## Introduction

North America's boreal forest supports billions of breeding birds from more than 300 species (Niemi et al. 1998). Increasing industrial development in this region (Hobson et al. 2013; Mahon et al. 2014), changing forest dynamics due to rapid rates of climate change (Stralberg et al. 2015), and numerous pressures during non-breeding periods (i.e., migration, wintering; Kirby et al. 2008) have led to concerns over the status of boreal avian populations. Recent studies suggest that boreal birds appear to have experienced among the steepest population declines of any avian group, owing to large declines of several previously abundant and widespread species (Rosenberg et al. 2016, 2019). There is therefore an urgent need to improve avian monitoring in the boreal forest (Cumming et al. 2010).

Population trends for most North American landbirds are derived from the North American Breeding Bird Survey (BBS), but this roadside survey has limited coverage in the mostly roadless core of the boreal zone. Consequently, the BBS samples a biased subset of boreal habitat (Van Wilgenburg et al. 2015), potentially leading to unrepresentative or biased trend estimates for populations of boreal bird species (Machtans et al. 2014). Many boreal-breeding species migrate to neotropical regions that are not adequately monitored by nonbreeding (i.e., 'wintering') surveys such as the Christmas Bird Count (CBC). Thus, while there are substantial data with which to estimate population status in select regions of northern forest, range-wide trends of most boreal species are lacking (Dunn et al. 2005).

There has long been interest in using standardized counts of migrating birds to evaluate population change for species that are not well-monitored during the non-migratory part of their life cycle (Francis and Hussell 1998; Dunn et al. 2005). Counting migrants for use in population monitoring is one goal of The Canadian Migration Monitoring Network, a collaborative initiative of bird observatories across Canada, Birds Canada, and Environment and Climate Change Canada (Canadian Migration Monitoring Network 2021). Similar migration count data are also collected by several long-running migration monitoring stations in the United States. However, migration monitoring data have typically been analyzed on a station-by-station basis due to lack of knowledge of the breeding origins of migrants passing each count site. This gap in knowledge precludes appropriate weighting of site-specific trends in combined analyses to derive regional and/or range-wide trends, while also hampering appropriate targeting of conservation actions.

Advances in probabilistic origin assignments using stable hydrogen ratios in feathers, have led to cost-efficient methods for broadly estimating the breeding ground origins of birds captured at migration count sites (Van Wilgenburg and Hobson 2011, Wunder 2012). The ability to assign birds to geographic

regions where they molted feathers provides a powerful opportunity to appropriately weight information across a network of migration monitoring stations and to estimate large-scale patterns in population trends (Crewe et al. 2016a). This is particularly relevant to boreal-breeding birds that are not well-monitored by other methods, thereby delivering enhanced information for conservation decision-making.

Blackpoll Warbler (*Setophaga striata*) is an abundant Nearctic-Neotropical migrant landbird species that breeds mostly in the northern boreal forest. Trends derived from roadside BBS surveys suggest populations have declined by over 90% since 1970; among the steepest declines of any landbird over that period (Rosenberg et al. 2016, Sauer et al. 2020). However, these trends estimates are potentially unreliable owing to the absence of BBS coverage throughout the core of the species' breeding range (Environment and Climate Change Canada 2019). Migration monitoring provides an attractive alternative method for tracking population status in boreal species, because populations from all parts of the breeding range pass through southern Canada and the eastern United States *en route* to and from their neotropical non-breeding areas.

Here, we develop a Bayesian integrated population model to estimate regional and national population trends using standardized daily counts of migrants from a series of monitoring stations, combined with estimates of the proportions of individuals coming from two geographic strata based on feather stable isotope assignments. Our method synthesizes trend information at a range-wide scale by weighting information across an international network of migration monitoring stations, thereby moving beyond station-by-station analyses. We apply this analysis to Blackpoll Warbler and compare the resulting trend estimates to those from conventional Breeding Bird Survey analysis. We also describe how this analytical framework can be applied to other migratory boreal species to generate estimates of range-wide population trends. We provide fully documented R code and recommendations for expanding this method to other avian species.

## Methods

### *Description of statistical model*

We developed a hierarchical model to estimate temporal (i.e. multi-decadal) patterns of population change within discrete geographic strata from which birds arriving at migration monitoring stations originated. The model simultaneously estimates 1) annual indices of bird abundance at each

migration monitoring station, 2) the proportion of birds at each station that have arrived from each stratum within each year, and 3) population trends within each stratum. We fit models separately to data collected during pre-breeding migration (i.e., the northward migration of breeding birds during North American spring season) and post-breeding migration (i.e., the southward migration of adult and newly fledged juvenile birds during the North American fall season). Equations and priors underlying the statistical model are included in Tables 1 and 2, and we describe logic of these equations below.

Quantities in the model are indexed by geographic stratum (*j*), year (*y*), monitoring station (*s*), and day of year (*d*). The highest level of the model (equation 1) describes the temporal pattern of population change in each geographic stratum *j* starting from a baseline year $y_0$. Our model assumes that abundance within each stratum ($X_{j,y}$) changes according to a log-linear trend.

The next level of the model (equation 2) describes an index of the expected number of migrants ($M_{j,s,y}$) arriving from each geographic stratum to each migration monitoring station during each year. $M_{j,s,y}$ is modeled as a product of annual stratum abundances ($X_{j,y}$) and station-level migration parameters ($\rho_{j,s}$) that describe the contribution of stratum *j* to station *s* in year *y*. The parameters $\rho_{j,s}$ are station-specific "capture rates" that convert the indices of abundance in each stratum to number of birds arriving at each station in each year. In cases where a migration monitoring station is known (or assumed) to exclusively capture birds from a subset of strata, the relevant parameters ($\rho_{j,s}$) can be fixed to zero for the strata that are not monitored by the station. We note that stratum-level intercept terms in the model ($X_{j,y_0}$), are not estimable with migration data alone because infinite combinations of $X_{j,y_0}$ and migration parameters ($\rho_{j,s}$) would be equally consistent with the observed data. Thus, we fix $X_{j,y_0} = 1$ such that the model estimates change relative to the start year within each stratum, ensuring all parameters are theoretically identifiable. This also implies the model cannot estimate relative abundances between strata and is only capable of estimating changes in relative abundance within each stratum. However, independent estimates of stratum relative abundances can be used to rescale $X_{j,y}$ in each stratum and calculate range-wide trends (see below).

We model an index of annual abundance at each station ($T_{s,y}$) as the sum of $M_{j,s,y}$ across $J$ strata multiplied by a yearly station-specific random effect ($\varepsilon_{s,y}$). The term $\varepsilon_{s,y}$ in equation 3 acknowledges that there is annual variation in the total number of migrants arriving at each station beyond that which is attributable to changes in $X_{j,y}$ (e.g., additional variation could be driven by year-to-year fluctuations in migration pathways).

The third level of the model (equation 4) distributes the $T_{s,y}$ migrants arriving at the monitoring station among days of the season. This component of the model is necessary because some monitoring stations are only operational for a subset of days per season; the model can therefore accommodate missing data within a season. Migration of individuals past monitoring stations is assumed to follow a symmetric seasonal pattern around a peak date. We therefore described the seasonal pattern of counts at each station using a normal probability density function $f(d, \mu_s, \sigma_s)$ that integrates to 1 across a season. The parameter $\mu_s$ is the ordinal date of the seasonal peak of migration at station $s$, while $\sigma_s$ describes the temporal dispersion of the migration period at that station (i.e., approximately 95% of the station's migration period occurs within $1.96\sigma_s$ on either side of $\mu_s$).

The final level of the model describes the observed daily count data (equation 5) and breeding origin estimates (equation 6). To describe migration count data, the number of birds counted on each day of the season ($d$) at each station ($s$) in each year ($y$) was modeled as an over-dispersed Poisson process with median equal to $T_{s,y} \times f(d, \mu_s, \sigma_s)$. We assumed log-normal variance for unexplained 'noise' in daily counts at each station (e.g., owing to weather conditions that affect daily migration behavior). Equation 5 also includes an offset equal to log(net hours) to account for spatio-temporal variation in monitoring effort. This also implies that $T_{s,y}$ can be compared to empirical count data (e.g., for goodness-of-fit evaluation) by dividing daily counts by effort and summing those quantities across a season.

We use a multinomial distribution to model breeding origins of migrants at the stations where they were collected (equation 6). In this equation, $\boldsymbol{Y_{s,y}}$ is a vector (denoted by bolded font) containing the number of sampled birds assigned to each of the $J$ strata at a station in a year. The vector of probabilities describing the multinomial distribution (i.e., $\frac{M_{j,s,y}}{\sum_1^J M_{j,s,y}}$ ) allow station-level dynamics to be linked to stratum-level dynamics.

Estimates of percent population change in a stratum relative to a baseline year (e.g., in this study, relative to the year 1998) can be calculated as:

$$PercentChange_{j,y_{end}} = 100 \times \frac{X_{j,y_{end}} - X_{j,y_0}}{X_{j,y_0}}, \quad (7)$$

where $X_{j,y_{end}}$ is an index of abundance in the final year ($y_{end}$). Following the North American BBS (Smith et al. 2014), we defined population trend as the geometric mean rate of change between two points in time, where:

$$Trend_j = 100 \times \left( \left( \frac{X_{j,y_{end}}}{X_{j,y_0}} \right)^{\frac{1}{y_{end}-y_0}} - 1 \right). \quad (8)$$

120

121 *Range-wide Trend Estimation*

122 Since the model requires that stratum-level intercept terms are fixed to a constant (such that
123 $\alpha_{0,j}$ and $\rho_{j,s}$ are identifiable), independent estimates of relative abundance in each stratum are required
124 to re-scale estimates of $X_{j,y}$ and thereby appropriately weight changes in abundance within each
125 stratum at a range-wide scale. This re-scaling of $X_{j,y}$ is necessary because a large positive trend in a
126 stratum that contains almost no birds will have a small effect on the continental trend compared to a
127 small negative trend in a stratum that contains a large majority of the population.

128 *Application to Blackpoll Warbler: data and model parameters*

129 We applied our statistical model to time series of standardized daily migration counts of
130 Blackpoll Warblers collected from 1998-2018 at permanent monitoring locations across Canada and the
131 United States (Figure 1, station-specific data in Appendix 1). Canadian sites are members of the
132 Canadian Migration Monitoring Network, which use station-specific standards to collect avian counts
133 during the pre-breeding (North American spring) and/or post-breeding (North American fall) migration
134 seasons. Count protocols differ among stations, but each site attempts to maintain relatively
135 standardized protocols and effort over time (Canadian Migration Monitoring Network 2021). The most
136 used count approaches include banding captures using mist nets and visual counts that record all birds
137 detected in a specified area during a specified time. Banding at most sites is for 6 hours starting a half-
138 hour pre-dawn at fixed net locations. The length of the daily survey period ranges from one hour (e.g.,
139 when counting birds along a fixed route) to more than six hours (e.g., continuous counts from a fixed
140 point). Data from U.S. bird observatories consist of daily captures with mist-nets, usually with fixed
141 locations and operated for similar hours as at Canadian sites for least 5 days per week during spring
142 and/or fall migrations. Because some U.S. stations had variable daily hours and/or number of nets, daily
143 number of net-hours was used as a statistical offset in analyses. Daily counts for Blackpoll Warbler were
144 available for 13 monitoring sites during pre-breeding migration and 18 sites in post-breeding migration
145 (Figure 1, Appendix 1). Data from each site were restricted to dates that were sampled in at least two-
146 thirds of years in which the station operated. Within those date limits, daily samples had to have been
147 made during at least 75% of days within local migration season for blackpolls. Specific days with <50

total net hours were omitted from analyses. For sites with notable changes in count effort between 1998 and 2018, data were limited to the years within which coverage was standardized.

We split the boreal breeding range of the Blackpoll Warbler into two geographic strata: 'West' and 'East' (Figure 1), based on feather stable hydrogen isotope analysis and other information in Dunn et al. (2023). Those authors defined three strata, but we combined the two covering all areas west of the Great Lakes into our single West stratum due to sample size considerations, and to ensure migrants could be assigned to each stratum with higher assignment accuracy. Breeding ground origins of individual Blackpoll Warbler can be confidently assigned to one or the other stratum based on stable hydrogen isotope ratios in their tail feathers ($\delta^2H_f$; Dunn et al. 2023). Boundaries of these strata are also consistent with *a priori* knowledge of Blackpoll Warbler migration routes, based on banding, geolocator, and additional stable isotope studies (DeLuca et al. 2015, 2019; Holberton et al. 2015; Morris et al. 2016; Covino et al. 2020).

We used the feather isotope results from Dunn et al. (2023) as data in our statistical population model to indicate the proportions of migrants at each capture site that originated from East and/or West strata in years where data were available. Stable hydrogen isotope ratios of feathers grown in parts of Alaska are isotopically similar to those expected in parts of the East stratum. However, because migrants captured west of Ontario can safely be assumed to have originated from the West stratum we fixed the West proportion for such stations at 100% without regard to isotope results. Elsewhere, we assigned values based on isotope results from within 100 km for which feathers had been collected and analyzed, though several sites had no data for estimating West *versus* East proportions of migrants. The values assigned to each site in each season and the basis for each selection are shown in Appendix 1.

We fit the statistical model in a Bayesian framework using JAGS version 4.3.0 (Plummer 2003), interfaced with the R programming language version 4.0.2 (R Core Team 2024) using the jagsUI library version 1.5.2 (Kellner 2021). We specified vague priors on most model parameters (Table 2) except $\sigma_s$ where we specified priors to facilitate better model convergence that were moderately informative, based on *a priori* knowledge of the typical "migration period" at stations. After a burn-in of 5,000 iterations, we stored every 50$^{th}$ iteration until we accumulated 2,000 posterior samples from each of three MCMC chains. We assessed chain convergence by visual examination of MCMC traceplots and by evaluating that the Gelman–Rubin convergence statistic was close to 1 for all model parameters. We also confirmed that effective sample sizes for monitored parameters were large enough (>1,000) to adequately characterize posterior distributions.

To calculate national trends for Blackpoll Warbler, we required independent estimates of the relative population abundances among strata to re-weight stratum-level trends at a continental scale. For this purpose, we used spatial abundance maps from the Boreal Avian Modeling Project, produced using boosted regression trees fit to a large dataset of point counts across North America that correct for variation in protocols among surveys (Solymos et al. 2013; Stralberg et al. 2015). The raster presents estimates of relative density in 1 km x 1 km pixels, as a prediction for the year 2011. We cropped the relative abundance raster to the boundary of each stratum, and summed pixel values to yield estimates of relative abundance. This map indicated that Blackpoll Warbler was 1.99 times more abundant in the West stratum than in the East and represents an estimate for the year 2011 (Appendix 2). We therefore re-scaled estimates of $X_{j,y}$ in each of the $j$ strata based on these values, using $\frac{X_{j,y}}{X_{j,2011}} \times RelAbund_{j,2011}$, where $RelAbund_{j,2011}$ was 1.99 and 1 for the West and East, respectively. To evaluate the sensitivity of national trend estimates to the choice of relative abundance raster, we also repeated this exercise for a raster obtained from the community science project eBird (Fink et al. 2023; Appendix 2), though we note that this raster is flagged by eBird as "low reliability" because estimates across large portions of the boreal range are extrapolations from distant and unrepresentative data points.

We assessed goodness-of-fit of model predictions by evaluating the correlations between observed seasonal totals at monitoring stations and the expected seasonal totals based on the fitted models (Appendix 3). We also conducted posterior predictive checks to confirm that the distribution of simulated counts based on the fitted statistical model were consistent with the distribution of observed counts at each station (Appendix 3). Using simulation, we also evaluated whether the migration model was able to generate unbiased estimates of regional population trends across a wide range of regional trend and station-level capture rate scenarios (Appendix 3).

*Comparison to trend estimates from the North American Breeding Bird Survey*

We compared regional trend estimates from our model to those derived from the North American Breeding Bird survey to evaluate similarities and differences between conventional breeding season analyses and our migration analysis. We fit a Bayesian hierarchical model to BBS time series from 1998 to 2018 (the same period as our migration analysis), using analytical strata implemented by the United States Geological Survey (USGS) for continental analysis. We specified a 'first difference' population process model (Link et al. 2017; Smith and Edwards 2021), which is widely used in standard

continental analysis of BBS data. We fit the model and extracted output using the 'bbsBayes' package in R (Edwards and Smith 2020), specifying a 50,000 iteration burn-in period, after which we stored every 100$^{th}$ posterior sample until we accumulated 2,000 posterior samples from each of 3 MCMC chains.

To derive regional trend estimates from the BBS, we assigned BBS analytical strata into East and West categories based on geographic overlap with the strata we used for migration analysis (illustrated in Appendix 4). This allowed us to calculate 'post hoc' synthetic estimates of regional population trends by calculating annual BBS indices from the fitted model across analytical strata that overlapped with the coarse East and West strata used for the migration monitoring analysis. Detailed methods for estimating population trajectories and trends within custom strata are described in the bbsBayes package (Edwards and Smith 2020).

**Results**

Simulations confirmed that the statistical model was capable of producing regional trend estimates that are identifiable, unbiased, and with appropriate credible interval coverage across a wide range of simulated population trends, and in situations where capture rates ($\rho_{j,s}$) are highly variable between stations, and migration counts are highly variable among years and among days within a season at each station (Appendix 3). Trends were also recoverable with highly incomplete isotope data, even when no stations were known *a priori* to exclusively monitor a single stratum.

Models unambiguously converged for empirical analysis of Blackpoll Warbler migration data in both pre-breeding and post-breeding seasons. Goodness-of-fit metrics for analysis of Blackpoll Warbler indicated a reasonable fit to the empirical data for most stations (Appendix 2), though posterior predictive checks suggested a poor fit for several station-year combinations due, owing to overdispersion in daily counts.

For Blackpoll Warbler populations breeding in the West stratum, migration monitoring in both seasons indicated similar population trends (Table 3, Figure 2). We detected strong evidence of population increases in the western stratum based on both pre-breeding migration and post-breeding migration (>0.99 probability of positive trend). In contrast, analysis of the BBS indicated a high probability that western populations had declined over the same period (0.97 probability of negative trend).

For the East stratum population, pre-breeding migration trends showed strong evidence of large declines (<0.01 probability of positive trend), with a median trend estimate of -4.6% per year leading to a decline of greater than 40% between 1998 and 2018. BBS also suggested there was a high probability (0.97) the trend was negative, and the magnitude of the BBS trend estimate was similar to that from pre-breeding migration monitoring (Figure 2, Table 3). The eastern population trend estimated from post-breeding migration was extremely imprecise because few stations captured exclusively eastern migrants during post-breeding migration (Appendix 1; Figure S1.4), and signals of eastern population changes were largely swamped by migrants originating from the western stratum.

Estimates of continental population trends depended on the source of relative abundance estimates among strata that were used to weight trend estimates in each stratum in a continental context. Figure 2 and Table 3 illustrate continental trend estimates that assume the population in the West stratum was 1.99 times larger than the East stratum in 2011, based on a population density raster obtained from Stralberg et al. (2015). If this relative abundance was accurate, the stratum-level trends we detected would have led to population increases in the West counter-balancing declines in the East, and the continental population would have remained approximately stable (based on pre-breeding migration) or have increased (based on post-breeding migration; Table 3). However, if relative abundance estimates from eBird are accurate, the continental population would have likely declined since 1998; though this source of relative abundance estimate is considered to be of low reliability (see Appendix 2; Fink et al. 2023). In contrast to migration monitoring, the BBS detected strong evidence of population declines in both strata, leading to strong evidence of continental declines (Table 3).

**Discussion**

Our study fulfills a longstanding need for North American landbird monitoring by facilitating large-scale population trend assessments for migratory species, particularly for the approximately 80 species that primarily breed in the core of the boreal forest (Dunn et al. 2005). Since many species are difficult to survey during the non-migratory periods of their life cycle, migration monitoring can be crucial in assessing the status of those species until other data become available. Our model synthesizes information from multiple sites (over 20 locations for the Blackpoll Warbler case study), and therefore represents an important step beyond analysis of migration trends on station-by-station basis (Canadian Migration Monitoring Network 2021). Variation in the number of migrants among stations is partitioned

by our model into contributions from differences in regional population trends, differences in station-level catchment, and multiple sources of unexplained spatial and temporal variation.

In general, confidence in a scientific finding is enhanced when multiple independent lines of inquiry converge on a shared conclusion through so-called 'methodological triangulation' (Heesen et al. 2019). For North American landbirds, there are now multiple survey programs from which population trends can potentially be estimated and compared. Large-scale, long-term, semi-structured community science programs such as the BBS and the Christmas Bird Count can provide valuable insights into population trends for species inhabiting the road-accessible portions of North America (Link et al. 2006; Sauer et al. 2020). Provincial and state breeding bird atlases (Dunn et al. 2008), along with a suite of other regional monitoring programs (e.g., Pavlacky et al. 2017; Hill 2023) can provide finer-scale inference in locations that may otherwise be data-deficient. eBird also provides estimates of local and regional population trends over the last 10 years across portions of North America, based on unstructured volunteer data analyzed with machine-learning algorithms designed to minimize biases inherent in opportunistically collected datasets (Fink et al. 2023). Range-wide migration monitoring, in combination with stable isotope analysis, can provide a powerful complement to these other surveys because it relies on a fundamentally different methodology: the capture of migrating birds using standardized protocols, assignment of those individuals to regional strata, and integrated analysis of data using hierarchical models. Migration monitoring has an additional advantage in that pre- and post-breeding migration seasons can be analyzed independently from each other (as in our current analysis), providing a further check on consistency in estimates. Consilience between population trend estimates from each migration season and other relevant breeding and non-breeding surveys could greatly strengthen confidence in species status and trend assessments. Conversely, lack of agreement may help identify weaknesses in a particular survey approach or suggest new avenues for research.

For Blackpoll Warbler, multiple lines of evidence indicate that severe population declines have recently occurred in eastern North America. Independent analysis of pre-breeding migration and BBS suggest that eastern populations have declined at rates of approximately -4% per year (Table 3). While trend estimates in the eastern stratum could not be precisely estimated from post-breeding migration, the analysis nevertheless suggested that declines were more likely to have occurred than population increases. Congruent with these findings, Mountain Birdwatch, which monitors montane breeding bird communities in the eastern United States, also estimated Blackpoll Warbler population declines of approximately -4.5% per year (80% CI = -5.2% to -3.8) from 2010 to 2023 (Hill 2023). Strong declines in

Blackpoll Warbler breeding occurrence were also detected in New Brunswick and Cape Breton Island, Nova Scotia, between the first (1990) and second (2010) Maritimes Breeding Bird Atlases (Stewart 2015). In contrast, eBird trend estimates suggest stable or increasing Blackpoll Warbler populations in eastern Canada between 2012 and 2022 (Fink et al. 2023); the cause of this discrepancy is unclear. Provincial breeding bird atlases currently underway in Ontario and Quebec are collecting boreal avian data using a nationally standardized framework (Van Wilgenburg et al. 2020) and are expected to provide critical insights into population changes that have occurred in eastern Canada since the early 2000s. Simultaneously, these efforts will provide another independent test of the ability of our model to track regional changes in population size.

Pre- and post-breeding migration analysis independently suggested that populations of Blackpoll Warbler west of Ontario, Canada have increased over a 20-year period. In contrast, the BBS suggested that populations have likely declined in this region over the same period. eBird also reports evidence of recent populations declines in western Canada and Alaska (Fink et al. 2023). However, data from BBS and eBird are heavily biased towards roadsides and human settlements and therefore sample an incomplete and biased portion of the western boreal forest (Machtans et al. 2014; Van Wilgenburg et al. 2015). BBS and eBird are also unable to estimate population trends in the Northwest Territories and Nunavut where long-term breeding season surveys are exceedingly sparse, but where a substantial proportion of the continental population of Blackpoll Warbler likely breeds (Appendix 2). The migration monitoring network likely captures substantial numbers of individuals from those regions, particularly during post-breeding migration when the entire continental population passes through eastern North America (Holberton et al. 2015). However, it is also possible that migration trends could be biased if sites monitoring “western” birds are poorly distributed such that the portion of the range they sample is unrepresentative of regional population dynamics.

While our model can estimate stratum-level trends, estimation of continental trends requires independent information describing the relative abundance of populations within each stratum to weight stratum trends appropriately in a broader context (see Methods; Appendix 2). Relative abundance maps based on breeding season surveys can be used for this purpose (e.g., Stralberg et al. 2015; Fink et al. 2023), but the reliability of those maps depends how representative the underlying data are spatially, temporally, and with respect to covariates that describe the species environmental niche. Unreliable abundance maps could lead to biased estimates of range-wide trends because stratum-level trends could be incorrectly weighted in a continental context. For example, the declines we detected in

the East stratum based on pre-breeding migration could either represent a massive continental decline (i.e., if based on eBird relative abundance maps), or could have been largely counter-balanced by population increases in the West (i.e., if based on density rasters derived from a large-scale boreal point count dataset; Stralberg et al. 2015). This emphasizes the critical need for rigorously designed, randomized surveys throughout the boreal zone to establish reliable baseline estimates of abundance (Van Wilgenburg et al. 2020), which will also greatly enhance the reliability of range-wide trend estimates from migration monitoring.

Our statistical model makes several additional assumptions that warrant consideration. First, migration counts at each station are driven by a complex product of both variation in regional population sizes and the probability that migrating individuals are observed at monitoring stations (Dunn 2005). Our model is unable to account for long-term directional changes in migration pathways and stopover behavior, which would impose spurious signals of population change. This occurs, for example, if migrants increasingly avoid particular monitoring stations when surrounding habitat changes considerably over time (Francis and Hussell 1998, Dunn 2005). However, network-wide trend analyses should be robust to this bias if many stations count birds from each region, reducing the influence of single stations on overall trend estimates. Additionally, site-specific migratory connectivity can be periodically reassessed to examine changes in the regions of the breeding range that are sampled by each station, as occurred for Blackpoll Warbler at some Great Lakes sites (Dunn et al. 2023). Yet, frequent analysis of feather isotopes across numerous stations could be cost or logistically prohibitive so alternative methods for estimating station catchment should also be considered (e.g., Meehan et al. 2022, Dunn et al. 2023). Second, migration counts are extremely stochastic and difficult to characterize parametrically (Crewe et al. 2016b). Unrealistic error distributions could result in poor goodness-of-fit metrics in some years at some stations (e.g., Appendix 3) with associated credible intervals on estimates that are too narrow. However, this phenomenon is not unique to our migration monitoring analysis and iterative improvements in model structure are almost always necessary for hierarchical analyses of long-term, large-scale monitoring data (Link et al. 2016). Third, the current structure of our population process model only includes a log-linear trend with each stratum and does not attempt to estimate random annual process variation in regional population sizes. This assumption may be justified for a wide-ranging, highly abundant species like Blackpoll Warbler, where large-scale regional trajectories derived from the BBS also resulted in relatively smooth trajectories (Figure 2). However, this would not be true for nomadic or irruptive populations. Use of our model for standardized hawk migration counts should consider that confidence intervals on trend estimates from our model for partial migrants, which

vary both in proportion of population migrating and in distance traveled, may be too narrow because random temporal variation in population dynamics is not fully accounted for. Future applications of our model our model could also consider attempting to fit more flexible temporal trajectory functions within strata to describe curvilinear changes in trajectories over time (Smith and Edwards 2021).

Application of our method to a larger number of boreal breeding species could yield novel insights into the population status of boreal-breeding migratory species. Our analytical approach could potentially be extended to other groups of birds that are more reliably covered on migration or at stopover sites, such as shorebirds (Smith et al. 2023) and raptors (Farmer et al. 2007) that are complete migrants, provided that the origins of birds can be estimated at migration survey locations. Our model could also potentially be extended to incorporate non-standardized migration count data from community science networks such as eBird, especially if there was a well-distributed roster of sites frequently-visited during migration seasons that could be targeted for daily coverage, conceptually adding thousands of migration monitoring 'stations' across the continent. This would have numerous advantages including more comprehensive coverage of migratory populations and reducing the influence of individual monitoring stations on regional trend estimates, but reliance on community science information requires careful screening of data and appropriate accounting of changes in observer effort over time (Fink et al. 2023). Continued efforts to implement and improve range-wide trend analyses from migration surveys holds considerable promise for North American landbird monitoring.

**Literature Cited**


Canadian Migration Monitoring Network. 2021. The Canadian Migration Monitoring Network - Réseau canadien de surveillance des migrations: Researching Canada's Landbirds for Twenty Years. *CMMN-RCSM Scientific Technical Report, 3.*

Covino, K., Morris, S., Shieldcastle, M., and Taylor, P., 2020. "Spring Migration of Blackpoll Warblers across North America." *Avian Conservation and Ecology 15* (1).

Crewe, T., Lepage, D., and Taylor, P., 2016a. Effect of sampling effort on bias and precision of trends in migration counts. The Condor, 118(1) pp.117-138.

Crewe, T., Taylor, P. and Lepage, D., 2016b. Temporal aggregation of migration counts can improve accuracy and precision of trends. *Avian Conservation and Ecology*, *11*(2).

Cumming, S.G., Bayne, E., Schmiegelow, F.K.A., Fontaine, T., Song, S.J. and Lefevre, K., 2010. Toward conservation of Canada's boreal forest avifauna: Design and application of ecological models at continental extents. Avian Conservation & Ecology*, 5(2).*

DeLuca, W.V., Woodworth, B.K., Rimmer, C.C., Marra, P.P., Taylor, P.D., McFarland, K.P., Mackenzie, S.A. and Norris, D.R., 2015. Transoceanic migration by a 12 g songbird. *Biology Letters*, *11*(4).

DeLuca, W. V., B. K. Woodworth, S. A. Mackenzie, A. E. M. Newman, H. A. Cooke, L. M. Phillips, N. E. Freeman, A. O. Sutton, L. Tauzer, and C. McIntyre. 2019. A boreal songbird's 20,000 km migration across North America and the Atlantic Ocean. Bulletin of the Ecological Society of America 100 (3): 1–5.

Dunn, A.M. and Weston, M.A., 2008. A review of terrestrial bird atlases of the world and their application. *Emu-Austral Ornithology*, *108*(1), pp.42-67.

Dunn, E.H., 2005. Counting migrants to monitor bird populations: State of the art. *USDA Forest Service General Technical Report PSW-GTR-191*, pp.712-717.

Dunn, E. H., B. L. Altman, J. Bart, C. J. Beardmore, H. Berlanga, P. J. Blancher, G. S. Butcher, D. W. Demarest, R. Dettmers, and W. C. Hunter. 2005. High Priority Needs for Range-Wide Monitoring of North American Landbirds. Partners in Flight Technical Series 2. Partners in Flight. https://www.partnersinflight.org/wp-content/uploads/2017/03/PIF-Technical-Series-02-Monitoring-Needs.pdf

Dunn, E.H., Kardynal, K.J., Covino, K.M., Morris, S.R., Holberton, R.L. and Hobson, K.A., 2023. Feather isotopes (δ 2 H f) and morphometrics reveal population-specific migration patterns of the Blackpoll Warbler (Setophaga striata). *Avian Conservation and Ecology*, *18*(2).

Edwards, B.P. and Smith, A.C., 2020. bbsBayes: An R package for hierarchical Bayesian analysis of North American breeding bird survey data. *Journal of Open Research Software 9(1).*

Environment and Climate Change Canada. 2019. "Blackpoll Warbler." Status of Birds in Canada. 2019. https://wildlife-species.canada.ca/bird-status/oiseau-bird eng.aspx?sY=2019&sL=e&sM=a&sB=BLPW.

Farmer, C.J., Hussell, D.J. and Mizrahi, D., 2007. Detecting population trends in migratory birds of prey. *The Auk*, *124*(3), pp.1047-1062.

Fink, D., T. Auer, A. Johnston, M. Strimas-Mackey, S. Ligocki, O. Robinson, W. Hochachka, L. Jaromczyk, C. Crowley, K. Dunham, A. Stillman, I. Davies, A. Rodewald, V. Ruiz-Gutierrez, C. Wood. 2023. eBird Status and Trends, Data Version: 2022; Released: 2023. *Cornell Lab of Ornithology, Ithaca, New York.* https://doi.org/10.2173/ebirdst.2022

Francis, C.M. and Hussell, D.J.T., 1998. Changes in numbers of land birds counted in migration at Long Point Bird Observatory, 1961-1997. *Bird Populations*, *4*, pp.37-66.

Heesen, R., Bright, L.K. and Zucker, A., 2019. Vindicating methodological triangulation. *Synthese*, *196*, pp. 3067-3081.

Hill, J.M., 2023. The State of the Mountain Birds Report: 2023. *Vermont Center for Ecostudies, White River Junction, VT.* https://mountainbirds.vtecostudies.org/. Accessed 1/15/2024.

Hobson, K.A., Wilson, A.G., Van Wilgenburg, S.L. and Bayne, E.M., 2013. An estimate of nest loss in Canada due to industrial forestry operations. *Avian Conservation & Ecology, 8(2).*

Hobson, K., Van Wilgenburg, S., Dunn, E., Hussell, D., Taylor, P. and Collister, D., 2015. Predicting origins of passerines migrating through Canadian migration monitoring stations using stable-hydrogen isotope analyses of feathers: a new tool for bird conservation. *Avian Conservation and Ecology*, *10(1).*

Holberton, Rebecca L., Steven L. Van Wilgenburg, Adrienne J. Leppold, and Keith A. Hobson, 2015. "Isotopic (Δ2Hf) Evidence of 'Loop Migration' and Use of the Gulf of Maine Flyway by Both Western and Eastern Breeding Populations of Blackpoll Warblers." *Journal of Field Ornithology 86 (3):* 213–28.

Kellner, K., 2021. jagsUI: a wrapper around rjags to streamline JAGS analyses. R package version 1.5.2.

Kirby, J.S., Stattersfield, A.J., Butchart, S.H., Evans, M.I., Grimmett, R.F., Jones, V.R., O'Sullivan, J., Tucker, G.M. and Newton, I., 2008. Key conservation issues for migratory land-and waterbird species on the world's major flyways. *Bird Conservation International*, *18*(S1), pp.S49-S73.

Link, W.A., Sauer, J.R. and Niven, D.K., 2006. A hierarchical model for regional analysis of population change using Christmas Bird Count data, with application to the American Black Duck. *The Condor*, *108*(1), pp.13-24.

Link, W. A., & Sauer, J. R. (2016). Bayesian cross-validation for model evaluation and selection, with application to the North American Breeding Bird Survey. *Ecology*, *97*(7), 1746-1758.

Link, W.A., Sauer, J.R. and Niven, D.K., 2017. Model selection for the North American Breeding Bird Survey: A comparison of methods. *The Condor: Ornithological Applications*, *119*(3), pp.546-556.

Machtans, C.S., Kardynal, K.J. and Smith, P.A., 2014. How well do regional or national Breeding Bird Survey data predict songbird population trends at an intact boreal site?. *Avian Conservation & Ecology*, *9*(1).

Mahon, C.L., Bayne, E.M., Sólymos, P., Matsuoka, S.M., Carlson, M., Dzus, E., Schmiegelow, F.K. and Song, S.J., 2014. Does expected future landscape condition support proposed population objectives for boreal birds?. *Forest Ecology and Management*, *312*, p.28-39.

Meehan, T.D., Saunders, S.P., DeLuca, W.V., Michel, N.L., Grand, J., Deppe, J.L., Jimenez, M.F., Knight, E.J., Seavy, N.E., Smith, M.A. and Taylor, L., 2022. Integrating data types to estimate spatial patterns of avian migration across the Western Hemisphere. *Ecological applications*, *32*(7).

Morris, S. R., Covino, K. M., Jacobs, J. D., & Taylor, P. D., 2016. Fall migratory patterns of the Blackpoll Warbler at a continental scale. *The Auk: Ornithological Advances*, *133*(1), 41-51.

Niemi, G., Hanowski, J., Helle, P., Howe, R., Mönkkönen, M., Venier, L. and Welsh, D., 1998. Ecological sustainability of birds in boreal forests. *Conservation Ecology*, *2*(2).

Pavlacky Jr, D.C., Lukacs, P.M., Blakesley, J.A., Skorkowsky, R.C., Klute, D.S., Hahn, B.A., Dreitz, V.J., George, T.L. and Hanni, D.J., 2017. A statistically rigorous sampling design to integrate avian monitoring and management within Bird Conservation Regions. PLoS One, 12(10).

Plummer, M., 2003, March. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, No. 125.10, pp. 1-10).

R Core Team. 2024. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rosenberg, K.V., Kennedy, J.A., Dettmers, R., Ford, R.P., Reynolds, D., Alexander, J.D., Beardmore, C.J., Blancher, P.J., Bogart, R.E., Butcher, G.S. and Camfield, A.F., 2016. Partners in Flight landbird conservation plan: 2016 revision for Canada and continental United States. *Partners in Flight Science Committee*, *119*.

Rosenberg, K.V., Dokter, A.M., Blancher, P.J., Sauer, J.R., Smith, A.C., Smith, P.A., Stanton, J.C., Panjabi, A., Helft, L., Parr, M. and Marra, P.P., 2019. Decline of the North American avifauna. *Science*, *366*(6461), pp.120-124.

Sauer, J.R., Link, W.A. and Hines, J.E., 2020. The North American breeding bird survey, analysis results 1966–2019. *US Geological Survey data release.*

Smith, A.C., Hudson, M.A.R., Downes, C. and Francis, C.M., 2014. Estimating breeding bird survey trends and annual indices for Canada: How do the new hierarchical Bayesian estimates differ from previous estimates? *The Canadian Field-Naturalist*, *128*(2), pp.119-134.

Smith, A.C. and Edwards, B.P., 2021. North American Breeding Bird Survey status and trend estimates to inform a wide range of conservation needs, using a flexible Bayesian hierarchical generalized additive model. *The Condor*, *123*(1).

Smith, P.A., Smith, A.C., Andres, B., Francis, C.M., Harrington, B., Friis, C., Morrison, R.G., Paquet, J., Winn, B. and Brown, S., 2023. Accelerating declines of North America's shorebirds signal the need for urgent conservation action. *Ornithological Applications*, *125*(2).

Sólymos, P., Matsuoka, S. M., Bayne, E. M., Lele, S. R., Fontaine, P., Cumming, S. G., ... & Song, S. J. (2013). Calibrating indices of avian density from non-standardized survey data: making the most of a messy situation. *Methods in Ecology and Evolution*, *4*(11), 1047-1058.

Stewart, R. L. M. 2015. Blackpoll Warbler, pp. 441 in Stewart, R.L.M., Bredin, K.A., Couturier, A.R., Horn, A.G., Lepage, D., Makepeace, S., Taylor, P., Villard, M.-A., and Whittam, R.M. (eds). 2015. Second Atlas of Breeding Birds of the Maritime Provinces. Bird Studies Canada.

Stralberg, D., Matsuoka, S.M., Hamann, A., Bayne, E.M., Sólymos, P., Schmiegelow, F.K.A., Wang, X., Cumming, S.G. and Song, S.J., 2015. Projecting boreal bird responses to climate change: the signal exceeds the noise. *Ecological Applications*, *25*(1), pp.52-69.

Van Wilgenburg, S.L. and Hobson, K.A., 2011. Combining stable-isotope (δD) and band recovery data to improve probabilistic assignment of migratory birds to origin. *Ecological Applications*, *21*(4), pp.1340-1351.

Van Wilgenburg, S.L., Beck, E.M., Obermayer, B., Joyce, T. and Weddle, B., 2015. Biased representation of disturbance rates in the roadside sampling frame in boreal forests: implications for monitoring design. *Avian Conservation & Ecology*, *10*(2).

Van Wilgenburg, S.L., Mahon, C.L., Campbell, G., McLeod, L., Campbell, M., Evans, D., Easton, W., Francis, C.M., Haché, S., Machtans, C.S. and Mader, C., 2020. A cost efficient spatially balanced hierarchical sampling design for monitoring boreal birds incorporating access costs and habitat stratification. *PLoS One*, *15*(6).

Wunder, M.B. 2012. Determining geographic patterns of migration and dispersal using stable isotopes in keratins. Journal of Mammalogy 93(2), pp.360-367.

## Tables

**Table 1.** Multi-level statistical model to estimate population trajectories in pre-defined geographic strata by integrating daily counts of migrants at a series of monitoring stations with estimates of breeding origins for a sample of migrants at a subset of stations. Equations are indexed by geographic strata (j), year (y), monitoring station (s), and day of year (d).

| Description | Equations | |
|---|---|---|
| *Stratum-level population process model:* | | |
| Log-linear population change within stratum *j*, starting from baseline year $y_0$. | $\log(X_{j,y}) = \log(X_{j,y_0}) + \beta_j \times (y-1)$ | (1) |
| *Migration process model:* | | |
| Number of migrants arriving at each station from each stratum controlled by "capture rate" parameter $\rho_{j,s}$. | $M_{j,s,y} = \rho_{j,s} X_{j,y}$ | (2) |
| Expected abundance migrants arriving at a station from all regions. Additional temporal variance ($\varepsilon_{s,y}$) is controlled by parameter $\sigma_\rho^2$. | $T_{s,y} \sim Lognormal(\log(\sum_1^J M_{j,s,y}), \sigma_\rho^2)$ | (3) |
| Seasonal temporal distribution of migrants arriving at the station follows a normal curve with a mean date $\mu_s$ and a standard deviation $\sigma_s$, where day of the year is indexed by $d$. | $\theta_{d,s,y} = T_{s,y} \times f(d, \mu_s, \sigma_s)$, where: $f(d, \mu_s, \sigma_s) = \frac{1}{\sigma_s\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{d-\mu_s}{\sigma_s})}$ | (4) |
| *Observation models:* | | |
| Observed number of migrants at each station on each day of year is Poisson distributed with log-normal overdispersion (controlled by $\omega_s^2$), and an offset for survey effort (e.g., number of hours nets were operational on a day). | $n_{d,s,y} \sim Poisson(\lambda_{d,s,y})$, where: $\log(\lambda_{d,s,y}) \sim Normal(\log(\theta_{d,s,y}) + offset_{d,s,y}, \omega_s^2)$ | (5) |
| Multinomial distribution describes the observed breeding origins for a sample of $n$ birds collected at a station in a given year. | $\boldsymbol{Y_{s,y}} \sim Multinomial\left(n_{s,y}, \left(\frac{M_{1,s,y}}{T_{s,y}}, \frac{M_{2,s,y}}{T_{s,y}}, \ldots, \frac{M_{J,s,y}}{T_{s,y}}\right)\right)$ | (6) |

519 **Table 2.** Specification of Bayesian priors for analysis of seasonal migration counts.

| *Parameter* | *Prior* | *Notes* |
|---|---|---|
| | *Stratum-level parameters:* | |
| $X_{j,y_0}$ | *Fixed to 1* | Fixing this term to 1 ensures $\rho_{j,s}$ terms are identifiable. Resultant estimates of $X_{j,y}$ can be rescaled outside of fitting procedure based on independent estimates of relative abundance across strata (e.g., based on a species distribution model describing spatial patterns in abundance). |
| $\beta_j$ | *Normal(0, 0.1$^2$)* | Log-linear temporal trend within stratum. |
| | *Station-level parameters:* | |
| $\rho_{j,s}$ | *Lognormal(0, 4$^2$)* | Migration parameters describing the expected annual contribution of migrants from stratum *j* to station *s.* Prior chosen to be highly vague (i.e., non-informative). |
| $\sigma_\rho$ | *Uniform(0,2)* | Magnitude of random year-to-year variation in station-level indices around the annual expected count |
| $\mu_s$ | *Uniform(1,360)* | Day of year at which the expected "peak" of migration occurs. |
| $\sigma_s$ | *Lognormal(2.2, 0.45$^2$)* | Describes temporal duration of the migration period within a season. Migration is assumed to follow a normal curve such that approximately 95% of birds arrive at station within $\mu_s \pm 1.96\sigma_s$. This prior implies an expectation that $\sigma_s$ is likely to be between 6 and 13, which allows for a wide but realistic range of migration windows at each monitoring station. |
| $\omega_s$ | *Uniform(0,2)* | Magnitude of extra-Poisson error in expected daily counts within a season. |

520

521 **Table 3.** Estimates of population trend and percent change relative to 1998 within each stratum. Values
522 are expressed as posterior median value followed by 95% equal-tailed credible interval in parentheses.

| Stratum | Source of trend estimate | 20-year trend | Prob trend is positive | % change since 1998 |
|---|---|---|---|---|
| West | Pre-breeding migration | +4.0 (1.5 to 6.9) | > 0.99 | +120 (36 to 279) |
| West | Post-breeding migration | +1.8 (-0.3 to 3.8) | 0.95 | +42 (-7 to 110) |
| West | Breeding Bird Survey | -2.7 (-5.3 to 0.3) | 0.03 | -43 (-66 to +6) |
| East | Pre-breeding migration | -4.8 (-7.3 to -2.4) | < 0.01 | -63 (-78 to -39) |
| East | Post-breeding migration | +0.2 (-13.2 to 6.3) | 0.51 | +3 (-94 to 239) |
| East | Breeding Bird Survey | -3.8 (-6.1 to -1.4) | < 0.01 | -54 (-71 to -24) |
| Continental | Pre-breeding migration | -0.4 (-0.4 to 2.7) | 0.65 | +8 (-27 to 69) |
| Continental | Post-breeding migration | +1.3 (-0.6 to 3.7) | 0.90 | +29 (-11 to 108) |
| Continental | Breeding Bird Survey | -3 (-5.1 to -0.7) | 0.01 | -46 (-65 to -13) |

523

## Figures

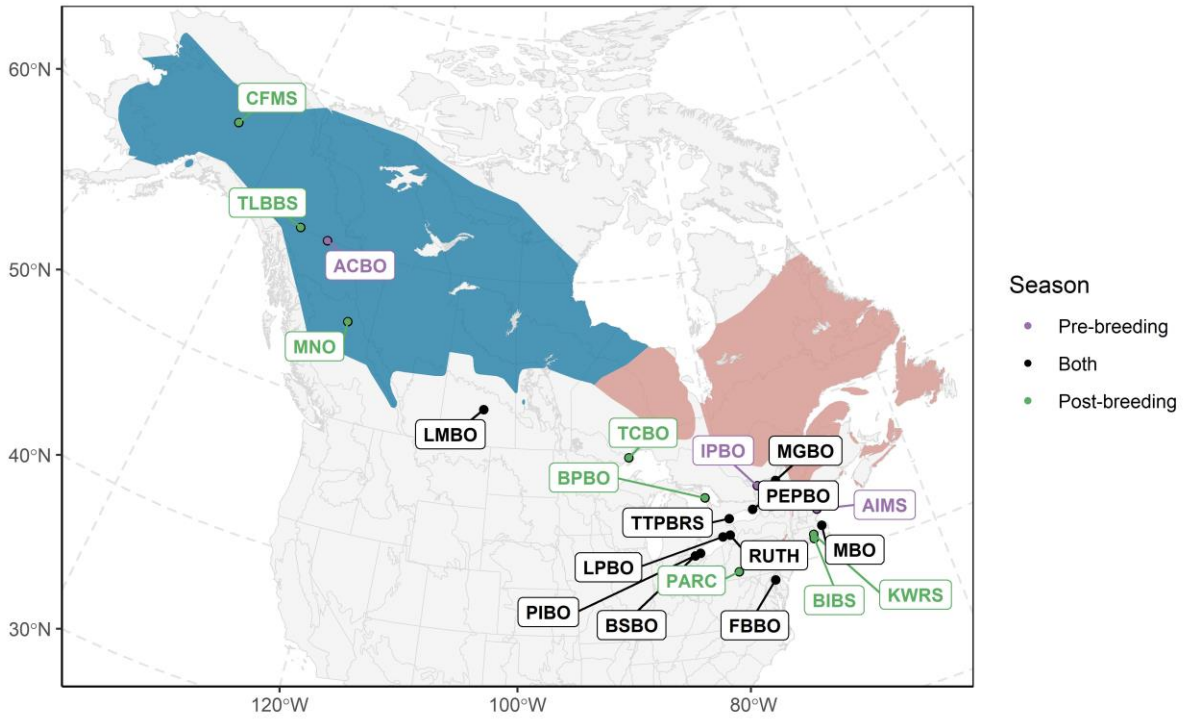



**Figure 1.** Boundaries of the two strata ("West" and "East") used for analysis of continental Blackpoll Warbler population trends from long-term migration monitoring data. Location labels for migration monitoring stations included in analysis are coloured to indicate data from pre-breeding migration only (purple), post-breeding (green) or both seasons (black).

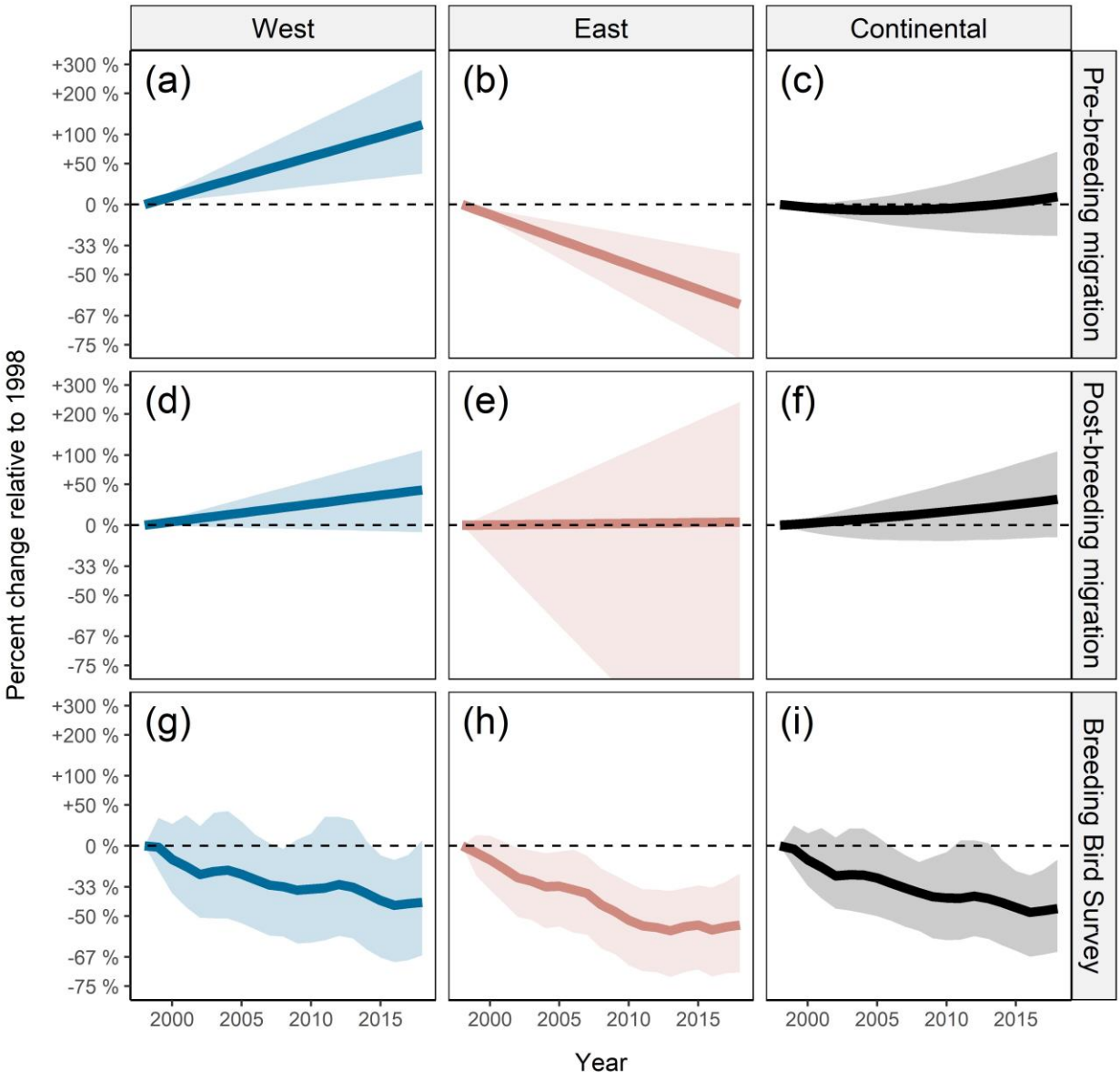



**Figure 2.** Estimated population trajectories of Blackpoll Warbler from 1998 to 2018 in East, West, and combined continental strata based on analysis of pre-breeding migration (top row), post-breeding migration (middle row), and North American Breeding Bird Survey (bottom row). Strata boundaries and stations contributing data to the analysis are illustrated in Figure 1. Dots and whiskers are posterior medians and 95% equal-tailed credible intervals, respectively.

# Supplementary Information

## Appendix 1: Station descriptions and examples of data structure

This appendix contains additional information describing the locations where data were collected (Table S1) and illustrates the structure of the data used for analysis (Figures S1.1 to S1.4).

**Table S1.** Locations of migration monitoring stations and summary of data contributing data to continental analysis of Blackpoll Warbler population trends. "Catchment Estimate Method" refers to the way in which migrants were assigned to particular geographic strata (i.e., either using stable isotopes in feathers, fixed based on knowledge of Blackpoll Warbler migration routes, or not estimated).

| Station code | Station name | Country | Lat | Lon | Year range | Mean annual count | Min annual count | Max annual count | Catchment Estimate Method |
|---|---|---|---|---|---|---|---|---|---|
| Pre-breeding ("spring") migration monitoring stations | | | | | | | | | |
| ACBO | Albert Creek Bird Observatory | CAN | 60.1 | -128.9 | 2007 - 2018 | 80.1 | 21 | 169 | West only |
| LMBO | Last Mountain Bird Observatory | CAN | 51.4 | -105.2 | 1998 - 2018 | 9.6 | 0 | 33 | West only |
| BSBO | Black Swamp Bird Observatory | USA | 41.6 | -83.2 | 1998 - 2018 | 78.4 | 20 | 176 | - |
| PIBO | Pelee Island Bird Observatory | CAN | 41.7 | -82.7 | 2003 - 2018 | 15.4 | 1 | 35 | - |
| LPBO | Long Point Bird Observatory | CAN | 42.6 | -80.3 | 1998 - 2018 | 261.6 | 68 | 577 | Isotopes |
| RUTH | Haldimand Bird Observatory - Ruthven | CAN | 42.6 | -79.5 | 1998 - 2018 | 20.8 | 1 | 70 | Isotopes |
| TTPBRS | Tommy Thompson Park Bird Research Station | CAN | 43.6 | -79.3 | 2005 - 2018 | 37.7 | 13 | 75 | Isotopes |
| PEPBO | Prince Edward Point Bird Observatory | CAN | 43.9 | -76.9 | 1998 - 2018 | 18.8 | 5 | 43 | Isotopes |
| FBBO | Foreman's Branch Bird Observatory | USA | 39.2 | -76.1 | 2006 - 2018 | 8.2 | 0 | 21 | - |
| IPBO | Innis Point Bird Observatory | CAN | 45.4 | -75.9 | 1998 - 2018 | 22.6 | 4 | 52 | - |
| MGBO | McGill Bird Observatory | CAN | 45.4 | -73.9 | 2006 - 2018 | 39.4 | 3 | 85 | Isotopes |
| AIMS | Appledore Island Migration Station | USA | 43 | -70.6 | 1998 - 2018 | 72 | 31 | 144 | East only |
| MBO | Manomet Observatory | USA | 41.9 | -70.5 | 1998 - 2018 | 13.6 | 1 | 34 | East only |
| Post-breeding ("fall") migration monitoring stations | | | | | | | | | |
| CFMS | Creamer's Field Migration Station | USA | 64.9 | -147.7 | 1998 - 2018 | 34 | 1 | 180 | West only |
| TLBBS | Teslin Lake Bird Banding Station | CAN | 60.2 | -133 | 2009 - 2018 | 159.4 | 89 | 277 | West only |
| MNO | Mackenzie Nature Observatory | CAN | 55.3 | -123.1 | 1998 - 2018 | 33.1 | 12 | 71 | West only |
| LMBO | Last Mountain Bird Observatory | CAN | 51.4 | -105.2 | 1998 - 2017 | 58.8 | 0 | 124 | West only |
| TCBO | Thunder Cape Bird Observatory | CAN | 48.3 | -88.9 | 1998 - 2018 | 112.5 | 10 | 448 | Isotopes |
| BSBO | Black Swamp Bird Observatory | USA | 41.6 | -83.2 | 1998 - 2018 | 536.8 | 117 | 1129 | - |
| PIBO | Pelee Island Bird Observatory | CAN | 41.7 | -82.7 | 2003 - 2018 | 113.6 | 39 | 256 | - |
| BPBO | Bruce Peninsula Bird Observatory | CAN | 45.2 | -81.4 | 2002 - 2018 | 18.9 | 4 | 80 | - |
| LPBO | Long Point Bird Observatory | CAN | 42.6 | -80.3 | 1998 - 2018 | 2560 | 613 | 5000 | Isotopes |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| RUTH | Haldimand Bird Observatory - Ruthven | CAN | 42.6 | -79.5 | 1998 - 2018 | 65.8 | 15 | 229 | Isotopes |
| TTPBRS | Tommy Thompson Park Bird Research Station | CAN | 43.6 | -79.3 | 2005 - 2018 | 50.1 | 13 | 99 | Isotopes |
| PARC | Powdermill Avian Research Center | USA | 40.2 | -79.3 | 1998 - 2018 | 33 | 18 | 92 | Isotopes |
| PEPBO | Prince Edward Point Bird Observatory | CAN | 43.9 | -76.9 | 2001 - 2018 | 142.4 | 58 | 371 | - |
| FBBO | Foreman's Branch Bird Observatory | USA | 39.2 | -76.1 | 2006 - 2018 | 17.8 | 4 | 68 | - |
| MGBO | McGill Bird Observatory | CAN | 45.4 | -73.9 | 2006 - 2018 | 26.4 | 3 | 55 | - |
| BIBS | Block Island Banding Station | USA | 41.2 | -71.6 | 1998 - 2018 | 15.8 | 0 | 100 | - |
| KWRS | Kingston Wildlife Research Station | USA | 41.5 | -71.5 | 1998 - 2018 | 21.1 | 0 | 65 | Isotopes |
| MBO | Manomet Observatory | USA | 41.9 | -70.5 | 1998 - 2018 | 87.8 | 14 | 303 | Isotopes |

Data at each migration monitoring station are collected daily. Daily counts of migrants, corrected for survey effort, are illustrated below. Note that for Canadian migration monitoring stations, daily effort was not available as "net-hours" but each station uses standardized protocols. Thus, to express $T_{s,y}$ (an index of summed daily counts per net hour across a season; see Table 1 in manuscript) on a similar scale to US stations, we assigned daily net-hours as the mean of US stations (257 in pre-breeding migration, and 232 for post-breeding migration).

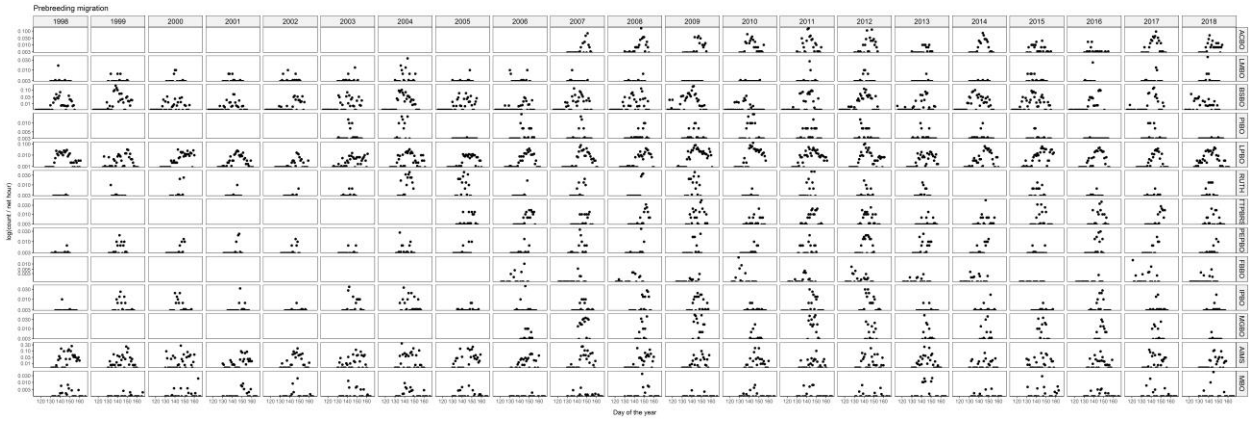



**Figure S1.1.** Number of birds counted per net-hour each day, at each station, within each year, during pre-breeding migration.

**Figure S1.2.** Number of birds counted per net-hour each day, at each station, within each year, during post-breeding migration. Stations are arranged from farthest west (top row) to farthest east (bottom row).

Stable isotopes of hydrogen within feather samples from migrating birds were analyzed and used to assign birds to geographic strata of origin. The numbers of migrants assigned to particular geographic strata, at each station within each year, were used as raw data in the Bayesian statistical model described in the main text of this manuscript (i.e., the breeding origin assignments represent $\boldsymbol{Y}_{s,y}$ in equation 6). Breeding origin assignments were used as data for stations within 250 km of the location where feather isotopes were collected. The resulting assignments are depicted in the figures below.

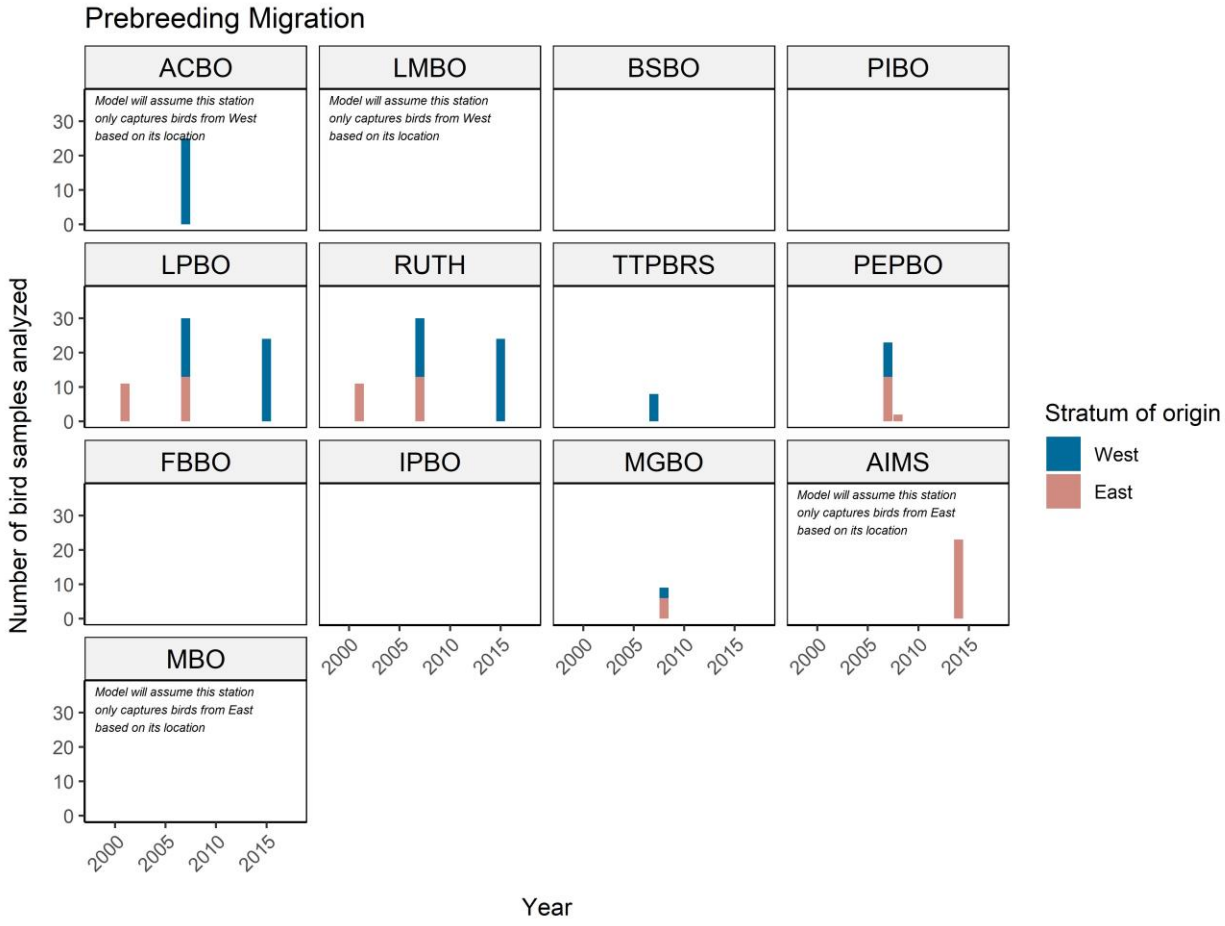



**Figure S1.3.** Number of birds assigned to discrete geographic strata based on analysis of stable isotopes of hydrogen in feather samples during pre-breeding ('Spring' migration), used as data ($\boldsymbol{Y}_{s,y}$ in equation 6) in the Bayesian statistical analysis in combination with time-series of migration counts. Some stations were also assumed to only capture birds from a single stratum in our statistical analysis, based on their geographic location. Panels are arranged from farthest west to farthest east when read from top left to bottom right.

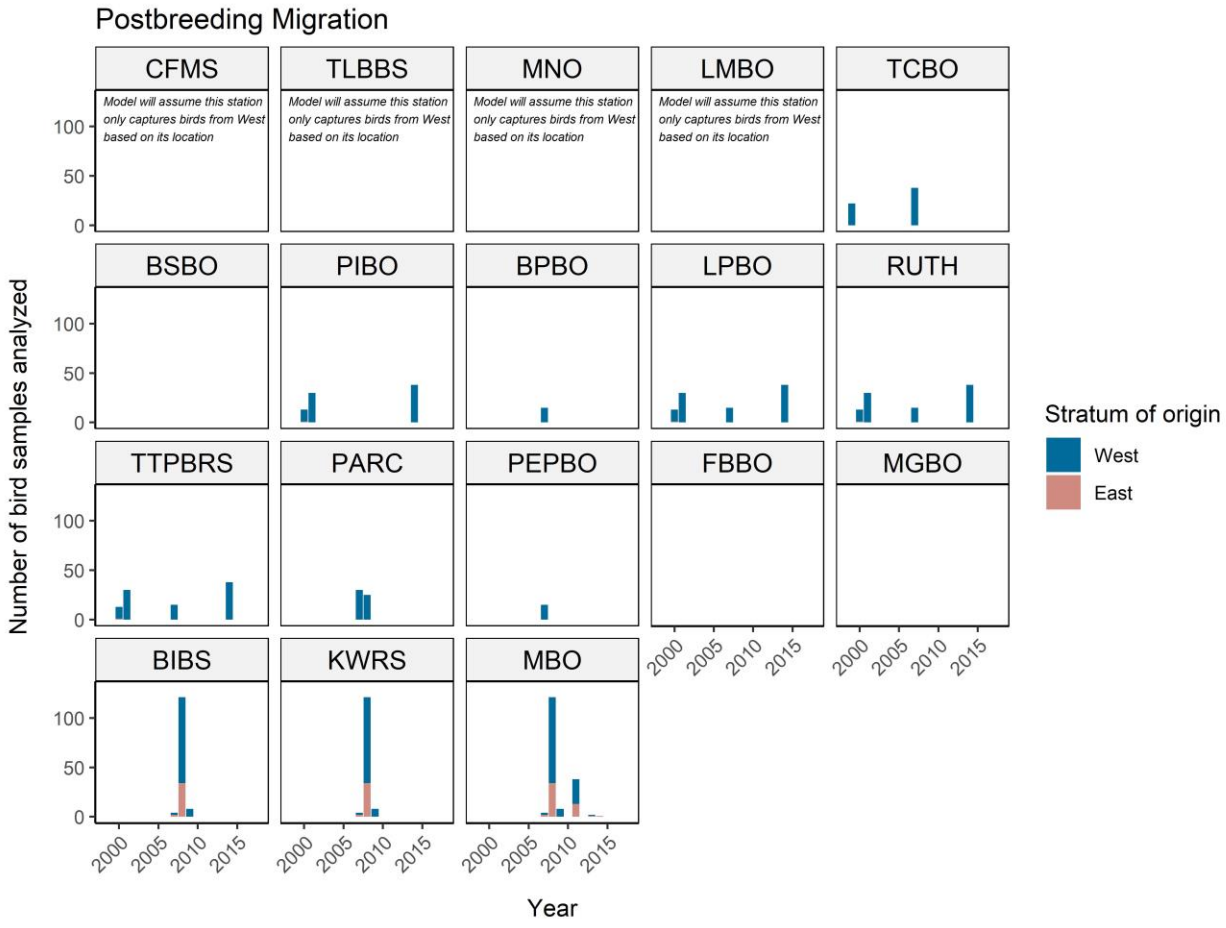



**Figure S1.4.** Number of birds at each station assigned to discrete geographic strata based on analysis of stable isotopes of hydrogen in feather samples during post-breeding ('Fall' migration), used as data ($\boldsymbol{Y}_{s,y}$ in equation 6) in the Bayesian statistical analysis in combination with time-series of migration counts. Some stations were also assumed to only capture birds from a single stratum in our statistical analysis, based on their geographic location. Panels are arranged from farthest west to farthest east when read from top left to bottom right.

## Appendix 2: Methods for estimating relative abundance of Blackpoll Warbler in each geographic stratum, and sensitivity of continental trend estimates to these methods

While the statistical model described in this study can estimate stratum-specific population trends, we require independent estimates of the relative abundance of birds within each geographic strata (which also must be referenced to a particular year) in order to calculate range-wide estimates of population trends from this modeling framework. This is necessary because trends within stratum must be "weighted" appropriately within a range-wide context; a large positive trend in a stratum that contains almost no birds will have a small effect on the continental trend compared to a small negative trend in a stratum that contains a large number of birds.

In the main text of this manuscript, we presented range-wide trends weighted according to relative abundance estimates from the boreal avian modeling (BAM) project. BAM produces maps of relative abundance across boreal Canada, referenced to the year 2011 (Stralberg et al. 2015), which we illustrate in Figure S2.1a. Pixels are 1 km x 1 km, and represent relative population density. Thus, by summing pixel values within each stratum we can generate estimates of relative abundance. This map suggests the Blackpoll Warbler population was 1.99 times more abundant in the West stratum than the East stratum in 2011, though we note this estimate excludes the Alaskan portion of their range.

However, to examine how an alternative source of relative abunance estimates would impact continental trend estimates, we calculated continental trends based on relative abundance estimates from eBird, referenced to the year 2022 (Fink et al. 2023). This map was downloaded using the 'ebirdst' package in R (Fink et al. 2023) and recreated in Figure S2.1b. However, eBird flags the relative abundance map for Blackpoll Warbler as "low reliability", because most of the predictions are extrapolations outside the spatial range of data availability, which is primarily based on opportunistically collected surveys at the margins of the species' boreal range. Based on this map, the relative abundance of Blackpoll Warbler in the East stratum was 1.46 times larger than the West stratum in 2022 (i.e., the opposite pattern predicted by BAM for the year 2011). Since our analysis only encompassed from the year 1998 to 2018, we applied this correction to 2018 such that the posterior mean of the population index for the East stratum was 1.46 times larger the West.

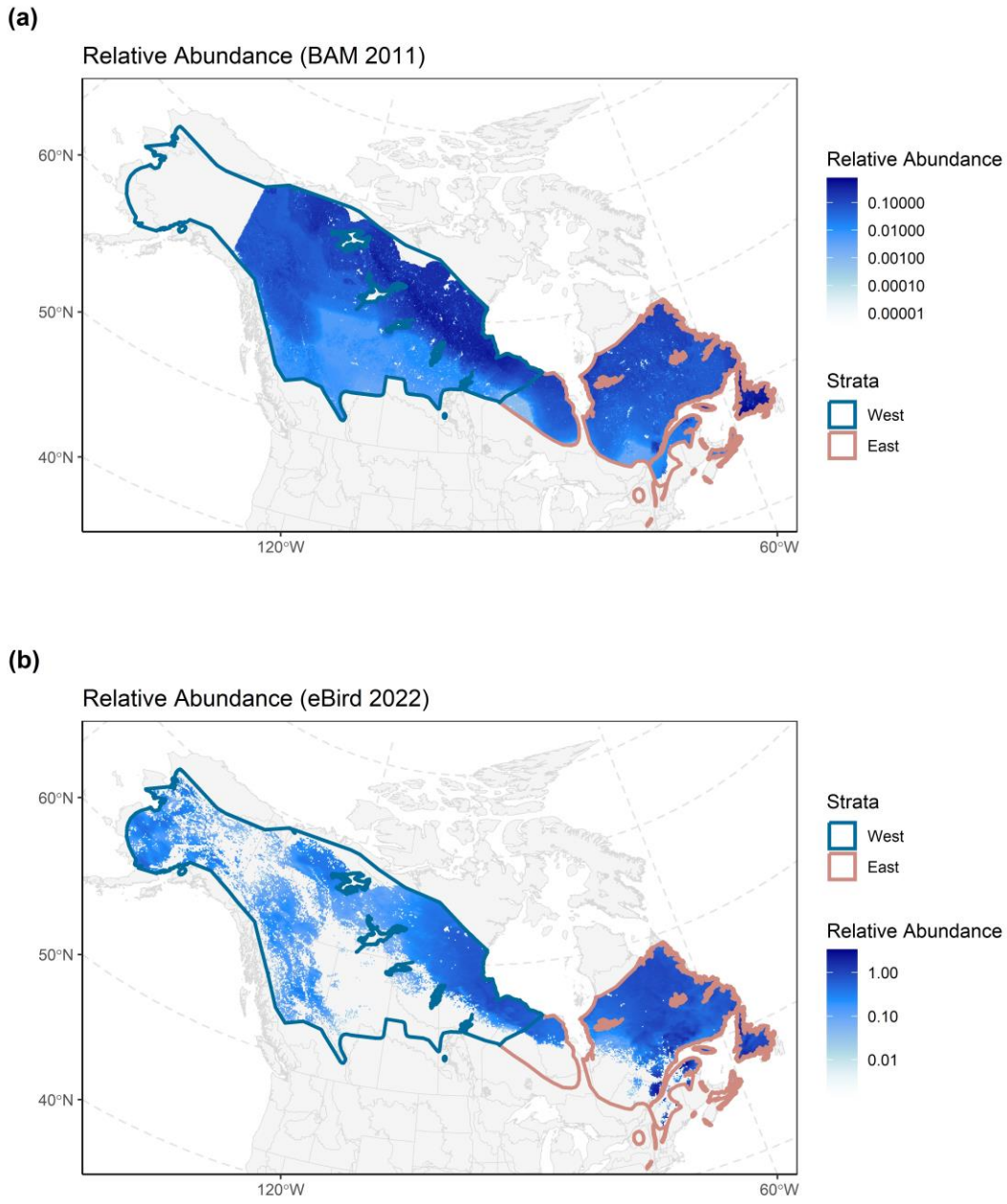



**Figure S2.1.** Relative abundance maps based on two independent sources of information. Panel a) based on predictions from the Boreal Avian Modeling project (Stralberg et al. 2015). Panel b) based on predictions from eBird (Fink et al. 2023).

While these differences have no effect on stratum-level population trend estimates, they affect the relative weighting of trends in a continental context. Below, we illustrate how re-weighting the stratum level trends (such that they achieve the appropriate relative abundance in the reference year

for each source of information) affects estimates of range-wide trend. Figure S2.2 illustrates population trajectories assuming BAM's relative abundance raster is a good approximation of density in 2011, such that the posterior mean of the West stratum's population index was 1.99 times larger than the East in 2011.

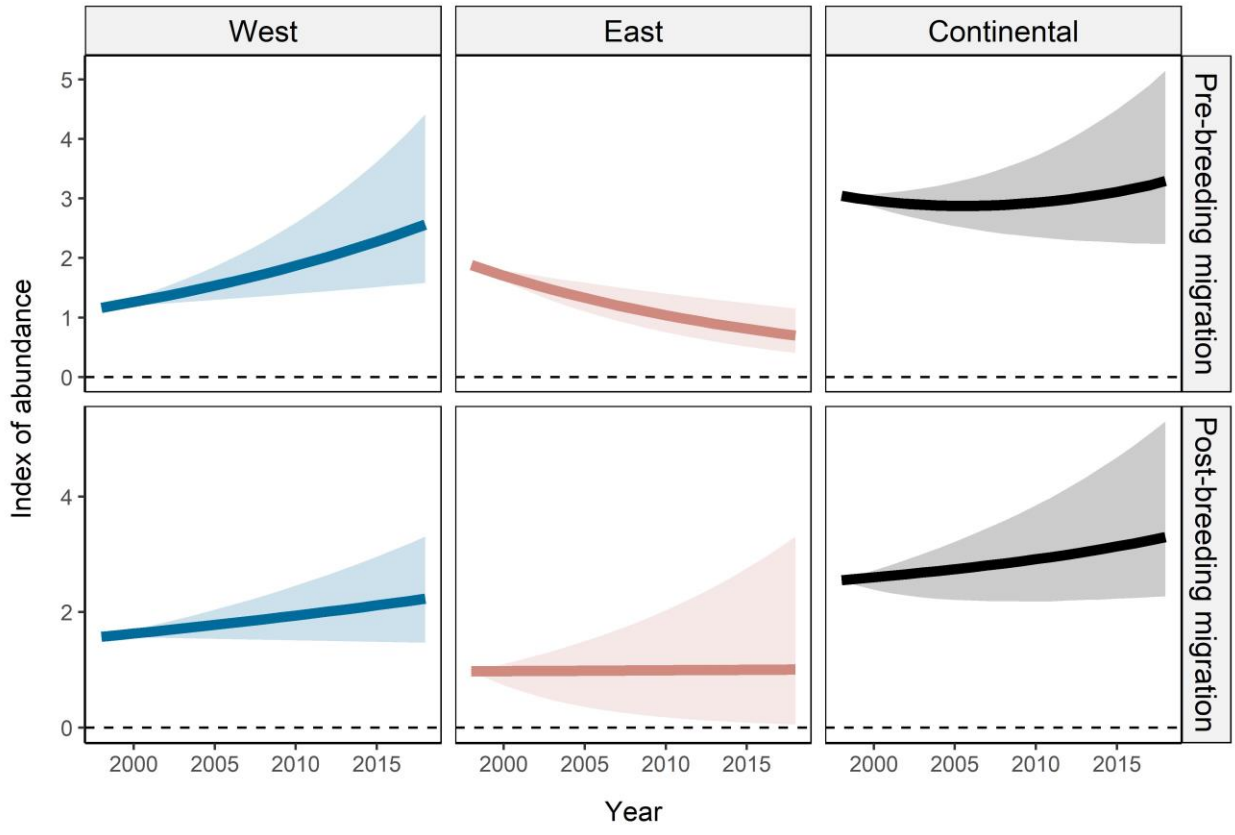



**Figure S2.2.** Population trajectories of Blackpoll Warbler assuming relative abundance was 1.99 times larger in the West stratum than the East stratum in 2011 (based on relative abundance estimates from the Boreal Avian Modeling Project). Solid lines represent posterior median of trajectories, ribbons represent 95% equal-tailed credible intervals.

Conversely, Figure S2.3 illustrates population trajectories assuming eBird's relative abundance raster is a good approximation of density in 2018, such that the posterior mean of the East stratum's population index was 1.46 times larger than the West in 2018.

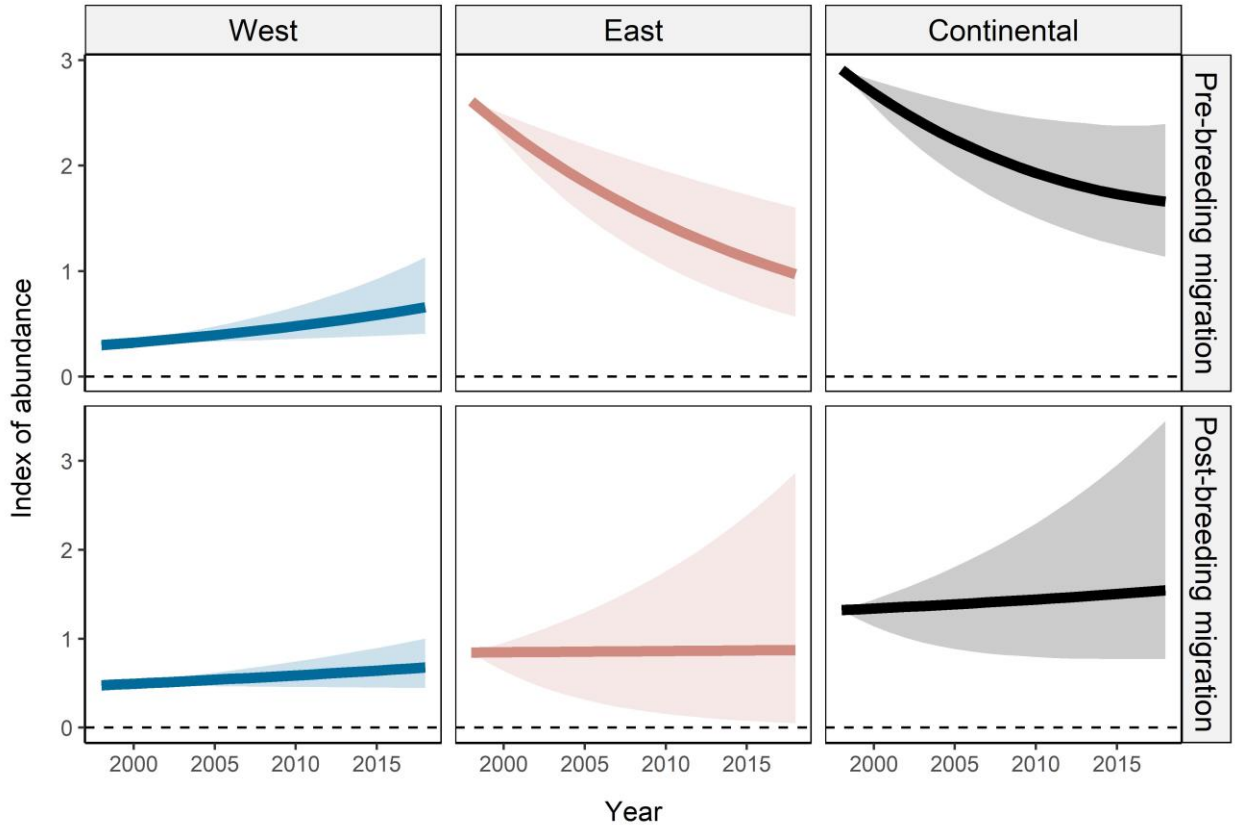



**Figure S2.3.** Population trajectories of Blackpoll Warbler assuming relative abundance was 1.46 times larger in the East stratum than the West stratum in 2018 (based on relative abundance estimates from eBird). Solid lines represent posterior median of trajectories, ribbons represent 95% equal-tailed credible intervals.

The resulting effects of these differences in the relative weighting of stratum-level trends on continental trend estimates is presented in Table S2.1.

**Table S2.1.** Continental trend estimates from pre- and post-breeding migration monitoring based on weighting stratum-level trends from two different sources of relative abundance information.

| Season of migration monitoring trend | Source of relative abundance estimate | 20-year trend | Prob trend is positive | % change since 1998 |
|---|---|---|---|---|
| Pre-breeding | BAM | -0.4 (-0.4 to 2.7) | 0.65 | +8 (-27 to 69) |
| Post-breeding | BAM | +1.3 (-0.6 to 3.7) | 0.90 | +29 (-11 to 108) |

| | | | | |
|---|---|---|---|---|
| Pre-breeding | eBird | -2.8 (-4.6 to -1) | 0 | -43 (-61 to -18) |
| Post-breeding | eBird | +0.8 (-2.7 to 4.9) | 0.61 | +17 (-42 to 161) |

644

645 *Literature Cited*


646 Fink, D., T. Auer, A. Johnston, M. Strimas-Mackey, S. Ligocki, O. Robinson, W. Hochachka, L. Jaromczyk,
647 C. Crowley, K. Dunham, A. Stillman, I. Davies, A. Rodewald, V. Ruiz-Gutierrez, C. Wood. 2023.
648 eBird Status and Trends, Data Version: 2022; Released: 2023. *Cornell Lab of Ornithology, Ithaca,*
649 *New York.* https://doi.org/10.2173/ebirdst.2022
650 Stralberg, D., Matsuoka, S. M., Hamann, A., Bayne, E. M., Sólymos, P., Schmiegelow, F. K. A., ... & Song,
651 S. J. (2015). Projecting boreal bird responses to climate change: the signal exceeds the
652 noise. *Ecological Applications*, *25*(1), 52-69.

## Appendix 3: Description of simulation analyses, and goodness-of-fit evaluation for empirical model

*Simulation description*

We conducted a series of simulations to evaluate whether the statistical model could accurately recover regional population trends, given noisy and incomplete datasets. We conducted 250 independent simulations, each 20 years in duration, wherein we generated random log-linear slopes from $Uniform(-0.1,0.1)$ for each of two hypothetical strata. The relative abundance in each stratum was set to 1 in the initial year of the simulations.

Simulations were intended to approximate the data collection process during pre-breeding migration in our empirical study. Simulations assumed there were 13 migration monitoring stations that counted birds in each year of study, and stations operated for the same durations as in the empirical analysis described in the main text. For each station, we selected model parameters that would lead to "realistic" annual counts and temporal variation in counts.

In each simulation, we randomly assigned migration parameters ($\rho_{j,s}$) to each of 13 migration monitoring stations by drawing from a uniform distribution on the log-scale between log(0.01) and log(3), resulting in a wide but realistic range of "capture rates" and numbers of seasonal migrants among stations; most $\rho_{j,s}$ in the empirical analysis were less than 0.5, though several stations had values larger than . For the remainder of model parameters, we used median parameter estimates from the empirical model fit to pre-breeding migration, to ensure that simulated datasets contained realistic values of process and observation variance. However, unlike the empirical analysis, we assumed that no stations were known *a priori* to exclusively monitor a single stratum. These simulations therefore represent an especially difficult monitoring situation in which must be estimated based only on noisy mixtures of birds from multiple regions at all stations simultaneously. In each simulation, we also assumed the data collection process was identical to that in the empirical analysis; several stations only operated for a limited number of years and days across the study duration, and feather isotopes were only collected at a subset of monitoring sites (also in a limited number of years at each station), corresponding to those in the empirical analysis. We used these parameters and data collection constraints to simulate new datasets of observed counts on each day of the season, at each monitoring station.

For each simulation, we fit the Bayesian statistical model to the simulated datasets and calculated estimates of trend to compare to the "true" (simulated) trends that generated the data. Critically, when we fit the Bayesian statistical model to the simulated data, we did not enter the

simulated values of regional abundance ($X_{j,y}$) or migration parameters ($\rho_{j,s}$) as data. Instead, the model estimated those parameters using only a small sample of simulated birds with "known" breeding origins at a subset of sites and years, combined with simulated daily migration counts at each station. For converged models (those with R-hat statistics less than 1.1 for all monitored parameters), we calculated the mean bias in estimates of stratum-level trends and coverage of the 95% credible intervals. If the model can recover stratum-level trends, we expected that bias would be minimal on average and the 95% credible intervals would contain the true (i.e., simulated) stratum-level trends in 95% of simulations.

Results of simulations confirmed that the model could accurately recover stratum-level population trends under a wide range of hypothetical population trends and migration scenarios. Models converged for 249 of 250 simulations. Estimates of regional trends were highly correlated with true (i.e., simulated) trends, bias in estimated trends was extremely small (mean = -0.0008), and 95% credible intervals overlapped with the true trends for 95.7% of estimates (477 of the 498 stratum-level trend estimates from the 249 converged model runs; Figure S4.1).

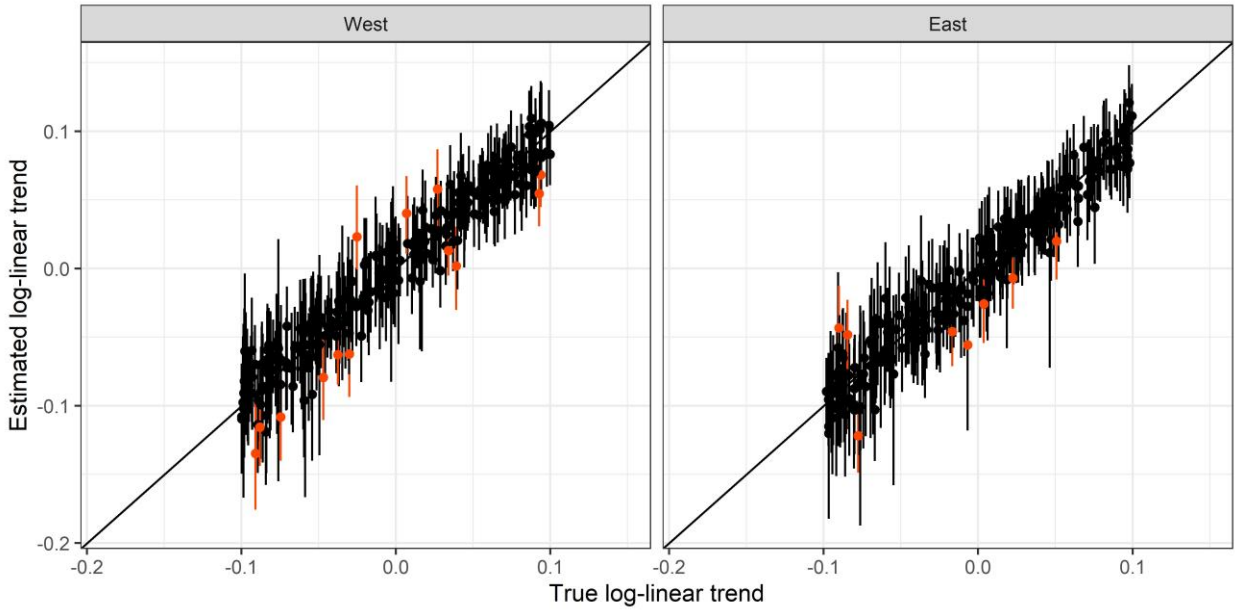



**Figure S4.1.** Results of simulation analysis, comparing true (i.e., simulated) log-linear trends to estimated trends for each of two hypothetical strata. Red dots/whiskers indicate simulations in which the 95% credible intervals on estimates did not contain true trends.

*Posterior Predictive Checks*

We used posterior predictive checks to confirm that the empirical datasets were "similar" to simulated datasets based on the fitted models. The logic of this exercise is discussed more fully in Kery and Royle 2016; pp 192-198) and is a common approach for evaluating goodness-of-fit for Bayesian hierarchical models with multi-level error structures.

In brief, for each iteration of the MCMC fitting algorithm, we calculated an expected value for each datapoint in the empirical dataset, and additionally, we generated an entire simulated dataset based on the estimated model parameters and variance components (i.e., a new dataset that was "perfectly consistent" with the fitted model). We then calculated a measure of the discrepancy between the expected values from the fitted model and 1) the empirical data, and 2) the simulated data. A mis-calibrated model can be diagnosed when discrepancy measures for the observed data are consistently higher or lower than the discrepancy between the simulated data and the model. The proportion of simulated datasets with lower discrepancy measures than the observed data is called the "Bayesian p-value", and a well-calibrated model will have a Bayesian p-value close to 0.5, while p-values close to either 0 or 1 (and far from 0.50) indicate model miscalibration. This method therefore determines whether the empirical dataset looks as if it was simulated from the fitted model.

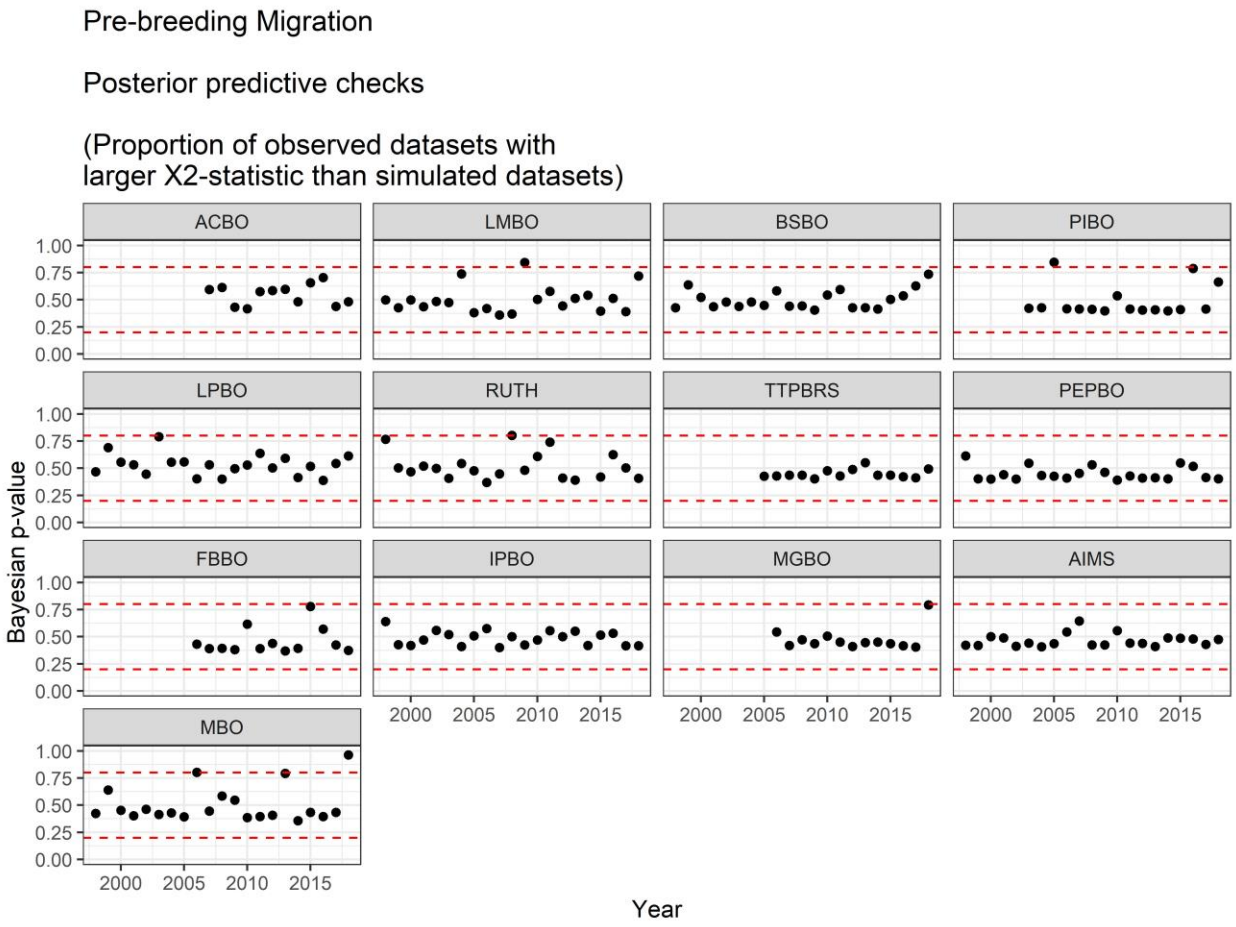
We calculated expected values for each datapoint as:

$$Expected_i = \exp\left(\log\left(T_{s,y}\right) + \log\left(f\left(d, \mu_s, \sigma_s\right)\right) + 0.5\omega_s^2\right).$$

Discrepancy measures were calculated as chi-squared statistics, where:

$$\chi_i^2 = \frac{(Count_i - Expected_i)^2}{Expected_i}.$$

To provide a measure of model "fit" for each station in each year of study, we summed these measures across all days of the season, for each monitoring station, within each year. We then calculated the proportion of simulated datasets with lower discrepancy statistics than the observed datasets for each station-year combination, within each season. Results for pre-breeding and post-breeding migration are depicted in Figures S4.2 and S4.3.




730

731 **Figure S4.2.** Comparison of discrepancy measures for observed counts and simulated counts at each
732 MCMC iteration, summed across all days of the season for each year, at each monitoring station for pre-
733 breeding migration. Discrepancy measures is X-squared residual between observed counts and
734 predicted count. Discrepancy for simulated count is summed squared residual between simulated
735 counts and predicted count. P-values close to 0.50 indicate that the model is reproducing the observed
736 distribution of data; values close to 0 or 1 indicate discrepancy is very different between simulated and
737 observed data. Red dashed lines are positioned at p-values of 0.2 and 0.8 to help visualize instances
738 where Bayesian p-values are indicating a potential lack of fit. Plots arranged in order from farthest west
739 to farthest east.

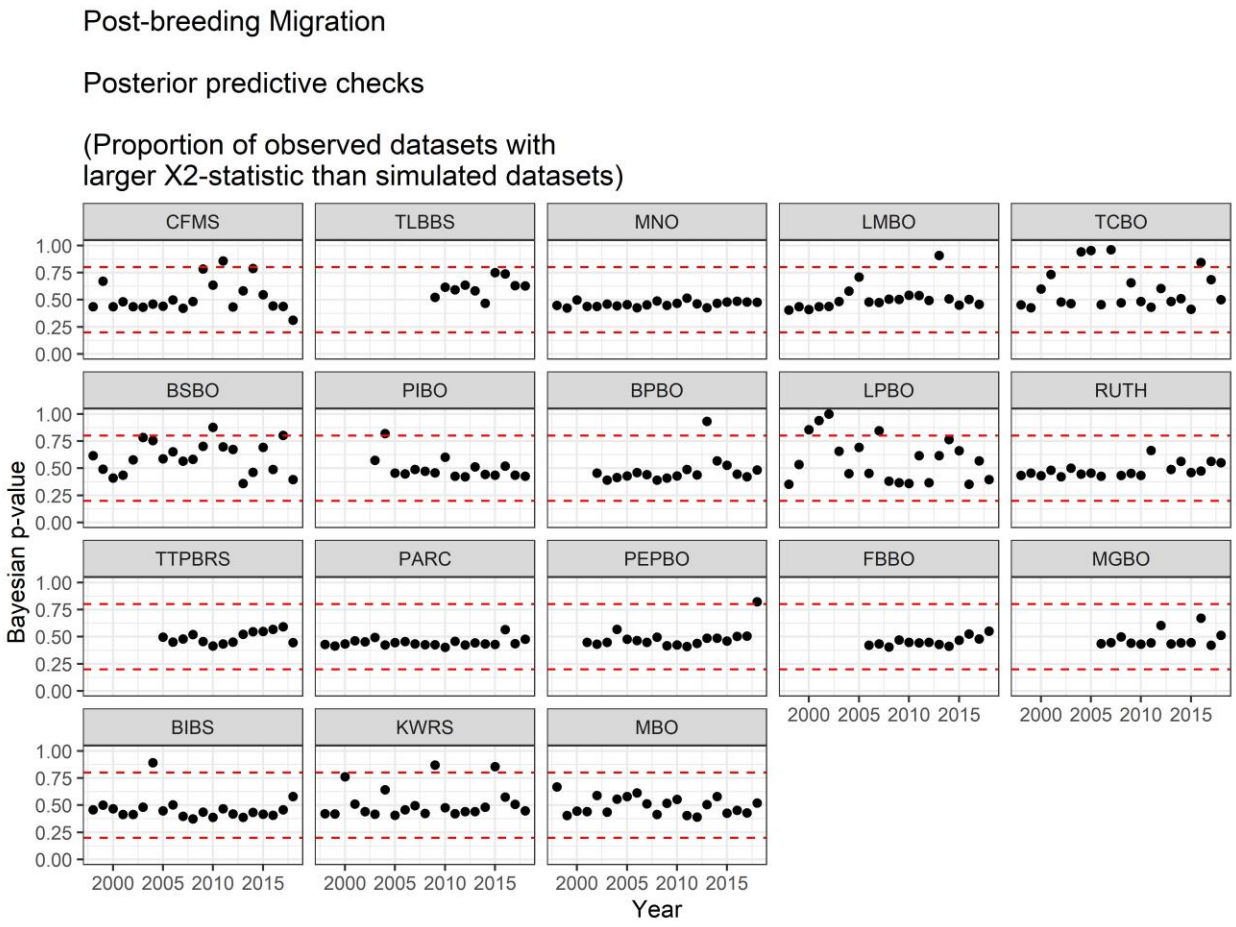



**Figure S4.3.** Comparison of discrepancy measures for observed counts and simulated counts at each MCMC iteration, summed across all days of the season for each year, at each monitoring station for post-breeding migration. Discrepancy measures is X-squared residual between observed counts and predicted count. Discrepancy for simulated count is summed squared residual between simulated counts and predicted count. P-values close to 0.50 indicate that the model is reproducing the observed distribution of data; values close to 0 or 1 indicate discrepancy is very different between simulated and observed data. Red dashed lines are positioned at p-values of 0.2 and 0.8 to help visualize instances where Bayesian p-values are indicating a potential lack of fit. Plots arranged in order from farthest west to farthest east.

We also plotted observed total counts versus expected total counts for each station, within each year of study. Annual expected counts were calculated as $\sum_{d=1}^{365}\left(\exp\left(log(T_{s,y})+log\left(\frac{1}{\sigma_s\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{d-\mu_s}{\sigma_s}\right)}\right)+offset_{d,s,y}+0.5\omega_s^2\right)\right)$. These comparisons are illustrated in figures S4.4 and S4.5, and mainly illustrate that the variance components in the model allow for accurate characterizations of temporal variation in seasonal totals at most stations, in most years.

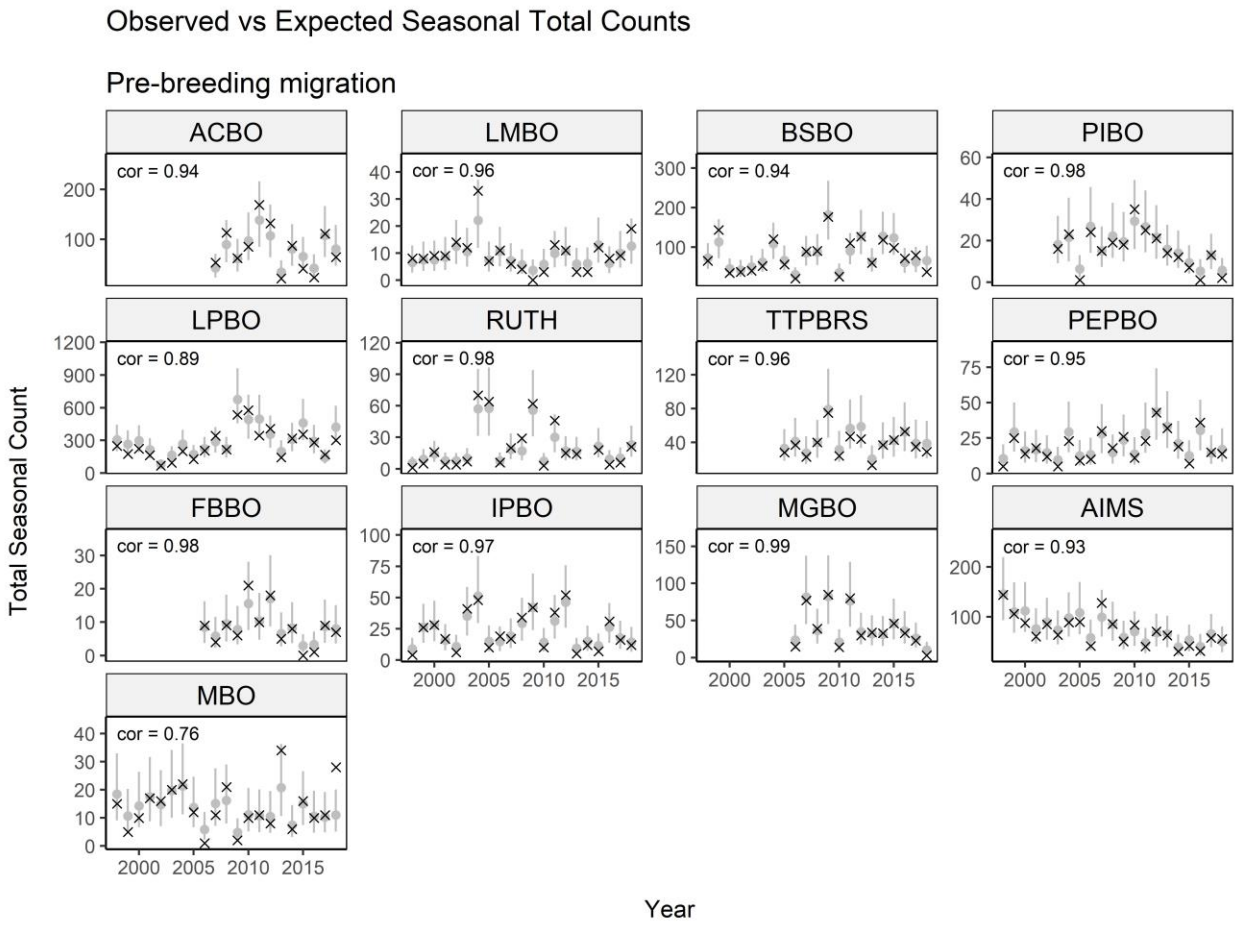



**Figure S4.4.** Comparison of observed (black "x") and expected (gray dots and whiskers) total seasonal counts in each season, at each monitoring station. Whiskers are 95% prediction intervals (i.e., 95% of observations are expected to fall inside those bounds). Pearson correlation between observed and expected (mean of posterior) counts are written in top left of each facet. Plots arranged in order from farthest west to farthest east.

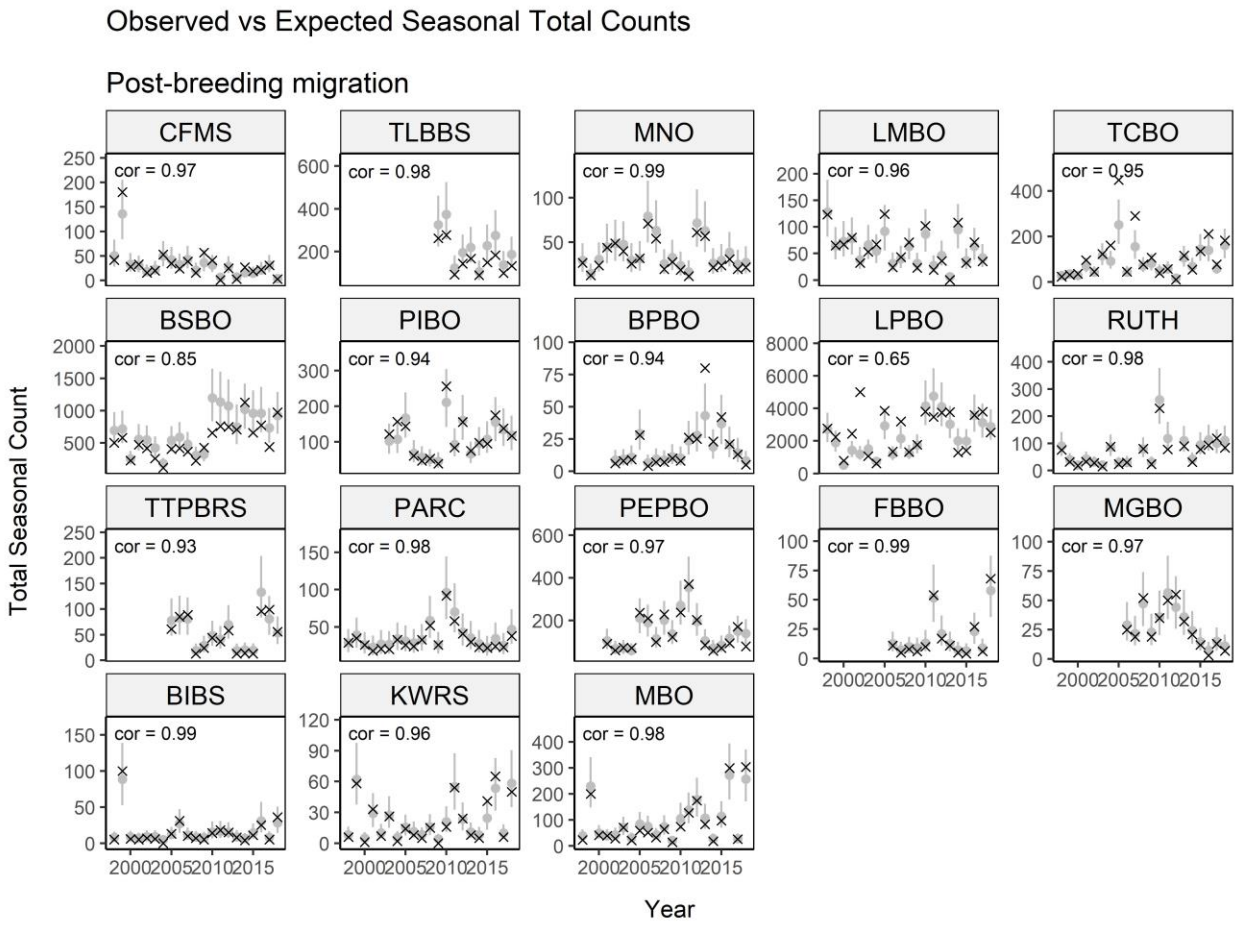



**Figure S4.5.** Comparison of observed (black "x") and expected (gray dots and whiskers) total seasonal counts in each season, at each monitoring station. Whiskers are 95% prediction intervals (i.e., 95% of observations are expected to fall inside those bounds). Pearson correlation between observed and expected (mean of posterior) counts are written in top left of each facet. Plots arranged in order from farthest west to farthest east.

*Literature Cited*

Kéry, M., & Royle, J. A. (2015). Applied Hierarchical Modeling in Ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 1: Prelude and Static Models. Academic Press.

## Appendix 4: Data and results of analysis of North American Breeding Bird Survey data

Strata and locations of routes contributing data to analysis of the North American Breeding Bird Survey (BBS) are illustrated in Figure S4.1. In total, Blackpoll Warblers were detected at 226 routes between 2000 and 2018, in 19 analytical strata.

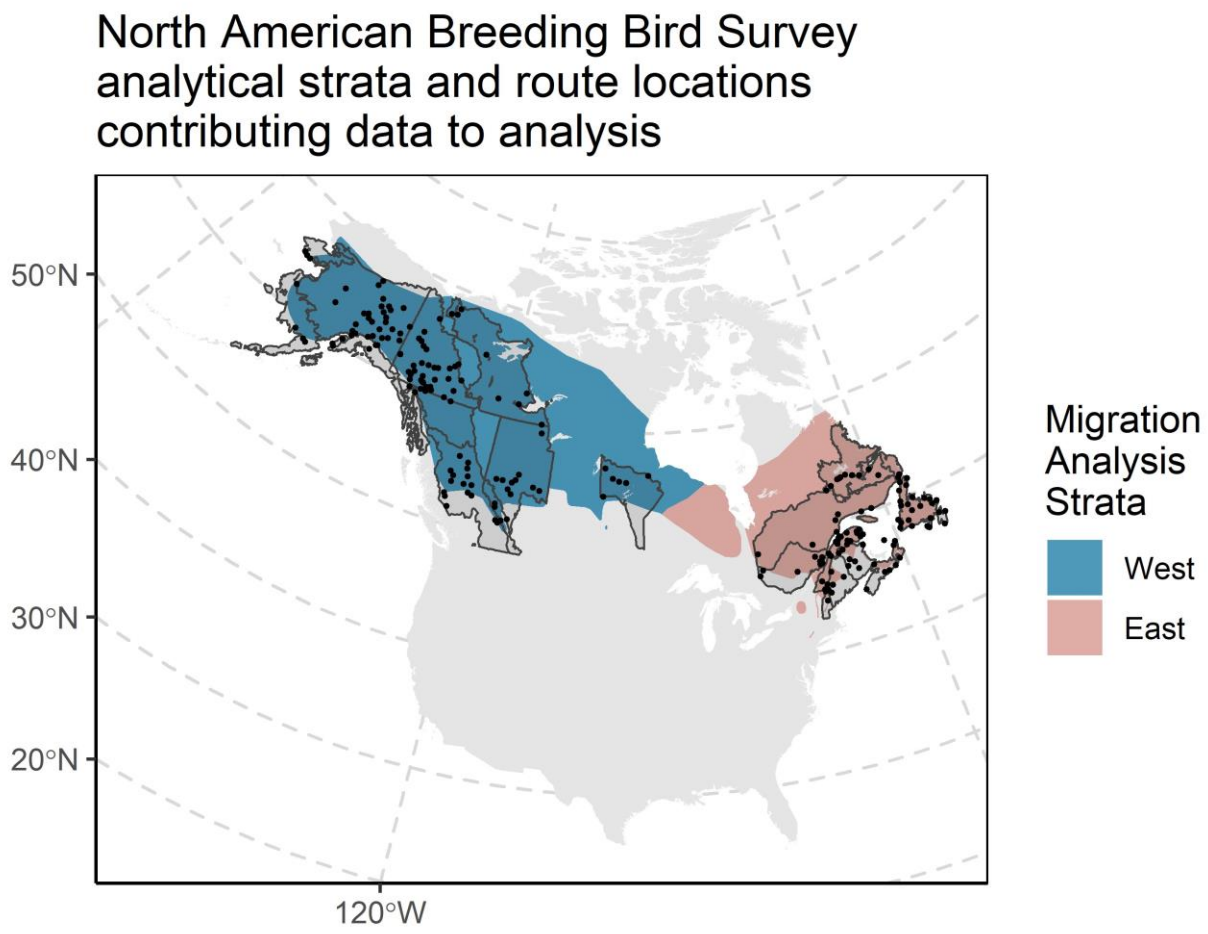



**Figure S4.1. Locations of routes contributing to continental trend analysis of Blackpoll Warbler between 2000 and 2018. USGS analytical strata are overlaid.**

North American Breeding Bird Survey data were analyzed using a 'first difference' model (Link et al. 2017; Smith et al. 2014), fit using the '*bbsBayes*' R package described in Edwards and Smith (2020). The estimated Blackpoll Warbler continental trajectory for the period of 2000-2018 is illustrated in figure S4.2.

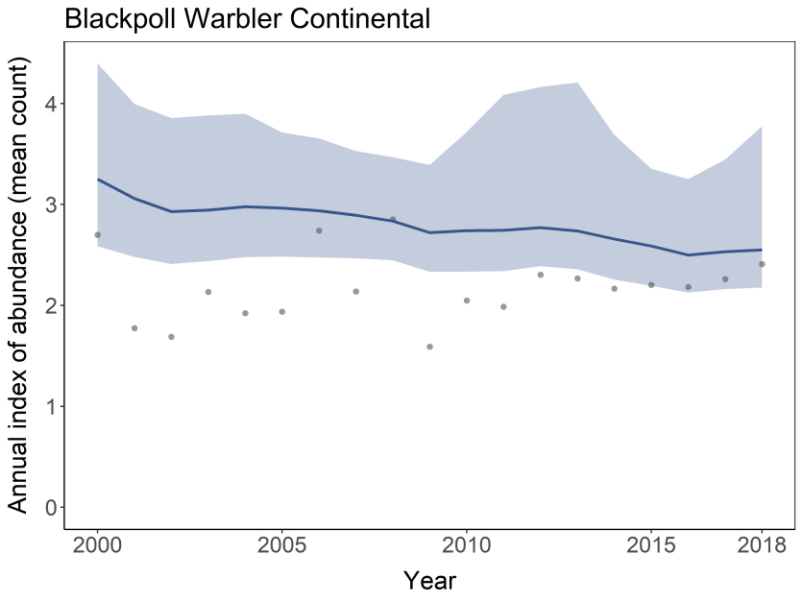



**Figure S4.2. Estimated continental trajectory of Blackpoll Warbler from 2000 to 2018 based on analysis of Breeding Bird Survey.**

The estimated trajectories of Blackpoll Warbler in eastern and western BBS strata are illustrated in Figure S3.3. Regional estimates were generated by aggregating area-weighted annual indices within USGS analytical strata, which were grouped in "east" and "west" categories based on geographic overlap with the corresponding strata used in migration analysis (see Figure S4.1). Methods for estimating trajectories within custom strata are described by Edwards and Smith (2020) and example code is available on https://github.com/bbsBayes/bbsBayes.

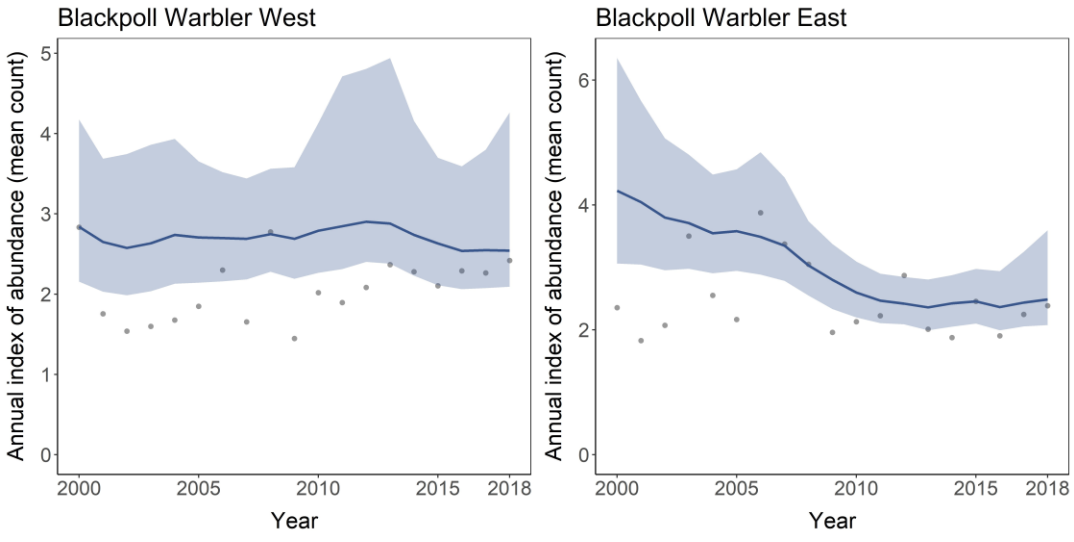



**Figure S4.3.** Estimated population trajectories of Blackpoll Warbler from 2000 to 2018 based on analysis of Breeding Bird Survey, by aggregating estimated annual indices in "eastern" and "western" strata.

*Literature Cited*

Edwards, B.P. and Smith, A.C., 2020. bbsBayes: An R package for hierarchical Bayesian analysis of North American breeding bird survey data. *Journal of Open Research Software 9(1).*

Link, W.A., Sauer, J.R. and Niven, D.K., 2017. Model selection for the North American Breeding Bird Survey: A comparison of methods. *The Condor: Ornithological Applications*, *119*(3), pp.546-556.

Smith, A.C., Hudson, M.A.R., Downes, C. and Francis, C.M., 2014. Estimating breeding bird survey trends and annual indices for Canada: How do the new hierarchical Bayesian estimates differ from previous estimates? *The Canadian Field-Naturalist*, *128*(2), pp.119-134.