ORIGINAL ARTICLE

Bayesian Skyline Plots disagree with range size changes based on Species Distribution Models for Holarctic birds

Eleanor F. Miller[1] | Rhys E. Green[1] | Andrew Balmford[1] |
Pierpaolo Maisano Delser[1] | Robert Beyer[1] | Marius Somveille[2] |
Michela Leonardi[1] | William Amos[1] | Andrea Manica[1]

[1]Department of Zoology, University of Cambridge, Cambridge, UK

[2]BirdLife International, Cambridge, UK

Correspondence
Eleanor F. Miller, Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK.
Email: em618@cam.ac.uk

Abstract

During the Quaternary, large climate oscillations impacted the distribution and demography of species globally. Two approaches have played a major role in reconstructing changes through time: Bayesian Skyline Plots (BSPs), which reconstruct population fluctuations based on genetic data, and Species Distribution Models (SDMs), which allow us to back-cast the range occupied by a species based on its climatic preferences. In this paper, we contrast these two approaches by applying them to a large data set of 102 Holarctic bird species, for which both mitochondrial DNA sequences and distribution maps are available, to reconstruct their dynamics since the Last Glacial Maximum (LGM). Most species experienced an increase in effective population size (N_e, as estimated by BSPs) as well as an increase in geographical range (as reconstructed by SDMs) since the LGM; however, we found no correlation between the magnitude of changes in N_e and range size. The only clear signal we could detect was a later and greater increase in N_e for wetland birds compared to species that live in other habitats, a probable consequence of a delayed and more extensive increase in the extent of this habitat type after the LGM. The lack of correlation between SDM and BSP reconstructions could not be reconciled even when range shifts were considered. We suggest that this pattern might be linked to changes in population densities, which can be independent of range changes, and caution that interpreting either SDMs or BSPs independently is problematic and potentially misleading.

KEYWORDS
Bayesian Skyline Plot, demographic history, Species Distribution Models

1 | INTRODUCTION

The Quaternary has been characterized by extensive cycles of glaciation and deglaciation. The legacy of these large-scale climate alterations is evident today in everything from species' genetic diversity to population structure (Svenning et al., 2015; Twitchett, 2006). Despite the profound impact of such changes on flora and fauna, there is limited quantitative evidence on which factors determined how different species fared during these cycles, or how species responded to post-glacial climate amelioration.

In the absence of direct evidence of species' prehistoric response to alterations in climatic conditions, sophisticated statistical methods have been developed to infer demographic history from more indirect sources. Different types of data, and consequently different approaches, are generally favoured by different fields. One of the most widely used genetic methods for inferring demographic history is the "skyline plot," a family of graphical, nonparametric methods introduced by Pybus et al. (2000). Grounded in the principles of Kingman's coalescent theory (Kingman, 1982), the "skyline-framework" uses DNA sequences to reconstruct a gene tree. The rate of coalescent events within it can be used to infer how the population size changed over time: periods of low coalescent rates imply a large population while a high density of coalescences implies a small population. Although skyline plots have been used to reconstruct demographic histories for extant and extinct species (Stiller et al., 2010), and across taxa that include vertebrates (Lu et al., 2012; Vignaud et al., 2014), invertebrates (Sanchez et al., 2016; Villalta et al., 2018) and even bacteria (Segawa et al., 2018), comparative multispecies studies are only now emerging (Burbrink et al., 2016).

Skyline plots have been used extensively to infer the response of species during the Last Glacial Maximum (LGM), and are often paired with climatic reconstructions to infer the changes in available habitat for a given species (Calderón et al., 2016; Foote et al., 2013; Lorenzen et al., 2012). One popular approach for reconstructing possible changes in available habitat for a species through time is bioclimatic Species Distribution Models (SDMs) (Elith & Leathwick, 2009). Modelling algorithms combine data on occurrences with environmental information to describe a species' current distribution and estimate how changes in environmental variables may have influenced this range through time. The underlying logic in linking these approaches is that, assuming limited population structure and appropriate sampling, a skyline plot could provide an indication of changes in total population size, which, in turn, may affect the extent of the range occupied by a species. However, the association between effective population sizes (N_e), as reconstructed by skyline plots, and species ranges is generally assumed, not tested.

There are many reasons why N_e might not be a good proxy for species ranges. Population structure is known to be a confounding effect, and the recommendation to counter it is to pool samples from multiple locations (Heller et al., 2013). However, even with this sampling scheme, there might be a mismatch in the two quantities if mean population density, and thus N_e, was affected by climate change differently to total range extent. For example, an increase in mean population density, and thus population size, might occur without a change in range, if the quality of habitat and its carrying capacity increased without a change in its extent (Figure 1a; Figures S1, S2) (Fordham et al., 2012). Given the positive relationship generally observed between range extent and mean local population density (Connor et al., 2000), N_e would also be expected to increase by a greater proportion than range extent under climatic amelioration. Another possibility is that, without substantial gene flow, skyline plots may reconstruct the population dynamics of the sampled locations rather than the whole species (Miller et al., 2018).

Another scenario that might lead to a disconnect between local N_e and range size arises during range shifts: sampled locations, suitable for a species today, might have been only marginally suitable in the past; that is, what is now the core area occupied by a species (where it is abundant, and sampling is more likely) might be inhabited by populations that in the past were at low densities because the area was only marginally suitable. Skylines from such populations would reveal a strong increase in N_e reflecting the local amelioration of conditions, irrespective of broader range changes (Figure 1b; Figures S1, S3). A similarly confounded signal will be found in the more extreme scenario where, as a result of a sizeable range shift, the sampled populations inhabit areas which were completely unsuitable in the past, and thus have undergone a founder event. Such populations would be characterized by an increase in N_e as they recovered from the local bottleneck associated with the founder event (Figure 1c; Figures S1, S4).
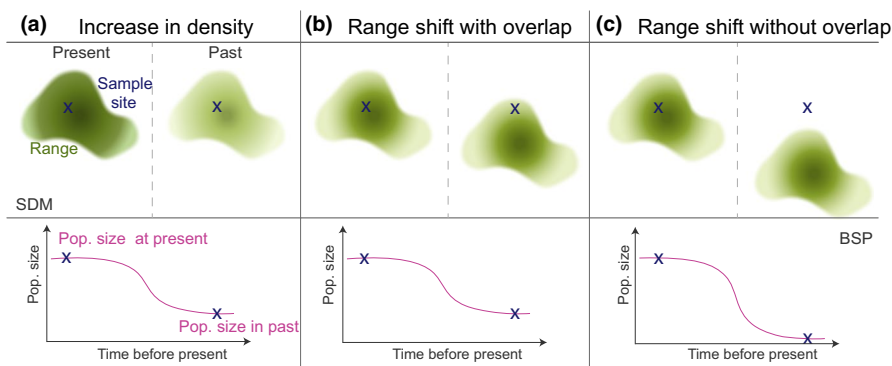
FIGURE 1 The top of each panel represents two schematic maps of the species range at the LGM (right) and present (left). Colour density within each virtual range shows population density. Blue crosses represent the geographically stable genetic sampling location. (a) Increasing population size without range change recovers an increasing BSP profile, as seen in the lower panel. (b) Core area today was only marginally suitable in the past, range size remains the same but local amelioration leads to a local increase in N_e. (c) Modern sampled populations inhabit areas outside the past species range. BSP recovers a steep increase in N_e associated with a founder event. See Material S1 for simulations supporting these expectations

The extent to which changes in population density and bottlenecks related to range shifts can override the signals linked to changes in range size and the overall metapopulation size is unknown. We collected from GenBank a comprehensive data set of mitochondrial DNA (mtDNA) sequences from many species of Holarctic birds, reconstructing their population dynamics using Bayesian Skyline Plots (BSPs). A simple prediction, based on the relative changes of habitat types as reconstructed from the pollen record (Allen et al., 2010), is that species associated with closed habitats (e.g., forests) should have increased since the LGM, whilst species from open and semiclosed habitats (grasslands and steppes) should show a decrease. However, species have more complex niche requirements than a simple association with a broad habitat type, and a more realistic prediction is that N_e should change in line with changes in extent of their geographical range. We therefore modelled changes in potential range size between the LGM and the present, using palaeoclimate reconstructions and SDMs, and investigated the relationship between changes in N_e and range size. Reconstructing individual ranges in the past is challenging because species may depend upon habitats whose extent is not easily reconstructed (e.g., wetlands); therefore, we also compared demographic profiles of species grouped by major biomes to investigate whether there was any consistent pattern in their response to climatic amelioration in the Holocene.

2 | MATERIALS AND METHODS

2.1 | Raw genetic data

We collated summary information on all available avian NADH dehydrogenase subunit 2 (*ND2*) and cytochrome b (*cytb*) sequences in GenBank, two of the most frequently uploaded avian mitochondrial genes, and screened for Holarctic species using the list of Voous (1977). Only species with more than 10 accessions for either gene were retained. When species had sufficient data for both *ND2* and *cytb*, sequences for each gene were extracted and handled as distinct data sets.

2.2 | Alignment

Sequence data for each species/gene combination come from multiple studies, so comparable regions were found by aligning sequences in MEGA (version 7.0; (Kumar et al., 2016)) using CLUSTALW (Thompson et al., 1994). Sequences for each taxon were trimmed to the longest common section between all samples. If inclusion of a single sequence required the loss of >200 bp from more than 50% of the other sequences, that sample was excluded as were all positions containing insertions, deletions or sequencing ambiguities. When studies uploaded one copy of each haplotype, associated frequency data were used to generate the appropriate number of copies in our database. If frequencies were unpublished the data were excluded. Sample sizes varied from 11 to 453 sequences per species, with lengths from 236 to 1,137 bp.

2.3 | Median joining networks (MJNs)

For each species/gene combination, we built an MJN in POPART (Leigh & Bryant, 2015). If the MJN contained long branches (single branch with 30 or more substitutions, indicative of profound population substructure) the species sampling location was reviewed. If clear geographical separation or grouping was found, the data were divided as appropriate and treated as discrete data sets. Single samples with >30 mutations on a branch were considered extreme outliers and dropped from alignments.

2.4 | Mutation rate

Recent work (Nabholz et al., 2016) proposes that body mass can be used to accurately calculate taxon-specific per year substitution rates and provides a correction factor for rate variation according to body mass as well as major mtDNA loci. We created data set-specific molecular evolution rates using average body mass data for each of our species (Dunning, 2007) together with the body mass/gene correction factors from Nabholz et al.'s (2016) "Calibration set 4" (3rd codon position). The equation was therefore $\frac{10^{(-0.247 * \log10(Mass) + 0.813)}}{100}$. A gene-specific correction factor (0.052 for *ND2*, −0.069 for *cytb*) was then applied (see Table S3). "Calibration set 4" was used as it includes younger species splits that should lead to estimates more appropriate for the within-species dynamics investigated here. Due to the uncertainty surrounding mutation rates, analyses based on "Calibration set 2" were also run (data not shown).

2.5 | BSP analysis

There now exist a range of related skyline plot methods (Ho & Shapiro, 2011). Here, we focus on BSP (Drummond et al., 2005) because its relative simplicity and inherent robustness make it popular in the field and particularly suitable for the quality of data sets in our study. For each data set, BSP analyses were implemented in BEAST2 (Bouckaert et al., 2014) using a strict clock with taxon-specific body mass/gene mutation rates, running 300 million steps sampled every 30,000 steps, with the first 10% discarded as burn-in. "*bModelTest*" was used to select the most appropriate site model and parameters for individual analyses (Bouckaert, 2015) whilst "*bGroupSizes*" and "*bPopSizes*" were set to three, and other parameters were left as default. Each analysis was run twice and convergence was verified by both a visual inspection of Markov chain Monte Carlo (MCMC) traces in TRACER v1.6 and confirming effective sample size values exceed 200 (Drummond & Rambaut, 2007). Demographic reconstructions were summarized in TRACER ("*Number of bins*" = 500), and plotted in R.

Sampling strategy can have important effects on BSP profiles (Chikhi et al., 2010; Heller et al., 2013; Städler et al., 2009). Geolocation information is only available for a small proportion of samples; for a subset of five species with large sample sizes and good geolocation coverage, we confirmed that using all samples provided

representative profiles compared to more structured subsampling (Materials S2, Figure S5A–E). Furthermore, we also confirmed (Materials S3, Figure S6) that, for these species, BSPs recovered profiles in line with those from Extended Bayesian Skyline Plots (EBSPs) (Heled & Drummond, 2008), which are more powerful but also require larger data sets (EBSPs were designed to take advantage of multilocus data).

2.6 | Inclusion criteria

Where data were available for both *ND2* and *cytb*, BSP profiles were compared along with summary statistics on each data set. When profiles agreed, the best supported data set was retained (e.g., largest sample size, longest sequences) to represent that species' history. If profiles were discordant but there was a clear disparity in the data sets' quality, we kept the profile from the stronger data set. If data quality did not vary but profiles were discordant, we conservatively rejected both profiles. For inclusion in further analysis, profiles needed to have a history deeper than 5000 years ago (ka) but shallower than 1 million years ago (deeper histories probably resulted from problematic fits). Also, because of the sample sizes and sequence lengths available for many species, we had to limit our BSP analysis to detect only one major size change (i.e., 3 change points). For this reason, if the largest change occurred before 60 ka, we excluded the species as we would have been blind to any smaller changes that occurred in the period of interest.

2.7 | Habitat classification

The species considered are associated with a wide range of habitat types, especially in terms of the predominant vegetation. Given the need for a small number of categories, no classification system can be perfect. We used the expert ornithological opinion of one of us (R.E.G.) to classify each species according to the major habitat with which it is currently associated, based upon descriptions from a standard work (BirdLife International & Handbook of the Birds of the World, 2018). After data-quality filtering, 138 species were available for classification. The habitat classes selected were Closed (forests), Semi-closed (shrubland and open woodlands), Open (grassland, montane and steppe) and Wetlands (freshwater wetlands). Some species could not be placed in them and there were other classes (e.g., Rivers) with five or fewer species; these were grouped into a separate category, "Other."

2.8 | Timing of expansion

Identifying a population's point of expansion from a BSP profile proved difficult given the wide range of shapes present: some population sizes changed little or gradually, while others showed sharp and/or multiple points of inflection. We attempted to develop algorithms that could automatically capture the inflection point but found them to be unreliable, and opted for visual scoring by E.F.M. Scoring was done prior to the running of SDMs, and thus blind to the expected changes from the range reconstructions.

2.9 | Size change

To compare the magnitude of estimated population expansions, the relative population size change between 21 ka and the present day was calculated. N_e at 21 ka was interpolated for each BSP profile and, where profiles were shorter than 21 ka, the population was assumed to be the size at the start of profile. Given the uncertainty of dating with molecular methods, and as there is a known risk of false signals in BSP profiles at very recent times (Heller et al., 2013), we also ran our analysis using size estimates from 60 ka/the earliest possible point (instead of 21 ka) to 5 ka (instead of 0 ka) (see Materials S4).

To compare average habitat profiles, we used a multispecies index (MSI) (Gregory et al., 2005). Here, equivalent adjacent time points are aligned and the geometric mean of the change in size calculated, based on all species with data. These average changes in size are then applied to a starting value of unity. Working back in time this creates directly comparable profiles where all groups reach a value of one at the present (see Material S5).

2.10 | Phylogenetic correction

All analyses included a phylogenetic correction based on the most complete molecular phylogeny of extant birds (www.birdtree.org, Jetz et al., 2012). Jetz et al. (2012) present two phylogenetic backbones, respectively Ericson and Hackett ("backbone-E" and "backbone-H"), considered equiprobable. For each backbone, we generated 1000 trees, randomly resolving polytomies in each, and repeated all analyses for each tree with a "*corPagel*" phylogenetic correction from the "ape" package in R. We also provide the results based ordinary least squares (i.e., without phylogenetic correction).

2.11 | Species Distribution Models

Our pipeline is provided as an R script in the Supporting Information.

2.11.1 | Climate reconstructions

To identify areas suitable for species through time, we used a 0.5° resolution data set for 19 bioclimatic variables (Net Primary productivity, Leaf Area Index and all the BioClim variables (Hijmans et al., 2005) excluding BIO2 and BIO3) that are available for the last 21,000 years in 1000-year time steps (Beyer et al., 2020). This data set was constructed from a combination of HadCM3 climate

simulations of the last 120,000 years, high-resolution HadAM3H simulations of the last 21,000 years and empirical present-day data. The data were downscaled and bias-corrected using the Delta Method (Beyer et al., 2019).

2.11.2 | Species data

Species occurrences were downloaded from the GBIF database (https://www.gbif.org) without any filtering (links in Table S4). Occurrences were then filtered based on coordinate accuracy (maximum error: 10 km), keeping only observations within the breeding and resident geographical ranges from Birdlife (BirdLife International, 2018). Points were regridded based on the palaeoclimatic reconstructions (0.5° x 0.5°) and, to work on presence/absence not frequency, only one presence per grid cell was kept.

For each species, this cleaned data set was used to select a subset of the bioclimatic variables available in the palaeoclimatic reconstructions (Beyer et al., 2020). We used a standardized approach in which we only removed highly correlated variables ($r > .7$), as these can lead to model instability (Guisan et al., 2017). A correlation matrix was constructed between the variables observed in the whole study area (Europe or North America). Uncorrelated variables (threshold = .7) were then selected by retaining the variable with the lowest overall correlation across the matrix where two values were highly correlated, leaving LAI, BIO4, BIO10, BIO14, BIO15, BIO18 and BIO19 in Eurasia and LAI, BIO5, BIO7, BIO8, BIO15 and BIO19 in North America.

To reduce the risk of geographical bias from opportunistic observations affecting the SDMs, the data set was thinned using the R package *spThin* (Aiello-Lammens et al., 2015), enforcing a minimum distance of 70 km between observations. Given the random nature of this process, it was repeated 100 times, with the result which kept the maximum number of observations retained for downstream analyses.

2.11.3 | Model fitting

SDM was performed using the R package *biomod2* (Thuiller et al., 2019) for all species with more than 10 occurrences after filtering and thinning (Stockwell & Peterson, 2002). The thinned data set was used as presences, the landmass of Eurasia or North America as background, and the same number of pseudo-absences as presences were randomly drawn five times from land outside the BirdLife resident and breeding masks, creating five independent data sets for further analysis. We found that drawing pseudo-absences in this way, from outside the masks, was the most effective strategy for retrieving SDMs consistent with species' modern-day ranges. By confirming that the estimated distributions recovered for the modern day were in accord with the BirdLife range predictions, we were confident that the modelled niche being projected into the past was as accurate as possible (Figure S7).

Following Bagchi et al. (2013), models were run independently for each of the five pseudoabsence data sets using four different algorithms: generalized linear models (GLMs), generalized boosting method (GBM), generalized additive models (GAMs) and random forest. Model evaluation was performed by spatial cross-validation (Roberts et al., 2017), namely splitting the data sets (presences and all five runs of pseudoabsences) based on latitudinal bands in America and longitudinal bands in Eurasia (Figure S8) with the R package *BlockCV* (Valavi et al., 2019), and using 4/5 of the splits to calibrate the model and the remaining 1/5 to evaluate it. As a data split containing only absences cannot be used for evaluation, and given the great variety of distribution of the species analysed, we maximized the probability of having at least some presences in all data splits by creating 15 spatial blocks encompassing the whole region of interest, either North America or Europe. Each block was given an ID, numbered sequentially 1–5 repeating, and the 15 blocks were assembled into these five working data splits grouped by the assigned ID numbers.

The models were run five times (once for each pseudoabsence run) for each of the four mentioned algorithms, using in turn four of the five defined data splits to calibrate and one to evaluate based on the TSS (threshold =0.7).

A full ensemble, combining all pseudo-absences sets and algorithms (Araújo & New, 2007), was built using only models with TSS > 0.7 averaged through four different statistics: mean, median, committee-average and weighted mean. The statistic showing the highest TSS was projected to either Eurasia or North America considering the present-day climate and the palaeoclimatic reconstruction for the LGM (defined as 21 ka). Such projections are visualized as maps, in which each cell shows the suitability for the species in a continuous range between 0 and 1, based on the threshold maximizing TSS.

2.11.4 | Range change and overlap

In order to calculate the range of each species, such continuous projections were transformed into binary (either 0—absence more likely, or 1—presence more likely) using 0.5 as a threshold. The binary projections were then used to estimate the climatically suitable area (km^2) for each species now and in the LGM. To do so, we reprojected the rasters to the Eckert IV equal-area pseudocylindrical projection setting the grid size to 50 × 50 km and multiplied the number of cells occupied in each period, and their overlap, by 2500 km^2 (cell area).

2.12 | Range size comparison between sample species and all Holarctic species

We used BirdLife data on the Extent of Occurrence (km^2) to define range size of all Holarctic bird species, those with a mean range latitude of above 20°N. We then compared this full data set to a subset of range sizes for the species included in our study using a Wilcoxon rank sum test.

3 | RESULTS

3.1 | Summary of available BSPs

A scan of GenBank yielded a preliminary data set of 208 species having a minimum of 10 individuals sequenced for either the *ND2* or *cytb* genes. Data sets were then discarded because of: insufficient haplotypes for demographic reconstruction (<5), insufficient sequence length (<200 bp), sequences across studies not from comparable gene sections, unpublished haplotype frequencies, inappropriate sampling strategies used by the original study (e.g., nonrandom) or extensive population substructure (full details in Materials and Methods). Application of these criteria left 167 data sets for BSP analyses. They were analysed with BEAST2 using a BSP and, with one exception (king eider, *Somateria spectabilis*), converged successfully. In BEAST2, we resized the "*bGroupSizes*" parameter, potentially constraining the level of detail recoverable in the profiles but providing comparable estimates between variable-quality data sets, analysed using the same settings.

Data for both *ND2* and *cytb* were available for 28 species. For 18 species, expansion times and profiles across them were consistent, and a single profile was selected to illustrate the population history. For 10 species, the two genes showed discordant demographic histories but, in seven cases, one data set was of appreciably lower quality (e.g., fewer samples, shorter sequences) and was removed. Three species were rejected because the two genes gave discordant profiles despite appearing to be of comparable quality (details in Table S1).

Further profiles had to be excluded as they were too deep (limit ≥ 1,000,000 years before the present, $n = 9$) or too short (limit ≤ 5,000 years before the present, $n = 4$) to be informative for the last deglaciation, or showed patterns of expansion/contraction that pre-dated the time period of interest, Marine Isotope Stage 3 (~60 ka, $n = 18$)(Van Meerbeeck et al., 2009). Notably, even among the accepted profiles, there remains great variation in depth due to the sparser and more stochastic branching patterns at the bases of the trees, which cause many profiles to truncate or to "flatline." Thus, the oldest population size estimates tend to be approximations because of the reduced information content impacting all profiles, and the need to use the points of truncation in short profiles or flatline states as the oldest size. For the two species where multiple lineages were identified (pine grosbeak, *Pinicola enucleator*, and horned lark, *Eremophila alpestris*), two separate BSP analyses were performed and an average of the estimates was taken for downstream analysis.

The above filters left 102 qualifying species BSP profiles. These species inhabit a wide range of habitats: Closed ($n = 43$), Open ($n = 17$), Semi-closed ($n = 25$), Wetlands ($n = 12$) and Other ($n = 5$). There was no indication that species associated with particular habitats were more/less likely to be excluded ($p = .77$, Fisher's exact test, excluding "Other" category as it had too few species for testing). Skyline profiles encompass a wide range of shapes, variously exhibiting single sharp inflection points, gradual changes in size and multiple points of change (see Figure S9). No significant differences were found in the proportion of *ND2/cytb* genes in each habitat type

($p = .78$, Fisher's exact test, excluding "Other" category), nor in the proportions of species from the Palearctic, Nearctic or Holarctic in each habitat type ($p = .10$, Fisher's exact test, excluding "Other" category).

3.2 | Direction and magnitude of demographic change

Only four out of 102 species showed an overall decrease in median N_e over time (21 ka to present), with 10 showing no sizeable change (<10% change), while all other species ($n = 88$) increased to some degree. The direction of change was not associated with habitat ($p = .3$, Fisher's exact test, excluding "Other" category), although the magnitude of change was (gls $p \leq .026$ for backbone-E and $p \leq .030$ for backbone-H, lm without phylogenetic correction $p = .028$; Figure 2a). This overall pattern of increase in N_e is rather extreme but could be the consequence of most species with sufficient samples in GenBank being relatively common and thus having thrived in the Holocene.

3.3 | Direction and magnitude of change in potential geographical range

To investigate the plausibility of changes in the extent of climatically suitable area contributing to the overwhelming majority of profiles showing an increase in N_e, we created individual SDMs for each of our species. We had credible SDMs for 96 species (Table S2); five species had to be excluded as there were insufficient observation points left for analysis after data thinning, and one species was rejected as its SDMs led to a present-day projection much larger than the observed range. For the valid SDMs, we generated ranges for the LGM (21 ka) and present day and quantified the changes in their size.

As was the case for N_e trajectories across the BSPs, many species showed an increase in reconstructed range extent since the LGM (76 of 96). However, the proportion of species showing an increase in range was significantly smaller than the proportion with increased N_e ($p = .009$, data subset to the 96 species for which both analyses were available). Whilst there was variation among groups of species associated with different habitats in the magnitude of change in range extent (gls $p \leq .0001$ for all 1,000 resolutions of backbone-E and backbone-H, lm without phylogenetic correction $p = .0001$, Figure 2b) and the direction of change in range extent ($p = .009$, Fisher's exact test, excluding "Other" category), there was no significant match between the direction of the trend in BSP and SDM reconstructions (Figure S10, Fisher's exact test $p = .572$). Neither was there a significant positive correlation across species between the signed magnitudes of the changes in the two measures (Figure 2c; gls $p \geq .994$ for all 1,000 resolutions of backbone-E and of backbone-H, lm without phylogenetic correction $p = .3707$). Furthermore, taking into account changes in the location of the range (i.e., the proportion
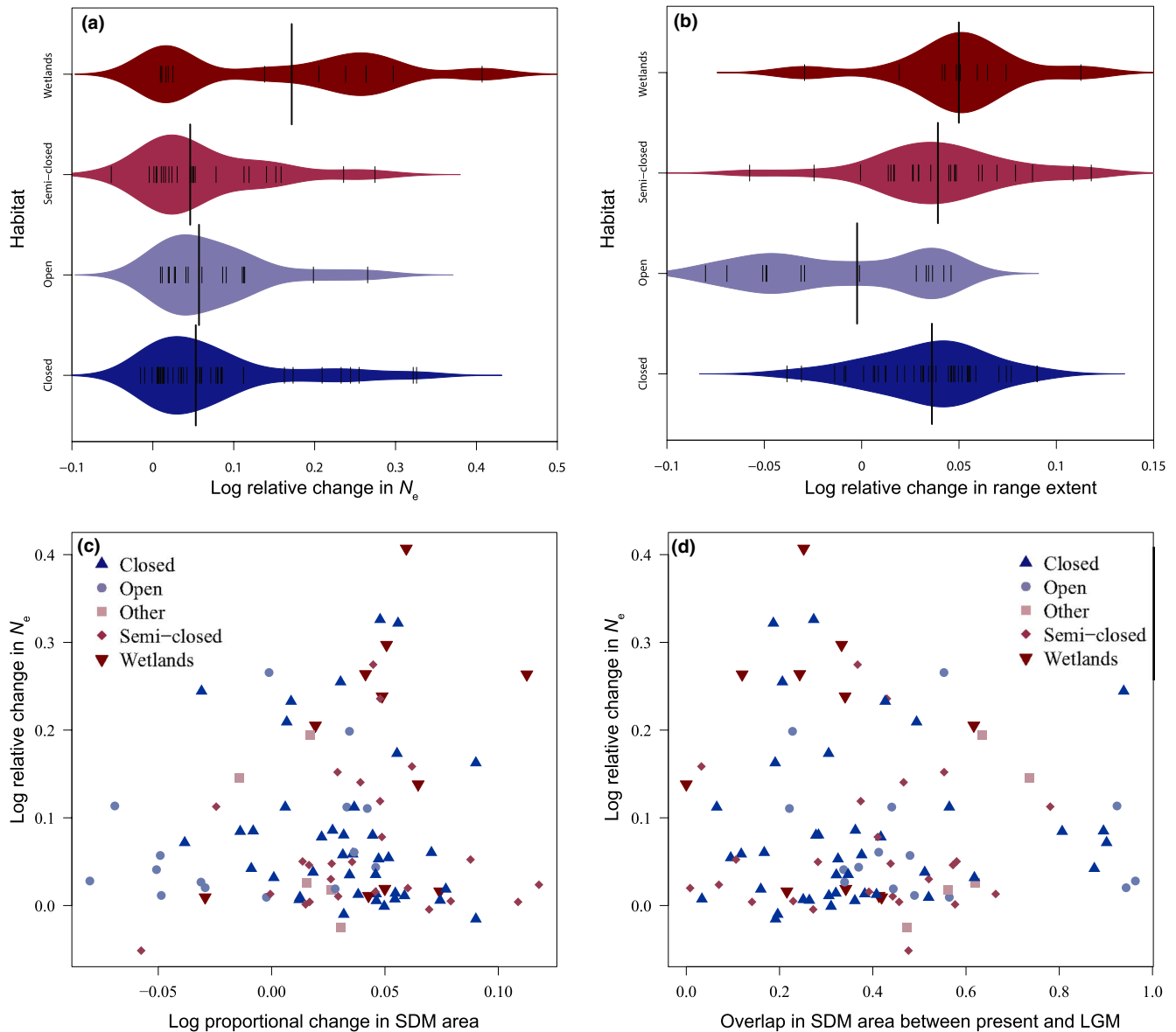
FIGURE 2 (a) Beanplot showing the log relative difference in N_e from 21 ka or start of the profile for species from each habitat type. Kernels represent density, small lines are individual species, and thick black lines are medians of species-specific values for the given habitat: Closed ($n = 43$), Open ($n = 17$), Semi-closed ($n = 25$), Wetlands ($n = 12$). (b) Beanplot showing the log relative difference in modelled range extent from 21 ka for species from each habitat: Closed ($n = 40$), Open ($n = 15$), Semi-closed ($n = 25$), Wetlands ($n = 11$). (c) Scatterplot of log ratio of N_e from 21 ka to present in relation to the log ratio of change in size of species' individual bioclimate SDM area from 21 ka to the present. (d) Scatterplot of log change in N_e in relation to the proportion of the species' modelled contemporary range that was also suitable during the LGM. In both scatterplots numbers of species are: Closed ($n = 40$), Open ($n = 15$), Other ($n = 5$), Semi-closed ($n = 25$), Wetlands ($n = 11$)

of overlap between LGM and present range) also failed to explain the changes in N_e as reconstructed by BSP (Figure 2D; gls $p \geq .734$ for backbone-E and backbone-H, lm without phylogenetic correction $p = .6917$).

3.4 | Timing of change

We next explored the relationship between the timing of the dominant population size change (since 60 ka) and habitat type (excluding

"Other" category, as it was heterogeneous and had five species). The timings of size change for each population in the four habitats are presented in Figure 3a. Major size change events in wetland-associated species tended to be more recent than for species from the other habitats (gls $p \leq .044$ in 1000 resolutions of backbone-E and backbone-H, lm without phylogenetic correction $p = .044$). Similar results were obtained excluding species which changed less than 10% in N_e (which might have added noise) (Figure S11). When using molecular evolution rates from Nabholz et al.'s (2016) "Calibration set 4," the timing of all expansions was generally consistent with a response

associated with the LGM. However, using the rate from "Calibration set 2" (which includes older nodes than set 4) would recover older expansion dates (data not shown); as such dates would correspond to periods of high ice coverage, they seem less likely.

For an alternative view of when the expansions occurred, we used an MSI. The MSI depicts normalized changes in size averaged across species within each habitat type for each time point (Figure 3b). Despite exploiting a different aspect of the BSP profile shape, mean change at each time point rather than a single mean date of maximum change, MSI profiles reveal a pattern that is strongly supportive of the previous result, where wetland species expand appreciably later than species in the other three habitats.

We also confirmed there was no correlation between body mass and timing of population size change ($p = .171$, Figure S12).
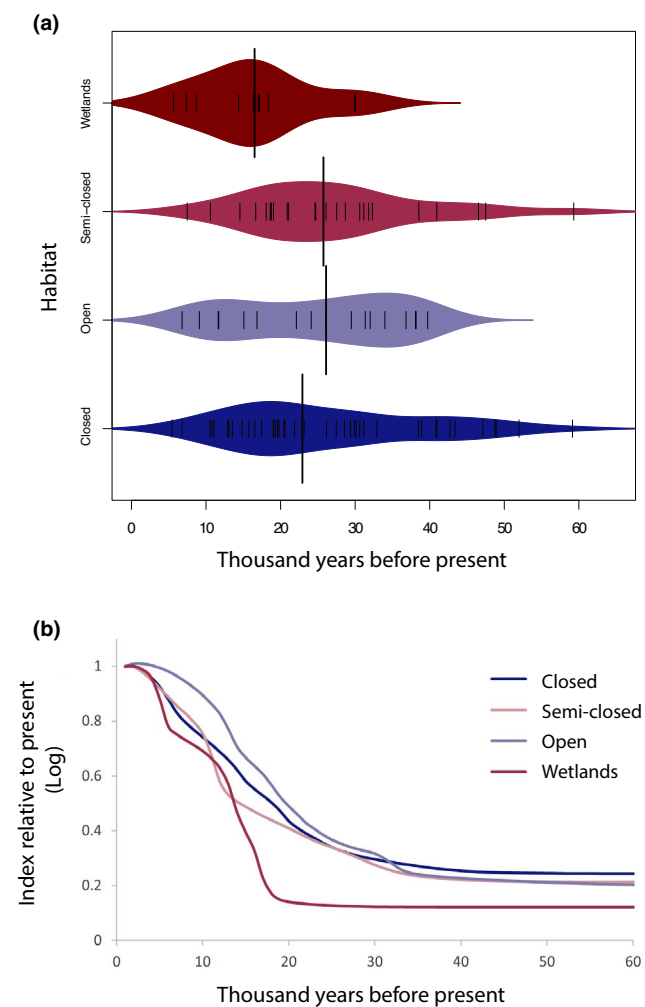

FIGURE 3 (a) Beanplot showing time of dominant effective population size change for species from each habitat type. Kernels represent density, and small lines the time of an individual species size change event (increase/decrease). Thick black line is median of species-specific change times for a given habitat: Closed ($n = 43$), Open ($n = 17$), Semi-closed ($n = 25$), Wetlands ($n = 12$). (b) MSI depicting normalized change in size averaged across species within each habitat for each time point

4 | DISCUSSION

We gathered a large collection of mtDNA data sets from many bird species to look for habitat-associated trends in population size through time. Although variable data quality may lead to uncertainties about the magnitude of any particular change in population size we detect, the direction of change is relatively robust (Grant, 2015; Grant et al., 2012). Of 102 species, only four species show an overall decrease in effective population size (N_e). Changes during the last deglaciation in the modelled extent of the geographical range also indicated increases for most species, though their proportion was lower than for N_e. However, we could find no association across species between the direction or magnitude of change in N_e and range reconstructions.

Species with large ranges at present are the probable "winners" in their response to climatic change and are more likely to show an increase in range and population size from the LGM. Widespread species might also be more likely to have been sampled, and indeed, we found that, based on BirdLife breeding range data, the species studied here have significantly larger modern-day ranges than Holarctic species as a whole (Wilcoxon rank sum $p < .001$, Figure S13). Most species sampled also showed an increase in range size based on SDMs. However, the proportion of expanding species according to SDMs was much lower than that observed for BSPs, thus failing to fully explain the ubiquity of expanding BSPs, and there was no statistical association between changes since the LGM as reconstructed by BSPs and SDMs.

Colonization bottlenecks during range shifts can, in principle, lead to an increase BSP irrespective of the overall change in range size, as long as migration is low enough (i.e., if the BSP captures the local dynamics in a given population/small geographical region rather than the whole range). However, if this mechanism were important, we would expect species exhibiting an increase in BSP despite a range contraction to be associated with large shifts. In fact, this is not the case for the majority of species with increasing BSP profiles and decreasing or stable SDM profiles, (Figure S14). Therefore, colonization bottlenecks do not seem to explain the ubiquity of expanding BSPs in our data set.

Increases in migration can also lead to an increasing BSP without any change in census population sizes. The potential role of migration in producing counterintuitive N_e estimates when assuming panmictic populations for a whole species has received much attention in the context of interpreting cross-coalescence (e.g., Multiple Sequentially Markovian Coalescent (MSMC)) profiles (Mazet et al., 2015). However, it is difficult to envisage a scenario where migration would increase significantly in the face of a range contraction; the effect of migration is more likely to be seen during an expansion, when previously isolated fragments are reconnected. Thus, it seems unlikely that migration can explain our results.

A final, more likely but difficult-to-test explanation is that population densities have increased since the LGM. Thus, even for species that have experienced a range contraction, there might have been changes in local population dynamics such that the average density

is higher at present. This decoupling of range extent and density makes interpretation of the SDMs and BSPs very challenging. To resolve them, there is a need to use species abundance models that explicitly predict population densities rather than presence/absence (Howard et al., 2014; Johnston et al., 2015; Potts & Elith, 2006). Whilst fitting such models is possible in principle, they have been little used because extensive population density information is rarely available. However, recent efforts such as those by the Cornell Lab of Ornithology (https://ebird.org/science/status-and-trends/; Auer et al., 2019) have started collating such information, opening a window to better understanding the link between range size and density. Our results, however, do raise a caveat in the interpretation of SMDs and BSPs for extinct species; arguably, the best strategy is to couple the two approaches, as only their combined results might provide a good overview of the fate of a species.

Besides the challenges of interpreting N_e reconstructions from BSPs, we also have to be cautious about the limitations of SDMs. SDMs assume that the range of a species is solely determined by a set of climatic variables that limits its expansion; in reality there are many other factors (such as availability of nesting sites, interactions with other species, availability of appropriate prey types) which might determine the distribution of a species (Araújo & Guisan, 2006). If those neglected factors are important enough and are not adequately predicted by (i.e., correlated with) a climatic variable, the SDMs will fail to properly reconstruct the past range. Furthermore, SDMs assume that the niche of a species has not changed through time, but some species might have adapted to different conditions over the last 20,000 years (Townsend, 2011; Veloz et al., 2012). Thus, we caution that some of the mismatch between BSPs and SDMs might also be attributed to shortcomings of the latter.

The timings we find for when population expansions occurred agree broadly with those of changes in climate after the LGM. Dates were based on mitochondrial mutation rates calibrated for body size and based on a calibration set that included relatively young species splits (Nabholz et al., 2016), and thus likely to give faster mutation rates (i.e., less affected by selection) that were appropriate for within-species analysis (Ho, 2007; Howell et al., 2003; Penny, 2005). Using a calibration set that included older splits (Nabholz et al., 2016) would lead to much older (well before the last glaciation), and thus less realistic changes in N_e. However, we strongly caution that mutation rates calibrated by bird body size, whilst the best available option for comparative analysis, are likely to be noisy, and individual species estimates should not be overly interpreted. Ideally, we would need taxon-specific mutation rates (Hope et al., 2014) which are not available for the number of species investigated here. Having said this, the fact that species from the same habitat tend to yield broadly similar profiles gives us confidence that the relative timings are robust, even if the absolute values have room for improvement.

We found that changes in N_e for wetland-associated species have occurred more recently, and have been significantly larger, than those from terrestrial habitats. Although the significance is marginal when based on point estimates for the date of most rapidly changing size, the finding is supported by MSI analysis, which also reveals

a similar pattern. Compared to many terrestrial habitats, wetlands tend to be less stable: factors such as local water levels, meltwater from retreating ice sheets, and soil erosion all play into wetland development and could have delayed their stabilization after the LGM. Although reconstruction of wetland environments and the modelling of wetland recovery is difficult (Fan & Miguez-Macho, 2011; Kaplan, 2002; Lafleur, 2008; Valdes et al., 2005), analysis of pollen across Eurasia shows that species associated with wetlands such as sphagnum moss and alder trees exhibit much later expansions compared with terrestrial species (Allen et al., 1999; Giesecke et al., 2017). Indeed, the expansion of alder relative to other trees (Allen et al., 1999) matches closely the later expansion we find for wetland vs. nonwetland birds.

BSP analysis is powerful but depends on a number of assumptions that are rarely met in real data, most notably the use of a random sample of individuals drawn from a panmictic population (Chikhi et al., 2010; Heller et al., 2013; Pannell, 2003). Consequently, most profiles should be seen as approximations that are easy to over-interpret (Grant, 2015; Miller et al., 2018). However, increasing numbers of public domain data sets allow studies based on characteristics averaged across multiple profiles constructed from species or populations that share habitats or traits. This approach has its own challenges and the data need stringent filtering. Here, we constructed and inspected network diagrams for each data set, identifying species with genetic outliers and evidence of strong population substructure to be divided or excluded. It is noteworthy that the use of public domain data sets can additionally restrict the level of control a study has on sampling scheme, a key tool for limiting the impact of violating the panmixia assumption (Chikhi et al., 2010; Heller et al., 2013). We explored the impact of sampling scheme on five of the data sets from this study and found the shape and trend of the median line to be generally robust to dramatic alterations in sampling scheme (Figure S5A–E). Although we recognize that structural biases undoubtedly exist in the data uploaded to GenBank, this gave us confidence that structural bias was not driving the demographic patterns recovered. However, the large number of species investigated allowed us to see a clear pattern of population expansion in almost all species following the LGM, irrespective of their range dynamics inferred from SDMs, and a tendency for the expansion to occur later in wetland species. The near-ubiquitous signal of expansion suggests a possible decoupling of range size and local densities, implying a need for carefully interpretation of BSP and SDMs to describe species-wide responses.

CONFLICT OF INTEREST

The authors declare no competing financial interests.

AUTHOR CONTRIBUTIONS

E.F.M. performed the analyses, produced the plots and wrote the manuscript. All SDMs were constructed by M.L. Expert ornithological knowledge was provided by R.E.G., who also proposed the MSI analysis. All authors provided supervision and comments on the manuscript.

DATA AVAILABILITY STATEMENT

Genetic sequence data used in this study were downloaded from the GenBank website —https://www.ncbi.nlm.nih.gov/genbank/. Species occurrence data were downloaded from the Global Biodiversity Information Facility (GBIF) database; DOIs are available for each data set in the Table S3.

ORCID

Eleanor F. Miller https://orcid.org/0000-0002-3213-5714
Rhys E. Green https://orcid.org/0000-0001-8690-8914
Andrew Balmford https://orcid.org/0000-0002-0144-3589
Pierpaolo Maisano Delser https://orcid.org/0000-0002-1844-1715
Marius Somveille https://orcid.org/0000-0002-6868-5080
Michela Leonardi https://orcid.org/0000-0001-8933-9374
William Amos https://orcid.org/0000-0002-0971-9914
Andrea Manica https://orcid.org/0000-0003-1895-450X

REFERENCES

Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B., & Anderson, R. P. (2015). spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography (Cop.)*, 38, 541–545. https://doi.org/10.1111/ecog.01132

Allen, J. R. M. et al (1999). Rapid environmental changes in southern Europe during the last glacial period. *Nature*, 400, 740–743. https://doi.org/10.1038/23432

Allen, J. R. M., Hickler, T., Singarayer, J. S., Sykes, M. T., Valdes, P. J., & Huntley, B. (2010). Last glacial vegetation of Northern Eurasia. *Quaternary Science Reviews*, 29, 2604–2618. https://doi.org/10.1016/j.quascirev.2010.05.031

Araújo, M. B., & Guisan, A. (2006). Five (or so) challenges for species distribution modelling. *Journal of Biogeography*, 33, 1677–1688. https://doi.org/10.1111/j.1365-2699.2006.01584.x

Araújo, M. B., & New, M. (2007). Ensemble forecasting of species distributions. *Trends in Ecology & Evolution*, 22, 42–47. https://doi.org/10.1016/j.tree.2006.09.010

Auer, T., Fink, D., & Strimas-Mackey, M. (2019). ebirdst: Tools for loading, plotting, mapping and analysis of eBird Status and Trends data products. R package version 0.1.0.

Bagchi, R. et al (2013). Evaluating the effectiveness of conservation site networks under climate change: Accounting for uncertainty. *Global Change Biology*, 19, 1236–1248. https://doi.org/10.1111/gcb.12123

Beyer, R., Krapp, M., & Manica, A. (2019). A systematic comparison of bias correction methods for paleoclimate simulations. *Climate of the Past Discussion*, 1–23, https://doi.org/10.5194/cp-2019-11

Beyer, R. M., Krapp, M., & Manica, A. (2020). High-resolution terrestrial climate, bioclimate and vegetation for the last 120,000 years. *Sci. Data*, 7, 1–9. https://doi.org/10.1038/s41597-020-0552-1

BirdLife International. (2018). *BirdLife International and Handbook of the Birds of the World. 2018. Bird species distribution maps of the world. Version 2018.1.* See http://datazone.birdlife.org/species/requestdis

BirdLife International and Handbook of the Birds of the World (2018). Bird species distribution maps of the world. Version 2018.1.

Bouckaert, R. (2015). bModelTest : Bayesian Site Model Selection for Nucleotide Data. *bioRxiv*, 1–17. https://doi.org/10.1101/020792

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C. H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 10, 1–6. https://doi.org/10.1371/journal.pcbi.1003537

Burbrink, F. T., Chan, Y. L., Myers, E. A., Ruane, S., Smith, B. T., & Hickerson, M. J. (2016). Asynchronous demographic responses to Pleistocene climate change in Eastern Nearctic vertebrates. *Ecology Letters*. https://doi.org/10.1111/ele.12695

Calderón, L. et al (2016). Genomic evidence of demographic fluctuations and lack of genetic structure across flyways in a long distance migrant, the European turtle dove. *BMC Evolutionary Biology*, 16, 237. https://doi.org/10.1186/s12862-016-0817-7

Chikhi, L., Sousa, V. C., Luisi, P., Goossens, B., & Beaumont, M. A. (2010). The confounding effects of population structure, genetic diversity and the sampling scheme on the detection and quantification of population size changes. *Genetics*, 186, 983–995. https://doi.org/10.1534/genetics.110.118661

Connor, E. F., Courtney, A. C., & Yoder, J. M. (2000). Individuals-area relationships: The relationship between animal population density and area. *Ecology*, 81, 734. https://doi.org/10.2307/177373

Drummond, A. J., & Rambaut, A. (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology*, 7, 214. https://doi.org/10.1186/1471-2148-7-214

Drummond, A. J., Rambaut, A., Shapiro, B., & Pybus, O. G. (2005). Bayesian coalescent inference of past population dynamics from molecular sequences. *Molecular Biology and Evolution*, 22, 1185–1192. https://doi.org/10.1093/molbev/msi103

Dunning, J. (2007). *CRC handbook of avian body masses*, 2nd ed. CRC Press.

Elith, J., & Leathwick, J. R. (2009). Species distribution models: Ecological explanation and prediction across space and time. *Annual Review of Ecology Evolution and Systematics*, 40, 677–697. https://doi.org/10.1146/annurev.ecolsys.110308.120159

Fan, Y., & Miguez-Macho, G. (2011). A simple hydrologic framework for simulating wetlands in climate and earth system models. *Climate Dynamics*, 37, 253–278. https://doi.org/10.1007/s00382-010-0829-8

Foote, A. D. et al (2013). Ancient DNA reveals that bowhead whale lineages survived Late Pleistocene climate change and habitat shifts. *Nature Communications*, 4, 1677. https://doi.org/10.1038/ncomms2714

Fordham, D. A. et al (2012). Plant extinction risk under climate change: are forecast range shifts alone a good indicator of species vulnerability to global warming? *Global Change Biology*, 18, 1357–1371. https://doi.org/10.1111/j.1365-2486.2011.02614.x

Giesecke, T., Brewer, S., Finsinger, W., Leydet, M., & Bradshaw, R. H. W. (2017). Patterns and dynamics of European vegetation change over the last 15,000 years. *Journal of Biogeography*, 44, 1441–1456. https://doi.org/10.1111/jbi.12974

Grant, W. S. (2015). Problems and cautions with sequence mismatch analysis and Bayesian skyline plots to infer historical demography. *Journal of Heredity*, 106, 333–346. https://doi.org/10.1093/jhered/esv020

Grant, W. S., Liu, M., Gao, T., & Yanagimoto, T. (2012). Limits of Bayesian skyline plot analysis of mtDNA sequences to infer historical demographies in Pacific herring (and other species). *Molecular Phylogenetics and Evolution*, 65, 203–212. https://doi.org/10.1016/j.ympev.2012.06.006

Gregory, R. D., Van Strien, A., Vorisek, P., Meyling, A. W. G., Noble, D. G., Foppen, R. P. B., & Gibbons, D. W. (2005). Developing indicators for

European birds. *Philosophical Transaction of the Royal Society B Biol. Sci.*, *360*, 269–288. https://doi.org/10.1098/rstb.2004.1602

Guisan, A., Thuiller, W., & Zimmermann, N. E. (2017). *Habitat suitability and distribution models*. Cambridge University Press. https://doi.org/10.1017/9781139028271

Heled, J., & Drummond, A. J. (2008). Bayesian inference of population size history from multiple loci. *BMC Evolutionary Biology*, *15*, 1–15. https://doi.org/10.1186/1471-2148-8-289.

Heller, R., Chikhi, L., & Siegismund, H. R. (2013). The confounding effect of population structure on Bayesian skyline plot inferences of demographic history. *PLoS One*, *8*, e62992. https://doi.org/10.1371/journal.pone.0062992

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, *25*, 1965–1978. https://doi.org/10.1002/joc.1276

Ho, S. Y. M. (2007). Calibrating molecular estimates of substitution rates and divergence times in birds. *Journal of Avian Biology*, *38*, 409–414. https://doi.org/10.1111/j.2007.0908-8857.04168.x

Ho, S. Y. W., & Shapiro, B. (2011). Skyline-plot methods for estimating demographic history from nucleotide sequences. *Molecular Ecology Resources*, *11*, 423–434. https://doi.org/10.1111/j.1755-0998.2011.02988.x

Hope, A. G., Ho, S. Y. W., Malaney, J. L., Cook, J. A., & Talbot, S. L. (2014). Accounting for rate variation among lineages in comparative demographic analyses. *Evolution (N., Y).* *68*, 2689–2700. https://doi.org/10.1111/evo.12469

Howard, C., Stephens, P. A., Pearce-Higgins, J. W., Gregory, R. D., & Willis, S. G. (2014). Improving species distribution models: The value of data on abundance. *Methods in Ecology and Evolution*, *5*, 506–513. https://doi.org/10.1111/2041-210X.12184

Howell, N., Smejkal, C. B., Mackey, D. A., Chinnery, P. F., Turnbull, D. M., & Herrnstadt, C. (2003). The Pedigree Rate of Sequence Divergence in the Human Mitochondrial Genome: There Is a Difference Between Phylogenetic and Pedigree Rates. *American Journal of Human Genetics*, *72*, 659–670. https://doi.org/10.1086/368264

Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K., & Mooers, A. O. (2012). The global diversity of birds in space and time. *Nature*, *491*, 444–448. https://doi.org/10.1038/nature11631

Johnston, A. et al (2015). Abundance models improve spatial and temporal prioritization of conservation resources. *Ecological Applications*, *25*, 1749–1756. https://doi.org/10.1890/14-1826.1

Kaplan, J. O. (2002). Wetlands at the Last Glacial Maximum: Distribution and methane emissions. *Geophysical Research Letters*, *29*, 3-1–3-4. https://doi.org/10.1029/2001GL013366

Kingman, J. F. C. (1982). On the genealogy of large populations. *Journal of Applied Probability*, *19*, 27. https://doi.org/10.2307/3213548

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution*, *33*, 1870–1874. https://doi.org/10.1093/molbev/msw054

Lafleur, P. M. (2008). Connecting Atmosphere and Wetland: Energy and Water Vapour Exchange. *Geography Compass*, *2*, 1027–1057. https://doi.org/10.1111/j.1749-8198.2007.00132.x

Leigh, J. W., & Bryant, D. (2015). Popart: Full-Feature Software for Haplotype Network Construction. *Methods in Ecology and Evolution*, *6*, 1110–1116. https://doi.org/10.1111/2041-210X.12410

Lorenzen, E. D. et al (2012). Species-specific responses of LateQuaternary megafauna to climate and humans. *Nature*, *479*, 359–364. https://doi.org/10.1038/nature10574

Lu, B., Zheng, Y., Murphy, R. W., & Zeng, X. (2012). Coalescence patterns of endemic Tibetan species of stream salamanders (Hynobiidae: Batrachuperus). *Molecular Ecology*, *21*, 3308–3324. https://doi.org/10.1111/j.1365-294X.2012.05606.x

Mazet, O., Rodríguez, W., & Chikhi, L. (2015). Demographic inference using genetic data from a single individual: Separating population

size variation from population structure. *Theoretical Population Biology*, *104*, 46–58. https://doi.org/10.1016/j.tpb.2015.06.003

Miller, E. F., Manica, A., & Amos, W. (2018). Global demographic history of human populations inferred from whole mitochondrial genomes. *Royal Society of Open Science*, *5*, 180543. https://doi.org/10.1098/rsos.180543

Nabholz, B., Lanfear, R., & Fuchs, J. (2016). Body mass-corrected molecular rate for bird mitochondrial DNA. *Molecular Ecology*, *25*, 4438–4449. https://doi.org/10.1111/mec.13780

Pannell, J. R. (2003). Coalescence in a Metapopulation with Recurrent Local Extinction and Recolonzation. *Evolution (N., Y).* *57*, 949–961. https://doi.org/10.1111/j.0014-3820.2003.tb00307.x.

Penny, D. (2005). Relativity for molecular clocks. *Nature*, *436*, 183–184. https://doi.org/10.1038/436183a

Potts, J. M., & Elith, J. (2006). Comparing species abundance models. *Ecological Modelling*, *199*, 153–163. https://doi.org/10.1016/j.ecolmodel.2006.05.025

Pybus, O. G., Rambaut, A., & Harvey, P. H. (2000). An integrated framework for the inference of viral population history from reconstructed genealogies. *Genetics*, *155*, 1429–1437.

Roberts, D. R. et al (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography (Cop.)*, *40*, 913–929. https://doi.org/10.1111/ecog.02881

Sanchez, G., Tomano, S., Yamashiro, C., Fujita, R., Wakabayashi, T., Sakai, M., & Umino, T. (2016). Population genetics of the jumbo squid Dosidicus gigas (Cephalopoda: Ommastrephidae) in the northern Humboldt Current system based on mitochondrial and microsatellite DNA markers. *Fisheries Research*, *175*, 1–9. https://doi.org/10.1016/j.fishres.2015.11.005

Segawa, T., Takeuchi, N., Fujita, K., Aizen, V. B., Willerslev, E., & Yonezawa, T. (2018). Demographic analysis of cyanobacteria based on the mutation rates estimated from an ancient ice core. *Heredity (Edinb)*, *1–12*, https://doi.org/10.1038/s41437-017-0040-3

Städler, T., Haubold, B., Merino, C., Stephan, W., & Pfaffelhuber, P. (2009). The impact of sampling schemes on the site frequency spectrum in nonequilibrium subdivided populations. *Genetics*, *182*, 205–216. https://doi.org/10.1534/genetics.108.094904

Stiller, M. et al (2010). Withering away-25,000 years of genetic decline preceded cave bear extinction. *Molecular Biology and Evolution*, *27*, 975–978. https://doi.org/10.1093/molbev/msq083.

Stockwell, D. R., & Peterson, A. T. (2002). Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, *148*, 1–13. https://doi.org/10.1016/S0304-3800(01)00388-X

Svenning, J.-C., Eiserhardt, W. L., Normand, S., Ordonez, A., & Sandel, B. (2015). The influence of paleoclimate on present-day patterns in biodiversity and ecosystems. *Annual Review of Ecology Evolution and Systematics*, *46*, 551–572. https://doi.org/10.1146/annurev-ecolsys-112414-054314

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*, 4673–4680. https://doi.org/10.1093/nar/22.22.4673

Thuiller, W., Georges, D., Engler, R., & Breiner, F. (2019) biomod2: Ensemble Platform for Species Distribution Modeling.

Townsend, P. A. (2011). Ecological niche conservatism: A time-structured review of evidence. *Journal of Biogeography*, *38*, 817–827. https://doi.org/10.1111/j.1365-2699.2010.02456.x

Twitchett, R. J. (2006). The palaeoclimatology, palaeoecology and palaeoenvironmental analysis of mass extinction events. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *232*, 190–213. https://doi.org/10.1016/j.palaeo.2005.05.019

Valavi, R., Elith, J., Lahoz-Monfort, J. J., & Guillera-Arroita, G. (2019). blockCV: An *r* package for generating spatially or environmentally separated folds for k-fold cross-validation of species distribution

models. *Methods in Ecology and Evolution*, 10, 225–232. https://doi.org/10.1111/2041-210X.13107

Valdes, P. J., Beerling, D. J., & Jonhson, C. E. (2005). The ice age methane budget. *Geophysical Research Letters*, 32, 1–4. https://doi.org/10.1029/2004GL021004

Van Meerbeeck, C. J., Renssen, H., & Roche, D. M. (2009). How did Marine Isotope Stage 3 and Last Glacial Maximum climates differ? - Perspectives from equilibrium simulations. *Climate of the Past*, 5, 33–51. https://doi.org/10.5194/cp-5-33-2009

Veloz, S. D., Williams, J. W., Blois, J. L., He, F., Otto-Bliesner, B., & Liu, Z. (2012). No-analog climates and shifting realized niches during the late quaternary: Implications for 21st-century predictions by species distribution models. *Global Change Biology*, 18, 1698–1713. https://doi.org/10.1111/j.1365-2486.2011.02635.x

Vignaud, T. M. et al (2014). Genetic structure of populations of whale sharks among ocean basins and evidence for their historic rise and recent decline. *Molecular Ecology*. https://doi.org/10.1111/mec.12754

Villalta, I., Amor, F., Galarza, J. A., Dupont, S., Ortega, P., Hefetz, A., Dahbi, A., Cerdá, X., & Boulay, R. (2018). Origin and distribution of desert ants across the Gibraltar Straits. *Molecular Phylogenetics and Evolution*, 118, 122–134. https://doi.org/10.1016/j.ympev.2017.09.026

Voous, K. H. (1977) List of recent Holarctic bird species. *Ibis (Lond. 1859).*

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Miller, E. F., Green, R. E., Balmford, A., Maisano Delser, P., Beyer, R., Somveille, M., Leonardi, M., Amos, W., & Manica, A. (2021). Bayesian Skyline Plots disagree with range size changes based on Species Distribution Models for Holarctic birds. *Molecular Ecology*, 30, 3993–4004. https://doi.org/10.1111/mec.16032