations, but only at much higher sample sizes ( $>200$  nonzero counts, but preferably  $>1,000$ ). Lengthening the count duration from 3 min to 5 or 10 min increased sample sizes of detections and improved the accuracy and precision of removal model estimates. Well-informed survey design combined with various forms of removal sampling are useful for accounting for availability bias in point counts, thereby improving population size estimates and allowing for better integration of disparate studies at large spatial scales.

## ACKNOWLEDGMENTS

This publication is a contribution of the Boreal Avian Modelling (BAM) Project, an international research collaboration on the ecology, management, and conservation of boreal birds. We acknowledge BAM's members, avian and biophysical data partners, and funding agencies (including Environment and Climate Change Canada and the U.S. Fish and Wildlife Service), listed in full at [www.borealbirds.ca/index.php/acknowledgements](http://www.borealbirds.ca/index.php/acknowledgements)

**Funding statement:** This project was funded by Environment and Climate Change Canada, the Natural Sciences and Engineering Research Council of Canada, U.S. Fish and Wildlife Service's Neotropical Migratory Bird Conservation Act, and the Alberta Biodiversity Monitoring Institute. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government. None of the funders had any influence on the content of the submitted or published manuscript. This paper has been peer reviewed and approved for publication consistent with USGS Fundamental Science Practices (<http://pubs.usgs.gov/circ/1367/>).

**Ethics statement:** This research was conducted in compliance with the *Ethical Guidelines for Statistical Practice*.

**Author contributions:** P.S., S.M.M., and E.M.B. conceived the idea and designed the study; P.S. analyzed the data; P.F. organized data for analyses; and all authors substantially contributed to writing the paper.

## LITERATURE CITED

- Allredge, M. W., K. H. Pollock, T. R. Simons, J. A. Collazo, and S. A. Shriner (2007a). Time-of-detection method for estimating abundance from point-count surveys. *The Auk* 124:653–664.
- Allredge, M. W., T. R. Simons, K. H. Pollock, and K. Pacifici (2007b). A field evaluation of the time-of-detection method to estimate population size and density for aural avian point counts. *Avian Conservation and Ecology* 2:art.13.
- Amundson, C. L., J. A. Royle, and C. M. Handel (2014). A hierarchical model combining distance sampling and time removal to estimate detection probability during avian point counts. *The Auk: Ornithological Advances* 131:476–494.
- Anderson, D. R., and R. S. Pospahala (1970). Correction of bias in belt transects of immobile objects. *The Journal of Wildlife Management* 34:141–146.
- Barker, N. K. S., P. C. Fontaine, S. G. Cumming, D. Stralberg, A. Westwood, E. M. Bayne, P. Sólymos, F. K. A. Schmiegelow, S. J. Song, and D. J. Rugg (2015). Ecological monitoring through harmonizing existing data: Lessons from the Boreal Avian Modelling Project. *Wildlife Society Bulletin* 39:480–487.
- Barker, R. J., and J. R. Sauer (1995). Statistical aspects of point count sampling. In *Monitoring Bird Populations by Point Counts* (C. J. Ralph, J. R. Sauer, and S. Droege, Technical Editors). USDA Forest Service General Technical Report PSW-GTR-149. pp. 125–130.
- Barker, R. J., J. R. Sauer, and W. A. Link (1993). Optimal allocation of point-count sampling effort. *The Auk* 110:752–758.
- Bart, J. (2005). Monitoring the abundance of bird populations. *The Auk* 122:15–25.
- Beaubien, E., and A. Hamann (2011). Spring flowering response to climate change between 1936 and 2006 in Alberta, Canada. *BioScience* 61:514–524.
- Bivand, R., and N. Lewin-Koh (2016). *maptools: Tools for reading and handling spatial objects*. R package version 0.8-39. <https://CRAN.R-project.org/package=maptools>
- Blancher, P. J., K. V. Rosenberg, A. O. Panjabi, B. Altman, A. R. Couturier, W. E. Thogmartin, and the Partners in Flight Science Committee (2013). *Handbook to the Partners in Flight Population Estimates Database*, version 2.0. Partners in Flight Technical Series No. 6. <http://pif.birdconservancy.org/PopEstimates/Userguide.aspx>
- Both, C., and M. E. Visser (2001). Adjustment to climate change is constrained by arrival date in a long-distance migrant bird. *Nature* 411:296–298.
- Both, C., C. A. M. Van Turnhout, R. G. Bijlsma, H. Siepel, A. J. Van Strien, and R. P. B. Foppen (2010). Avian population consequences of climate change are most severe for long-distance migrants in seasonal habitats. *Proceedings of the Royal Society B* 277:1259–1266.
- Brandt, J. P. (2009). The extent of the North American boreal zone. *Environmental Reviews* 17:101–161.
- Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas (2001). *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*. Oxford University Press, Oxford, UK.
- 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. *The Auk* 125:932–938.
- Cumming, S. G., K. Lefevre, E. Bayne, T. Fontaine, F. K. A. Schmiegelow, and S. J. Song (2010). Toward conservation of Canada's boreal forest avifauna: Design and application of ecological models at continental extents. *Avian Conservation and Ecology* 5:art.8.
- Delbart, N., T. Le Toan, L. Kergoat, and V. Fedotova (2006). Remote sensing of spring phenology in boreal regions: A free of snow-effect method using NOAA-AVHRR and SPOT-VGT data (1982–2004). *Remote Sensing of Environment* 101:52–62.
- Dickinson, J. L., B. Zuckerberg, and D. N. Bonter (2010). Citizen science as an ecological research tool: Challenges and benefits. *Annual Review of Ecology, Evolution, and Systematics* 41:149–172.
- Diefenbach, D. R., M. R. Marshall, J. A. Mattice, and D. W. Brauning (2007). Incorporating availability for detection in estimates of bird abundance. *The Auk* 124:96–106.
- Etterson, M. A., G. J. Niemi, and N. P. Danz (2009). Estimating the effects of detection heterogeneity and overdispersion on trends estimated from avian point counts. *Ecological Applications* 19:2049–2066.
- Farnsworth, G. L., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines, and J. R. Sauer (2002). A removal model for estimating detection probabilities from point-count surveys. *The Auk* 119:414–425.
- Glennie, R., S. T. Buckland, and L. Thomas (2015). The effect of animal movement on line transect estimates of abundance. *PLOS One* 10:e0121333.
- Hampton, S. E., C. A. Strasser, J. J. Tewksbury, W. K. Gram, A. E. Budden, A. L. Batcheller, C. S. Duke, and J. H. Porter (2013). Big data and the future of ecology. *Frontiers in Ecology and the Environment* 11:156–162.



- Handel, C. M., S. A. Swanson, D. A. Nigro, and S. M. Matsuoka (2009). Estimation of avian population sizes and species richness across a boreal landscape in Alaska. *The Wilson Journal of Ornithology* 121:528–547.
- Hayne, D. W. (1949). An examination of the strip census method for estimating animal populations. *The Journal of Wildlife Management* 13:145–157.
- Jewell, N. P. (1982). Mixtures of exponential distributions. *Annals of Statistics* 10: 479–484.
- Johnson, D. H. (2008). In defense of indices: The case of bird surveys. *The Journal of Wildlife Management* 72:857–868.
- Kelling, S., W. M. Hochachka, D. Fink, M. Riedewald, R. Caruana, G. Ballard, and G. Hooker (2009). Data-intensive science: A new paradigm for biodiversity studies. *BioScience* 59:613–620.
- Kendeigh, S. C. (1944). Measurement of bird populations. *Ecological Monographs* 14:67–106.
- Kociolek, A. V., A. P. Clevenger, C. C. St. Clair, and D. S. Proppe (2011). Effects of road networks on bird populations. *Conservation Biology* 25:241–249.
- Leopold, A. (1933). *Game Management*. Charles Scribner's Sons, New York, NY, USA.
- Linkosalo, T., R. Häkkinen, J. Terhivuo, H. Tuomenvirta, and P. Hari (2009). The time series of flowering and leaf bud burst of boreal trees (1846–2005) support the direct temperature observations of climatic warming. *Agricultural and Forest Meteorology* 149:453–461.
- 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. *The Auk* 129:268–282.
- 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. *The Condor: Ornithological Applications* 116:599–608.
- McKenney, D., D. Price, P. Papadapol, M. Siltanen, and K. Lawrence (2006). High-resolution climate change scenarios for North America. *Frontline: Forestry Research Applications Technical Note No. 107*, Natural Resources Canada, Great Lakes Forestry Centre, Canadian Forest Service, Sault Ste. Marie, ON, Canada.
- Meeus, J. (1999). *Astronomical Algorithms*, second edition. Wilmann-Bell, Richmond, VA, USA.
- Moran, P. A. P. (1951). A mathematical theory of animal trapping. *Biometrika* 38:307–311.
- Nichols, J. D., L. Thomas, and P. B. Conn (2009). Inferences about landbird abundance from count data: Recent advances and future directions. In *Environmental and Ecological Statistics, Volume 3: Modeling Demographic Processes in Marked Populations* (D. L. Thomson, E. G. Cooch, and M. J. Conroy, Editors). Springer Science and Business Media, New York, NY, USA. pp. 201–235.
- O'Connor, R. J., E. Dunn, D. H. Johnson, S. L. Jones, D. Petit, K. Pollock, C. R. Smith, J. L. Trapp, and E. Welling (2000). A programmatic review of the North American Breeding Bird Survey: Report of a peer review panel to the U.S. Geological Survey Patuxent Wildlife Research Center, Laurel, MD, USA. <http://www.pwrc.usgs.gov/bbs/bbsreview/bbsfinal.pdf>
- Poole, A. (Editor) (2012). *The Birds of North America*. Cornell Laboratory of Ornithology, Ithaca, NY, USA. <https://birdsna.org>
- Ralph, C. J., S. Droege, and J. R. Sauer (1995). Managing and monitoring birds using point counts: Standards and applications. In *Monitoring Bird Populations by Point Counts* (C. J. Ralph, J. R. Sauer, and S. Droege, Technical Editors). USDA Forest Service General Technical Report PSW-GTR-149. pp. 161–169.
- R Core Team (2017). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/>
- Reidy, J. L., F. R. Thompson, III, and J. W. Bailey (2011). Comparison of methods for estimating density of forest songbirds from point counts. *The Journal of Wildlife Management* 75:558–568.
- Rosenberg, K. V., and P. J. Blancher (2005). Setting numerical population objectives for priority landbird species. In *Bird Conservation Implementation and Integration in the Americas: Proceedings of the Third International Partners in Flight Conference* (C. J. Ralph and T. D. Rich, Editors). USDA Forest Service General Technical Report PSW-GTR-191. pp. 57–67.
- Seber, G. A. F. (1982). *Estimation of Animal Abundance and Related Parameters*. Macmillan, New York, NY, USA.
- Simons, T. R., K. H. Pollock, J. M. Wettroth, M. W. Alldredge, K. Pacifici, and J. Brewster (2009). Sources of measurement error, misclassification error, and bias in auditory avian point count data. In *Environmental and Ecological Statistics, Volume 3: Modeling Demographic Processes in Marked Populations* (D. L. Thomson, E. G. Cooch, and M. J. Conroy, Editors). Springer Science and Business Media, New York, NY, USA. pp. 237–254.
- Sólymos, P., and S. R. Lele (2016). Revisiting resource selection probability functions and single-visit methods: Clarification and extensions. *Methods in Ecology and Evolution* 7:196–205.
- 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.
- Sólymos, P., M. Moreno, and S. R. Lele (2018). detect: Analyzing wildlife data with detection error. R package version 0.4-1. <https://cran.r-project.org/package=detect>
- Stacies, C. A., V. Ingalls, and T. W. Sherry (2006). Singing behavior varies with breeding status of American Redstarts (*Setophaga ruticilla*). *The Wilson Journal of Ornithology* 118:439–451.
- Thogmartin, W. E. (2010). Sensitivity analysis of North American bird population estimates. *Ecological Modelling* 221:173–177.
- Thogmartin, W. E., F. P. Howe, F. C. James, D. H. Johnson, E. T. Reed, J. R. Sauer, and F. R. Thompson, III (2006). A review of the population estimation approach of the North American Landbird Conservation Plan. *The Auk* 123:892–904.
- Twedt, D. J. (2015). Estimating regional landbird populations from enhanced North American Breeding Bird Surveys. *Journal of Field Ornithology* 86:352–368.
- Wand, M. (2015). KernSmooth: Functions for kernel smoothing supporting Wand & Jones (1995). R package version 2.23-15. <https://CRAN.R-project.org/package=KernSmooth>

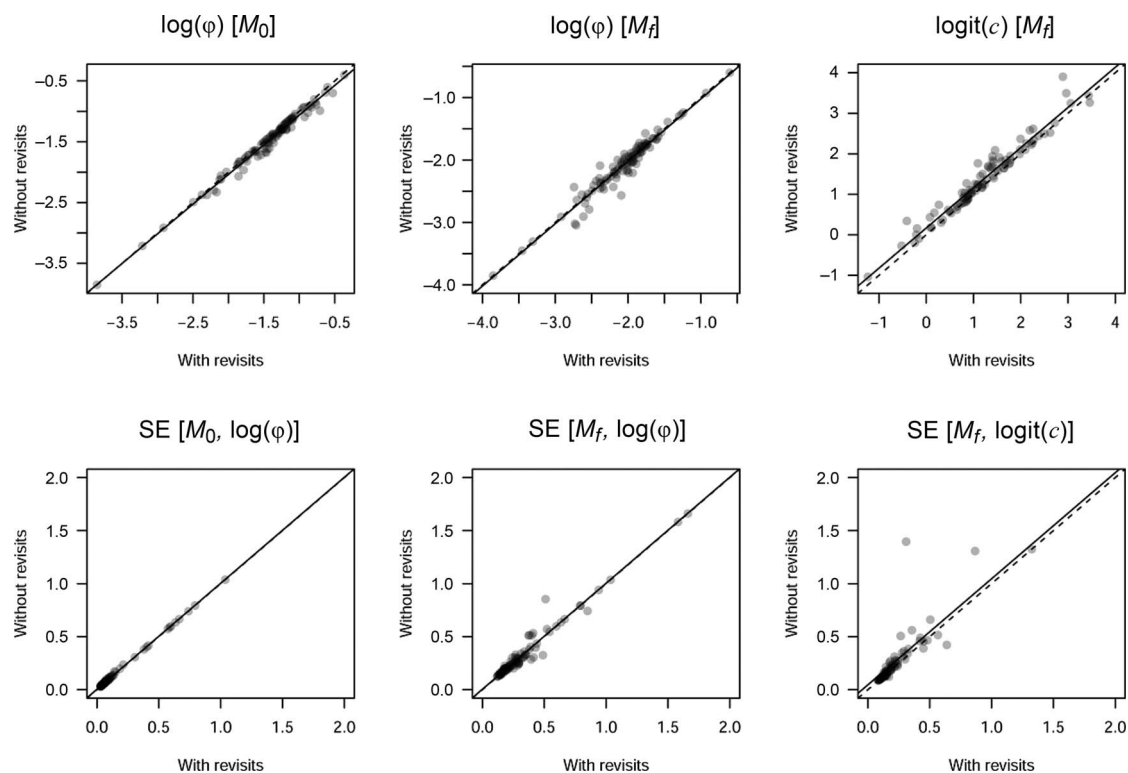


Wand, M. P., and M. C. Jones (1995). Kernel Smoothing. Chapman and Hall, London, UK.

Wilson, D. M., and J. Bart (1985). Reliability of singing bird surveys: Effects of song phenology during the breeding season. *The Condor* 87:69–73.

Zippin, C. (1956). An evaluation of the removal method of estimating animal populations. *Biometrics* 12:163–189.

Zippin, C. (1958). The removal method of population estimation. *The Journal of Wildlife Management* 22:82–90.



**APPENDIX FIGURE 6.** Comparing estimates of removal model parameters based on data containing revisits vs. data without revisits (within or between years) to the same point count location. Sample sizes were kept constant (1,000 or the unique locations with detections, whichever was smallest); time-invariant conventional ( $M_0$ ) and finite mixture ( $M_f$ ) removal model parameters (first row) and their standard errors (second row) were estimated 200 times for each of 106 species using resampling without replacement. Dots correspond to the medians of the 200 replicates. The continuous line corresponds to the fitted linear model, and the dashed line indicates the expected 1:1 relationship. The fitted lines are close to the 1:1 relationship, indicating that using data from revisits did not bias our parameter estimates and had no effect on the standard errors of the estimates; therefore, our conditional independence assumption holds for our dataset, and our inferences and sample size recommendations based on revisit data are unaffected.

**APPENDIX TABLE 2.** Common and scientific names of the species used in this study evaluating the probability of availability during point count surveys of terrestrial landbird species in northern North America, with number of nonzero counts ( $n$ ) from the full dataset of point count surveys, along with parameter estimates of availability ( $\phi$ : singing rate;  $c$ : proportion of infrequent singers) from time-invariant models (conventional:  $M_0$ ; finite mixture:  $M_f$ ).

Common name	Scientific name	$n$	$M_0$	$M_f$	
			$\phi$	$\phi$	$c$
Ruffed Grouse	<i>Bonasa umbellus</i>	2,730	0.2740	0.2307	0.8863
Willow Ptarmigan	<i>Lagopus lagopus</i>	181	0.0571	0.0570	1.0000
Wild Turkey	<i>Meleagris gallopavo</i>	180	0.2228	0.1300	0.8264
Mourning Dove	<i>Zenaida macroura</i>	1,235	0.2339	0.0744	0.7871
Yellow-billed Cuckoo	<i>Coccyzus americanus</i>	196	0.1547	0.1547	1.0000
Black-billed Cuckoo	<i>Coccyzus erythrophthalmus</i>	412	0.1256	0.0320	0.9364
Ruby-throated Hummingbird	<i>Archilochus colubris</i>	289	0.0595	0.0173	0.9846
Killdeer	<i>Charadrius vociferus</i>	758	0.2154	0.0923	0.7964
Upland Sandpiper	<i>Bartramia longicauda</i>	364	0.1635	0.0425	0.8867
Long-billed Curlew	<i>Numenius americanus</i>	439	0.2897	0.0989	0.6703
Marbled Godwit	<i>Limosa fedoa</i>	705	0.2625	0.1348	0.7389
Rock Sandpiper	<i>Calidris ptilocnemis</i>	141	0.2329	0.1438	0.8775
Wilson's Snipe	<i>Gallinago delicata</i>	5,358	0.2807	0.1246	0.7181
Spotted Sandpiper	<i>Actitis macularia</i>	498	0.1765	0.1140	0.8861
Solitary Sandpiper	<i>Tringa solitaria</i>	556	0.0994	0.0739	0.9689
Lesser Yellowlegs	<i>Tringa flavipes</i>	914	0.1652	0.1652	1.0000
Willet	<i>Tringa semipalmata</i>	738	0.2339	0.1274	0.7887
Greater Yellowlegs	<i>Tringa melanoleuca</i>	922	0.2558	0.1128	0.7544
Belted Kingfisher	<i>Megasceryle alcyon</i>	132	0.0663	N/A	N/A
Red-bellied Woodpecker	<i>Melanerpes carolinus</i>	419	0.1149	N/A	N/A
Yellow-bellied Sapsucker	<i>Sphyrapicus varius</i>	9,655	0.1547	0.0939	0.9036
Downy Woodpecker	<i>Picoides pubescens</i>	1,201	0.1085	0.0202	0.9591
Hairy Woodpecker	<i>Picoides villosus</i>	2,063	0.1332	0.0711	0.9190
American Three-toed Woodpecker	<i>Picoides dorsalis</i>	421	0.1569	0.1354	0.9544
Black-backed Woodpecker	<i>Picoides arcticus</i>	384	0.1970	0.1514	0.9028
Northern Flicker	<i>Colaptes auratus</i>	3,502	0.1217	0.0702	0.9334
Pileated Woodpecker	<i>Dryocopus pileatus</i>	2,386	0.1222	0.0535	0.9281
Olive-sided Flycatcher	<i>Contopus cooperi</i>	1,495	0.2807	0.1937	0.7949
Western Wood-Pewee	<i>Contopus sordidulus</i>	913	0.3061	0.1429	0.6788
Eastern Wood-Pewee	<i>Contopus virens</i>	5,181	0.3059	0.1421	0.6879
Yellow-bellied Flycatcher	<i>Empidonax flaviventris</i>	4,652	0.2721	0.1468	0.7481
Alder Flycatcher	<i>Empidonax alnorum</i>	7,226	0.3205	0.2127	0.7415
Least Flycatcher	<i>Empidonax minimus</i>	9,394	0.4174	0.1809	0.5275
Hammond's Flycatcher	<i>Empidonax hammondi</i>	416	0.2809	0.2809	1.0000
Eastern Phoebe	<i>Sayornis phoebe</i>	250	0.1859	0.0401	0.8722
Great Crested Flycatcher	<i>Myiarchus crinitus</i>	1,944	0.1678	0.0669	0.8760
Eastern Kingbird	<i>Tyrannus tyrannus</i>	427	0.2436	0.1523	0.8126
Yellow-throated Vireo	<i>Vireo flavifrons</i>	988	0.1995	0.0848	0.8338
Blue-headed Vireo	<i>Vireo solitarius</i>	3,846	0.2200	0.0935	0.7980
Philadelphia Vireo	<i>Vireo philadelphicus</i>	643	0.3471	0.1891	0.6552
Warbling Vireo	<i>Vireo gilvus</i>	2,427	0.3330	0.1130	0.6031
Red-eyed Vireo	<i>Vireo olivaceus</i>	30,338	0.3767	0.1755	0.5901
Gray Jay	<i>Perisoreus canadensis</i>	8,443	0.1519	0.0953	0.9090
Blue Jay	<i>Cyanocitta cristata</i>	10,494	0.1567	0.0959	0.9055
Black-billed Magpie	<i>Pica hudsonia</i>	1,227	0.2313	0.1247	0.7944
American Crow	<i>Corvus brachyrhynchos</i>	8,314	0.2572	0.1315	0.7611
Common Raven	<i>Corvus corax</i>	6,150	0.1935	0.1127	0.8557
Horned Lark	<i>Eremophila alpestris</i>	1,959	0.4921	0.1600	0.3909
Tree Swallow	<i>Tachycineta bicolor</i>	1,060	0.2271	0.0650	0.8019
Bank Swallow	<i>Riparia riparia</i>	146	0.2073	0.1374	0.8536
Cliff Swallow	<i>Petrochelidon pyrrhonota</i>	88	0.2265	N/A	N/A
Barn Swallow	<i>Hirundo rustica</i>	224	0.1259	0.0594	0.9215
Black-capped Chickadee	<i>Poecile atricapillus</i>	7,405	0.1697	0.1162	0.9044
Boreal Chickadee	<i>Poecile hudsonicus</i>	2,078	0.1389	0.1128	0.9530
Red-breasted Nuthatch	<i>Sitta canadensis</i>	8,103	0.1600	0.0774	0.8854
White-breasted Nuthatch	<i>Sitta carolinensis</i>	1,735	0.1496	0.0250	0.9280

APPENDIX TABLE 2. Continued.

Common name	Scientific name	<i>n</i>	<i>M</i> <sub>0</sub>	<i>M</i> <sub>f</sub>	
			φ	φ	<i>c</i>
Brown Creeper	<i>Certhia americana</i>	2,930	0.1803	0.1098	0.8795
House Wren	<i>Troglodytes aedon</i>	2,180	0.4176	0.1387	0.5134
Winter Wren	<i>Troglodytes hiemalis</i>	9,450	0.3478	0.1653	0.6240
Sedge Wren	<i>Cistothorus platensis</i>	254	0.2858	0.1137	0.7312
Marsh Wren	<i>Cistothorus palustris</i>	126	0.3757	0.0320	0.7333
Blue-gray Gnatcatcher	<i>Polioptila caerulea</i>	534	0.2718	0.0514	0.7574
Golden-crowned Kinglet	<i>Regulus satrapa</i>	3,987	0.2007	0.1266	0.8570
Ruby-crowned Kinglet	<i>Regulus calendula</i>	11,782	0.2862	0.1724	0.7468
Arctic Warbler	<i>Phylloscopus borealis</i>	139	0.2256	0.1144	0.7932
Eastern Bluebird	<i>Sialia sialis</i>	99	0.0707	0.0144	0.9825
Townsend's Solitaire	<i>Myadestes townsendi</i>	92	0.1234	0.1234	1.0000
Veery	<i>Catharus fuscescens</i>	10,332	0.2652	0.1489	0.7679
Gray-cheeked Thrush	<i>Catharus minimus</i>	489	0.2145	0.1409	0.8440
Swainson's Thrush	<i>Catharus ustulatus</i>	19,357	0.3266	0.1761	0.6594
Hermit Thrush	<i>Catharus guttatus</i>	17,623	0.3228	0.1656	0.6724
Wood Thrush	<i>Hylocichla mustelina</i>	1,235	0.3051	0.1380	0.6861
American Robin	<i>Turdus migratorius</i>	16,001	0.2332	0.1447	0.8219
Varied Thrush	<i>Ixoreus naevius</i>	2,313	0.2517	0.2516	1.0000
Gray Catbird	<i>Dumetella carolinensis</i>	819	0.2126	0.0925	0.8136
Brown Thrasher	<i>Toxostoma rufum</i>	407	0.2839	0.1285	0.7128
European Starling	<i>Sturnus vulgaris</i>	251	0.3268	0.0983	0.6104
Bohemian Waxwing	<i>Bombycilla garrulus</i>	116	0.0554	0.0554	1.0000
Cedar Waxwing	<i>Bombycilla cedrorum</i>	2,335	0.1143	0.0441	0.9373
House Sparrow	<i>Passer domesticus</i>	227	0.4480	0.1069	0.4336
American Pipit	<i>Anthus rubescens</i>	336	0.1885	0.0495	0.8554
Sprague's Pipit	<i>Anthus spragueii</i>	1,074	0.5236	0.1259	0.3526
Evening Grosbeak	<i>Coccothraustes vespertinus</i>	627	0.1205	0.0531	0.9295
Pine Grosbeak	<i>Pinicola enucleator</i>	213	0.0990	N/A	N/A
Purple Finch	<i>Haemorhous purpureus</i>	1,122	0.1104	0.0656	0.9465
Red Crossbill	<i>Loxia curvirostra</i>	133	0.2605	0.1037	0.7269
White-winged Crossbill	<i>Loxia leucoptera</i>	1,961	0.1287	0.1056	0.9594
Pine Siskin	<i>Spinus pinus</i>	2,973	0.1738	0.0306	0.8934
American Goldfinch	<i>Spinus tristis</i>	1,876	0.1136	0.0759	0.9494
Lapland Longspur	<i>Calcarius lapponicus</i>	210	0.3021	0.1707	0.7759
Chestnut-collared Longspur	<i>Calcarius ornatus</i>	535	0.5270	0.1728	0.3469
McCown's Longspur	<i>Rhynchophanes mccownii</i>	192	0.3554	0.0803	0.5752
Eastern Towhee	<i>Pipilo erythrophthalmus</i>	836	0.2705	0.0835	0.7299
American Tree Sparrow	<i>Spizelloides arborea</i>	1,059	0.2394	0.2394	1.0000
Chipping Sparrow	<i>Spizella passerina</i>	15,060	0.2816	0.1142	0.7003
Clay-colored Sparrow	<i>Spizella pallida</i>	4,469	0.4619	0.1492	0.4463
Field Sparrow	<i>Spizella pusilla</i>	187	0.1825	0.1001	0.8681
Vesper Sparrow	<i>Pooecetes gramineus</i>	3,128	0.5324	0.1608	0.3649
Savannah Sparrow	<i>Passerculus sandwichensis</i>	5,048	0.4665	0.1737	0.4490
Grasshopper Sparrow	<i>Ammodramus savannarum</i>	149	0.5503	0.2108	0.3472
Baird's Sparrow	<i>Ammodramus bairdii</i>	651	0.5525	0.2202	0.3531
LeConte's Sparrow	<i>Ammodramus leconteii</i>	1,312	0.5880	0.1555	0.4611
Nelson's Sparrow	<i>Ammodramus nelsoni</i>	109	0.2599	0.1252	0.7533
Fox Sparrow	<i>Passerella iliaca</i>	2,243	0.2118	0.1795	0.9207
Song Sparrow	<i>Melospiza melodia</i>	3,961	0.2912	0.1643	0.7355
Lincoln's Sparrow	<i>Melospiza lincolni</i>	5,930	0.3054	0.1585	0.7016
Swamp Sparrow	<i>Melospiza georgiana</i>	2,677	0.2512	0.0973	0.7605
White-throated Sparrow	<i>Zonotrichia albicollis</i>	28,598	0.3012	0.1417	0.6852
White-crowned Sparrow	<i>Zonotrichia leucophrys</i>	2,752	0.2656	0.2628	0.9917
Golden-crowned Sparrow	<i>Zonotrichia atricapilla</i>	299	0.2010	0.2010	1.0000
Dark-eyed Junco	<i>Junco hyemalis</i>	10,061	0.2222	0.1983	0.9404
Yellow-headed Blackbird	<i>Xanthocephalus xanthocephalus</i>	444	0.4409	0.2277	0.5533
Western Meadowlark	<i>Sturnella neglecta</i>	2,663	0.6920	0.1980	0.2181
Baltimore Oriole	<i>Icterus galbula</i>	572	0.2227	0.0322	0.8473
Red-winged Blackbird	<i>Agelaius phoeniceus</i>	4,060	0.3943	0.1549	0.5374

APPENDIX TABLE 2. Continued.

Common name	Scientific name	<i>n</i>	<i>M</i> <sub>0</sub>	<i>M</i> <sub>f</sub>	
			φ	φ	<i>c</i>
Brown-headed Cowbird	<i>Molothrus ater</i>	3,694	0.2042	0.0960	0.8223
Rusty Blackbird	<i>Euphagus carolinus</i>	420	0.1569	0.0777	0.8956
Brewer's Blackbird	<i>Euphagus cyanocephalus</i>	346	0.2822	N/A	N/A
Common Grackle	<i>Quiscalus quiscula</i>	153	0.2247	0.1217	0.8037
Ovenbird	<i>Seiurus aurocapilla</i>	34,929	0.3870	0.1797	0.5774
Northern Waterthrush	<i>Parkesia noveboracensis</i>	2,890	0.2517	0.1307	0.7644
Golden-winged Warbler	<i>Vermivora chrysoptera</i>	933	0.2143	0.0680	0.8144
Blue-winged Warbler	<i>Vermivora cyanopectus</i>	154	0.2341	0.0361	0.8174
Black-and-white Warbler	<i>Mniotilta varia</i>	7,578	0.2439	0.1114	0.7702
Tennessee Warbler	<i>Oreothlypis peregrina</i>	15,533	0.4508	0.1311	0.4354
Orange-crowned Warbler	<i>Oreothlypis celata</i>	3,081	0.2048	0.2048	1.0000
Nashville Warbler	<i>Oreothlypis ruficapilla</i>	14,703	0.3345	0.1531	0.6494
Connecticut Warbler	<i>Oporornis agilis</i>	1,677	0.4305	0.2018	0.5186
Mourning Warbler	<i>Geothlypis philadelphia</i>	6,558	0.2999	0.1536	0.7048
Common Yellowthroat	<i>Geothlypis trichas</i>	7,698	0.3092	0.1588	0.6961
American Redstart	<i>Setophaga ruticilla</i>	7,266	0.3108	0.1432	0.6719
Cape May Warbler	<i>Setophaga tigrina</i>	1,905	0.2760	0.1020	0.7063
Northern Parula	<i>Setophaga americana</i>	3,095	0.2363	0.0936	0.7782
Magnolia Warbler	<i>Setophaga magnolia</i>	9,264	0.2796	0.1291	0.6952
Bay-breasted Warbler	<i>Setophaga castanea</i>	2,561	0.2935	0.0688	0.6728
Blackburnian Warbler	<i>Setophaga fusca</i>	5,530	0.2694	0.1376	0.7494
Yellow Warbler	<i>Setophaga petechia</i>	3,678	0.3289	0.1985	0.6977
Chestnut-sided Warbler	<i>Setophaga pensylvanica</i>	10,324	0.2968	0.1583	0.7196
Blackpoll Warbler	<i>Setophaga striata</i>	658	0.2628	0.0936	0.7093
Black-throated Blue Warbler	<i>Setophaga caerulescens</i>	1,125	0.2432	0.0916	0.7273
Palm Warbler	<i>Setophaga palmarum</i>	3,193	0.3360	0.1074	0.6114
Pine Warbler	<i>Setophaga pinus</i>	2,226	0.2386	0.1252	0.7923
Yellow-rumped Warbler	<i>Setophaga coronata</i>	24,714	0.2871	0.1506	0.7170
Townsend's Warbler	<i>Setophaga townsendi</i>	545	0.1235	0.1235	1.0000
Black-throated Green Warbler	<i>Setophaga virens</i>	8,422	0.2873	0.1400	0.7149
Canada Warbler	<i>Cardellina canadensis</i>	3,851	0.3001	0.1542	0.7009
Wilson's Warbler	<i>Cardellina pusilla</i>	2,127	0.2541	0.2065	0.8799
Scarlet Tanager	<i>Piranga olivacea</i>	3,276	0.1924	0.0881	0.8430
Western Tanager	<i>Piranga ludoviciana</i>	3,238	0.2895	0.1331	0.6775
Northern Cardinal	<i>Cardinalis cardinalis</i>	691	0.2976	0.0786	0.6895
Rose-breasted Grosbeak	<i>Pheucticus ludovicianus</i>	9,770	0.2555	0.1249	0.7578
Indigo Bunting	<i>Passerina cyanea</i>	1,217	0.2399	0.1140	0.7812



**APPENDIX TABLE 3.** Migratory status and best-supported models for probability of availability during point count surveys of terrestrial landbird species in northern North America, based on comparing the 3 classes of time-varying removal models (see Table 1 for model IDs). The column 'Best' considers all 3 classes (model ID is <class> <model ID>) and time adjustment values (PIF's [Partners in Flight] time adjustment:  $T_{adj}$ ; and inverse availability based measures:  $U_{inv}$ ). Migratory status: WR = winter resident; SD = short-distance migrant; LD = long-distance migrant. Model classes:  $M_0^{\phi}$  = conventional with time-varying rate parameter;  $M_f^{\phi}$  = finite mixture with time-varying rate parameter; and  $M_f^c$  = finite mixture with time-varying proportion of infrequent singers.

Common name	Migratory status	Best $M_0^{\phi}$	Best $M_f^{\phi}$	Best $M_f^c$	Best: All	$T_{adj}$	$U_{inv}$
Ruffed Grouse	WR	6	13	7	$M_f^c$ 7	N/A	1.89
Willow Ptarmigan	WR	8	7	0	$M_f^{\phi}$ 7	1.27	89.69
Wild Turkey	SD	7	13	13	$M_f^c$ 13	1.27	3.20
Mourning Dove	SD	11	11	11	$M_f^c$ 11	1.06	2.57
Yellow-billed Cuckoo	LD	0	0	0	$M_0^{\phi}$ 0	1.27	2.69
Black-billed Cuckoo	LD	0	1	0	$M_f^{\phi}$ 1	1.16	5.45
Ruby-throated Hummingbird	LD	10	3	0	$M_f^{\phi}$ 3	1.27	39.43
Killdeer	SD	10	3	10	$M_f^c$ 10	N/A	2.59
Upland Sandpiper	LD	1	10	1	$M_f^c$ 1	N/A	6.62
Long-billed Curlew	SD	11	11	11	$M_f^c$ 11	N/A	2.23
Marbled Godwit	SD	13	13	4	$M_f^{\phi}$ 13	N/A	2.93
Rock Sandpiper	SD	9	9	9	$M_f^c$ 9	N/A	1.28
Wilson's Snipe	LD	6	2	6	$M_f^c$ 6	N/A	2.10
Spotted Sandpiper	LD	3	0	4	$M_f^c$ 4	N/A	3.20
Solitary Sandpiper	LD	4	13	4	$M_f^c$ 4	N/A	10.57
Lesser Yellowlegs	SD	7	7	0	$M_0^{\phi}$ 7	N/A	2.57
Willet	LD	10	9	4	$M_f^{\phi}$ 9	N/A	2.40
Greater Yellowlegs	SD	6	7	3	$M_f^{\phi}$ 7	N/A	8.89
Belted Kingfisher	SD	0	N/A	N/A	N/A	1.27	5.55
Red-bellied Woodpecker	WR	1	N/A	N/A	N/A	2.25	3.43
Yellow-bellied Sapsucker	SD	6	6	7	$M_f^c$ 7	N/A	3.24
Downy Woodpecker	WR	4	13	4	$M_f^{\phi}$ 13	11.62	27.08
Hairy Woodpecker	WR	0	1	0	$M_f^{\phi}$ 1	1.21	4.19
American Three-toed Woodpecker	WR	9	9	13	$M_f^c$ 13	1.14	2.72
Black-backed Woodpecker	WR	10	10	10	$M_f^{\phi}$ 10	2.54	2.67
Northern Flicker	SD	1	7	7	$M_f^c$ 7	1.45	4.05
Pileated Woodpecker	WR	6	5	12	$M_f^{\phi}$ 5	1.06	5.12
Olive-sided Flycatcher	LD	10	4	13	$M_f^c$ 13	1.13	1.83
Western Wood-Pewee	LD	2	5	0	$M_f^{\phi}$ 5	1.27	2.16
Eastern Wood-Pewee	LD	12	12	12	$M_f^c$ 12	1.92	1.88
Yellow-bellied Flycatcher	LD	6	7	3	$M_f^c$ 3	N/A	1.97
Alder Flycatcher	LD	8	8	7	$M_f^c$ 7	25.88	1.71
Least Flycatcher	LD	13	13	13	$M_f^c$ 13	1.61	1.49
Hammond's Flycatcher	LD	3	3	0	$M_0^{\phi}$ 3	1.60	1.93
Eastern Phoebe	SD	2	2	2	$M_f^c$ 2	1.19	4.66
Great Crested Flycatcher	LD	1	1	1	$M_f^c$ 1	1.34	3.58
Eastern Kingbird	LD	4	12	13	$M_f^c$ 13	1.19	1.64
Yellow-throated Vireo	LD	3	1	9	$M_f^c$ 9	1.27	3.15
Blue-headed Vireo	SD	7	5	7	$M_f^c$ 7	1.27	2.52
Philadelphia Vireo	LD	12	2	2	$M_f^{\phi}$ 2	1.27	1.65
Warbling Vireo	LD	3	10	3	$M_f^c$ 3	1.27	1.82
Red-eyed Vireo	LD	11	7	11	$M_f^c$ 11	1.27	1.55
Gray Jay	WR	1	14	7	$M_f^c$ 7	1.27	3.47
Blue Jay	WR	13	4	13	$M_f^c$ 13	1.42	3.36
Black-billed Magpie	WR	4	9	4	$M_f^c$ 4	1.40	2.86
American Crow	SD	11	7	11	$M_f^c$ 11	1.07	2.09
Common Raven	WR	13	13	4	$M_f^c$ 4	1.42	2.78
Horned Lark	SD	13	3	7	$M_f^c$ 7	1.37	1.70
Tree Swallow	LD	13	5	4	$M_f^c$ 4	N/A	2.93
Bank Swallow	LD	2	10	10	$M_f^c$ 10	1.47	2.41
Cliff Swallow	LD	7	N/A	N/A	N/A	25.88	2.03
Barn Swallow	LD	3	0	3	$M_0^{\phi}$ 3	1.91	3.30
Black-capped Chickadee	WR	9	11	13	$M_f^c$ 13	1.60	3.00
Boreal Chickadee	WR	14	4	4	$M_f^c$ 4	1.24	3.83
Red-breasted Nuthatch	WR	2	13	13	$M_f^c$ 13	1.48	3.49
White-breasted Nuthatch	WR	12	12	10	$M_f^c$ 10	1.27	8.38

APPENDIX TABLE 3. Continued.

Common name	Migratory status	Best $M_0^{\phi}$	Best $M_f^{\phi}$	Best $M_f^c$	Best: All	$T_{adj}$	$U_{inv}$
Brown Creeper	WR	3	1	10	$M_f^c$ 10	1.13	2.76
House Wren	LD	8	13	13	$M_f^c$ 13	1.52	1.65
Winter Wren	SD	12	11	12	$M_f^c$ 12	1.51	1.62
Sedge Wren	SD	12	9	10	$M_f^c$ 10	1.27	1.89
Marsh Wren	SD	6	2	2	$M_f^{\phi}$ 2	1.13	2.21
Blue-gray Gnatcatcher	SD	2	4	2	$M_f^{\phi}$ 4	1.14	2.48
Golden-crowned Kinglet	SD	7	1	13	$M_f^c$ 13	1.36	2.59
Ruby-crowned Kinglet	SD	6	8	3	$M_f^c$ 3	1.27	1.94
Arctic Warbler	WR	0	4	0	$M_f^{\phi}$ 4	1.34	6.05
Eastern Bluebird	SD	7	1	7	$M_f^c$ 7	1.10	6,788,866.26
Townsend's Solitaire	SD	2	2	0	$M_0^{\phi}$ 2	1.27	1.50
Veery	LD	5	11	5	$M_f^c$ 5	1.27	2.00
Gray-cheeked Thrush	LD	0	1	0	$M_f^{\phi}$ 1	1.15	2.63
Swainson's Thrush	LD	14	13	14	$M_f^c$ 14	N/A	1.70
Hermit Thrush	SD	13	4	4	$M_f^c$ 4	1.06	1.79
Wood Thrush	LD	7	4	7	$M_f^{\phi}$ 4	1.27	1.56
American Robin	SD	7	7	8	$M_f^c$ 8	1.19	2.27
Varied Thrush	SD	7	7	0	$M_0^{\phi}$ 7	1.89	1.92
Gray Catbird	SD	7	5	4	$M_f^{\phi}$ 5	1.14	2.62
Brown Thrasher	SD	3	0	3	$M_f^c$ 3	1.88	2.46
European Starling	SD	7	7	7	$M_f^c$ 7	1.23	2.11
Bohemian Waxwing	WR	1	1	0	$M_0^{\phi}$ 1	2.47	5.98
Cedar Waxwing	SD	10	3	2	$M_f^{\phi}$ 3	1.33	5.61
House Sparrow	WR	12	0	2	$M_f^c$ 2	1.67	1.36
American Pipit	LD	3	10	0	$M_f^{\phi}$ 10	1.19	1.56
Sprague's Pipit	SD	12	1	1	$M_f^{\phi}$ 1	1.27	1.41
Evening Grosbeak	WR	13	13	4	$M_0^{\phi}$ 13	1.19	3.37
Pine Grosbeak	WR	9	N/A	N/A	N/A	1.16	3.89
Purple Finch	SD	7	4	7	$M_f^c$ 7	1.14	3.79
Red Crossbill	WR	12	10	10	$M_f^c$ 10	1.29	2.23
White-winged Crossbill	WR	8	7	8	$M_f^c$ 8	1.75	7.43
Pine Siskin	SD	14	12	4	$M_f^c$ 4	1.10	5.46
American Goldfinch	SD	1	1	0	$M_f^{\phi}$ 1	1.71	3.97
Lapland Longspur	SD	13	13	4	$M_f^{\phi}$ 13	1.57	3.37
Chestnut-collared Longspur	SD	11	2	7	$M_f^{\phi}$ 2	1.12	1.36
McCown's Longspur	SD	11	5	5	$M_f^{\phi}$ 5	1.15	3.35
Eastern Towhee	SD	10	2	9	$M_f^c$ 9	1.32	2.26
American Tree Sparrow	SD	11	11	0	$M_0^{\phi}$ 11	1.27	2.18
Chipping Sparrow	SD	13	4	11	$M_f^c$ 11	1.68	2.08
Clay-colored Sparrow	LD	13	7	13	$M_f^c$ 13	1.21	1.59
Field Sparrow	SD	4	0	4	$M_f^c$ 4	1.77	3.94
Vesper Sparrow	SD	6	5	5	$M_f^c$ 5	N/A	1.47
Savannah Sparrow	SD	14	7	13	$M_f^c$ 13	N/A	1.38
Grasshopper Sparrow	SD	4	4	4	$M_f^{\phi}$ 4	1.13	1.70
Baird's Sparrow	SD	4	4	4	$M_f^c$ 4	3.36	1.48
LeConte's Sparrow	SD	14	10	10	$M_f^c$ 10	1.33	1.38
Nelson's Sparrow	LD	2	2	2	$M_f^{\phi}$ 2	1.45	2.57
Fox Sparrow	SD	8	7	7	$M_0^{\phi}$ 8	2.46	2.41
Song Sparrow	SD	2	5	2	$M_f^{\phi}$ 5	N/A	1.84
Lincoln's Sparrow	SD	7	7	7	$M_f^c$ 7	2.10	1.97
Swamp Sparrow	SD	6	11	12	$M_f^c$ 12	1.27	2.23
White-throated Sparrow	SD	14	14	14	$M_f^c$ 14	1.18	1.87
White-crowned Sparrow	SD	12	12	4	$M_f^c$ 4	1.27	1.73
Golden-crowned Sparrow	SD	6	3	0	$M_0^{\phi}$ 6	19.50	6.51
Dark-eyed Junco	SD	13	13	13	$M_f^c$ 13	1.27	2.15
Yellow-headed Blackbird	LD	1	3	10	$M_f^{\phi}$ 3	1.50	1.34
Western Meadowlark	SD	12	4	4	$M_f^c$ 4	1.27	1.25
Baltimore Oriole	LD	0	9	4	$M_f^{\phi}$ 9	1.03	3.45
Red-winged Blackbird	SD	13	7	13	$M_f^c$ 13	N/A	1.75
Brown-headed Cowbird	SD	13	13	13	$M_f^c$ 13	1.16	2.83
Rusty Blackbird	SD	3	4	7	$M_f^{\phi}$ 4	N/A	10,022.05

APPENDIX TABLE 3. Continued.

Common name	Migratory status	Best $M_0^{\text{p}}$	Best $M_f^{\text{p}}$	Best $M_f^c$	Best: All	$T_{adj}$	$U_{inv}$
Brewer's Blackbird	SD	12	N/A	N/A	N/A	1.71	1.75
Common Grackle	SD	3	1	3	$M_f^{\text{p}}$ 1	1.17	2.72
Ovenbird	LD	7	7	7	$M_f^c$ 7	1.12	1.50
Northern Waterthrush	LD	7	8	7	$M_f^c$ 7	1.21	2.32
Golden-winged Warbler	LD	4	0	4	$M_f^c$ 4	1.53	3.50
Blue-winged Warbler	LD	3	2	0	$M_f^{\text{p}}$ 2	1.41	6.05
Black-and-white Warbler	LD	7	13	13	$M_f^c$ 13	1.27	2.18
Tennessee Warbler	LD	6	7	7	$M_f^c$ 7	1.27	1.49
Orange-crowned Warbler	SD	6	8	0	$M_f^{\text{p}}$ 8	1.83	1.85
Nashville Warbler	LD	7	9	8	$M_f^c$ 8	1.58	1.77
Connecticut Warbler	LD	10	10	10	$M_f^c$ 10	1.15	1.41
Mourning Warbler	LD	6	12	11	$M_f^{\text{p}}$ 12	1.14	1.82
Common Yellowthroat	SD	4	2	4	$M_f^c$ 4	1.18	1.78
American Redstart	LD	13	10	7	$M_f^c$ 7	1.27	1.82
Cape May Warbler	LD	6	2	11	$M_f^c$ 11	1.09	2.23
Northern Parula	LD	3	4	5	$M_f^{\text{p}}$ 4	1.07	2.66
Magnolia Warbler	LD	5	5	5	$M_f^c$ 5	1.51	1.96
Bay-breasted Warbler	LD	5	2	2	$M_f^c$ 2	1.72	2.31
Blackburnian Warbler	LD	12	2	10	$M_f^c$ 10	1.42	2.02
Yellow Warbler	LD	7	7	8	$M_f^c$ 8	10.00	1.88
Chestnut-sided Warbler	LD	4	4	4	$M_f^c$ 4	1.15	1.88
Blackpoll Warbler	LD	7	13	7	$M_f^c$ 7	1.27	2.50
Black-throated Blue Warbler	LD	0	9	10	$M_f^{\text{p}}$ 9	1.17	2.47
Palm Warbler	LD	7	13	13	$M_f^c$ 13	1.74	1.90
Pine Warbler	SD	4	4	4	$M_f^{\text{p}}$ 4	2.32	1.92
Yellow-rumped Warbler	SD	7	7	7	$M_f^c$ 7	1.66	1.97
Townsend's Warbler	LD	6	14	0	$M_0^{\text{p}}$ 6	N/A	2.22
Black-throated Green Warbler	LD	5	7	7	$M_f^c$ 7	1.14	1.87
Canada Warbler	LD	7	2	7	$M_f^c$ 7	1.29	1.77
Wilson's Warbler	LD	13	13	13	$M_f^c$ 13	1.27	1.92
Scarlet Tanager	LD	4	0	4	$M_f^c$ 4	20.00	2.40
Western Tanager	LD	5	2	12	$M_f^c$ 12	1.27	1.99
Northern Cardinal	WR	2	2	2	$M_f^c$ 2	1.60	2.24
Rose-breasted Grosbeak	LD	14	13	10	$M_f^c$ 10	1.27	2.10
Indigo Bunting	LD	4	5	4	$M_f^c$ 4	1.53	1.86