

Youngflesh et al.

1 *Running Head*: Coloniality, stochasticity, and phenology

2 Rethinking ‘normal’: The role of stochasticity in the phenology of a synchronously
3 breeding seabird

4

5 Casey Youngflesh^{1†}, Stephanie Jenouvrier^{2,3}, Jefferson T. Hinke⁴, Lauren DuBois⁵, Judy St.
6 Leger⁵, Wayne Z. Trivelpiece⁶, Susan G. Trivelpiece⁶, Heather J. Lynch¹

7 ¹Department of Ecology and Evolution, Stony Brook University, Stony Brook, NY 11790 USA

8 ²Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543 USA

9 ³Centre d’Etudes Biologiques de Chizé, UMR 7372 Centre National de la Recherche
10 Scientifique/Univ La Rochelle, FR-79360 Villiers en Bois, France

11 ⁴Antarctic Ecosystem Research Division, Southwest Fisheries Science Center, National Marine
12 Fisheries Service, National Oceanic and Atmospheric Administration, La Jolla, CA 92037 USA

13 ⁵SeaWorld, San Diego, CA 92109 USA

14 ⁶5959 Shoreline Highway, Bolinas, CA 94924 USA

15

16 †Corresponding author: casey.youngflesh@stonybrook.edu

17

18

19

20

21

22

23

24

25 ***This is the peer-reviewed version of the following article:***

26 Youngflesh C, Jenouvrier S, Hinke JT, et al. Rethinking “normal”: The role of stochasticity in the phenology
27 of a synchronously breeding seabird. *J Anim Ecol.* 2018;00:1–9.

28 ***which has been published in final form at <https://doi.org/10.1111/1365-2656.12790>. This article may be***
29 ***used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.***

30 **SUMMARY/ABSTRACT**

- 31 1. Phenological changes have been observed in a variety of systems over the past century.
32 There is concern that, as a consequence, ecological interactions are becoming
33 increasingly mismatched in time, with negative consequences for ecological function.
- 34 2. Significant spatial heterogeneity (inter-site) and temporal variability (inter-annual) can
35 make it difficult to separate intrinsic, extrinsic, and stochastic drivers of phenological
36 variability. The goal of this study was to understand the timing and variability of
37 breeding phenology of Adélie penguins under fixed environmental conditions, and to use
38 those data to identify a ‘null model’ appropriate for disentangling the sources of variation
39 in wild populations.
- 40 3. Data on clutch initiation were collected from both wild and captive populations of Adélie
41 penguins. Clutch initiation in the captive population was modeled as a function of year,
42 individual, and age to better understand phenological patterns observed in the wild
43 population.
- 44 4. Captive populations displayed as much inter-annual variability in breeding phenology as
45 wild populations, suggesting that variability in breeding phenology is the norm and thus
46 may be an unreliable indicator of environmental forcing. The distribution of clutch
47 initiation dates was found to be moderately asymmetric (right skewed) both in the wild
48 and in captivity, consistent with the pattern expected under social facilitation.
- 49 5. The role of stochasticity in phenological processes has heretofore been largely ignored.
50 However, these results suggest that inter-annual variability in breeding phenology can
51 arise independent of any environmental or demographic drivers and that synchronous
52 breeding can enhance inherent stochasticity. This complicates efforts to relate

53 phenological variation to environmental variability in the wild. Accordingly, we must be
54 careful to consider random forcing in phenological processes, lest we fit models to data
55 dominated by random noise. This is particularly true for colonial species where breeding
56 synchrony may outweigh each individual's effort to time breeding with optimal
57 environmental conditions. Our study highlights the importance of identifying appropriate
58 null models for studying phenology.

59 *Keywords: Antarctica, Adélie penguin, Bayesian hierarchical model, climate change, coloniality,*
60 *phenological mismatch, Pygoscelis adeliae, stochasticity, synchrony*

61

62

63

64

65

66

67

68

69

70

71 **INTRODUCTION**

72 There is concern that ecological interactions are becoming increasingly mismatched in time as a
73 result of climate change-driven shifts in the timing of key life history events such as migration,
74 foraging, and breeding (Thackeray et al. 2016). Such mismatches may result in decreased fitness
75 (Cushing 1974, Visser and Both 2005) with long-term repercussions for population dynamics
76 (Ludwig et al. 2006, Miller-Rushing et al. 2010, McLean et al. 2016). These asynchronies have
77 been observed in a diverse range of taxa, including arthropods (Winder and Schindler 2004, Both
78 et al. 2009), birds (Both et al. 2009, Visser et al. 2012), fish (Durant et al. 2005), and mammals
79 (Post and Forchhammer 2008) and in a variety of ecological systems (Kerby et al. 2012 and
80 references therein). To better understand and predict how phenological change might impact
81 ecological systems, it is important to recognize what factors drive the timing of these life-history
82 events.

83 Photoperiod has been identified as a proximate driver of the timing of key life history events
84 in both plants and animals (animal breeding, animal migration, plant flowering; Hay 1990,
85 Temte and Temte 1993, Bradley et al. 1999, Helm 2009, Zerbe et al. 2012). Studies of
86 phenological variation (inter-annual variation and/or long-term trends), however, have focused
87 on abiotic environmental factors such as temperature and precipitation (Visser et al. 2009,
88 Thackeray et al. 2016), or biotic factors such as body condition, which may reflect
89 environmental conditions and/or prey availability (Bêty et al. 2003; see review in Dawson 2008).
90 Variability in population-level phenology arises by some combination of three factors: (1)
91 extrinsic drivers: changes in environmental (both abiotic and biotic) conditions, including
92 changes in other components of the ecological community (e.g., prey, predators, mutualists); (2)
93 intrinsic drivers: fixed differences among individuals coupled with demographic turnover within

94 the population; and (3) stochastic variation: seemingly random variation in the timing of
95 breeding of individuals (unrelated to known intrinsic or extrinsic factors) that propagates up to
96 population-level metrics of phenology. Consistent with usage by Lande et al. (2003), we define
97 stochasticity in this context as variability that is either truly random or, at least, appears random
98 with respect to factors relevant to the question of interest. While numerous studies have focused
99 on intrinsic and extrinsic drivers of phenology, there has been relatively little attention paid to
100 the role of stochasticity. This lack of attention to stochastic factors is due, in part, to the fact that
101 wild populations are subject to fluctuating environmental conditions, making it difficult to
102 identify forces independent of measureable extrinsic and intrinsic factors.

103 The scale on which data are collected (i.e., population-level vs. individual-level) also
104 contributes to the challenge of identifying stochasticity. Most studies of phenology in wild
105 populations use population-level summary statistics (such as first or mean timing of a life-history
106 event in a population) due to difficulties associated with collecting individual-level data. Metrics
107 that capture the first instance of an event of interest are known to be problematic due to their
108 sensitivity to population size and sampling frequency (Miller-Rushing et al. 2008). While more
109 robust to these influences, population mean metrics can be affected by other factors such as age
110 (Ainley et al. 1983, Ainley 2002) and random variation among individuals (Crawley and
111 Akhteruzzaman 1988) that, through shifts in demographic composition, can impact population-
112 level statistics. For instance, if older individuals breed earlier, an aging population will display
113 an apparent trend towards earlier breeding (Lewis et al. 2012). Similarly, any random subset of
114 individuals may, by chance, have an earlier or later average phenology than another such random
115 subset of individuals.

116 Individual-level data, in either wild or captive populations, allows us to track the role of
117 demographic turnover (age and individual effects) on phenological response. An even more
118 comprehensive understanding of the role for stochastic factors in driving phenological variability
119 can be developed by paring studies of wild populations with studies of populations under fixed
120 environmental conditions (Lambrechts et al. 1999, Dunne et al. 2003, Visser et al. 2009).
121 Phenological variability in captive populations kept under fixed conditions represents a null
122 model against which interannual variability in wild populations can be compared. Without such a
123 control group, it may be difficult to determine the extent to which phenological variability is
124 driven by environmental variability or, alternatively, inherent stochasticity. An estimate of inter-
125 annual variation under fixed environmental conditions is therefore of value when interpreting
126 phenological studies of wild populations. Assessing variability under these conditions may be
127 particularly important for colonially breeding species that may rely on social cues to synchronize
128 breeding (Ims 1990b).

129 Here we used data collected from both wild (representing naturally variable environmental
130 conditions) and captive (representing fixed environmental conditions) populations of Adélie
131 penguins *Pygoscelis adeliae*, to identify the role of stochastic factors in driving breeding
132 phenology. Adélie penguins are a site faithful, highly colonial species that inhabit the Antarctic
133 continent and surrounding islands. Several studies (Barbraud and Weimerskirch 2006,
134 Emmerson et al. 2011, Hinke et al. 2012, Lynch et al. 2012, Youngflesh et al. 2017) have sought
135 to understand the conditions associated with breeding phenology in Adélie penguins. However,
136 despite considerable effort to collect and analyze long-term phenological data, our understanding
137 of what drives phenology in this species and the potential role that stochastic factors might play,
138 is limited. Our aim was to address three principal questions in this study: 1) how variable is

139 Adélie penguin breeding phenology in the absence of environmental variability; 2) can variation
140 in breeding phenology under fixed environmental conditions be explained by individual variation
141 and age structure, or is there substantial residual random variation that remains unexplained; and
142 3) what implications do our findings have for interpreting variability in wild populations of
143 synchronously breeding colonial species? In sum, to what extent have we been overemphasizing
144 the role of exogenous environmental forcing in the phenological variability of some wild
145 populations?

146

147 **MATERIALS AND METHODS**

148 *Description of data*

149 Individual-level data on clutch initiation date (CID – the date in which the first egg is laid in each
150 nest) were obtained for all individuals in a captive Adélie penguin population at SeaWorld San
151 Diego from 1992 - 2015 (89 penguins in total). The number of breeding females in each year
152 varied from 12 to 37 throughout the course of the study. The youngest birds to breed were 2
153 years of age, while the oldest were 45 years of age. Most penguins (67 of the 89) were born in
154 captivity – individuals born before 1984 were taken as chicks from wild populations in the Ross
155 Sea region of Antarctica in 1976.

156 All captive penguins were associated with unique identifiers, and metrics for each individual
157 were tracked through time. Temperature at the facility was kept at a constant -4°C to -2°C year-
158 round. Feeding regime of the captive birds did not change over the study period. Nesting
159 materials (stones) were provided at the same time in each year. Photoperiod for the exhibit

160 lighting mimicked that of 77°S latitude (though minimal exhibit lighting is required during the
161 winter period), representative of the southern limit of the Adélie penguin breeding range.
162 Seasonal variations in lighting are accounted for in the lighting regime and are constant from
163 year to year. The viewing area is setup in such a way to minimize light exposure to the penguin
164 enclosure. The penguin exhibit at SeaWorld underwent a 6-month renovation in 2005. The
165 captive penguin population used in this study was kept in a separate enclosure during this
166 renovation, which may have impacted breeding phenology in this year.

167 Data on CID were also obtained for a unique set of 100 individuals each year in a wild
168 population of Adélie penguins located at Admiralty Bay, Antarctica (62.2°S, 58.4°W) from
169 1986-2012 using methodologies outlined in Hinke et al. (2012). Individual-level data across
170 years was not available for the wild population. Wild penguins are typically younger than those
171 in captivity, generally less than 20 years of age (Ainley et al. 1983, Ainley 2002).

172 *Statistical analysis of individual phenology in marked captive penguins*

173 For the captive Adélie penguin time series, a hierarchical Bayesian approach was used to model
174 CID (y_{ij}), with year (i) and individual (j) as random effects (α and β , respectively) and age as a
175 fixed effect (γ):

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma * AGE_{ij} + \varepsilon_{ij} \quad (1)$$

$$\alpha_i \sim N(0, \sigma^2_{year})$$

$$\beta_j \sim N(0, \sigma^2_{individual})$$

$$\varepsilon_{ij} \sim N(0, \sigma^2_{model})$$

176 where μ represents the intercept, AGE represents the age of the female, and ε represents the error
177 term. This model was used to determine whether variation in y_{ij} was driven by the α parameter
178 (variance attributed to a colony-wide effect that varies among years), the β parameter (variance
179 attributed to fixed differences between individuals), the AGE covariate, or the error term (ε).
180 Individual data used in this statistical framework allowed us to account for potential confounding
181 factors related to differences in phenology due to individual identity (or quality) and age, and
182 provide a more complete understanding of these processes than is possible using population-level
183 summary statistics.

184 Models were fit using the R package ‘rjags’ (Plummer 2013), an interface to JAGS (Plummer
185 2003), in the R statistical environment (R Core Team 2016). Normal priors were used for α , β ,
186 γ , and ε . Broad Gamma priors were used for all precision ($\tau = \frac{1}{\sigma^2}$) parameters (shape = 0.01,
187 rate = 0.01). Posterior distributions were derived from three chains with 5,000 samples (after
188 thinning every other draw) following a ‘burn-in’ period of 40,000 draws and an adaptation
189 period of 5000 draws. Model convergence was assessed through a visual analysis of the posterior
190 chains, in addition to the use of the Gelman-Rubin convergence diagnostic (Brooks and Gelman
191 1998). All models unambiguously converged. Parameter estimate plots were generated using the
192 ‘MCMCvis’ package (Youngflesh 2016) while other plots were generated using the ‘ggplot2’
193 package (Wickham 2009) in the R statistical environment.

194 *Statistical analysis of population-level phenology in captive and wild penguins*

195 No information on individual phenology across years was available for the wild population (i.e.,
196 no information on the β parameter or AGE covariate), leaving it ambiguous as to which
197 component was contributing to the variability in y_{ij} (CID of individual j in year i). Therefore, to

198 directly compare the captive and wild Adélie populations, we considered population-level
199 aggregate summaries of phenology in both populations. The median colony CID in each year
200 (median across individuals; denoted y_i .) was calculated for both the captive and wild Adélie
201 penguin populations. Between year variation in median phenology, $\sigma_{between}^2 = var(y_i)$,
202 included variation from all sources (year, individual, age, and unexplained variance as captured
203 by ε). Note that in contrast to σ_{model}^2 in Eq 1, $\sigma_{between}^2$ includes variation due to age and
204 individual identity as well as residual stochastic variation unrelated to these factors. For both the
205 captive and wild populations, variation within year i , $\sigma_{i,within}^2 = var(y_{ij})$, was used as a
206 measure of year-specific, within-population breeding synchrony. We also note that while
207 differences in (simulated) latitude may have generated a fixed difference in photoperiod between
208 the captive and wild populations, our analysis examined only within-site inter-annual variability
209 in breeding phenology.

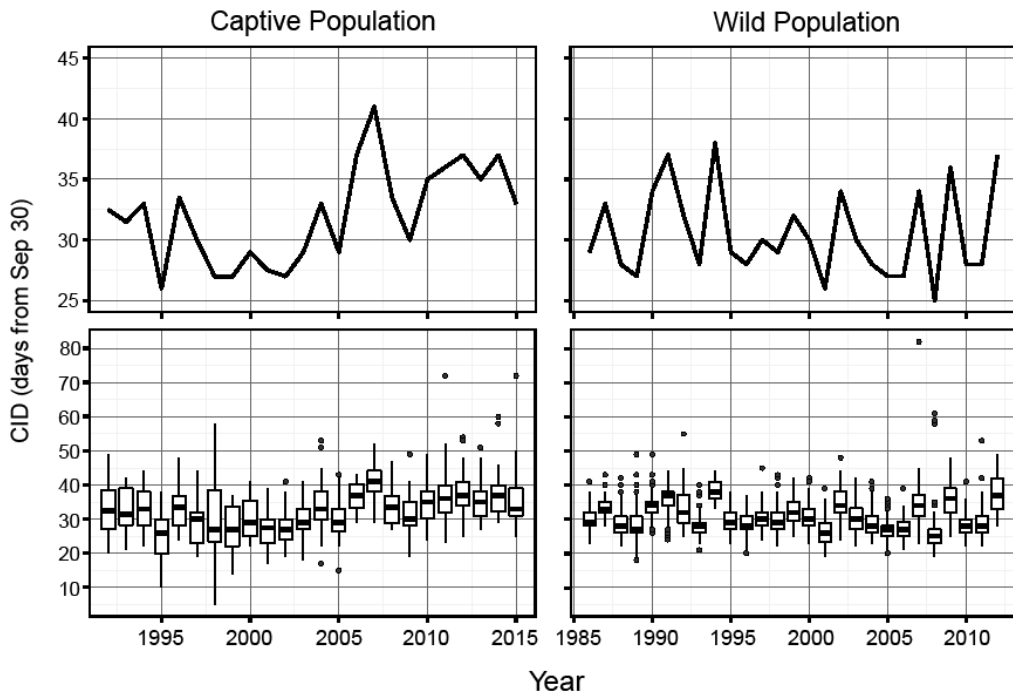
210 To investigate whether individual breeding dates within a colony were distributed
211 symmetrically around a population mean, as might be expected *a priori* if individuals were
212 acting independently, CID values in each year were standardized ($z_i = \frac{y_i - \bar{y}_i}{sd(y_i)}$) and aggregated
213 across years to be analyzed for skewness using a D'Agostino skewness test (D'Agostino 1970).
214 All analyses were performed in the R statistical environment (R Core Team 2016).

215

216 RESULTS

217 *Inter-annual and intra-annual variance in CID*

218 Inter-annual variance of median colony CID was similar between the captive
 219 ($\sigma^2_{between-captive} = 15.8 [SE = 4.7]$) and wild ($\sigma^2_{between-wild} = 13.5 [SE = 3.7]$) populations
 220 (Fig. 1). This degree of phenological variability is comparable to that seen in other taxa
 221 (Appendix S1). Within year, individual birds were relatively synchronous in both populations,
 222 with greater synchrony among individuals in the wild population compared to the captive
 223 population ($E[\sigma^2_{within-wild}] < E[\sigma^2_{within-captive}]$; Welch two-sample t-test, $t = 4.99$, $df = 28.5$, p
 224 < 0.001) (Fig. 1; Appendix S1).



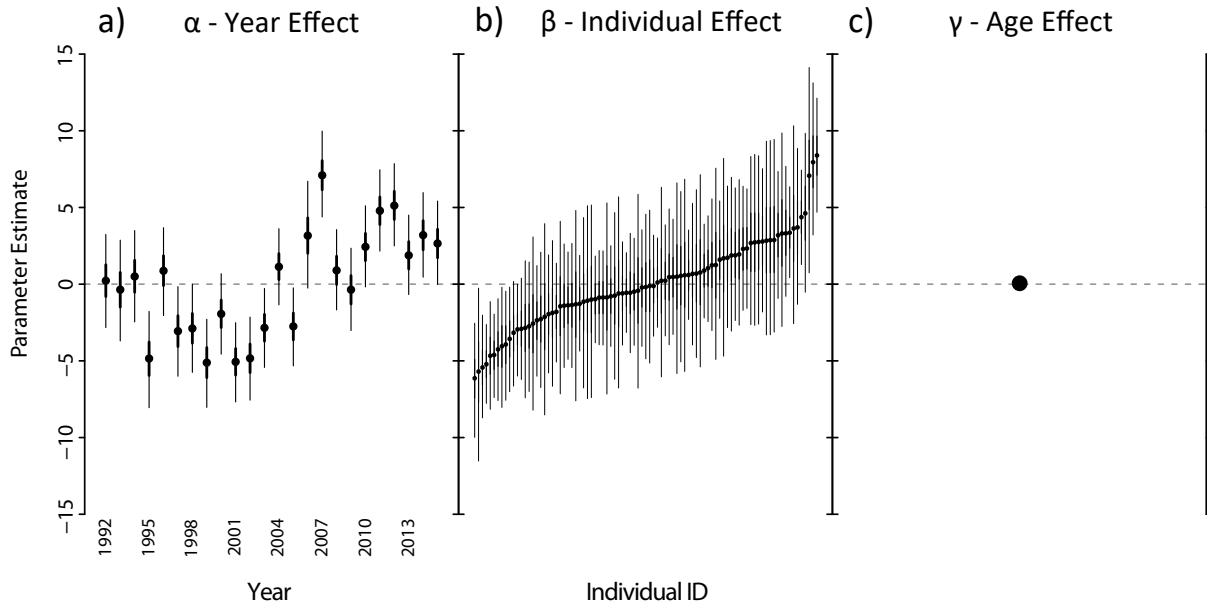
225
 226 **Fig. 1.** Mean colony breeding phenology (top panels) and distribution of individual CID in each
 227 year (bottom panels). The bold lines in the box-and-whisker plots represent the median CID
 228 while boxes represent the 25th and 75th quantiles. The top and bottom of the whiskers are $1.5 \times$
 229 IQR (inter-quartile range) from the upper and lower boxes, respectively. Data beyond this range

230 are plotted as points. The number of data points per year in the captive population ranged from
231 13 to 38. One hundred data points were collected for each year in the wild population.

232

233 *Individual phenology in marked captive penguins*

234 A strong year effect (α) on CID was apparent for the captive penguin population (Fig. 2a;
235 $median(\sigma_{year}^2) = 13.81$ [Appendix S1]). Significant year to year variation was found, even
236 when accounting for the effects of age and individual. Inter-annual variation is still substantial
237 when accounting for first-order autocorrelation (Appendix S2). Random effects due to individual
238 (β) were notable (as evidenced by the spread of β parameter estimates), with some individuals
239 breeding consistently earlier/later than other individuals (Fig. 2b; $median(\sigma_{individual}^2) = 13.66$
240 [Appendix S1]). Over the 24-year study period, 16 different individuals were the first to lay eggs
241 in a particular year (in 3 of these years, multiple individuals laid eggs on the same day). No
242 biologically significant effect of age (γ) on CID was found (median posterior estimate = 0.06
243 days/year); the 95% credible interval overlapped 0 (Fig. 2c). Variance unaccounted for by the
244 model is represented by $median(\sigma_{model}^2) = 34.96$ (Appendix S1).



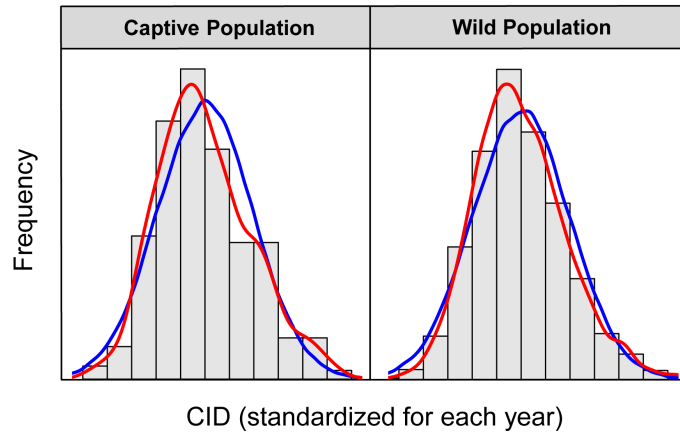
245

246 **Fig. 2.** Posterior estimates for the captive population parameters: (a) year effect – α ; (b)
 247 individual effect – β ; and (c) age effect – γ parameters (see Eq 1). Black circles represent
 248 posterior medians. Thicker lines represent 50% credible intervals while thinner lines represent
 249 95% credible intervals. Error bars for the γ parameter are obscured by the point itself.

250

251 *Intra-annual variance in CID*

252 The distributions of CID for both captive and wild populations (Fig. 3) had a small but
 253 statistically significant right skew ($g_1^{captive} = 0.54$ [SE = 0.10]; $g_1^{wild} = 0.79$, [SE = 0.04];
 254 D’Agostino test, $p < 0.001$).



255

256 **Fig. 3.** Distribution of CIDs for the captive (left) and wild (right) populations. Data were
 257 standardized and aggregated across years. Colored lines depict the kernel density estimates on
 258 the distribution. Red lines represent the actual distributions of CID, while the blue lines represent
 259 the normal distributions generated using the actual mean and variance of the CID distribution.

260

261 **DISCUSSION**

262 Despite prior expectations of lower inter-annual variance in breeding phenology under fixed
 263 environmental conditions, we found that both the captive and wild Adélie penguin populations
 264 exhibit a similar degree of inter-annual fluctuation in breeding phenology. Stochasticity appears
 265 to play a substantial role in determining the timing of Adélie penguin breeding within the
 266 window of time dictated by biological and physical constraints. We suggest that stochasticity at
 267 the individual level is amplified by the importance of breeding synchrony among individuals,
 268 producing the observed inter-annual variance.

269 *Inherent inter-annual variation*

270 Previous studies have suggested that photoperiod (Dawson 2008), abiotic conditions such as
271 rainfall (Leitner et al. 2003, Deviche et al. 2006) and temperature (Both et al. 2004, Torti &
272 Dunn 2005, Visser et al. 2009), biotic conditions such as food availability (Reynolds et al. 2003),
273 or some combination of these factors, are important in regulating phenology in a number of bird
274 species. While these conditions may define a larger envelope of time in which successful
275 breeding may occur, our analysis of a captive Adélie penguin population shows that external
276 drivers neither explain nor are required to generate substantial inter-annual variability in
277 breeding phenology (i.e., a large year effect). Inter-annual fluctuations are apparent even under
278 constant environmental conditions and after accounting for the effects of individual variation and
279 age (Fig. 2a). Other potential factors that we thought might have influenced penguin breeding
280 phenology, such as number of breeders in a given season, the timing of nesting material
281 availability, and potential effects of the 2005 captive population exhibit renovation, were
282 investigated but ultimately determined to have only minor effects (Appendix S2).

283 Beyond the random effect of year, some fixed differences in the timing of breeding among
284 individuals in the captive population do exist – that is, some individuals breed earlier than others
285 on average (Fig. 2b). However, the relatively minor shifts in the composition of the population
286 each year cannot explain the inter-annual fluctuations in the breeding phenology of captive
287 individuals. While previous work has suggested that age plays a role in determining Adélie
288 penguin breeding phenology in the wild (Ainley et al. 1983, Ainley 2002), the effect of age on
289 breeding phenology was minimal in the captive population studied here (Fig. 2c). This is not
290 surprising, given that one hypothesized mechanism by which age might impact breeding
291 phenology is through an individual's ability to navigate back to the breeding colony following
292 the overwintering period. Older individuals, being more experienced, are thought to be able to

293 find their way back to the breeding colonies more quickly, particularly through substantial sea
294 ice (Ainley et al. 1983). With no migration in a captive population, we would expect age and
295 experience to have a substantially smaller impact on breeding phenology.

296 *Interplay between synchrony and stochasticity*

297 Adélie penguins are highly synchronous breeders within a population in a given season (Fig. 1;
298 see also Ainley 2002), and previous work has shown that increased synchrony among individuals
299 leads to increased breeding success (Youngflesh et al. 2017). Where colonial breeding is driven
300 by predator avoidance, individuals that breed not only in the same area but also at the same time
301 as others in the population would be at an advantage (Darling 1938, Young 1994). Penguins
302 breeding during the same time period in the same location can more easily defend nests against
303 aerial predators, a principal threat to Adélie penguin chicks (Young 1994). Synchronous
304 breeding also results in an overwhelming influx of potential prey resources for species that prey
305 on penguins; such ‘predator-swamping’ can facilitate higher breeding success (Ims 1990a). We
306 hypothesize that the fitness consequences of breeding in sync with conspecifics may outweigh
307 the importance of matching optimal environmental conditions within the environmentally-driven
308 time envelope in which reproduction can occur successfully (as suggested in Hinke et al. 2012,
309 Youngflesh et al. 2017). We suggest that the importance of this phenomenon for any given
310 species reflects a balance between the expected fitness advantages of synchronous breeding
311 among conspecifics and the fitness advantages of breeding during some (environmentally-
312 determined) optimal period of time – the latter of which may be more difficult to assess for an
313 individual with limited information.

314 In this way, predators indirectly influence the breeding phenology of Adélie penguins – a
315 top-down process. Abiotic (e.g., photoperiod, temperature) and/or bottom-up processes (e.g.,
316 resource availability, organism physiological condition) are typically the focus of phenological
317 studies. However, top-down forcing (via antagonists such as predators and pathogens) has been
318 identified as an important process in the regulation of a number of phenological processes,
319 including absolute phenology (Elzinga et al. 2007, Galloway and Burgess 2012), phenological
320 synchrony (Hatchwell 1991, Sinclair et al. 2000), and the rate at which offspring develop
321 (Vonesh 2005). These processes often operate in concert with abiotic and/or bottom-up processes
322 (as suggested by Varpe et al. 2007, Burr et al. 2016), painting a complex picture of multiple
323 ecological determinants of phenological processes. While the mechanisms for top-down control
324 on phenology are diverse, these patterns may be more common among colonial breeders given
325 the importance of phenological synchrony for predator avoidance in many species (Ims 1990b).

326 While the precise mechanism that regulates this highly synchronous behavior in Adélie
327 penguins is unknown, social cues have been found to drive courtship and copulation (Waas 1988,
328 1991, 1995, Waas et al. 2000), and result in more synchronized breeding patterns (Setiawan et al.
329 2007). This has been demonstrated in a number of birds (Burger 1979, Danchin 1988, Waas et al.
330 2005), including other species of penguins, as well as in mammals (McClintock 1978, Scott
331 1986, Berger 1992). Vocalizations (Clark et al. 2012), exposure to courtship displays (Lehrman
332 and Friedman 1969), and chemical cues (McClintock 1978) have all been demonstrated as
333 proximate behavioral mechanisms by which this social facilitation occurs.

334 Slightly right skewed distributions of CID are apparent within each year – very late breeders
335 in this right skewed distribution contribute to a relatively large residual variance in our model.
336 This skew is also consistent with, though by no means proves conclusively, a role for social

337 facilitation in the timing of breeding. Even a simple model in which breeding is accelerated
338 proportional to the number of pairs that have recently initiated breeding creates a right skewed
339 distribution for clutch initiation (Appendix S1). This notion of facilitation is further supported by
340 the strong relationship (70% of variance explained) between first CID (the first breeders in each
341 year) and median CID in both the captive and wild populations (Appendix S1). Right skewed
342 distributions of phenological events are apparent in many species of both birds and plants
343 (Thomson 1980, Sparks et al. 2005, Wilson 2013) in magnitudes similar to those observed here
344 (Appendix S1). Environmental factors may play a role in driving this pattern for many species, a
345 hypothesis difficult to exclude if only wild data were available. In this case, however, both a
346 right skewed distribution and a strong relationship between first breeding and median breeding
347 are apparent under controlled conditions, despite inter-annual variability in the overall timing of
348 breeding.

349 It should be noted that the captive penguin population in this study exhibited less synchrony
350 among individuals in a given year than did the wild population (Fig. 1; Appendix S1). One
351 possible explanation for this relates to colony size; larger populations in the wild may, through
352 increased colony noise, better facilitate the transmission of social cues (Waas et al. 2000). Other
353 possibilities include the notion that the degree of synchrony is a plastic trait that may be relaxed
354 under the predator-free conditions of captivity. Future work, including manipulative
355 experiments, is required to understand the precise behavioral mechanisms regulating synchrony
356 in this species.

357 To be clear, we are not suggesting that synchronous breeding is necessary to observe the
358 impacts of stochastic factors on breeding phenology at the individual level, but rather that the
359 importance of synchrony in colonial species may facilitate the propagation of stochasticity from

360 the individual-level to that of the population. Social cues that facilitate synchrony (see above)
361 may encourage individuals to initiate breeding once other individuals in the colony have done so.
362 This cascading effect of synchronous breeding may drive a shift in the average phenology of the
363 population that is largely uncoupled from any environmental trigger.

364 *In ignoring the role of stochasticity, have we been modeling noise?*

365 In the search for causal drivers for phenology, the role of inherent stochasticity has been largely
366 ignored in the existing phenological literature. Our study has important implications for studying
367 patterns of phenology across all animal systems as it highlights the difficulty of teasing out the
368 extent to which the environment may, or may not, be driving variation in phenology. The
369 observed levels of inter-annual variability under fixed environmental conditions are of a similar
370 magnitude to those seen in the wild (Fig. 1; Appendix S1, S2) – a surprising result. Previous
371 studies of Adélie penguin phenology at Admiralty Bay found October mean air temperature
372 (thought to be related to snow melt and nest site availability) to be the most important
373 environmental determinant of CID in a model selection framework (Hinke et al. 2012, Lynch et
374 al. 2012). Our analyses of the same CID data (with additional years beyond what was used in the
375 original studies) do not contradict these original findings, but suggest that environmental drivers,
376 such as temperature, are layered on top of substantial ‘built in’ variability. Put another way, our
377 null model for phenological studies should not, by default, be one of stasis.

378 Whether the drivers that influence a pair’s ‘decision’ to breed are unknown, or unknowable,
379 the implication is that phenological variability of this scale can be generated in the absence of
380 environmental variability. This finding is of practical importance, as it highlights the difficulty in
381 identifying external causal drivers of phenological events. Our study shows that stochastic

382 variation in penguin breeding phenology, with respect to the environmental factors addressed
383 here, may be the rule rather than the exception and does not require external forcing from the
384 environment. This is similar to stochastic outcomes of individuals, whereby identical individuals,
385 experiencing identical conditions, will differ with respect to lifespan and fecundity (Caswell
386 2011). This role of stochasticity may partially explain previous findings of a relatively weak
387 relationship between Antarctic seabird breeding phenology and environmental forcing (Barbraud
388 and Weimerskirch 2006).

389 One way to address the inherent convolution of environmental variability and inherent
390 stochasticity is through the study of captive populations in controlled conditions. While studies
391 focusing on plants more often include experiments in controlled environments to distinguish the
392 role of multiple factors on phenology (Cleland et al. 2007), there are fewer studies elucidating
393 inter-annual changes in phenology in captive animals (but see Lambrechts et al. 1999, Visser et
394 al. 2009). Captive animal populations are not, of course, perfect replicas of wild populations. For
395 instance, penguins in captive populations are kept in enclosed spaces and do not undergo an
396 overwinter migration. Captive populations are also limited in size, often smaller than what might
397 be observed in the wild. Accordingly, we must be cautious not to overextend the analogy
398 between captive and wild populations. Nevertheless, studies focusing on populations in
399 controlled conditions can provide a wealth of information on phenology in the absence of all
400 environmental factors deemed potentially important in wild populations. Identification and
401 tracking of individual animals further allows for straightforward estimation of age effects and
402 random individual variation. Ultimately these studies can provide a reasonable null model
403 against which to assess factors contributing to variation in wild populations. Studies involving
404 experimental manipulation and/or transplantation of organisms to new environments (e.g., Helm

405 2009) may provide additional power to disentangle the factors controlling breeding phenology.
406 Our study highlights the challenges of understanding the factors driving phenology in wild
407 populations, and reminds us to take caution in ascribing causality when we are unaware of the
408 degree of inherent variation in the response variable of interest.

409

410 **AUTHORS' CONTRIBUTIONS**

411 CY, HJL, and SJ conceived the ideas and designed the methodology; JTH, LD, JSL, WZT, and SGT
412 collected the data; CY analyzed the data; CY and HJL led the writing of the manuscript. All authors
413 contributed to drafts and gave final approval for publication.

414

415 **ACKNOWLEDGEMENTS**

416 We acknowledge the many fieldworkers and SeaWorld staff that contributed to the collection of data used
417 in this study. We also acknowledge the authors of, and commenters on, the blog Dynamic Ecology for
418 insights into stochasticity in ecological systems. Funding to HJL and CY was provided by the National
419 Science Foundation Grant OPP/GSS 1255058, and to SJ, HJL, and CY, by the National Science
420 Foundation Grant 1341474. Data from Admiralty Bay was made possible with support from the US
421 Antarctic Marine Living Resources program, the Lenfest Oceans Program at the Pew Charitable Trusts,
422 and many prior grants from the National Science Foundation (including grant #1016936 to WZT/SGT).
423 The authors declare no conflicts of interest.

424

425 **DATA ACCESSIBILITY**

Youngflesh et al.

426 Data used in this study are available from the Dryad Digital Repository

427 <https://doi.org/10.5061/dryad.23sv1> (Youngflesh et al. 2017).

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443 **REFERENCES**

- 444 Ainley, D. (2002) *The Adélie Penguin: Bellwether of Climate Change*. Columbia University
445 Press, New York, NY, USA.
- 446 Ainley, D.G., LeResche, R.E. & Sladen, W.J.L. (1983) *Breeding Biology of the Adélie Penguin*.
447 University of California Press, Los Angeles, CA, USA.
- 448 Barbraud, C. & Weimerskirch, H. (2006) Antarctic birds breed later in response to climate
449 change. *Proceedings of the National Academy of Sciences*, **103**, 6248–6251.
- 450 Berger, J. (1992) Facilitation of reproductive synchrony by gestation adjustment in gregarious
451 mammals: A new hypothesis. *Ecology*, **73**, 323–329.
- 452 Bêty, J., Gauthier, G. & Giroux, J. (2003) Body condition, migration, and timing of reproduction
453 in snow geese: A test of the condition-dependent model of optimal clutch size. *The*
454 *American Naturalist*, **162**, 110–121.
- 455 Both, C., Artemyev, A.V., Blaauw, B., Cowie, R.J., Dekhuijzen, A.J., Eeva, T., Enemar, A.,
456 Gustafsson, L., Ivankina, E.V., Jarvinen, A., Metcalfe, N.B., Nyholm, N.E.I., Potti, J.,
457 Ravussin, P.-A., Sanz, J.J., Silverin, B., Slater, F.M., Sokolov, L.V., Torok, J., Winkel,
458 W., Wright, J., Zang, H. & Visser, M.E. (2004) Large-scale geographical variation
459 confirms that climate change causes birds to lay earlier. *Proceedings of the Royal Society*
460 *B: Biological Sciences*, **271**, 1657–1662.
- 461 Both, C., van Asch, M., Bijlsma, R.G., van den Burg, A.B. & Visser, M.E. (2009) Climate
462 change and unequal phenological changes across four trophic levels: constraints or

- 463 adaptations? *Journal of Animal Ecology*, **78**, 73–83.
- 464 Bradley, N.L., Leopold, A.C., Ross, J. & Huffaker, W. (1999) Phenological changes reflect
465 climate change in Wisconsin. *Proceedings of the National Academy of Sciences*, **96**,
466 9701–9704.
- 467 Brooks, S.P. & Gelman, A. (1998) General methods for monitoring convergence of iterative
468 simulations. *Journal of Computational and Graphical Statistics*, **7**, 434.
- 469 Burger, J. (1979) Colony size: a test for breeding synchrony in herring gull (*Larus argentatus*)
470 colonies. *The Auk*, 694–703.
- 471 Burr, Z.M., Varpe, Ø., Anker-Nilssen, T., Erikstad, K.E., Descamps, S., Barrett, R.T., Bech, C.,
472 Christensen-Dalsgaard, S., Lorentsen, S.-H., Moe, B., Reiertsen, T.K. & Strøm, H. (2016)
473 Later at higher latitudes: large-scale variability in seabird breeding timing and
474 synchronicity. *Ecosphere*, **7**, e01283.
- 475 Caswell, H. (2011) Beyond R0: demographic models for variability of lifetime reproductive
476 output. *PloS one*, **6**, e20809.
- 477 Clark, J.A., Haseley, A., Van Genderen, G., Hofling, M. & Clum, N.J. (2012) Increasing
478 breeding behaviors in a captive colony of Northern Bald Ibis through conspecific acoustic
479 enrichment. *Zoo Biology*, **31**, 71–81.
- 480 Cleland, E.E., Chuine, I., Menzel, A., Mooney, H.A. & Schwartz, M.D. (2007) Shifting plant
481 phenology in response to global change. *Trends in ecology & evolution*, **22**, 357–365.
- 482 Crawley, M.J. & Akhteruzzaman, M. (1988) Individual variation in the phenology of oak trees

- 483 and its consequences for herbivorous insects. *Functional Ecology*, **2**, 409.
- 484 Cushing, D. (1974) The natural regulation of fish populations. *Sea Fisheries Research* (ed F.R.
485 Harden-Jones), pp. 399–412, John Wiley and Sons, New York, NY, USA.
- 486 D’Agostino, R.B. (1970) Transformation to normality of the null distribution of g_1 . *Biometrika*,
487 **57**, 679.
- 488 Danchin, E. (1988) Social interactions in kittiwake colonies: social facilitation and/or favourable
489 social environment. *Animal Behaviour*, **36**, 443–451.
- 490 Darling, F.F. (1938) *Bird Flocks and the Breeding Cycle*. Cambridge University Press,
491 Cambridge, UK.
- 492 Dawson, A. (2008) Control of the annual cycle in birds: endocrine constraints and plasticity in
493 response to ecological variability. *Philosophical Transactions of the Royal Society B:*
494 *Biological Sciences*, **363**, 1621–1633.
- 495 Deviche, P., Small, T., Sharp, P. & Tsutsui, K. (2006) Control of luteinizing hormone and
496 testosterone secretion in a flexibly breeding male passerine, the Rufous-winged Sparrow,
497 *Aimophila carpalis*. *General and Comparative Endocrinology*, **149**, 226–235.
- 498 Dunne, J.A., Harte, J. & Taylor, K.J. (2003) Subalpine meadow flowering phenology responses
499 to climate change: integrating experimental and gradient methods. *Ecological*
500 *Monographs*, **73**, 69–86.
- 501 Durant, J.M., Hjermann, D.O., Anker-Nilssen, T., Beaugrand, G., Myrsetrud, A., Pettorelli, N. &
502 Stenseth, N.C. (2005) Timing and abundance as key mechanisms affecting trophic

- 503 interactions in variable environments. *Ecology Letters*, **8**, 952–958.
- 504 Elzinga, J.A., Atlan, A., Biere, A., Gigord, L., Weis, A.E. & Bernasconi, G. (2007) Time after
505 time: flowering phenology and biotic interactions. *Trends in Ecology & Evolution*, **22**,
506 432–439.
- 507 Emmerson, L., Pike, R. & Southwell, C. (2011) Reproductive consequences of environment-
508 driven variation in Adélie penguin breeding phenology. *Marine Ecology Progress Series*,
509 **440**, 203–216.
- 510 Galloway, L.F. & Burgess, K.S. (2012) Artificial selection on flowering time: influence on
511 reproductive phenology across natural light environments. *Journal of Ecology*, **100**, 852–
512 861.
- 513 Hatchwell, B.J. (1991) An experimental study of the effects of timing of breeding on the
514 reproductive success of common guillemots (*Uria aalge*). *The Journal of Animal*
515 *Ecology*, **60**, 721.
- 516 Hay, R.K.M. (1990) The influence of photoperiod on the dry matter production of grasses and
517 cereals. *New Phytologist*, **116**, 233–254.
- 518 Helm, B. (2009) Geographically distinct reproductive schedules in a changing world: costly
519 implications in captive Stonechats. *Integrative and Comparative Biology*, **49**, 563–579.
- 520 Hinke, J., Polito, M., Reiss, C., Trivelpiece, S. & Trivelpiece, W. (2012) Flexible reproductive
521 timing can buffer reproductive success of *Pygoscelis* spp. penguins in the Antarctic
522 Peninsula region. *Marine Ecology Progress Series*, **454**, 91–104.

- 523 Ims, R.A. (1990a) On the adaptive value of reproductive synchrony as a predator-swamping
524 strategy. *The American Naturalist*, **136**, 485–498.
- 525 Ims, R.A. (1990b) The ecology and evolution of reproductive synchrony. *Trends in ecology &*
526 *evolution*, **5**, 135–140.
- 527 Kerby, J.T., Wilmer, C.C. & Post, E. (2012) Climate change, phenology and the nature of
528 consumer–resource interactions: advancing the match/mismatch hypothesis. *Trait-*
529 *mediated Indirect Interactions: Ecological and Evolutionary Perspectives* (eds T.
530 Ohgushi, O.J. Schmitz, and R.D. Holt), pp. 508–525. Cambridge University Press,
531 Cambridge, UK.
- 532 Lande, R., Engen, S. & Saether, B. (2003) *Stochastic Population Dynamics in Ecology and*
533 *Conservation*. Oxford University Press, Oxford, UK.
- 534 Lambrechts, M.M., Perret, P., Maistre, M. & Blondel, J. (1999) Do experiments with captive
535 non–domesticated animals make sense without population field studies? A case study
536 with blue tits’ breeding time. *Proceedings of the Royal Society of London B: Biological*
537 *Sciences*, **266**, 1311–1315.
- 538 Lehrman, D.S. & Friedman, M. (1969) Auditory stimulation of ovarian activity in the ring dove
539 (*Streptopelia risoria*). *Animal Behaviour*, **17**, 494–497.
- 540 Leitner, S., Van’t Hof, T.J. & Gahr, M. (2003) Flexible reproduction in wild canaries is
541 independent of photoperiod. *General and comparative endocrinology*, **130**, 102–108.
- 542 Lewis, S., Nussey, D.H., Wood, A.G., Croxall, J.P. & Phillips, R.A. (2012) Intrinsic

- 543 determinants of a population trend in timing of breeding in the wandering albatross.
544 *Oikos*, **121**, 2061–2071.
- 545 Ludwig, G.X., Alatalo, R.V., Helle, P., Linden, H., Lindstrom, J. & Siitari, H. (2006) Short- and
546 long-term population dynamical consequences of asymmetric climate change in black
547 grouse. *Proceedings of the Royal Society B: Biological Sciences*, **273**, 2009–2016.
- 548 Lynch, H.J., Fagan, W.F., Naveen, R., Trivelpiece, S.G. & Trivelpiece, W.Z. (2012) Differential
549 advancement of breeding phenology in response to climate may alter staggered breeding
550 among sympatric pygoscelid penguins. *Marine Ecology Progress Series*, **454**, 135–145.
- 551 McClintock, M.K. (1978) Estrous synchrony and its mediation by airborne chemical
552 communication (*Rattus norvegicus*). *Hormones and behavior*, **10**, 264–276.
- 553 McLean, N., Lawson, C.R., Leech, D.I. & van de Pol, M. (2016) Predicting when climate-driven
554 phenotypic change affects population dynamics. *Ecology Letters*, **19**, 595–608.
- 555 Miller-Rushing, A.J., Hoyer, T.T., Inouye, D.W. & Post, E. (2010) The effects of phenological
556 mismatches on demography. *Philosophical Transactions of the Royal Society B:*
557 *Biological Sciences*, **365**, 3177–3186.
- 558 Miller-Rushing, A.J., Inouye, D.W. & Primack, R.B. (2008) How well do first flowering dates
559 measure plant responses to climate change? The effects of population size and sampling
560 frequency. *Journal of Ecology*, **96**, 1289–1296.
- 561 Plummer, M. (2003) *JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs*
562 *Sampling*. Proceedings of the Third International Workshop on Distributed Statistical

- 563 Computing. R Project for Statistical Computing, Vienna, Austria.
- 564 Plummer, M. (2013) *rjags: Bayesian graphical models using MCMC*. [https://CRAN.R-](https://CRAN.R-project.org/package=rjags)
- 565 project.org/package=rjags
- 566 Post, E. & Forchhammer, M.C. (2008) Climate change reduces reproductive success of an Arctic
- 567 herbivore through trophic mismatch. *Philosophical Transactions of the Royal Society B:*
- 568 *Biological Sciences*, **363**, 2367–2373.
- 569 R Core Team. (2016) *R: A Language and Environment for Statistical Computing*. R Foundation
- 570 for Statistical Computing, Vienna, Austria. www.r-project.org
- 571 Reynolds, S.J., Schoech, S.J. & Bowman, R. (2003) Nutritional quality of prebreeding diet
- 572 influences breeding performance of the Florida scrub-jay. *Oecologia*, **134**, 308–316.
- 573 Scott, M.P. (1986) The timing and synchrony of seasonal breeding in the marsupial, *Antechinus*
- 574 *stuartii*: Interaction of environmental and social cues. *Journal of Mammalogy*, **67**, 551–
- 575 560.
- 576 Setiawan, A.N., Davis, L.S., Darby, J.T., Lokman, P.M., Young, G., Blackberry, M.A., Cannell,
- 577 B.L. & Martin, G.B. (2007) Effects of artificial social stimuli on the reproductive
- 578 schedule and hormone levels of yellow-eyed penguins (*Megadyptes antipodes*).
- 579 *Hormones and Behavior*, **51**, 46–53.
- 580 Sinclair, A.R.E., Mduma, S.A.R. & Arcese, P. (2000) What determines phenology and
- 581 synchrony of ungulate breeding in Serengeti? *Ecology*, **81**, 2100.
- 582 Sparks, T.H., Bairlein, F., Bojarinova, J.G., Huppopp, O., Lehikoinen, E.A., Rainio, K., Sokolov,

- 583 L.V. & Walker, D. (2005) Examining the total arrival distribution of migratory birds.
584 *Global Change Biology*, **11**, 22–30.
- 585 Temte, J.L. & Temte, J. (1993) Photoperiod defines the phenology of birth in captive California
586 sea lions. *Marine Mammal Science*, **9**, 301–308.
- 587 Thackeray, S.J., Henrys, P.A., Hemming, D., Bell, J.R., Botham, M.S., Burthe, S., Helaouet, P.,
588 Johns, D.G., Jones, I.D., Leech, D.I., Mackay, E.B., Massimino, D., Atkinson, S., Bacon,
589 P.J., Brereton, T.M., Carvalho, L., Clutton-Brock, T.H., Duck, C., Edwards, M., Elliott,
590 J.M., Hall, S.J.G., Harrington, R., Pearce-Higgins, J.W., Høye, T.T., Kruuk, L.E.B.,
591 Pemberton, J.M., Sparks, T.H., Thompson, P.M., White, I., Winfield, I.J. & Wanless, S.
592 (2016) Phenological sensitivity to climate across taxa and trophic levels. *Nature*, **535**,
593 241–245.
- 594 Thomson, J.D. (1980) Skewed flowering distributions and pollinator attraction. *Ecology*, **61**,
595 572–579.
- 596 Torti, V.M. & Dunn, P.O. (2005) Variable effects of climate change on six species of North
597 American birds. *Oecologia*, **145**, 486–495.
- 598 Varpe, Ø., Jørgensen, C., Tarling, G.A. & Fiksen, Ø. (2007) Early is better: seasonal egg fitness
599 and timing of reproduction in a zooplankton life-history model. *Oikos*, **116**, 1331–1342.
- 600 Visser, M.E. & Both, C. (2005) Shifts in phenology due to global climate change: the need for a
601 yardstick. *Proceedings of the Royal Society B: Biological Sciences*, **272**, 2561–2569.
- 602 Visser, M.E., Holleman, L.J.M. & Caro, S.P. (2009) Temperature has a causal effect on avian

- 603 timing of reproduction. *Proceedings of the Royal Society B: Biological Sciences*, **276**,
604 2323–2331.
- 605 Visser, M.E., te Marvelde, L. & Lof, M.E. (2012) Adaptive phenological mismatches of birds
606 and their food in a warming world. *Journal of Ornithology*, **153**, 75–84.
- 607 Vonesh, J.R. (2005) Egg predation and predator-induced hatching plasticity in the African reed
608 frog, *Hyperolius spinigularis*. *Oikos*, **110**, 241–252.
- 609 Waas, J.R. (1988) Acoustic displays facilitate courtship in little blue penguins, *Eudyptula minor*.
610 *Animal Behaviour*, **36**, 366–371.
- 611 Waas, J.R. (1991) Do little blue penguins signal their intentions during aggressive interactions
612 with strangers? *Animal Behaviour*, **41**, 375–382.
- 613 Waas, J. (1995) Social stimulation and reproductive schedules: does the acoustic environment
614 influence the egg-laying schedule in penguin colonies. *Penguins: Ecology and*
615 *Management* (ed. P. Dann, I. Norman & P. Peilly), 111–137.
- 616 Waas, J.R., Caulfield, M., Colgan, P.W. & Boag, P.T. (2000) Colony sound facilitates sexual and
617 agonistic activities in royal penguins. *Animal Behaviour*, **60**, 77–84.
- 618 Waas, J.R., Colgan, P.W. & Boag, P.T. (2005) Playback of colony sound alters the breeding
619 schedule and clutch size in zebra finch (*Taeniopygia guttata*) colonies. *Proceedings of*
620 *the Royal Society B: Biological Sciences*, **272**, 383–388.
- 621 Wickham, H. (2009) *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York,
622 NY, USA.

- 623 Wilson, W. (2013) A deeper statistical examination of arrival dates of migratory breeding birds
624 in relation to global climate change. *Biology*, **2**, 742–754.
- 625 Winder, M. & Schindler, D.E. (2004) Climate change uncouples trophic interactions in an
626 aquatic ecosystem. *Ecology*, **85**, 2100–2106.
- 627 Young, E. (1994) *Skua and Penguin: Predator and Prey*. Cambridge University Press,
628 Cambridge, UK.
- 629 Youngflesh, C. (2016) *MCMCvis: Tools to visualize, Manipulate, and Summarize MCMC*
630 *output*. <https://CRAN.R-project.org/package=MCMCvis>
- 631 Youngflesh, C., Jenouvrier, S., Hinke, J.T., DuBois, L., St. Leger, J., Trivelpiece, W.Z.,
632 Trivelpiece, S.G. & Lynch, H.J. (2017). Data from: Rethinking ‘normal’: The role of
633 stochasticity in the phenology of a synchronously breeding seabird. Dryad Digital
634 Repository. <https://doi.org/10.5061/dryad.23sv1>
- 635 Youngflesh, C. Jenouvrier, S., Li, Y., Ji, R., Ainley, D.G., Ballard, G., Barbraud, C., Delord, K.,
636 Dugger, K.M., Emmerson, L.M., Fraser, W.R., Hinke, J.T., Lyver, P.O’B., Olmastroni,
637 S., Southwell, C., Trivelpiece, S.G., Trivelpiece, W.Z., Lynch, H.J. (2017) Circumpolar
638 analysis of the Adélie penguin reveals the importance of environmental variability in
639 phenological mismatch. *Ecology*, **98**, 940-951.
- 640 Zerbe, P., Clauss, M., Codron, D., Bingaman Lackey, L., Rensch, E., Streich, J.W., Hatt, J.-M. &
641 Müller, D.W.H. (2012) Reproductive seasonality in captive wild ruminants: implications
642 for biogeographical adaptation, photoperiodic control, and life history. *Biological*
643 *Reviews*, **87**, 965–990.

Appendix S1: Study R/JAGS code with results

Rethinking ‘normal’: The role of stochasticity in the phenology of a synchronously breeding seabird

Casey Youngflesh, Stephanie Jenouvrier, Jefferson T. Hinke, Lauren DuBois, Judy St. Leger, Wayne Z. Trivelpiece, Susan G. Trivelpiece, Heather J. Lynch

Initial set up

Packages

```
#install packages if they don't exist - then load them
if('pacman' %in% rownames(installed.packages()) == FALSE)
{
  install.packages('pacman')
}

pacman::p_load(ggplot2,
               reshape2,
               dplyr,
               rjags,
               moments,
               MCMCvis)
```

Inter-annual variation in CID

Load data

```
setwd('Data')

captive_data <- read.csv('Captive_CID.csv', header = TRUE)
wild_data <- read.csv('Wild_CID.csv', header = TRUE)
```

Function to calculate intra-annual median and var

```
srt.fun <- function(IN)
{
  #IN <- SD_lay
  yrs <- unique(IN$YEAR)

  OUT <- c()
  for(i in min(yrs):max(yrs))
  {
    #i <- 1993
    temp <- filter(IN, YEAR == i)
    t_md <- median(temp$J_CID)
    t_var <- var(temp$J_CID)
    temp2 <- data.frame(YEAR = i, MEDIAN = t_md, VAR = t_var)
    OUT <- rbind(OUT, temp2)
  }
}
```

```
  return(OUT)
}
```

Function to calculate SE of variance

$$\sigma_{s^2} = s^2 * \sqrt{2/(n-1)}$$

```
se_var <- function(data)
{
  OUT <- var(data)*sqrt(2/(length(data)-1))
  return(OUT)
}
```

Captive inter-annual variance - $var(y_{i.})$

```
captive_md_sd <- srt.fun(captive_data)
```

```
(c_med <- var(captive_md_sd$MEDIAN))
```

```
## [1] 15.83288
```

Captive standard error of variance

```
se_var(captive_md_sd$MEDIAN)
```

```
## [1] 4.668862
```

Wild inter-annual variance - $var(y_{i.})$

```
wild_md_sd <- srt.fun(wild_data)
```

```
(w_med <- var(wild_md_sd$MEDIAN))
```

```
## [1] 13.49003
```

Wild standard error of variance

```
se_var(wild_md_sd$MEDIAN)
```

```
## [1] 3.741461
```

Inter-annual variation in CID from literature

Inter-annual variation in CID in the Adélie penguin population are of a similar magnitude to those seen in other bird species.

```
setwd('Data')
```

```
lit_data <- read.csv('Lit_data.csv', header = TRUE)
```

Both et al. 2009 (Fig. 2) - Blue tit, Great tit, Pied flycatcher, Coal tit, Sparrowhawk

#detrrend because we are interested in the interannual variance - trends would conflate this estimate

```
fit_BT_B_2009 <- lm(lit_data$BT_B_2009 ~ lit_data$YEAR)
```

```
res_BT_B_2009 <- residuals(fit_BT_B_2009)
```

```
v_BT_B_2009 <- var(res_BT_B_2009)
```

```
fit_GT_B_2009 <- lm(lit_data$GT_B_2009 ~ lit_data$YEAR)
```

```
res_GT_B_2009 <- residuals(fit_GT_B_2009)
```

```
v_GT_B_2009 <- var(res_GT_B_2009)
```

```

fit_PF_B_2009 <- lm(lit_data$PF_B_2009 ~ lit_data$YEAR)
res_PF_B_2009 <- residuals(fit_PF_B_2009)
v_PF_B_2009 <- var(res_PF_B_2009)

fit_CT_B_2009 <- lm(lit_data$CT_B_2009 ~ lit_data$YEAR)
res_CT_B_2009 <- residuals(fit_CT_B_2009)
v_CT_B_2009 <- var(res_CT_B_2009)

#no need to detrend - no significant change
v_SH_B_2009 <- var(lit_data$SH_B_2009, na.rm = TRUE)

```

Valtonen et al. 2017 (Fig. 2) - Common redstart, Great tit, Pied flycatcher

```

#no need to detrend - no significant change
v_CR_V_2017 <- var(lit_data$CR_V_2017, na.rm = TRUE)

v_GT_V_2017 <- var(lit_data$GT_V_2017, na.rm = TRUE)

v_PF_V_2017 <- var(lit_data$PF_V_2017, na.rm = TRUE)

```

Species	Publication	Metric	Inter-annual Variance
Blue tit	Both et al. 2009	Hatch date	11.94
Great tit	Both et al. 2009	Hatch date	12.82
Pied flycatcher	Both et al. 2009	Hatch date	4.22
Coal tit	Both et al. 2009	Hatch date	25.11
Sparrowhawk	Both et al. 2009	Hatch date	9.74
Common redstart	Valtonen et al. 2017	Clutch initiation date	10.65
Great tit	Valtonen et al. 2017	Clutch initiation date	15.62
Pied flycatcher	Valtonen et al. 2017	Clutch initiation date	10.26
Adélie penguin (captive)	This publication	Clutch initiation date	15.83
Adélie penguin (wild)	This publication	Clutch initiation date	13.49

Adélie penguins exhibit degrees of inter-annual variance in phenology comparable to other species.

Plot median CID over time - Captive

```

plt_t <- data.frame(YEAR = 1992:2015, MD_CID = captive_md_sd$MEDIAN)
plt <- melt(plt_t, id = 'YEAR')

#CAPTIVE PLOT
ggplot(plt, aes(YEAR, value)) +
  geom_line(size = 1.2, col = 'red') +
  theme_bw() +
  ggtitle('Captive penguin breeding phenology') +
  xlab('Year') +
  ylab('CID (days from Sep 30)') +
  coord_cartesian(xlim = c(1990, 2015)) +
  coord_cartesian(ylim = c(25, 45)) +
  scale_x_continuous(breaks = seq(1990, 2015, by = 5)) +
  scale_y_continuous(breaks = c(25, 30, 35, 40, 45)) +

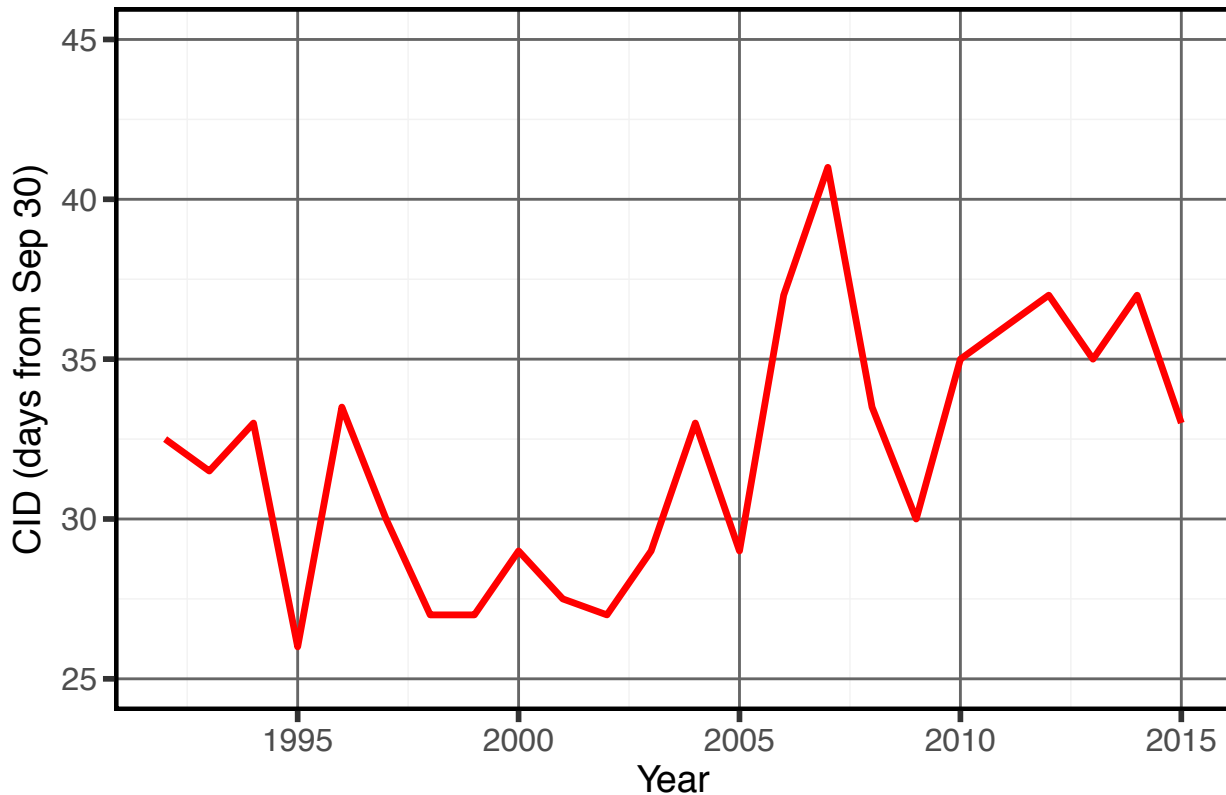
```

```

theme(
  axis.text = element_text(size = 12), #axis label size
  axis.title = element_text(size = 14),
  panel.grid.major = element_line(color = 'gray40'), #lower # is darker
  panel.grid.minor = element_line(color = 'gray95'),
  panel.background = element_blank(),
  panel.border = element_rect(fill = NA, color= 'black', size = 1.5),
  axis.ticks.length= unit(0.15, 'cm'), #length of axis tick
  axis.ticks = element_line(size = 1)
)

```

Captive penguin breeding phenology



Plot median CID over time - Wild

```

plt_t2 <- data.frame(YEAR = 1986:2012, MD_CID = wild_md_sd$MEDIAN)
plt2 <- melt(plt_t2, id = 'YEAR')

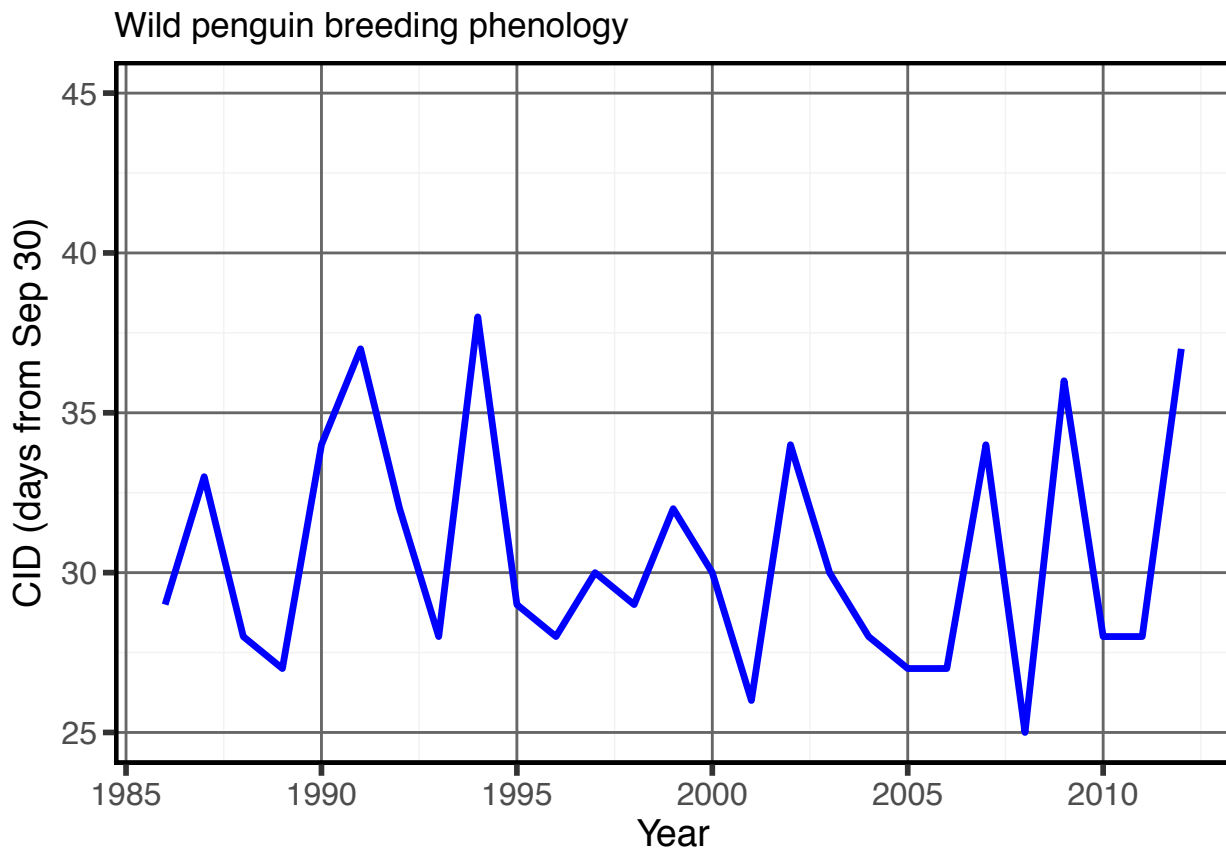
#WILD PLOT
ggplot(plt2, aes(YEAR, value)) +
  geom_line(size = 1.2, col = 'blue') +
  theme_bw() +
  ggtitle('Wild penguin breeding phenology') +
  xlab('Year') +
  ylab('CID (days from Sep 30)') +
  coord_cartesian(xlim = c(1984, 2012)) +

```

```

coord_cartesian(ylim = c(25, 45)) +
scale_x_continuous(breaks = seq(1985, 2012, by = 5)) +
scale_y_continuous(breaks = c(25, 30, 35, 40, 45)) +
theme(
  axis.text = element_text(size = 12), #axis label size
  axis.title = element_text(size = 14),
  panel.grid.major = element_line(color = 'gray40'), #lower # is darker
  panel.grid.minor = element_line(color = 'gray95'),
  panel.background = element_blank(),
  panel.border = element_rect(fill = NA, color = 'black', size = 1.5),
  axis.ticks.length = unit(0.15, 'cm'), #length of axis tick
  axis.ticks = element_line(size = 1)
)

```



Determine which female is the first to lay in each year (including ties for first)

```

FEM <- c()
for (i in 1992:2015)
{
  #i <- 1997
  temp <- filter(captive_data, YEAR == i)
  pos <- which(temp$J_CID == min(temp$J_CID))
  t_data <- temp[pos,]
}

```

```
FEM <- rbind(FEM, t_data)
}
```

16 different 'leaders' in 24 years

```
length(unique(FEM$FEMALE_ID))
```

```
## [1] 16
```

Intra-annual variation in CID

```
#t-test to determine if intra-annual variation differs between captive and wild populations
t.test(captive_md_sd$VAR, wild_md_sd$VAR)
```

```
##
## Welch Two Sample t-test
##
## data: captive_md_sd$VAR and wild_md_sd$VAR
## t = 4.9893, df = 28.488, p-value = 2.733e-05
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## 19.51181 46.65717
## sample estimates:
## mean of x mean of y
## 50.75654 17.67205
```

Number of breeders

```
len_fun <- function(IN)
{
  yrs <- range(IN$YEAR)

  LEN <- c()
  for(i in yrs[1]:yrs[2])
  {
    #i <- 1992
    temp <- filter(IN, YEAR == i)
    t1 <- dim(temp)[1]
    tb <- c(i, t1)
    LEN <- rbind(LEN, tb)
  }

  return(LEN)
}
```

```
#range
range(len_fun(captive_data)[,2])
```

```
## [1] 12 37
```

Hierarchical model - captive population

JAGS model

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma * AGE_{ij} + \epsilon_{ij}$$

$$\alpha_i \sim N(0, \sigma_{year}^2)$$

$$\beta_j \sim N(0, \sigma_{individual}^2)$$

$$\epsilon_{ij} \sim N(0, \sigma_{model}^2)$$

```
setwd('Data')

AGE_mat <- read.csv('AGE_mat.csv', header= TRUE)
CID_mat <- read.csv('CID_mat.csv', header= TRUE)

DATA <- list(
  y = CID_mat,
  yr = 1:NCOL(CID_mat),
  ind = 1:NROW(CID_mat),
  age = AGE_mat,
  N = NCOL(CID_mat), #columns are year in matrix
  M = NROW(CID_mat)) #rows are individuals

#-----#
#model

#alpha = YEAR - random
#beta = INDIVIDUAL - random
#gamma = AGE - fixed

setwd('JAGS')

{
sink("captive.jags")

cat("
  model {

    for(t in 1:N)
    {
      for(i in 1:M)
      {
        y[i,t] ~ dnorm(mu.g[i,t], tau)
        mu.g[i,t] <- mu + alpha[yr[t]] + beta[ind[i]] + gamma*age[i,t]
      }
    }

    #priors
```

```

#year
for(t in 1:N)
{
alpha[t] ~ dnorm(0, tau.year)
}

#individual
for(i in 1:M)
{
beta[i] ~ dnorm(0, tau.ind)
}

#mu, gamma, and tau
mu ~ dnorm(0, 0.001)
gamma ~ dnorm(0, 0.001)
tau ~ dgamma(0.01, 0.01)
var.model <- 1/tau #convert from precision to variance

#hyperparameters
tau.year ~ dgamma(0.01, 0.01)
var.year <- 1/tau.year
tau.ind ~ dgamma(0.01, 0.01)
var.ind <- 1/tau.ind

}","fill = TRUE)

sink()
}

```

Run model

```

#-----#
#Starting values

Inits <- function() {list(alpha = rep(rnorm(1),
                                   ncol(CID_mat)),
                          beta = rep(rnorm(1),
                                   nrow(CID_mat)),
                          tau = rgamma(1, 1),
                          mu = rnorm(1),
                          gamma = rnorm(1),
                          tau.year = rgamma(1, 1),
                          tau.ind = rgamma(1, 1))}

#-----#
#Parameters to track

Pars <- c('alpha', 'beta', 'gamma', 'var.year', 'var.ind', 'var.model')

# Inputs for MCMC -----#

```



```

n_adapt <- 5000 # number for initial adapt
n_burn <- 40000 # number burnin
n_draw <- 10000 # number of final draws to make
n_thin <- 2 # thinning rate
n_chain <- 3 # number of chains

Rhat_max <- 1.1 # max allowable Rhat (close to 1 = convergence)
n_max <- 1e7 # max allowable iterations

#-----#
#Run model

jm = jags.model(data = DATA,
                file = "JAGS/captive.jags",
                inits = Inits,
                n.chains = 3,
                n.adapt = n_adapt)

update(jm, n.iter = n_burn)

out = coda.samples(jm,
                  n.iter = n_draw,
                  variable.names = Pars,
                  thin = n_thin)

#extra draws to ensure convergence
n_total <- n_burn + n_draw
n_extra <- 0
while(max(MCMCsummary(out)[,6]) > Rhat_max &
      n_total < n_max)
{
  out <- update(out,
                n.iter = n_draw,
                n.chains = n_chain,
                n.thin = n_thin)

  n_extra <- n_extra + n_draw
  n_total <- n_total + n_draw
}

n_final <- n_draw/n_thin

```

Inferences were derived from 5000 samples drawn following an adaptation period of 5000 draws, and a burn-in period of 4×10^4 draws using 3 chains and a thinning rate of 2.

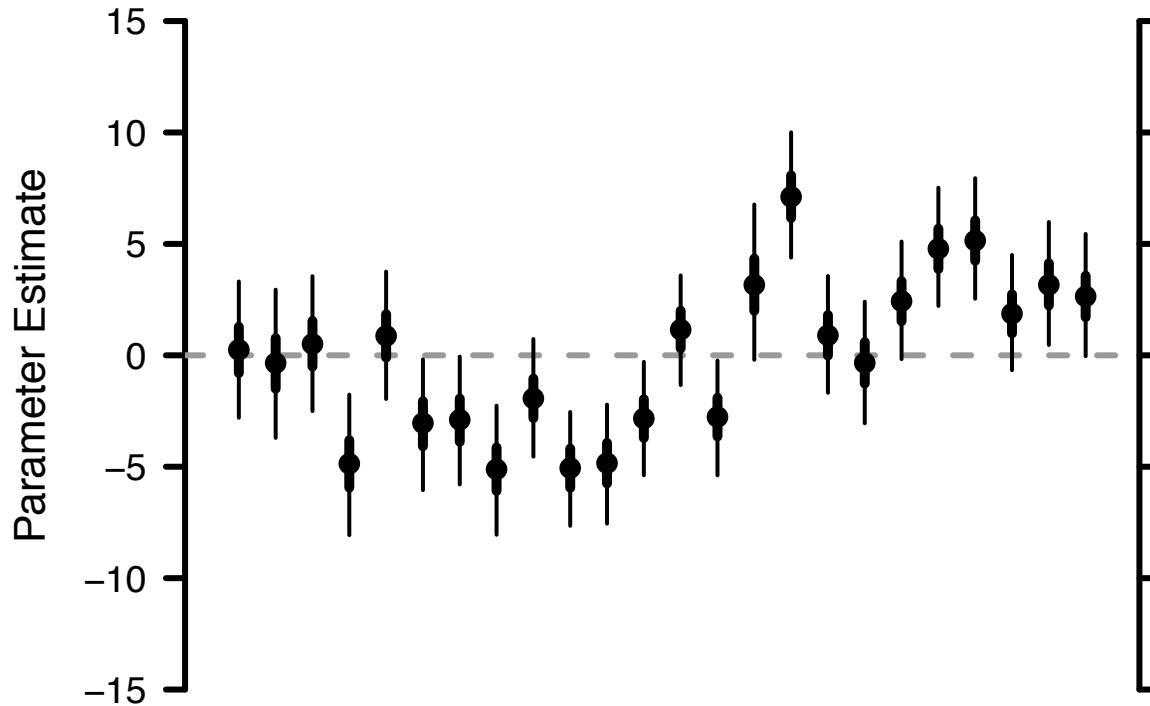
Year effect (alpha)

```

MCMCplot(out,
          params = 'alpha',

```

```
rank = FALSE,
labels = NULL,
horiz = FALSE,
ref_ovl = FALSE,
ylim = c(-15, 15))
```



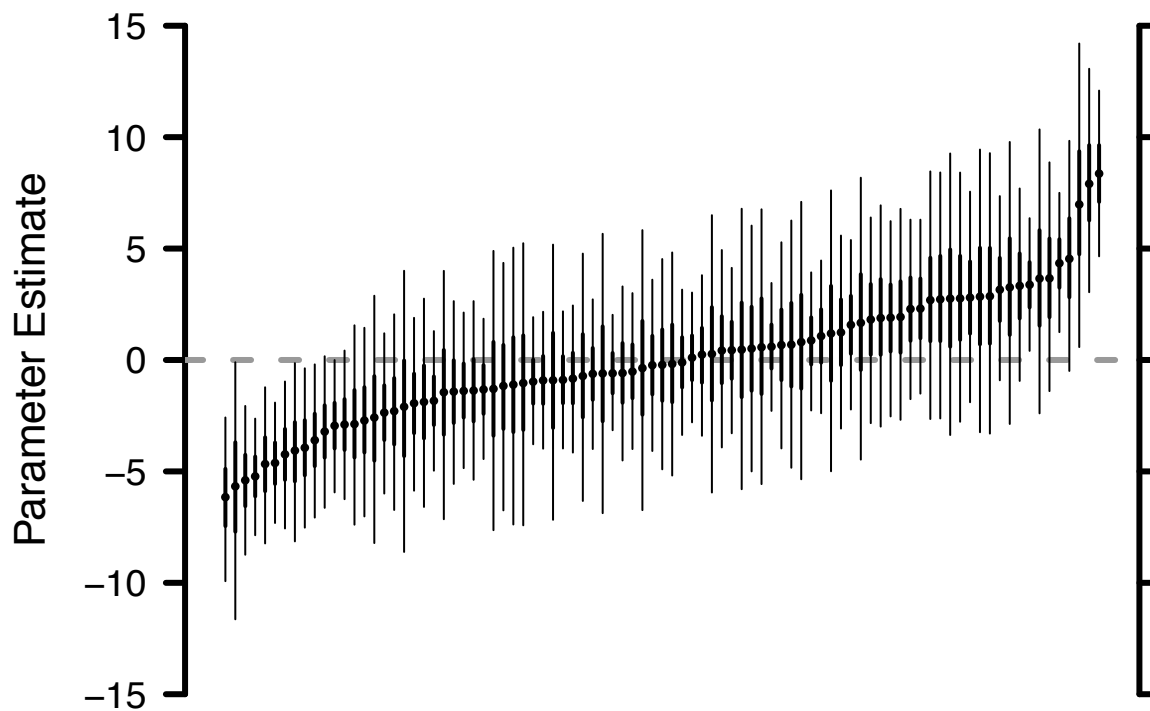
```
MCMCsummary(out,
  params = 'alpha')
```

##		mean	sd	2.5%	50%	97.5%	Rhat
##	alpha[1]	0.24	1.55	-2.80	0.25	3.31	1
##	alpha[2]	-0.36	1.69	-3.70	-0.35	2.95	1
##	alpha[3]	0.52	1.54	-2.50	0.51	3.55	1
##	alpha[4]	-4.87	1.60	-8.07	-4.86	-1.76	1
##	alpha[5]	0.87	1.46	-1.96	0.87	3.76	1
##	alpha[6]	-3.07	1.50	-6.05	-3.04	-0.18	1
##	alpha[7]	-2.90	1.45	-5.80	-2.89	-0.06	1
##	alpha[8]	-5.12	1.47	-8.05	-5.12	-2.26	1
##	alpha[9]	-1.93	1.33	-4.55	-1.93	0.73	1
##	alpha[10]	-5.07	1.31	-7.65	-5.06	-2.54	1
##	alpha[11]	-4.85	1.36	-7.55	-4.84	-2.21	1
##	alpha[12]	-2.84	1.30	-5.38	-2.84	-0.29	1
##	alpha[13]	1.15	1.25	-1.33	1.15	3.59	1
##	alpha[14]	-2.78	1.30	-5.39	-2.77	-0.23	1
##	alpha[15]	3.18	1.76	-0.20	3.16	6.77	1
##	alpha[16]	7.13	1.43	4.39	7.11	10.01	1
##	alpha[17]	0.90	1.34	-1.68	0.90	3.56	1
##	alpha[18]	-0.33	1.39	-3.05	-0.34	2.41	1
##	alpha[19]	2.44	1.34	-0.17	2.43	5.11	1
##	alpha[20]	4.80	1.35	2.22	4.78	7.52	1
##	alpha[21]	5.16	1.37	2.54	5.15	7.95	1

```
## alpha[22]  1.87 1.31 -0.67  1.87  4.50  1
## alpha[23]  3.19 1.40  0.47  3.16  5.98  1
## alpha[24]  2.65 1.39 -0.04  2.65  5.45  1
```

Individual effect (beta)

```
MCMCplot(out,
  params = 'beta',
  rank = TRUE,
  labels = NULL,
  horiz = FALSE,
  thick_sz = 2,
  thin_sz = 1,
  med_sz = .6,
  ref_ovl = FALSE,
  ylim = c(-15, 15))
```



```
MCMCsummary(out,
  params = 'beta')
```

##	mean	sd	2.5%	50%	97.5%	Rhat
## beta[1]	2.28	2.07	-1.76	2.29	6.30	1
## beta[2]	-2.71	2.16	-7.02	-2.72	1.44	1
## beta[3]	-0.36	3.19	-6.74	-0.37	5.83	1
## beta[4]	-2.96	1.52	-5.94	-2.95	0.00	1
## beta[5]	1.74	3.21	-4.47	1.67	8.18	1
## beta[6]	4.60	2.62	-0.49	4.54	9.83	1
## beta[7]	8.37	1.90	4.65	8.36	12.09	1
## beta[8]	-2.18	3.22	-8.61	-2.10	4.00	1
## beta[9]	-0.23	1.96	-4.09	-0.23	3.60	1
## beta[10]	-6.18	1.89	-9.93	-6.16	-2.58	1

```

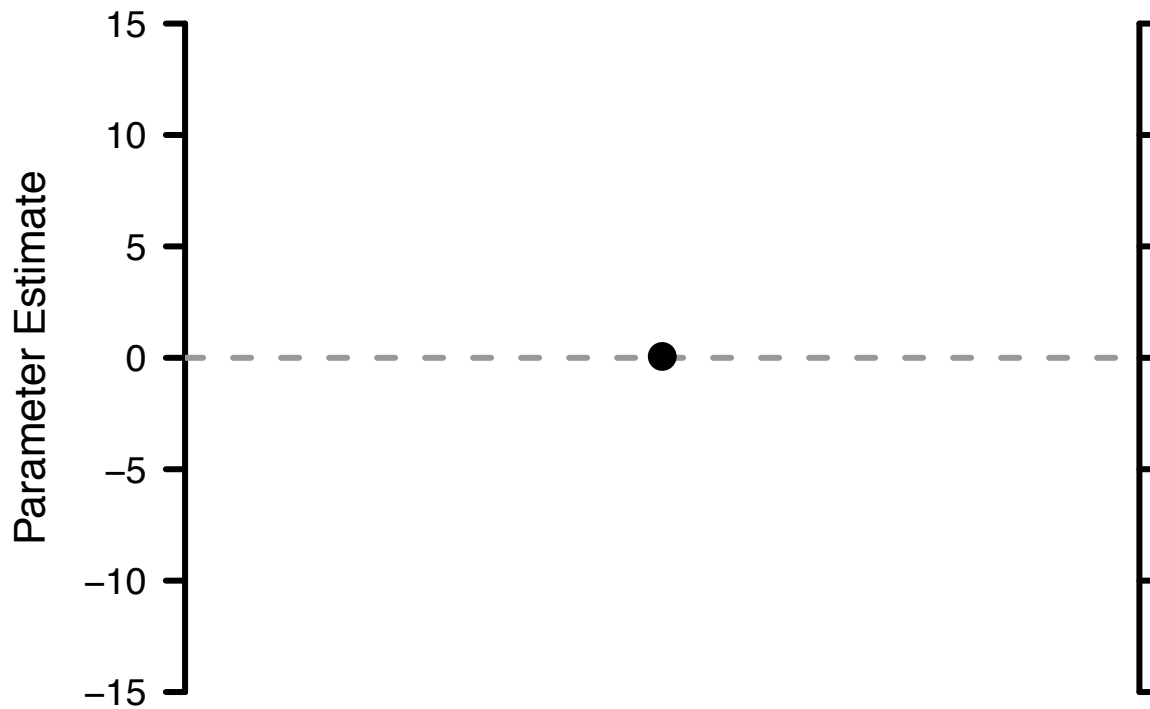
## beta[11] -0.91 1.58 -3.98 -0.92 2.16 1
## beta[12] -1.47 2.83 -7.15 -1.45 4.00 1
## beta[13] 0.86 1.58 -2.27 0.88 3.93 1
## beta[14] -0.51 1.78 -4.00 -0.52 3.00 1
## beta[15] -0.89 1.57 -3.97 -0.89 2.18 1
## beta[16] 1.08 1.76 -2.39 1.07 4.46 1
## beta[17] -1.83 1.61 -4.97 -1.83 1.30 1
## beta[18] -2.88 2.27 -7.39 -2.88 1.56 1
## beta[19] -0.63 1.71 -4.00 -0.62 2.72 1
## beta[20] 2.80 2.40 -1.90 2.80 7.55 1
## beta[21] -3.60 1.75 -7.07 -3.60 -0.20 1
## beta[22] 7.09 3.48 0.58 6.98 14.20 1
## beta[23] 3.69 2.61 -1.40 3.66 8.87 1
## beta[24] 1.22 3.20 -4.99 1.19 7.61 1
## beta[25] -4.63 1.38 -7.32 -4.62 -1.91 1
## beta[26] -0.96 1.45 -3.78 -0.97 1.92 1
## beta[27] 2.80 3.21 -3.37 2.76 9.27 1
## beta[28] -1.12 3.18 -7.38 -1.11 5.04 1
## beta[29] -0.71 2.81 -6.33 -0.72 4.77 1
## beta[30] -1.04 3.21 -7.42 -1.04 5.24 1
## beta[31] -0.58 1.32 -3.15 -0.60 2.03 1
## beta[32] 0.67 2.34 -3.97 0.67 5.28 1
## beta[33] 0.70 2.82 -4.84 0.69 6.26 1
## beta[34] -3.22 1.74 -6.64 -3.21 0.17 1
## beta[35] -1.95 1.98 -5.87 -1.94 1.89 1
## beta[36] -3.95 1.84 -7.52 -3.94 -0.37 1
## beta[37] -4.25 1.69 -7.56 -4.24 -0.97 1
## beta[38] -5.40 1.71 -8.74 -5.40 -2.06 1
## beta[39] -0.93 3.17 -7.17 -0.91 5.18 1
## beta[40] -5.73 2.96 -11.64 -5.67 -0.10 1
## beta[41] -0.10 1.66 -3.37 -0.11 3.16 1
## beta[42] -1.32 1.62 -4.45 -1.33 1.85 1
## beta[43] -4.11 2.01 -8.14 -4.07 -0.14 1
## beta[44] 1.81 2.35 -2.84 1.81 6.40 1
## beta[45] -5.23 1.33 -7.86 -5.22 -2.63 1
## beta[46] 0.26 3.15 -5.95 0.27 6.50 1
## beta[47] 0.10 1.49 -2.80 0.10 3.03 1
## beta[48] 2.89 3.21 -3.30 2.86 9.28 1
## beta[49] 3.71 3.22 -2.39 3.65 10.35 1
## beta[50] 0.22 1.83 -3.40 0.24 3.80 1
## beta[51] 4.34 1.59 1.25 4.34 7.50 1
## beta[52] 0.60 1.47 -2.29 0.60 3.45 1
## beta[53] -2.30 2.23 -6.73 -2.30 2.05 1
## beta[54] -1.32 3.18 -7.64 -1.30 4.89 1
## beta[55] 2.75 2.84 -2.65 2.69 8.46 1
## beta[56] -1.37 2.06 -5.38 -1.37 2.64 1
## beta[57] -2.90 1.69 -6.25 -2.90 0.42 1
## beta[58] 2.79 2.82 -2.77 2.76 8.41 1
## beta[59] 3.37 1.51 0.41 3.38 6.36 1
## beta[60] 2.34 1.99 -1.51 2.31 6.30 1
## beta[61] 2.90 3.24 -3.24 2.84 9.44 1
## beta[62] -1.42 2.10 -5.56 -1.42 2.64 1
## beta[63] -1.20 2.83 -6.75 -1.16 4.36 1
## beta[64] 0.43 1.90 -3.29 0.44 4.13 1

```

```
## beta[65]  1.91 2.55 -2.99  1.88  6.94  1
## beta[66] -2.62 2.83 -8.22 -2.58  2.88  1
## beta[67] -1.38 1.79 -4.85 -1.39  2.13  1
## beta[68]  0.47 3.19 -5.80  0.46  6.79  1
## beta[69]  3.33 3.23 -2.87  3.26  9.78  1
## beta[70]  7.95 2.53  3.04  7.91 13.07  1
## beta[71]  0.45 2.24 -3.92  0.42  4.93  1
## beta[72]  3.34 2.17 -0.94  3.33  7.70  1
## beta[73] -0.16 2.57 -5.19 -0.18  4.83  1
## beta[74] -1.89 2.40 -6.59 -1.88  2.75  1
## beta[75] -0.63 3.19 -6.87 -0.60  5.66  1
## beta[76]  0.51 2.81 -5.00  0.51  6.03  1
## beta[77] -0.22 2.40 -4.91 -0.22  4.53  1
## beta[78]  1.96 2.41 -2.69  1.93  6.78  1
## beta[79]  2.77 2.83 -2.62  2.72  8.42  1
## beta[80] -4.68 1.79 -8.23 -4.67 -1.22  1
## beta[81] -0.82 1.68 -4.15 -0.84  2.45  1
## beta[82]  1.88 2.21 -2.53  1.90  6.23  1
## beta[83]  1.24 2.20 -3.08  1.24  5.58  1
## beta[84]  0.82 3.16 -5.35  0.80  7.10  1
## beta[85]  3.17 2.10 -0.91  3.15  7.35  1
## beta[86] -2.37 1.84 -5.99 -2.36  1.20  1
## beta[87]  0.61 3.16 -5.57  0.57  6.76  1
## beta[88]  1.58 1.94 -2.22  1.57  5.38  1
## beta[89] -0.58 2.00 -4.51 -0.59  3.30  1
```

Age effect (gamma)

```
MCMCplot(out,
  params = 'gamma',
  rank = TRUE,
  labels = NULL,
  horiz = FALSE,
  med_sz = 2,
  ref_ovl = FALSE,
  ylim = c(-15, 15))
```



```
MCMCsummary(out,
  params = 'gamma')
```

```
##      mean  sd  2.5%  50% 97.5% Rhat
## gamma 0.06 0.04 -0.03 0.06  0.15   1
```

Variance estimates (inverse precision)

```
MCMCsummary(out,
  params = c('var.year', 'var.ind', 'var.model'))
```

```
##          mean  sd  2.5%  50% 97.5% Rhat
## var.year 14.83 5.50  7.16 13.81 28.43   1
## var.ind  14.09 3.80  7.95 13.66 22.74   1
## var.model 35.05 2.26 30.92 34.96 39.75   1
```

Skewness test and plots

Captive

```
#scale and aggregate data

sc_agg_fun <- function(IN)
{
  yrs <- range(IN$YEAR)

  OUT <- c()
  for(i in yrs[1]:yrs[2])
```

```

{
  temp <- filter(IN, YEAR == i)
  s_CID <- scale(temp$J_CID, scale = TRUE)

  t.out <- cbind(temp, s_CID)
  OUT <- rbind(OUT, t.out)
}
return(OUT)
}

cap_sk <- sc_agg_fun(captive_data)

```

Skew determined using D'Agostino test

```

skew_captive <- agostino.test(cap_sk$s_CID)$statistic[1]
agostino.test(cap_sk$s_CID)

```

```

##
## D'Agostino skewness test
##
## data: cap_sk$s_CID
## skew = 0.54072, z = 5.24950, p-value = 1.525e-07
## alternative hypothesis: data have a skewness

```

Standard error skew

```

#standard error of skewness function
ses <- function(n)
{
  sqrt((6*n*(n-1))/((n-2)*(n+1)*(n+3)))
}

```

```

length_captive <- dim(cap_sk)[1]

ses(length_captive)

```

```
## [1] 0.09774553
```

Wild

```
wild_sk <- sc_agg_fun(wild_data)
```

Skew

```

skew_wild <- agostino.test(wild_sk$s_CID)$statistic[1]
agostino.test(wild_sk$s_CID)

```

```

##
## D'Agostino skewness test
##
## data: wild_sk$s_CID
## skew = 0.79228, z = 16.09200, p-value < 2.2e-16
## alternative hypothesis: data have a skewness

```

Standard error skew

```
length_wild <- dim(wild_sk)[1]
```

```
ses(length_wild)
```

```
## [1] 0.04376183
```

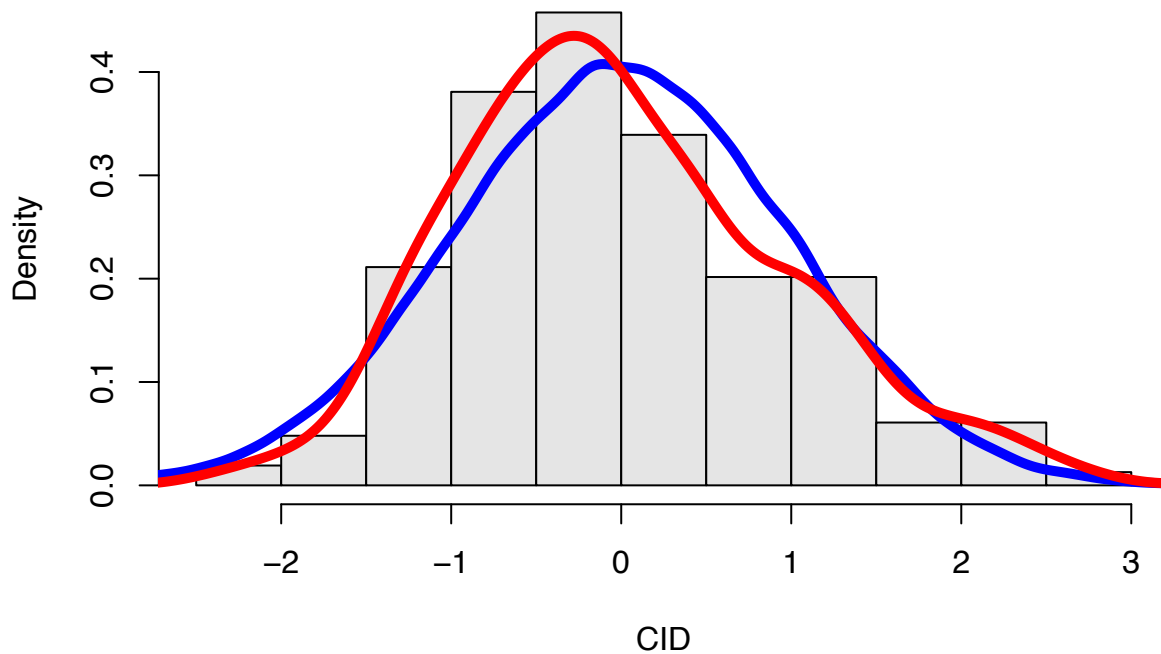
Simulate normal breeding distribution given true CID mean and sd - Plot

Captive

```
m_cap <- mean(cap_sk$s_CID, na.rm = TRUE)
sd_cap <- sd(cap_sk$s_CID, na.rm = TRUE)
cap_rd <- rnorm(100000, mean = m_cap, sd = sd_cap)
```

```
hist(cap_sk$s_CID, prob = TRUE,
     main = 'Breeding distribution - Captive',
     xlab = 'CID', ylab = 'Density', col = 'grey90',
     xlim = c(-2.5, 3),
     breaks = 15)
lines(density(cap_rd), col = 'blue', lwd = 5)
lines(density(cap_sk$s_CID, na.rm = TRUE), col = 'red', lwd = 5)
```

Breeding distribution – Captive



Wild

```
m_wild <- mean(wild_sk$s_CID, na.rm=TRUE)
sd_wild <- sd(wild_sk$s_CID, na.rm=TRUE)
wild_rd <- rnorm(100000, mean= m_wild, sd= sd_wild)
```

```
hist(wild_sk$s_CID, prob=TRUE,
     main='Breeding distribution - Wild',
```

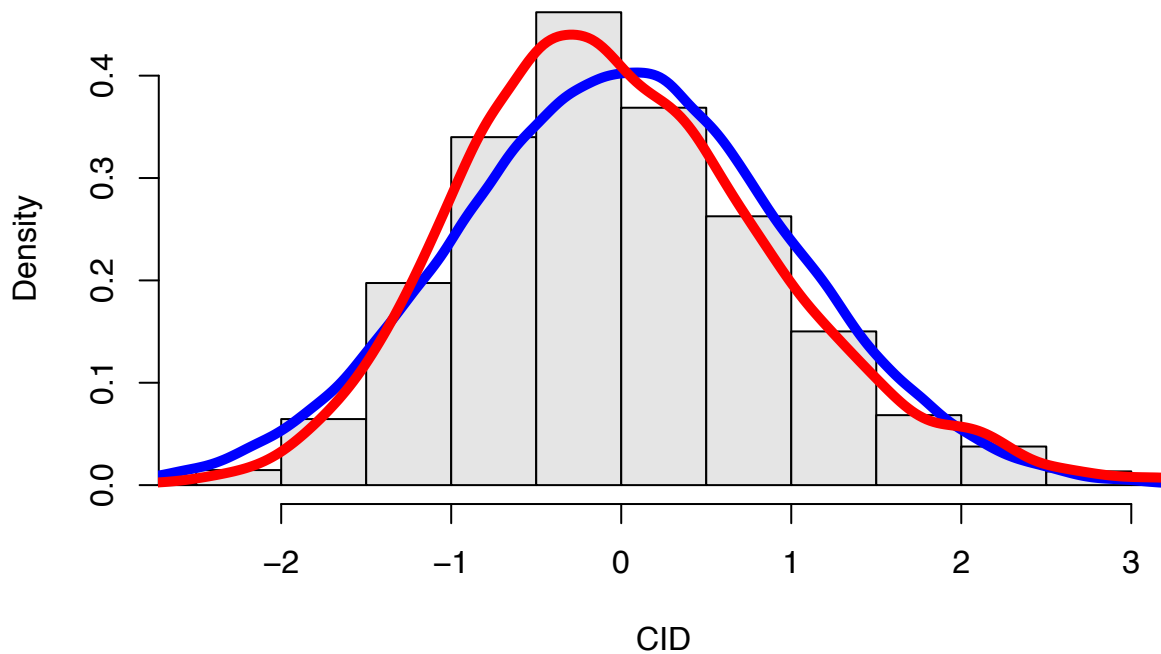


```

xlab= 'CID', ylab= 'Density', col = 'grey90',
xlim = c(-2.5, 3),
breaks = 25)
lines(density(wild_rd), col='blue', lwd=5)
lines(density(wild_sk$s_CID, na.rm=TRUE), col='red', lwd=5)

```

Breeding distribution – Wild



First CID as a predictor for median CID

Fit a linear model to examine the predictive power of first CID for median CID in a given year.

#remove first values when calculating median for that year

```

#CAPTIVE
YRS_C <- unique(captive_data$YEAR)

OUT_C <- c()
for (i in 1:length(YRS_C))
{
  #i <- 1
  temp <- filter(captive_data, YEAR == YRS_C[i])
  min_temp <- min(temp$J_CID)
  min_ind <- which(temp$J_CID == min(temp$J_CID))
  med_temp <- median(temp$J_CID[-min_ind])

  t_OUT_C <- c(YRS_C[i], min_temp, med_temp)
  OUT_C <- rbind(OUT_C, t_OUT_C)
}

```

```

fit_C <- summary(lm(OUT_C[,3] ~ OUT_C[,2]))

#WILD
YRS_W <- unique(wild_data$YEAR)

OUT_W <- c()
for (i in 1:length(YRS_W))
{
  #i <- 1
  temp <- filter(wild_data, YEAR == YRS_W[i])
  min_temp <- min(temp$J_CID)
  min_ind <- which(temp$J_CID == min(temp$J_CID))
  med_temp <- median(temp$J_CID[-min_ind])

  t_OUT_W <- c(YRS_W[i], min_temp, med_temp)
  OUT_W <- rbind(OUT_W, t_OUT_W)
}

fit_W <- summary(lm(OUT_W[,3] ~ OUT_W[,2]))

```

Captive model $r^2 = 0.7$

Wild model $r^2 = 0.7$

Explanatory power is high and very similar between the captive and wild populations.

Skew estimates from literature

Data on skew of phenological data from Thomson 1980, Sparks et al. 2005, Wilson 2013 from plant and bird species. Data is composed of various phenological measures, including first flowering for plants and first arrival for birds (which differs from CID used in this study).

```

setwd('Data')

lit_skew <- read.csv('Lit_skew.csv', header = TRUE)

mn_skew_lit <- mean(lit_skew$Skew)

```

mean skew = 0.43

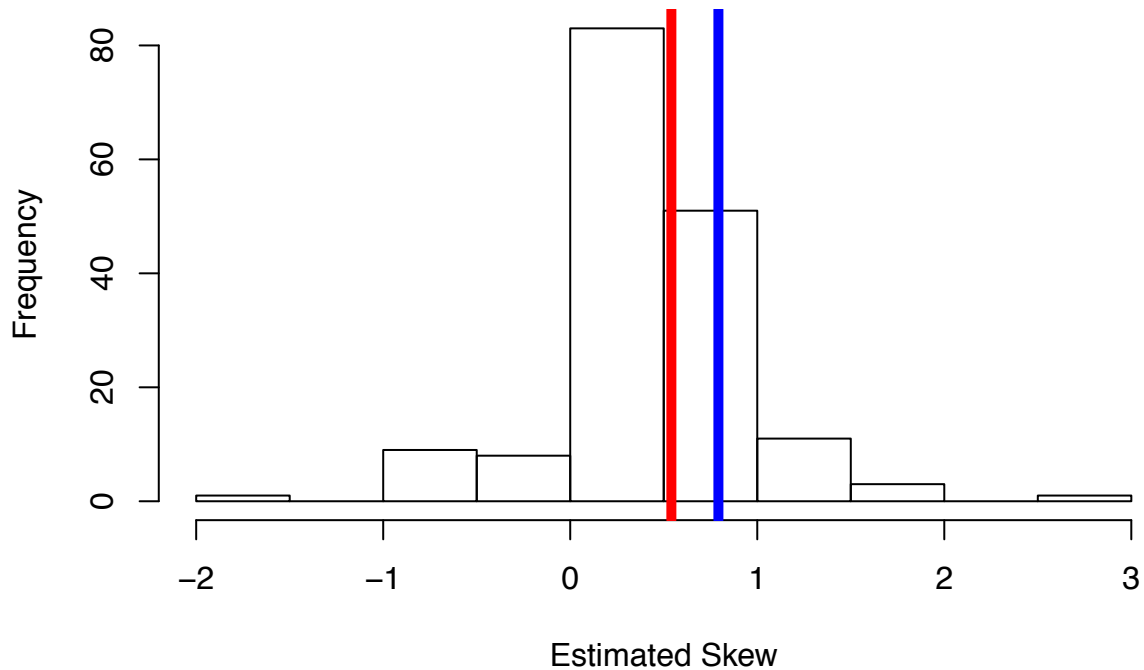
Histogram of skew from literature

```

hist(lit_skew$Skew,
     main = 'Red = Captive; Blue = Wild',
     xlab = 'Estimated Skew')
abline(v = skew_captive, col = 'red', lwd = 5)
abline(v = skew_wild, col = 'blue', lwd = 5)

```

Red = Captive; Blue = Wild



```
per_captive <- length(which(lit_skew$Skew > skew_captive))/NROW(lit_skew)
per_wild <- length(which(lit_skew$Skew > skew_wild))/NROW(lit_skew)
```

Percent literature values greater than captive population skew = 37%

Percent literature values greater than wild population skew = 17%

Skew for Adélie penguin CID is slightly higher than the skew estimated for most populations.

Simulation of right skew breeding distribution

Simple model to show that a right skew distribution can be produced when breeding is accelerated proportional to the number of pairs that have recently initiated breeding.

```
#simulate a normal distirbution (breeding in the absence of conspecifics)
set.seed(1)
time <- round(rnorm(1000, 50, 10))
time <- sort(time)
#no significant skew
agostino.test(time)
```

```
##
## D'Agostino skewness test
##
## data: time
## skew = -0.01578, z = -0.20520, p-value = 0.8374
## alternative hypothesis: data have a skewness
```

```
#simulate 'contagion effect' of breeding
for (i in min(time):max(time))
{
```

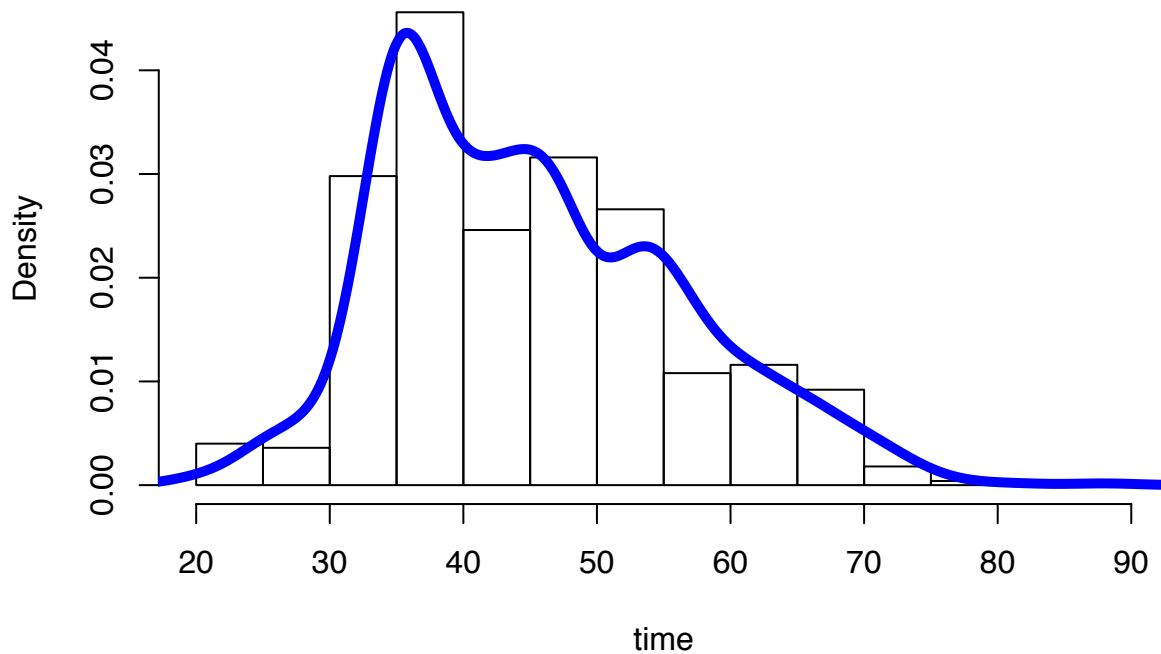
```

time[time == i] <- time[time == i] - sum(as.numeric((time > (i - 3)) & (time < i)))/4
}

hist(time, prob = TRUE)
lines(density(time), col = 'blue', lwd = 5)

```

Histogram of time



```

agostino.test(time)

##
## D'Agostino skewness test
##
## data: time
## skew = 0.55355, z = 6.74630, p-value = 1.516e-11
## alternative hypothesis: data have a skewness

```

Literature Cited

Both, C., van Asch, M., Bijlsma, R.G., van den Burg, A.B. & Visser, M.E. (2009) Climate change and unequal phenological changes across four trophic levels: constraints or adaptations? *Journal of Animal Ecology*, 78, 73–83.

Sparks, T.H., Bairlein, F., Bojarinova, J.G., Huppopp, O., Lehikoinen, E.A., Rainio, K., Sokolov, L.V. & Walker, D. (2005) Examining the total arrival distribution of migratory birds. *Global Change Biology*, 11, 22–30.

Thomson, J.D. (1980) Skewed flowering distributions and pollinator attraction. *Ecology*, 61, 572–579.

Valtonen, A., Latja, R., Leinonen, R. & Pöysä, H. (2017) Arrival and onset of breeding of three passerine birds in eastern Finland tracks climatic variation and phenology of insects. *Journal of Avian Biology*, 48, 785–795.

Wilson, W. (2013) A deeper statistical examination of arrival dates of migratory breeding birds in relation to global climate change. *Biology*, 2, 742–754.

Appendix S2: Additional factors that may affect Adélie penguin breeding phenology

Rethinking ‘normal’: The role of stochasticity in the phenology of a synchronously breeding seabird

Casey Youngflesh, Stephanie Jenouvrier, Jefferson T. Hinke, Lauren DuBois, Judy St. Leger, Wayne Z. Trivelpiece, Susan G. Trivelpiece, Heather J. Lynch

The effect of breeding in year t-1 on breeding in year t

We conducted a post-hoc analysis on the captive population data to investigate temporal autocorrelation in the posterior estimates for the year effect (Fig. 2a). Plotting clutch initiation date (CID) in year t against CID in year t-1, the effect of this temporal inertia on CID is apparent (Fig S1).

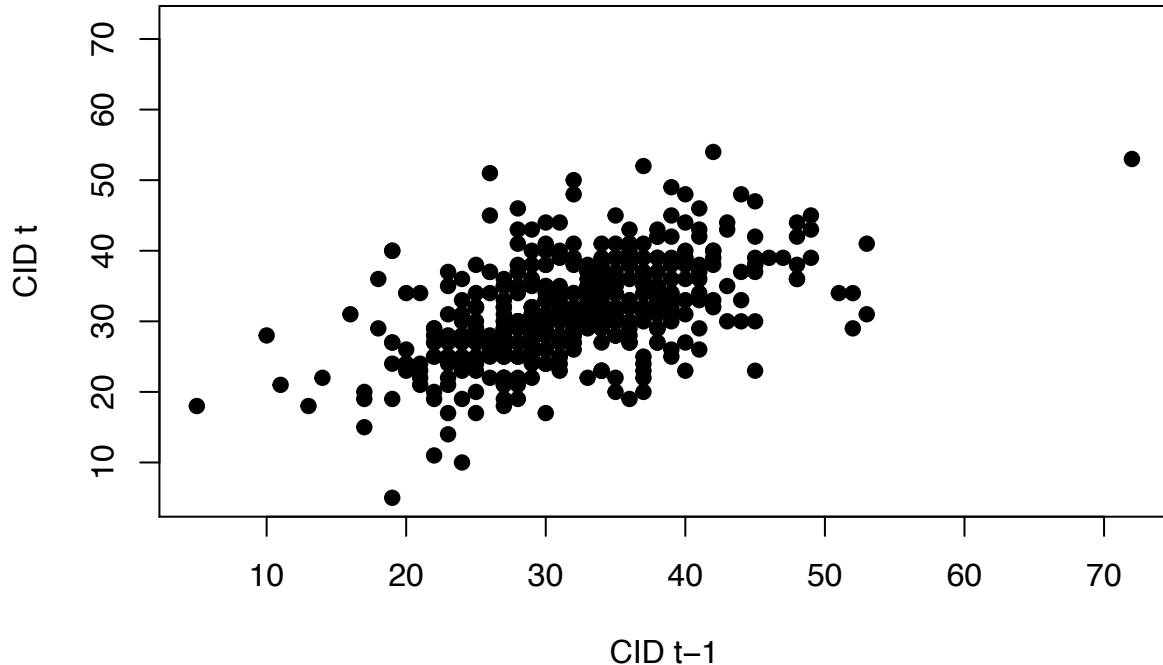


Figure S1: CID for all individuals in the captive population (see Methods) in year t plotted against CID in year t-1 ($r^2 = 0.27$).

To further investigate this pattern, we fit a hierarchical Bayesian model, similar to Eq 1, but with an added autoregressive term, to account for possible temporal autocorrelation.

$$y_{ij} = \mu + \alpha_i + \pi * y_{i-1j} + \beta_j + \gamma * AGE_{ij} + \epsilon_{ij} \quad (S1)$$

$$\alpha_i \sim N(0, \sigma_{year}^2)$$

$$\pi \sim N(0, 1000)$$

$$\beta_j \sim N(0, \sigma_{individual}^2)$$

$$\gamma \sim N(0, 1000)$$

$$\epsilon_{ij} \sim N(0, \sigma_{model}^2)$$

where y_{ij} represents CID, μ represents the intercept, α represents the year effect, β represents the individual effect, π represents the autoregressive term, γ represents the effect of age, AGE represents the age of the female penguin, and ϵ represents the error term.

The model was fit using the R package ‘R2jags’ (Su and Yajima 2015), to interface with JAGS (Plummer 2003) in the R statistical environment (R Development Core Team 2016). Broad Gamma priors were used for all precision ($\tau = \frac{1}{\sigma^2}$) parameters (shape = 0.01, rate = 0.01). Inferences were derived from 20,000 samples drawn following a ‘burn-in’ period of 30,000 draws, using a thinning rate of 2 and 3 chains. Model convergence was assessed through a visual analysis of the posterior chains, in addition to the use of the Gelman-Rubin convergence diagnostic (Brooks and Gelman 1998). All models unambiguously converged. Parameter estimates plots were generated using the ‘MCMCvis’ package (Youngflesh 2016) in the R statistical environment.

RESULTS

Even when accounting for temporal autocorrelation, the year effect appears strong (Fig S2, S3). This suggests that even when controlling for the effect of previous year, the stochastic component to CID across years is still apparent. Inclusion of the autoregressive term does not substantially impact the posterior results of the other parameters. Individual effects are similar between the two models, with estimates closer to 0 when including the autoregressive term (Fig S5, S6). The effect of age is also similar between the two models (Fig S7, S8).

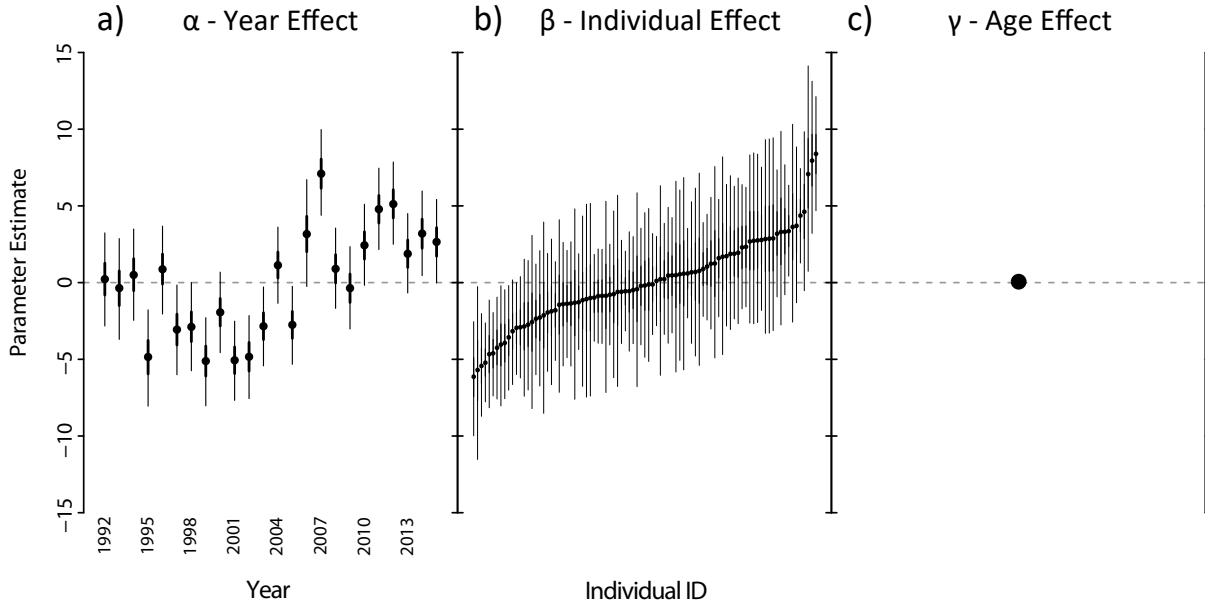


Figure S2: Posterior estimates for the captive population parameters: (a) year effect – α ; (b) individual effect – β ; and (c) age effect – γ parameters for the original model, that does not include the auto-regressive term (see Eq 1). Black circles represent posterior medians. Thicker lines represent 50% credible intervals while thinner lines represent 95% credible intervals. Error bars for the γ parameter are not visible.

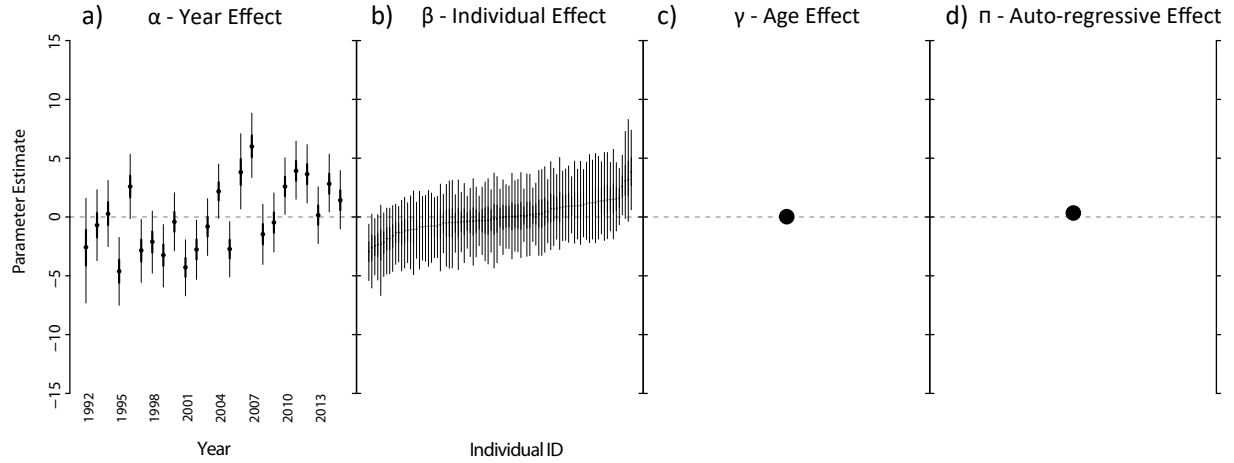


Figure S3: Posterior estimates for the captive population parameters: (a) year effect – α ; (b) individual effect – β ; (c) age effect – γ ; and (d) autoregressive effect - π parameters for the model that includes the auto-regressive term (see Eq S1). Black circles represent posterior medians. Thicker lines represent 50% credible intervals while thinner lines represent 95% credible intervals. Error bars for the γ and π parameters are not visible.

Our primary concern was to determine whether autocorrelation in the CID of individual penguins across years could explain the year effect apparent in the original model (Eq 1). We show that temporal inertia in CID does exist, though it does not explain the year-to-year stochasticity in CID. The reason for this temporal autocorrelation may be due to physiological factors but is beyond the scope of this study. Our model results show a strong year effect in CID is still apparent when accounting for this potentially confounding factor.

The effect of breeding population size on CID

We were also interested to know if the number of breeders in a particular year impacted the breeding phenology of the captive penguins. To further investigate this pattern, we fit a hierarchical Bayesian model, similar to Eq 1, but with an added term for the number of breeders in each year.

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma * AGE_{ij} + \zeta * NB_i + \epsilon_{ij} \quad (S2)$$

$$\alpha_i \sim N(0, \sigma_{year}^2)$$

$$\beta_j \sim N(0, \sigma_{individual}^2)$$

$$\gamma \sim N(0, 1000)$$

$$\zeta \sim N(0, 1000)$$

$$\epsilon_{ij} \sim N(0, \sigma_{model}^2)$$

where y_{ij} represents CID, μ represents the intercept, α represents the year effect, β represents the individual effect, γ represents the effect of age, AGE represents the age of the female penguin, ζ represents the effect of the number of breeders, NB represents the number of breeders in each year, and ϵ represents the error term.

The model was fit using the same methodology denoted above.

RESULTS

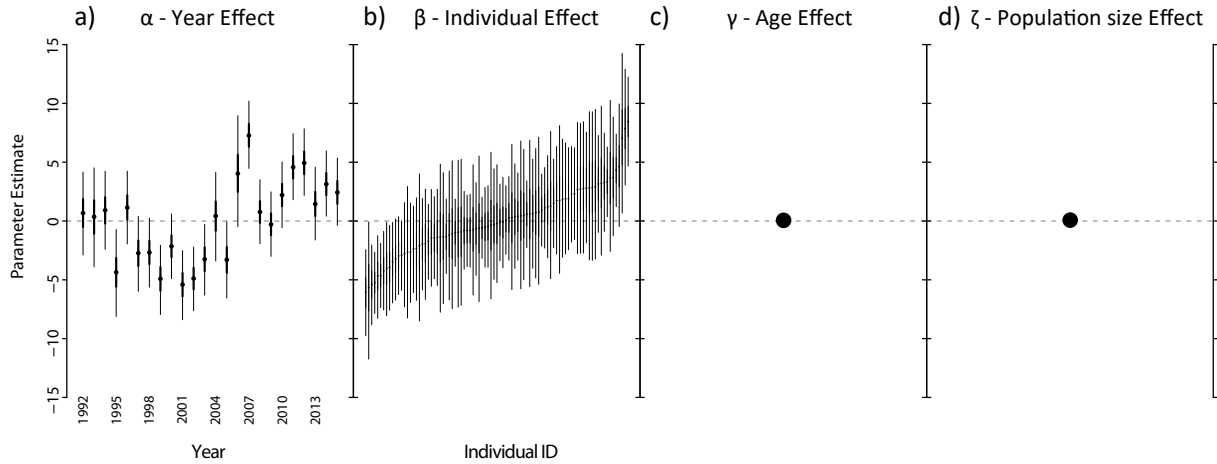


Figure S4: Posterior estimates for the captive population parameters: (a) year effect – α ; (b) individual effect – β ; (c) age effect – γ ; and (d) effect of population size – ζ (see Eq S2). Black circles represent posterior medians. Thicker lines represent 50% credible intervals while thinner lines represent 95% credible intervals. Error bars for the γ and ζ parameters are not visible.

It does not appear that population size has a strong effect on captive penguin CID. The parameter estimate for ζ is small and a year-effect is still apparent.

The effect of the timing of nesting material availability on breeding phenology

We initially hypothesized that the timing of nesting material availability may also impact CID. Adélie penguin rely on the availability of stones to build nests before laying eggs during the breeding season. We compared the ‘waiting time’ of a wild Adélie penguin population at Cape Crozier (77.45°S, 169.20°E) (i.e., the time between the first arrival at the colony and the first clutch initiation date [data presented in Table 3.2 of Ainley et al. (1983)]), to the ‘waiting time’ of the captive population (i.e., the time between when nesting materials are first made available to the penguins [available for 2007-2015] and the first clutch initiation date).

Mean time between the first arrival at the colony and the first egg laid in a wild population was found to be 13.7 days (Ainley et al. 1983).

Mean time between when nesting materials first became available and the first egg laid in this captive population was found to be 41.3 days.

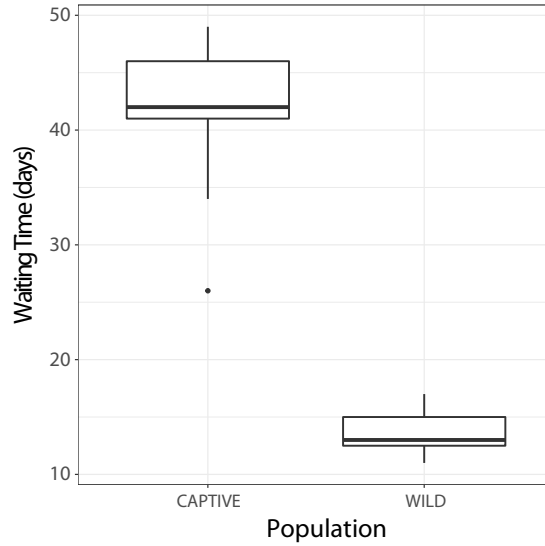


Figure S5: Barchart showing distribution of waiting times for both the captive Adélie penguin population (time between first availability of nesting materials and first CID) and the wild Adélie penguin population observed by Ainley et al. (1983) (time between first arrival at the colony and first CID). This analysis was conducted using the mean CID in the captive population as opposed to the min CID with the same results.

The mean waiting time is more than 3 times higher in captivity than the wild. This suggests that availability of nesting materials is not a barrier for breeding for Adélie penguins in captivity.

Potential effects of 2005 exhibit renovation on phenology

CID appears slightly delayed following the captive exhibit renovation in 2005. To investigate whether this impacted our results, we standardized median CID for 1992-2004 and from 2006-2015 (removing 2005 to account for any impact that the renovation may have had in that year), and calculated the variance for the time series.

To further investigate the effect that any potential step-change due to the 2005 renovation may have had on CID, we fit a hierarchical Bayesian model, similar to Eq 1, but with an added term for pre- and post-renovation identity.

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma * AGE_{ij} + \kappa_l * RID_i + \epsilon_{ij} \quad (S3)$$

$$\alpha_i \sim N(0, \sigma_{year}^2)$$

$$\beta_j \sim N(0, \sigma_{individual}^2)$$

$$\gamma \sim N(0, 1000)$$

$$\kappa_l \sim N(0, 1000)$$

$$\epsilon_{ij} \sim N(0, \sigma_{model}^2)$$

where y_{ij} represents CID, μ represents the intercept, α represents the year effect, β represents the individual effect, γ represents the effect of age, AGE represents the age of the female penguin, κ represents the effect of pre/post renovation identity, RID represents whether a particular year belongs in the pre or post renovation period, and ϵ represents the error term.

The model was fit using the same methodology denoted above.

RESULTS

$$\sigma_{RID}^2 = 7.59$$

$$\sigma_{AdmiraltyBay}^2 = 13.5$$

When standardizing the pre and post renovation periods, the overall variance of the captive time series decreases, leading to a slightly lower variance than that seen at Admiralty Bay. This is well within the range of variance in CID found at other wild Adélie penguin colonies in Antarctica (Youngflesh et al. 2017), however, confirming that variance in CID in this captive population is as great as that seen in the wild.

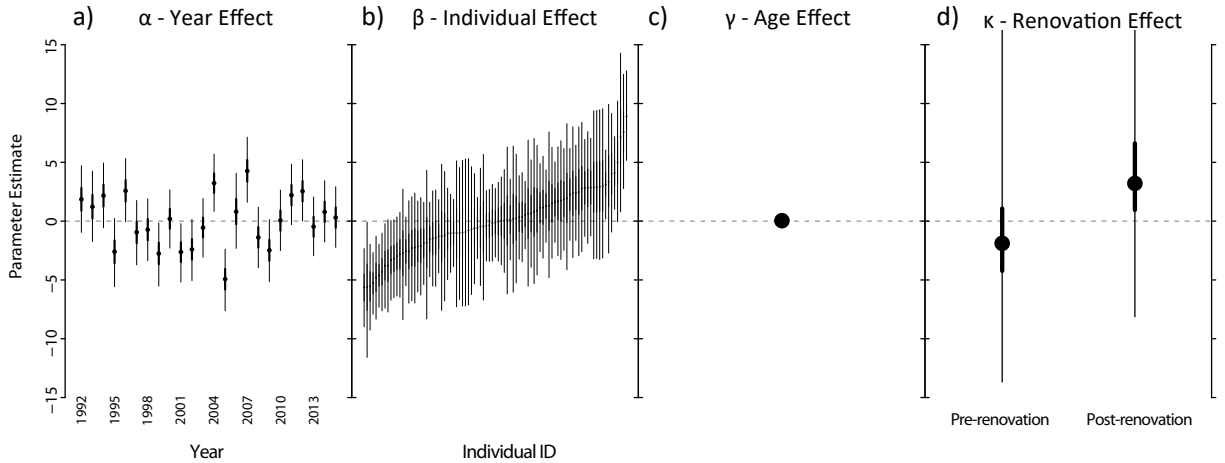


Figure S6: Posterior estimates for the captive population parameters: (a) year effect – α ; (b) individual effect – β ; (c) age effect – γ ; and (d) effect of period id (pre/post renovation) - κ (see Eq S3). Black circles represent posterior medians. Thicker lines represent 50% credible intervals while thinner lines represent 95% credible intervals. Error bars for the γ parameter are not visible.

None of the factors examined changed our overall conclusions. Year effects are slightly weaker when adding the κ term compared to the original model (Fig S2), however are still prominent. Parameter estimates for κ are relatively close to zero, with large credible intervals. Even when accounting for any potential effect that the renovation may have had on CID, interannual variability in CID appears to be the norm.

Literature Cited

Ainley, D. G., R. E. LeResche, & Sladen W. J. L. (1983) *Breeding Biology of the Adélie Penguin*. University of California Press, Los Angeles, CA, USA.

Plummer, M. (2003) *JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling. Proceedings of the Third International Workshop on Distributed Statistical Computing*. R Project for Statistical Computing, Vienna, Austria.

R Core Team. (2016) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. www.r-project.org

Su, Y., & Yajima, M. (2015) *R2jags: Using R to Run "JAGS."* cran.r-project.org/web/packages/R2jags/index.html

Youngflesh, C. (2016) *MCMCvis: Tools to visualize, Manipulate, and Summarize MCMC output*. cran.r-project.org/web/packages/MCMCvis/index.html

Youngflesh, C., Jenouvrier, S., Li, Y., Ji, R., Ainley, D.G., Ballard, G., Barbraud, C., Delord, K., Dugger, K.M., Emmerson, L.M., Fraser, W.R., Hinke, J.T., Lyver, P.O'B., Olmastroni, S., Southwell, C., Trivelpiece, S.G., Trivelpiece, W.Z., & Lynch, H.J. (2017) Circumpolar analysis of the Adélie penguin reveals the importance of environmental variability in phenological mismatch. *Ecology*, 98, 940-951.