Title: Rising role of phenotypic plasticity explains changes in avian body size distributions

Authors: Joshua Kenneth Robertson Tabh[1*], Pierre Nyquist[2], Andreas Nord[1]

Affiliations: [1] Section for Evolutionary Ecology and Infections Biology,
Department of Biology, Lund University.
Sölvegatan 37, 223 62 Lund, Sweden

[2] Section for Applied Mathematics and Statistics,
Department of Mathematical Sciences,
Chalmers University of Technology.
Chalmers Tvärgata 3, Göteborg, Sweden

Contact Information: joshua.tabh@biol.lu.ca

Conflicts of interest: The authors have no competing interests to report.

Funding statement: Funding for J.K.R.T. was provided by the Wenner-Gren foundation (UPD2021-0038), the Sven och Lily Lawski foundation (20240523), Stiftelsen Lunds Djurskyddsfond, and Lars Hierta Minne (20221121). Funding for P.N. was provided by the Swedish Research Council (018-07050 and 2023-03484) and a Wallenberg AI, Autonomous Systems and Software Program (WASP) awarded through the Knut and Alice Wallenberg Foundation. A.N. was funded by the Swedish Research Council (2020-04686 and 2024-05362).

Data availability statement: All statistical code used in this study will be available at datadryad upon review. All data analysed in this study are freely available through the MAPS program (https://www.birdpop.org/pages/maps.php) and the United States Geological Survey (USGS) Bird Banding Lab (https://www.usgs.gov/labs/bird-banding-laboratory/data).

Acknowledgements: We thank Cecilia Lundberg, Jan-Åke Nilsson, Fredrik Andreasson, Elin Persson, Martin Stjerman and Erik Svensson for their useful insights during the construction of this manuscript.

Summary

Natural selection overwhelmingly favours the largest within animal species. Why average body sizes in the world's temperate birds have declined is therefore both mechanistically confusing and ecologically worrying. We argue that understanding the causes of these declines requires knowledge of how complete, body mass probability distributions are moving across space and time. Using data from 159 North American bird species – spanning 26 families, 53 degrees of latitude, and 25 years of observation since the early 1990s – we show that body mass distributions are not only downgrading, but also becoming narrower, and rapidly losing their upper tails (i.e. the largest individuals), with the most dramatic changes occurring at species' southern range limits. Combining fitness modelling with a novel, distributional variant of the Price equation revealed that these changes are increasingly explained by phenotypic plasticity and not poleward range-shifts, or selection for or against certain body masses within species. Growing contributions of genetic drift are also dismissed. Together, these results show that disproportionate losses of the largest body sizes best explain declining averages within species, and that these losses may well be reversible. Diminishing resource accessibility from rising breeding season temperatures, particularly at lower range limits, are most consistent with these findings.

Introduction

"The Big Snake [Bashe] eats elephants and after three years it disgorges their bones. Gentlemen take a dose of this snake so that they will never have heart disease or illnesses of the belly."

- Shanhai jing, chapter 10

"[The behemoth] moveth his tail like a cedar: the sinews of his stones are wrapped together. His bones are as strong pieces of brass; his bones are like bars of iron. He is the chief of the ways of God..."

-Job 40:17-19

"We're not big, but we're small."

- Stuart MacLean, *in re* The Vinyl Cafe

Although a motley collection, this triple exergue reveals the almost salient connection between extreme size – be it large or small – and power and greatness in the human perception. While the reason for this connection is debatable, one might argue that its roots lie in evolutionary reality. Among animals, body size is almost always under positive selection, owing to its myriad competitive and reproductive advantages - so much so that its average within lineages tends to increase with evolutionary time (a phenomenon known as "Copes Rule"; Kingsolver and Pfennig, 2004). By contrast, the sheer costs of a large size, including high resource demands and generally diminished rates of reproduction (Blueweiss et al, 1978), can lead small sizes to prevail in the game of fitness under certain conditions (Blankenhorn, 2000). But despite such evident links between size and fitness, our knowledge about how and why size varies at population-levels, and across space and near-time, is far from complete.

Since as early as the 20th century, the average body size of many bird and mammalian species has diminished (Searing et al, 2023; Youngflesh et al, 2022; McQueen et al, 2024), with mean declines ranging between ~1% per decade (Docampo et al, 2019; Youngflesh et al, 2022; Weeks et al, 2020), and up to ~12% per decade (Rioux Paquette et al, 2014) in punctuated periods. That these declines coincide with increasing global temperatures has raised reasonable speculation about a climate-based cause (i.e. by shifting size-specific thermoregulatory costs; Gardner et al, 2011; Searing et al, 2023), and indeed, taxonomic "dwarfing" events across earlier warming periods (e.g. the Palaeocene/Eocene Thermal Maximum, 55.8 mya) support this (Rankin et al, 2015). We now know, however, that the causative picture is rich (Nord et al, 2024) and other drivers, including resource declines (Eichenberg et al, 2021) – particularly for insectivorous species (Hallmann et al, 2017; Wagner, 2020; but see Müller

et al, 2024) – and changes in species assemblage composition (Smith et al, 2018; Martins et al, 2023) are also at play. Yet even with this developing, causative picture, exactly how size declines have been mediated (i.e. via natural selection, phenotypic plasticity, or both) is hazy. Trends at northern latitudes point to relaxed selection against small body size as winters warm (the “warming winter” hypothesis: Gardner et al, 2019; Ozgul et al, 2009; Bosco et al, 2023; but see Shipley et al, 2022), but direct changes in selection gradients, or genetic determinants of size, are scarce (Siepielski et al, 2019; Teplitsky et al, 2008; Prokosch et al, 2019). Moreover, stunting effects of poor resource accessibility (McNutt and Gusset, 2012; Rioux Paquette et al. 2014) and high temperatures during development (Shipley et al, 2022; Corregidor-Castro et al, 2023) implicate phenotypic plasticity, and not selection, in driving size shifts. While both processes are likely operative, opinions on the relative contributions of each are divided.

We argue that a major barrier to understanding how size declines have occurred has been a failure to examine them at a complete, distributional level. Longitudinal studies of mammalian and avian size have overwhelmingly focussed on changes to means (Searing et al, 2023; Youngflesh et al, 2022; McQueen et al, 2024), and less commonly, variance (Zheng et al, 2023). Although valuable starting points, trends in central tendency and dispersion generally lack the diacrisis needed for critical hypothesis testing. For example, both relaxed selection on small body size and plastic reductions of realised body size (whether from resource limitations or high developmental temperatures) equally result in declining means; so too should negative selection on large body size, whether from high thermoregulatory costs (Gardner et al, 2011) or high energetic costs of maintaining such a size (Fig. 1). Distinctions in variance alone are also unclear. When viewed holistically, however, each of these mechanisms predict unique and testable differences on species-specific, body size distributions (Fig. 1). Should changes in whole body size distributions be clarified, we argue that the mechanisms driving size declines may therefore be deduced.

Using a dataset monitoring morphology of over 150 avian species across their North American ranges, we quantified how body size distributions (measured as body mass distributions) have changed in recent decades (Fig. 2A). Changes in distributions – allowed to vary in skewness, variance, and higher moments – were then used to infer the most likely mediator of widely-observed morphological changes. Our final dataset included 758,326 body mass measurements from 159 avian species ($\bar{x}$ ± s.d. = 4,769 ± 7,498 per species), 26 families ($\bar{x}$ ± s.d. = 29,166 ± 49,603 per family), and 26 years (1992-2018), collected during the reproductive period in summer. Next, to more explicitly compare the

contributions of natural selection and phenotype plasticity toward contemporary size shifts, we generated species-specific, body mass selection gradients (linking size with survivorship probability) from North American ringing records and temporal symmetry (Pradel) models, then partitioned the effects of these selection gradients on distributional changes using a novel adaptation of the classic Price equation (Price, 1970). In its original form, the Price equation describes the degree to which a change in a species' mean phenotype is explained by linear effects of selection and so-called "transmission" effects (which combine plasticity and drift; Price, 1970). We adapted the Price equation into a 'distributional' analogue that partitions changes in a trait's full probability distribution function between linear or non-linear selection gradients and complementary transmission effects. When paired with distributional comparisons, this adapted equation offers a more robust and flexible opportunity to pinpoint the precise determinants of phenotypic changes through time, such as avian size.

Results and Discussion

Body size distributions are contracting and shifting.

Across all species and families assessed, body mass has predominantly declined (β = -0.012, or -1.2% per decade [posterior probability > 99.9%]), with average mass for 64.2% of species becoming smaller at their median latitudes. Percent changes in average mass ranged from -3.76% (the Black-throated Sparrow, *Amphispiza bilineata*) to +2.46% (the Carolina Wren, *Thryothorus ludovicianus*) per decade, aligning with previous estimates of size shifts in both North American and European songbirds (Youngflesh et al, 2022; Bosco et al, 2024). Coinciding with these declines, variation is body mass also decreased nearly ubiquitously (83.0% of species), averaging at a loss of 0.01 standard deviations (s.d. = 0.02) per decade (Figs. 2B-2C). This widespread and relatively rapid loss of size variance in North American birds strongly contrasts the increasing trend in global, terrestrial mammals (Zheng et al, 2023), and points to a disparity of mechanisms driving body size changes across animal groups. Whether this attrition of variance in birds is underpinned by a broader decline in genetic variation (Makvandi-Nejad et al, 2012; Carbeck et al, 2023) is not clear, but holds substantial implications for long-term, species adaptability (Lacy, 1997). Answering this question should therefore be an urgent priority.

Beyond shifts in central tenancy and dispersion, changes in the skewness of a phenotype can reveal crucial information about whether an adaptive optimum is moving, and how that movement is occurring (Jones et al. 2012; Urban et al. 2013; Débarre et al. 2015). We found that body size

distributions have also become increasingly centralised, with historic left-skewness contracting in the majority of species (69.8%; Fig. 2C) owing to disproportionate losses above the mean (-0.17% of total probability density /decade). When viewed alongside changing means and variance, this illuminates a broad-scale downgrading – or left-shifting – of body mass distributions in our analysed species, not from an increasing presence of small body sizes, but from a decreasing presence of large body sizes. Supporting this, maximum observed body size across species shrunk at a rate comparable to average body size (0.38%/decade; s.d. = 0.06%; β = -0.012 [>99.9%]; Fig. 3A), while minimum observed body size remained unchanged across our 26 year observation period (β = 0.01 [23.1%]; Fig. 3B). These findings demonstrate that relaxed mortality risk on small size – stemming from warming winters (Ozgul et al, 2009; Gardner et al, 2019; Bosco et al, 2023) – is insufficient to explain the size declines observed across extant avian species. If distributional shifts were indeed driven by relaxed selection against small individuals, we would instead expect maximum body size to remain stable and minimum body size to diminish, the precise opposite of our observations.

To understand the generality of shifts in size distributions, and better dissect their proximate mediators, we asked how they might have differed across the North American continent. We found that exactly how body size distribution have changed in recent years depended strongly on space and species-specific characteristics. Although shifts in average size and size variance did not differ across latitudes (Fig. 4; mean: β = -0.006 [46.4%]; variance: $\beta \approx 0$ [6.2%]), shifts in variance did vary by relative latitude within species' ranges (Fig. 4; mean: β = 0.008 [73.1%]; variance: β = 0.001 [82.5%]), indicating a possibly common, ultimate driver with intensity dependent on species' characteristics (i.e. thermal tolerance or ecological niche) and not location itself. At southern-most observation points, variance declined at rates over-doubling those at northern-most observation points, reaching average loss rates as high as 13.3% per decade (relative to 6.1%) while changes in skewness remained unchanged ($\beta \approx 0$ [5.6%]).

While it is possible that northward movement from southern populations (Brommer et al, 2012) has depleted some body size variance at southern range limits, such movement should also yield concurrent increases in variance at northern range limits (owing from an influx of comparatively small individuals from southern populations; Bergmann, 1847). A lack of this paired signal indicates that poleward movement – the driver of taxonomic dwarfing in early, prehistoric warming periods (the Palaeocene/Eocene Thermal Maximum; Rankin et al, 2019) – is not sufficient to explain our observed

shifts in size distributions through near-time. This is particularly significant, as it reveals that the mechanisms governing contemporary size changes in birds are not consistent with those governing size changes in earlier eras. Instead, the comparatively faster declines in variance at southern range limits raise increased prevalence of extreme heat *per se* – highest at these locations – as a probable, common driver of distributional shifts, as already evidenced in some mammalian (Mason et al, 2014; Stobo-Wilson et al, 2020) and avian (Shipley et al, 2022; Corregidor-Castro et al, 2023) populations. Still, given that changes in land-use tend to be most severe at lower latitudes (e.g. Popatov et al, 2022), we cannot rule out the possibility of more rapid habitat and resource losses at southern range limits speeding local, distributional contractions. However, the similarity of distributional changes between low and high latitude species (indicated by the lack of a latitudinal effect across species; mean: $\beta = -0.006$ [46.4%]; variance: $\beta \approx 0$ [6.2%]) indicates that this is likely insufficient to explain trends within species ranges.

Environmental effects are increasingly shaping body size distributions

Effects of body size on annual survivorship probability were predominantly quadratic across species and implicated stabilising selection as the dominant selective process shaping body size distributions through time ($mass^2$: $\beta = -0.186$ [97.4%]; mass: $\beta = 0.001$ [3.2%], $mass^3$: $\beta = -0.070$ [2.9%]; controlling for variations in capture effort). Our analyses show that the intensity of this stabilising selection – and thus, selection against body size extremes – has weakened ($mass^2 \times$ decade: $\beta = 0.208$ [85.8%]), although considerable variability exists across species (53.3% of species with non-zero interaction terms at ≥85% posterior probability). If changes in selection gradients are mediating our changes in body size distributions, this weakening selection against body size extremes implies that size variance should be increasing, not decreasing, as we observed (Figs. 2B-2C). This stark mismatch led us to evaluate the precise contributions of natural selection toward distributional shifts over time, using a novel Price equation (discussed above).

Our Price equation revealed that, since the beginning of our study period, the ability of selection gradients to predict changes in body size distributions has decreased ($\beta = -0.56$ [>99.9%]; Fig. 5A). Moreover, the strength of any detected selective effects, or their "penetrance," on distributional changes has significantly waned ($\beta = -0.80$ [>99.9%]; Fig. 5B). Phylogenetic effects of both changes were limited (Pagel's $\lambda = 0.03$ [s.d. = 0.02] and 0.03 [s.d. = 0.02] respectively). Modelling the magnitude of selective effects linearly through time showed that, by 1997 (one half-decade into our monitoring

period), non-selective effects – including phenotypic plasticity and genetic drift – had surpassed these selective effects as the largest determinant of body size distributions within species (Fig. 5B; averaged across species). Given many North American bird species have experienced large population declines in recent decades (Rosenberg et al, 2019), we wondered if this rising prominence of non-selective effects could merely be explained by strengthening impacts of genetic drift in recent years (an expected outcome of decreasing effective population sizes; Masel, 2011). To test this, we compared changes in selective magnitude between species with increasing and decreasing population sizes across our last half-decade of observation (2010-2015), when the prominence of non-selective effects was clear (Figs. 5A-5B) . In both species groupings, declines in selective magnitude were remarkably similar (Fig. 5C). Thus, rising non-selective effects are more likely explained by an increasing prominence of phenotypic plasticity, and not by increases in genetic drift.

Implications

Our results are consistent with temperature-mediated, phenotypic plasticity driving morphometric changes in North American birds. While similar conclusions have been reached for select species (Teplitsky et al, 2008; Shipley et al, 2022), our findings are the first to confirm a generalised mechanistic basis across a suite of avian species (>150), with explicit support from an expanded and trait-wide Price equation. By lacking evidence for a selective basis (Fig. 5), we strongly question whether size shifts seen in these species may be considered adaptive (as also doubted on physiological grounds; Nord et al, 2024; Tabh et al, 2025).

Exactly how plasticity is operating to alter size distributions in North American birds is unclear. While increasing developmental temperatures – particularly at southern range limits (DuRant et al, 2019; but see O'Connor et al, 2022) – may be directly truncating body sizes of young by maturity (reviewed in Weeks et al, 2022 and Tabh et al, 2023), this mechanism is unable to explain the limited reduction of body size probability below medians (Fig. 2B). Alternatively, rising ambient temperatures may limit resources available for growth, either directly by changing the distribution and abundance of prey species (Ozgul et al, 2009), or indirectly by limiting offspring provisioning or viable foraging time (Mason et al, 2014; Bourne et al, 2021). In this way, climate-mediated resource limitation may well explain our observed reduction in maximal body size among birds (Fig. 3A), as well as the slight loss of body size probability below medians if shortened foraging times motivate a need for increased, internal resource storage (i.e. as fat; Andrews et al, 2021). Regardless, the reduction in body size and its

variance reported here are likely to have meaningful consequences on environmental resilience and future persistence (Kingsolver and Pfennig, 2004). Understanding whether plastic shifts in size distributions are offering hitherto unknown functional benefits will be a crucial next step toward predicting such consequences.

Figures

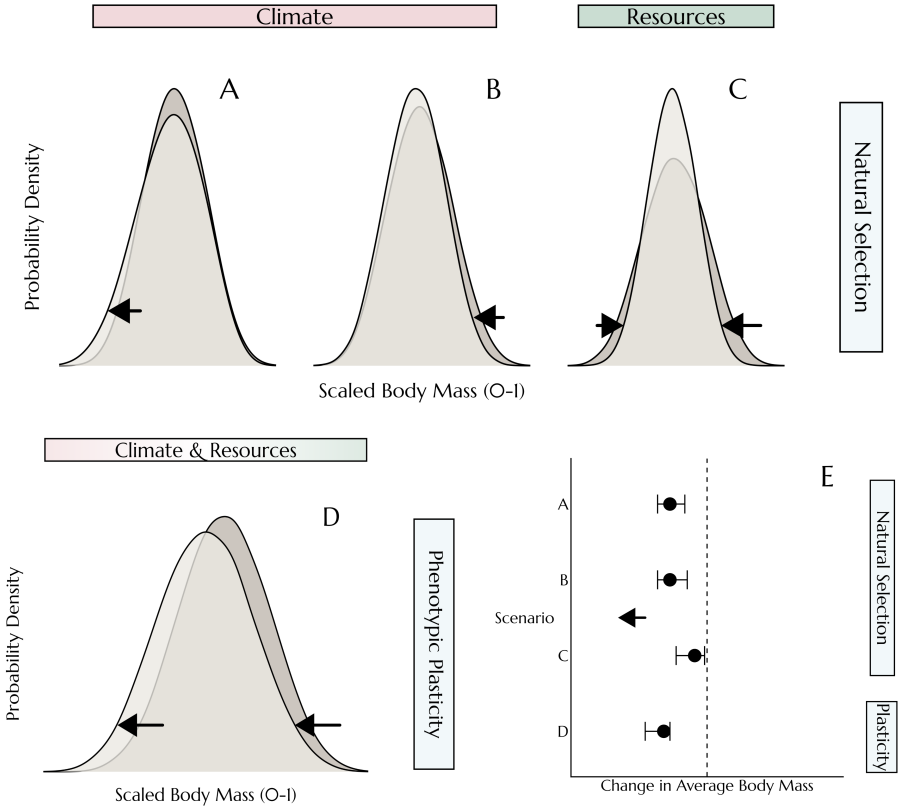

Figure 1. Outcomes of mechanistic hypotheses explaining changes in body size in response to climate warming ("Climate") and resource limitation ("Resources"). Numerous mechanistic hypotheses generate equivalent predictions about how body size averages should change. However, we show that predictions regarding changes in body size probability distributions vary. Panels **A** and **B** display predictions for two common hypotheses regarding how climate warming may diminish body size averages: (**A**) by relaxing mortality risk on small sizes (Ozgul et al, 2009), or (**B**) by increasing selection against large body size through high thermoregulatory costs (Gardner et al, 2011). Panel **C** displays predicted outcomes of limited resource availability on size distributions, specifically by imposing costs on both large sizes owing to high resource demands, and small size by limiting competitive ability simultaneously, but with costs being hypothetically largest for larger sizes. Last, panel **D** displays hypothetical outcomes for combined climate warming and resource limitation, acting through phenotypic plasticity. Here, climate warming truncates size at maturity through direct effects of higher developmental temperatures (Weeks et al, 2022; Tabh et al, 2023), while reduced resources shift maximal attainable size downward. Differential outcomes of each hypothesis highlight the diacratic value of testing predictions at the distributional level. Panel **E** shows the similar change in averages

from each hypothesis (or “scenario”) described in the previous panels.

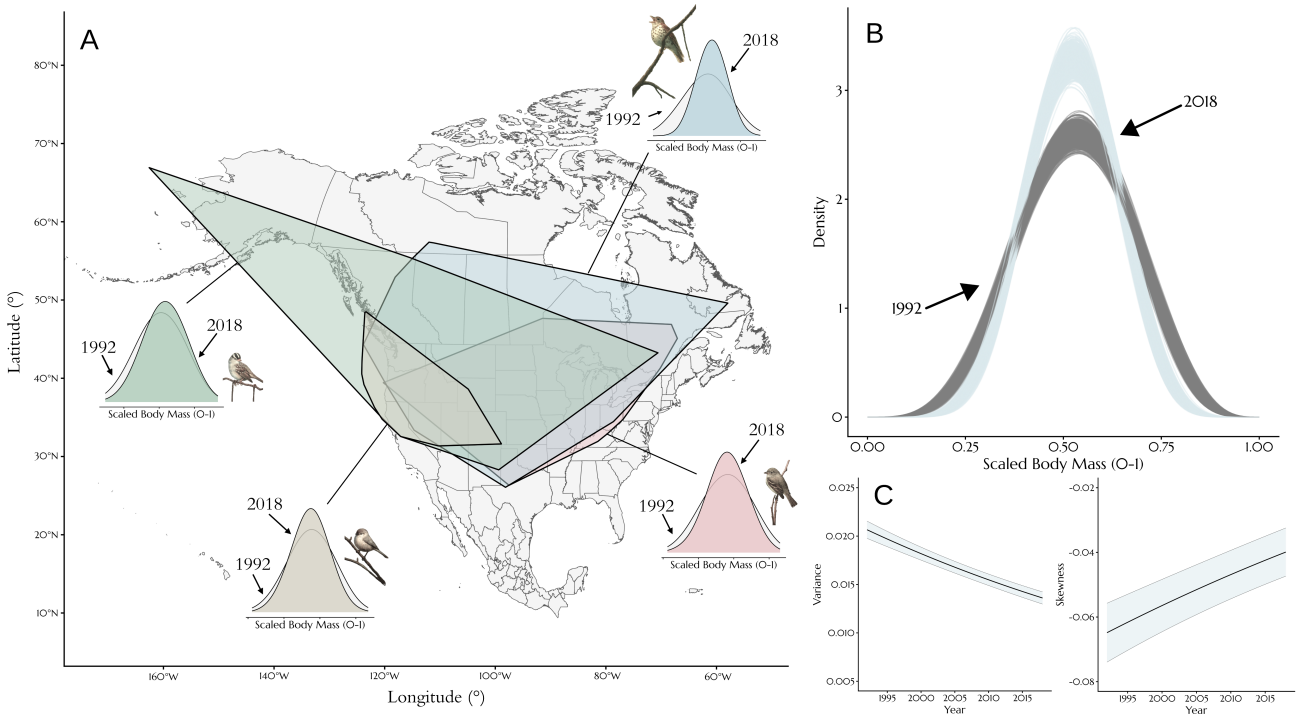

Figure 2. Declines in average size and size variance, and increases in size skewness, across 159 North American bird species. Body size measurements (body mass) were collected across naturally-occurring species ranges and their distributions modelled across latitudes within and across ranges. A sample of species and distributional shifts at range medians is shown in panel **A** (complete set of distributional shifts in SFig. 1). **B.** Estimated changes in average body size distributions between 1992 (grey) and 2018 (light blue) at median latitudes per species range, for all species in the data set. Lines indicate outcomes from distinct posterior samples drawn from a Bayesian mixed effects model. **C**. Changes in variance and skewness of body size distributions respectively, as estimates from outcomes of a Bayesian mixed effects model. Lines represent predicted trends, and bands (blue) represent standard deviations around predicted trends. Body size measurements are scaled between 0 and 1 per species and sex (where applicable).

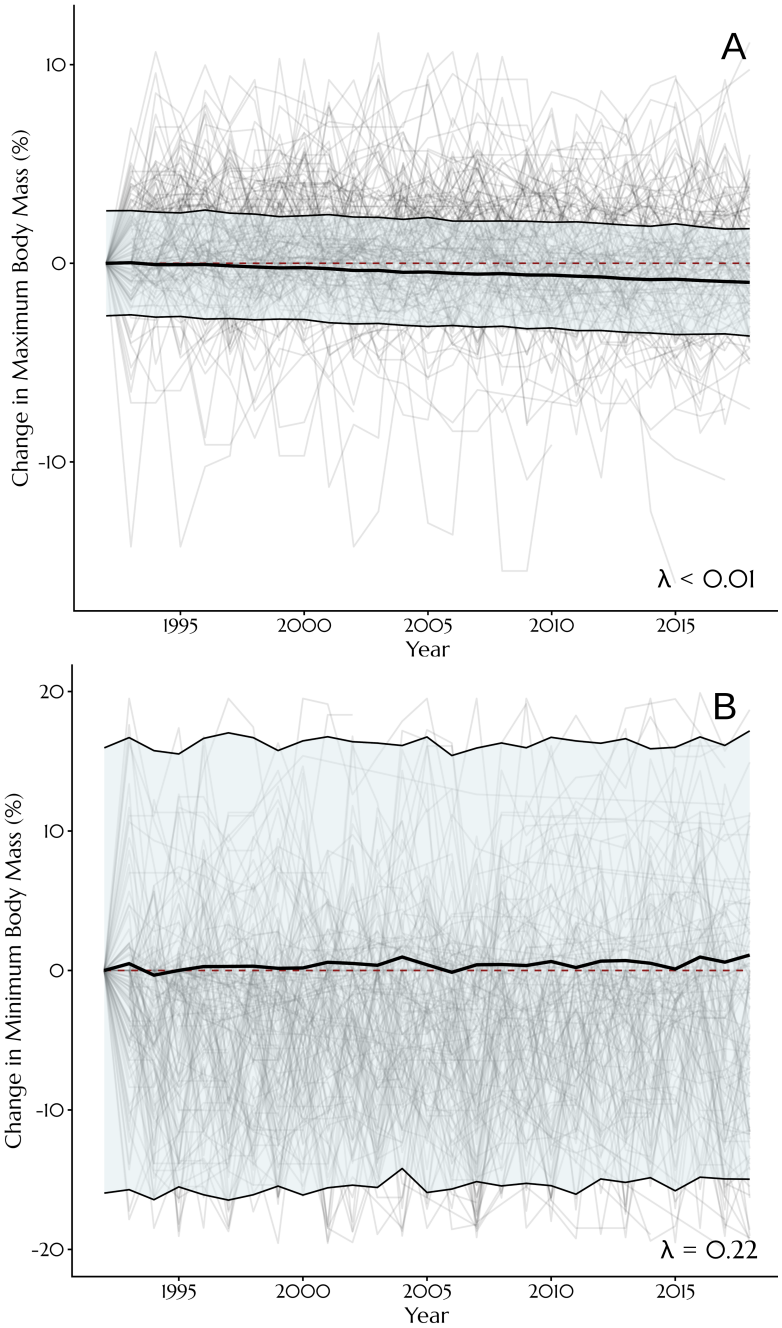

Figure 3. Declines in maximal body size, and stability of minimum body size across 159 North American bird species. Maximum (**A**) and minimum (**B**) body sizes represent maximum and minimum observed body mass, per species, and were natural-log transformed before use in Bayesian mixed effects model. Relative changes from baseline (1992) are displayed for adequate cross-species comparison. Grey lines represent single-species trends, while black lines indicate model fits. Bands (blue) display ± 1 standard deviation around trend lines. Red dashed lines display the 0-change line. In

all panels, λ represents Pagel's lambda.

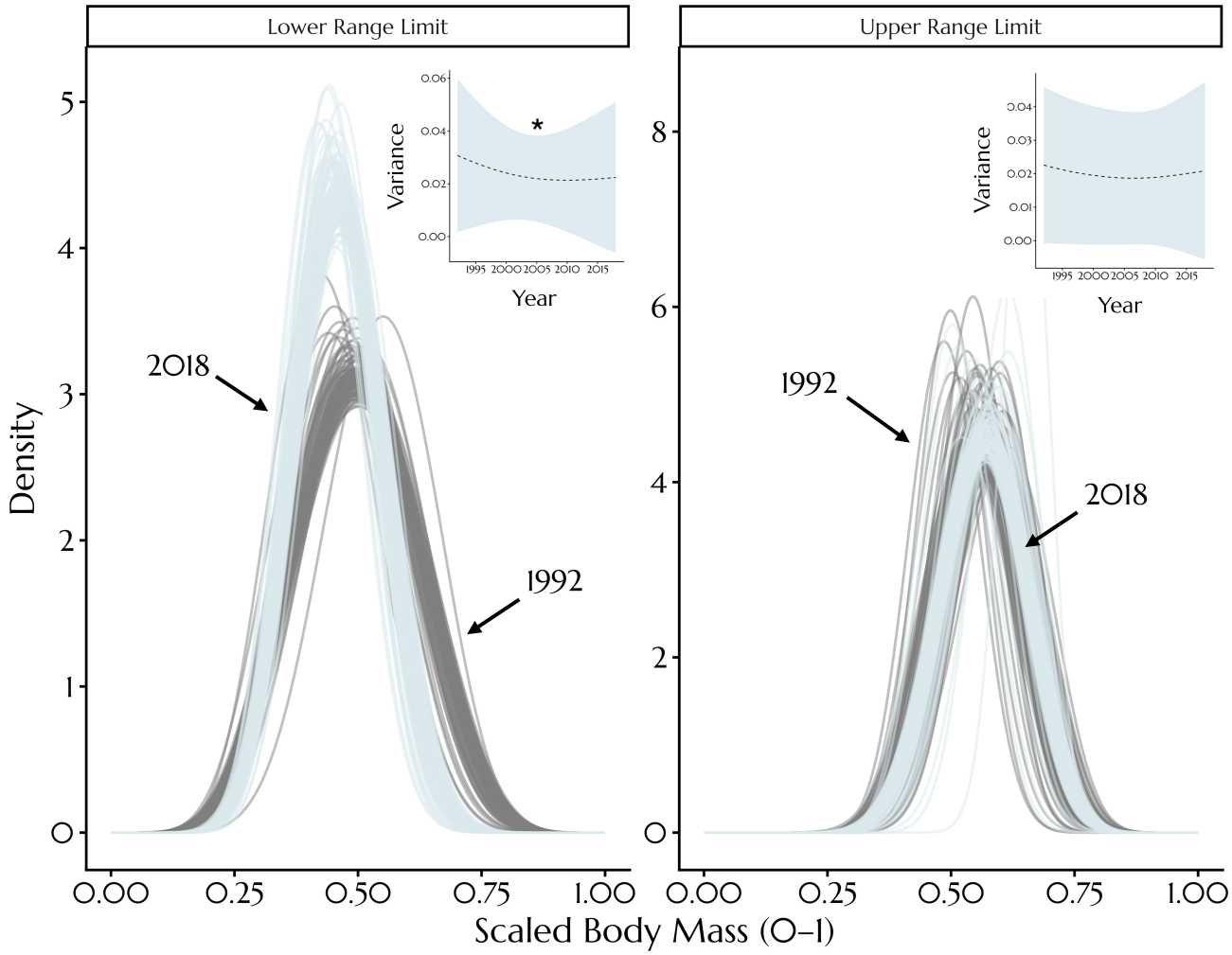

Figure 4. Declines in body size variance differ across species ranges. Temporal declines in body mass variance correlate with latitude within, but not across, species (β = 0.001 [82.5%]), with southern range limits experiencing the greatest losses of variance. Declines in average body mass are also greater at southern range limits, but with greater uncertainty (latitudinal effect: β = 0.008 [73.1%]). Lines in the large panels represent predicted body mass distributions across species, at southern-most and northern-most observation points (for each species individually). Blue lines indicate estimates for 2018 and grey lines indicate estimates for 1992; each individual line represents a posterior draw from a Bayesian beta-distributional model. Average changes in variance across time and species for each region (range limit) are shown in small insets. The asterisk indicates a trend with over 80% certainty.

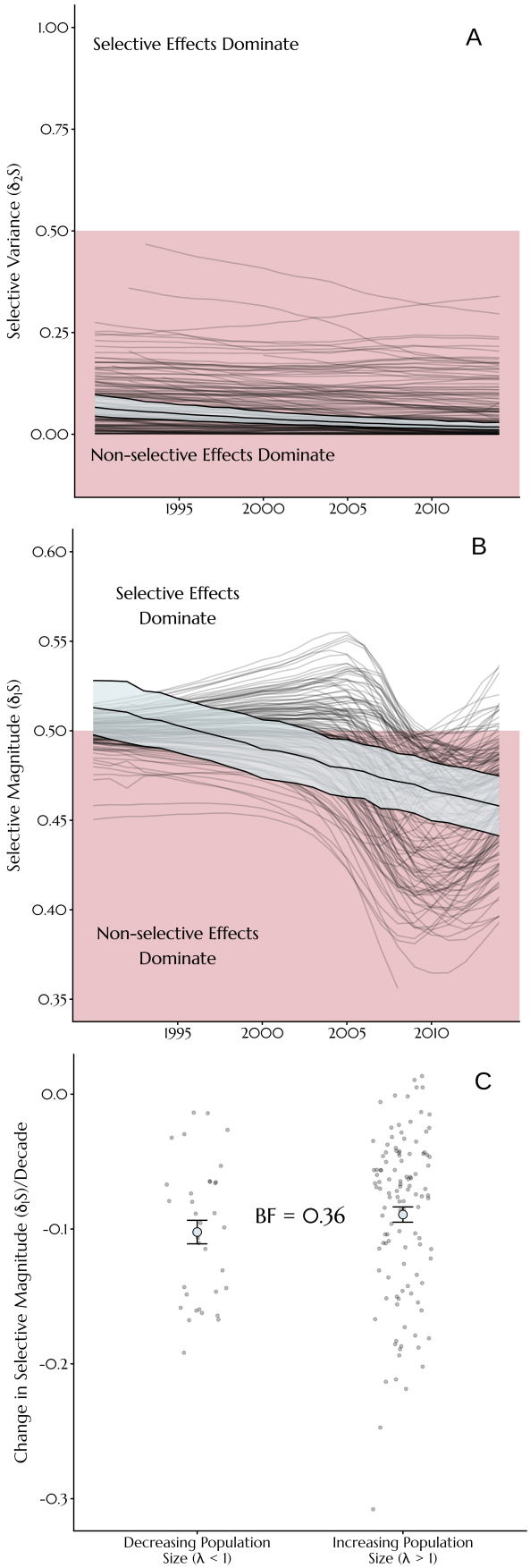

Figure 5. Phenotypic plasticity is increasingly defining body size distributions in North American birds (151 species). Selective variance (δ_2S; **A**) – derived from a distributional variant of the Price equation (Price, 1970) – describes the ability of a species' body size selection gradient to explain changes in its full, body size distributions (similar to a Pearson R^2). Values range between 0 and 1, with values above 0.5 indicating that natural selection is better able to explain distributional shifts than non-selective effects. Selective magnitude (δ_1S; **B**) describes the total extent to which detected, selective effects contribute to a change in a body size distributions. Here, values above 0.5 indicate that these selective effects explain more than half of total distributional movement. Grey lines in **A** and **B** display trends for individual species. Black lines and blue bands display linear trend lines ± 1 standard deviation respectively, derived from Bayesian mixed effects models. **C**. Effect of population change (growth or decline) on linear changes in selective magnitude through time (1992-2015). Grey dots display values per species, blue dots display means, and errorbars display ± 1 standard error above and below means. Selection gradients and estimates of population change are derived from a temporal symmetry ("Pradel") capture recapture model.

References

Kingsolver, J.G., Pfennig, D.W. Individual-level selection as a cause of Cope's rule of phyletic size increase. *Evol.* 58(7), 1608-1612. (2004).

Blueweiss, L., Fox, H., Kudzma, V., Nakashima, D., Peters, R., Sams, S. Relationships between body size and some life history parameters. *Oecol.* 37(2), 257-272. (1978).

Blanckenhorn, W.U. The evolution of body size: what keeps organisms small? *Quart. Rev. Biol.* 75(4), 385-407. (2000).

Searing, K.B., Lomolino, M.V., Rozzi, R. Melting climates shrink North American small mammals. *Proc. Natl. Acad. Sci.* 120(50), e2310855120. (2023).

Youngflesh, C., Saracco, J.F., Siegel, R.B., Tingley, M.W. Abiotic conditions shape spatial and temporal morphological variation in North American birds. *Nat. Ecol. Ecol.* 6(12), 1860-1870. (2022).

Weeks, B.C. et. al. Shared morphological consequences of global warming in North American migratory birds. *Ecol. Lett.* 23(2), 316-325. (2020).

Docampo, M. Moreno, S., Santoro S. Marked reduction in body size of a wood mouse population in less than 30 years. *Mamm. Biol.* 95(1), 127-134 (2019).

McQueen, A. et. al. Shorebirds are shrinking and shape-shifting: Declining body size and lengthening bills in the past half-century. *Ecol. Lett.* 27(12), e14513. (2024).

Gardner, J. L., Peters, A., Kearney, M.R., Joseph, L., Heinsohn, R. Declining body size: a third universal response to warming? *TREE* 26(6), 285-291. (2011).

Rankin, B.D. et al. The extended Price equation quantifies species selection on mammalian body size across the Palaeocene/Eocene Thermal Maximum. *Proc. Roy. Soc. B. Biol. Sci.* 282(1812), 20151097. (2015).

Eichenberg, D. et. al. Widespread decline in Central European plant diversity across six decades. *Glob. Change Biol.* 27(5), 1097-1110. (2021).

Hallmann, C A. et. al. More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PloS One* 12(10), e0185809. (2017).

Wagner, D.L. Insect declines in the Anthropocene. *Ann. Rev. Entomol.* 65(1), 457-480. (2020).

Müller, J. et. al. Weather explains the decline and rise of insect biomass over 34 years. *Nature* 628(8007), 349-354. (2024).

Ozgul, A., Tuljapurkar, S., Benton, T.G., Pemberton, J.M., Clutton-Brock, T.H., Coulson, T. The dynamics of phenotypic change and the shrinking sheep of St. Kilda. *Science* 325(5939), 464-467.

(2009).

Gardner, J.L. et. al. Australian songbird body size tracks climate variation: 82 species over 50 years. *Proc. Roy. Soc. B.* 286(1916), 20192258. (2019).

Shipley, J.R., Twining, C.W., Taff, C.C., Vitousek, M.N., Winkler, D.W. Selection counteracts developmental plasticity in body-size responses to climate change. *Nat. Clim. Change* 12(9), 863-868. (2022).

Bosco, L., Otterbeck, A., Fransson, T., Lindén, A., Piha, M., Lehikoinen, A. Increasing winter temperatures explain body size decrease in wintering bird populations of Northern Europe—But response patterns vary along the spatioclimatic gradient. *Glob. Ecol. Biogeog.* 32(12), 2100-2110. (2023).

Nord, A., Persson, E., Tabh, J.K., Thoral, E. Shrinking body size may not provide meaningful thermoregulatory benefits in a warmer world. *Nat. Ecol. Evol.* 8(3), 387-389. (2024).

Smith, F.A., Elliott Smith, R.E., Lyons, S.K., Payne, J.L. Body size downgrading of mammals over the late Quaternary. *Science* 360(6386), 310-313. (2018).

Martins, I.S. et. al. Widespread shifts in body size within populations and assemblages. *Science* 381(6662), 1067-1071. (2023).

Siepielski, A.M. et. al. No evidence that warmer temperatures are associated with selection for smaller body sizes. *Proc. Roy. Soc. B. Biol. Sci.* 286(1907), 20191332. (2019).

Teplitsky, C., Mills, J.A., Alho, J.S., Yarrall, J.W., Merilä, J. Bergmann's rule and climate change revisited: Disentangling environmental and genetic responses in a wild bird population. *Proc. Natl. Acad. Sci.* 105(36), 13492-13496. (2008).

Prokosch, J., Bernitz, Z., Bernitz, H., Erni, B., Altwegg, R. Are animals shrinking due to climate change? Temperature-mediated selection on body mass in mountain wagtails. *Oecol.* 189(3), 841-849. (2019).

Corregidor-Castro, A. et. al. Experimental nest cooling reveals dramatic effects of heatwaves on reproduction in a Mediterranean bird of prey. *Glob. Change Biol.* 29(19), 5552-5567. (2023).

McNutt, J. W., Gusset, M. Declining body size in an endangered large mammal. *Biol. J. Linnean Soc.* 105(1), 8-12. (2012).

Rioux Paquette, S., Pelletier, F., Garant, D., Bélisle, M. Severe recent decrease of adult body mass in a declining insectivorous bird population. *Proc. Roy. Soc. B. Biol. Sci.* 281(1786), 20140649. (2014).

Youngflesh, C., Saracco, J.F., Siegel, R.B., Tingley, M.W. Reply to: Shrinking body size may not provide meaningful thermoregulatory benefits in a warmer world. *Nat. Ecol. Ecol.* 8(3), 390-391. (2024).

Zheng, S., Hu, J., Ma, Z., Lindenmayer, D., Liu, J. Increases in intraspecific body size variation are

common among North American mammals and birds between 1880 and 2020. *Nat. Ecol. Evol.* 7(3), 347-354. (2023).

Price, G. Covariance and selection. *Nature* 227(1), 520-21. (1970).

Makvandi-Nejad, S. at al. Four loci explain 83% of size variation in the horse. *PloS One*, *7*(7), e39929. (2012).

Carbeck, K et. al. Candidate genes under selection in song sparrows co-vary with climate and body mass in support of Bergmann's Rule. *Nat. Comm.* 14(1), 6974. (2023).

Lacy, R.C. Importance of genetic variation to the viability of mammalian populations. *J. Mammol.* 78(2), 320-335. (1997).

Jones, A.G., Bürger, R., Arnold, S.J., Hohenlohe, P.A., Uyeda, J.C. The effects of stochastic and episodic movement of the optimum on the evolution of the G-matrix and the response of the trait mean to selection. *J. Evol. Biol*. 25(11): 2210–2231. (2012).

Urban, M.C., Bürger, R., Bolnick, D.I. Asymmetric selection and the evolution of extraordinary defences. *Nat. Comm.* 4(1): 2085. (2013).

Débarre, F., Yeaman, S., Guillaume, F. Evolution of quantitative traits under a migration-selection balance: when does skew matter? *Am. Nat.* 186(S1): S37-S47. (2015).

DuRant, S.E., Willson, J.D., Carroll, R.B. Parental effects and climate change: will avian incubation behavior shield embryos from increasing environmental temperatures? Integr. *Comp. Biol.* 59(4), 1068-1080. (2019).

O'Connor, R.S., et. al. Warming in the land of the midnight sun: breeding birds may suffer greater heat stress at high-versus low-Arctic sites. *Proc. Roy. Soc. B.* 289(1981), 20220300. (2022).

Mason, T.H., Apollonio, M., Chirichella, R., Willis, S.G., Stephens, P.A. Environmental change and long-term body mass declines in an alpine mammal. *Front. Zool.* 11(1), 69. (2014).

Bourne, A.R., Ridley, A.R., Spottiswoode, C.N., Cunningham, S.J. Direct and indirect effects of high temperatures on fledging in a cooperatively breeding bird. Bahv. *Ecol*. 32(6), 1212-1223. (2021).

Andrews, C., Zuidersma, E., Verhulst, S., Nettle, D., Bateson, M. Exposure to food insecurity increases energy storage and reduces somatic maintenance in European starlings (*Sturnus vulgaris*). *Roy. Soc. Open Sci.* 8(9), 211099. (2021).

Stobo-Wilson, A.M., Cremona, T., Murphy, B.P., Carthew, S.M. Geographic variation in body size of five Australian marsupials supports Bergmann's thermoregulation hypothesis. *J. Mammol.* 101(4), 1010-1020. (2020).

Potapov, P., et. al. The global 2000-2020 land cover and land use change dataset derived from the Landsat archive: first results. *Frontiers in Remote Sensing*, *3*, 856903. (2022).

Brommer, J.E., Lehikoinen, A., Valkama, J. The breeding ranges of central European and arctic bird species move poleward. *PLoS One,* 7(9): e43648. (2020).

Rosenberg, K.V., et al. Decline of the North American avifauna. *Science,* 366(6461): 120-124. (2019).

Masel, J. Genetic drift. *Curr. Biol.,* 21(20): R837-R838. (2011).

Tabh, J.K.R, et. al. Limited evidence that body size shrinking and shape-shifting alleviate thermoregulatory pressures in a warmer world. *Comm. Biol.* 8(1): 707. (2025).

Weeks, B.C. et al. Temperature, size and developmental plasticity in birds. *Biol. Lett.* 18(12): 20220357. (2022).

Tabh, J.K., Nord, A. Temperature-dependent developmental plasticity and its effects on Allen's and Bergmann's rules in endotherms. *Integr. Comp. Biol.* 63(3): 758-771. (2023).

Materials and Methods

1.0 Body mass data

Body mass data for North American birds were collected from the Monitoring Avian Productivity and Survivorship (MAPS) program (DeSante et al, 2004), as organised and collated by the Institute of Bird Populations. The MAPS program is a collaborative program that combines avian morphometric data from over 400 species captured across 1265 sampling locations, between the years of 1992 and 2018. Sampling locations in these years range from a minimum latitude of 26.065° to a maximum latitude of 69.382°, and are widespread from the west to east of the North American continent (-165.669° to -57.606°; see Youngflesh et al, 2022). For our study, species inclusion depended solely on the sufficiency of measurements to estimate time-specific body mass distributions. As such, data from all monitored species were included a priori unless removed owing to insufficiency (see paragraph below).

To remove bias from species misidentifications or body mass measurement errors, all body mass values exceeding a species' median ± 4 median absolute deviations were removed (similar to Youngflesh et al, 2022). Next, to avoid the influence of migration behaviour on body mass measurements (and thus, the generality of our results), we conservatively filtered data to include only measurements obtained between the months of June and August, where migration was unlikely to occur. Last, to ensure that samples sizes were sufficient to estimate yearly deviations from intercept distributions, we removed (1) all measurement years for a species where < 20 individuals were captured, and (2) all species where < 4 measurement years at the previous criteria. In total, our analysis included 758,326 body mass measurements from across 159 species ($\bar{x}$ ± s.d. = 4,769 ± 7,498 per species), 26 families ($\bar{x}$ ± s.d. = 29,166 ± 49,603 per family), and 26 years (1992-2018).

1.1 Modelling shifts in species body mass across time

Phenotype distributions in nature can vary in skewness, modality (i.e. unimodal vs. bimodal) and higher moments, as well as central tendency and dispersion (Kapteyn, 1916; Pick et al, 2022; Pollo et al, 2025). The beta probability density function ("pdf") is flexible pdf able to capture variation in each for these metrics by adjusting the values of two shape parameters, α and β (or alternatively: μ, the distributional mean, and Φ, a scaling parameter). To model changes in body mass distributions, we therefore generalised mass distributions per species using a beta pdf, with body mass first scaled between > 0 and < 1 (specifically, between $1.0\times10^{-5} – 1 - 1.0\times10^{-5}$) per species. To adjust for possible sexual size dimorphism within species, scaling was conducted within sex whenever sex was reliably

determinable at capture (>75% of captures).

To estimate temporal changes in species' body mass distributions, we used a Bayesian mixed effects model with scaled body mass as the beta-distributed response variable, and with μ (on the logit scale) and Φ (on the natural log scale) varying by model predictors. Here, sampling year was included as a continuous, population-level predictor (or "fixed effect"), and was median-centred and scaled by a factor of ten to ease model fitting. Given that body mass of many species is known to vary by latitude (Ashton, 2002; a phenomenon known as "Bergmann's Rule"; Bergmann, 1847), both a species' median latitude, and an individual's relative latitude for its species (here, in ° from the species' median) were included as additional, population-level predictors, with each parameters interacting with year to account for recent changes in species' ranges (e.g. Hitch and Leberg, 2007). Last, to estimate distributional changes species-specifically, species identity was included as a group-level predictors of μ and Φ, with effects of year and relative latitude on μ and Φ also allowed to vary by species. The role of phylogenetic covariance on shaping distributional changes was evaluated in subsequent analyses (see section 4.2).

Priors for our distributional model were informed by the findings of others (e.g. Ashton, 2002; Meiri and Dayan, 2003; Van Buskirk et al, 2010; Weeks et al, 2020; Youngflesh et al, 2022) and by preliminary data visualisations. For population-level intercepts on μ and Φ, we used normally-distributed priors with means of 0.15 and 2.7 respectively, and standard deviations of 0.25, representing expected parameter estimates from cross-species histograms. For the effects of latitude and relative latitude within species, priors were also normally distributed and assumed increases in distributional means with each variable (μ: latitude ~ $\mathcal{N}$[mean = 0.05, s.d. = 0.1], relative latitude within species ~ $\mathcal{N}$[mean = 0.05, s.d. = 0.1]), but no concurrent change in variance across variables (Φ: latitude across and within species ~ $\mathcal{N}$[mean = 0, s.d. = 0.05]). For our interactions between year and either latitude or relative latitude within species, all priors were conservative and assumed no effect on distributional means or variance (μ: $\mathcal{N}$[mean = 0, s.d. = 0.2] for both terms; Φ: $\mathcal{N}$[mean = 0, s.d. = 0.1] for both terms). At the species-level, priors were exponentially-distributed and as follows: μ and Φ: variation in intercepts ~ exponential(λ = 5), variation in slopes by year ~ exponential(λ = 10), and variation in slopes by relative latitude ~ exponential(λ = 10).

Our complete model was constructed in Stan (version 2.32.2) and executed in R (version 4.4.1; R Core Team, 2025) using the R package "cmdstanR" (Gabry et al, 2025). Posteriors were estimated using

Hamiltonian Monte Carlo (HMC) sampling, with 20,000 sample iterations – thinned by a factor of fourty – drawn across 4 chains. The first 10,000 iterations from each chain were discarded as warm-up. To ease model initiation, a preliminary model run was conducted using meanfield automatic differentiation variational inference (ADVI; Kucukelbir et al, 2015), and approximate posterior medians used to inform initial parameter values of our HMC chains. For this initial run, we used an η value (a step-size weighting variable) of 1.0, and 100,000 sample iterations. Chain mixing of our final model was evaluated by calculating Gelman-Rubin statistics ($\hat{R}$; Gelman and Rubin, 1992) contrasting them to 1.0, while autocorrelation within chains was evaluated by inspecting effective sample size to sample size ratios (N_{eff}/N) per parameter. All $\hat{R}$ values fell between 1.00 – 1.02 and mean N_{eff}/N values equaled 0.93 (s.d. = 0.12), indicating clear chain mixing and limited within-chain autocorrelation. Posterior predictive fits were diagnosed visually.

1.2 Deriving variance and skewness from species' mass distributions

For each species and observation year, we calculated the variance (σ^2) and skewness (γ_1) of body mass distributions, as derived from our distributional model (described above). For comparative purposes, each characteristic was also calculated at the minimum, median, and maximum observed latitude for each species. To do so, posterior predictions of μ and Φ for each species and context (n = 1000 each) were drawn from our distributional model, converted to standard beta-distribution shape parameters α and β ($\alpha = \mu \times \Phi$; $\beta = [1- \mu] \times \Phi$), and loaded into standard forms equations as follows:

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2 \cdot (\alpha+\beta+1)}, \tag{1}$$

and

$$\gamma_1 = \frac{2 \cdot (\beta-\alpha)\sqrt{(\alpha+\beta+1)}}{(\alpha+\beta+2) \cdot \sqrt{(\alpha\beta)}}. \tag{2}$$

Resultant σ^2 and γ_1 values were then averaged per predictive scenario, and used to visualise and numerically quantify changes in each across time.

1.3 Modelling changes in minimum and maximum body size through time

To further our understanding of how body size distributions have, or have not, shifted through time, we also tested whether species' observed minimum and maximum size measurements (here, natural log-

transformed body mass) have changed across the same observation period (1992-2018). To do so, we constructed two Bayesian linear mixed effects models, one with each species' observed minimum body size per year as the response variable, and the other with each species' observed maximum body size per year as the response variable. In both models, error was treated as Gaussian, and we included year as the sole population-level predictor (median-centred and scaled by 10 as above), and both species and year-by-species as group-level predictors (intercept and slope respectively). Further, to correct for phylogenetic non-independence between species, we also allowed our model intercept and effect of year to vary by the phylogenetic covariance between species. Here, phylogenetic covariation was calculated from a consensus tree constructed using a recent genetically-derived phylogeny of birds (McTavish et al, 2024) via Beast2's TreeAnnotator (Bouckaert et al, 2014; burn-in = 25% and node heights representing medians). A full description of covariance calculation is provided elsewhere (Tabh et al, 2023).

Priors for both models described above were liberal and weakly informative. Specifically, for our model predicting minimum observed size, priors for our population-level intercept and effect of year were normal, with means of 2.35 and 0 respectively, and standard deviations of 1 and 0.5 respectively. For our model predicting maximum observed size, priors for these parameters were largely identical, however, the mean of our intercept prior was set to 3. For both models, priors for all group-level effects were exponential with lambda values of 0.5. Sampling for each model was conducted across 4 HMC chains, totalling 50,000 iterations per chain, thinned by a factor of 20 and with the first 10,000 iterations discarded as warmup. $\hat{R}$ values averaged at 1 (s.d. = 0.001), while N_{eff}/N values averaged 0.660 (s.d. = 0.305) across both models.

2.0 Estimating annual survivorship at the species- and individual-level

2.1 Capture-recapture model overview

We next sought to estimate how body mass relates to fitness across species (i.e. via body mass selection gradients), and whether this relationship – if any – has changed through time. This estimation is a necessary step toward testing whether changes in body mass selection gradients alone are sufficient to explain recent shifts in body mass distributions (see section 3.0). Hypotheses using selection to explain morphometric shifts in birds and mammals largely focus on mass-specific changes to survivorship and not reproduction (see Gardner et al, 2011). As such, we chose to use annual survivorship probability as our metric of fitness in our selection gradients.

To estimate annual survivorship probability (ϕ) at the species level, we used a hierarchical and Bayesian temporal symmetry model (or "Pradel model"; Pradel, 1996), adjusted to account for the presence of transients who are unlikely to be detected within a sampling area twice (adapted from Telenský et al, 2024). Briefly, this model uses both forward and reverse time to simultaneously estimate ϕ, and an annual "seniority" probability (γ; i.e. that an individual captured at time t was present in the capture area at time t-1) from capture-recapture records, with each parameter being contingent on a probability of detection (p; reviewed in Sandercock et al, 2020), and a probability of residency within the sampling area (r; or alternatively). In this model, rates of population change (λ) directly represent the ratio of ϕ to γ (Pradel, 1996). This relationship was particularly useful for our study, because prior estimates of λ are often available, and reparameterising the model to directly sample λ instead of γ allowed us to improve the estimational precision of ϕ.

2.2 Generalising the Pradel model for variations in capture effort

One known shortcoming of Pradel models is the assumption that all individuals have equal probability of capture, whether previously captured or not. This implies that capture effort is constant across both space and time in a sampled population – an assumption that is difficult to satisfy when records are collated across independently organised observation sites (e.g. the MAPS data set). To extend the utility of the Pradel model, we therefore explicitly estimated mean annual capture effort (χ) per species and year, then accounted for effects of that effort on p (see section 2.3). Because effort required for detection may differ between naive and previously captured individuals (Riddle et al, 2010), we estimated χ as latent determinant of two partial effort metrics: (1) the total number of observation days in a given year, per species, weighted by the probability of detection at each observation site (χ_C; a metric of first capture effort), and (2) the average number of possible observation days before detection, per individual of a species, weighted by the probability of capture across active sampling sites per year (χ_R; a metric of recapture effort). More specifically, we assumed:

$$\chi_{C_{sp,t}} = \sum_{cs=1}^{n_{cs}} N_{days_{cs,t}} \cdot \omega_{cs \vee sp}, \tag{3}$$

$$\chi_{R_{sp,t}} = \frac{\sum_{i=1}^{n_{sp,t}} \sum_{cs=1}^{n_{cs}} N_{days_{cs,t}} \cdot \omega_{cs \vee sp,i}}{n_{sp,t}}, \tag{4}$$

and

$$\begin{aligned} \ln\left(\chi_C+1\right) &\sim N\left(\beta_C\,\chi,\sigma_C\right) \\ \ln\left(\chi_R+1\right) &\sim N\left(\beta_R\,\chi,\sigma_R\right) \end{aligned}, \tag{5}$$

where *sp* represents species, *i* represents individual identity, *t* represents year, *cs* represents a capture site, $n_{sp,t}$ represents the number of individuals for a given species captured in year *t*, n_{cs} represents the number of capture stations used in our study, $N_{days_{cs,t}}$ represents the number of active days at a given capture site and year (defined as the number of days where at least one individual of any species was observed) and $\omega_{cs|sp,i}$ or $\omega_{cs|sp}$ indicates the probability of capture at capture site *cs* given a combination of species *sp* and individual *i*.

To first derive $\omega_{cs,sp,}$ we divided the number of species observations at a capture site *cs* by the total number of observations across sites for species *sp*. To next derive $\omega_{cs|sp,i}$, we integrated Gaussian kernel densities drawn across all known capture-recapture distances (in km) per species, with lower limits of 0 and upper limits of the straight-line distances between an individual's initial capture site and possible recapture sites. Kernel bandwidths were set to 0.5 km, and as such, we assumed:

$$\omega_{cs_{i,sp}} = \int_0^{d_i} \left[\frac{1}{n_{d,sp}} \sum_{j=1}^{n_{d,sp}} \frac{1}{0.5 \cdot \sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{d_{sp}-d_{j,sp}}{0.5}\right)^2} \right] dd_{sp} \tag{6}$$

where d_i indicates the Haversine distance (in km) between the previous capture site of individual *i* and capture site *cs*, and $n_{d,sp}$ indicates the number of known capture-recapture distances for species *sp*. For the calculation of both χ_C and χ_R, possible capture days represented only those during which a given species was expected to be present at a given site, in a given year. Relevant ranges were set by first modelling the very first known capture dates (within a year) per species as a function of Bayesian, Gaussian processes with latitude and longitude as predictors (n basis functions = 5; kernel = $\sigma^2 \cdot \exp\left(\frac{-(x-x')^2}{2l^2}\right)$, priors = flat; ε = Gaussian). Observation days across sites that preceded the expected first capture date from these models (the posterior medians), or preceded the observed first capture date in that year for a species (whichever was earliest) were then excluded from our estimates of χ_C and χ_R. Last possible observation days used in χ_C and χ_R calculation were set to the final observation date for a given species and site, across years.

Once χ_C and χ_R were calculated, latent and "true" effort values (χ) per species *sp* and year *t* were first approximated using a closed, single-iteration Newton-Raphson as follows:

$$\nabla_1(\chi_{sp,t}) = \frac{-\beta_C n_{sp,t} \cdot (\chi_{C_{sp,t}} - \beta_C \chi_{sp,t})}{\widetilde{\sigma_C^2}} - \frac{\beta_R n_{sp,t} \cdot (\chi_{R_{sp,t}} - \beta_R \chi_{sp,t})}{\widetilde{\sigma_R^2}} + \chi_{sp,t}$$

$$\frac{d}{d\chi_{sp,t}} \nabla_1(\chi_{sp,t}) = \frac{\beta_C^2 n_{sp,t}}{\widetilde{\sigma_C^2}} + \frac{\beta_R^2 n_{sp,t}}{\widetilde{\sigma_R^2}} + 1 \qquad (7)$$

$$\chi_{sp,t}^{(i+1)} = \chi_{sp,t}^{(i)} - \frac{\nabla_1(\chi_{sp,t})\chi_{sp,t}^{(i)}}{\frac{d}{d\chi}\nabla_1(\chi_{sp,t})}, i = 1, 2$$

$$\chi_{sp,t} = \widehat{\chi_{sp,t}} = \chi_{sp,t}^{(2)}$$

where $\widetilde{\sigma_C}$ and $\widetilde{\sigma_R}$ represent the measured variance of χ_C and χ_R receptively when the alternate effort values (i.e. χ_R and χ_C) are near their means (±0.05). β_C and β_R were estimated using a simultaneous and multivariate Bayesian, linear model following relationships described in eq. 5, with σ_C and σ_R being set as $\widetilde{\sigma_C}$ and $\widetilde{\sigma_R}$ respectively, scaled by the square-root of $n_{sp,t}$ (thus, assuming increased precision when the number of captures was high); where $n_{sp,t}$ for a given year *t* was less than 1, scaling was ignored. For this model, priors for β_C and β_R were broad and normal, with means and variance of 1. Final estimate of χ across species and years represented posterior medians from this model (run across 4 HMC chains, each with 5,000 warm-up iterations, 5,000 sampling iterations, and a thinning factor of 10), and positive correlations between χ and both χ_R and χ_C were confirmed by visualisation.

2.3 Pradel model specifications and capture-recapture data augmentation

All vital parameters in our Pradel model were estimated from independent functions. First, probability of detection (*p*) was modelled as a direct and linear function of effort (χ), with both average *p* and effects of χ on *p* allowed to vary across species. Here, time was excluded as a determinant of *p* since we assumed that the majority of temporal effects on *p* would be explained by temporal variations in χ. Second, rates of population change (λ) were assumed to vary stochastically across sampling years, with variation in λ across years differing by species. Similar to λ, probability of residency (*r*) was also assumed to vary species-specifically across sampling years. However, we recognise that *r* is likely to be influenced by which observation sites were active in a given year – a variable with strong temporal inertia that is not truly stochastic. For this reason, *r* values across years were modelled as first-order

random walks per species. Last, annual probability of survivorship (ϕ) was modelled as a linear function of time (i.e. by year, scaled as described for our distributional model), with average ϕ, and the effects of time on ϕ, varying by species. A linear effect of time on ϕ was chosen to ease cross-species comparisons in future analyses. Together, formulae predicting ϕ, γ, p and r were therefore as follows:

$$
\begin{aligned}
p &\sim g^{-1}\left(\left[\beta_0+v_{0sp}\right]+\left[\beta_1+v_{1sp}\right]\cdot\chi\right)\\
\lambda &\sim 2\cdot g^{-1}\left(\beta_0+v_{0sp,year}\right)\\
r &\sim g^{-1}\left(\left[\beta_0+v_{0sp}\right]+\left[\sigma_{rw}\sum_{year=2}^{year}v_{1sp}\right]\right)\\
\phi &\sim g^{-1}\left(\left[\beta_0+v_{0sp}\right]+\left[\beta_1+v_{1sp}\right]\cdot t\right)
\end{aligned}
\tag{8}
$$

where *sp* indicates species, g^{-1} indicates the inverse logit function (thus, fixing select parameters on the probability scale), *year* indicates the sampling year, with 1990 equaling 1, *t* indicates the median-centred and scaled year, σ_{rw} indicates a cross-species, random-walk scaling factor, and v_{sp} indicates adjustments to population-level coefficients for species *sp*, or, in the case of *r*, the species-specific random-walk coefficient. Inverse-logit predictors of λ are multiplied by 2 to allow λ to range between realistic extremes: 0 (a complete population crash) and 2 (a population doubling).

Priors for all cross-species predictors of model parameters were broad and moderately informed by estimates of others (DeSante and Kaschube, 2009). For our cross-species intercept of *p*, we used a normal prior with a mean of 0 and standard deviation of 1.25, while for our effect of χ on *p*, we used a skew-normal prior with an ξ of 0.5, ω of 1.5, and α of 5 (thus assuming a positive value). For both group-level effects of species on *p*, and the relationship between χ on *p*, we used half-normal priors with means of 0.1 and 0.05 respectively, and standard deviations of 0.15. Next, for our intercepts of λ and *r*, we again used normally-distributed priors with means of 0 and -1 respectively (thus, assuming no population change and relatively low probability of residency), and standard deviations of 0.1 and 1 respectively. Our group-level effect of species-year on λ, and our random walk scalar were both modelled with exponential priors, with scale parameters (Λ) of 10 and 7.5 respectively. Last, for our cross-species intercept of ϕ and effect of scaled year on ϕ, we used normally-distributed priors with means of 0 (setting reference ϕ at approximately 0.5 on the probability scale), and standard deviations of 0.75 and 0.1 respectively, while group-level effects of species on each were modelled with half-normal priors with means of 0.1 and 0.01 respectively, and standard deviations of 0.2 and 0.15 respectively.

For this analysis, we included capture-recapture data from 1,620,420 individuals across 151 species, covering 26 families, 1786 capture sites, and spanning 25 years of observation (from 1990-2015). Primary data were obtained from the MAPS programme (DeSante et al, 2004) and supplemented with additional capture data from the United States Geological Survey (USGS) Bird Banding Laboratory database. Execution of our Pradel model was achieved in Stan (version 2.32.2) via the R package "cmdstanR" (Gabry et al, 2025). Initial values for the HMC chains (n= 4) in our model were extracted from posteriors of an initial model executed using meanfield ADVI (as described in section 1.1). In total, our final Pradel model was run across 4 HMC chains, and using 20,000 sampling iterations, thinned by a factor of 40, with the first 10,000 discarded as warm-up. All $\hat{R}$ values from this model fell between 1.00 – 1.03, indicating adequate chain mixing, and effective sample size to sample size ratios averaged 0.98 (± s.d. = 0.10), indicating minimal within-chain autocorrelation.

2.4 Deriving individual-level survivorship

Across individuals, the probability of recovering a captured individual (*R*) in a given year *j* may be described as:

$$R_{j,sp} = r_{j,sp} \cdot \phi_{j,sp} \cdot \left(\prod_{t=t_0+1}^{j-1} \phi_{t,sp} \cdot \left(1 - p_{t,sp}\right) \right) \cdot p_{j,sp} \tag{9}$$

where t_0 indicates an individual's year of last capture, and p_j, ϕ_j, and r_j represent averages for an individual's species (*sp*) in year *j* (see Telenský et al, 2024). Using Bayes' Rule, the probability that an individual survives the year of its first capture (defined as $\phi_{j,t1}$) in the absence of subsequent recaptures can therefore be expressed as:

$$\left(\phi_{t1} \middle| R'\right)_{j,sp} = \frac{\left(1 - p_{j,sp} \cdot r_{j,sp}\right) \cdot \phi_{j,sp}}{\left(1 - p_{j,sp} \cdot r_{j,sp}\right) \cdot \phi_{j,sp} + \left(1 - \phi_{j,sp}\right)} \ . \tag{10}$$

For subsequent years (where residency in the sample area is therefore assumed to be confirmed), φ in the absences of recaptures may then be generalised as:

$$\left(\phi_{tx}\middle|R'\right)_{j,sp}=\left(\phi_{t1}\middle|R'\right)_{j,sp}\cdot\left(\prod_{T=1}^{x}\frac{\phi_{j+T,sp}\cdot\left(1-p_{j+T,sp}\right)}{\phi_{j+T,sp}\cdot\left(1-p_{j+T,sp}\right)+\left(1-\phi_{j+T,sp}\right)}\right)x\in\{2,3...L\} \qquad (11)$$

where L indicates the maximum number of years remaining in an individual of species *sp*'s expected lifespan. After L has been reached, ϕ is then assumed to be 0.

To estimate ϕ across all known individuals and species, we used eq. 11 while also assuming that any future recapture events for individual i implies that $\phi = 1$ in year j. Here, $p_{j,sp}$, $r_{j,sp}$ and $\phi_{j,sp}$ were predicted from our Pradel model (posterior medians), with capture effort – required to predict $p_{j,sp}$ – representing individual estimates in place of species-level averages, and $r_{j,sp}$ equaling 1 if an individual was previously captured at least once (therefore, confirming their residency status). To determine L, we first calculated each species' apparent longevity (i.e. their maximum lifespan; *sensu* Ma and Gladyshev, 2017) from capture-recapture records as the maximum time between capture and recapture events plus the age at first capture (assumed to be 1 if unknown) within that species. L for a given individual was then set to its species' apparent longevity less the number of years between an individual's estimated age from capture records (again, assumed to be 1 if unknown).

2.5 Constructing selection gradients

To determine: (1) how body mass relates to fitness across species, and (2) how this relationship may have changed through time, we combined annual survivorship probabilities with scaled body mass measures to construct time-varying selection gradients, following standard approaches (Lande and Arnold, 1983). Specifically, we constructed a Bayesian linear model with individual-level survivorship estimates (ϕ; from section 2.4), divided by their species-specific means (Lande and Arnold, 1983) as our response variable, and relative body mass (scaled between 0 and 1 within species and sexes, then centred at 0.5), year (scaled by a factor of 10 and median centred, as described above), and the interaction between relative body mass and year as population-level predictors. Hypotheses seeking to explain shifting morphology in birds often assume non-linear patterns of selection (Fig. 1). Thus, to allow for both linear and non-linear effects of body mass on ϕ, we also included squared and cubic forms of relative body mass as predictors, with year interacting with each level. Further, to account for possible spatial variations in slopes of ϕ by year, we also included latitude and longitude of an observation as a Gaussian process predictor (n = 3 basis functions; kernel = $\sigma^2\cdot\exp\left(\frac{-(x-x')^2}{2l^2}\right)$.

Finally, to consider possible variation in body mass, temporal, and spatial effects on ϕ by species, we included group-level intercepts of species on ϕ, as well as group-level slopes of mass, time, the interaction between each, and space (our Gaussian process) per species (where Gaussian processes were constructed group-wise), assuming no correlation between each. We assumed a Student's T distributed error for this model, owing to high excess kurtosis apparent in data visualisations.

Priors for this model were weak and conservative, assuming no effect of body mass and year on survivorship (intercept ~ $\mathcal{N}$[mean = 1, s.d. = 1], body mass ~ $\mathcal{N}$[mean = 0, s.d. = 4], body mass^2 ~ $\mathcal{N}$[mean = 0, s.d. = 16], body mass^3 ~ $\mathcal{N}$[mean = 0, s.d. = 16], year ~ $\mathcal{N}$[mean = 0, s.d. = 0.5], body mass × year $\mathcal{N}$[mean = 0, s.d. = 2.5], body mass^2 × year $\mathcal{N}$[mean = 0, s.d. = 2.5], body mass^3 × year $\mathcal{N}$[mean = 0, s.d. = 2.5], all group-level effects of species ~ exponential[Λ = 2.5], ℓ[Gaussian process length-scale parameter] ~ exponential[Λ = 0.5], standard deviation of Gaussian process ~ Gamma[α = 2, β = 0.75]). Error in this model was assumed to be near-Gaussian (ν ~ Gamma[α = 2, β = 1.25]).

Only observations where both ϕ and body mass were available in a give year were used to construct our selection gradient model. Final data used in this analysis thus included 85,880 observations collected between 1992 and 2014, and spanning across 150 species and 26 families. Our model was run across 4 HMC chains, again using 14,000 sampling iterations thinned by a factor of 10, with the first 8,000 discarded as warm-up. $\hat{R}$ and Neff/N values for all coefficients fell between 1.00 – 1.01, and 0.55 - 1.08 respectively.

4.0 Partitioning effects of selection and plasticity on changes in body mass distributions

To partition the contributions of natural selection and phenotypic plasticity toward changes in body mass distributions across species, we constructed a novel, "functional" Price equation that accepts phenotypic change functions on the left side of the equation, instead of changes in mean phenotypes. Briefly, the Price equation is an exact calculation that describes the change in the average value of a trait within a population or species. In its original form, this change is defined as the sum of the scaled covariance between that trait and relative fitness (a "selection term"), and the combined effects of plasticity, drift, and mutational effects (often defined as the "transmission", or "non-selective" term; Price 1970). While a useful starting point, this equation does not allow for partitioning of selective and non-selective effects toward more complex changes in phenotypes within a population, for example, in skewness and kurtosis. Moreover, by defining the selection term as covariance between two continuous variables, Price's equation thus does not allow for integration of, or distinguishment

between, forms of "variance selection" (i.e. stabilising selection, disruptive selection, or even the relaxation of selection), or otherwise non-linear selection. Our functional Price equation corrects for both shortcomings, and as such, offers opportunities to apply Price's identity to answer questions about the role of selection in shaping traits at a holistically level.

Our functional Price equation is derived as follows: setting z as a trait value, and p_1 and p_2 as probability density functions (pdfs) across the set of z values (Z) at time-points 1 and 2 respectively, the change in a trait's pdf [$\Delta p(z)$] between two adjacent time-points is given as:

$$\Delta p(z) = p_2(z) - p_1(z) \tag{12}$$

or equivalently:

$$\Delta p(z) = \left(p'_1(z) - p_1(z)\right) + \left(p_2(z) - p'_1(z)\right) \tag{13}$$

where $p'_1(z)$ represents an intermediate pdf, and the first and second summands represent changes in $\Delta p(z)$ driven by selective (henceforth ∇_{NS}) and non-selective effects (∇_{EC}) respectively (*sensu* Price 1972). Setting f_1 as a selection gradient (i.e. a function relating a given trait value to its relative fitness value), and $\mathbb{E}_{p_1}[f_1]$ as the expectation of f_1 under p_1, $p'_1(z)$ may be explicitly defined as:

$$p'_1(z) = \frac{f_1(z)\, p_1(z)}{\mathbb{E}_{p_1}[f_1]} \tag{14}$$

and ∇_{NS} therefore as:

$$\begin{aligned} \nabla_{NS} &= \frac{f_1(z)\, p_1(z)}{\mathbb{E}_{p_1}[f_1]} - p_1(z) \\ &= \left(\frac{f_1(z)}{\mathbb{E}_{p_1}[f_1]} - 1\right) p_1(z) \end{aligned} . \tag{15}$$

Next, using a displacement kernel (K_D) to describe the probability of a given trait value (z) being obtained from another trait value (y) within the same set Z (i.e. via plasticity or imperfect inheritance), ∇_{EC} is defined as:

$$\nabla_{EC} = \int_{y \in Z} K_D(z|y)\, p'_1(y) - p'_1(z)\, dy, \tag{16}$$

or:

$$\nabla_{EC} = \int_{y \in Z} \left[K_D(z|y) - \delta(z-y) \right] p'_1(y)\, dy \tag{17}$$

where δ represents the Dirac delta (setting $p'_1(y) = \int p'_1(z)\,\delta(z-y)\,dy$). Combining eq. 15 and 17, and expanding $p'_1(y)$, we obtain the near-final form:

$$\Delta p(z) = \left(\frac{f_1(z)}{\mathbb{E}_{p_1}[f_1]} - 1 \right) p_1(z) + \frac{\int_{y \in Z} p_1(y) f_1(y) \left[K_D(z|y) - \delta(z-y) \right] dy}{\mathbb{E}_{p_1}[f_1]}; \tag{18}$$

Given K_D is often unknown, the non-selective term (∇_{EC}) is not always calculable and may only be solved given a known change in a probability distribution [$\Delta p(x)$] and a known selection gradient f_1. In these cases, setting the non-selective term to ε_1, eq. 18 is further simplified to a more useful form:

$$\Delta p(z) = \left(\frac{f_1(z)}{\mathbb{E}_{p_1}[f_1]} - 1 \right) p_1(z) + \epsilon_1(z). \tag{19}$$

To evaluate and compare the relative contributions of both the selection and transmission term toward $\Delta p(x_2)$, we offer two approaches: one in L_1 form (analogous to regression slopes) and one in L_2 form (analogous to Pearson R^2 values). For our L_1 solution, contributions of the selective and non-selective terms (set as $\partial_1 S$ and $\partial_1 E$ respectively) are defined as:

$$\partial_1 S = 1 - \frac{\left(\int \left| \Delta p(z) - \left(\frac{f_1(z)}{\mathbb{E}_{p_1}[f_1]} - 1 \right) p_1(z) \right| dz \right)}{\int |\Delta p(z)|\, dz}; \partial_1 E = 1 - \partial_1 S, \tag{20}$$

or more stably as:

$$\partial_1 S = \frac{\int |\nabla_{NS}| dz}{\int |\nabla_{NS}| dz + \int |\nabla_{EC}| dz} \partial_1 E = 1 - \partial_1 S, \tag{21}$$

where ∇_{EC} is obtained from $\Delta p(z)$ - $\nabla_{NC.}$

For our L_2 solution, these contributions (set as ∂_2S and ∂_2E respectively) are thus:

$$\partial_2 S = \frac{\left(\int \Delta p(z) \cdot \left(\left(\frac{f_1(z)}{\mathbb{E}_{p_1}[f_1]} - 1 \right) p_1(z) \right) p_1(z) dz \right)^2}{\left(\int (\Delta p(z))^2 p_1(z) dz \right) \cdot \left(\int \left[\left(\frac{f_1(z)}{\mathbb{E}_{p_1}[f_1]} - 1 \right) p_1(z) \right]^2 p_1(z) dz \right)}; \partial_2 E = 1 - \partial_2 S. \tag{22}$$

4.1 Validating Price derivations

To validate our functional Price equation, we generated counter-factual body mass pdfs per species and year, where changes in these pdfs are explained purely by estimated selection gradients operative at that time. If contributions of selection toward a change in a mass pdf are high (as estimated by our Price equation), we should expect that the difference between this counter-factual mass pdf and an observed, "true" mass pdf for the same year and species are small. To formally test this prediction, we first constructed counter-factual mass pdfs as follows:

$$h(x)_{sp,t}' = \frac{f(x-0.5)_{sp,t-1} \cdot Beta(x|\alpha_{sp,t-1}, \beta_{sp,t-1})}{\int_0^1 \left(f(x-0.5)_{sp,t-1} \cdot Beta(x|\alpha_{sp,t-1}, \beta_{sp,t-1}) \right) dx} \tag{23}$$

where x indicates a scaled body mass, $f(x)$ again represents a selection gradient for a given species sp and year t (where x was centred at 0; see section 2.5), α and β are posterior medians predicted per species and year at a species' median latitude (from our distributional model), and $h(x)'$ represents the counter-factual mass pdf. To approximate the precise form of $h(x)'$ – a requirement for KL divergence calculation – we used linear interpolation occurring across 250 evenly spaces points. KL divergences between $h(x)'$ and "true" body mass pdfs (at equivalent time points and species) were then calculated using standard methods:

$$KL_{sp,t} = \int_0^1 h(x)'_{sp,t} \cdot \ln\left(\frac{h(x)_{sp,t}'}{Beta(x|\alpha_{sp,t}, \beta_{sp,t})}\right) dx \quad (24)$$

where *KL* represents the KL divergence for a given species *sp* and time *t*. KL divergence values were then plotted against ∂_1S from our functional Price equation at equivalent time-points per species to confirm a negative relationship. We then confirmed this negative relationship by modelling natural-log transformed KL divergence values as a linear function of ∂_1S using simple, Bayesian linear regression, with weak priors on the effect of either selective contribution term ($\mathcal{N}$[mean = 0, s.d. = 1]) and on model intercepts ($\mathcal{N}$[mean = -5, s.d. = 2.5]). This model was run using 4 HMC chains, sampled with 50,000 iterations, thinned by a factor of 20, and with the first 10,000 iterations removed as warm-up. $\hat{R}$ and N_{eff}/N values for all parameters ranged between 0.99 and 1.01, and 0.93 and 1.01 and respectively.

Our ∂_1S term (termed "selective magnitude") negatively correlated with the KL divergence between true body mass distributions and expected (or "counterfactual") distributions, derived from selective effects alone (β = -1.91 [>99.9%]). More specifically, expected body mass distributions – solely defined by selective effects – were more similar to true body mass distributions when ∂_1S was high. This findings confirms that the selective contribution derived from our Price equation (∇_{NS}) indeed positively reflects that magnitude of estimated selective effects.

4.2 Evaluating shifts in contributions of natural selection toward body mass distributions

To evaluate whether and how the relative contributions of natural selection toward defining a species' body mass distribution may be changing, we calculated ∂_1S and ∂_2S for all species with defined body mass selection gradients (obtained in section 2.5; setting latitude and longitude as the geographical centre of each species' range, and extracting posterior medians) and time-specific body mass distributions (obtained in section 1.1; assuming median latitudes per species) between the years of 1990 and 2015. In total, this included data from 150 species and 26 families spanning from 1992-2014. Given that both ∂_1S and ∂_2S are directly solutions and not statistical estimates, functions of each across time have a precise polynomial form. However, we were interested in understanding the general, linear trajectories of these polynomial forms. As such, once calculated, we constructed two Bayesian linear models with ∂_1S and ∂_2S as beta-distributed response variables, and with year (median-centred and scaled by decade) as the sole population-level predictor in both models. To further understand the role of phylogeny in shaping temporal changes in ∂_1S and ∂_2S, we allowed both intercepts and temporal

slopes to vary by phylogenetic covariance between species, as group-level effects. Calculation of phylogenetic covariation is described in section 1.3. Priors for model intercepts we normal with means of 0 and standard deviations of 1 in natural log space (thus assuming 50% of mass distributional shifts explained by selection, in raw space). For effects of time on ∂_1S and ∂_2S, we also assumed normally-distributed priors with means of 0 and standard deviations of 1. Last, for effects of phylogenetic covariance on intercepts and time-dependent slopes, we used exponential priors with lambda values (Λ) liberally set to 0.5. Similar to other models, these models were run using 10,000 sampling iterations across 4 HMC chains, each thinned by a factor of 10, and with the first 5,000 iterations in each chain being removed as warm-up.

5.0 Statistical reporting

All model coefficients represent posterior modes, unless otherwise stated. Posterior probabilities represent the smallest, symmetrical highest posterior density intervals around posterior modes that include 0.

References

DeSante, D.F., Saracco, J.F., O'Grady, D.R., Burton, K.M., Walker, B.L. Methodological considerations of the Monitoring Avian Productivity and Survivorship (MAPS) program. *Stud. Avian Biol. 29*(1):28-45. (2004).

Youngflesh, C., Saracco, J.F., Siegel, R.B., Tingley, M.W. Abiotic conditions shape spatial and temporal morphological variation in North American birds. *Nat. Ecol. Ecol.* 6(12), 1860-1870. (2022).

Kapteyn, J.C. Skew frequency curves in biology and statistics. *Recueil Trav. Bot. Néerl.* 13(2), 105-157. (1916).

Pick, J.L., Lemon, H.E., Thomson, C.E., Hadfield, J.D. Decomposing phenotypic skew and its effects on the predicted response to strong selection. *Nat. Evol. Ecol.* 6(6), 774-785. (2022).

Pollo, P. et al. Beyond sex differences in mean: meta-analysis of differences in skewness, kurtosis, and correlation. *EcoEvoRxiv.* (2025).

Ashton, K. G. Patterns of within species body size variation of birds: strong evidence for Bergmann's rule. *Glob. Ecol. Biogeog. 11*(6):505-523. (2002).

Bergmann, C. Über die Verhältnisse der warmeokonomiedie Thiere zu ihrer grösse. Göttinger Stud. 3(1):359-708. (1847).

Hitch, A.T., Leberg, P.L. Breeding distributions of North American bird species moving north as a result of climate change. *Cons. Biol. 21*(2):534-539. (2007).

Meiri, S., Dayan, T. On the validity of Bergmann's rule. *J. Biogeog. 30*(3):331-351. (2003).

Van Buskirk, J., Mulvihill, R.S., Leberman, R.C. Declining body sizes in North American birds associated with climate change. *Oikos 119*(6):1047-1055. (2010).

Weeks, B.C. et. al. Shared morphological consequences of global warming in North American migratory birds. *Ecol. Lett.* 23(2), 316-325. (2020).

Kucukelbir, A., Ranganath, R., Gelman, A., Blei, D. (2015). Automatic variational inference in Stan. *Advances in Neural Information Proc. Syst. 28*(1):1-9. (2015).

Gabry J, Češnovar R, Johnson A, Bronder S *cmdstanr: R Interface to 'CmdStan'*. R package version 0.9.0, https://discourse.mc-stan.org. (2025).

Gelman, A., Rubin, D.B. Inference from iterative simulation using multiple sequences. *Stat. Sci.*, *7*(4):457-472. (1992).

Pradel, R. Utilization of capture-mark-recapture for the study of recruitment and population growth rate. *Biomet. 52*(1):703-709. (1996).

Telenský, T., Storch, D., Klvaňa, P., Reif, J. Extension of Pradel capture–recapture survival-recruitment model accounting for transients. *Methods Ecol. Evol. 15*(2):388-400. (2024).

Sandercock, B.K., Murray, D.L., Sandercock, B.K. Mark-recapture models for estimation of demographic parameters. In D.L. Murray & B.K. Sandercock (Eds.) *Population Ecology in Practice* (pp. 157-190). John Wiley & Sons Ltd. (2020).

Riddle, J.D., Mordecai, R.S., Pollock, K.H., Simons, T.R. Effects of prior detections on estimates of detection probability, abundance, and occupancy. *Auk* 127(1), 94-99. (2010).

DeSante, D.F., Kaschube, D.R. The monitoring avian productivity and survivorship (MAPS) program 2004, 2005, and 2006 report. *Bird Pop. 9*(1):86-169. (2009).

Ma, S., Gladyshev, V.N. Molecular signatures of longevity: Insights from cross-species comparative studies. *Sem. Cell Devel. Biol. 70*(1):190-203. (2017).

Lande, R., Arnold, S.J. The measurement of selection on correlated characters. *Evol. 37*(6):1210-1226. (1983).

Price, G. Selection and covariance. *Nature* 227(1):520-521. (1970).

McTavish, E.J. et al. A complete and dynamic tree of birds. *BioRxiv*, 2024-05. (2024).

Bouckaert, et. al. BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comp. Biol. 10*(4):e1003537. (2014).

Tabh, J. K., Hartjes, M., Burness, G. Endotherms trade body temperature regulation for the stress response. *Proc. Roy. Soc. B. 290*(2010):20231251. (2023).