

1 Environmental drivers of varying selective optima in  
2 a small passerine: a multivariate, multiepisodic  
3 approach  
4

5 January 9, 2019

6 Marlène Gamelon<sup>1,\*</sup> Jarle Tufto<sup>2,\*</sup> Anna L.K. Nilsson<sup>3</sup> Kurt Jerstad<sup>4</sup> Ole W. Røstad<sup>5</sup>  
7 Nils C. Stenseth<sup>1,3</sup> Bernt-Erik Sæther<sup>1</sup>

8 1. Centre for Biodiversity Dynamics CBD, Department of Biology, Norwegian Univer-  
9 sity of Science and Technology, 7491 Trondheim, Norway

10 2. Centre for Biodiversity Dynamics CBD, Department of Mathematical Sciences, Nor-  
11 wegian University of Science and Technology, 7491 Trondheim, Norway

12 3. Centre for Ecological and Evolutionary Synthesis CEES, Department of Biosciences,  
13 University of Oslo, 0316 Oslo, Norway

14 4. Jerstad Viltforvaltning, Aurebekksveien 61, 4516 Mandal, Norway

15 5. Faculty of Environmental Sciences and Natural Resource Management, Norwegian  
16 University of Life Sciences, 1432 Ås, Norway.

17 \* Correspondence to: marlene.gamelon@ntnu.no and jarle.tufto@ntnu.no

18 **Running title:** Varying selection in the wild

19 **Acknowledgements:** We thank all the persons involved in the field work. We also  
20 warmly thank Luis-Miguel Chevin, Ophélie Ronce, Michael Morrissey and three anony-  
21 mous referees for helpful comments on previous drafts and Hans J. Skaug for useful  
22 discussions. This work was partly supported by the Directorate for Nature Manage-  
23 ment (Norwegian Environment Agency) and the Research Council of Norway through

24 its Centres of Excellence funding scheme (project number 223257 to CBD and 179569 to  
25 CEES).

26 **Author contributions:** K.J. and O.W.R. contributed to field work and data collec-  
27 tion; A.L.K.N. provided data on egg-laying dates; M.G., J.T., N.C.S. and B.E.S. conceived  
28 the study; J.T. developed the statistical model; M.G. and J.T. conducted the analyses;  
29 M.G. and J.T. wrote the manuscript and all authors contributed to revisions on later  
30 manuscript versions.

31 **Data archiving:** All the necessary data and scripts to carry out the analyses pre-  
32 sented here are available in the dryad digital repository <https://doi.org/10.5061/dryad.k871qr2>

33 In changing environments, phenotypic traits are shaped by numerous agents  
34 of selection. The optimal phenotypic value maximizing the fitness of an indi-  
35 vidual thus varies through time and space with various environmental covari-  
36 ates. Selection may differ between different life cycle stages and act on cor-  
37 related traits inducing changes in the distribution of several traits simultane-  
38 ously. Despite increasing interests in environmental sensitivity of phenotypic  
39 selection, estimating varying selective optima on various traits throughout  
40 the life cycle, while considering (a)biotic factors as potential selective agents  
41 has remained challenging. Here, we provide a statistical model to measure  
42 varying selective optima from longitudinal data. We apply our approach to  
43 analyse environmental sensitivity of phenotypic selection on egg-laying date  
44 and clutch size throughout the life cycle of a white-throated dipper popula-  
45 tion. We show the presence of a joint optimal phenotype that varies over the  
46 35-yr period, being dependent on altitude and temperature. We also find that  
47 optimal laying date is density-dependent, with high population density favor-  
48 ing earlier laying dates. By providing a flexible approach, widely applicable  
49 to free-ranging populations for which long-term data on individual pheno-  
50 types, fitness and environmental factors are available, our study improves the  
51 understanding of phenotypic selection in varying environments.

52 Key words: Clutch size, density dependence, egg-laying date, fluctuating  
53 environment, selection episode.

54 Environments, through variation in habitats, competition or predation, are heteroge-  
55 neous at both temporal and spatial scales. Such changes in biotic and abiotic conditions  
56 impose selection on wild populations (Bell, 2010), leading phenotypic traits to be con-  
57 stantly shaped and reshaped by the environment and numerous agents of natural selection  
58 (Endler, 1986). Although often ignored, population density is one of these potential selec-  
59 tive agents (Sæther et al., 2016). For a long time, evolution has been considered too slow  
60 and too weak to leave a signature in ecological dynamics (Slobodkin, 1961). However, it  
61 is now widely accepted that rates of evolution can be rapid and strong (Pelletier et al.,  
62 2007; Ozgul et al., 2009, 2010; Pemberton, 2010; Bell, 2010; Schoener, 2011). Interest-  
63 ingly, observations of phenotypic selection in free-ranging populations also indicate that  
64 evolution may vary through space (Endler, 1977; Hereford, 2009; Siepielski et al., 2013;  
65 Hedrick et al., 1976) and time (Bell, 2010; Siepielski et al., 2009; Morrissey and Hadfield,  
66 2012). A landmark case study of varying selection is the beak size variation in Dar-  
67 win’s finches in response to droughts. While drought events have favored large beaks well  
68 adapted to large seeds, high precipitation have selected for smaller beak sizes particularly  
69 useful for consuming small and soft seeds (Grant and Grant, 2002) such that the optimal  
70 phenotypic value maximizing fitness is moving as a result of fluctuating environmental  
71 conditions (Charlesworth, 1993; Tufto, 2015; Chevin et al., 2015).

72 Chevin et al. (2015) provided a method for estimating varying phenotypic selection  
73 from measurements of a fitness-related trait across time. It assesses the support for  
74 stabilizing selection and for an optimal phenotype possibly influenced by environmen-  
75 tal covariates (hereafter called environmental sensitivity of selection sensu Chevin et al.  
76 (2010)) and random effects autocorrelated across years. Using great tits (*Parus major*)  
77 as a case study, this work showed autocorrelated variations in the optimal egg-laying  
78 date that maximizes the number of offspring surviving to the fledgling stage. In addition,  
79 the optimal date was well predicted by spring temperature. This study left some ques-  
80 tions unanswered, in particular how to measure varying phenotypic selection on multiple  
81 correlated traits and also throughout multiple episodes of selection.

82 Indeed, natural selection does not operate on a single trait but acts jointly and cor-  
83 relatively on multiple characters, and the environment causes this complex selection to  
84 change in a more or less predictable way. This very fundamental and widely acknowledged  
85 vision of how adaptation to changing environment proceeds still fails to be detected in a

86 comprehensive way. The classical approach of Lande and Arnold (1983) allows estimating  
87 variation in linear and quadratic selection gradients acting on multiple traits including  
88 correlational selection (Phillips and Arnold, 1989b; Sinervo and Svensson, 2002) over  
89 time (Engen et al., 2012). This multivariate selection analysis thus provides important  
90 information on the direction, shape and strength of selection acting on multiple pheno-  
91 typic traits over time by linking relative fitness to trait values. However, relating such  
92 phenotypic selection gradients to environmental factors may lead to incomplete repre-  
93 sentations of how the environment interacts with the trait-fitness relationships (Hunter  
94 et al., 2018). Indeed, this variation not only reflects variation in the fitness function  
95 (i.e. the relationship between individual expected fitness and individual phenotype, see  
96 Walsh and Morrissey (2018)), but is complicated by the response to selection generated  
97 by the varying fitness function and phenotypic changes caused by other evolutionary  
98 forces. As a simple illustration, if the mean phenotype tracks a varying optimum almost  
99 perfectly (e.g. through plasticity), little variation in selection gradients will be detected  
100 using the traditional method of Lande and Arnold (1983). Conversely, an evolutionary  
101 force such as genetic drift will generate varying gradients, even if the fitness function  
102 is constant (Chevin and Haller, 2014). In other words, considering phenotypic selection  
103 coefficients such as selection gradients or differentials alone does not necessarily allow one  
104 to characterize how the fitness function has changed, because changes in the distribution  
105 of phenotype can change selection coefficients, independently of changes in the fitness  
106 function.

107 Natural selection does not operate on a single episode but the strength and the di-  
108 rection of selection on a trait may change from one life cycle stage to another (Chevin  
109 et al., 2017; Engen et al., 2011). This was recognized more than thirty years ago by  
110 Arnold and Wade (1984) who highlighted the need to measure selection through separate  
111 episodes of selection across the life cycle. However, if selection is estimated separately  
112 for each life cycle segment as in Engen et al. (2012), this leads to a loss of parsimony  
113 when different episodes are similarly influenced by the same environmental covariates or  
114 random processes.

115 Here, we analyze fluctuating fitness functions through time and space in a Norwe-  
116 gian white-throated dipper population (*Cinclus cinclus*) by extending the approach from  
117 Chevin et al. (2015). In particular, we explore the dynamics of selective optima through

118 time and space on two key fitness-related traits in such a small passerine (Newton, 1998),  
119 namely egg-laying date and clutch size. For many breeding females, information on egg-  
120 laying date and clutch size is available annually, thus providing the required data to  
121 develop a multivariate approach. We evaluate the effects of biotic and abiotic factors  
122 such as weather conditions, altitude and densities as well as random unobserved drivers  
123 on spatio-temporal variation in selective optima of the two traits. To make efficient use  
124 of all the data, we estimate varying phenotypic selection through several episodes of se-  
125 lection in a single joint model, from egg to fledgling stage (episode 1), from fledgling to  
126 recruit stage (episode 2) and via mother survival (episode 3), thanks to the availability  
127 of individual-based data from birth to death.

128 Importantly, instead of modelling variation in selection gradients, we model fluctua-  
129 tions in the fitness function directly. Parameter estimates from our model can thus more  
130 easily be related to the theory on varying selection, both in time (Bull, 1987; Lande and  
131 Shannon, 1996; Lande, 2007; Tufto, 2015) and space (Kirkpatrick and Barton, 1997).  
132 While much of this theory predominantly deals with simple life histories with discrete,  
133 non-overlapping generations, our approach provides a statistical model for estimating  
134 varying selection acting in an age-structured population by applying a stochastic trait-  
135 dependent Leslie matrix (Caswell, 2001), including covariates such as population density.  
136 Further theoretical work will be needed to understand the evolutionary response to se-  
137 lection described by our modelling approach, through some integrated measure of overall  
138 lifetime fitness. Still, in our dipper case study, even in the absence of such theory, we  
139 find that the observed pattern in mean phenotypic trait values, qualitatively behaves as  
140 expected in response to our estimated spatial and temporal variation in selective optima.

## 141 Methods

### STUDY SPECIES AND DATA COLLECTION

142 The studied population is located in the river system of Lyngdalselva in southern Norway  
143 ( $58^{\circ}08' - 58^{\circ}40'N$ ,  $6^{\circ}56' - 7^{\circ}20'E$ ). The white-throated dipper is a short-lived passerine bird  
144 distributed in mountainous regions across the Palearctic. It depends on open water for  
145 foraging and running water for nesting. The amount of ice during the winter thus influ-  
146 ences the availability of feeding and breeding habitats explaining why survival and fecun-

147 dity rates are affected by mean winter temperature (temp) (December-February) of the  
148 whole region called Sørlandet ([http://www.yr.no/sted/Norge/Vest-Agder/Audnedal/  
149 Konsmo~6051/klima.vinter.html](http://www.yr.no/sted/Norge/Vest-Agder/Audnedal/Konsmo~6051/klima.vinter.html)) (Sæther et al., 2000; Loison et al., 2002; Nilsson  
150 et al., 2011a; Gamelon et al., 2017). Demographic rates are also strongly density-  
151 regulated and annual estimates of the number of breeding females in the population  
152 (variable dens) are available from a previous study (Gamelon et al., 2017). For each year  
153  $t = 1979, 1980, \dots, 2013$ , all breeding sites were visited during the nest building period to  
154 identify breeding pairs and record occupied nests. Assuming that covariates are missing  
155 at random, we based the analysis on a subset of  $i = 1, 2, \dots, 546$  breeding events with  
156 non-missing covariates out of a total of 1880 observed breeding events. The altitude ( $alt_i$ )  
157 (ranging from sea level to an altitude of about 600 m above sea level) and the identity of  
158 each breeding territory ( $k = 1, 2, \dots, 167$ ) were recorded. During visits in the breeding  
159 season, ringed mothers were identified and unringed mothers given a ring to allow future  
160 identifications ( $j = 1, 2, \dots, 375$ ). For each breeding pair, the egg-laying date (date of  
161 first egg laying) and the clutch size ( $y_{0i}$ ) was determined (Nilsson et al., 2011b). We  
162 express egg-laying date as the number of days elapsed since 1<sup>st</sup> of March ( $z_i$ ). On av-  
163 erage, twenty-two days later, fledglings that had survived were ringed and their number  
164 recorded ( $y_{1i}$ ). Finally, the next season, that is, on average 343 days later, a number  $y_{2i}$   
165 of ringed fledglings were recorded as recruited to the breeding population if they were  
166 caught breeding. Mothers caught again breeding the following year were recorded as  
167 having survived ( $y_{3i} = 1$ ); otherwise, they were considered as dead ( $y_{3i} = 0$ ) (Fig. 1).  
168 Therefore, survival on episodes 2 and 3 ( $p_{2i}$  and  $p_{3i}$ ) correspond to apparent survival, i.e.  
169 the probability for a female to survive and stay in the population until the next breeding  
170 season. Survival on the first episode ( $p_{1i}$ ) corresponds to true survival, because there  
171 is no possible dispersal during the first episode. Note that the annual recapture rate is  
172 high during the studied period, ranging between 88 and 92% (estimates obtained in a  
173 previous work (Gamelon et al., 2017)), meaning that virtually all the females alive were  
174 caught breeding. Moreover, age of the mothers ( $a_i$ ) was determined. The oldest breeding  
175 female recorded in our population was 10 years of age. Note that multiple-clutching was  
176 sometimes observed in that population. Thus, we also recorded the total number of eggs  
177 produced by a female a given year.

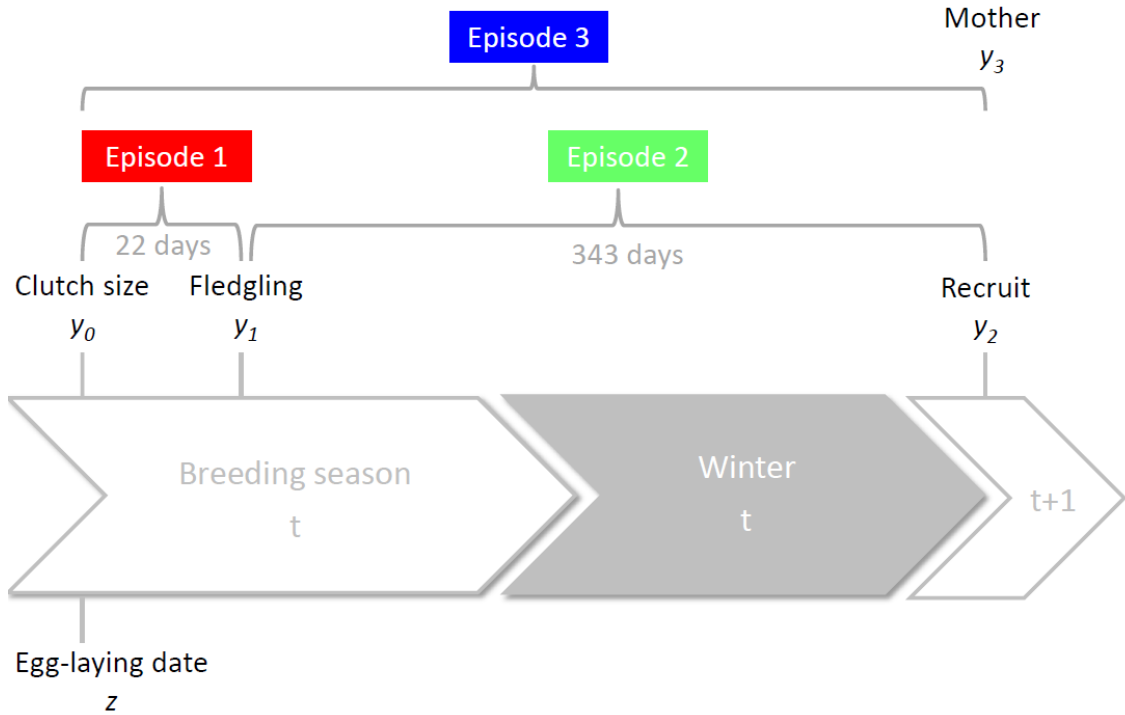


Figure 1: Dipper life cycle. Episode 1 corresponds to the episode from egg to fledgling stage (in red), episode 2 from fledgling to recruit stage (in green), and episode 3 corresponds to adult female stage (in blue).

## STATISTICAL MODEL

178 Using the above data, our aim is to estimate varying selective optima, extending the  
 179 method of Chevin et al. (2015) to multiple traits (individual laying dates and clutch  
 180 sizes) and to multiple episodes of selection (egg-to-fledgling survival  $s = 1$ , fledgling-  
 181 to-recruit survival  $s = 2$ , and adult female survival  $s = 3$ ) through associated survival  
 182 probabilities  $p_{si}$ ,  $s = 1, 2, 3$  (Fig. 1). A more technical discussion of other differences  
 183 from the Chevin et al. (2015) method is given in Appendix A.1. We emphasize that  
 184 our model of stabilizing selection does not necessarily imply a fitness optimum within  
 185 the range of observed phenotypic values in any particular generation. Instead, as in  
 186 theoretical models such as Lande and Shannon (1996); Hansen (1997); Bürger (1999),  
 187 overall individual fitness is a strictly decreasing function on both sides of some optimal  
 188 trait value possibly located outside this range. The particular model we implement, at  
 189 least when survival is low such that the model becomes approximately Gaussian, also



190 implies that selection acts to reduce the phenotypic variance and that selection gradients  
 191 (sensu Lande and Arnold, 1983) change linearly with the distance from the assumed  
 192 optimum.

193 Before describing in detail how we implement stabilizing selection, we describe more  
 194 generally how covariates and random effects (some of which induce varying selection)  
 195 need to enter into the model given that the different episodes differ in duration. As  
 196 recommended by Ergon et al. (2017), we model all three survival probabilities  $p_{si}$  only  
 197 indirectly via effects of covariates and random effects on the hazard function. More  
 198 specifically, we allow a non-constant hazard (instantaneous mortality rate) with respect  
 199 to age but assume that covariates and random effects act with a constant multiplicative  
 200 effect on this rate (via a log link) within each episode of selection. Thus, the hazard for  
 201 an offspring or an adult female at age  $a$  is

$$\lambda_{si}(a) = \lambda_0(a) \exp(\eta_{si}). \quad (1)$$

202 Here,  $\eta_{si}$  is a (non-linear) predictor containing fixed and random effects on the hazard  
 203 associated with the  $i$ 'th breeding event during selection episode  $s = 1, 2, 3$ , and  $\lambda_0(a)$  is a  
 204 baseline hazard affecting all individuals, possibly varying with age  $a$  within each interval.  
 205 The survival probabilities associated with each of the three episodes of selection are given  
 206 by

$$\begin{aligned} p_{si} &= \exp \left( - \int_{a_{s-1}}^{a_s} \lambda_0(a) \exp(\eta_{si}) da \right) \\ &= \exp \left( - \exp(\eta_{si}) \bar{\lambda}_s (a_s - a_{s-1}) \right), \end{aligned} \quad (2)$$

207 where  $a_{s-1}$  and  $a_s$  is the age at the beginning and end of selection episode  $s$ . Note how  $\bar{\lambda}_s$   
 208 is the mean of the possibly non-constant baseline hazard  $\lambda_0(a)$  during selection episode  
 209  $s$ .

210 To model selection on the laying date  $z_i$  and on clutch size  $y_{0i}$ , we in turn assume that  
 211 the (time-averaged) hazard of individual  $i$  during each selection episode has the form

$$\exp(\eta_{si}) \bar{\lambda}_s = \exp \left( \eta_{si}^{(\beta)} + \frac{1}{2\omega_s^2} (z_i - \eta_{si}^{(\theta)})^2 \right). \quad (3)$$

212 with the parameter  $\omega_s$  determining the rate of proportional increase in the hazard with  
 213 increasing deviations of the phenotypic laying date  $z_i$  from the optimal laying date  $\eta_{si}^{(\theta)}$ .  
 214  $\omega_s$  thus corresponds to the width of the fitness peak (smaller  $\omega_s$  causes stronger stabilizing

215 selection). Here

$$\eta_{si}^{(\theta)} = \theta_{0s} + \theta_{\text{dens},s} \text{dens}_i + \theta_{\text{temp},s} \text{temp}_i + \theta_{\text{alt},s} \text{alt}_i + \zeta_t \quad (4)$$

216 is a linear sub-predictor determining the optimal laying date during episode  $s$  containing  
 217 possible effects of environmental covariates such as population density ( $\text{dens}_i$ ), winter  
 218 temperature ( $\text{temp}_i$ ) and altitude ( $\text{alt}_i$ ) as well as a random effect term  $\zeta_t$  as in Chevin  
 219 et al. (2015). Similarly,

$$\begin{aligned} \eta_{si}^{(\beta)} &= \beta_{0s} + \beta_{\text{clutchsize},s} y_{0i} \\ &+ \beta_{\text{temp},s} \text{temp}_i + \beta_{\text{dens},s} \text{dens}_i + \beta_{\text{alt},s} \text{alt}_i + \beta_{\text{age},s} a_i \\ &+ \sigma_s u_{t_i} + \tau_s v_{k_i} + \kappa_s w_{j_i} + x_{a_i,s} \end{aligned} \quad (5)$$

220 is another linear sub-predictor determining the hazard at the optimal laying date con-  
 221 taining effects of covariates as well as a number of random effects (details are given in  
 222 the next 4 paragraphs and Appendix A.1). The above regression coefficients must not  
 223 be confused with the selection gradient  $\beta$  as defined by Lande and Arnold (1983). All  
 224 parameters possibly differ between episodes  $s = 1, 2, 3$ , but can also be constrained to the  
 225 same value for different subsets of episodes. Importantly, this facilitates the formulation  
 226 of more parsimonious model alternatives in cases where the evidence for any difference  
 227 between episodes is small. Note that the log of mean baseline hazard has been absorbed  
 228 in the possibly episode-dependent intercept  $\beta_{0s} = \ln \bar{\lambda}_s$  in (5).

229 Before going through the details of the linear predictors in 4 and 5, note first that  
 230 the expected number of recruits produced by a given female (the fecundities  $f_a$  in the  
 231 first row of a pre-breeding census Leslie matrix), assuming that a single clutch is laid, are  
 232 given by products of clutch size  $y_0$  and the survival probabilities for the two first episodes  
 233 of selection  $f_a(y_{0i}, z_i) = y_{0i} p_{1i} p_{2i}$ . These fecundities are important fitness components  
 234 and correspond to the number of young produced during the breeding season in year  $t$   
 235 that have survived until the next year  $t + 1$  (see Fig. 1). Even without a quadratic effect  
 236 of clutch size in (5), provided that increasing clutch sizes translates to a reduction in  
 237 overall survival during episodes  $s = 1$  or  $2$  ( $\beta_{\text{clutchsize},s}$  sufficiently positive), it follows that  
 238 these fecundities are maximized for some intermediate clutch size, as expected through  
 239 the trade-off between offspring number and offspring survival (Smith and Fretwell, 1974).  
 240 Note that the model may predict an optimum located outside the range of observed  
 241 phenotypic values. Thus, our model specifies a joint optimum for the two phenotypic

242 traits laying date and clutch size for which  $f_a(y_0, z)$  is maximized. To obtain a more  
 243 standard measure of the strength of stabilizing selection acting jointly on both traits, we  
 244 evaluated the matrix of second derivatives of  $\ln f_a(y_0, z)$  at the joint optimum to obtain  
 245 the parameters of the Gaussian approximation of the fitness function (appearing in many  
 246 theoretical models, e.g. Phillips and Arnold (1989a); Chevin (2013); Tufto (2017)). We  
 247 report the widths  $\omega_{y_0}$  and  $\omega_z$  of this approximation with respect to each trait (in units  
 248 of number of eggs and number of days, respectively), analogous to standard deviations  
 249 of Gaussian distributions.

250 Temporal covariates and temporal random effects appearing in 4 and 5 translate to  
 251 variation from year to year in the optimal laying date and clutch size, respectively. To  
 252 model possibly correlated fluctuations in the joint optimum as in Chevin (2013) as well  
 253 as autocorrelation across time (as in Lande and Shannon, 1996; Lande, 2009; Tufto, 2015;  
 254 Chevin et al., 2017), the random effects representing yearly variation in overall survival  
 255  $u_t$  (5) and variation in the optimal laying date  $\zeta_t$  (4) are assumed to follow a first-order  
 256 vector autoregressive VAR(1) process

$$\begin{bmatrix} u_t \\ \zeta_t \end{bmatrix} = \mathbf{\Phi} \begin{bmatrix} u_{t-1} \\ \zeta_{t-1} \end{bmatrix} + \mathbf{w}_t, \quad (6)$$

257 where  $\mathbf{\Phi}$  is a  $2 \times 2$  matrix of autoregressive coefficients and  $\mathbf{w}_t$  is bivariate normal  $N(0, \mathbf{\Sigma})$   
 258 white noise. This only specifies the autocorrelation matrix function (see Wei, 2006, ch.  
 259 16.1) of the process  $(u_t, \zeta_t)$ . But as long as the variance of  $u_t$  is small, optimal clutch size  
 260 will be approximately linearly dependent on  $u_t$  and so the autocorrelation matrix function  
 261 of the joint optimal clutch size and laying date will be almost identical to that of  $(u_t, \zeta_t)$ .  
 262 Correlation between  $u_t$  and  $\zeta_t$  can arise either through  $\mathbf{\Sigma}$ ,  $\mathbf{\Phi}$  or both having non-zero  
 263 off-diagonal entries. If  $\mathbf{\Phi}$  and  $\mathbf{\Sigma}$  are both diagonal, this simplifies to two independent  
 264 AR(1) processes and if all entries of  $\mathbf{\Phi}$  are zero,  $u_t$  and  $\zeta_t$  are simple independent and  
 265 identically normally distributed (iid) white noise processes. We parameterize this part  
 266 of the model in terms of  $\mathbf{\Phi}$ , the white noise correlation  $\rho = \Sigma_{12}/\sqrt{\Sigma_{11}\Sigma_{22}}$ , the stationary  
 267 variance  $\sigma_\zeta^2$  of  $\zeta_t$ , and with the stationary variance of  $u_t$  in (5) set equal to one but with  
 268 separate parameters  $\sigma_s$  representing the potentially different effects of  $u_t$  on the three  
 269 selection episodes. Note also that additional correlation in variation of the joint optimum  
 270 as well as autocorrelation across time can be induced through the temporal fixed effect  
 271 covariates appearing in each linear predictor.

272 Correlated optima discussed above are distinct from correlational selection. Two kinds  
273 of correlational selection can be accommodated in our non-Gaussian model for  $f_a(y_{0i}, z_i)$ ,  
274 either by adding  $z_i y_{0i}$  as a covariate in (5) making the optimal clutch size dependent  
275 on laying date (first kind), or by adding clutch size  $y_{0i}$  as an additional covariate in (4)  
276 making the optimal laying date dependent on clutch size (second kind).

277 To model the effect of the age of breeding female, we consider models where the log  
278 hazard during the different episodes is either independent of age (a single term  $\beta_{0,s}$  in (5)),  
279 linearly dependent on age (an additional term  $\beta_{\text{age},s} \text{age}_i$  in (5)) or non-linearly dependent  
280 on age. Instead of modelling non-linear age-dependence parametrically (using for example  
281 quadratic, Gompertz, piecewise linear or two-parameter Weibull models (Gaillard et al.,  
282 2004; Marzolin et al., 2011)), we use a non-parametric approach: we model the age effects  
283 through a set of correlated random effects, more precisely as a second-order random  
284 walk. This is a commonly used method for smoothing data and modelling response  
285 functions (Green and Silverman, 1994; Rue and Held, 2005). It provides a simple and  
286 flexible way to model the hazard variations as a function of age, and thus to explore  
287 senescence. Under the second order random walk model, the joint distribution of the age  
288 effects  $x_{s,1}, x_{s,2}, \dots, x_{s,10}$  (last term in (5), Fig. 2, upper plot, blue curve) is specified by  
289 assuming that the second order differences  $\Delta^2 x_{s,a}$  are independently normally distributed  
290 with zero mean and variance  $\nu_s$ . The parameter  $\nu_s$  (estimated along with the random  
291 age effects themselves) thus controls the magnitude of these second order differences  
292 (analogous to the second order derivative) and hence the smoothness of the resulting  
293 function. Further details on the implementation of the model and the two last random  
294 effects terms appearing in 5 are given in Appendix A.1.

295 Apart from the non-linear predictor, our model based on the assumption of multi-  
296 plicative effects on the hazard corresponds to a generalized linear mixed model with a  
297 log-log link (or complementary log-log) function with the log of the length of the se-  
298 lection episode included as an offset variable. Surprisingly, this link and its associated  
299 assumption of multiplicative effects on the hazard is rarely used in ecology and evolu-  
300 tion. However, this is a standard assumption in many models in survival analysis (Ergon  
301 et al., 2017), for example in Cox proportional hazards model, and seems more reasonable  
302 from a biological point of view. We must point out, however, that the magnitude of  
303 the variation in optimal clutch size induced by temporal covariates and random effects,

304 depends, to some extent, on our choice of link function. For the alternative logit link, the  
305 induced variations implied by the model would have been somewhat smaller and would  
306 go to zero in the limit of low survival where the logit tends to a log link. The logit and  
307 other link functions such as the probit, however, would not lead to regression coefficients  
308 having a common interpretation across episodes of selection of different length (Ergon  
309 et al., 2017). Hence, those link functions would not facilitate the fitting of sometimes  
310 more parsimonious null models for which these regression coefficients are constrained to  
311 a common value across several selection episodes.

Table 1: Parameters estimated with the best model retained (Table 2). Displayed are the meaning of the parameters, their notations and their estimates ( $\pm$  standard errors) for the three episodes of selection.

Meaning	Parameters	Symbol	Selection episode			Unit
			Egg to fledgling $s = 1$	Fledgling to recruit $s = 2$	Adult female $s = 3$	
Intercept, mortality		$\hat{\beta}_{0,s}$	$1.95 \pm 0.40$	$0.66 \pm 0.28$	0	
Effect of clutch size on mortality		$\hat{\beta}_{\text{clutchsize},s}$	$-0.170 \pm 0.080$	$0.076 \pm 0.055$	0	
Effect of winter temperature on mortality		$\hat{\beta}_{\text{temp},s}$	0	0	$-0.243 \pm 0.035$	$^{\circ}\text{C}^{-1}$
Effect of population density on mortality		$\hat{\beta}_{\text{dens},s}$	0	0	$0.0096 \pm 0.0028$	
Effect of altitude on mortality		$\hat{\beta}_{\text{alt},s}$	$0.00053 \pm 0.00028$	$0.00053 \pm 0.00028$	0	$\text{m}^{-1}$
Intercept, optimal egg-laying date		$\hat{\theta}_{0,s}$	$40.9 \pm 9.6$	$42.1 \pm 5.5$	0	d
Effect of population density on optimal date		$\hat{\theta}_{\text{dens},s}$	$-0.41 \pm 0.14$	$-0.41 \pm 0.14$	0	d
Effect of winter temperature on optimal date		$\hat{\theta}_{\text{temp},s}$	0	$8.9 \pm 3.5$	0	$\text{d } ^{\circ}\text{C}^{-1}$
Linear deterministic trend on optimal date		$\hat{\theta}_{t,s}$	$-1.37 \pm 0.59$	$-1.37 \pm 0.59$	0	$\text{d year}^{-1}$
Effect of altitude on optimal date		$\hat{\theta}_{\text{alt},s}$	$0.100 \pm 0.029$	$0.100 \pm 0.029$	0	$\text{d m}^{-1}$
Quadratic effect of egg-laying date		$\hat{\omega}_s$	$45 \pm 12$	$45 \pm 12$	0	
SD of random year effect		$\hat{\sigma}_s$	0	$0.144 \pm 0.050$	$0.144 \pm 0.050$	
SD of random territory effect		$\hat{\tau}_s$	$0.180 \pm 0.054$	$0.180 \pm 0.054$	$0.180 \pm 0.054$	
Beta-binomial dispersion parameter		$\hat{\gamma}_s$	$1.879 \pm 0.137$	$1.093 \pm 0.085$		
SD of second order random age effect		$\hat{\nu}_s$	0	0	$0.36 \pm 0.17$	

## 312 Results and discussion

313 Our statistical model applied to the dipper allows us to select the most parsimonious  
314 model (Table 1) among all the tested ones (Table 2). In this section, we explore in detail  
315 all the effects retained in the best model and discuss their implications.

Table 2: Model selection - Displayed are all the tested models derived from the best model retained, the difference  $\Delta\text{AIC}$  (respectively  $\Delta p$ ) in AIC (respectively in number of parameters  $p$ ) between each model and the best one and their description. Subscripts indicate the selection episode under consideration, i.e. 1 for the first episode from egg to fledgling stage, 2 for the second episode from fledgling to recruit stage and 3 for the third episode corresponding to the adult stage.

	$\Delta\text{AIC}$	$\Delta p$	Description
A	0	0	Best model, see Table 1
B	2.44	2	$\beta_{\text{age},1} \neq 0, \beta_{\text{age},2} \neq 0$ (different linear age effect for episodes 1 and 2)
C	1.97	1	$\beta_{\text{age},1} \neq 0$ (linear age effect for episode 1)
D	0.66	1	$\beta_{\text{age},1} = \beta_{\text{age},2} \neq 0$ (common linear age effect for episodes 1 and 2)
E	0.49	1	$\beta_{\text{age},2} \neq 0$ (linear age effect for episode 2)
F	0.94	1	$\beta_{\# \text{eggs},3}$
G	1.96	1	$\phi_{11} \neq 0$ ( $u_t \sim AR(1)$ )
H	2	1	$\sigma_{\zeta} > 0$ and $\phi_{12} \neq 0$ ( $u_t$ dependent on $\zeta_{t-1}$ )
I	2.97	2	$\beta_{\text{clutchsize}/\text{date},1} \neq 0, \beta_{\text{clutchsize}/\text{date},2} \neq 0$ (correlational selection of first kind)
J	1.56	1	$\beta_{\text{clutchsize}/\text{date},1} = \beta_{\text{clutchsize}/\text{date},2} \neq 0$ (correlational selection of first kind)
K	16.73	-6	Remove stab. sel. on $z$ (quadratic term in eq. 3) for episodes 1 and 2
L	0.5	1	$\omega_1 \neq \omega_2$ (strength of stabilizing selection different for episodes 1 and 2)
M	2.02	1	$\theta_{0,3} \neq 0$ (stabilizing selection for episode 3)
N	5.73	-1	$\theta_t = 0$ (remove trend in optimal laying date)
O	1.21	1	$\theta_{\text{clutchsize},1} = \theta_{\text{clutchsize},2} \neq 0$ (correlational selection of second kind)
P	2	1	$\sigma_{\zeta} > 0$ (random effect on optimal laying date)
Q	4	2	$\sigma_{\zeta} > 0$ and $\phi_{22} \neq 0$ ( $\zeta_t \sim AR(1)$ )
R	3.93	2	$\sigma_{\zeta} > 0$ and $\phi_{21} \neq 0$ ( $\zeta_t$ dependent on $u_{t-1}$ )
S	9.06	-1	$\theta_{\text{temp},2} = 0$ (remove temperature effect on optimal laying date)
T	7.31	-1	$\theta_{\text{dens},1} = \theta_{\text{dens},2} = 0$ (remove density effect on optimal laying date)
U	9.88	-1	$\theta_{\text{alt},1} = \theta_{\text{alt},2} = 0$ (remove altitude effect on laying date)
V	16.9	0	All age effects modelled as second order random walks
W	2.48	-2	$\beta_{\text{clutchsize},1} = \beta_{\text{clutchsize},2} = 0$
X	45.88	-1	$\beta_{\text{temp},3} = 0$
Y	0.7	1	$\beta_{\text{temp},2} \neq 0$
Z	24.39	0	$\beta_{\text{temp},2} = \beta_{\text{temp},3}$
AA	10.39	-1	$\beta_{\text{dens},3} = 0$
AB	1.46	1	$\beta_{\text{dens},2} \neq 0$
AC	0.98	1	$\beta_{\text{dens},1} \neq 0$
AD	11.62	0	$\beta_{\text{dens},1} \neq 0, \beta_{\text{dens},3} = 0$
AE	14.34	0	$\beta_{\text{dens},1} = \beta_{\text{dens},3} \neq 0$
AF	7.71	0	$\beta_{\text{dens},2} = \beta_{\text{dens},3}$
AG	12.36	0	$\beta_{\text{dens},2} \neq 0, \beta_{\text{dens},3} = 0$
AH	10.4	0	$\beta_{\text{dens},1} = \beta_{\text{dens},2} = \beta_{\text{dens},3}$
AI	2.66	2	$\beta_{\text{dens},1} \neq 0, \beta_{\text{dens},2} \neq 0, \beta_{\text{dens},3} \neq 0$
AJ	1.88	-1	$\beta_{\text{alt},1} = \beta_{\text{alt},2} = 0$
AK	2.9	1	$\beta_{\text{alt},1} = 0, \beta_{\text{alt},3} \neq 0$
AL	1.97	1	$\beta_{\text{alt},1} \neq 0, \beta_{\text{alt},2} \neq 0$
AM	1.97	1	$\beta_{\text{alt},3} \neq 0$
AN	3.41	0	$\beta_{\text{alt},2} = 0$
AO	0.94	0	$\beta_{\text{alt},1} = 0$
AP	2.11	0	$\beta_{\text{alt},1} = 0, \beta_{\text{alt},2} = \beta_{\text{alt},3} \neq 0$
AQ	3.74	0	$\beta_{\text{alt},1} = \beta_{\text{alt},2} = 0, \beta_{\text{alt},3} \neq 0$
AR	1.27	0	$\beta_{\text{alt},1} = \beta_{\text{alt},2} = \beta_{\text{alt},3}$
AS	3.95	2	$\beta_{\text{alt},1} \neq 0, \beta_{\text{alt},2} \neq 0, \beta_{\text{alt},3} \neq 0$
AT	1.03	1	$\beta_{\text{clutchsize}/\text{date},1} \neq 0$ (correlational selection of first kind)
AU	2	1	$\beta_{\text{clutchsize}/\text{date},2} \neq 0$ (correlational selection of first kind)
AV	2.13	1	$\theta_{0,1} \neq 0, \theta_{0,2} \neq 0, \theta_{0,3} \neq 0$
AW	1.99	1	$\theta_{t,1} \neq \theta_{t,2}$
AX	2.14	2	$\theta_{\text{clutchsize},1} \neq 0, \theta_{\text{clutchsize},2} \neq 0$ (correlational selection of second kind)
AY	2.03	0	$\theta_{\text{dens},1} = 0$
AZ	7.54	0	$\theta_{\text{dens},2} = 0$
BA	1.44	1	$\theta_{\text{dens},1} \neq 0, \theta_{\text{dens},2} \neq 0$
BB	2.96	0	$\theta_{\text{alt},1} = 0$
BC	11.38	0	$\theta_{\text{alt},2} = 0$
BD	1.32	1	$\theta_{\text{alt},1} \neq 0, \theta_{\text{alt},2} \neq 0$
BE	1.99	1	$\sigma_2 \neq 0, \sigma_3 \neq 0$
BF	2	1	$\sigma_1 \neq 0, \sigma_2 = \sigma_3 \neq 0$
BG	4	2	$\sigma_1 \neq 0, \sigma_2 \neq 0, \sigma_3 \neq 0$
BH	2.26	2	$\tau_1 \neq 0, \tau_2 \neq 0, \tau_3 \neq 0$
BI	1.8	1	$\tau_1 \neq 0, \tau_2 = 0, \tau_3 \neq 0$
BJ	-0.18	0	$\tau_1 = \tau_3 \neq 0, \tau_2 = 0$
BK	0.28	1	$\tau_1 = \tau_3 \neq 0, \tau_2 \neq 0$
BL	5.2	3	$\kappa_1 \neq 0, \kappa_2 \neq 0, \kappa_3 \neq 0$
BM	3.2	2	$\kappa_1 = 0, \kappa_2 \neq 0, \kappa_3 \neq 0$
BN	1.4	1	$\kappa_1 = \kappa_3 = 0, \kappa_2 \neq 0$
BO	1.28	1	$\kappa_2 = \kappa_3 \neq 0, \kappa_1 = 0$
BP	3.96	2	$\sigma_{\zeta} > 0$ and $\phi_{11} \neq 0$ and $\phi_{22} \neq 0$ ( $u_t$ and $\zeta_t \sim AR(1)$ )
BQ	4	2	$\sigma_{\zeta} > 0$ and $\phi_{12} \neq 0$ and $\phi_{21} \neq 0$
BR	7.96	4	$\sigma_{\zeta} > 0$ and $\phi_{11}, \phi_{12}, \phi_{21}$ and $\phi_{22} \neq 0$

## AGE EFFECTS ON SURVIVAL

316 The best model indicates no effect of mother age on offspring survival from egg to fledgling  
317 stage (episode 1) and from fledgling to recruit stage (episode 2) (see last row in Table 1,  
318 Fig. 2). Considering for instance a linear effect of female age on offspring survival does  
319 not improve the model fit (models B to E, Table 2). Indeed, model B that includes a  
320 different linear age effect for episodes 1 and 2 provides estimates not significantly different  
321 from zero ( $\hat{\beta}_{\text{age},1} = -0.009 \pm 0.039$  for episode 1 and  $\hat{\beta}_{\text{age},2} = -0.029 \pm 0.024$  for episode  
322 2). Because the probability of rearing a chick is independent of mother age, our findings  
323 indicate no senescence in maternal care. Note also that we do not detect senescence in  
324 clutch size (results not shown here). Mean survival on the first episode is close to 83%  
325 and drops around 7% on the second episode (Fig. 2). This high mean survival on episode  
326 1 compared to episode 2 (about 10 times higher) simply results from the much longer  
327 duration of the second episode. At the adult stage (episode 3), survival is age-dependent,  
328 increasing from age 1 to 4 and decreasing from age 4 onwards. This result indicates  
329 actuarial senescence (Fig. 2), caused by a progressive loss of cellular and physiological  
330 functions late in life (Williams, 1957; Hamilton, 1966). Senescence is pervasive in the wild  
331 (Nussey et al., 2013), and we provide here additional evidence in a short-lived species.

## ENVIRONMENTAL EFFECTS ON SURVIVAL

332 As expected for this species that strongly depends on open water for foraging, warmer  
333 winters favor adult survival (see third row in Table 1). Therefore, adult females are  
334 more likely to survive, stay and breed the next breeding season in the population when  
335 the winters are mild. Apparent survival from fledgling to recruit stage is not affected  
336 by winter temperatures (model Y, estimate = 0.043 (SE: 0.042)). In accordance with  
337 previous studies (Gamelon et al., 2017; Nilsson et al., 2011a; Sæther et al., 2000; Loison  
338 et al., 2002), high population density increases competition among individuals and thus  
339 mortality (and possibly dispersal rate), especially at the adult stage (see fourth row in  
340 Table 1). Including an effect of density on survival on other episodes does not improve  
341 the model fit (models AA to AI, Table 2). However, it is noteworthy that a model  
342 including an additional effect of density on episode 1 is close to the best model (model  
343 AC), but the low effect size indicates no density-dependent mortality on this first episode  
344 ( $\hat{\beta}_{\text{dens},1} = -0.003 \pm 0.003$ ). We thus do not find any evidence for density-dependent

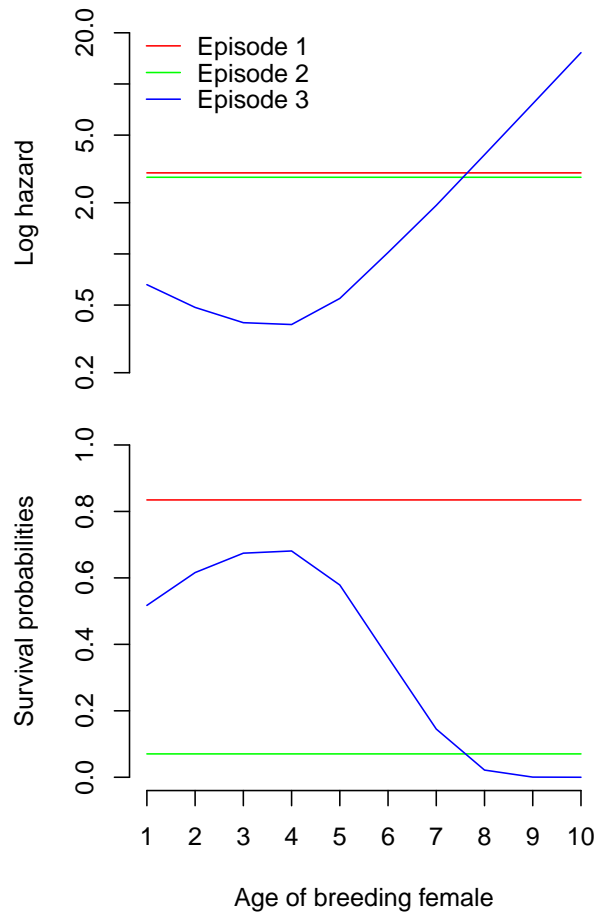


Figure 2: The log of the hazard (per year) (upper plot) and the resulting probability of surviving each episode (lower plot) as a function of mother age. Mean hazard and survival are computed at optimal egg-laying date and at the most frequent clutch size (5 eggs), at mean altitude, population density and winter temperature for each episode.

345 mortality at the offspring stage, contrary to some other passerine bird species such as  
 346 great tit (Reed et al., 2013a; Sæther et al., 2016) for which low densities are generally  
 347 associated with high offspring survival. Notice also that high altitudes negatively affect  
 348 offspring survival from egg to fledgling stage and also from fledgling to recruit stage (see  
 349 fifth row in Table 1).

### SURVIVAL AND VARYING SELECTIVE OPTIMA FOR CLUTCH SIZE

350 Mortality on episode 1 decreases in large clutches (see second row in Table 1, first axis  
 351 on Fig. 3A). However, because of high overall survival during the first short selection  
 352 episode, this does not translate to strong selection for large clutch sizes. Low mortality  
 353 on episode 1 in large clutches also means that females with large clutches have offspring  
 354 with the highest survival in the nest. This suggests individual heterogeneity, with some



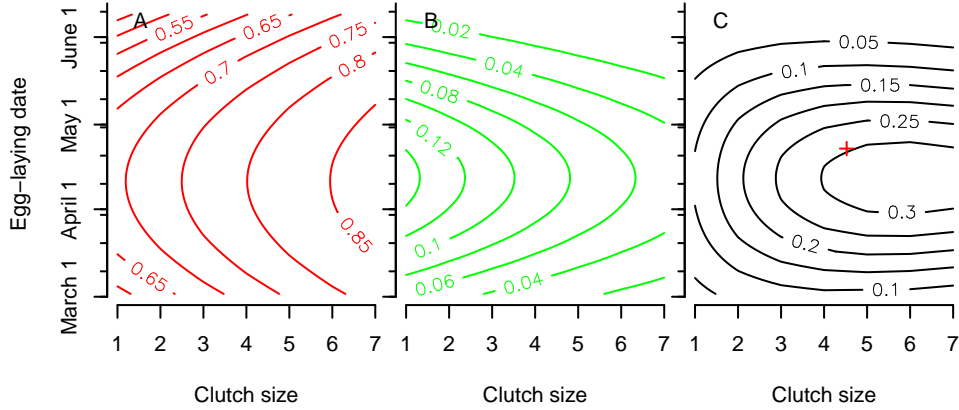


Figure 3: Estimated A) survival rate on episode 1 (from egg to fledgling stage, in red), B) survival rate on episode 2 (from fledgling to recruit stage, in green), and C) total fecundity rate (i.e. total survival rate from egg to recruit stage x clutch size, in black) as a function of egg-laying dates and clutch sizes. Mean vital rates are computed at average altitude, population density and winter temperature. The red cross indicates mean clutch size and mean egg-laying date observed during the study period. The minor tick marks are at the 10<sup>th</sup>, 20<sup>th</sup> and 30<sup>th</sup> of each month.

355 females performing better than others in terms of reproductive success. It is noteworthy  
 356 that females that lay a large number of eggs, not only during a given breeding event but  
 357 generally during the whole breeding season, do not pay direct survival costs. Indeed,  
 358 adding a term for the effect of the total number of eggs laid during the whole breeding  
 359 season in (5) does not provide any improvement as indicated by model F (Table 2,  $\hat{\beta}_{\text{eggs},3} =$   
 360  $-0.062 \pm 0.061$ ). For episode 2, the best model indicates that large clutch sizes increase  
 361 mortality (Table 1, Fig. 3B). This might be due to malnutrition and reduced parental  
 362 care (Noordwijk et al., 1980). As a result of lower survival during the second episode,  
 363 this translates to overall survival over the two first episodes combined decreasing with  
 364 increasing clutch size. Therefore, laying too many eggs is associated with increasing  
 365 offspring mortality. But obviously, laying too few eggs is not a successful breeding tactic  
 366 for a female either (Lack, 1954; Boyce and Perrins, 1987; Both et al., 2000). Here,  
 367 we provide evidence for an optimal clutch size maximizing offspring survival and more  
 368 generally overall fecundity rate estimated to be 6.69 eggs. The width of the fitness  
 369 function with respect to clutch sizes,  $\omega_{y_0}$ , is estimated to be between 4.89 and 5.89 eggs  
 370 in different years, being approximately proportional to the optimum in any given year.  
 371 The fitness function is also somewhat asymmetric with respect to clutch size (Fig. 3C).  
 372 This high value for the width of the fitness function indicates weak stabilizing selection

Table 3: Observed frequencies of different clutch sizes in the population.

Clutch size	0	1	2	3	4	5	6	7
Frequency	57	1	52	69	186	636	192	4

373 for optimal clutch size, as also illustrated in Fig. 3C.

374 We find that this optimal clutch size varies through time between 6.19 and 7.53 eggs  
 375 (Fig. 4A). This agrees quite well with observed clutch sizes in the population, mainly  
 376 ranging between 4 and 6 eggs (Table 3). The mean observed clutch size of 4.52 eggs is  
 377 somewhat smaller than the estimated overall optimum. However, this observed mean is  
 378 influenced by the left tail of the frequency distribution, which might reflect total failure  
 379 of some females. Interestingly, the observed mean clutch size maximizes total fecundity  
 380 rate (Fig. 3C).

381 According to our model, the temporal variation in optimal clutch size is induced by  
 382 random variations in survival during episode 2 modelled by the random effect term  $\sigma_2 u_t$   
 383 in (5) ( $\hat{\sigma}_s = 0.144 \pm 0.050$  for episode  $s = 2$ , Table 1) such that larger clutches turn out to  
 384 be favoured in years with high survival. Modelling  $u_t$  as an autoregressive process does  
 385 not improve the model (Table 2, model G), that is, we find no evidence for autocorrelation  
 386 in these variations.

387 Finally, the models including correlational selection between egg-laying date and  
 388 clutch size on maximum survival on the first two episodes do not perform better than  
 389 the best model (models I and J, Table 2). There is thus no evidence for correlational  
 390 selection between clutch size and egg-laying date on offspring survival.

391 It should be noted that we have treated clutch size as a trait on which selection  
 392 operates, rather than as a fitness component. This is reasonable because in many bird  
 393 populations, recruitment (function of clutch size) has a limited influence on the popula-  
 394 tion growth rate (see Sæther et al. (2016) for a comparative analysis). Instead, a large  
 395 proportion of the temporal variance in population change of temperate passerines is due  
 396 to variation in survival (Gould and Nichols, 1998; Sæther et al., 2004), dipper being no  
 397 exception (Loison et al., 2002). It is therefore relevant not to consider clutch size as a  
 398 fitness component but rather as a trait under selection.

## SURVIVAL AND VARYING SELECTIVE OPTIMA FOR EGG-LAYING DATE

399 We find that egg-laying date is under stabilizing selection. Indeed, removing stabilizing  
400 selection in (3) for episodes 1 and 2 does not provide any improvement in AIC (model K,  
401 Table 2). Interestingly,  $\omega_s$ , that determines the strength of stabilizing selection, is similar  
402 for episodes 1 and 2 (see eleventh row in Table 1). Indeed, estimating two different values  
403 ( $\hat{\omega}_1 = 33.37 \pm 9.1$  and  $\hat{\omega}_2 = 49.11 \pm 12.8$ ) does not improve the model fit (model L, Table  
404 2). Given the much longer duration of episode 2, however, most of the selection happens  
405 during this episode. The approximate width of the overall fitness function  $\omega_z$  varies  
406 between 27.76 and 32.01 days reflecting strong stabilizing selection on egg-laying date,  
407 as illustrated in Fig. 3C. Adding stabilizing selection for episode 3 does not improve the  
408 model fit (model M, Table 2), meaning laying eggs early or late in the breeding season  
409 has no effect on mother survival. In contrast, in a tropical parrot, the green-rumped  
410 parrotlet (*Forpus passerinus*), there is strong selection on egg-laying date through adult  
411 survival, such that females that breed early in the season exhibit lower survival to the  
412 next breeding season (Tarwater and Beissinger, 2013). This discrepancy between our  
413 finding and Tarwater and Beissinger (2013) might be explained by the contrasting life  
414 history strategies of these two species. Indeed, parrotlets have a much slower pace of life  
415 than the dipper. In slow-living species, adult survival has the highest contribution to  
416 population growth rate and is thus expected to be particularly canalized (Gaillard and  
417 Yoccoz, 2003). This explains why viability selection is important in slow-living species like  
418 parrotlets compared to short-lived species such as dipper. As a consequence, our results  
419 indicate that the strong stabilizing selection on egg-laying date operates only through  
420 offspring survival.

421 The timing of egg laying and more generally the timing of reproduction is critical  
422 in many species (Price et al., 1988) because hatching/birth should match with good  
423 environmental conditions in terms of weather and/or food resources. Otherwise, offspring  
424 survival may be jeopardized. Our findings provide evidence for an optimal laying date in  
425 the dipper with an overall mean egg-laying date maximizing survival from egg to fledgling  
426 stage estimated to be  $\hat{\theta}_{0,1} = 40.9 \pm 9.6$  days after the 1<sup>st</sup> of March (i.e. mid-April, Fig. 3A,  
427 see sixth row in Table 1) and with an overall mean egg-laying date maximizing survival  
428 from fledgling to recruit stage estimated to be  $\hat{\theta}_{0,2} = 42.1 \pm 5.5$  days after the 1<sup>st</sup> of March  
429 (Fig. 3B, Table 1). This translates to a mean egg-laying date maximizing total fecundity

430 rate estimated to be April 12 (Fig. 3C). This mean optimal date is slightly earlier than the  
431 mean laying date actually observed over the study period (April 22 for  $n = 741$  clutches  
432 for which information on egg-laying date was available, second axis, Fig. 3C).

433 Observed egg-laying dates have advanced at a rate of  $0.15 \pm 0.07$  days/year during  
434 the 35-year period (Fig. 4B, grey dots). Interestingly, we find a significant trend towards  
435 earlier optimal dates at a rate of  $\hat{\theta}_{t,s} = 1.37 \pm 0.59$  days/year (Table 1). Removing the  
436 trend worsens model fit considerably (model N, Table 2). This estimate seems somewhat  
437 large and would imply an advance in the optimum of 47 days over the course of the study.  
438 The lower confidence limit of 7.24 days for this advance (based on approximate normality  
439 of  $\hat{\theta}_{t,s}$ ) seems more reasonable and comparable to the observed phenotypic change of 8.8  
440 days on average (ranging from 4 to 17 days), recorded in other bird populations over a  
441 25-year period (Crick et al., 1997).

442 Importantly, in addition to the linear trend, the estimated optimal laying date (at the  
443 average altitude) varies over years between February 8 and May 24 (Fig. 4B, black line)  
444 as a result of the effects of winter temperature and population density. Removing the  
445 effect of environmental covariates on the optimal laying date worsens model fit (models S  
446 and T, Table 2) thus providing strong evidence for environmental sensitivity of selection.  
447 In particular, when the winter following the breeding season is warm, the optimal laying  
448 date maximizing survival from fledgling to recruit stage (i.e. episode 2) is delayed by  
449  $\hat{\theta}_{\text{temp},2} = 8.9 \pm 3.5$  days/ $^{\circ}\text{C}$  (see eighth row in Table 1). One can hypothesize that when  
450 the subsequent winter is mild, offspring produced at late laying dates during the previous  
451 breeding season are disproportionately more likely to survive, thus generating selection  
452 for later laying dates. While including the same effect of winter temperature on the  
453 optimal laying date of episode 1 led to a slight improvement in AIC, such a model would  
454 clearly be biologically unrealistic as survival on episode 1 has to be causally independent  
455 of the following winter conditions (Fig. 1). This model alternative was thus excluded  
456 from consideration (see Burnham and Anderson, 2002, ch. 6.8.7). Under our best model,  
457 different optima for episodes 1 and 2 are therefore estimated (Fig. 4B, red and green  
458 curves). The overall optimum (black curve) is approximately an average of the optima  
459 for each of the two episodes, weighted by the respective strength of stabilizing selection for  
460 each episode (this approximation would be exact if the fitness functions for each episode  
461 were exactly Gaussian). Indeed, as can be seen from Fig. 4B, the overall optimum always

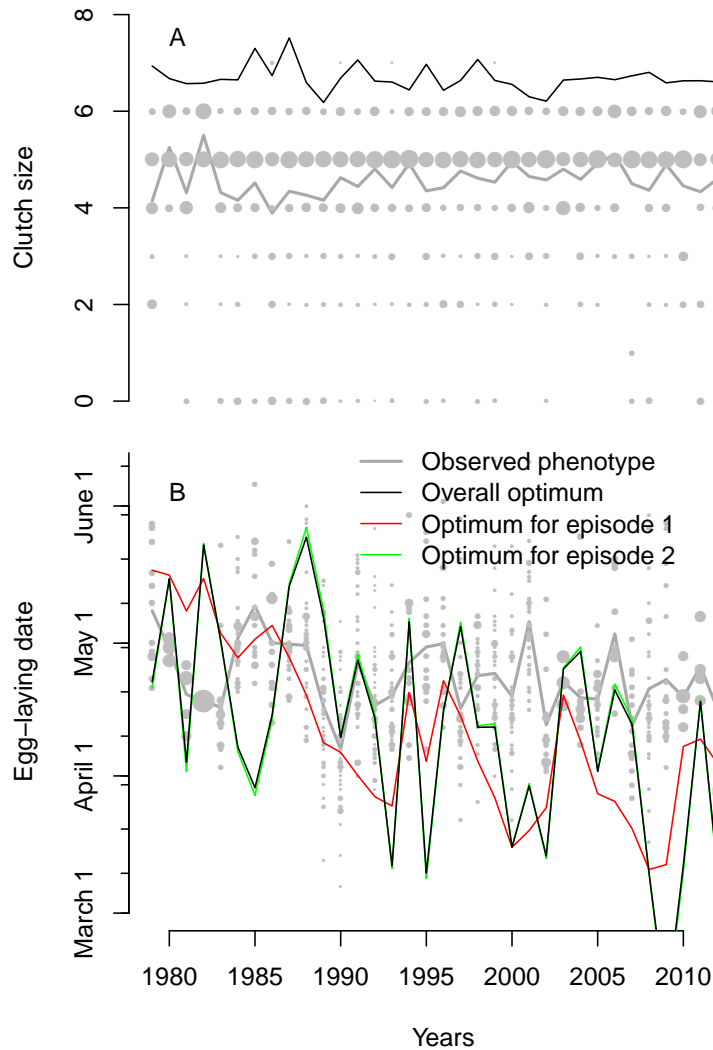


Figure 4: Estimated annual optimal clutch size (plot A) and laying dates (plot B) for the best model (Table 1). In plot B, the red and green curves are the estimated optimal laying dates in terms of survival during episodes 1 and 2, respectively, and the black curve the optimal laying date for both episodes combined. Note that the overall optima (black curve) nearly coincide with the optima of the episode 2. All optima are estimated at the average altitude and for the most frequent clutch size (5 eggs). Grey lines represent mean phenotypic values and the size of the grey dots the frequencies of different phenotypes in the total population. The minor tick marks in plot B are located as in Fig. 3.

462 falls between the green and red curve but much closer to the green curve given the much  
 463 stronger stabilizing selection during episode two.

464 Interestingly, previous studies on other passerine bird species have shown that selec-  
 465 tion on egg-laying date often depends on the timing of the peak in caterpillars, the main  
 466 food resources, itself closely related to spring temperatures (e.g. in great tits, (Chevin  
 467 et al., 2015; Visser et al., 2006; Reed et al., 2013b)). Here, we do not find statistical evi-  
 468 dence for the effect of spring conditions (date of ice break-up) on the optimal egg-laying

469 date. Based on the upper confidence limit, the regression coefficient is considerably  
470 smaller than its expected value (see Appendix A.2), suggesting that other environmental  
471 factors are more important for the dipper’s optimal laying date. This discrepancy be-  
472 tween our findings and previous works may be explained by the difference in the biology  
473 of these species. While great tit strongly depends on insect availability during spring,  
474 food resources are available on a larger time window (over spring and summer) for the  
475 dipper.

476 We also find that optimal egg-laying date is density-dependent, with high densities  
477 favoring earlier optimal date (see seventh row in Table 1). Despite a large amount of the-  
478 oretical and experimental works on the role of density as a selective agent (Charlesworth,  
479 1994; Engen et al., 2013; De Lisle and Rowe, 2013; Shaw, 1986), empirical evidence in the  
480 wild remains scarce (but see Sæther et al. (2016) for great tits *Parus major* and Hunter  
481 et al. (2018) for Soay sheep). In green-rumped parrotlets for instance, late breeding dates  
482 are selected for when the density is low, through enhanced adult survival (Tarwater and  
483 Beissinger, 2013). Similarly, a recent study has shown that in North American red squir-  
484 rels (*Tamiasciurus hudsonicus*), high population densities might increase the strength of  
485 selection for earlier birth dates (Fisher et al., 2017), favoring successful recruitment for  
486 juveniles. Here, our findings provide additional support for a key role of density as a  
487 selective agent in the wild.

488 Having included the temporal covariates winter temperature and population density,  
489 we do not find any latent variations in optimal laying date as indicated by the lack of  
490 improvement in AIC when including  $\zeta_t$  (corresponding to random variation in the optimal  
491 laying date) as a iid random effect (model P, Table 2). We also considered including  $\zeta_t$   
492 distributed as an AR(1) process ( $\phi_{22} \neq 0$ , model Q), cross-correlated with  $u_{t-1}$  (corre-  
493 sponding to variation in overall survival) ( $\phi_{21} \neq 0$ , model H) or cross-correlated with  $u_{t+1}$   
494 ( $\phi_{12} \neq 0$ , model R) but neither of these model alternatives led to any improvement in  
495 AIC.

496 Through their estimated joint effect on the optimal laying date, winter temperature  
497 and population density induce autocorrelation on the deviations of the optimal laying  
498 dates from the estimated linear trend. For episode 1, only influenced by population  
499 density and no additional latent random process, the autocorrelation function of the  
500 optimum is identical to that of population density, with a significant autocorrelation of

501 0.58 at lag 1 (Figure A.3). For episode 2, despite also being influenced by population  
502 density in addition to winter temperature, the resulting optimum given by  $\hat{\theta}_{\text{dens},2}\text{dens}_t +$   
503  $\hat{\theta}_{\text{temp},2}\text{temp}_t$  exhibited no significant autocorrelation, mainly because of the larger effect  
504 of winter temperature (exhibiting no autocorrelation, Fig. A.2) accounting for 75% of  
505 the total variance in the optimum. Since the optimal laying date for episodes 1 and 2  
506 combined almost coincides with the optimum for episode 2 (Fig. 4B, black and green  
507 curves), the same applies to the corresponding autocorrelation function.

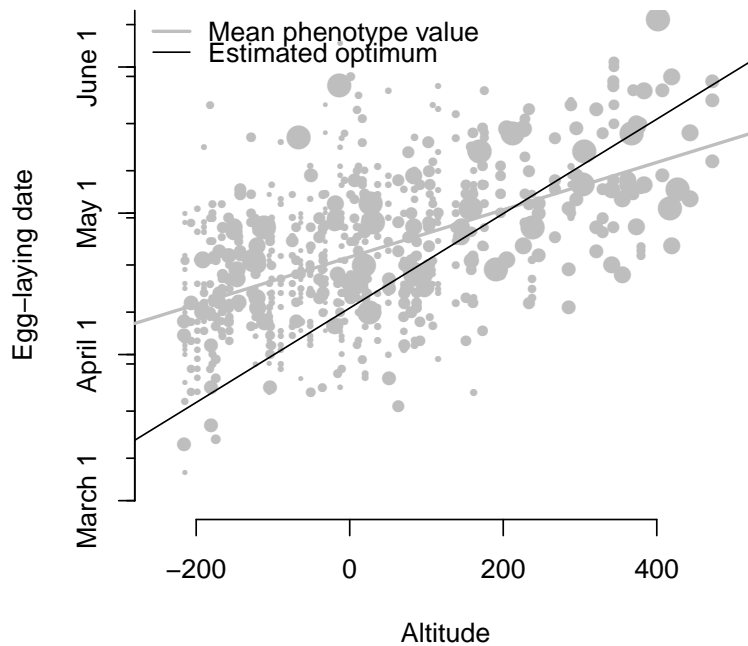


Figure 5: Estimated optimal (black line) and observed mean (grey line) egg-laying dates as a function of altitude (mean centered), together with observed egg-laying dates (grey dots). Dot sizes illustrates the number of individuals exhibiting the same laying date at a given altitude. Estimated optimal egg-laying dates are computed at average population density and winter temperature. The minor tick marks are located as in Fig. 3.

508 Finally, in addition to the dependency on winter temperature and population density,  
509 we find that optimal and observed egg-laying dates depend on altitude (model U, Table  
510 2), occurring later at high altitudes (see tenth row in Table 1, black line on Fig. 5).  
511 This result gives clear support for spatial variation in phenotypic selection and provides  
512 empirical evidence for adaptation along an environmental gradient. To what extent this  
513 can be explained by the altitudinal gradient in temperatures would require fine-grained  
514 data on local temperature, unfortunately not available. Interestingly, the slope in the  
515 mean phenotype based on the egg-laying dates in the complete data (grey line in Fig. 5,

516  $n = 741$ ) is significantly shallower than the estimated slope in the optimum (black line  
517 in Fig. 5, one sided  $Z$ -test,  $p$ -value = 0.041). Thus, laying tends to occur too late at  
518 low altitudes and too early at high altitude relative to the estimated optimal laying date.  
519 This suggests maladaptation at the extreme ends of the gradient, possibly caused by  
520 gene flow as in Kirkpatrick and Barton's (1997) model of evolution of species' range or  
521 by imperfect plasticity (Tufto, 2000; Chevin and Lande, 2011; Gienapp et al., 2014).

## MODEL WITHOUT TEMPORAL COVARIATES

522 Our statistical approach can be used when environmental covariates involved in the vary-  
523 ing optima are missing. As an illustration, we removed population density and winter  
524 temperatures as temporal covariates as well as the trend in optimal egg-laying dates and  
525 tested different models for the random effects instead (see Table A.1). The best model  
526 (see Fig A.1 and Table A.2) included both latent variations in overall mortality (through  
527 the  $u_t$  term inducing variation in optimal clutch size as before) and in optimal laying  
528 date (through a significant  $\zeta_t$  term, model BT vs. best model,  $\hat{\sigma}_\zeta = 19 \pm 12$  days), jointly  
529 following a vector autoregressive process (see (6)). In line with the absence of autocorre-  
530 lation in the optima induced by the temporal covariates for model A in Table 2 (Fig A.3),  
531 we found no evidence for autocorrelation in optima via  $u_t$  and  $\zeta_t$  for the model without  
532 temporal covariates (models BU and BV in Table A.1). Surprisingly, this model includes  
533 a negative autoregressive coefficient  $\hat{\phi}_{12} = -0.45 \pm 0.95$ , making  $\zeta_{t-1}$  (optimal laying date  
534 in year  $t - 1$ ) negatively correlated with  $u_t$  (the hazard in year  $t$ ). A possible explanation  
535 is that an unknown temporal covariate influences the optimal laying date with a delayed  
536 effect on survival.

537 The moving optimal laying dates and clutch sizes estimated with the best models  
538 with and without temporal covariates are generally in accordance (Fig. 4 vs. Fig A.1).  
539 The parameter estimates provided by the model without temporal covariates are close  
540 to the ones of our best model for the effects of clutch size and altitude on mortality  
541 (comparison between Table 1 and Table A.2). However, including temporal covariates  
542 improves the precision of most parameter estimates (Table 1 vs Table A.2). In particular,  
543 the precision of the estimates of optimal egg-laying date strongly differs between the two  
544 models, being equal to  $\hat{\theta}_{0,1} = 22 \pm 29$  days after the 1<sup>st</sup> of March (i.e. March  $30 \pm 29$   
545 days) when excluding temporal covariates and to  $40.9 \pm 9.6$  days (i.e. April  $10 \pm 9.6$



546 days) for the first episode and to  $\hat{\theta}_{0,2} = 42.1 \pm 5.5$  days (i.e. April  $12 \pm 5.5$  days) for the  
547 second episode when including them. Similarly,  $\omega_s$  for episodes 1 and 2 is estimated to be  
548  $45 \pm 12$  and  $81 \pm 45$  for the best models with and without temporal covariates (translating  
549 to widths  $\omega_z$  of the Gaussian approximation of the fitness function ranging from 28 to  
550 32 days and 39 to 62 days, respectively). This difference also explains the shift in the  
551 estimated optimal laying dates (Fig. 4B vs. Fig A.1B) since the location of the optimum  
552 is, to some extent, estimated through extrapolation of the quadratic model of the effect  
553 of laying date on the hazard beyond the range of observed phenotypic values (see (3)).

## 554 Conclusion

555 Understanding how life history traits vary in time and space and determining the selective  
556 forces behind this variation is one of the central issues in ecology and evolution. Instead  
557 of working within the classical framework of Lande and Arnold (1983), we have extended  
558 the statistical glm-like approach used in several works (Janzen and Stern, 1998; Shaw  
559 and Geyer, 2010; Chevin et al., 2015), by allowing movements of the underlying fitness  
560 function. These movements can be induced through observed environmental covariates  
561 or latent processes following different types of plausible autoregressive models. This  
562 extension leads to a non-linear latent variable statistical model, efficiently handled thanks  
563 to modern statistical software. A particular advantage of our approach is that all the data  
564 for several episodes are utilized in a single joint model. Also, correlational selection on  
565 multiple traits and correlated optima are accommodated as possible model alternatives.

566 The statistical approach we have used leads to models of stabilizing selection that  
567 are different from the standard Gaussian fitness function frequently used in theoretical  
568 models. Our approach is similar to how a quadratic effect of a trait in the logistic  
569 regression approach of Janzen and Stern (1998) translates into a non-Gaussian, plateauing  
570 fitness function if survival is high. Our view is that such non-Gaussian functions are  
571 more realistic because they reflect the constraint that survival probabilities necessarily  
572 have an upper bound of one. The difference from our approach is in the exact link  
573 function used (the logit link versus the loglog link corresponding to our proportional  
574 hazard assumption). The same type of argument can be made for how non-Gaussian  
575 stabilizing selection on clutch size emerges in our model. This being said, if survival

576 is low, for example during selection on laying date in episode  $s = 2$ , and not strongly  
577 dependent on clutch size, the resulting fitness function is well approximated by a Gaussian  
578 function with widths  $\omega_z$  and  $\omega_{y0}$  easily derived from the basic parameters of the model  
579 (Table 1). It should also be noted that when estimated optima fall outside the range  
580 of observed phenotypic values, the existence of an optimum is not an inference drawn  
581 from the data alone. Instead, we assume that an optimum exists and this is a reasonable  
582 assumption. For instance, based on a priori biological knowledge, we know that the  
583 reproductive success of a female that starts breeding too early in the season under harsh  
584 winter conditions will be low. The locations of optima are in turn estimated based on  
585 the most parsimonious model of the fitness curvature supported by the data.

586 In our dipper case study, using model selection criteria to choose between a large  
587 number of alternative models, we find evidence for varying selective optima on two key life-  
588 history traits. Spatio-temporal variation in optimal laying dates is induced by variation in  
589 altitude, winter temperature and population density. Optimal clutch sizes, resulting from  
590 an estimated trade-off between offspring number and survival, exhibit similar random  
591 variations over time. Selection on these traits mainly operates through low survival from  
592 fledgling to recruit stage (episode 2), paralleling the key role of this life stage as a driver  
593 of fluctuations in avian population dynamics (Sæther et al., 2016). For adult survival, we  
594 clearly show that senescence occurs but we find no cost of large clutch sizes or selection  
595 on laying dates via adult survival.

596 Our statistical approach can also be used when environmental covariates involved in  
597 the varying optima are missing. Indeed, the moving optimal laying dates and clutch  
598 sizes estimated with the models with and without temporal covariates are generally in  
599 accordance (Fig. 4 vs. Fig A.1). However, including relevant environmental covariates  
600 is important to improve the overall precision of parameter estimates and of course, to  
601 identify the agents of selection.

602 Our main objective in the present study has been to estimate the pattern of varying  
603 selective optima acting at various life stages. We are not aware of any simple theory for  
604 how this translates to variation in selection acting over the whole lifespan of an individual.  
605 It is noteworthy that the variation in mean phenotypes through space and time in many  
606 respects are qualitatively similar to variation in the estimated optima. However, we  
607 cannot conclude that these patterns in mean phenotypes are consistent with the pattern

608 of varying selection estimated by our method. Further work is needed to quantify the  
609 expected evolutionary response that is expected from our model. Given the complications  
610 of age-structured and density-dependent models (Engen and Saether, 2017), this is likely  
611 feasible only via simulation-based approaches.

## 612 Literature Cited

613 Agresti, A., 2002. *Categorical Data Analysis*. Wiley.

614 Arnold, S. J. and M. J. Wade, 1984. On the measurement of natural and sexual selection:  
615 theory. *Evolution* 38:709–719.

616 Bell, G., 2010. Fluctuating selection: the perpetual renewal of adaptation in variable  
617 environments. *Philosophical transactions of the Royal Society of London. Series B,*  
618 *Biological sciences* 365:87–97.

619 Both, C., J. M. Tinbergen, and M. E. Visser, 2000. Adaptive density dependence of avian  
620 clutch size. *Ecology* 81:3391–3403.

621 Boyce, M. S. and C. M. Perrins, 1987. Optimizing great tit clutch size in a fluctuating  
622 environment. *Ecology* 68:142–153.

623 Bull, J., 1987. Evolution of phenotypic variance. *Evolution* 41:303–315.

624 Bürger, R., 1999. Evolution of genetic variability and the advantage of sex and recombina-  
625 tion in changing environments. *Genetics* 153:1055–69.

626 Burnham, K. and D. Anderson, 2002. *Model Selection and Multimodel Inference: A*  
627 *Practical Information-Theoretic Approach*, second edition. Springer, New York.

628 Caswell, H., 2001. *Matrix Population Models*. Sinauer, Sunderland, Massachusetts.

629 Charlesworth, B., 1993. Directional selection and the evolution of sex and recombination.  
630 *Genetical Research* 61:205–224.

631 ———, 1994. *Evolution in age-structured populations*. 2 edition ed. Cambridge Univer-  
632 sity Press, Cambridge England ; New York.

- 633 Chevin, L.-M., 2013. Genetic constraints on adaptation to a changing environment.  
634 *Evolution* 67:708–721.
- 635 Chevin, L.-M., O. Cotto, and J. Ashander, 2017. Stochastic Evolutionary Demography  
636 under a Fluctuating Optimum Phenotype. *Am. Nat.* 190:786–802.
- 637 Chevin, L.-M. and B. Haller, 2014. The temporal distribution of linear gradients under  
638 selection for an optimum. *Evolution* 68:3381–3394.
- 639 Chevin, L.-M. and R. Lande, 2011. Adaptation to marginal habitats by evolution of  
640 increased phenotypic plasticity. *J. Evol. Biol.* 24:1462–76.
- 641 Chevin, L.-M., R. Lande, and G. M. Mace, 2010. Adaptation, plasticity, and extinction  
642 in a changing environment: towards a predictive theory. *PLoS biology* 8:e1000357.
- 643 Chevin, L.-M., M. E. Visser, and J. Tufto, 2015. Estimating the variation, autocorrelation'  
644 and environmental sensitivity of phenotypic selection. *Evolution* 69:2319–2332.
- 645 Crick, H. Q. P., C. Dudley, D. E. Glue, and D. L. Thomson, 1997. UK birds are laying  
646 eggs earlier. *Nature* 388:526.
- 647 De Lisle, S. P. and L. Rowe, 2013. Correlated Evolution of Allometry and Sexual Dimor-  
648 phism across Higher Taxa. *Am. Nat.* 182:630–9.
- 649 Endler, J. A., 1977. Geographic variation, speciation, and clines. Princeton University  
650 Press.
- 651 ———, 1986. Natural selection in the wild. Princeton University Press, Princeton, N.J.
- 652 Engen, S., R. Lande, and B.-E. Saether, 2013. A quantitative genetic model of r- and  
653 K-selection in a fluctuating population. *Am. Nat.* 181:725–36.
- 654 Engen, S., R. Lande, and B.-E. Sæther, 2011. Evolution of a plastic quantitative trait in  
655 an age-structured population in a fluctuating environment. *Evolution* 65:2893–2906.
- 656 Engen, S. and B.-E. Saether, 2017. r- and K-selection in fluctuating populations is  
657 determined by the evolutionary trade-off between two fitness measures: Growth rate  
658 and lifetime reproductive success. *Evolution* 71:167–173.

659 Engen, S., B.-E. Saether, T. Kvalnes, and H. Jensen, 2012. Estimating fluctuating selec-  
660 tion in age-structured populations. *J. Evol. Biol.* 25:1487–99.

661 Ergon, T., Ø. Borgan, C. R. Nater, and Y. Vindenes, 2017. Study mortality with hazard  
662 rates, not probabilities. *bioRxiv* .

663 Fisher, D. N., S. Boutin, B. Dantzer, M. M. Humphries, J. E. Lane, and A. G. McAdam,  
664 2017. Multilevel and sex-specific selection on competitive traits in North American red  
665 squirrels. *Evolution* 71:1841–1854.

666 Gaillard, J.-M., A. Viallefont, A. Loison, and M. Festa-Bianchet, 2004. Assessing senes-  
667 cence patterns in populations of large mammals. *Animal Biodiversity and Conservation*  
668 27:47–58.

669 Gaillard, J.-M. and N. G. Yoccoz, 2003. Temporal variation in survival of mammals: a  
670 case of environmental canalization? *Ecology* 84:3294–3306.

671 Gamelon, M., V. Grøtan, A. L. K. Nilsson, S. Engen, J. W. Hurrell, K. Jerstad,  
672 A. Phillips, O. W. Røstad, T. Slagsvold, B. Walseng, N. C. Stenseth, and B.-E. Sæther,  
673 2017. Interactions between demography and environmental effects are important de-  
674 terminants of population dynamics. *Science Advances* 3:e1602298.

675 Gienapp, P., T. E. Reed, and M. E. Visser, 2014. Why climate change will invariably  
676 alter selection pressures on phenology. *Proceedings. Biological sciences / The Royal*  
677 *Society* 281:755–766.

678 Gould, W. R. and J. D. Nichols, 1998. Estimation of temporal variability of survival in  
679 animal populations. *Ecology* 79:2531–2538.

680 Grant, P. and B. Grant, 2002. Unpredictable evolution in a 30-year study of Darwin’s  
681 finches. *Science* 296:707–711.

682 Green, P. J. and B. W. Silverman, 1994. *Nonparametric Regression and Generalized*  
683 *Linear Models: A roughness penalty approach*. Chapman & Hall.

684 Hamilton, W. D., 1966. The moulding of senescence by natural selection. *Journal of*  
685 *theoretical biology* 12:12–45.

- 686 Hansen, T. F., 1997. Stabilizing selection and the comparative analysis of adaptation.  
687 *Evolution* 51:1341–1351.
- 688 Hedrick, P., M. Ginevan, and E. Ewing, 1976. Genetic polymorphism in heterogeneous  
689 environments. *Annual Review of Ecology and ...* Pp. 1–32.
- 690 Hereford, J., 2009. A quantitative survey of local adaptation and fitness trade-offs. *Amer-*  
691 *ican naturalist* 173:579–88.
- 692 Hunter, D. C., J. M. Pemberton, J. G. Pilkington, and M. B. Morrissey, 2018. Quantifica-  
693 tion and decomposition of environment-selection relationships. *Evolution* 72:851–866.
- 694 Janzen, F. and H. Stern, 1998. Logistic regression for empirical studies of multivariate  
695 selection. *Evolution* 52:1564–1571.
- 696 Kirkpatrick, M. and N. H. Barton, 1997. Evolution of a Species' Range. *Am. Nat.*  
697 150:1–23.
- 698 Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. Bell, 2016. TMB: Automatic  
699 Differentiation and Laplace Approximation. *Journal of Statistical Software* 70:1–21.
- 700 Lack, D. L., 1954. *The natural regulation of animal numbers*. Clarendon Press.
- 701 Lande, R., 2007. Expected relative fitness and the adaptive topography of fluctuating  
702 selection. *Evolution* 61:1835–1846.
- 703 ———, 2009. Adaptation to an extraordinary environment by evolution of phenotypic  
704 plasticity and genetic assimilation. *J. Evol. Biol.* 22:1435–46.
- 705 Lande, R. and S. J. Arnold, 1983. The measurement of selection on correlated characters.  
706 *Evolution* 37:1210–1226.
- 707 Lande, R. and S. Shannon, 1996. The role of genetic variation in adaptation and popu-  
708 lation persistence in a changing environment. *Evolution* 50:434–437.
- 709 Loison, A., B.-E. Sæther, K. Jerstad, and O. W. Røstad, 2002. Disentangling the sources  
710 of variation in the survival of the European dipper. *Journal of Applied Statistics*  
711 29:289–304.

712 Marzolin, G., A. Charmantier, and O. Gimenez, 2011. Frailty in state-space models:  
713 Application to actuarial senescence in the Dipper. *Ecology* 92:562–567.

714 Morrissey, M. B. and J. D. Hadfield, 2012. Directional selection in temporally replicated  
715 studies is remarkably consistent. *Evolution* 66:435–42.

716 Newton, I., 1998. *Population Limitation in Birds*. Academic Press.

717 Nilsson, A. L. K., E. Knudsen, K. Jerstad, O. W. Røstad, B. Walseng, T. Slagsvold, and  
718 N. C. Stenseth, 2011a. Climate effects on population fluctuations of the white-throated  
719 dipper *Cinclus cinclus*. *The Journal of Animal Ecology* 80:235–243.

720 Nilsson, A. L. K., L. C. Stige, K. Jerstad, O. W. Røstad, T. Slagsvold, E. Knudsen,  
721 B. Walseng, and N. C. Stenseth, 2011b. To make the most of what we have: extracting  
722 phenological data from nestling measurements. *International Journal of Biometeorology*  
723 55:797–804.

724 Noordwijk, A. J. V., J. H. V. Balen, and W. Scharloo, 1980. Genetic and environmental  
725 variation in clutch size of the great tit (*Parus major*). *Netherlands Journal of Zoology*  
726 31:342–372.

727 Nussey, D. H., H. Froy, J.-F. Lemaître, J.-M. Gaillard, and S. N. Austad, 2013. Senescence  
728 in natural populations of animals: widespread evidence and its implications for bio-  
729 gerontology. *Ageing Research Reviews* 12:214–225.

730 Ozgul, A., D. Z. Childs, M. K. Oli, K. B. Armitage, D. T. Blumstein, L. E. Olson,  
731 S. Tuljapurkar, and T. Coulson, 2010. Coupled dynamics of body mass and population  
732 growth in response to environmental change. *Nature* 466:482–485.

733 Ozgul, A., S. Tuljapurkar, T. G. Benton, J. M. Pemberton, T. H. Clutton-Brock, and  
734 T. Coulson, 2009. The dynamics of phenotypic change and the shrinking sheep of St.  
735 Kilda. *Science* 325:464–467.

736 Pelletier, F., T. Clutton-Brock, J. Pemberton, S. Tuljapurkar, and T. Coulson, 2007. The  
737 evolutionary demography of ecological change: linking trait variation and population  
738 growth. *Science* 315:1571–1574.

739 Pemberton, J. M., 2010. Evolution of quantitative traits in the wild: mind the ecology.  
740 Philosophical Transactions of the Royal Society B: Biological Sciences 365:2431–2438.

741 Phillips, P. and S. Arnold, 1989a. Visualizing multivariate selection. Evolution 43:1209–  
742 1222.

743 Phillips, P. C. and S. J. Arnold, 1989b. Visualizing multivariate selection. Evolution  
744 43:1209–1222.

745 Price, T., M. Kirkpatrick, and S. Arnold, 1988. Directional selection and the evolution  
746 of breeding date in birds. Science 240:798–799.

747 Reed, T. E., V. Grøtan, S. Jenouvrier, B.-E. Sæther, and M. E. Visser, 2013a. Population  
748 growth in a wild bird is buffered against phenological mismatch. Science 340:488–91.

749 Reed, T. E., S. Jenouvrier, and M. E. Visser, 2013b. Phenological mismatch strongly  
750 affects individual fitness but not population demography in a woodland passerine. J.  
751 Anim. Ecol. 82:131–44.

752 Rue, H. and L. Held, 2005. Gaussian Markov random fields: theory and applications.  
753 Chapman and Hall/CRC press, London.

754 Rue, H., S. Martino, and N. Chopin, 2009. Approximate Bayesian inference for latent  
755 Gaussian models by using integrated nested Laplace approximations. J. Roy. Stat. Soc.  
756 B 71:319–392.

757 Sæther, B.-E., J. Tufto, S. Engen, K. Jerstad, O. W. Røstad, and J. E. Skåtan, 2000.  
758 Population dynamical consequences of climate change for a small temperate songbird.  
759 Science 287:854–856.

760 Sæther, B.-E., M. E. Visser, V. Grøtan, and S. Engen, 2016. Evidence for  $r$  - and  $K$   
761 -selection in a wild bird population: a reciprocal link between ecology and evolution.  
762 Proceedings of the Royal Society B: Biological Sciences 283:20152411.

763 Schoener, T. W., 2011. The newest synthesis: understanding the interplay of evolutionary  
764 and ecological dynamics. Science 331:426–429.

765 Shaw, R. G., 1986. Response to density in a wild population of the perennial herb *salvia*  
766 *lyrata*: Variation among families. Evolution 40:492–505.



767 Shaw, R. G. and C. J. Geyer, 2010. Inferring fitness landscapes. *Evolution* 64:2510–2520.

768 Shumway, R. H. and D. S. Stoffer, 2011. *Time Series Analysis and Its Applications*.  
769 Springer. URL <http://link.springer.com/10.1007/978-1-4419-7865-3>.

770 Siepielski, A. M., J. D. DiBattista, and S. M. Carlson, 2009. It’s about time: the temporal  
771 dynamics of phenotypic selection in the wild. *Ecology letters* 12:1261–76.

772 Siepielski, A. M., K. M. Gotanda, M. B. Morrissey, S. E. Diamond, J. D. DiBattista, and  
773 S. M. Carlson, 2013. The spatial patterns of directional phenotypic selection. *Ecology*  
774 *Letters* 16:1382–92.

775 Sinervo, B. and E. Svensson, 2002. Correlational selection and the evolution of genomic  
776 architecture. *Heredity* 89:329–338.

777 Slobodkin, L. B., 1961. *Growth and regulation of animal populations*. Holt, Rinehart  
778 and Winston.

779 Smith, C. C. and S. D. Fretwell, 1974. The Optimal Balance between Size and Number  
780 of Offspring. *Am. Nat.* 108:499–506.

781 Sæther, B.-E., V. Grøtan, S. Engen, T. Coulson, P. R. Grant, M. E. Visser, J. E.  
782 Brommer, B. Rosemary Grant, L. Gustafsson, B. J. Hatchwell, K. Jerstad, P. Karell,  
783 H. Pietiäinen, A. Roulin, O. W. Røstad, and H. Weimerskirch, 2016. Demographic  
784 routes to variability and regulation in bird populations. *Nature Communications*  
785 7:12001.

786 Sæther, B.-E., W. J. Sutherland, and S. Engen, 2004. Climate influences on avian popu-  
787 lation dynamics. Pp. 185 – 209, *in Birds and Climate Change, Advances in Ecological*  
788 *Research*, vol. 35. Academic Press. URL [http://www.sciencedirect.com/science/](http://www.sciencedirect.com/science/article/pii/S0065250404350099)  
789 [article/pii/S0065250404350099](http://www.sciencedirect.com/science/article/pii/S0065250404350099).

790 Tarwater, C. E. and S. R. Beissinger, 2013. Opposing selection and environmental vari-  
791 ation modify optimal timing of breeding. *Proceedings of the National Academy of*  
792 *Sciences* 110:15365–15370.

793 Tufto, J., 2000. The evolution of plasticity and nonplastic spatial and temporal adapta-  
794 tions in the presence of imperfect environmental cues. *Am. Nat.* 156:121–130.

- 795 ———, 2015. Genetic evolution, plasticity and bet-hedging as adaptive responses to tem-  
796 porally autocorrelated fluctuating selection: A quantitative genetic model. *Evolution*  
797 69:2034–2049.
- 798 ———, 2017. Domestication and fitness in the wild: A multivariate view. *Evolution*  
799 71:2262–2270.
- 800 Visser, M. E., L. J. M. Holleman, and P. Gienapp, 2006. Shifts in caterpillar biomass phe-  
801 nology due to climate change and its impact on the breeding biology of an insectivorous  
802 bird. *Oecologia* 147:164–72.
- 803 Walsh, B. and M. B. Morrissey, 2018. Individual fitness and the measurement of uni-  
804 variate selection. Pp. 1104–1106, *in* Walsh and Lynch, eds. *Evolution and Selection of*  
805 *Quantitative Traits*. Oxford University Press.
- 806 Wei, W. W. S., 2006. *Time Series Analysis: Univariate and Multivariate Methods*.  
807 Pearson Addison Wesley.
- 808 Williams, G. C., 1957. Pleiotropy, natural selection, and the evolution of senescence.  
809 *Evolution* 11:398–411.

## 810 Appendix A Appendix

### A.1 STATISTICAL MODEL DETAILS

811 To accommodate overdispersion in the number of surviving fledglings  $y_{1i}$  and in the  
812 number of surviving recruits  $y_{2i}$ , we use a beta-binomial distribution for each episode  
813  $s = 1, 2$ , parameterized in terms of the survival probability  $p_{si}$  and a dispersion parameter  
814  $\gamma_s$  (the factor by which the variance is inflated relative to the simpler binomial model).  
815 Conditional on the initial clutch size  $y_{0i}$ , the joint distribution of number of fledglings  $y_{1i}$   
816 and number of recruits  $y_{2i}$  is then

$$\begin{aligned}
 p(y_{1i}, y_{2i}) &= p(y_{1i})p(y_{2i}|y_{1i}) \\
 &= \prod_{s=1}^2 \text{betabin}(y_{si}; y_{s-1,i}, p_{si}, \gamma_s),
 \end{aligned}
 \tag{A.1}$$

817 where  $\text{betabin}(y; n, p, \gamma)$  is the probability mass function of the beta-binomial distribu-  
 818 tion. Similarly, the survival of the adult female  $y_{3i}$  associated with breeding event  $i$  is  
 819 modelled as Bernoulli distributed with parameter  $p_{3i}$ .

820 The other random effects  $v_k$  and  $w_j$  appearing in (5), based on territory and female  
 821 identity, are included to model possible positive correlation between number of fledglings  
 822 and recruits produced by the same territory and female in different years. We param-  
 823 eterized the model such that they are iid standard normal and like  $u_t$ , their effect on  
 824 the different episodes  $s = 1, 2, 3$  are potentially different, depending on whether their  
 825 standard deviations  $\tau_s$  and  $\kappa_s$  differ between episodes.

826 In practice, the expected number of recruits produced as function of clutch size  $y_0$   
 827 and laying date  $z$ ,  $f_a(y_0, z)$  (Fig. A.1C), was computed by evaluating

$$\begin{aligned}
 f_a(y_0, z) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_0 p_1(y_0, z, v, w) p_2(y_0, z, v, w) f(v) f(w) dv dw \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y_0 \exp\left(-\sum_{s=1}^2 (a_s - a_{s-1}) \exp\left(\eta_s(y_0, z, v, w)\right)\right) f(v) f(w) dv dw,
 \end{aligned}
 \tag{A.2}$$

828 using numerical integration (R-package `cubature`), thus integrating out the above non-  
 829 temporal random territory and female identity effects  $v$  and  $w$ . Here  $f$  is the standard  
 830 normal probability density function,  $\eta_s(y_0, z, v, w)$  is the non-linear predictor given by  
 831 equations (3), (4) and (5) (a function of  $y_0, z, v$  and  $w$ ), and  $a_s - a_{s-1}$  the duration of  
 832 episode  $s$ . In the terminology of generalized linear mixed models, this gives us the so called  
 833 marginal as opposed to conditional model (Agresti, 2002, section 12.2.2) with respect to  $v$   
 834 and  $w$ . Fixed effect covariates were set to their mean values and other temporal random  
 835 effects were set to zero (in Fig. 3) or to their estimated values (in Fig. 4) (giving us a  
 836 conditional model with respect to  $\zeta_t$  and  $u_t$ ). Computing the arithmetic mean fitness in  
 837 this way, averaging over these random effects rather than considering the fitness function  
 838 conditional on their values, appears to be the most reasonable approach, at least under the  
 839 assumption of hard selection (non-local density regulation occurring after locally varying  
 840 selection, see Tufto (2015) appendix S3). The joint optimum of  $f_a(y_0, z)$  (Figs. 4 and A.1)  
 841 and the second derivatives of its log was also computed numerically (R function `optim`).

842 To implement the model, software for fitting generalized linear mixed models cannot  
 843 be used, because  $\eta_{si}$  in (3) is non-linear in the parameters and the random effects. In-  
 844 stead, we used Template Model Builder (Kristensen et al., 2016) (R-package `TMB`) which

845 provides a general framework for fitting complex, non-linear, random effects (latent vari-  
846 able) models. Briefly, the user defines the joint likelihood for the data and the random  
847 effects as a C++ template function. Based on this template, TMB generates a function  
848 computing the Laplace approximation of the marginal likelihood. This is, in turn, maxi-  
849 mized numerically to obtain the maximum likelihood estimates of the model parameters.  
850 Biologically meaningful models derived from variations of (4) and (5) were fitted (Table  
851 2). We selected the best model using the Akaike information criteria AIC (Burnham and  
852 Anderson, 2002) among all tested models and recovered the estimates of all parameters  
853 together with their associated standard errors. Explanatory variables  $alt_i$ ,  $temp_i$  and  
854  $dens_i$  were mean centered.

855 An important advantage of the modelling framework provided by TMB is that models  
856 that are non-linear in the parameters and random effects can be fitted with little effort  
857 from the user perspective. In contrast, Chevin et al. (2015) used the INLA R-package  
858 (Rue et al., 2009) to estimate a log-linear model for the Poisson mean with a constant  
859 quadratic term representing stabilizing selection and a random autocorrelated effect on  
860 the slope. This translates to autocorrelated fluctuations in the moving optimal phenotype.  
861 This approach requires modelling variation in the elevation of the fitness function in  
862 different years through a fixed effect on the intercept (the term  $\mu_t$  in their Eq. (3)). If  
863 instead, variation in elevation was modelled through a random effect on the intercept,  
864 this would translate to an undesirable and unrealistic quadratic relationship between the  
865 expected fitness at the optimum and the location of the optimum. In contrast, the TMB  
866 framework we have used here allows more biologically realistic models to be fitted with  
867 the inclusion of random effects acting directly on the fitness maximum (or equivalently,  
868 on the minimum of the hazard function) and on its location. Modelling variation in the  
869 elevation of fitness functions across years in this way makes better use of the data as some  
870 of the information contained in mean survival in a given year to some extent, depending  
871 on the magnitude of the random effect on the fitness maximum, is informative about the  
872 location of the optimum.

873 Another important improvement over the Chevin et al.'s (2015) method is the inclu-  
874 sion of individual level covariates (such as altitude) directly influencing the location of  
875 the phenotypic optimum at the individual level. This makes the overall predictor (3)  
876 non-linear also in the fixed effect parameters (the parameter  $\theta_{alt,s}$  in 4). In the INLA-

877 based approach of Chevin et al. (2015), a term for this could naively be included as an  
878 interaction term between trait value and altitude in the linear predictor. But again, this  
879 would translate to an unrealistic quadratic relationship between the maximum of the  
880 fitness function and the fitness optima experienced by individuals at different altitudes.

## A.2 EFFECT OF SPRING CONDITIONS ON OPTIMAL EGG-LAYING DATE

881 As the species mainly feed underwater, we used the timing of ice break-up during spring  
882 as a measure of spring conditions. Ice cover break-up was defined as when there no longer  
883 was a connective ice layer across the southern end of the lake Lygne located in the middle  
884 of the dipper system, where the outlet is 1<sup>st</sup> of March was set as day 1 and dates were  
885 sequentially numbered. This information was available from 1979 to 2009. From 2010 to  
886 2013, the timing of ice break-up was set to its mean observed between 1979 and 2009,  
887 i.e. 46 days after 1<sup>st</sup> of March.

888 From the best model retained (Table 2), we tested an additional effect of the timing of  
889 ice break-up (denoted *spring*) on optimal egg-laying date, similar for episodes 1 and 2 (i.e.  
890  $\theta_{\text{spring},1} = \theta_{\text{spring},2} \neq 0$ ,  $\Delta\text{AIC}=0.89$ ). We found that  $\hat{\theta}_{\text{spring},1} = \hat{\theta}_{\text{spring},2} = -0.229 \pm 0.276$ ,  
891 in the opposite direction and different from the theoretical value of  $\theta_{\text{spring},s} = 1$  expected  
892 if the optimal laying date occurs at a fixed number of days after ice break up. While  
893 there may still be an effect that is not detected because of low statistical power, the effect  
894 would have to be quite small, based on the upper approximate confidence 95% confidence  
895 limit of 0.31. We also considered additional model alternatives by excluding the effect  
896 of mean winter temperatures (i.e.  $\theta_{\text{spring},1} = \theta_{\text{spring},2} \neq 0$ ,  $\theta_{\text{temp},2} = 0$ ). Once again, this  
897 model did not show any improvement ( $\Delta\text{AIC}=9.16$ ). In addition, we tested an effect of  
898 the timing of ice break-up on optimal egg-laying date, different for episodes 1 and 2 (i.e.  
899  $\theta_{\text{spring},1} \neq 0$ ,  $\theta_{\text{spring},2} \neq 0$ ,  $\Delta\text{AIC}=2.23$ ). We also evaluated the effect of the timing of ice  
900 break-up on optimal egg-laying date on episode 1 only (i.e.  $\theta_{\text{spring},1} \neq 0$ ,  $\Delta\text{AIC}=0.23$ )  
901 and on episode 2 only (i.e.  $\theta_{\text{spring},2} \neq 0$ ,  $\Delta\text{AIC}=1.99$ ). Finally, we tested the effect of the  
902 timing of ice break-up on optimal egg-laying date on episode 1 only while excluding the  
903 effect of mean winter temperatures (i.e.  $\theta_{\text{spring},1} \neq 0$ ,  $\theta_{\text{temp},2} = 0$ ,  $\Delta\text{AIC}=8.19$ ). None of  
904 these models improved the fit.

Table A.1: Model selection - Displayed are all the tested models derived from the best model without temporal covariates (i.e. best model without population density and winter temperatures, see Table A.2 for a list of parameters included in the best model), the difference  $\Delta\text{AIC}$  (respectively  $\Delta p$ ) in AIC (respectively in number of parameters  $p$ ) between each model and the best one and their description.

	$\Delta\text{AIC}$	$\Delta p$	Description
BS	0	0	Best model
BT	6.74	-1	$\phi_{12} = 0$
BU	1.79	1	$\phi_{11} \neq 0$
BV	1.76	1	$\phi_{22} \neq 0$
BW	1.59	1	$\phi_{21} \neq 0$
BX	7.1	0	$\phi_{11} \neq 0, \phi_{12} = 0$
BY	5.35	0	$\phi_{22} \neq 0, \phi_{12} = 0$

Table A.2: Parameters estimated with an alternative model without any temporal covariates (population density and winter temperatures) but with autocorrelated random effects. Displayed are the meaning of the parameters, their notations and their estimates ( $\pm$  standard errors) for the three episodes of selection.

Parameters	Symbol	Selection episode			Unit
		Egg to fledgling $s = 1$	Fledgling to recruit $s = 2$	Adult female $s = 3$	
Meaning					
Intercept, mortality	$\hat{\beta}_{0,s}$	$1.94 \pm 0.40$	$0.68 \pm 0.29$	0	
Effect of clutch size on mortality	$\hat{\beta}_{\text{clutchsize},s}$	$-0.170 \pm 0.080$	$0.082 \pm 0.055$	0	
Effect of altitude on mortality	$\hat{\beta}_{\text{alt},s}$	$0.00084 \pm 0.00053$	$0.00084 \pm 0.00053$	0	$\text{m}^{-1}$
Intercept, optimal egg-laying date	$\hat{\theta}_{0,s}$	$22 \pm 29$	$22 \pm 29$	0	d
Effect of altitude on optimal date	$\hat{\theta}_{\text{alt},s}$	$0.20 \pm 0.12$	$0.20 \pm 0.12$	0	$\text{d m}^{-1}$
Quadratic effect of egg-laying date	$\hat{\omega}_s$	$81 \pm 45$	$81 \pm 45$	0	
SD of random year effect	$\hat{\sigma}_s$	0	$0.032 \pm 0.067$	$0.032 \pm 0.067$	
SD of random territory effect	$\hat{\tau}_s$	$0.181 \pm 0.055$	$0.181 \pm 0.055$	$0.181 \pm 0.055$	
Beta-binomial dispersion parameter	$\hat{\gamma}_s$	$1.915 \pm 0.139$	$1.072 \pm 0.084$		
SD of second order random age effect	$\hat{\nu}_s$	0	0	$0.34 \pm 0.17$	
VAR(1) autoregressive parameters	$\hat{\Phi}$	0	$-0.45 \pm 0.95$		
		0	0		
SD of latent fluctuations in optimum date	$\hat{\sigma}_\zeta$	$19 \pm 12$			d

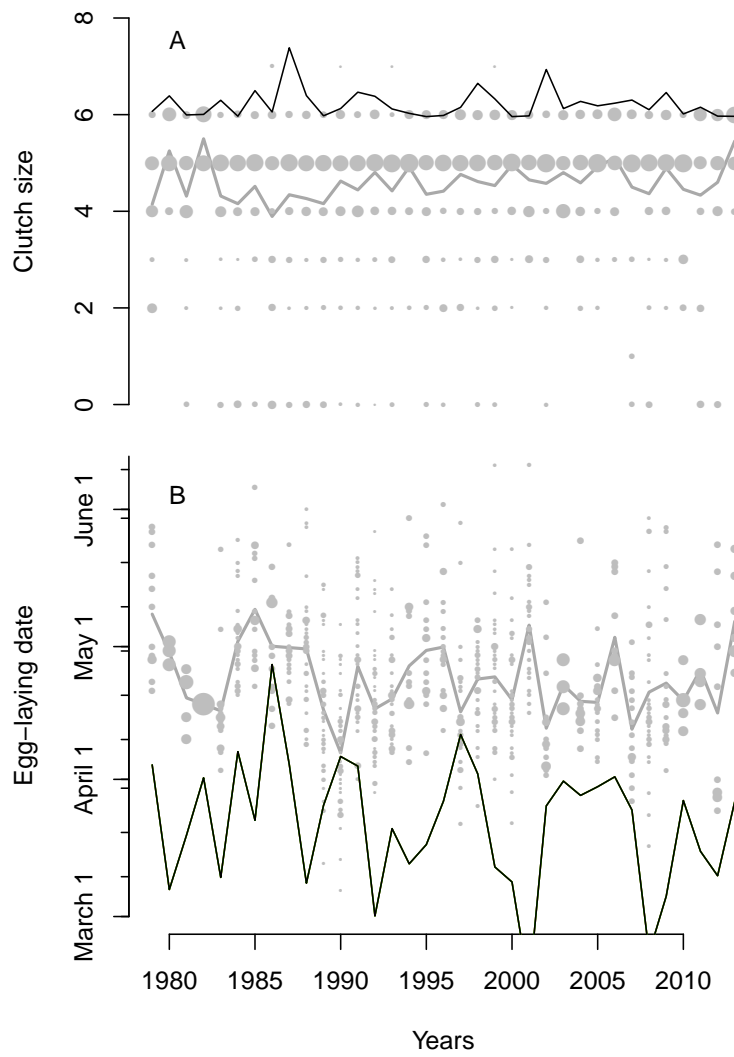


Figure A.1: Estimated annual optimal clutch size (plot A) and laying dates (plot B) for the alternative model without temporal covariates (Table A.2).

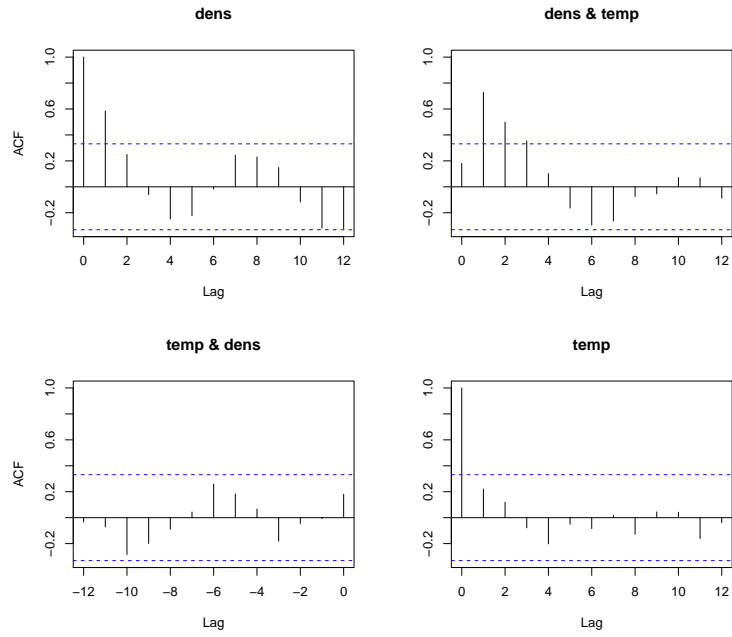


Figure A.2: Sample autocorrelation- and cross-autocorrelation functions (see e.g. Shumway and Stoffer, 2011, Defs. 1.14 and 1.41) for the time series of observed population densities and temperatures. The off-diagonal plots displays estimates of  $\text{corr}(\text{dens}_t, \text{temp}_{t-k})$ .

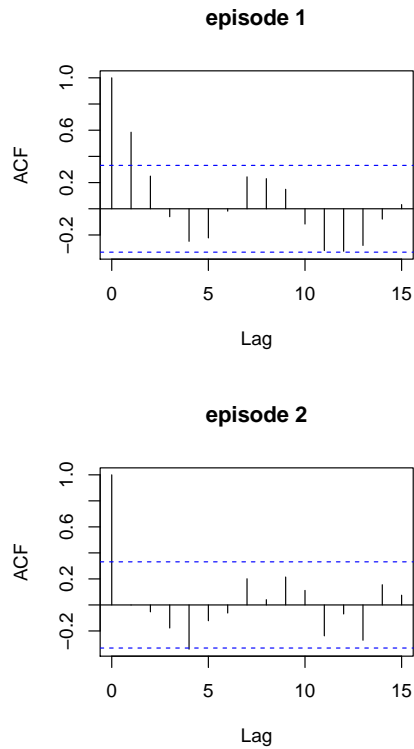


Figure A.3: Estimated autocorrelation functions of optimal laying dates  $\hat{\theta}_{\text{dens},s} \text{dens}_t + \hat{\theta}_{\text{temp},s} \text{temp}_t$  induced by population density and winter temperature for episodes  $s = 1$  and  $s = 2$ .