

Running title: Morphological change in Atlantic Forest birds

Weak and species-specific morphological change in Atlantic Forest birds across two decades of warming

Eduardo S.A. Santos^{1*}, Ayumi Mizuno¹, Santiago Ortega¹, Gustavo Burin²

¹ Collaboration for Open Science and Synthesis in Ecology and Evolution, Department of Biological Sciences, University of Alberta, Edmonton, AB, T6G 2E9, Canada

² Department of Biological and Environmental Sciences, University of Gothenburg, Sweden

ORCID: Eduardo S.A. Santos: 0000-0002-0434-3655 | Ayumi Mizuno: 0000-0003-0822-5637 | Santiago Ortega: 0000-0002-3518-276X | Gustavo Burin: 0000-0001-7851-0666

* **Correspondence:** Eduardo S.A. Santos (esantos2@ualberta.ca)

Running title: *Morphological change in Atlantic Forest birds*

Abstract

Birds are widely expected to shrink as the climate warms, but few tropical studies span decades. In the central Amazon, mean body mass fell in all 77 species studied and a third of species lengthened their wings. We analysed 24 years (1995–2018) of morphology from 73 passerine species (12,571 live-capture records; ATLANTIC BIRD TRAITS) across the Brazilian Atlantic Forest, a region that warmed by about 0.25°C per decade. The records came from 47 research programmes whose sites changed over time, so we modelled contributor and municipality explicitly and checked by simulation that these models recover trends planted in the real sampling structure. Unadjusted models suggested a steep wing-length decline and increases in bill width and trait variability. Once sampling turnover was modelled, the wing decline was less than half as steep (–1.0% per decade, 95% credible interval spanning zero), and the bill-width and variability changes disappeared. Body mass declined by 0.9% per decade, close to the mean Amazonian rate of 1.0%, but its interval spanned zero (–2.8 to +1.0%). Responses varied strongly among species, and most species showed no change distinguishable from zero. Wings did not lengthen, and among individuals measured for both traits, mass tended to be retained relative to wing length. Neither trait was associated with interannual temperature anomalies. Atlantic Forest passerines thus show neither the consistent mass loss across species nor the wing elongation reported from the Amazon, although their average mass change is compatible with the Amazonian rate. Compiled trait databases can extend long-term monitoring to under-studied tropical regions, but only when sampling structure is modelled: ignoring it here would have more than doubled the estimated wing decline.

Keywords: body size, climate change, phenotypic variation, macroecology, Atlantic Forest, phylogenetic comparative analysis

Plain Language Summary

As global temperatures rise, many animals are expected to become smaller or change shape to cope with heat. In the intact Amazon rainforest, every bird species studied became lighter over four decades, and many grew longer wings. We asked whether birds in another tropical forest, the Brazilian Atlantic Forest, changed in the same way, using 24 years of field measurements (1995–2018) of 73 perching-bird species. On average, Atlantic Forest birds became lighter at about the same rate as Amazonian birds, but the estimate was uncertain and species went in both directions rather than all declining together. Wings did not grow longer, and bills did not change. The data came from many research groups whose field sites changed over the years, and simpler analyses exaggerated how much the birds had changed: allowing for who collected the

data, and where, more than halved the apparent shortening of wings. Computer simulations confirmed that this correction recovers real trends rather than hiding them.

1. Introduction

Climate warming can alter animal morphology through several pathways: direct thermoregulatory selection, changes in resource availability, and developmental or evolutionary responses to a shifting environment (Gardner et al. 2011; Sheridan and Bickford 2011; Teplitsky and Millien 2014). These mechanisms have generated important predictions, most familiarly that endotherms should become smaller in warmer environments, because a higher surface-area-to-volume ratio aids metabolic heat dissipation (Bergmann's rule, Bergmann 1847; Mayr 1956; Ashton 2002), and that heat-dissipating extremities such as bills should enlarge (Allen's rule, Allen 1877; Ryding et al. 2021). Both rules were formulated to describe spatial patterns, across latitudes and climates; the expectation that animals should shrink, or that their appendages should lengthen, as a given locality warms is a temporal extension of them and a separate hypothesis (Gardner et al. 2011; Teplitsky and Millien 2014). Empirical support for such general responses is, however, mixed. Broad syntheses indicate that morphological responses to contemporary warming vary substantially among traits, species, and ecological contexts rather than following a single directional pattern. Radchuk et al. (2019), synthesising phenotypic responses across wild populations, recovered no consistent systematic morphological response to warming. Understanding this heterogeneity requires determining whether the patterns documented in well-studied systems extend to ecological and geographic settings that are underrepresented in long-term datasets.

Multi-decadal morphological analyses have predominantly drawn on temperate, often migratory, bird species sampled through museum collections or long-term ringing stations in the Northern Hemisphere (e.g., Weeks et al. 2020). Tropical communities, where avian diversity peaks, are far less represented, so it remains unclear how general the reported patterns are. Against this background, Jirinec et al. (2021) provides a valuable tropical investigation: analysing four decades (1979–2019) of standardised field measurements from 77 non-migratory understory species in the continuous primary rainforest of the central Amazon (Biological Dynamics of Forest Fragments Project, BDFFP), they documented widespread declines in body mass alongside wing elongation in roughly a third of species, and interpreted the resulting reduction in wing loading as an adaptive shift towards flight economy in an increasingly warm understory. Whether that pronounced Amazonian response is representative of tropical forest birds more broadly, or is instead context-dependent, is the question we address here.

The Brazilian Atlantic Forest offers an informative contrast. It is a global biodiversity hotspot that differs from the continuous lowland central Amazon in almost every

respect that might matter: it spans steep topographic and latitudinal gradients, and it has lost over 80% of its original forest cover, leaving the remainder divided among thousands of small, isolated fragments (Ribeiro et al. 2009). It is therefore a useful test of whether the Amazonian pattern generalises to a tropical forest system with a very different landscape and thermal structure.

Two additional axes may help characterise whatever response is present. First, warming may act on phenotypic variance as well as on trait means, either widening among-individual variation through developmental instability or relaxed selection, or narrowing it through directional canalisation (Teplitsky and Millien 2014; Debat and David 2001; Badyaev 2005). Second, morphological change may be driven indirectly through food supply rather than directly through thermoregulation (Schmidt and Jensen 2005; Teplitsky et al. 2008). If so, species that rely on different foods need not respond in the same way, and diet, from predominantly frugivorous or granivorous to strictly insectivorous, may help explain why some species change and others do not.

Testing this requires long series, and in most tropical regions these exist only as compilations of records from many research programmes. Trait databases such as ATLANTIC BIRD TRAITS (Rodrigues et al. 2019) now assemble decades of field measurements across whole biomes, and mixed models can separate a temporal trend from the identity of the programme and site that supplied each record. This separation matters because programmes and sites rarely persist through a record. If those sampled late in a series differ from those sampled early, in measurement practice, protocol or habitat, a compilation can show a temporal trend that no population has undergone. How strongly such turnover can bias trends in compiled morphological data is seldom quantified.

Here, we use 24 years (1995–2018) of morphological data from 73 passerine species (12,571 live-capture records; ATLANTIC BIRD TRAITS; [Figure 1](#)) to ask whether Atlantic Forest passerines show consistent morphological change across a period of regional warming. We characterise that response in three focal traits—body mass, wing length, and bill width—and then ask how it is structured: whether mass and wing length change in concert on the same individuals or depart from geometric isometry, whether morphology covaries with interannual temperature anomalies, whether residual phenotypic variability has changed, and whether diet explains variation among species. Because the data were compiled from dozens of research programmes whose sites shifted over time, we model contributor and municipality explicitly, test by simulation whether these models recover known trends in the real sampling structure, and compare the resulting rates with those of the central Amazon on a common scale.

2. Results

2.1 The region warmed while contributing programmes and sites changed

The analytical dataset comprised 73 passerine species and 12,571 adult live-capture records (8,478 wing records across 72 species; *Herpsilochmus pectoralis* has no wing measurements in the compilation and enters only the analyses of the other traits), gathered between 1995 and 2018 at 455 coordinate localities spanning the latitudinal and elevational range of the Brazilian Atlantic Forest, from sea level to over 1,500 m (Figure 1). Filtering criteria are given in the Methods.

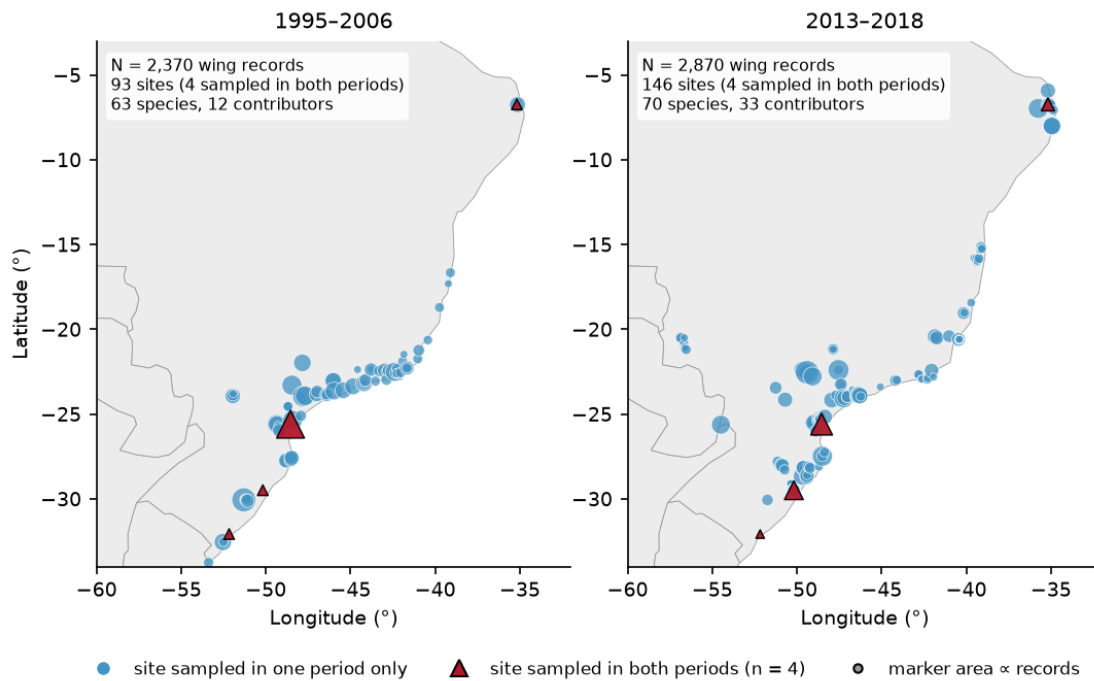


Figure 1: Geographic distribution and multi-decadal sampling of 73 passerine bird species across the Brazilian Atlantic Forest domain. Sampling localities are displayed separately for the early (1995–2006; $n = 2,370$ wing records, 93 coordinate sites, 12 contributors) and late (2013–2018; $n = 2,870$ wing records, 146 coordinate sites, 33 contributors) sampling periods across the biome. Points represent sampling localities coloured by mean wing length, illustrating the broad latitudinal and elevational range covered by the dataset.

Historical climate data, available for 357 of these localities (the remainder fall outside the WorldClim land mask; see Methods), documented regional warming over the study period: mean annual temperature rose by $+0.246^{\circ}\text{C}/\text{decade}$ (95% CI: $+0.076$ to $+0.415^{\circ}\text{C}/\text{decade}$), representing an average warming of $+0.56^{\circ}\text{C}$ across capture sites between 1995 and 2018. In contrast, precipitation ($-44.7 \text{ mm}/\text{decade}$, 95% CI: -123.1 to

+33.7 mm) and drought indices (SPEI-12: -0.256, 95% CI: -0.617 to +0.105) showed no clear directional trend.

These records were contributed by 47 research programmes working in 153 municipalities (42 and 130 in the wing sample), and both the set of active programmes and the geographic focus of sampling shifted substantially over the 24 years (Figure 1; Supplementary Fig. S1).

2.2 Wing length declined weakly once sampling turnover was modelled, and responses varied among species

Model-free comparisons of per-species mean trait values between the early (1995–2006) and late (2013–2018) sampling periods, the lower and upper quartiles of capture year, showed substantial variation across species (Figure 2): wing length decreased in 43 species and increased in 18 (Figure 2 A), body mass decreased in 36 and increased in 26 (Figure 2 B), and bill width decreased in 13 and increased in 31 (Figure 2 C).

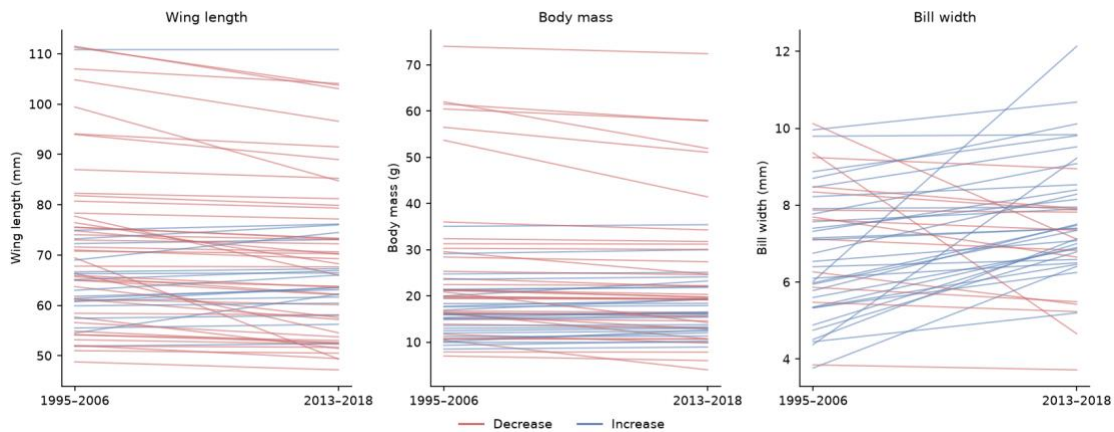


Figure 2: Species-level trajectories of wing length, body mass, and bill width between early and late sampling periods. Paired per-species means between the early (1995–2006) and late (2013–2018) periods for (A) wing length ($n = 61$ species), (B) body mass ($n = 62$ species), and (C) bill width ($n = 44$ species) across Atlantic Forest passerines. Lines connect period means for each species, coloured by direction of shift (red = decrease, blue = increase), illustrating the raw interspecific variation across the assemblage.

Unadjusted models suggested a relatively steep decline in wing length over time (baseline model: -1.8 mm/decade, 95% CrI: -2.8 to -0.9; -2.6%/decade; $P(\beta < 0) > 0.99$), but the estimated trend weakened substantially after accounting for contributor, municipality and the other sources of sampling turnover. In the fully adjusted model, the assemblage-level trajectory was modest and its credible interval overlapped zero: -0.7 mm/decade (95% CrI: -1.6 to +0.1; -1.0%/decade; $P(\beta < 0) = 0.96$; Figure 3). The estimate was similar on complete cases (-0.6 mm/decade, 95% CrI: -1.3 to +0.2) and for log wing length (-0.9%/decade, 95% CI: -2.0 to +0.2%; REML check), and slightly

steeper, with an interval excluding zero, after removing 6 anomalous sampling blocks found by a post hoc screen (-0.9 mm/decade, 95% CI: -1.6 to -0.1 ; Methods; Supplementary Tables S1 and S5). Letting species slopes covary phylogenetically left the estimate similar but its interval much wider (-0.5 mm/decade, 95% CrI: -2.3 to $+1.4$), and the phylogenetic share of slope variance was poorly constrained ($H^2 = 0.24$, 95% CrI: 0.01 to 0.69), so the data do not tell us how far species' responses follow phylogeny.

Separating the year effect within and between contributors helped to locate some of the attenuation. Capture year varied substantially within contributors: 8 of the 42 contributors, holding 3,382 wing records, sampled over 10 or more years, and 42% of the variance in capture year lay within contributors. The within-contributor slope was -0.3 mm/decade (95% CrI: -1.6 to $+1.0$; $P(\beta < 0) = 0.69$); its interval includes both zero and the fully adjusted assemblage trend, so it cannot distinguish no change from a modest decline. Contributors whose records came, on average, from later years recorded shorter wings (-2.6 mm per decade of contributor mean year, 95% CrI: -4.9 to -0.2 ; $P(\beta < 0) = 0.98$), a term estimated from only 42 contributor means. Municipalities offer less temporal replication: only 18 of 130 municipalities (2,497 wing records) were sampled over 10 or more years, and 26% of the variance in capture year lay within municipalities. The corresponding within-municipality slope (REML adjustment) was -1.5 mm/decade (95% CI: -3.4 to 0.4). The decompositions are therefore consistent with much of the steep unadjusted decline in wing length reflecting changes in sampling structure, but they do not establish how much temporal change remains within research programmes or sites.

To check that the adjusted model neither absorbs a real trend nor leaves an artefact in place, we planted known trends in 200 simulated data sets per scenario that kept the real records, species, individuals, contributors, municipalities and years (Methods). With contributor and municipality effects independent of year, the fully adjusted model recovered a planted decline of -1.00 mm/decade almost exactly (mean estimate -0.99). With contributor effects that drifted with contributors' mean capture year, calibrated to the gap between the real unadjusted and adjusted estimates, and no planted trend, the baseline model returned -1.46 mm/decade and the fully adjusted model -0.10 . Across scenarios, the 95% intervals of the fully adjusted model covered the planted value in 92–96% of simulations. They excluded zero in 68–79% of simulations with a planted decline of 1 mm/decade (about 1.4% of wing length), in 22–40% with 0.5 mm/decade, and in 6–8% with no trend (Supplementary Table S4). The Atlantic Forest bird data could therefore detect a wing decline of about 1 mm/decade with moderate power, but not one half that size.

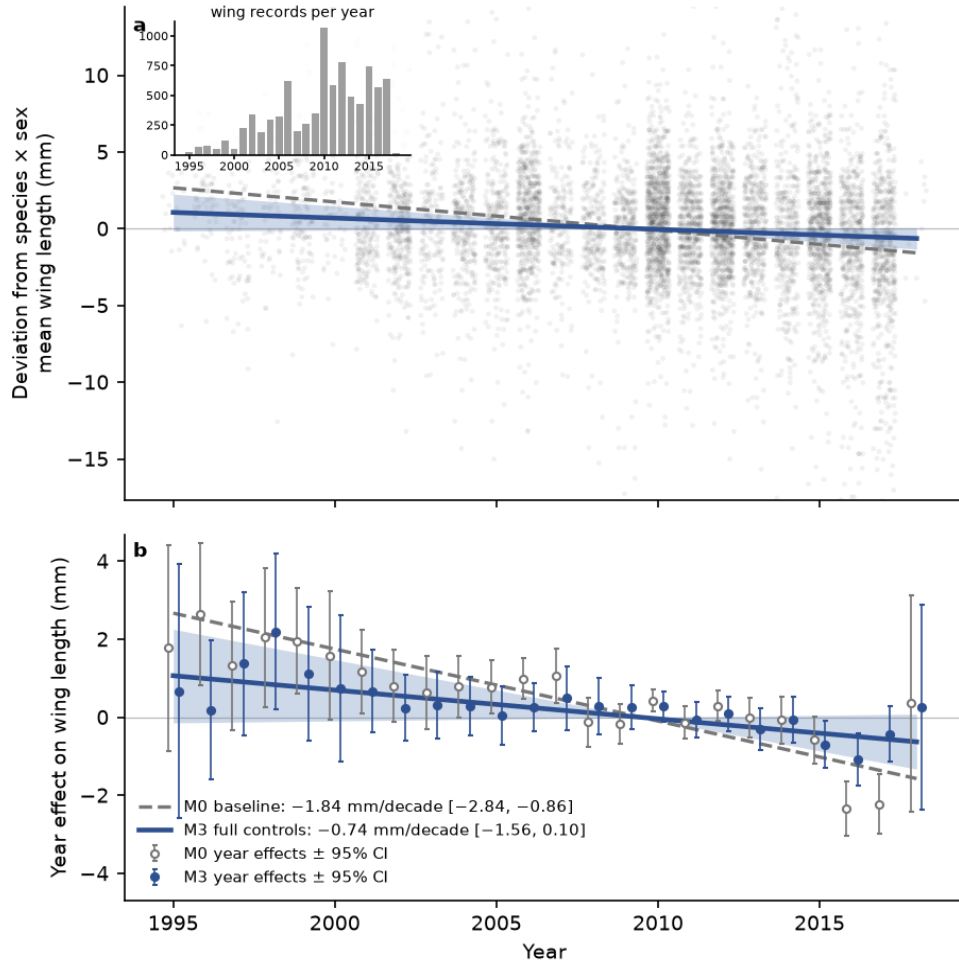


Figure 3: Modelling sampling turnover more than halves the estimated wing-length decline. (A) Within-species-centred wing length over time across 8,478 records from 72 species. The dashed line represents the baseline model (-1.8 mm/decade); the solid line and band represent the fully adjusted model (-0.7 mm/decade, 95% CrI spanning zero). Inset displays wing records per year. (B) Effects of each capture year on wing length from the baseline (open circles) and fully adjusted (filled circles) models with year as a factor, centred on the record-weighted mean year, with 95% confidence intervals (glmmTMB REML, pooled over 50 phylogenetic trees); the fitted linear trends of panel A are repeated. Unlike raw annual means, these intervals account for the clustering of records within species, contributors and municipalities. Linear trends and their bands are posterior means and 95% credible intervals pooled over 20 phylogenetic trees.

Species-specific slopes from the fully adjusted model showed pronounced interspecific heterogeneity (Figure 4): 49 of 72 species with wing measurements had negative posterior means, but 57 of 72 had 95% credible intervals overlapping zero. Only 13 species declined with intervals excluding zero, and 2 increased. The median species slope was -0.55 mm/decade. Overall, wing-length trajectories were weakly negative at the assemblage level but highly heterogeneous among species, with substantial uncertainty around the average trend.

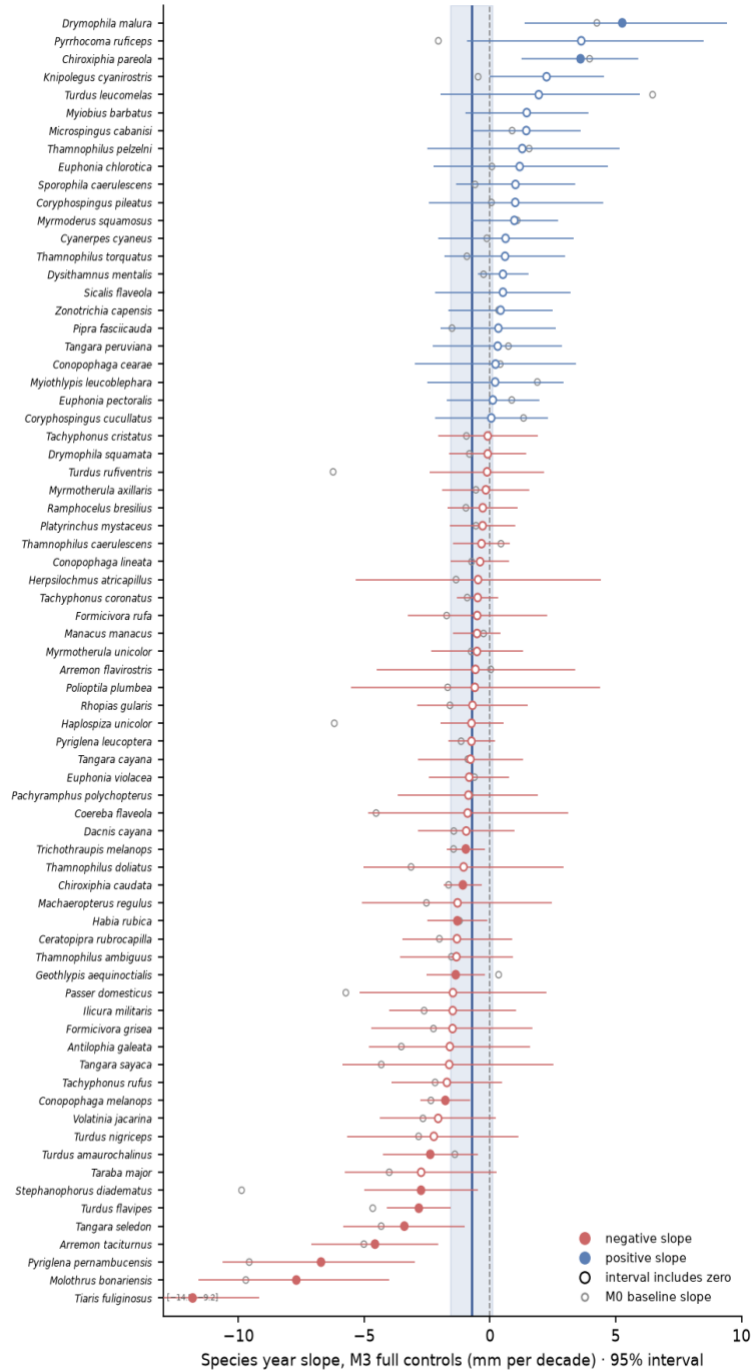


Figure 4: Interspecific heterogeneity in decadal wing-length trajectories under the fully adjusted model. Forest plot of species-specific slopes (mm/decade) across 72 passerine species, ordered by rate of change. Horizontal bars show 95% credible intervals; filled circles denote species whose intervals exclude zero (red = negative, blue = positive), while open circles denote species whose intervals span zero ($n = 57$ species). Shaded vertical band and solid blue line indicate the assemblage year slope (-0.7 mm/decade, 95% CrI: -1.6 to $+0.1$). Hollow grey circles indicate baseline-model slopes.

2.3 Body mass showed no clear directional change

Body mass showed no clear directional change across the 24-year record ($n = 11,077$ records). The baseline estimate was modest ($-1.4\%/decade$, 95% CrI: -3.3 to $+0.4$; $P(\beta < 0) = 0.93$). Under full adjustment it was $-0.9\%/decade$ (95% CrI: -2.8 to $+1.0$; $P(\beta < 0) = 0.83$), close in proportional terms to the fully adjusted wing-length trend ($-1.0\%/decade$) but less precise, and within contributors it was $-0.6\%/decade$ (95% CrI: -2.0 to $+0.9$). Unlike wing length, the body-mass estimate changed little from the unadjusted to the fully adjusted model (Supplementary Table S2), and its interval still allows a decline of up to 2.8% per decade.

2.4 Mass tended to be retained relative to wing length

We next asked how the two traits changed together, using 7,409 shared capture events in which wing length and body mass were measured on the same individual, across 72 species. Under strict geometric isometry mass scales with the cube of wing length, so we modelled the isometry contrast $\Delta_{iso} = \ln M - 3 \ln L$ (Methods): it stays constant if birds shrink or grow without changing shape, increases if mass is retained relative to wing length, and decreases if birds become lighter for their wing length. We call this last pattern slenderisation, the direction of change reported for the central Amazon.

On these shared records, under full adjustment, wing length changed by $-0.8\%/decade$ (95% CrI: -1.9 to $+0.2\%$) and body mass by $-0.5\%/decade$ (95% CrI: -2.2 to $+1.3\%$; [Figure 5 A](#)). Strict isometry would predict a mass change of $-2.5\%/decade$ from the wing trend alone. Fitted directly, the isometry contrast changed by $+2.3\%/decade$ (95% CrI: -0.5 to $+5.1\%$; $P(\beta > 0) = 0.94$), and by $+2.4\%/decade$ (95% CrI: -0.5 to $+5.3\%$; $P(\beta > 0) = 0.95$) under the parsimonious adjustment set. The direct estimate differs slightly from the contrast implied by the separate wing and mass fits ($+2.0\%/decade$), because each model estimates its own species slopes and random-effect variances; the direct fit is the test of isometry. Both estimates point towards mass being retained relative to the wing, but in both the credible interval includes zero, so the shared-individual data do not exclude strict geometric isometry. Because body mass changed little, the contrast mostly reflects the uncertain wing decline. Phylogenetically covarying species slopes gave a similar estimate with a wider interval ($+1.9\%/decade$, 95% CrI: -2.9 to $+6.6\%$).

In bivariate morphospace ([Figure 5 B](#)), the assemblage estimate lies close to stasis and above the isometry line, not in the quadrant of longer wings and lower mass that corresponds to the direction of change reported for the central Amazon (Jirinec et al. 2021), although the credible intervals of both traits include zero. Only 8 of 72 species had posterior mean slopes in that quadrant.

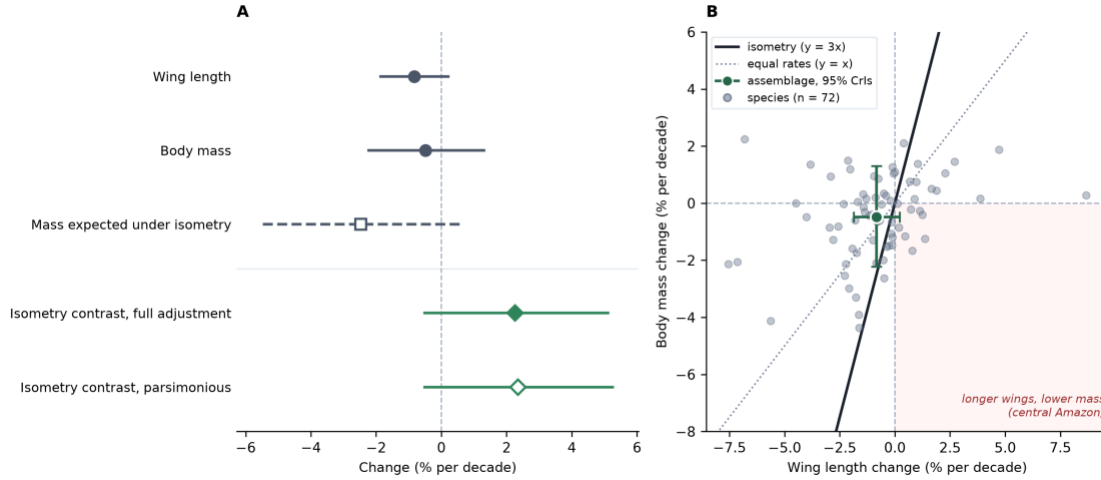


Figure 5: Bivariate temporal allometry and test of geometric isometry on shared individuals. (A) Decadal rates of morphological change (% per decade) on 7,409 shared records (72 species) with simultaneous wing length and body mass measurements, under full adjustment. Points and horizontal bars indicate posterior means and 95% credible intervals, pooled over 20 phylogenetic trees, for wing length, body mass, the mass change expected under cubic geometric scaling ($M \propto L^3$), and the isometry contrast ($\Delta_{iso} = \ln M - 3 \ln L$, fitted directly under the full and the parsimonious adjustment sets). (B) Bivariate morphological state space (Δ Wing vs. Δ Mass). The solid black line indicates strict geometric isometry ($y = 3x$), the dotted line indicates equal rates ($y = x$), and the shaded quadrant (longer wings, lower mass) marks the direction of change reported for the central Amazon (Jirinec et al. 2021), which is shown as a direction only (the rates are compared in the Discussion). The green point is the Atlantic Forest assemblage estimate with the marginal 95% credible intervals of wing and mass change; the two traits were fitted in separate models, so no joint credible region is shown. Semi-transparent circles are species posterior means ($n = 72$).

2.5 Annual temperature anomalies were not associated with the wing–mass relationship

If the wing–mass shift tracks thermal conditions, individuals sampled in warmer-than-usual years should be heavier or lighter for their wing length than birds caught in cooler years. We tested this with the local annual temperature anomaly at each capture (ΔT_{local} ; Methods).

We found no clear evidence for it. Under the fully adjusted set used for the temporal models, the isometry contrast was not associated with the annual anomaly ($\beta = +0.0123$ per $^{\circ}\text{C}$, 95% CrI: -0.0080 to $+0.0328$; $P(\beta > 0) = 0.88$), and neither were wing length ($\beta = -0.01$ mm per $^{\circ}\text{C}$, 95% CrI: -0.42 to $+0.39$) or log body mass ($\beta = +0.0052$ per $^{\circ}\text{C}$, 95% CrI: -0.0091 to $+0.0197$). The anomaly effect on the contrast was positive in every specification, but its interval always included zero, including with calendar year in the model ($\beta = +0.0094$, 95% CrI: -0.0102 to $+0.0290$), with a detrended anomaly, with a locality-year random intercept and with a baseline over every year from 1995 to 2018 (Methods; Supplementary Tables S3 and S6).

Anomalies of the year before capture, or of the two preceding years, were not associated with the contrast either (lag 1: $\beta = +0.0015$, 95% CI: -0.0191 to $+0.0222$; lags 1–2: $\beta =$

+0.0048, 95% CI: -0.0278 to +0.0375; REML, 50 trees), nor with wing length or body mass. We therefore make no claim that interannual thermal variation drives any change in the wing-mass relationship.

2.6 An apparent increase in bill width disappeared once contributors and municipalities were modelled

Raw species-level trajectories for bill width displayed both increases and decreases across species (Figure 2 C). In a baseline model without contributor controls, bill width appeared to increase (+0.49 mm/decade, 95% CrI: +0.12 to +0.86; $P(\beta > 0) > 0.99$, $n = 3,209$), as Allen-type expectations would predict.

However, when contributor and municipality were accounted for, this apparent increase disappeared (-0.06 mm/decade, 95% CrI: -0.37 to +0.26). In the fully adjusted model, bill width remained flat: -0.05 mm/decade (95% CrI: -0.36 to +0.27; $n = 3,205$). Temporal change within the same contributor was similarly flat (+0.04 mm/decade, 95% CrI: -0.38 to +0.47). The apparent baseline increase was therefore not robust to adjustment for contributor and spatial sampling structure, which were added together and whose contributions we cannot separate.

2.7 An apparent rise in variability disappeared once contributors and municipalities were modelled

In the baseline heteroscedastic model, the residual standard deviation of wing length appeared to increase by +9.8%/decade (95% CrI: +6.5 to +13.2%; Methods). With contributor and municipality intercepts in both the mean and the residual-SD models it was -6.5%/decade (95% CrI: -13.7 to +1.2%), and its sign depended on whether species intercepts also entered the residual-SD model (+3.5%/decade, 95% CI: -4.7 to +12.2%; REML check). Residual variance was not associated with local temperature (-0.5% per SD of temperature, 95% CrI: -8.7 to +8.4%).

2.8 Diet did not clearly explain which species changed

In the baseline model, wing-length trajectories appeared to depend on the share of invertebrates in the diet (EltonTraits; year $\times$ diet interaction $\beta = +0.66$ mm per SD of year per SD of diet, 95% CrI: +0.20 to +1.12). With contributor and municipality intercepts the interaction halved ($\beta = +0.32$, 95% CrI: -0.03 to +0.67), and comparing invertebrate-dominated with other species gave a similarly uncertain result ($\beta = +0.38$, 95% CrI: -0.36 to +1.14; Figure 6). If anything, species eating more fruit and seeds declined more steeply: wing length changed by -1.81 mm/decade (95% CrI: -3.04 to -0.59) at the 10th percentile of invertebrate share (10% invertebrates) and by -0.14

mm/decade (95% CrI: -1.36 to $+1.09$) at the 90th (100%). Because the interaction itself is uncertain, we cannot exclude a diet effect of moderate size in either direction.

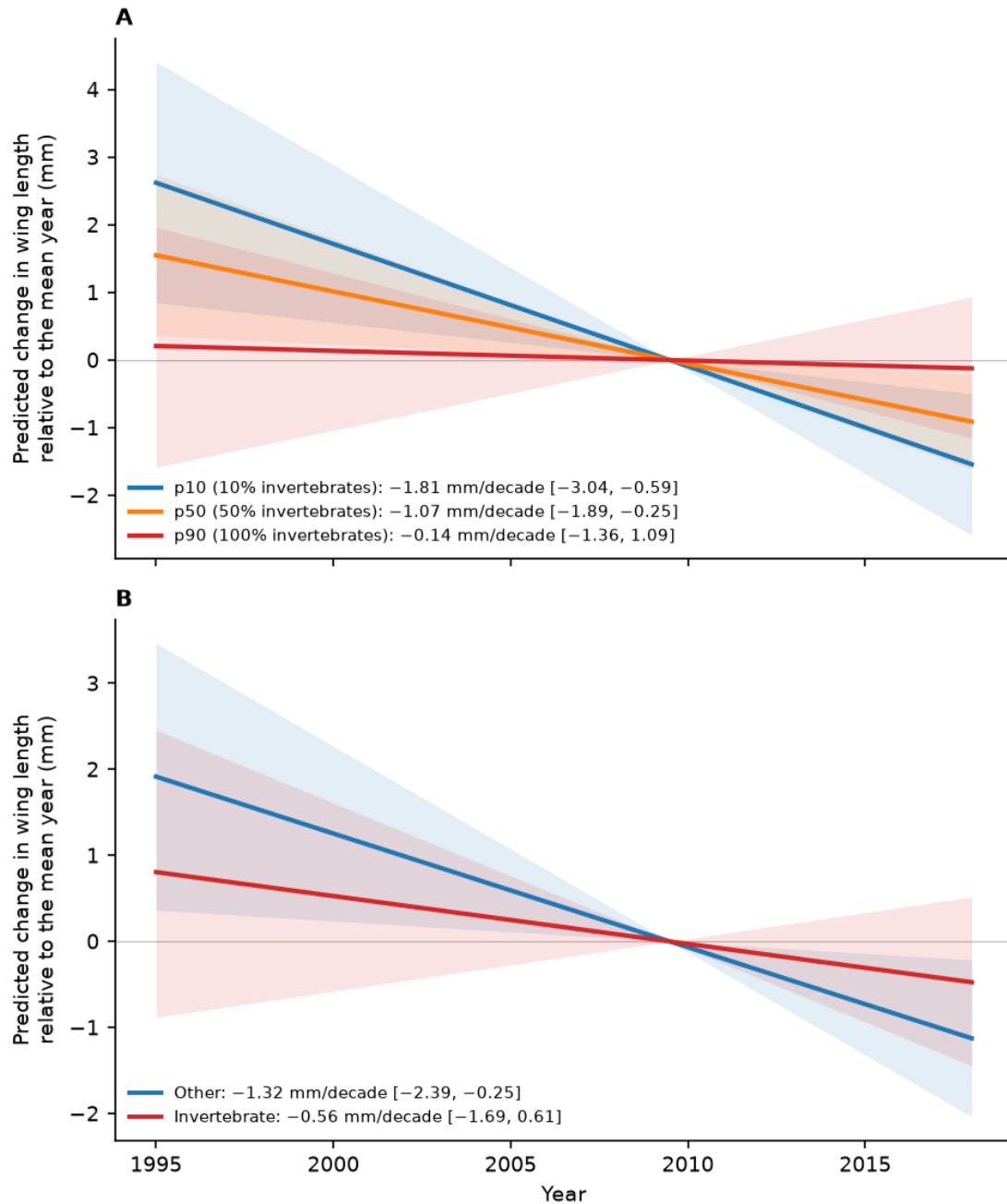


Figure 6: Decadal wing-length change as a function of diet (share of invertebrates in the diet). Both panels show the models with contributor and municipality intercepts; lines are the predicted change in wing length relative to the mean sampling year, computed from the posterior mean of the year slope, and shaded ribbons are the 95% credible intervals of that slope (posterior pooled over 20 phylogenetic trees), so they widen away from the mean year. Legend values give each slope in mm/decade with its 95% credible interval. (A) Continuous interaction model, evaluated at the 10th (p10, blue), 50th (p50, orange) and 90th (p90, red) species-weighted percentiles of invertebrate diet proportion. (B) Category model, for invertebrate-dominated and other species.

3. Discussion

Across 73 passerine species and 24 years of regional warming in the Brazilian Atlantic Forest, morphological change was weak and varied among species. Once turnover in contributors and sites was modelled, the average body-mass decline was close to the mean rate reported from the central Amazon, but its interval included zero and species did not decline together; wings did not lengthen, and bill width did not change. Much of the change suggested by unadjusted models reflected how the compilation was assembled, and simulations showed that modelling this structure recovers known trends rather than absorbing them.

3.1 Interspecific variation and the wing–mass relationship

The heterogeneity among species (Figure 4) has a practical implication: an assemblage-level trend near zero can conceal species changing in opposite directions, so community averages are a weak basis for judging whether, or how, a regional avifauna is responding to warming. Large multi-species analyses elsewhere likewise report species-specific responses, even where the average trend is clear (Teplitsky and Millien 2014; Youngflesh et al. 2022; Dubiner and Meiri 2022). Identifying which species change, and what they have in common, is therefore more informative than refining the assemblage mean, and will require longer species-level series than most contributors to this compilation provide. Diet was one candidate. In our data, diet did not clearly separate species that changed from those that did not, and if anything species feeding mostly on fruit and seeds declined in wing length more than strict insectivores. Species-level diet shares are coarse, however, and do not capture seasonal or local shifts in what birds eat, so a role for diet in structuring responses remains open.

The wing–mass result bears on what kind of change occurred. If birds had simply become smaller, mass would have fallen roughly three times faster than wing length; instead, mass tended to be retained relative to wing length, although strict isometry cannot be excluded. Even a clear departure in the contrast would reject only proportional co-variation of the two traits; it would not on its own establish a change in whole-body shape. This does not indicate a general reduction in structural size and, together with the uncertain change in mass, gives no sign of the slenderisation documented in the Amazon (Jirinec et al. 2021). One reason the two traits can change independently is that they measure different things. Wing length in live passerines is a measurement of the primary feathers (Jenni and Winkler 1989), which are renewed at each moult and whose growth responds to nutritional condition (Grubb 2006), and wing shape can change over decades as landscapes alter the demands placed on mobility (Desrochers 2010). Body mass, by contrast, combines structural size with short-

term condition. In a 70-year museum series, Dubiner and Meiri (2022) similarly found that most species changed in either mass or body length but seldom in both, whereas changes in wing length were few and inconsistent. Modest wing change without proportional mass change is therefore as compatible with shifts in feather growth, wear, or the condition of birds at capture as with a change in body size.

3.2 Contrasting macroecological dynamics: the Atlantic Forest and the central Amazon

On a common scale, the Atlantic Forest and the central Amazon differ less in the average rate of mass change than in its consistency and in wing length. Across the 68 passerines of Jirinec et al. (2021), the most directly comparable multi-decadal Neotropical benchmark, mass changed by -1.0% /decade on average (species range -1.9 to -0.1% ; Methods), close to our fully adjusted estimate, and per degree of local warming the two are also similar (Atlantic Forest -3.7% per $^{\circ}\text{C}$, 95% CrI: -11.3 to $+3.9$; central Amazon -5.0 to -3.1% per $^{\circ}\text{C}$). In the Amazon, however, all 77 species had lower mean mass by 2019 and wings lengthened by $+0.3\%$ /decade on average, so wing loading fell across the understorey community. In the Atlantic Forest, 36 of 62 species decreased in mass and 26 increased between the early and late periods (Figure 2 B), and the fully adjusted wing estimate (-1.0% /decade, 95% CrI: -2.2 to $+0.1$) has an upper bound that only just reaches the Amazonian mean. Our interval for body mass is wide, so the data cannot say whether the two regions differ in their average rate. Other tropical studies are closer to our result: in Panama, wing length changed weakly and in both directions (Wilcox et al. 2024), century-scale trends in lowland Colombia were mixed (Pérez-Amaya et al. 2025), and Afromontane understorey birds became larger (Neate-Clegg et al. 2024).

The Amazonian response is therefore one of several tropical outcomes. Ours is a contrast between two studies, not a controlled comparison: the two datasets differ in landscape, sampling design, temporal coverage, species composition and measurement structure, and part of the difference may lie in these features. Three ecological hypotheses, which our data cannot discriminate, could also contribute. First, the Amazonian study took place in continuous forest, where longer wings and lower wing loading reduce the cost of flight (Norberg 1990; Jirinec et al. 2021), whereas the Atlantic Forest survives as small fragments (Ribeiro et al. 2009). Wing morphology predicts movement across such landscapes (Sheard et al. 2020; Claramunt et al. 2022), so movement between fragments could favour different changes in wings and mass. Second, the Atlantic Forest spans steep gradients from sea level to over 2,000 m, where birds can track warming upslope (Freeman et al. 2018) or in buffered microhabitats (De Frenne et al. 2019; Scheffers et al. 2014), which would weaken thermoregulatory

pressure on body size. Third, birds in fragmented forest already experience wide, edge-driven thermal variation (Murcia 1995; Ewers and Banks-Leite 2013) and may tolerate warming without directional change. We did not measure wing shape, microclimate or edge exposure, so these hypotheses remain untested.

The contrast may also reflect study design: a single continuous 40-year ringing programme (BDFFP) versus a multi-site compilation assembled from many programmes. Our within-contributor decomposition does not rule this out. The within-contributor slope, which uses only temporal variation inside each research programme, was centred near zero for wing length (-0.3 mm/decade; complete cases: $+0.1$ mm/decade) and body mass (-0.6% /decade), but it was estimated imprecisely, and its interval for wing length also included the fully adjusted assemblage trend. Moreover, restricting the comparison to within a programme removes confounding from turnover among research programmes, but not change over time in the field workers, localities, or protocols operating within a programme. A sampling explanation for part of the contrast therefore remains open.

3.3 Evaluating ecogeographical rules in tropical endotherms

Our findings bear on the temporal extension of classic ecogeographical rules in tropical endotherms under climate change. Bergmann's and Allen's rules describe spatial patterns; the expectation that warming through time should drive reductions in body size, generalised from Bergmann's rule (Gardner et al. 2011; Sheridan and Bickford 2011), is a separate hypothesis, and a temporal null result bears on that hypothesis, not on the spatial rules themselves. In our 24-year dataset of 73 tropical passerine species, the average body-mass decline (-0.9% /decade) was similar to the Amazonian rate, but its interval spanned zero despite $+0.25^{\circ}\text{C}$ /decade of regional warming. The data therefore neither confirm nor exclude a warming-driven decline of the size reported from the Amazon, although the Atlantic Forest bird data make a decline steeper than about 3% per decade unlikely.

Several considerations explain why the decline hypothesis is a weak prediction over this time span, even under measurable warming. First, the warming recorded at our capture localities, about 0.56°C over the study period, is small relative to the thermal variation that the same species experience across their elevational and latitudinal ranges and between forest microhabitats (De Frenne et al. 2019; Scheffers et al. 2014), so any thermoregulatory advantage of a smaller body over this period would be slight. Second, Bergmann's rule is itself far from universal: it holds in a majority, but not all, of the bird and mammal taxa examined (Meiri and Dayan 2003), and reported recent changes in vertebrate body size are inconsistent in direction and prone to methodological artefacts (Yom-Tov and Geffen 2011). Where temporal declines in avian size have been detected,

they come from very large samples and vary among species (Youngflesh et al. 2022; Dubiner and Meiri 2022). Third, body size depends on food supply and conditions during development as well as on temperature (Teplitsky et al. 2008; Teplitsky and Millien 2014), and these influences need not act in the same direction as warming. Finally, a heritable response within 24 years would require strong and consistent selection, and evidence for genetic adaptation to climate change remains scarce relative to plastic responses (Merilä and Hendry 2014). These considerations do not turn the absence of a decline into a confirmed prediction. They explain why a 24-year record may not resolve a decline of this size, and our result bears on the temporal decline hypothesis, not on whether body size has a thermoregulatory role.

Similarly, Allen's rule posits that appendages enlarge in warmer environments to facilitate heat dissipation (Allen 1877; Ryding et al. 2021). While unadjusted models suggested a temporal increase in bill width, this apparent trend did not survive adjustment for contributor and municipality, and bill width showed no clear change. In tropical passerines, bill morphology is heavily constrained by foraging biomechanics, trophic specialisation, and vocal performance, which may override subtle thermoregulatory selection for larger bills (Podos 2001; Tattersall et al. 2017).

3.4 Sampling structure in compiled trait databases

Our study highlights the importance of modelling provenance and sampling structure when analysing multi-decadal compilations in macroecology. Large-scale compilations such as ATLANTIC BIRD TRAITS (Rodrigues et al. 2019) are indispensable for macroecological and macroevolutionary research, providing taxonomic and geographic scope that single-site studies cannot achieve. However, because compiled databases combine measurements across dozens of independent research teams, sampling localities, protocols, and seasons, apparent multi-decadal trends can be substantially confounded by sampling turnover.

In our dataset, unadjusted models suggested steep declines in wing length, increases in bill width, and expanding trait variability. Once sampling structure was modelled, the estimated wing decline fell from -1.8 to -0.7 mm/decade, the apparent bill expansion disappeared, and the apparent variance inflation was no longer supported. The size of the wing attenuation is itself uncertain: resampling contributors, the difference between the adjusted and unadjusted slopes ranged from -0.23 to $+2.53$ mm/decade (95% interval), although the adjusted slope was the shallower one in 90% of 200 resamples. Because contributor identity, sampling geography, protocol, and season are not fully separable in a compilation of this kind, we do not attribute this attenuation to any single source; what the comparison establishes is that the unadjusted estimates were substantially influenced by how the dataset was assembled. Simulations in the real

sampling structure show that the adjusted model recovers planted trends without absorbing them, and that the baseline model is biased when contributor effects drift with time. They cannot rule out turnover that the modelled grouping factors and covariates do not capture, so a residual sampling explanation remains possible. The main inference is therefore methodological: when trends are estimated from compiled trait databases, contributor, municipality and the other features of sampling need to be modelled explicitly, and unadjusted and adjusted estimates reported side by side. Trait compilations can therefore support macroecological inference, provided that sampling structure is modelled explicitly and its influence reported.

3.5 Conclusion

In summary, 24 years of warming in the Brazilian Atlantic Forest were accompanied by an average body-mass decline similar in rate to that of the central Amazon, but not by its consistency across species or by wing elongation. Once sampling structure was modelled, wing length and body mass both showed small declines (-1.0% and -0.9% per decade) whose credible intervals included zero, bill width did not change, and body mass tended to be retained relative to wing length, although strict geometric isometry could not be excluded. For body mass, the interval (-2.8 to $+1.0\%$ per decade) makes a sustained decline steeper than about 3% per decade unlikely, and our simulations show that a wing decline of about 1 mm per decade would usually have been detected. The difference from the Amazon raises the possibility that landscape fragmentation, topographic complexity and ecological context shape how consistently tropical birds respond to warming, a hypothesis that our study motivates but cannot itself test, and it may also reflect differences in study design. More generally, compiled trait databases can extend long-term morphological monitoring to under-studied tropical regions, but only when turnover in contributors and sites is modelled: ignoring it here would have more than doubled the estimated wing decline.

4. Materials and Methods

4.1 Trait database filtering and analytical sample derivation

We extracted avian morphological measurements from the ATLANTIC BIRD TRAITS database (Rodrigues et al. 2019), a comprehensive data compilation comprising 72,483 morphometric entries across the Brazilian Atlantic Forest domain, of which passerines represent 83% (60,066 records). To ensure that temporal analyses reflect genuine living phenotypic variation rather than methodological or preservation artefacts, we implemented a multi-stage filtering protocol.

First, we restricted the dataset strictly to live field captures, excluding museum study skins. Museum preparation induces post-mortem shrinkage in feathers, bill keratin, and skeletal dimensions (typically 1–3% reduction in linear dimensions upon drying; Weeks et al. (2020)). Because museum specimens in regional collections are heavily concentrated in earlier decades of the 20th century, retaining study skins introduces artificial positive temporal trends. Second, we restricted records to adult individuals (Age == "Adult"), excluding juveniles and fledglings whose flight feathers and bills are incompletely developed. Third, we filtered records spatially to sampling localities falling within a 20-km buffer of the Brazilian Atlantic Forest biome boundary (IBGE/MMA official shapefile), eliminating captures situated in adjacent Cerrado or Caatinga enclaves. The 20-km allowance is the tolerance already applied in the ATLANTIC BIRD TRAITS compilation itself, and it accommodates coordinate imprecision and the fact that the mapped biome boundary is a generalisation of a gradual ecotone; it is not a biological buffer and no result depends on its exact width. Fourth, we bounded the temporal window to 1995–2018 (24 calendar years), representing the era of active multi-project mist-netting in the biome with consistent capture documentation. We did not filter on migratory status: the compilation does not record it, and we therefore make no claim that the focal species are strictly resident. The sample comprises the passerines captured by Atlantic Forest mist-netting programmes, which may include partial or altitudinal migrants.

Finally, to guarantee robust parameter estimation for species-level random intercepts and random slopes, we restricted the primary analytical sample to species with ≥ 30 live-capture records and a temporal sampling span of ≥ 5 years. This yielded an analytical dataset of 73 passerine species and 12,571 live-capture records (8,478 wing-length records across 72 species; one species, *Herpsilochmus pectoralis*, was sampled exclusively for bill and tarsus traits and lacked wing and mass records in the compilation). Wing length was coalesced across database columns, prioritising right wing length when specified, followed by left wing length and unspecified wing length (conc.wing.length).

4.2 Contributor and municipality effects

Two features of a multi-source compilation can confound temporal trends: turnover in who supplied the records and in where they were taken. Differences between observers can be as large as the temporal changes of interest here (Yezerinac et al. 1992). The field-measurer column of ATLANTIC BIRD TRAITS is empty for every record, but the principal investigator (Main_researcher) is recorded for all of them (42 in the wing sample). We call this grouping the *contributor*: the research group or project that supplied the record. A contributor random effect absorbs whatever the records of one programme share, which includes measurement conventions but also protocols,

habitats, localities and species composition, so the attenuation of a trend after adding it does not isolate measurement error.

Both contributors and sites turned over. Only 6 of the 42 contributors sampled in both the early (1995–2006) and late (2013–2018) periods, the lower and upper quartiles of capture year, and 12 and 33 contributors were active in each (Supplementary Fig. S1); the middle half of records is left out of these descriptive comparisons only, and all models use every year. Sampling also moved north (mean latitude -24.8° to -21.9°) and upslope (median altitude 175 m to 489 m), along gradients on which body size often varies. All models therefore included standardised latitude, and models beyond the baseline added longitude and elevation.

Every model beyond the baseline included random intercepts for contributor and municipality. The two are correlated, because projects usually worked in particular municipalities: 85.9% of municipalities in the shared sample were sampled by a single contributor. They are nonetheless partly crossed (14% of municipalities in the wing dataset had more than one contributor, and 64% of contributors worked in more than one municipality), so their variances can be estimated separately, and both were clearly non-zero (contributor SD 3.47 mm; municipality SD 1.52 mm). The municipality term matters for the temporal inference: municipality alone accounts for 74% of the variance in capture year in the wing dataset, against 58% for contributor, so without it differences among sites sampled early and late can be read as temporal change.

4.3 Model sequence and within-between decomposition

To evaluate how dataset structure influences temporal trends, we fitted a ladder of increasingly controlled mixed-effects models. We refer to each model by a descriptive name; the short codes used in the analysis scripts and in Supplementary Tables S1 and S2 are given in parentheses.

- **Baseline** (M0): replicating published unadjusted model structures, with sex, calendar year (scaled_yr, centred and divided by its standard deviation, SD = 5.02 years), latitude (scaled_lat), species-specific random intercepts and year slopes, and phylogenetic covariance among the species intercepts.
- **+ contributor & municipality** (M1): adding random intercepts for contributor identity (1 | Main_researcher), to absorb systematic differences among research programmes, including but not limited to measurement conventions, and for municipality (1 | Municipality), to absorb local differences among sampled sites.
- **+ wing side, longitude & elevation** (M2): adding fixed effects for the protocol proxy (wing_col: right wing, left wing, or side unspecified; Supplementary Fig.

S2), standardised longitude (*scaled_long*), and standardised elevation (*scaled_elev*).

- **Fully adjusted** (M3): adding fixed effects for seasonal quarter (DJF: Dec–Feb; MAM: Mar–May; JJA: Jun–Aug; SON: Sep–Nov) to account for feather wear across the annual cycle, active primary moult status (*Molt*), and a random intercept for the individual bird ($1 \mid \text{ID_individual}$) (*Ring* $\times$ *Species*) to account for repeated captures of banded individuals.
- **Year within vs between contributors** (M5src; Mundlak within-between decomposition): to separate temporal change occurring within a single contributor from compositional turnover across contributors, calendar year was partitioned into the contributor mean capture year ($\bar{Y}_j = \text{yr_src_mean}$) and individual capture deviations from that contributor mean ($Y_{ij} - \bar{Y}_j = \text{yr_within_src}$) (Mundlak 1978; Bell et al. 2019), on the fully adjusted specification. The within-contributor coefficient β_{within} measures whether birds contributed by the *same* research programme changed over time, and is therefore not driven by turnover in which programmes were active. It does not, however, remove change over time in the individual field workers, localities, or protocols operating within a programme.
- **Robustness and sensitivity models** (Supplementary Note S5): the fully adjusted model on complete cases (M3_cc), because the primary sample keeps records with an imputed altitude or an unrecorded season or moult as explicit levels; on first captures only (M7); and without 6 anomalous contributor $\times$ municipality $\times$ year blocks (193 records; M3_noanom). That screen flagged blocks of at least 10 records whose mean deviation from species $\times$ sex means exceeded 6 mm in either direction; we set the threshold after inspecting the data, so it is a post hoc sensitivity check, not part of the primary analysis. We also decomposed year within and between municipalities (M5).

4.4 Bayesian phylogenetic mixed models across 20 topologies

Species names were matched to the tips of the phylogeny of McTavish et al. (2025) with the matching cascade of *prepR4pcm* (Nakagawa et al. 2026), and all 73 species were placed. To carry phylogenetic uncertainty into every estimate, we drew 20 trees from its posterior distribution and computed the phylogenetic correlation matrix *A* of each under Brownian motion.

All models were Gaussian linear mixed models fitted with a purpose-written Stan program (Carpenter et al. 2017). Each contained the fixed effects of its specification, species random intercepts and year slopes, and the random intercepts of its specification for contributor, municipality and individual bird. The species intercepts

had a phylogenetic and a non-phylogenetic component, H^2 being the phylogenetic share of their variance; species year slopes were independent among species. Each model was sampled on each tree, and the draws from the 20 trees were pooled, which gives the mixture posterior over trees (Nakagawa and Villemereuil 2019). No sampler divergences occurred in any of the 40 models, and the largest rank-normalised $\hat{R}$ of any quoted parameter on any tree was 1.010 (Vehtari et al. 2021). The likelihood, priors, sampler settings and diagnostics are given in Supplementary Note S5.

We report posterior means with 95% equal-tailed credible intervals (CrI) and, for temporal effects, the posterior probability that the effect is negative, $P(\beta < 0)$: the share of posterior draws below zero, which is a direct statement of how probable a decline is given the model and data. Values near 0.5 indicate no evidence for either direction.

As an independent check, every model except the detrended-anomaly and locality-year checks was also fitted by restricted maximum likelihood with the propto covariance structure of `glmmTMB` v1.1.14 (Williams et al. 2025) and pooled across 50 trees with Rubin’s rules (Nakagawa and Villemereuil 2019; Nakagawa et al. 2026). For the temporal coefficients, the two engines differed by at most 0.09 posterior standard deviations (Supplementary Tables S1 and S2; Supplementary Note S5).

4.5 Body mass and bill width modelling and testing Allen’s rule

To evaluate whether temporal shifts occurred in body mass and extremity dimensions, we applied the sequential control ladder to log-transformed body mass ($n = 11,077$) and bill width ($n = 3,205$). Bill width is the appendage measure used here because it is the extremity trait with the broadest coverage in the compilation and because bill surface is a major avian site of non-evaporative heat exchange, which is the dimension Allen-type predictions concern. It is, however, a single and imperfect proxy: tarsus length is recorded for fewer individuals, and we did not analyse bill depth or surface area, so our test addresses one extremity dimension rather than heat-dissipating surface as a whole. The fully adjusted models incorporated sex, latitude, longitude, elevation, season, species-specific random intercepts (with phylogenetic covariance) and year slopes, and random intercepts for contributor, municipality and individual bird identity; moult is not recorded for enough of these records to be included.

To evaluate Allen’s rule (Allen 1877; Ryding et al. 2021), which predicts that extremities enlarge in warmer conditions to facilitate non-evaporative heat dissipation, we tested whether bill width exhibited a positive temporal slope under contributor and spatial controls.

4.6 Direct test of geometric isometry

To evaluate whether body shape changed over time, we analysed records with simultaneous measurements of wing length and body mass on the same individual bird. Under geometric isometry, body mass scales with the cube of linear dimensions ($M \propto L^3$), so a bird shrinking uniformly without changing shape loses mass at about three times the proportional rate of its linear shrinkage ($\% \Delta M \approx 3 \times \% \Delta L$).

We tested departure from isometry by fitting the log-scale contrast $\Delta_{\text{iso}} = \ln M - 3 \ln L$ directly as the response. Its year slope is zero under strict isometry whatever the separate trends of the two traits, and fitting it directly yields its posterior without depending on the covariance between two separately estimated slopes. A positive slope indicates that body mass declined less than cubic scaling of the wing predicts (mass conservation or stoutening); a negative slope indicates disproportionate mass loss or wing elongation (slenderisation). The contrast was fitted under two adjustment sets: a parsimonious set (sex, year, latitude, species intercepts with phylogenetic covariance and species year slopes, and random intercepts for contributor and municipality), and the fully adjusted set, which adds longitude, elevation, season, moult and individual identity. To show the two sides of the contrast, $\ln L$ and $\ln M$ were also fitted separately under the fully adjusted set. Year slopes on log scales are expressed as percentage change per decade, $100 [\exp(10\beta/SD_{\text{year}}) - 1]$. These models use the 7,409 of the 7,577 shared records with longitude, elevation and season recorded. They do not require the local temperature anomaly, which is available for a subset of 7,070 records (next section); on that subset the isometry contrast gave a similar estimate (+2.3%/decade, 95% CrI: -0.5 to +5.2%).

4.7 Interannual temperature anomalies

Each record was linked to the mean annual temperature at its coordinates in its capture year, from the WorldClim historical monthly weather series (monthly tmin/tmax, CRU-TS 4.09 downscaled with WorldClim 2.1, 2.5 arc-min, Harris et al. 2020; Fick and Hijmans 2017), not the WorldClim climate normals. The local anomaly (ΔT_{local}) is that value minus its mean over all records at the locality, that is, over the years in which the locality was sampled, weighted by captures. This baseline removes most of the regional warming trend: the anomaly correlates with capture year at 0.07, against 0.39 for a baseline over every year from 1995 to 2018. The land mask leaves no value at 98 of the 455 coordinate localities, which lie about 10 km from the nearest valid cell and are not gap-filled. Of the 7,577 shared records, 353 from 11 such localities have no climate value; these are coastal (median altitude 35 m against 268 m) and come from 6 contributors, so they are not a random subset. A further 154 records lacked longitude, altitude or season,

leaving 7,070 records for the thermal models. Regional climate trends were estimated from the locality-by-year values with linear mixed models (Supplementary Note S2).

We related wing length, log body mass and the isometry contrast to ΔT_{local} under the fully adjusted set of the temporal models, with species anomaly slopes. Because an anomaly measured over a warming period correlates with year, and because birds caught at one locality in one year share their anomaly, we also fitted the isometry model with calendar year added, with a detrended anomaly, with a locality-year random intercept and with the full-series baseline. The thermal models were also refitted with the anomaly of the year before capture and the mean of the two years before (REML across 50 trees; Supplementary Note S4 and Tables S3 and S6). The birds are adults of unknown age, so these lags do not identify a developmental window, and for wing length the relevant exposure, conditions during the last moult, cannot be resolved.

4.8 Phenotypic variability: heteroscedastic models

Residual phenotypic variability was modelled with continuous heteroscedastic models, which estimate temporal change in the residual standard deviation from the full continuous time series while adjusting for covariates. Alongside the model for the mean, the residual standard deviation σ_i of each record had its own log-linear model:

$$\log(\sigma_i) = \gamma_0 + \gamma_1 t_i + \gamma_2 \text{Sex}_i [+ f_{k[i]} + g_{j[i]}]$$

where γ_1 is the change in log residual SD per SD of year, reported as percentage change per decade. In the baseline version the mean model contained sex, year, latitude and the species terms; in the adjusted version both the mean model and the residual-SD model received random intercepts for contributor (f_k) and municipality (g_j), and the mean model also longitude, elevation and season. A third version added standardised local mean annual temperature (`scaled_tmean`) to the residual-SD model to test whether residual variance tracked thermal conditions. Individual identity is not included in these models, so the residual SD is record-level: it combines among-individual differences, within-individual variation across recaptures (the wing sample includes 373 birds captured more than once), and measurement error, and should not be read as among-individual variability alone.

4.9 Diet

Species' diet was described by the share of invertebrates in the diet, extracted from the EltonTraits 1.0 database (Wilman et al. 2014) (`Diet-Inv`, % of diet consisting of invertebrates), which places species along a gradient from mostly plant-based diets (fruit, seeds, nectar) to strict insectivory. In continuous interaction models, `Diet-Inv` was standardised (z-score) and interacted with calendar year (`scaled_yr *`

diet_inv_std). Wing-length trajectories were projected across the 10th (p10), 50th (p50), and 90th (p90) species-weighted percentiles of invertebrate share. In categorical models, diet was collapsed into Invertebrate-dominated ($\geq 50\%$ invertebrate diet) versus Other (frugivores, nectarivores, granivores, omnivores). Both were fitted as a baseline (sex, latitude, species terms) and with random intercepts for contributor and municipality.

4.10 Recovery simulation

To test whether the fully adjusted model can recover a trend in this sampling structure, we simulated wing lengths on the real wing sample, keeping every record's species, sex, individual, contributor, municipality, year and covariates. A generating model with the M3 structure was fitted to the real data (lme4 REML, without the phylogenetic term, which structures species intercepts rather than the year slope). Each simulated response combined its fixed effects, with the year slope replaced by a planted value of 0, -0.5 or -1.0 mm/decade, with fresh species, individual and residual draws at the fitted variances. Contributor and municipality effects followed one of three scenarios: fresh draws independent of year (clean); the fitted conditional modes, which keep their real association with year (real); or fresh draws plus a contributor shift proportional to the contributor's mean capture year, scaled so that the shift alone reproduced the gap between the real baseline and fully adjusted estimates (artefact). Each of the 200 data sets per scenario and planted value was fitted with M0, M1, M3 and the within-contributor model, and we recorded bias, root-mean-square error, coverage of the 95% interval and the share of intervals excluding zero (Supplementary Table S4). To put an interval on the attenuation of the wing trend, we also resampled contributors with replacement (200 draws, all records of a drawn contributor kept and relabelled) and refitted M0 and M3 to each draw.

4.11 Comparison with the central Amazon

To put our estimates on the same scale as those of Jirinec et al. (2021), we used their supplementary Table S6, which gives each species' model-estimated mean mass and wing length in 1980 and 2019 and the change between them. The Amazonian benchmark is the mean change across the 68 passerine species (for body mass, the 77 species including non-passerines give the same value to one decimal place), divided by 3.9 decades. Per- $^{\circ}\text{C}$ rates divide this benchmark by the warming at their site, which the authors report as $+1.00^{\circ}\text{C}$ in the wet season and $+1.65^{\circ}\text{C}$ in the dry season since 1966; our per- $^{\circ}\text{C}$ rates divide the fully adjusted rates by the regional warming rate (Results). Because the benchmark averages species-level estimates and the two studies differ in design, the comparison is descriptive.

4.12 Use of AI tools

Generative AI tools (Anthropic Claude, accessed through Claude Code) were used during the analysis and preparation of this manuscript. Their use covered writing and revising analysis scripts in R, executing those scripts, drafting and revising manuscript text, and auditing the analysis code for reproducibility. No AI tool was used to generate, alter or impute data, or to produce figures from anything other than the analysis outputs described above. All analyses were re-executed from the archived scripts and all reported values are read programmatically from the saved model objects. All authors checked the final text and take responsibility for its content.

4.13 Reproducible software environment

The Bayesian models were fitted on Linux servers in R 4.5.3 (R 4.6.1 for the all-record isometry models and the trees refitted with longer chains) with CmdStan 2.39.0 through `cmdstanr` (v0.9.0), using the Stan program archived as `Analysis/scripts/stan/phylo_lmm_marginal.stan`. All other analyses and manuscript generation were conducted in R 4.6.0 with `glmmTMB` (v1.1.14), `prepR4pcm` (v0.5.0.9000), `metafor` (v5.0-1), `lme4` (v2.0-1), `ape` (v5.8-1), `phytools` (v2.5-2), `sf` (v1.1-2) and `terra` (v1.9-27), alongside Python (3.11 or later) with `matplotlib` and `pandas` for the figures. The complete environment, giving the exact version and source of every package, is recorded in `renv.lock` at the root of the archived repository and can be restored with `renv::restore()`; `Analysis/output/session_info.txt` gives the same information in human-readable form. Each saved `glmmTMB` result object additionally records the R and engine version under which it was fitted. All analysis scripts, curated datasets and the pooled posterior summaries of every fitted model are archived, with the outputs needed to reproduce every reported value, in the project repository (Zenodo, DOI: [10.5281/zenodo.23087760](https://doi.org/10.5281/zenodo.23087760)).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

E.S.A.S. conceived the study, curated the data, designed and performed the analyses, produced the figures, and wrote the original draft. A.M., S.O. and G.B. contributed to the methodology, validated the analyses, and reviewed and edited the manuscript. All authors approved the final version.

Funding

G.B. was funded by the European Union's Horizon research and innovation programme under the Marie Skłodowska-Curie grant CAUSATION (Grant Agreement ID: 101150734). A.M. was supported by a Japan Society for the Promotion of Science (JSPS) Overseas Research Fellowship. E.S.A.S. and S.O. were supported by a Canada Excellence Research Chair (CERC-2022-00074).

Data availability statement

All primary data analysed here are from the previously published, openly available compilations of morphological records from ATLANTIC BIRD TRAITS (Rodrigues et al. 2019). Climate data are from the WorldClim historical monthly weather series (CRU-TS 4.09 downscaled, Harris et al. 2020; Fick and Hijmans 2017). Phylogenies were drawn from the posterior distribution of McTavish et al. (2025). Dietary data are from EltonTraits 1.0 (Wilman et al. 2014).

Analysis code, the derived analytical datasets, and the pooled posterior summaries needed to reproduce every value reported here are archived at Zenodo (<https://doi.org/10.5281/zenodo.23087760>).

Ethics

This study is a secondary analysis of previously published, openly available records. No new animal capture or handling was undertaken. Capture and ringing permits for the primary data are those held by the contributing research programmes compiled in ATLANTIC BIRD TRAITS (Rodrigues et al. 2019).

Acknowledgements

We thank João C.T. Menezes for contributions to the initial design of this study and for comments on the manuscript. We also thank Guilherme R. R. Brito for his comments about maturation time for the temporal lag analysis. We thank Shinichi Nakagawa for thoughtful comments on a previous version of this manuscript. We are grateful to the many field researchers and research groups whose long-term mist-netting efforts are compiled in the ATLANTIC BIRD TRAITS database, without which this work would not have been possible.

References

- Allen, Joel Asaph. 1877. "The Influence of Physical Conditions in the Genesis of Species." *Radical Review* 1: 108–40.
- Ashton, Kyle G. 2002. "Patterns of Within-Species Body Size Variation of Birds: Strong Evidence for Bergmann's Rule." *Global Ecology and Biogeography* 11 (6): 505–23. <https://doi.org/10.1046/j.1466-822X.2002.00313.x>.
- Badyaev, Alexander V. 2005. "Stress-Induced Variation in Evolution: From Behavioural Plasticity to Genetic Assimilation." *Proceedings of the Royal Society B: Biological Sciences* 272 (1566): 877–86. <https://doi.org/10.1098/rspb.2004.3045>.
- Bell, Andrew, Malcolm Fairbrother, and Kelvyn Jones. 2019. "Fixed and Random Effects Models: Making an Informed Choice." *Quality & Quantity* 53 (2): 1051–74. <https://doi.org/10.1007/s11135-018-0802-x>.
- Bergmann, Carl. 1847. "Über Die Verhältnisse Der Wärmeökonomie Der Thiere Zu Ihrer Grösse." *Göttinger Studien* 3: 595–708. <https://www.digitale-sammlungen.de/en/view/bsb10306637?page=1>.
- Carpenter, Bob, Andrew Gelman, Matthew D. Hoffman, et al. 2017. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 76 (1): 1–32. <https://doi.org/10.18637/jss.v076.i01>.
- Claramunt, Santiago, Milly Hong, and Adriana Bravo. 2022. "The Effect of Flight Efficiency on Gap-Crossing Ability in Amazonian Forest Birds." *Biotropica* 54 (4): 860–68. <https://doi.org/10.1111/btp.13109>.
- De Frenne, Pieter, Florian Zellweger, Francisco Rodríguez-Sánchez, et al. 2019. "Global Buffering of Temperatures Under Forest Canopies." *Nature Ecology & Evolution* 3: 744–49. <https://doi.org/10.1038/s41559-019-0842-1>.
- Debat, Vincent, and Patrice David. 2001. "Mapping Phenotypes: Canalization, Plasticity and Developmental Stability." *Trends in Ecology & Evolution* 16 (10): 555–61. [https://doi.org/10.1016/S0169-5347\(01\)02266-2](https://doi.org/10.1016/S0169-5347(01)02266-2).
- Desrochers, André. 2010. "Morphological Response of Songbirds to 100 Years of Landscape Change in North America." *Ecology* 91 (6): 1577–82. <https://doi.org/10.1890/09-2202.1>.
- Dubiner, Shahar, and Shai Meiri. 2022. "Widespread Recent Changes in Morphology of Old World Birds, Global Warming the Immediate Suspect." *Global Ecology and Biogeography* 31 (4): 791–801. <https://doi.org/10.1111/geb.13474>.

- Ewers, Robert M., and Cristina Banks-Leite. 2013. "Fragmentation Impairs the Microclimate Buffering Effect of Tropical Forests." *PLOS ONE* 8 (3): e58093. <https://doi.org/10.1371/journal.pone.0058093>.
- Fick, Stephen E., and Robert J. Hijmans. 2017. "WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37 (12): 4302–15. <https://doi.org/10.1002/joc.5086>.
- Freeman, Benjamin G., Micah N. Scholer, Viviana Ruiz-Gutierrez, and John W. Fitzpatrick. 2018. "Climate Change Causes Upslope Shifts and Mountaintop Extirpations in a Tropical Bird Community." *Proceedings of the National Academy of Sciences* 115: 11982–87. <https://doi.org/10.1073/pnas.1804224115>.
- Gardner, Janet L., Anne Peters, Michael R. Kearney, Leo Joseph, and Robert Heinsohn. 2011. "Declining Body Size: A Third Universal Response to Warming?" *Trends in Ecology & Evolution* 26 (6): 285–91. <https://doi.org/10.1016/j.tree.2011.03.005>.
- Grubb, Thomas C. 2006. "Can Reduced Nutritional Condition Really Cause Narrowed Growth Bars?" In *Ptilochronology*. Oxford University Press. <https://doi.org/10.1093/oso/9780199295500.003.0003>.
- Harris, Ian, Timothy J. Osborn, Phil Jones, and David Lister. 2020. "Version 4 of the CRU TS Monthly High-Resolution Gridded Multivariate Climate Dataset." *Scientific Data* 7 (1): 109. <https://doi.org/10.1038/s41597-020-0453-3>.
- Jenni, Lukas, and Rafael Winkler. 1989. "The Feather-Length of Small Passerines: A Measurement for Wing-Length in Live Birds and Museum Skins." *Bird Study* 36 (1): 1–15. <https://doi.org/10.1080/00063658909476996>.
- Jirinec, Vitek, Ryan C. Burner, Bruna R. Amaral, et al. 2021. "Morphological Consequences of Climate Change for Resident Birds in Intact Amazonian Rainforest." *Science Advances* 7 (46): eabk1743. <https://doi.org/10.1126/sciadv.abk1743>.
- Mayr, Ernst. 1956. "Geographical Character Gradients and Climatic Adaptation." *Evolution* 10 (1): 105–8. <https://doi.org/10.2307/2406103>.
- McTavish, Emily Jane, Jeff A. Gerbracht, Mark T. Holder, et al. 2025. "A Complete and Dynamic Tree of Birds." *Proceedings of the National Academy of Sciences* 122 (18): e2409658122. <https://doi.org/10.1073/pnas.2409658122>.
- Meiri, Shai, and Tamar Dayan. 2003. "On the Validity of Bergmann's Rule." *Journal of Biogeography* 30: 331–51. <https://doi.org/10.1046/j.1365-2699.2003.00837.x>.

- Merilä, Juha, and Andrew P. Hendry. 2014. "Climate Change, Adaptation, and Phenotypic Plasticity: The Problem and the Evidence." *Evolutionary Applications* 7: 1–14. <https://doi.org/10.1111/eva.12137>.
- Mundlak, Yair. 1978. "On the Pooling of Time Series and Cross Section Data." *Econometrica* 46 (1): 69–85. <https://doi.org/10.2307/1913646>.
- Murcia, Carolina. 1995. "Edge Effects in Fragmented Forests: Implications for Conservation." *Trends in Ecology & Evolution* 10 (2): 58–62. [https://doi.org/10.1016/s0169-5347\(00\)88977-6](https://doi.org/10.1016/s0169-5347(00)88977-6).
- Nakagawa, Shinichi, Santiago Ortega, Ayumi Mizuno, et al. 2026. *prepR4pcm: Prepare Data and Trees for Phylogenetic Comparative Methods*. <https://github.com/itchyshin/prepR4pcm>.
- Nakagawa, Shinichi, and Pierre de Villemereuil. 2019. "A General Method for Simultaneously Accounting for Phylogenetic and Species Sampling Uncertainty via Rubin's Rules in Comparative Analysis." *Systematic Biology* 68 (4): 632–41. <https://doi.org/10.1093/sysbio/syy089>.
- Neate-Clegg, Montague H. C., Morgan W. Tingley, and William D. Newmark. 2024. "Afro-montane Understory Birds Increase in Body Size over Four Decades." *Ecography* 2024: e07106. <https://doi.org/10.1111/ecog.07106>.
- Norberg, Ulla M. 1990. *Vertebrate Flight: Mechanics, Physiology, Morphology, Ecology and Evolution*. Springer-Verlag. <https://doi.org/10.1007/978-3-642-83848-4>.
- Pérez-Amaya, Natalia, Montague H. C. Neate-Clegg, Natalia Ocampo-Peñuela, Carlos Daniel Cadena, Juliana Soto-Patiño, and Andrés M. Cuervo. 2025. "Century-Long Shifts in Body Size and Proportions of Lowland Rainforest Birds." *Evolutionary Ecology* 39: 475–94. <https://doi.org/10.1007/s10682-025-10343-6>.
- Podos, Jeffrey. 2001. "Correlated Evolution of Morphology and Vocal Signal Structure in Darwin's Finches." *Nature* 409: 185–88. <https://doi.org/10.1038/35051570>.
- Radchuk, Viktoriia, Thomas Reed, Céline Teplitsky, et al. 2019. "Adaptive Responses of Animals to Climate Change Are Most Likely Insufficient." *Nature Communications* 10 (1): 3109. <https://doi.org/10.1038/s41467-019-10924-4>.
- Ribeiro, Milton Cezar, Jean Paul Metzger, Alexandre Camargo Martensen, Flavio Jorge Ponzoni, and Marcia Makiko Hirota. 2009. "The Brazilian Atlantic Forest: How Much Is Left, and How Is the Remaining Forest Distributed? Implications for Conservation." *Biological Conservation* 142 (6): 1141–53. <https://doi.org/10.1016/j.biocon.2009.02.021>.

- Rodrigues, Rodolpho Credo, Érica Hasui, Julia Camara Assis, et al. 2019. "ATLANTIC BIRD TRAITS: A Data Set of Bird Morphological Traits from the Atlantic Forests of South America." *Ecology* 100 (6): e02647. <https://doi.org/10.1002/ecy.2647>.
- Ryding, Sara, Marcel Klaassen, Glenn J. Tattersall, Janet L. Gardner, and Matthew R. E. Symonds. 2021. "Shape-Shifting: Changing Animal Morphologies as a Response to Climatic Warming." *Trends in Ecology & Evolution* 36 (11): 1036–48. <https://doi.org/10.1016/j.tree.2021.07.006>.
- Scheffers, Brett R., David P. Edwards, Arvin Diesmos, Stephen E. Williams, and Theodore A. Evans. 2014. "Microhabitats Reduce Animal's Exposure to Climate Extremes." *Global Change Biology* 20: 495–503. <https://doi.org/10.1111/gcb.12439>.
- Schmidt, Niels, and Per Jensen. 2005. "Concomitant Patterns in Avian and Mammalian Body Length Changes in Denmark." *Ecology and Society* 10 (2). <https://doi.org/10.5751/ES-01372-100205>.
- Sheard, Catherine, Montague H. C. Neate-Clegg, Nico Alioravainen, et al. 2020. "Ecological Drivers of Global Gradients in Avian Dispersal Inferred from Wing Morphology." *Nature Communications* 11: 2463. <https://doi.org/10.1038/s41467-020-16313-6>.
- Sheridan, Jennifer A., and David Bickford. 2011. "Shrinking Body Size as an Ecological Response to Climate Change." *Nature Climate Change* 1 (8): 401–6. <https://doi.org/10.1038/nclimate1259>.
- Tattersall, Glenn J., Bassel Arnaout, and Matthew R. E. Symonds. 2017. "The Evolution of the Avian Bill as a Thermoregulatory Organ." *Biological Reviews* 92 (3): 1630–56. <https://doi.org/10.1111/brv.12299>.
- Teplitsky, Celine, and Virginie Millien. 2014. "Climate Warming and Bergmann's Rule Through Time: Is There Any Evidence?" *Evolutionary Applications* 7 (1): 156–68. <https://doi.org/10.1111/eva.12129>.
- Teplitsky, Céline, James A. Mills, Jussi S. Alho, John W. Yarrall, and Juha Merilä. 2008. "Bergmann's Rule and Climate Change Revisited: Disentangling Environmental and Genetic Responses in a Wild Bird Population." *Proceedings of the National Academy of Sciences* 105 (36): 13492–96. <https://doi.org/10.1073/pnas.0800999105>.
- Vehtari, Aki, Andrew Gelman, Daniel Simpson, Bob Carpenter, and Paul-Christian Bürkner. 2021. "Rank-Normalization, Folding, and Localization: An Improved $\hat{R}$ for Assessing Convergence of MCMC (with Discussion)." *Bayesian Analysis* 16 (2): 667–718. <https://doi.org/10.1214/20-BA1221>.

Weeks, Brian C., David E. Willard, Marketa Zimova, et al. 2020. "Shared morphological consequences of global warming in North American migratory birds." *Ecology Letters* 23 (2): 316–25. <https://doi.org/10.1111/ele.13434>.

Wilcox, Rebecca C., Thomas J. Benson, Jeffrey D. Brawn, and Corey E. Tarwater. 2024. "Observed Declines in Body Size Have Differential Effects on Survival and Recruitment, but No Effect on Population Growth in Tropical Birds." *Global Change Biology* 30: e17455. <https://doi.org/10.1111/gcb.17455>.

Williams, Coralie, Maeve McGillicuddy, Szymon M. Drobniak, Benjamin M. Bolker, David I. Warton, and Shinichi Nakagawa. 2025. "Fast Phylogenetic Generalised Linear Mixed-Effects Modelling Using the glmmTMB R Package." *bioRxiv*, ahead of print. <https://doi.org/10.64898/2025.12.20.695312>.

Wilman, Hamish, Jonathan Belmaker, Jennifer Simpson, Carolina de la Rosa, Marcelo M Rivadeneira, and Walter Jetz. 2014. "EltonTraits 1.0: Species-Level Foraging Attributes of the World's Birds and Mammals: Ecological Archives E095-178." *Ecology* 95 (7): 2027–27. <https://doi.org/10.1890/13-1917.1>.

Yezerinac, Stephen M., Stephen C. Loughheed, and Paul Handford. 1992. "Measurement Error and Morphometric Studies: Statistical Power and Observer Experience." *Systematic Biology* 41 (4): 471–82. <https://doi.org/10.1093/sysbio/41.4.471>.

Yom-Tov, Yoram, and Eli Geffen. 2011. "Recent Spatial and Temporal Changes in Body Size of Terrestrial Vertebrates: Probable Causes and Pitfalls." *Biological Reviews* 86 (2): 531–41. <https://doi.org/10.1111/j.1469-185x.2010.00168.x>.

Youngflesh, Casey, James F. Saracco, Rodney B. Siegel, and Morgan W. Tingley. 2022. "Abiotic Conditions Shape Spatial and Temporal Morphological Variation in North American Birds." *Nature Ecology & Evolution* 6: 1860–70. <https://doi.org/10.1038/s41559-022-01893-x>.

Supplementary Material: Weak and species-specific morphological change in Atlantic Forest birds across two decades of warming

1. Supplementary Notes

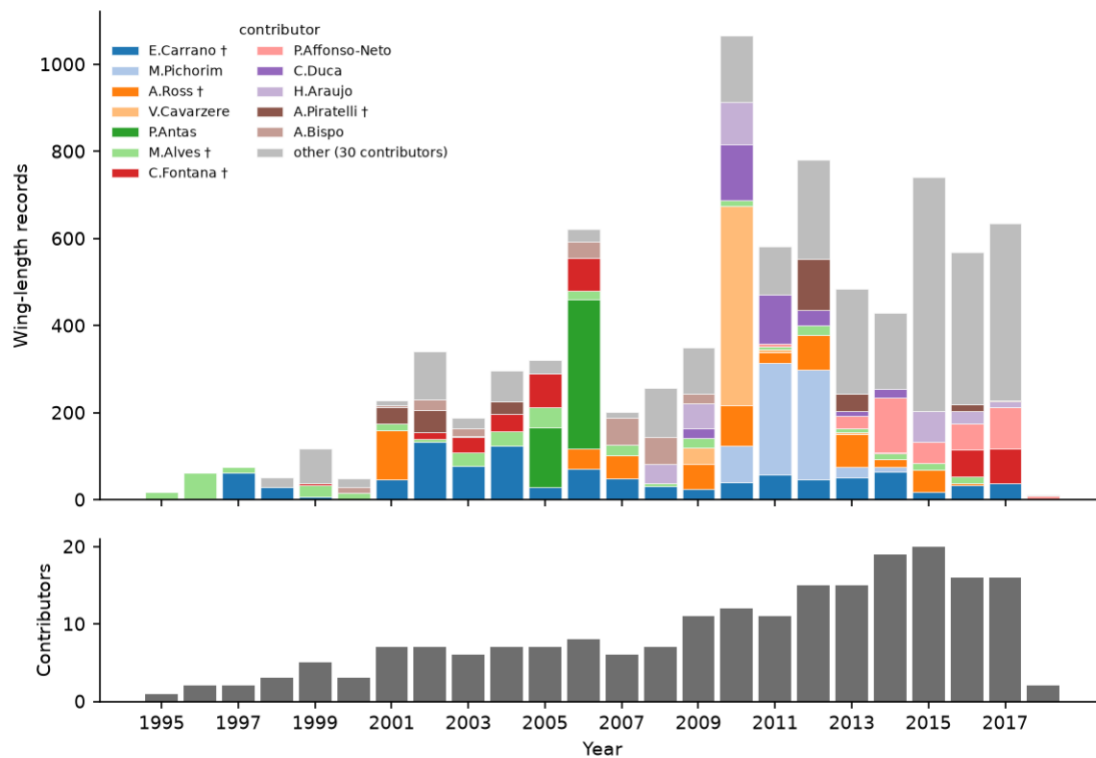
1.1 Supplementary Note S1: Macroecological Provenance and Contributor Turnover

Large-scale biodiversity trait compilations aggregate measurements across numerous independent research initiatives spanning wide geographical territories and multiple decades (Rodrigues et al. 2019). In our analytical sample from the ATLANTIC BIRD TRAITS database, 47 contributors (research groups, identified by their principal investigator) carried out field mist-netting in 153 municipalities between 1995 and 2018 (42 contributors and 130 municipalities in the wing sample; Fig. S1).

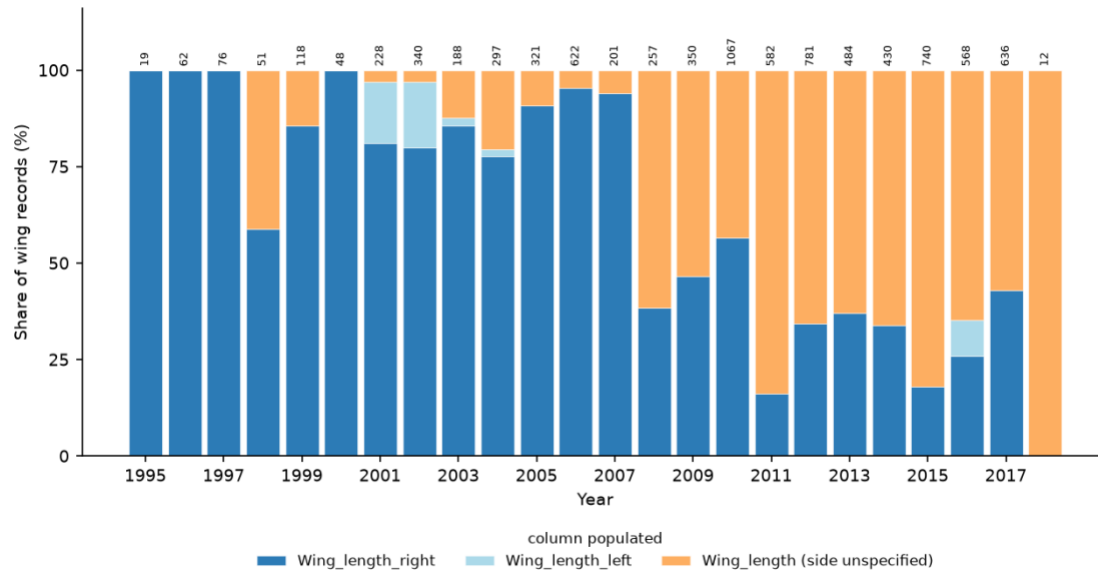
A provenance audit revealed substantial compositional turnover among active research teams over time: only 6 of the 42 contributors in the wing sample sampled birds in both the early period (1995–2006; 12 active contributors) and the late period (2013–2018; 33 active contributors; Fig. S1). Furthermore, spatial sampling expanded northward and into higher-elevation montane habitats during the latter half of the time series. In avian morphometrics, differences between observers can be of the same order as the temporal changes of interest (Yezerinac et al. 1992), and the column that supplied the analysed wing length also changed over time (Fig. S2). When contributor identity is left uncontrolled, compositional turnover can mimic temporal phenotypic change: the unadjusted wing-length decline (-1.8 mm/decade) falls to -0.7 mm/decade once contributor, municipality and the other covariates are modelled (main text; REML ladder in Table S1), although the size of that attenuation is itself uncertain (Table S4), and the apparent trends in bill width and trait variance do not persist.

To resolve genuine biological change from sampling turnover, all models beyond the baseline incorporate random intercepts for contributor and municipality ($(1 \mid \text{Main_researcher})$, $(1 \mid \text{Municipality})$). Within-between random effects models (Mundlak REWB, Mundlak 1978; Bell et al. 2019) do not detect directional change

within contributors, but estimate it imprecisely (within-contributor wing slope: -0.3 mm/decade, 95% CI: -1.6 to 1.0 ; complete cases: 0.1 mm/decade, 95% CI: -1.2 to 1.3 ; REML).



Supplementary Figure S1. Temporal distribution of live-capture passerine wing records across contributors (1995–2018). (Top) Number of wing records per calendar year, stacked by the 12 largest research contributors (and an aggregated “Other (30)” category). Daggers (†) mark contributors that sampled birds in both the early (1995–2006) and late (2013–2018) periods. (Bottom) Number of active research contributors per calendar year.



Supplementary Figure S2. Share of wing records by the wing-length column that supplied the analysed (coalesced) value, by year (priority: right, then left, then side-unspecified). Sample sizes (n) above bars are the annual numbers of wing records.

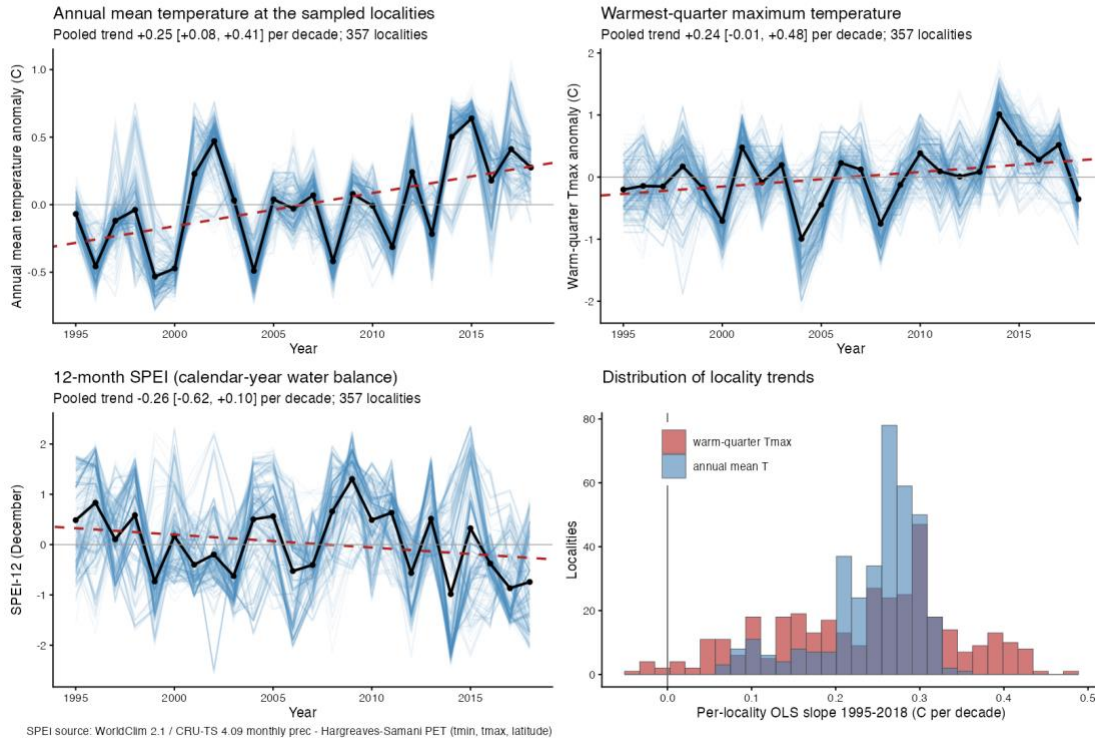
1.2 Supplementary Note S2: Regional Environmental and Climatic Trends (1995–2018)

We extracted monthly climate data from the WorldClim 2.1 historical monthly weather series (CRU-TS 4.09 downsampled to 2.5 arc-min, Harris et al. 2020; Fick and Hijmans 2017) at the capture localities. Of the 455 coordinate localities, 357 had raster values; the other 98 fall outside the land mask. SPEI-12 is the standardised 12-month climatic water balance, precipitation minus potential evapotranspiration (Vicente-Serrano et al. 2010); we calculated potential evapotranspiration with the Hargreaves–Samani method, took the December value of each year, and used 1990–2018 as the reference period. Trends were estimated from the locality-by-year values with linear mixed models on year, with random intercepts and slopes for locality and a random intercept for year.

Across sampling localities, mean annual temperature increased by $+0.246^{\circ}\text{C}/\text{decade}$ (95% CI: $+0.076$ to $+0.415^{\circ}\text{C}/\text{decade}$), an average warming of $+0.56^{\circ}\text{C}$ between 1995 and 2018 (Fig. S3, top left). Warmest-quarter maximum temperature rose at a similar rate but less certainly ($+0.24^{\circ}\text{C}/\text{decade}$, 95% CI: -0.01 to $+0.48$; top right). Total annual precipitation (-44.7 mm/decade, 95% CI: -123.1 to $+33.7$) and SPEI-12 (-0.256 units/decade, 95% CI: -0.617 to $+0.105$; bottom left) showed no clear directional trend. Regional environmental change across our capture sites was therefore characterised mainly by warming, without a clear multi-decadal change in moisture availability.

Climate at the bird sampling localities, 1995-2018

WorldClim 2.1 historical monthly weather (CRU-TS 4.09 downscaled, 2.5m), thin lines = 357 coordinate localities with data (98 more fall outside the land mask)
black = across-locality mean, dashed = pooled trend from the locality random-slope model with a year random effect (95% CI in panel subtitles)

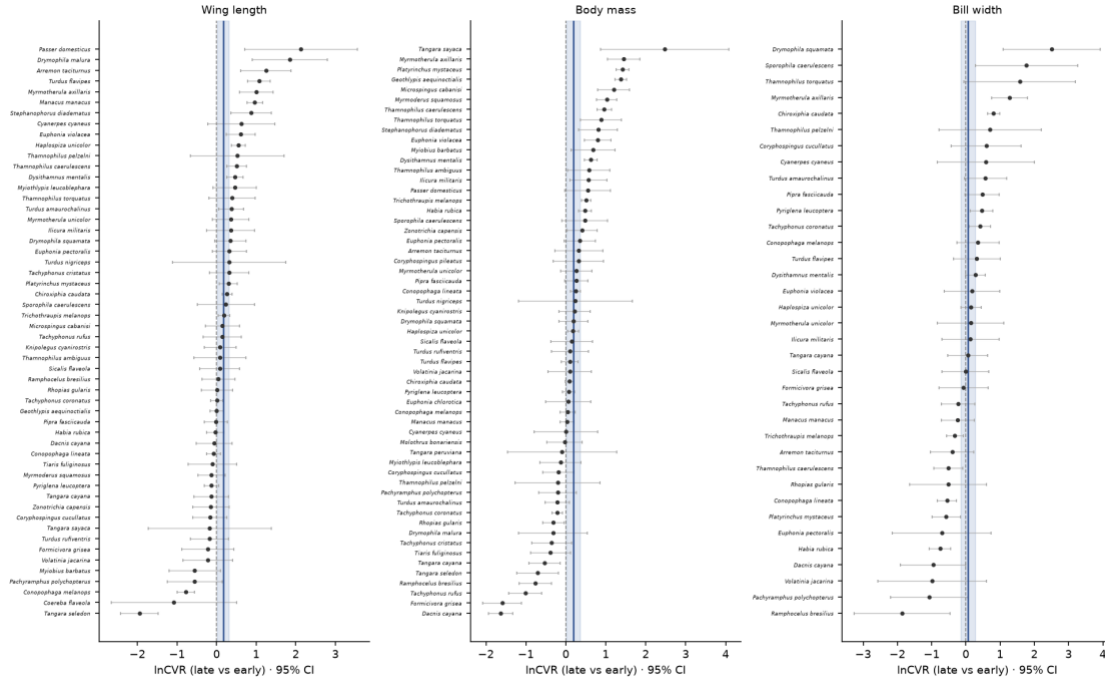


Supplementary Figure S3. Climate at the 357 coordinate localities with WorldClim values (1995–2018): annual mean temperature (top left), warmest-quarter maximum temperature (top right), December SPEI-12 (bottom left), and the distribution of per-locality temperature slopes (bottom right). Thin lines show individual localities, the black line the across-locality mean, and the dashed line the pooled random-slope trend (with a year random effect; 95% CI in the panel subtitles).

1.3 Supplementary Note S3: Species-Level Phenotypic Variability (lnCVR)

The main text reports among-individual phenotypic variability from continuous heteroscedastic dispersion models. As a complementary species-level summary, we computed log coefficient-of-variation ratios (lnCVR) with small-sample bias correction (Nakagawa et al. 2015) for every species with a standard deviation in both the early (1995–2006) and late (2013–2018) periods (at least two records per period), pooling effect sizes in random-effects meta-analytic models using metafor (Viechtbauer 2010). Pooled across contributors, variability appeared to increase for wing length and body mass (wing length lnCVR = +0.177, 95% CI: 0.031 to 0.323, $k = 54$; body mass +0.193, 95% CI: 0.021 to 0.366, $k = 57$; bill width +0.068, 95% CI: -0.146 to +0.282, $k = 36$). For wing length, restricted to species $\times$ sex cells measured by the same contributor in both periods

(at least five records per period; $k = 15$ cells), the apparent expansion was no longer evident ($\ln\text{CVR} = -0.25$, 95% CI: -0.51 to 0.02). This subset is small and compositionally narrow: 7 of the 15 cells come from one contributor, and the estimate depends on the threshold (-0.11 with at least two records per period, $k = 58$; -0.42 with at least ten, $k = 4$).



Supplementary Figure S4. Interspecific patterns of among-individual phenotypic variability ($\ln\text{CVR}$) between early and late sampling periods. Forest plots display the species-level log coefficient-of-variation ratio ($\ln\text{CVR}$) comparing late (2013–2018) versus early (1995–2006) periods for (A) wing length ($k = 54$ species), (B) body mass ($k = 57$ species), and (C) bill width ($k = 36$ species) across Atlantic Forest passerines. Points and error bars represent species estimates and 95% confidence intervals; vertical lines and shaded bands represent random-effects meta-analytic means ($\ln\text{CVR} = +0.177$ for wing length; $+0.193$ for body mass; $+0.068$ for bill width).

1.4 Supplementary Note S4: Temperature-Anomaly Models Across Specifications

The main text reports no clear association between morphology and the local annual temperature anomaly (ΔT_{local}) under the full M3 adjustment set. Because an anomaly measured over a warming study period is itself correlated with calendar year, and because all birds captured at one locality in one year share an anomaly value, we fitted five specifications (Table S3). All are produced by `Analysis/scripts/referee_reruns.R` and stored in `Analysis/output/referee_reruns/referee_reruns.rds`.

In the parsimonious models containing both terms (E1a), calendar year rather than the anomaly carries the signal for wing length and the isometry contrast; at full M3 adjustment (E3b), neither term is clearly associated with any trait. In the parsimonious fits the detrended anomaly (E1b) is associated with the isometry contrast ($p = 0.041$), in line with the consistently positive but uncertain sign reported in the main text. The locality-year intercept, included only in E2a, widens the anomaly standard errors by about 1.3 to 1.5 times relative to E1a, so the other rows may be somewhat over-precise. We did not use a cluster-robust variance estimator: the random-effects structure is non-nested, which available implementations (`clubSandwich` CR2) do not support, so locality-year clustering is handled with a random intercept instead.

A further caveat is one of exposure window rather than statistics. Wing length is fixed at moult, so the mean temperature of the calendar year in which a bird happened to be captured is not a plausible causal exposure for its wing length. Any thermal mechanism would have to operate during feather growth, which these data cannot resolve.

1.5 Supplementary Note S5: Model Specification, Sampling and Convergence

1.5.1 Robustness and sensitivity models

- **Fully adjusted, complete cases** (M3_cc): the fully adjusted specification on complete cases ($n = 8,282$). The primary sample ($n = 8,478$) does not discard records with incomplete covariates; instead, missingness is encoded explicitly. Records lacking a recorded altitude take a value imputed from other records at the same coordinates, and season and moult carry an explicit unrecorded level rather than being dropped. The complete-case model removes the 196 records with an imputed altitude or an unrecorded season, so the contrast between the two isolates the effect of that encoding. Unrecorded moult is kept as an explicit level in both models: it applies to 17% of wing records, and 8 of the 42 contributors never recorded it, so dropping these records would remove whole contributors and confound the complete-case contrast with contributor composition.
- **Fully adjusted, first captures only** (M7): restricted to first captures ($n = 7,791$) to exclude any potential bias from recapture histories.
- **Fully adjusted, anomalous blocks removed** (M3_noanom): re-fitted by REML after excluding 6 anomalous sampling blocks (193 records). Blocks are contributor $\times$ municipality $\times$ year cells. For each record we took the deviation of its wing length from its own species $\times$ sex mean, averaged those deviations within each cell, and flagged cells with at least 10 records whose mean deviation exceeded 6 mm in absolute value. We set the 6 mm threshold after inspecting the

data, so this screen is a post hoc sensitivity check and not part of the primary analysis; the threshold is roughly 1.5 residual standard deviations of the fully adjusted fit and about 8% of a typical wing; a shift of that size shared across an entire contributor-site-year block is more consistent with a protocol or data-entry change than with phenotypic variation. The screen is sign-blind: of the flagged blocks, 4 lie below and 2 above species means.

- **Year within vs between municipalities** (M5; REML): the same decomposition as M5src with municipality in place of contributor, to separate change within sampled sites from differences between sites sampled at different times.

1.5.2 Likelihood, priors and sampling

All models were Gaussian linear mixed models fitted in a Bayesian framework with a purpose-written Stan program (Carpenter et al. 2017). For the fully adjusted wing model, the length y_i of record i is

$$y_i = \mathbf{x}_i^\top \boldsymbol{\beta} + a_{s[i]} + b_{s[i]} t_i + c_{k[i]} + m_{j[i]} + u_{r[i]} + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2),$$

where $\mathbf{x}_i$ holds the fixed effects (sex, standardised year t_i , latitude, longitude, elevation, season, moult and wing side), a_s and b_s are the intercept and year-slope deviations of species s , and c_k , m_j and u_r are normally distributed random intercepts for contributor k , municipality j and individual bird r , each with its own standard deviation. The other models differ only in the terms they include. The species intercepts have a phylogenetic and a non-phylogenetic component,

$$\mathbf{a} = \sigma_a \left(\sqrt{H^2} \mathbf{L} \mathbf{z}_{\text{phy}} + \sqrt{1 - H^2} \mathbf{z}_{\text{sp}} \right), \quad \mathbf{L} \mathbf{L}^\top = \mathbf{A},$$

with $\mathbf{z}_{\text{phy}}$, $\mathbf{z}_{\text{sp}}$ standard normal and $\mathbf{L}$ the spectral square root of $\mathbf{A}$, so that $\mathbf{a}$ has covariance $\sigma_a^2 [H^2 \mathbf{A} + (1 - H^2) \mathbf{I}]$ and H^2 is the phylogenetic share of among-species variance in the intercepts. Species year slopes b_s are modelled as independent among species rather than as a Brownian process. The individual intercepts are integrated out of the likelihood analytically: the records of one bird are jointly normal with covariance $\sigma^2 \mathbf{I} + \tau^2 \mathbf{1} \mathbf{1}^\top$, whose inverse and determinant have closed forms, so the thousands of individual effects are never sampled. Together with sampling the phylogenetic effects in the eigenbasis of $\mathbf{A}$ (adapting marginalised spectral code by S. Ortega), this makes fitting across many trees feasible.

The response was centred and scaled before fitting and all estimates were back-transformed to the original scale. Priors were weakly informative on that standardised scale: $\mathcal{N}(0, 5)$ for fixed effects, half-Student- $t(3, 0, 1)$ for all standard deviations, Beta(1, 1) for H^2 , and $\mathcal{N}(0, 1)$ for the coefficients of the residual-SD model (below). Each model was sampled separately on each tree with CmdStan 2.39.0 through cmdstanr 0.9.0: four

chains of 1,000 warm-up and 500 sampling iterations (2,000 warm-up and 2,000 sampling iterations for bill width), with `adapt_delta` = 0.95 and a maximum tree depth of 12. The posteriors from the 20 trees were combined by pooling their draws, which gives the mixture posterior over trees and so carries phylogenetic uncertainty into every estimate (Nakagawa and Villemereuil 2019). No sampler divergences occurred in any of the 40 models fitted. For the parameters quoted in the text (year, temperature-anomaly and interaction terms, and the phylogenetic share and correlation of the species slopes), the largest rank-normalised $\hat{R}$ on any single tree was 1.010 (none of the 1,120 parameter-by-tree combinations exceeded 1.01). Trees on which any of these parameters initially exceeded 1.01 were refitted with 2,000 warm-up and 2,000 sampling iterations per chain and thinned by four, so that each keeps the same weight in the tree mixture; the replaced fits and the list of refitted trees are archived. The smallest per-tree bulk and tail effective sample sizes were 219 and 304, and summed over the 20 trees they were at least 12,000 and 16,794, respectively (Vehtari et al. 2021). The among-species standard deviation of the intercepts mixed more slowly (largest per-tree $\hat{R}$ 1.047; smallest per-tree bulk effective sample size 104); it is a nuisance parameter here and none of our conclusions rests on it.

1.5.3 Agreement with REML and phylogenetic slopes

As an independent check on the estimation engine, every model except the detrended-anomaly and locality-year checks (see *Interannual temperature anomalies*) was also fitted by restricted maximum likelihood with the `propto` covariance structure of `glmmTMB` v1.1.14 (Williams et al. 2025) and pooled across 50 trees with Rubin’s rules (Nakagawa and Villemereuil 2019; Nakagawa et al. 2026). For the temporal coefficients reported here, posterior means and REML estimates differed by at most 0.09 posterior standard deviations, and posterior SDs were 1.00–1.09 times the REML standard errors, as expected when the uncertainty in the variance components is propagated. The complete REML control ladder is given in Supplementary Tables S1 and S2. As a sensitivity analysis, we refitted the fully adjusted wing model and the isometry model with Brownian species slopes correlated with the species intercepts, a structure that `propto` cannot represent. For the fully adjusted wing model we also compared, by REML likelihood ratio on each tree, the model with independent species slopes against one that adds a phylogenetic slope component (boundary-corrected p). The larger model failed to converge on 10 of the 50 trees; on the remaining 40, none gave $p < 0.05$ (median $p = 0.46$).

2. Supplementary Tables

Supplementary Table S1. REML phylogenetic mixed models of wing length (glmmTMB, pooled across 50 trees; Wald 95% CI), evaluating temporal trajectories across sequential control tiers. M1a is shown to separate the contributor term from the municipality term, which M1 adds together.

Model specification	β (mm/SD-yr)	95% CI	Rate (mm/dec)	N
M0: Baseline (sex + latitude + phylogeny)	-0.92	[-1.40, -0.44]	-1.8	8,478
M1a: + Contributor intercept only	-0.68	[-1.08, -0.28]	-1.4	8,478
M1: + Contributor and municipality intercepts	-0.53	[-0.92, -0.13]	-1.1	8,478
M2: M1 + Wing-column proxy + Longitude + Altitude	-0.50	[-0.90, -0.10]	-1.0	8,478
M3: Full controls (M2 + Season + Moults + Individual)	-0.37	[-0.78, 0.04]	-0.7	8,478
M3_cc: Full controls on complete cases (N = 8,282)	-0.29	[-0.65, 0.08]	-0.6	8,282
M3_noanom: Full controls without the anomalous sampling blocks	-0.43	[-0.82, -0.05]	-0.9	8,285
M5src: Within-contributor temporal slope (Mundlak REWB)	-0.16	[-0.80, 0.48]	-0.3	8,478
M5: Within-municipality temporal slope (Mundlak REWB)	-0.74	[-1.68, 0.21]	-1.5	8,478
M4: Separate trend per contributor (mean slope)	-0.88	[-1.76, 0.01]	-1.7	8,478
M3_longrun: Six long-running contributors only	0.09	[-0.53, 0.72]	0.2	3,017
M6: Unknown-sex replication (full M3 controls)	-0.69	[-1.21, -0.17]	-1.4	3,532
M7: First captures only (no recaptures)	-0.38	[-0.78, 0.02]	-0.8	7,791

Supplementary Table S2. Body mass temporal models across control tiers (REML check on the Bayesian estimates, glmmTMB pooled across 50 phylogenetic trees). Moults is not included because it is not recorded for enough of the body-mass records; M3 and the within-contributor model use the complete-case sample.

Model specification	β (per SD-yr)	95% CI	Rate (%/dec)	N
Body mass M0 Baseline (sex + latitude + phylogeny)	-0.0072	[-0.0165, 0.0021]	-1.4%	11,256
Body mass M1 (+ Contributor and municipality intercepts)	-0.0056	[-0.0164, 0.0051]	-1.1%	11,256
Body mass M3 Full controls (M1 + Longitude + Altitude + Season + Individual)	-0.0046	[-0.0138, 0.0047]	-0.9%	11,077
Body mass Mundlak within-contributor slope	-0.0032	[-0.0104, 0.0040]	-0.6%	11,077

Model specification	β (per SD-yr)	95% CI	Rate (%/dec)	N
		0.0041]		

Supplementary Table S3. Temperature-anomaly associations across specifications (REML, glmmTMB). The anomaly effect is per °C and the year effect per SD of year; the isometry contrast is $\ln M - 3 \ln L$. E1a, E1b and E2a are fits without phylogeny, with contributor and municipality intercepts (a simpler control set than M3); E2a adds a locality-year intercept. E3 and E3b use the full M3 adjustment set across 50 phylogenetic trees; **Trees** is the number of trees whose fits converged and were pooled (E3 converged on fewer than 50 trees for body mass and the isometry contrast, whereas every E3b fit used all 50).

Specification	Response	Term	Estimate	95% CI	p	N	Trees
E1a: anomaly + year	Wing length (mm)	ΔT (local annual anomaly, °C)	-0.02200	[-0.405, 0.361]	0.9100	7,224	1
E1a: anomaly + year	Wing length (mm)	Year (per SD)	-0.53800	[-0.967, -0.109]	0.0140	7,224	1
E1b: detrended anomaly	Wing length (mm)	ΔT detrended (°C)	-0.16300	[-0.546, 0.22]	0.4000	7,224	1
E1a: anomaly + year	log body mass	ΔT (local annual anomaly, °C)	0.00727	[-0.0062, 0.0207]	0.2900	7,224	1
E1a: anomaly + year	log body mass	Year (per SD)	-0.00189	[-0.0121, 0.00832]	0.7200	7,224	1
E1b: detrended anomaly	log body mass	ΔT detrended (°C)	0.00640	[-0.00695, 0.0198]	0.3500	7,224	1
E1a: anomaly + year	Isometry contrast ($\ln M - 3 \ln L$)	ΔT (local annual anomaly, °C)	0.01570	[-0.00379, 0.0352]	0.1100	7,224	1
E1a: anomaly + year	Isometry contrast ($\ln M - 3 \ln L$)	Year (per SD)	0.01990	[0.00579, 0.0341]	0.0057	7,224	1
E1b: detrended anomaly	Isometry contrast ($\ln M - 3 \ln L$)	ΔT detrended (°C)	0.02010	[0.000833, 0.0394]	0.0410	7,224	1
E2a: + locality-year intercept	Wing length (mm)	ΔT (local annual anomaly, °C)	-0.32900	[-0.902, 0.243]	0.2600	7,224	1
E2a: + locality-year intercept	log body mass	ΔT (local annual anomaly, °C)	0.00574	[-0.012, 0.0235]	0.5300	7,224	1
E2a: + locality-year intercept	Isometry contrast ($\ln M - 3 \ln L$)	ΔT (local annual anomaly, °C)	0.02570	[-0.00277, 0.0542]	0.0770	7,224	1
E3: M3 adjustment	Wing length (mm)	ΔT (local annual anomaly, °C)	0.01040	[-0.369, 0.39]	0.9600	7,070	50
E3b: M3 adjustment + year	Wing length (mm)	ΔT (local annual anomaly, °C)	0.11200	[-0.272, 0.496]	0.5700	7,070	50
E3b: M3 adjustment + year	Wing length (mm)	Year (per SD)	-0.29400	[-0.701, 0.113]	0.1600	7,070	50
E3: M3 adjustment	log body mass	ΔT (local annual	0.00485	[-0.00847,	0.4800	7,070	38

Specification	Response	Term	Estimate	95% CI	p	N	Trees
		anomaly, °C)		0.0182]			
E3b: M3 adjustment + year	log body mass	ΔT (local annual anomaly, °C)	0.00596	[-0.00749, 0.0194]	0.3900	7,070	50
E3b: M3 adjustment + year	log body mass	Year (per SD)	-0.00179	[-0.0104, 0.00686]	0.6900	7,070	50
E3: M3 adjustment	Isometry contrast (ln M – 3 ln L)	ΔT (local annual anomaly, °C)	0.01170	[-0.00777, 0.0312]	0.2400	7,070	48
E3b: M3 adjustment + year	Isometry contrast (ln M – 3 ln L)	ΔT (local annual anomaly, °C)	0.00945	[-0.0102, 0.0292]	0.3500	7,070	50
E3b: M3 adjustment + year	Isometry contrast (ln M – 3 ln L)	Year (per SD)	0.01120	[-0.00248, 0.0248]	0.1100	7,070	50

Supplementary Table S4. Recovery of planted wing-length trends in the real sampling structure (200 simulated data sets per scenario and planted value; lme4 REML). Clean: contributor and municipality effects drawn independently of year. Real: their conditional modes from the model fitted to the real data. Calibrated artefact: independent draws plus a contributor shift proportional to the contributor's mean capture year, sized to reproduce the real gap between the baseline and fully adjusted estimates. M5src gives the within-contributor slope. Estimates in mm/decade. A cluster bootstrap over contributors (200 draws) gave a 95% interval of -0.23 to 2.53 mm/decade for the difference between the M3 and M0 slopes.

Scenario	Planted (mm/dec)	Model	Mean estimate	Bias	RMSE	Coverage of 95% CI	Share of CIs excluding 0
Clean	0.0	M0	-0.52	- 0.52	1.15	0.55	0.45
Clean	0.0	M1	-0.33	- 0.33	0.52	0.88	0.12
Clean	0.0	M3	-0.07	- 0.07	0.41	0.94	0.06
Clean	0.0	M5src	-0.08	- 0.08	0.43	0.95	0.04
Clean	-0.5	M0	-1.07	- 0.57	1.18	0.56	0.55
Clean	-0.5	M1	-0.75	- 0.25	0.47	0.93	0.42
Clean	-0.5	M3	-0.49	0.01	0.40	0.95	0.21
Clean	-0.5	M5src	-0.51	- 0.01	0.44	0.94	0.22
Clean	-1.0	M0	-1.47	- 0.47	1.09	0.61	0.67
Clean	-1.0	M1	-1.25	- 0.25	0.48	0.90	0.86

Scenario	Planted (mm/dec)	Model	Mean estimate	Bias	RMSE	Coverage of 95% CI	Share of CIs excluding 0
Clean	-1.0	M3	-0.99	0.01	0.41	0.94	0.69
Clean	-1.0	M5src	-0.98	0.02	0.44	0.94	0.64
Real conditional modes	0.0	M0	-1.11	- 1.11	1.16	0.32	0.69
Real conditional modes	0.0	M1	-0.42	- 0.42	0.56	0.84	0.16
Real conditional modes	0.0	M3	-0.21	- 0.21	0.44	0.94	0.07
Real conditional modes	0.0	M5src	-0.16	- 0.16	0.45	0.93	0.07
Real conditional modes	-0.5	M0	-1.59	- 1.09	1.15	0.38	0.95
Real conditional modes	-0.5	M1	-0.89	- 0.39	0.55	0.84	0.57
Real conditional modes	-0.5	M3	-0.68	- 0.18	0.43	0.93	0.40
Real conditional modes	-0.5	M5src	-0.64	- 0.14	0.43	0.95	0.31
Real conditional modes	-1.0	M0	-2.05	- 1.05	1.12	0.39	1.00
Real conditional modes	-1.0	M1	-1.35	- 0.35	0.52	0.85	0.93
Real conditional modes	-1.0	M3	-1.14	- 0.14	0.42	0.93	0.79
Real conditional modes	-1.0	M5src	-1.07	- 0.07	0.43	0.94	0.69
Calibrated artefact	0.0	M0	-1.46	- 1.46	1.74	0.28	0.72
Calibrated artefact	0.0	M1	-0.37	- 0.37	0.56	0.83	0.17
Calibrated artefact	0.0	M3	-0.10	- 0.10	0.44	0.93	0.07
Calibrated artefact	0.0	M5src	-0.04	- 0.04	0.45	0.93	0.07
Calibrated artefact	-0.5	M0	-2.16	- 1.66	1.94	0.26	0.87
Calibrated artefact	-0.5	M1	-0.82	- 0.32	0.51	0.88	0.52
Calibrated artefact	-0.5	M3	-0.55	- 0.05	0.41	0.94	0.27
Calibrated artefact	-0.5	M5src	-0.49	0.01	0.44	0.94	0.21
Calibrated artefact	-1.0	M0	-2.56	- 1.56	1.83	0.24	0.94

Scenario	Planted (mm/dec)	Model	Mean estimate	Bias	RMSE	Coverage of 95% CI	Share of CIs excluding 0
Calibrated artefact	-1.0	M1	-1.29	- 0.29	0.51	0.88	0.85
Calibrated artefact	-1.0	M3	-1.02	- 0.02	0.43	0.94	0.69
Calibrated artefact	-1.0	M5src	-0.97	0.03	0.47	0.93	0.60

Supplementary Table S5. Wing-length ladder with millimetre and log-scale responses (glmmTMB REML with the phylogenetic species intercept, Rubin-pooled over the trees with a positive-definite Hessian; Wald 95% CI). Log-scale slopes are expressed as % per decade.

Model	Response	Rate	95% CI	Unit	Trees
M0: Baseline	log wing length	-2.54	[-3.85, - 1.21]	%/decade	50
M0: Baseline	wing length (mm)	-1.84	[-2.79, - 0.88]	mm/decade	50
M1: + Contributor and municipality	log wing length	-1.37	[-2.45, - 0.29]	%/decade	50
M1: + Contributor and municipality	wing length (mm)	-1.05	[-1.84, - 0.27]	mm/decade	29
M3: Full controls	log wing length	-0.88	[-2.00, 0.25]	%/decade	50
M3: Full controls	wing length (mm)	-0.74	[-1.55, 0.07]	mm/decade	50
M5src: Within-contributor slope	log wing length	-0.26	[-2.06, 1.57]	%/decade	50
M5src: Within-contributor slope	wing length (mm)	-0.32	[-1.59, 0.96]	mm/decade	50
M3_noanom: Full controls without anomalous blocks	log wing length	-1.12	[-2.11, - 0.12]	%/decade	44
M3_noanom: Full controls without anomalous blocks	wing length (mm)	-0.86	[-1.62, - 0.10]	mm/decade	50

Supplementary Table S6. Temperature-anomaly models under two baselines (full M3 adjustment, glmmTMB REML, phylogenetic species intercept, Rubin-pooled; $n = 7,070$). The primary anomaly subtracts the locality's mean over its own records (correlation with capture year 0.07); the alternative subtracts the locality's unweighted mean over every year from 1995 to 2018 (correlation 0.39), so it retains the regional warming trend and is shown with and without calendar year.

Response	Model	Term	Estimate	95% CI	Trees
Wing length (mm)	dT	ΔT , baseline over the locality's records	0.01040	[-0.369, 0.39]	50

Response	Model	Term	Estimate	95% CI	Trees
Wing length (mm)	dT + year	ΔT , baseline over the locality's records	0.11200	[-0.272, 0.496]	50
Wing length (mm)	dT + year	Year (per SD)	-0.29400	[-0.701, 0.113]	50
Wing length (mm)	dT_full	ΔT , baseline over 1995–2018	-0.41600	[-1.21, 0.381]	50
Wing length (mm)	dT_full + year	ΔT , baseline over 1995–2018	-0.03740	[-0.393, 0.318]	50
Wing length (mm)	dT_full + year	Year (per SD)	-0.28600	[-0.693, 0.121]	50
log body mass	dT	ΔT , baseline over the locality's records	0.00485	[-0.00847, 0.0182]	38
log body mass	dT + year	ΔT , baseline over the locality's records	0.00596	[-0.00749, 0.0194]	50
log body mass	dT + year	Year (per SD)	-0.00179	[-0.0104, 0.00686]	50
log body mass	dT_full	ΔT , baseline over 1995–2018	0.00235	[-0.0112, 0.0159]	50
log body mass	dT_full + year	ΔT , baseline over 1995–2018	0.00458	[-0.00776, 0.0169]	50
log body mass	dT_full + year	Year (per SD)	-0.00192	[-0.0106, 0.00679]	50
Isometry contrast (ln M – 3 ln L)	dT	ΔT , baseline over the locality's records	0.01170	[-0.00777, 0.0312]	48
Isometry contrast (ln M – 3 ln L)	dT + year	ΔT , baseline over the locality's records	0.00945	[-0.0102, 0.0292]	50
Isometry contrast (ln M – 3 ln L)	dT + year	Year (per SD)	0.01120	[-0.00248, 0.0248]	50
Isometry contrast (ln M – 3 ln L)	dT_full	ΔT , baseline over 1995–2018	0.02010	[9.28e-05, 0.04]	50
Isometry contrast (ln M – 3 ln L)	dT_full + year	ΔT , baseline over 1995–2018	0.01670	[-0.00149, 0.0349]	50
Isometry contrast (ln M – 3 ln L)	dT_full + year	Year (per SD)	0.01000	[-0.00362, 0.0237]	50

References

- Bell, Andrew, Malcolm Fairbrother, and Kelvyn Jones. 2019. "Fixed and Random Effects Models: Making an Informed Choice." *Quality & Quantity* 53 (2): 1051–74.
<https://doi.org/10.1007/s11135-018-0802-x>.
- Carpenter, Bob, Andrew Gelman, Matthew D. Hoffman, et al. 2017. "Stan: A Probabilistic Programming Language." *Journal of Statistical Software* 76 (1): 1–32.
<https://doi.org/10.18637/jss.v076.i01>.

- Fick, Stephen E., and Robert J. Hijmans. 2017. "WorldClim 2: New 1-Km Spatial Resolution Climate Surfaces for Global Land Areas." *International Journal of Climatology* 37 (12): 4302–15. <https://doi.org/10.1002/joc.5086>.
- Harris, Ian, Timothy J. Osborn, Phil Jones, and David Lister. 2020. "Version 4 of the CRU TS Monthly High-Resolution Gridded Multivariate Climate Dataset." *Scientific Data* 7 (1): 109. <https://doi.org/10.1038/s41597-020-0453-3>.
- Mundlak, Yair. 1978. "On the Pooling of Time Series and Cross Section Data." *Econometrica* 46 (1): 69–85. <https://doi.org/10.2307/1913646>.
- Nakagawa, Shinichi, Santiago Ortega, Ayumi Mizuno, et al. 2026. *prepR4pcm: Prepare Data and Trees for Phylogenetic Comparative Methods*. <https://github.com/itchyshin/prepR4pcm>.
- Nakagawa, Shinichi, Robert Poulin, Kerrie Mengersen, et al. 2015. "Meta-Analysis of Variation: Ecological and Evolutionary Applications and Beyond." *Methods in Ecology and Evolution* 6 (2): 143–52. <https://doi.org/10.1111/2041-210X.12309>.
- Nakagawa, Shinichi, and Pierre de Villemereuil. 2019. "A General Method for Simultaneously Accounting for Phylogenetic and Species Sampling Uncertainty via Rubin's Rules in Comparative Analysis." *Systematic Biology* 68 (4): 632–41. <https://doi.org/10.1093/sysbio/syy089>.
- Rodrigues, Rodolpho Credo, Érica Hasui, Julia Camara Assis, et al. 2019. "ATLANTIC BIRD TRAITS: A Data Set of Bird Morphological Traits from the Atlantic Forests of South America." *Ecology* 100 (6): e02647. <https://doi.org/10.1002/ecy.2647>.
- Vehtari, Aki, Andrew Gelman, Daniel Simpson, Bob Carpenter, and Paul-Christian Bürkner. 2021. "Rank-Normalization, Folding, and Localization: An Improved $\hat{R}$ for Assessing Convergence of MCMC (with Discussion)." *Bayesian Analysis* 16 (2): 667–718. <https://doi.org/10.1214/20-BA1221>.
- Vicente-Serrano, Sergio M., Santiago Beguería, and Juan I. López-Moreno. 2010. "A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index." *Journal of Climate* 23 (7): 1696–718. <https://doi.org/10.1175/2009JCLI2909.1>.
- Viechtbauer, Wolfgang. 2010. "Conducting Meta-Analyses in R with the metafor Package." *Journal of Statistical Software* 36 (3): 1–48. <https://doi.org/10.18637/jss.v036.i03>.
- Williams, Coralie, Maeve McGillicuddy, Szymon M. Drobniak, Benjamin M. Bolker, David I. Warton, and Shinichi Nakagawa. 2025. "Fast Phylogenetic Generalised Linear

Mixed-Effects Modelling Using the glmmTMB R Package.” *bioRxiv*, ahead of print.
<https://doi.org/10.64898/2025.12.20.695312>.

Yezerinac, Stephen M., Stephen C. Loughheed, and Paul Handford. 1992. “Measurement Error and Morphometric Studies: Statistical Power and Observer Experience.” *Systematic Biology* 41 (4): 471–82. <https://doi.org/10.1093/sysbio/41.4.471>.