Phenological mismatch strongly affects individual fitness but not population demography in a woodland passerine

Thomas E. Reed[1], Stephanie Jenouvrier[2] & Marcel E. Visser[1]

[1]Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), P.O. Box 50, 6700AB Wageningen, The Netherlands

[2] Woods Hole Oceanographic Institution, Woods Hole, MA, 02543, United States of America.

Corresponding author email: t.reed@nioo.knaw.nl

Keywords: timing, mismatch, synchrony, climate, demographic, vital rates, fitness, selection gradient.

Abstract

1. Populations are shifting their phenology in response to climate change, but these shifts are often asynchronous among interacting species. Resulting phenological mismatches can drive simultaneous changes in natural selection and population demography, but the links between these interacting processes are poorly understood.

2. Here we analyse 37 years of data from an individual-based study of great tits (*Parus major*) in the Netherlands and use mixed-effects models to separate the within- and across-year effects of phenological mismatch between great tits and caterpillars (a key food source for developing nestlings) on components of fitness at the individual and population levels..

3. Several components of individual fitness were affected by individual mismatch (i.e. late breeding relative to the caterpillar food peak date), including the probability of double-brooding, fledgling success, offspring recruitment probability, and the number of recruits. Together these effects contributed to an overall negative relationship between relative fitness and laying dates, i.e. selection for earlier laying on average.

4. Directional selection for earlier laying was stronger in years where birds bred on average later than the food peak, but was weak or absent in years where the phenology of birds and caterpillars matched (i.e. no population mismatch).

5. The mean number of fledglings per female was lower in years when population mismatch was high, in part because fewer second broods were produced. Population mismatch had a weak effect on the mean number of recruits per female, and no effect on mean adult survival, after controlling for the effects of breeding density and the quality of the autumnal beech (*Fagus sylvatica*) crop.

44   6.   These findings illustrate how climate-change-induced mismatch can have strong

45   effects on the relative fitness of phenotypes within years, but weak effects on mean

46   demographic rates across years. We discuss various general mechanisms that

47   influence the extent of coupling between breeding phenology, selection and

48   population dynamics in open populations subject to strong density regulation and

49   stochasticity.

50

51

52

53

54

55

56

57

58

59

60

61

62

Introduction

Natural selection is an on-going phenomenon in dynamic environments (Endler 1986). Temporal variations in extrinsic factors (e.g. climate, habitat, interspecific competition) and intrinsic factors (e.g. intraspecific competition for food or nest sites) drive phenotypic selection, which typically fluctuates in magnitude, form (Kingsolver *et al.* 2001; Bell 2010) and sometimes sign (Siepielski, DiBattista & Carlson 2009; but see Morrisey & Hadfield 2012). Stochastic environmental variation also directly influences age/stage-specific average reproduction and survival, and hence population demography (Coulson *et al.* 2001; Lande, Engen & Saether 2003; Jenouvrier *et al.* 2012). Until relatively recently, however, factors influencing the nature and strength of connections between natural selection and population dynamics have received little empirical attention (Saccheri & Hanski 2006; Kokko & Lopez-Sepulcre 2007).

Natural selection and population demography are both affected by individual variation in survival and reproductive success (Clutton-Brock 1998; Metcalf & Pavard 2007). The crucial difference is that selection is driven by differences in the *relative* fitness of individuals with different trait values, whereas population demography is shaped by variation in the *absolute* performance of individuals. It follows, therefore, that selection may influence population demography in situations where selection among alternative phenotypes alters mean survival or fecundity at the population level (Saccheri & Hanski 2006; Coulson, Tuljapurkar & Childs 2010). Charlesworth (1971; 1994) showed how population dynamic responses can be critically sensitive to selection on some life-history traits, but not others, depending on where in the life cycle selection occurs relative to population regulation. For example, in species with extended parental care such as altricial birds and mammals, selection arising from variation in breeding success (number of young raised to independence) might

87   not be expected to impact demography much if the survival of offspring post-independence is

88   higher in years where average breeding success is lower, because of reduced intra-cohort

89   competition. Similarly, variation in breeding success might have weak effects on population

90   demography if annual recruitment is driven more by exogenous factors (e.g. climate) during

91   the non-breeding season (Saether, Sutherland & Engen 2004).

92   The need to understand links between individual fitness, natural selection and

93   population demography has become an issue of applied importance in the face of widespread,

94   human-induced alterations to natural environments (Kinnison & Hairston 2007). Climate

95   change, for example, is thought to represent perhaps the biggest threat to global biodiversity

96   (Thomas *et al.* 2004; Malcolm *et al.* 2006), yet we know surprisingly little about how changes

97   in climate translate into changes in local selective pressures and how these, in turn, influence

98   the demographic responses of populations (Reed, Schindler & Waples 2010). One critical

99   pathway via which changes in climate potentially influence fitness is phenology (Jenouvrier

100   & Visser 2011), i.e. the timing of life cycles in relation to key environmental factors. In

101   seasonal environments, life history events such as annual reproduction or migration are

102   typically scheduled to coincide with favourable periods, for example benign weather

103   conditions or seasonal peaks in food abundance. In many regions, these favourable periods are

104   shifting as the climate changes, and species are adjusting their phenology (Parmesan & Yohe

105   2003; Root *et al.* 2003). Rates of phenological change have typically been observed to be

106   unequal across functional groups (Thackeray *et al.* 2010), however, leading to mismatches

107   between interacting species such as predators and prey (Visser & Both 2005). Ostensibly,

108   mismatch should entail negative fitness consequences for the consumer, yet relatively little is

109   known about the evolutionary and demographic implications (Both 2010; Miller-Rushing *et*

110   *al.* 2010; Heard, Riskin & Flight 2011).

111   In birds, mismatches have been shown or hypothesised to occur in a range of species

112   for which synchronisation of breeding with narrow seasonal food peaks is important

113   (reviewed by Both 2010, Visser, te Marvelde & Lof 2011). Under the so-called 'mismatch

114   hypothesis (Drever & Clark 2007; Dunn *et al.* 2011), fitness is lower for females breeding

115   both earlier and later than the seasonal food peak, although fitness need not peak exactly

116   when breeding coincides with the food peak given that other selective pressures can be

117   involved (Visser, te Marvelde & Lof 2011; Lof *et al.* 2012). Climate change has led to an

118   increase in positive mismatch years (late breeding relative to seasonal food peaks) for

119   woodland birds in temperate regions, as spring/summer warming has tended to advance food

120   peaks faster than avian phenology (Visser, Both & Lambrechts 2004; Jones & Cresswell

121   2010). While increasing mismatch has been linked to population declines in some species

122   (e.g. long distance migrants, Both *et al.* 2006; Both *et al.* 2010), evidence for negative fitness

123   effects has been mixed in others (Eeva, Veistola & Lehikoinen 2000; Drever & Clark 2007;

124   Shultz *et al.* 2009; Dunn *et al.* 2011; Vatka, Orell & Rytkönen 2011).

125   Here we explore relationships between phenological mismatch and components of

126   fitness at the individual and population levels in great tits (*Parus major* L.), to better

127   understand the various mechanisms by which climate effects on phenology simultaneously

128   influence natural selection and population demography. Across Europe, populations of great

129   tits have exhibited variable phenological responses to large-scale changes in spring

130   temperature since 1980 (Visser *et al.* 2003). Great tits rely heavily on caterpillars during the

131   breeding season to feed their chicks (van Balen 1973; Naef-Daenzer, Naef-Daenzer & Nager

132   2000; Mols, van Noordwijk & Visser 2005; Wilkin, King & Sheldon 2009), and in some

133   habitats (e.g. oak forests) caterpillar biomass typically shows a pronounced, narrow seasonal

134   peak in late spring/early summer (Visser, Holleman & Gienapp 2006). Caterpillar

135   development is strongly affected by temperature, and great tits at mid-latitudes use predictive

136 cues such as early spring temperatures (Visser, Holleman & Caro 2009; Schaper *et al.* 2012)

137 to adjust their egg-laying dates in line with fluctuations in the seasonal peak in caterpillar

138 biomass. In our Hoge Veluwe study population in the Netherlands, advancements in laying

139 dates in response to warmer springs have been insufficient to keep pace with stronger

140 advancements in caterpillar phenology, and the population now breeds much later relative to

141 the seasonal caterpillar peak (Visser 2008). While previous studies on this population have

142 examined selection on laying dates (Visser *et al.* 1998; Visser et al. 2006; Gienapp, Postma &

143 Visser 2006), the effects of mismatch on population demography have not been explored in

144 detail, which requires separating within-year effects on individual fitness from between-year

145 effects on average fitness.

146   The aims of this paper were therefore threefold: (1) To explore the impact of

147 phenological mismatch on components of individual fitness, (2) to explore the effects of mean

148 mismatch on population mean vital rates, and (3) to link the individual and population impacts

149 by estimating annual selection differentials and testing for an association with population

150 mean mismatch.

151

152

153

154

155

156

157

Materials and methods

Study area and field methods

The data analysed come from a long-term, individual-based demographic study of great tits (*Parus major*) at the Hoge Veluwe National Park in the Netherlands (52° 02' 07" N 5° 51' 32" E). The study area consists of mixed pine-deciduous woodland on poor sandy soils. A large block of pure pine plantation was included from 1955 to 1972, but this was damaged by a severe storm in the winter of 1972/1973. Here we focus on the years 1973-2011, when the study area included only mixed coniferous-deciduous woodland. The study area remained the same size across this period and the number of nest boxes was approximately constant, although some were replaced or moved as the study progressed. A surplus of nest boxes was provided to ensure that availability of artificial nest sites did not limit population size (the ratio of nest boxes to breeding females was approximately 3:1, on average). The study area is surrounded by a matrix of potentially suitable breeding habitat for great tits, thus the population is open to immigration and emigration.

During the breeding season (April to June/July), nest boxes were visited at least once per week. The number of eggs or nestlings present was counted at each visit. When the nestlings were 7-10 days old, the parents were caught on the nest using a spring trap. Parents already ringed were identified and unringed birds were given a metal ring with a unique number. Young were ringed on day 7. Female great tits are capable of producing a second brood each season (i.e. laying a second clutch and raising a new brood after successful fledging of the first brood), although the frequency of double-brooding in this population has declined in recent decades (Husby, Kruuk & Visser 2009). A small but variable proportion of breeding females each year were not caught, primarily those that desert their clutches early in the breeding attempt. Unknown females were not included in the survival analyses, as their

182 survival to future breeding seasons could not be determined. Recapture probability was very

183 high in females (average = 98.7%) and males (average = 95.5%). Female recapture

184 probability did not exhibit any trends over time ($P = 0.460$, Fig. S2a) or any association with

185 population mean mismatch ($P = 0.425$, Fig. S2b). Male recapture probability also did not

186 exhibit any trends over time ($P = 0.839$, Fig. S2a) or association with population mean

187 mismatch ($P = 0.588$, Fig. S2b). Therefore, we did not include recapture probability in our

188 survival analyses.

189 In some years, brood size manipulation experiments were carried out that affected

190 fledgling production or recruitment probability. Manipulated broods were excluded from all

191 analyses. Data from the 1991 breeding season were also excluded, as this was an anomalous

192 year where a late frost resulted in a very late caterpillar food peak (Visser *et al.* 1998). The

193 analysed dataset consisted of 3472 records of 2599 females breeding in 37 years. 560 of these

194 records were of unknown females. The average number of breeding records per known female

195 was 1.43.

196 Dates of the peak in caterpillar biomass were estimated for 1985 – 2010 from frass fall

197 samples in the Hoge Veluwe. The most predominant species in our system are the winter

198 moth (*Operophtera brumata*) and the oak leaf roller (*Tortrix virirdana*), although caterpillars

199 of several other species are also present. The annual caterpillar peak is well predicted by mean

200 temperatures from 8th March – 17th May ($r^2 = 0.80$), and this relationship was used to predict

201 caterpillar peaks from 1973 to 1984. For full details see Visser et al. (1998) and Visser *et al.*

202 (2006). The basic patterns presented in the results were similar when the analyses were

203 restricted to the years where food peaks were measured directly, so we include all years in the

204 final analysis.

205

Statistical analyses

Effects of mismatch on individual and population-level fitness components

Food demands of great tit nestlings are highest approximately 9-10 days after hatching

(Royama 1966; Gebhardt-Henrich 1990; Keller & van Noordwijk 1994; Mols, van Noordwijk

& Visser 2005) and females strive to match nestling energy requirements to the period when

caterpillars are plentiful. The mismatch between a female's breeding time and the timing of

the food peak was defined as the difference between the laying date of her first clutch and the

food peak date, plus 30 days (i.e. individual mismatch = laying date + 30 – food peak date).

Laying dates are given as April-days (1 April is April-day 1, 24 May is April-day 54). This

mismatch metric essentially measures laying dates relative to the food peak, but the constant

value of 30 days was added to in order to make the values more easily-interpretable. Great tits

in our study population typically lay 9 eggs and incubate them for 12 days, hence nestling

food requirements peak approximately 30 days (9 + 12 + 9) after laying of the first egg. Thus,

according to this metric, a female laying too early relative to the food peak would have a

negative value for individual mismatch (IM), a female laying too late would have a positive

IM value, while a female who lays on the date such that her chicks are 9 days old at the food

peak would have an IM value of 0 (see Fig.1). We stress that this is purely an operational

definition of mismatch; we do not assume that fitness is highest for females with an IM of 0.

Annual population mismatch (PM) was defined simply as the arithmetic average of IM

values each year (Fig. 1). This difference between the mean phenology of birds and the food

peak is only a proxy for true population-level mismatch, of course, but it does provide a

straightforward, easily calculable metric comparable to previous studies on this (Nussey *et al.*

2005; Visser et al. 2006) and other species (Visser & Both 2005). See the aentary Material for

more discussion of the pros and cons of our mismatch measure and potential alternatives.

230   Generalized linear mixed effects models (GLMMs) were used to examine variation in

231   fitness components in relation to individual- and population-level mismatch simultaneously.

232   We separated IM from PM effects by standardizing IM within years (by subtracting year-

233   specific PM values from IM values) and including both standardized IM and PM as fixed

234   effects in the GLMMs. Thus, the fixed effect of PM measures the across-year effect of

235   average mismatch, while the fixed effect of standardized IM effectively quantifies the within-

236   year effect of individual breeding time *relative to the mean breeding time that year*. This is

237   directly analogous to 'within-subject centering', a technique used in mixed-effects models to

238   distinguish within-individual from between-individual effects (van de Pol & Wright 2009).

239   Individual- and population-level effects of mismatch are illustrated graphically in separate

240   figures (Figs.2 and 3), but the predicted effects themselves are estimated in the same GLMMs

241   (see Table 1).

242   For each breeding record included in the GLMM analyses, mismatch was defined on

243   the basis of first clutches (n=3472 breeding records where the laying date of the first clutch

244   was known), but fledglings and recruits produced from second clutches were included in the

245   fitness calculations. The following fitness components were examined: (a) the probability of

246   double-brooding, (b) clutch size of the first clutch (c) probability of producing zero fledglings

247   that season (including those from second broods), (d) number of fledglings produced, given

248   that one or more chicks were raised, (e) probability of recruitment (the total number of

249   offspring per female surviving to breed themselves in subsequent years, divided by the total

250   number of fledglings she produced that year), (f) total number of recruits, (g) female local

251   survival (the probability that a female parent survives between year t and t+1, i.e. was

252   observed as a breeder the following year), and (h) male local survival. Fitness components

253   measured as probabilities (probability of double-brooding, probability of producing zero

254   fledglings, offspring recruitment, adult survival) were analyzed using GLMMs with logit-link

255   functions and binomial errors. Fitness components measured as counts (clutch size, number of

256   fledglings, number of recruits) were analyzed using GLMMs with Poisson errors and log-link

257   functions. The distribution of total number of fledglings per female is strongly zero-inflated,

258   as many females fail to raise any chicks each year. Hence, the probability of producing zero

259   fledglings was analyzed separately to the number of fledglings produced given than one or

260   more chicks were fledged. In the case of recruitment and adult survival, death cannot be

261   distinguished from permanent emigration from the study area; thus we effectively model

262   apparent local recruitment and survival.

263   For each fitness component, the full models contained the following fixed effects:

264   intercept, standardized individual mismatch (hereafter IM′, with the prime symbol indicating

265   the standardization relative to PM), a quadratic effect of IM′, PM, mother age class as a 2-

266   level factor (first time-breeder or experienced breeder), breeding density (annual number of

267   first clutches), and the interactions mother age $\times$ (IM′ + IM′^{2}), PM $\times$ (IM′ + IM′^{2}), and

268   breeding density $\times$ (IM′ + IM′^{2}). Quadratic effects of IM′ were included as we suspected that

269   both breeding too early or too late relative to the food peak might negatively impact fitness.

270   The interaction PM $\times$ (IM′ + IM′^{2}) tested whether the potentially nonlinear effects of IM′

271   varied as a function of PM (e.g., fitness differences between early and late laying females

272   might be larger in years where the population breeds too late on average). Mother age and the

273   interactions with IM′ and IM′^{2} were included to examine potential differences in the

274   relationships between fitness components and IM′ for inexperienced versus experienced

275   breeders. Demographic studies of great tits typically find that first-years females lay later,

276   produce smaller clutches, and recruit fewer offspring than older age classes (Perrins & Moss

277   1974; Harvey *et al.* 1979; Jarvinen 1991). Note that age information was not available for the

278   560 records of unknown females. Breeding density was included as a continuous covariate as

279   previous studies have documented strong density dependence at various stages in the great tit

280   life history (e.g. Dhondt, Kempenaers & Adriaensen 1992; Both, Visser & Verboven 1999)

281   and on overall numbers (Saether *et al.* 1998; Grøtan *et al.* 2009). The interaction breeding

282   density $\times$ (IM$'$ + IM$'^2$) was included to test whether the (potentially nonlinear) effects of

283   mismatch depended on breeding density. In GLMMs (e) to (h) we also included the

284   explanatory variable beech crop index (BCI) as a factor with three levels, 3 being the highest.

285   BCI quantifies the amount of beech nuts available in winter on a 3 point scale, and also

286   correlates with the crop size of other tree species in the region (see Perdeck, Visser & Van

287   Balen 2000 for further details). Beech nuts are an important winter food source affecting the

288   overwinter survival of juveniles and adults alike (Perrins 1965; Clobert *et al.* 1988; Grøtan *et*

289   *al.* 2009). The interaction BCI $\times$ (IM$'$ + IM$'^2$) was included in these models to test whether the

290   effects of individual mismatch depended on the quality of the beech crop that year.

291   Random effects of female identity and year were included in all GLMMs. Models

292   were fitted in R using the function *glmer* in the package *lme4*. We used a backwards stepwise

293   model simplification procedure, sequentially removing non-significant fixed-effect terms

294   ($P > 0.05$, where P values correspond to the z-values reported by *glmer*) starting with higher-

295   order terms (first interactions involving quadratic terms, then linear terms), to yield minimum

296   adequate models. We stress that the goal of these GLMMs was not to explain as much

297   variation in each fitness component as possible using all possible candidate explanatory

298   variables, but rather to characterize the relationships with phenological mismatch while

299   correcting for key covariates known *a priori* to be important. Testing for significant

300   interactions between individual mismatch and year-specific covariates (PM, density, BCI)

301   also provides insights into the mechanisms underlying population-level relationships (or lack

302   thereof) between mismatch and demographic rates. Overall raw relationships between

303   demographic rates and year (i.e. not correcting for environmental variables) are presented in

304   Fig. S1.

305

Selection analyses

307 Selection differentials, defined as the covariance between phenotype and relative fitness,

308 quantify the strength of directional selection on a trait (Lande & Arnold 1983). We used the

309 number of locally-recruiting offspring per female as a measure of individual (annual) fitness.

310 Fitness was converted to relative fitness by dividing by the mean number of recruits each

311 year. Laying date, the phenological trait assumed to be under selection, was standardized

312 within years to a mean of zero and a standard deviation (SD) of one by subtracting the annual

313 mean and dividing by the annual SD. Each year t, a standardized estimate of annual

314 directional selection (standardized linear selection differential, β_t) can then be obtained as the

315 slope of the regression of relative fitness on standardized laying dates. To explore which

316 environmental factors best explained variation in annual directional selection, we regressed

317 the β_t estimates against PM, PM^2, breeding density, BCI, and age composition (the ratio of

318 first-time breeding females to experienced breeders). Data points in this multiple regression

319 were weighted by $1/[(\text{standard error of } \beta_t)^2]$, to account for the fact that β_t estimates in some

320 years were based on a small number of recruits (e.g. four recruits from the 1984 breeding

321 season) and therefore much less certain than years with more recruits (e.g. 105 in 1976). We

322 predicted that reproductive output might be lower, on average, in years where selection was

323 stronger. To test this, we regressed the annual mean number of recruits per female against β_t

324 values and their square.

325 We also estimated standardized nonlinear selection differentials, given as twice the

326 quadratic coefficient in a regression of relative fitness on standardized laying date +

327 standardized laying date^2. Note that quadratic regression coefficients and their standard errors

328 must be doubled to obtain point estimates of annual nonlinear selection differentials (hereafter

14

329   γ_t) and their uncertainty (Stinchcombe *et al.* 2008). We also tested for relationships between γ_t

330   and PM, PM^2, breeding density, BCI, and age composition, weighting the annual data points

331   by $1/[($standard error of $\gamma_t)^2]$.

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

Results

Effects of mismatch on individual and population-level fitness components

348
349
350 Within years, the probability that an individual female attempted a second brood was

351 nonlinearly related to IM, with relatively early females (negative IM′ values) being more

352 likely to attempt a second brood (Fig.2A, linear effect: P=0.728; quadratic effect: P=0.003;

353 estimates ± SE and sample sizes are provided in Table 1). First-time breeders were less likely

354 to attempt a second brood compared with experienced breeders (P=0.005). Across years, the

355 mean probability of double-brooding brooding was negatively related to average mismatch

356 (P<0.001, Fig. 3A) and to breeding density ($P < 0.001$, Table 1A). The negative relationship

357 between probability of double-brooding and IM′ was also stronger in years where PM was

358 larger (IM′ × PM interaction term: P<0.001) and when breeding density was higher (IM′ ×

359 density interaction: P=0.014, Table 1A).

360 Females breeding late relative to the food peak laid significantly fewer eggs (i.e. a

361 negative effect of IM′: P<0.001, Fig. 2B, Table 1B). There was no across-year relationship

362 between mean clutch size and PM (Fig. 3B), but annual mean clutch size was negatively

363 related to breeding density (P<0.001, Table 1B). Females that bred late relative to the food

364 peak were more likely to fail to raise any fledglings (Fig. 2C; linear effect of IM′: P<0.001;

365 quadratic effect of IM′: P<0.001; Table 1C). While there was no overall effect of PM on mean

366 probability of producing zero fledglings (Fig. 3C), the effect of IM′ was stronger in years

367 where PM was larger (Table 1C; IM′ × PM interaction: P<0.001). Among those females that

368 did fledge chicks, there was a negative quadratic relationship between the number fledged and

369 IM′ (linear effect: P<0.001; quadratic effect: P=0.003; Table 1D, Fig. 1D). First-time

370 breeders fledged fewer chicks than experienced breeders (P<0.001; Table 1D). Across years,

371   the mean number of fledglings per female was negatively related to PM (P=0.019, Fig. 3D)

372   and breeding density ($P < 0.001$, Table 1D).

373   Within years, recruitment probability was negatively related to IM (linear effect of

374   IM′: P<0.001; Fig. 2E), with the relationship being stronger in years where average mismatch

375   was larger (Table 1E; IM′ × PM interaction: P<0.001). Across years, there was no

376   relationship between average recruitment probability and PM (P=0.151; Fig. 3E), a negative

377   relationship with breeding density (P<0.001), and a positive relationship with BCI (Table 1E).

378   A higher proportion of fledglings recruited in years where BCI was medium or high (two or

379   three, on the 3-point scale) compared to years where BCI was low (one on the 3-point scale).

380   The total number of recruits per female was negatively related to IM′ within years (Fig. 2F;

381   linear effect of IM′: P<0.001; negative quadratic effect of IM′: P=0.044; Table 1F). Across

382   years, there was a weak negative relationship between the mean number of recruits per female

383   and PM (P=0.038, Fig. 3F), a negative relationship with breeding density (P<0.001), and a

384   positive relationship with BCI (Table 1F). First-time breeders produced fewer recruits than

385   experienced breeders (P=0.032; Table 1F). The negative relationship between the number of

386   recruits per female and IM was stronger in years where PM was larger (Table 1F; IM′ × PM

387   interaction: P=0.001).

388   Female adult survival was not related to mismatch within years, although there was a

389   non-significant negative trend (P=0.068, Fig. 2G). There was no relationship between mean

390   female survival and PM across years (Fig. 3G), while there was a negative effect of breeding

391   density (P=0.003) and a positive effect of BCI (Table 1G). Similarly, there was no

392   relationship between male adult survival and IM′ within years (Fig. 2F) or PM across years

393   (Fig. 3F). Mean adult survival for males was negatively related to breeding density (P=0.003)

394   and positively related to BCI (Table 1F).

395

396   *Selection analyses*

397   When data from all years were pooled, there was an overall negative relationship between

398   relative fitness (the number of recruits relative to the annual mean) and standardized laying

399   date, i.e. directional selection for earlier egg-laying (overall standardized selection differential

400   = -0.198 ± 0.035 [standard error], $t = $ -5.658, $P < 0.001$, df = 3470). The annual point

401   estimates for the strength of directional selection (i.e. β_t values) varied considerably from year

402   to year, but were negative in most years (Fig. S3a). There was a negative quadratic

403   relationship between β_t and the annual population mismatch (Fig. 4; $\beta_t = $ -0.133 -0.007 × PM -

404   0.002 × PM^2; linear term: $P = 0.277$; quadratic term: $P = 0.020$; overall model: $F_{(2,34)} = $

405   6.273, $P = 0.005$). Directional selection was stronger in years where birds bred on average

406   later than the food peak, but was weak or absent in years where the synchrony between birds

407   and caterpillars was high or negative (Fig. 4). Density, BCI, and age composition did not have

408   significant effects on β_t. There was no relationship between the annual mean number of

409   recruits and β_t (linear effect: $P = 0.445$; quadratic effect: $P = 0.358$).

410   Nonlinear selection was apparent in many years (Fig. S3c), but the form of this

411   selection varied from concave (negative quadratic selection, reduced fitness for early as well

412   as late breeders) to convex (positive quadratic selection, all but the very earliest birds fare

413   poorly). There was no significant relationship between the strength of quadratic selection and

414   PM, although there was a non-significant positive trend ($P = 0.107$), i.e. the relationship

415   between relative fitness and laying date appeared to be more convex in years where most of

416   the population bred too late relative to the food peak (Fig. S3d).

417

418

419

Discussion

420

421 In this study we explored relationships between climate, demography and natural selection in

422 a great tit population that has experienced significant spring warming in recent decades. This

423 warming has led to an increasing mismatch between the phenology of the birds and the

424 seasonal peak in caterpillar abundance, the primary food source for nestlings. In the 1970s,

425 typical breeding times closely matched the caterpillar biomass peak, but since then a

426 mismatch of almost two weeks has developed (Fig. S3b) – many pairs now breed too late to

427 profit fully from the short period in summer when caterpillars are plentiful (Visser *et al.* 1998;

428 Nussey *et al.* 2005; Visser et al. 2006). This trophic asynchrony has imposed directional

429 selection for earlier breeding (Fig. 4), and while laying dates have responded through

430 phenotypic plasticity and possibly some microevolution (Gienapp, Postma & Visser 2006),

431 the rate of advance has been much slower than that of caterpillar phenology. Similar

432 mismatches are likely developing in many populations of temperate woodland bird species

433 that are experiencing rapid spring warming (Leech & Crick 2007), yet very little is known

434 about the demographic and evolutionary consequences (Both 2010; Heard, Riskin & Flight

435 2011).

436 Our primary goal in this study was to characterize relationships at both the individual

437 and population levels between fitness components and mismatch. In doing so, we provide a

438 comprehensive analysis of the various ways in which mismatch can affect individual

439 performance and how these translate into signatures (or lack thereof) of climate change at the

440 level of population demography. The results illustrate how phenological mismatch can be

441 associated with strong phenotypic selection while having relatively weak or no apparent

442 effects on key population vital rates (recruitment, adult survival) across years. This highlights

443 the importance of distinguishing conceptually between the effects of mismatch on individual

444 (relative) performance and those on mean productivity or other population-level parameters,

445 and we show how this can be achieved statistically using generalized linear mixed models.

446 Our results also suggest that caution is advisable when extrapolating individual-level

447 relationships to the population level and vice versa, a general problem of statistical and logical

448 inference in hierarchical systems known as 'ecological fallacy' (Robinson 1950; van de Pol &

449 Wright 2009).

450

451 STRONG INDIVIDUAL-LEVEL BUT WEAK POPULATION EFFECTS OF MISMATCH

452 At the individual level, strong negative effects of mismatch, sometimes curvilinear, were

453 detected for all fitness components examined except adult survival. In any given year, females

454 breeding late relative to the seasonal peak in caterpillar biomass (i.e. females with positive

455 values of individual mismatch) were less likely to produce a second brood, laid smaller

456 clutches, and were more likely to fledge no offspring (Fig. 2A-C). Among those females that

457 did manage to raise some chicks to fledging, those breeding late relative to the food peak

458 fledged fewer chicks (Fig. 2D), and these chicks in turn were less likely to recruit (Fig. 2E).

459 The net result was that females laying relatively early produced more recruits (Fig. 2F) and

460 hence their relative fitness was on average higher than that of late-laying females.

461 Despite these pronounced individual-level effects, across-year relationships between

462 mean demographic rates (i.e. annual averages for each fitness component) and population-

463 level mismatch were either much weaker or entirely absent (Fig. 3). For example, annual

464 variation in the mean number of recruits per female – the demographic rate that most strongly

465 influences population fluctuations in this species (van Balen 1980) – was large and driven

466 mostly by density effects and stochastic fluctuations in beech crop (Table 1F). Hence, the

467 mismatch signal was not obvious at the population level for this demographic rate (Fig. 3F)

468   and only statistically significant once breeding density, beech crop and additional stochastic

469   variation due to unknown environmental factors (captured by the 'year' random effect) were

470   formally accounted for in the GLMM. A similarly weak negative relationship between the

471   annual mean number of recruits and phenological asynchrony with caterpillars was found for

472   a UK population of great tits (Charmantier *et al.* 2008).

473   Similar patterns were found for the number of fledglings: a strong negative curvilinear

474   relationship with mismatch at the individual-level (Fig. 2D), but a much weaker negative

475   linear relationship at the population level, with lots of scatter (Fig. 3D). Some of this

476   interannual variation in fledgling production was accounted for by negative density

477   dependence and fluctuations in age composition (Table 1D). The remaining unexplained

478   variation could be due to many factors, for example direct climatic influences on chick

479   mortality; our goal was not to explain as much variation in demographic rates as possible, but

480   rather to understand the mechanisms and extent to which mismatch affects demographic

481   performance. This level of understanding facilitates the development and parameterization of

482   ecologically-realistic population models, which can then be used to predict possible effects of

483   climate change on population dynamics.

484   Several processes could explain why effects of breeding season mismatch on mean

485   demographic rates were weak, despite strong within-year, among-individual effects. First,

486   reductions in the reproductive output of individuals breeding late relative to the food peak

487   might be offset by increases in early birds, for example if young fledged early in the season

488   experience less-intense competition for food in years of high population mismatch because of

489   the higher mortality of late broods. While we do not have direct evidence for this, we did find

490   a significant interaction between PM and IM in the model of recruitment probability (Table

491   1E): the negative effect of IM was stronger in years of large PM, which is consistent with a

492   scenario of frequency-dependent benefits of early fledging. Inspection of the annual

493   relationships between relative reproductive success and standardised laying dates also

494   revealed that the relative success of the earliest females has increased more over the study

495   period than that of the latest females has decreased, which again suggests a role for frequency

496   or density dependence. However, there were no significant interactions between IM and

497   density in the GLMMs for the number of fledglings (Table 1D), probability of recruitment

498   (Table 1E), or number of recruits (Table 1F), nor was there was any overall relationship

499   between annual linear selection differentials and mean breeding density (e.g. stronger

500   selection for earlier breeding in high density years). The annual number of first clutches in the

501   whole study area might be too coarse a measure of density to capture the relevant competition

502   effects, although relative fledging mass might be more important than relative fledging date

503   *per se* in this regard (Both *et al.* 1999).

504   Second, negative fitness effects of mismatch during the breeding season might be

505   counterbalanced by improved survival at other times of the year, for example if winters

506   become less severe because of global warming (Saether *et al.* 2000; Jenouvrier *et al.* 2006).

507   We find no evidence in our study population for increases over time in juvenile or adult

508   survival (Fig. S1); if anything, there was a marginally non-significant negative trend ($P =$

509   0.081) in adult female survival across the study period (Fig. S1G), which might be related to

510   increased competition associated with a higher influx of immigrants (Reed & Visser,

511   unpublished). Reductions in the total number of fledglings produced in years of large

512   population mismatch could also be followed by improved average post-fledgling survival, via

513   density-dependent feedbacks, dampening the effects of mismatch on mean recruitment

514   success. If this were true, however, we would also expect to find a significant statistical

515   interaction between breeding density and individual-level mismatch on recruitment

516   probability, but this was not observed (Table 1E).

517   The third, and in our opinion most likely, explanation for the weaker-than-expected

518   effects of population mismatch on the mean number of fledglings and recruits, is that

519   mismatch signals are simply difficult to detect at the population level because of high

520   environmental stochasticity in these demographic rates. Year-to-year fluctuations in the

521   survival of juvenile and adult great tits are strongly affected by the quality of the autumnal

522   beech crop (Perdeck *et al.* 2000, Grøtan *et al*. 2009) and by winter severity (Kluijver 1951;

523   van Balen 1980), which adds considerable 'environmental noise' to any underlying influence

524   of mismatch. Detecting mismatch effects on demographic rates thus becomes an issue of

525   statistical power, which can easily be confirmed by simulations based on the observed

526   individual-level relationships and between-year stochastic variance in fitness components

527   (results not shown). This conclusion is itself biologically interesting: we have almost four

528   decades of data on great tit demography, a period across which substantial spring warming

529   occurred, yet we find very weak effects of mismatch on mean recruitment rates and no effects

530   on adult survival. This suggests that very long time series, very strong climatic change, or

531   both will be required to observe significant effects of phenological mismatch on population

532   demography, although this of course will depend on the life history and ecology of the species

533   being considered.

534

535   EFFECTS OF MISMATCH ON NATURAL SELECTION

536   Estimating selection differentials provides further insight into links between individual-level

537   and population-level processes. The individual-level analyses (Fig. 2) showed that timing of

538   breeding relative to the seasonal peak in caterpillar biomass has a strong effect on individual

539   relative fitness in our study population. If synchrony with the food peak was the only selective

540   pressure and mean synchrony had not changed over time, then one would expect the fitness

541 curves to more bell-shaped, with lower fitness for both relatively early and relatively late

542 females (i.e. stabilising selection). Indeed, fledging success and fledging mass in great tits

543 tend to be lower both before and after the food peak, at least for first broods (Verboven,

544 Tinbergen & Verhulst 2001; Visser, Holleman & Gienapp 2006). When negative and positive

545 mismatch years are considered separately, the relationship between the number of recruits and

546 IM is more obviously bell-shaped (Fig. S4). However, synchrony with the food peak is not the

547 only selective factor (see below), and average mismatch has increased significantly over time

548 in our study population (Fig. S3). Considering all years together, the overall net effect is

549 directional selection for earlier laying dates.

550   The current study is purely correlational and therefore we cannot exclude the

551 possibility that factors other than timing relative to the food peak (e.g. phenotypic quality

552 effects, seasonal changes in other factors) are responsible for the observed relationships. The

553 relationship with clutch size (Fig. 2B), for example, is probably driven by the fact that early

554 layers *per se* tend to produce larger clutches (Perrins 1970), rather than any causal effect of

555 caterpillar availability given that eggs are laid well before the food peak. Alternatively,

556 females might actively adjust their clutch size (and hence their reproductive effort) in

557 response to environmental cues that predict subsequent caterpillar biomass (Verboven,

558 Tinbergen & Verhulst 2001). The causal effects of caterpillar availability are better

559 established for the relationships between fledgling success and mismatch (Verboven *et al.*

560 2001) and local recruitment and fledging date (Verboven & Visser 1998). Note that we do not

561 account for individual variation in clutch size when calculating IM, which could introduce a

562 potential bias into our estimation of the relationships between IM and fledging/recruitment

563 success, given that late breeders tend to lay smaller clutches. However, the patterns remain

564 largely unchanged when clutch size variation was taken into account (Fig. S5). Thus we chose

565 to account only for laying date variation when calculating IM, given that the primary timing

566 decision for a female is when to initiate egg-laying, not how many eggs to lay (the latter being

567 more related to parental investment decisions).

568   We found that directional selection was stronger in years where birds bred on average

569 later than the food peak, but was weak or absent in years where there was little population

570 mismatch (Fig.4, see also van Noordwijk, McCleery & Perrins 1995; Charmantier *et al.*

571 2008). However, we stress that mismatch is not the only selective pressure affecting laying

572 dates and hence perfect synchrony with the food peak is not necessarily optimal. For example,

573 the interests of chicks and parents need not coincide exactly and females might be

574 constrained, or unwilling, to breed at the optimal date in terms of chick survival prospects

575 because of high costs of producing and incubating eggs early in the season when it is still cold

576 and food is scarce (Perrins 1970; Visser & Lessells 2001). Being 'adaptively mismatched' by

577 a few days might therefore be optimal from the perspective of parental fitness (Visser, te

578 Marvelde & Lof 2011), particularly if day-to-day variation in temperature is high (Lof *et al.*

579 2012). Optimal laying dates may also depend on trade-offs between the fitness benefits of

580 synchronising the first brood with the food peak on the one hand, and reduced probability of

581 producing a second brood (Fig. 2A), on the other (Verboven, Tinbergen & Verhulst 2001). In

582 addition to these selective processes, females laying too early relative to the food peak may

583 have higher-than-expected fitness simply because they are in better body condition, and thus

584 measured fitness curves need not be bell-shaped.

585   In conclusion, we show that in years of large population mismatch, in which a high

586 proportion of females breed too late relative to the food peak, *relative* fitness differences

587 among females breeding at different dates are large, but the average *absolute* fitness is similar

588 to years where population mismatch is smaller or absent. Thus, phenological mismatch

589 appears to have strong effects on selection pressures, but weak effects on key demographic

590 rates. This result suggests that climatic influences on evolutionary and population dynamics

591 might be uncoupled in this population, at least for the trait we considered and within the

592 observed range of spring warming. However, it would be premature to conclude that future

593 climate change does not pose a threat to this population, as reductions in vital rates could

594 unfold rapidly if mismatch increases beyond a certain point.

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

References:

614

615 Bell, G. (2010) Fluctuating selection: the perpetual renewal of adaptation in variable environments.
616   *Philosophical Transactions of the Royal Society B-Biological Sciences,* **365,** 87-97.
617 Both, C. (2010) Food availability, mistiming, and climatic change. *Effects of climate change on birds*
618   (eds A.P. Moller, W. Fiedler & P. Berthold). Oxford University Press.
619 Both, C., Bouwhuis, S., Lessells, C.M. & Visser, M.E. (2006) Climate change and population declines in
620   a long-distance migratory bird. *Nature,* **441,** 81-83.
621 Both, C., Van Turnhout, C.A.M., Bijlsma, R.G., Siepel, H., Van Strien, A.J. & Foppen, R.P.B. (2010) Avian
622   population consequences of climate change are most severe for long-distance migrants in
623   seasonal habitats. *Proceedings of the Royal Society B: Biological Sciences,* **277,** 1259-1266.
624 Both, C., Visser, M.E. & Verboven, N. (1999) Density-dependent recruitment rates in great tits: the
625   importance of being heavier. *Proceedings of the Royal Society of London Series B-Biological*
626   *Sciences,* **266,** 465-469.
627 Charlesworth, B. (1971) Selection in density-regulated populations. *Ecology,* **52,** 469-&.
628 Charlesworth, B. (1994) *Evolution in age-structured populations*. Cambridge University Press.
629 Charmantier, A., McCleery, R.H., Cole, L.R., Perrins, C., Kruuk, L.E.B. & Sheldon, B.C. (2008) Adaptive
630   phenotypic plasticity in response to climate change in a wild bird population. *Science,* **320,**
631   800-803.
632 Clobert, J., Perrins, C.M., McCleery, R.H. & Gosler, A.G. (1988) Survival rate in the great tit *Parus*
633   *major* L. in relation to sex, age, and immigration status. *Journal of animal ecology,* **57,** 287-
634   306.
635 Clutton-Brock, T. (1998) *Reproductive success: studies of invididual variation in contrasting breeding*
636   *systems*. University of Chicago Press, Chicago.
637 Coulson, T., Catchpole, E.A., Albon, S.D., Morgan, B.J.T., Pemberton, J.M., Clutton-Brock, T.H.,
638   Crawley, M.J. & Grenfell, B.T. (2001) Age, sex, density, winter weather, and population
639   crashes in Soay sheep. *Science,* **292,** 1528-1531.
640 Coulson, T., Tuljapurkar, S. & Childs, D.Z. (2010) Using evolutionary demography to link life history
641   theory, quantitative genetics and population ecology. *Journal of animal ecology,* **79,** 1226-
642   1240.
643 Dhondt, A.A., Kempenaers, B. & Adriaensen, F. (1992) Density-dependent clutch size caused by
644   habitat heterogeneity. *Journal of animal ecology,* **61,** 643-648.
645 Drever, M.C. & Clark, R.G. (2007) Spring temperature, clutch initiation date and duck nest success: a
646   test of the mismatch hypothesis. *Journal of animal ecology,* **76,** 139-148.
647 Dunn, P.O., Winkler, D.W., Whittingham, L.A., Hannon, S.J. & Robertson, R.J. (2011) A test of the
648   mismatch hypothesis: How is timing of reproduction related to food abundance in an aerial
649   insectivore? *Ecology,* **92,** 450-461.
650 Eeva, T., Veistola, S. & Lehikoinen, E. (2000) Timing of breeding in subarctic passerines in relation to
651   food availability. *Canadian Journal of Zoology,* **78,** 67-78.
652 Endler, J.A. (1986) *Natural selection in the wild*. Princeton University Press.
653 Gebhardt-Henrich, S. (1990) Temporal and spatial variation in food availability and its effect on
654   fledgling size in the great tit. *Population biology of passerine birds* (eds J. Blondel, A. Gosler,
655   J.-D. Lebreton & R.H. McCleery), pp. 175–186. Springer, Berlin, Heidelberg, New york.
656 Gienapp, P., Postma, E. & Visser, M.E. (2006) Why breeding time has not responded to selection for
657   earlier breeding in a songbird population. *Evolution,* **60,** 2381-2388.
658 Grøtan, V., Saether, B.E., Engen, S., van Balen, J.H., Perdeck, A.C. & Visser, M.E. (2009) Spatial and
659   temporal variation in the relative contribution of density dependence, climate variation and
660   migration to fluctuations in the size of great tit populations. *Journal of animal ecology,* **78,**
661   447-459.
662 Harvey, P.H., Greenwood, P.J., Perrins, C.M. & Martin, A.R. (1979) Breeding success of great tits
663   (*Parus major*) in relation to age of male and female parent. *Ibis,* **121,** 216-219.

664   Heard, M.J., Riskin, S.H. & Flight, P.A. (2011) Identifying potential evolutionary consequences of
665   climate-driven phenological shifts. *Evolutionary Ecology*, 1-9.
666   Husby, A., Kruuk, L.E.B. & Visser, M.E. (2009) Decline in the frequency and benefits of multiple
667   brooding in great tits as a consequence of a changing environment. *Proceedings of the Royal*
668   *Society B-Biological Sciences,* **276,** 1845-1854.
669   Jarvinen, A. (1991) A meta-analytic study of the effects of female age on laying date and clutch size in
670   the great tit (*Parus major*) and the pied flycatcher *(Ficedula hypoleuca)*. *Ibis,* **133,** 62-66.
671   Jenouvrier, S., Holland, M., Stroeve, J., Barbraud, C., Weimerskirch, H., Serreze, M. & Caswell, H.
672   (2012) Effects of climate change on an emperor penguin population: analysis of coupled
673   demographic and climate models. *Global Change Biology*.
674   Jenouvrier, S. & Visser, M.E. (2011) Climate change, phenological shifts, eco-evolutionary responses
675   and population viability: toward a unifying predictive approach. *International Journal of*
676   *Biometeorology,* **55,** 905-919.
677   Jones, T. & Cresswell, W. (2010) The phenology mismatch hypothesis: are declines of migrant birds
678   linked to uneven global climate change? *Journal of animal ecology,* **79,** 98-108.
679   Keller, L.F. & van Noordwijk, A.J. (1994) Effects of local environmental conditions on nestling growth
680   in the great tit (*Parus major* L.). *Ardea,* **82,** 349-362.
681   Kingsolver, J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., Hoang, A., Gibert,
682   P. & Beerli, P. (2001) The strength of phenotypic selection in natural populations. *American*
683   *Naturalist,* **157,** 245-261.
684   Kinnison, M.T. & Hairston, N.G. (2007) Eco-evolutionary conservation biology: contemporary
685   evolution and the dynamics of persistence. *Functional Ecology,* **21,** 444-454.
686   Kluijver, H.N. (1951) The population ecology of the Great Tit, *Parus m. major* L. . *Ardea* **39**.
687   Kokko, H. & Lopez-Sepulcre, A. (2007) The ecogenetic link between demography and evolution: can
688   we bridge the gap between theory and data? *Ecology Letters,* **10,** 773-782.
689   Lande, R. & Arnold, S.J. (1983) The measurement of selection on correlated characters. *Evolution,* **37,**
690   1210-1226.
691   Lande, R., Engen, S. & Saether, B.-E. (2003) Stochastic population dynamics in ecology and
692   conservation. *Stochastic population dynamics in ecology and conservation.***,** i-x, 1-212.
693   Leech, D.I. & Crick, H.Q.P. (2007) Influence of climate change on the abundance, distribution and
694   phenology of woodland bird species in temperate regions. *Ibis,* **149,** 128-145.
695   Lof, M.E., Reed, T.E., McNamara, J.M. & Visser, M.E. (2012) Timing in a fluctuating environment:
696   environmental variability and asymmetric fitness curves can lead to adaptively mismatched
697   avian reproduction. **In Press**.
698   Malcolm, J.R., Liu, C.R., Neilson, R.P., Hansen, L. & Hannah, L. (2006) Global warming and extinctions
699   of endemic species from biodiversity hotspots. *Conservation Biology,* **20,** 538-548.
700   Metcalf, C.J.E. & Pavard, S. (2007) Why evolutionary biologists should be demographers. *Trends in*
701   *ecology & evolution,* **22,** 205-212.
702   Miller-Rushing, A.J., Hoye, T.T., Inouye, D.W. & Post, E. (2010) The effects of phenological
703   mismatches on demography. *Philosophical Transactions of the Royal Society B-Biological*
704   *Sciences,* **365,** 3177-3186.
705   Mols, C.M.M., van Noordwijk, A.J. & Visser, M.E. (2005) Assessing the reduction of caterpillar
706   numbers by Great Tits Parus major breeding in apple orchards. *Ardea,* **93,** 259-269.
707   Morrisey, M.B. & Hadfield, J.D. (2012) Directional selection in temporally replicated studies is
708   remarkably consistent. *Evolution,* **66,** 435–442.
709   Naef-Daenzer, L., Naef-Daenzer, B. & Nager, R.G. (2000) Prey selection and foraging performance of
710   breeding Great Tits Parus major in relation to food availability. *Journal of Avian Biology,* **31,**
711   206-214.
712   Nussey, D.H., Postma, E., Gienapp, P. & Visser, M.E. (2005) Selection on heritable phenotypic
713   plasticity in a wild bird population. *Science,* **310,** 304-306.
714   Parmesan, C. & Yohe, G. (2003) A globally coherent fingerprint of climate change impacts across
715   natural systems. *Nature,* **421,** 37-42.

716 Perdeck, A.C., Visser, M.E. & Van Balen, J.H. (2000) Great Tit Parus major survival, and the beech-crop
717   cycle. *Ardea,* **88,** 99-108.
718 Perrins, C. (1970) The timing of birds' breeding seasons. *Ibis,* **112,** 242-255.
719 Perrins, C.M. (1965) Population fluctuations and clutch size in the great tit *Parus major* L. *Journal of*
720   *animal ecology,* **34,** 601-647.
721 Perrins, C.M. & Moss, D. (1974) Survival of young great tits in relation to age of female parent. *Ibis,*
722   **116,** 220-224.
723 Reed, T.E., Schindler, D.E. & Waples, R.S. (2010) Interacting effects of phenotypic plasticity and
724   evolution on population persistence in a changing climate. *Conservation Biology,* **25,** 56-63.
725 Robinson, W.S. (1950) "Ecological Correlations and the Behavior of Individuals". . *American*
726   *Sociological Review,* **15,** 351–357.
727 Root, T.L., Price, J.T., Hall, K.R., Schneider, S.H., Rosenzweig, C. & Pounds, J.A. (2003) Fingerprints of
728   global warming on wild animals and plants. *Nature,* **421,** 57-60.
729 Royama, T. (1966) Factors governing feeding rate, food requirement and brood size of nestling Great
730   Tits *Parus major Ibis,* **108,** 313-&.
731 Saccheri, I. & Hanski, I. (2006) Natural selection and population dynamics. *Trends in ecology &*
732   *evolution,* **21,** 341-347.
733 Saether, B.E., Engen, S., Islam, A., McCleery, R. & Perrins, C. (1998) Environmental stochasticity and
734   extinction risk in a population of a small songbird, the great tit. *American Naturalist,* **151,**
735   441-450.
736 Saether, B.E., Sutherland, W.J. & Engen, S. (2004) Climate influences on avian population dynamics.
737   *Birds and Climate Change* (eds A.P. Moller, W. Fielder & P. Berthold), pp. 185-209.
738 Schaper, S.V., Dawson, A., Sharp, P.J., Gienapp, P., Caro, S.P. & Visser, M.E. (2012) Increasing
739   temperature, not mean temperature, is a cue for avian timing of reproduction. *American*
740   *Naturalist,* **179,** E55-E69.
741 Shultz, M.T., Piatt, J.F., Harding, A., Kettle, A.B. & Van Pelt, T.I. (2009) Timing of breeding and
742   reproductive performance in murres and kittiwakes reflect mismatched seasonal prey
743   dynamics. *Marine Ecology Progress Series,* **393,** 247-258.
744 Siepielski, A.M., DiBattista, J.D. & Carlson, S.M. (2009) It's about time: the temporal dynamics of
745   phenotypic selection in the wild. *Ecology Letters,* **12,** 1261-1276.
746 Stinchcombe, J.R., Agrawal, A.F., Hohenlohe, P.A., Arnold, S.J. & Blows, M.W. (2008) Estimating
747   nonlinear selection gradients using quadratic regression coefficients: Double or nothing ?
748   *Evolution,* **62,** 2435-2440.
749 Thackeray, S.J., Sparks, T.H., Frederiksen, M., Burthe, S., Bacon, P.J., Bell, J.R., Botham, M.S.,
750   Brereton, T.M., Bright, P.W., Carvalho, L., Clutton-Brock, T., Dawson, A., Edwards, M., Elliott,
751   J.M., Harrington, R., Johns, D., Jones, I.D., Jones, J.T., Leech, D.I., Roy, D.B., Scott, W.A., Smith,
752   M., Smithers, R.J., Winfield, I.J. & Wanless, S. (2010) Trophic level asynchrony in rates of
753   phenological change for marine, freshwater and terrestrial environments. *Global Change*
754   *Biology,* **16,** 3304-3313.
755 Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus,
756   B.F.N., de Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S.,
757   Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O.L. & Williams, S.E.
758   (2004) Extinction risk from climate change. *Nature,* **427,** 145-148.
759 van Balen, J.H. (1973) A comparative study of the breeding ecology of the great tit *Parus major* in
760   different habitats. *Ardea,* **61,** 1-93.
761 van Balen, J.H. (1980) Population fluctuations of the Great Tit and feeding conditions in winter.
762   *Ardea,* **68,** 143-164.
763 van de Pol, M.V. & Wright, J. (2009) A simple method for distinguishing within- versus between-
764   subject effects using mixed models. *Animal Behaviour,* **77,** 753-758.
765 van Noordwijk, A.J., McCleery, R.H. & Perrins, C.M. (1995) Selection for the timing of great tit
766   breeding in relation to caterpillar growth and temperature. *Journal of animal ecology,* **64,**
767   451-458.

768 Vatka, E., Orell, M. & Rytkönen, S. (2011) Warming climate advances breeding and improves
769   synchrony of food demand and food availability in a boreal passerine. *Global Change Biology*.
770 Verboven, N., Tinbergen, J.M. & Verhulst, S. (2001) Food, reproductive success and multiple breeding
771   in the Great Tit Parus major. *Ardea,* **89,** 387-406.
772 Verboven, N. & Visser, M.E. (1998) Seasonal variation in local recruitment of great tits: the
773   importance of being early. *Oikos,* **81,** 511-524.
774 Visser, M.E. (2008) Keeping up with a warming world; assessing the rate of adaptation to climate
775   change. *Proceedings of the Royal Society B-Biological Sciences,* **275,** 649-659.
776 Visser, M.E., Adriaensen, F., van Balen, J.H., Blondel, J., Dhondt, A.A., van Dongen, S., du Feu, C.,
777   Ivankina, E.V., Kerimov, A.B., de Laet, J., Matthysen, E., McCleery, R., Orell, M. & Thomson,
778   D.L. (2003) Variable responses to large-scale climate change in European Parus populations.
779   *Proceedings of the Royal Society of London Series B-Biological Sciences,* **270,** 367-372.
780 Visser, M.E. & Both, C. (2005) Shifts in phenology due to global climate change: the need for a
781   yardstick. *Proceedings of the Royal Society B-Biological Sciences,* **272,** 2561-2569.
782 Visser, M.E., Both, C. & Lambrechts, M.M. (2004) Global climate change leads to mistimed avian
783   reproduction. *Advances in Ecological Research,* **35,** 89-110.
784 Visser, M.E., Holleman, L.J.M. & Caro, S.P. (2009) Temperature has a causal effect on avian timing of
785   reproduction. *Proceedings of the Royal Society B-Biological Sciences,* **276,** 2323-2331.
786 Visser, M.E., Holleman, L.J.M. & Gienapp, P. (2006) Shifts in caterpillar biomass phenology due to
787   climate change and its impact on the breeding biology of an insectivorous bird. *Oecologia,*
788   **147,** 164-172.
789 Visser, M.E. & Lessells, C.M. (2001) The costs of egg production and incubation in great tits (Parus
790   major). *Proceedings of the Royal Society of London Series B-Biological Sciences,* **268,** 1271-
791   1277.
792 Visser, M.E., te Marvelde, L. & Lof, M.E. (2011) Adaptive phenological mismatches of birds and their
793   food in a warming world *Journal of Ornithology,* **Early online,** DOI: 10.1007/s10336-10011-
794   10770-10336.
795 Visser, M.E., van Noordwijk, A.J., Tinbergen, J.M. & Lessells, C.M. (1998) Warmer springs lead to
796   mistimed reproduction in great tits (Parus major). *Proceedings of the Royal Society of London
797   Series B-Biological Sciences,* **265,** 1867-1870.
798 Wilkin, T.A., King, L.E. & Sheldon, B.C. (2009) Habitat quality, nestling diet, and provisioning
799   behaviour in great tits Parus major. *Journal of Avian Biology,* **40,** 135-145.

800

801

802

803

804

805

806

807

808

809   The following Supporting Information is available for this article online:

810   Appendix S1: Potential limitations of the mismatch measure

811   Fig. S1: Population-level plots of mean demographic rates versus year

812   Fig. S2: Annual recapture probability as a function of year and population-level mismatch

813   Fig. S3: Temporal trends in linear selection, nonlinear selection and population mismatch.

814   Fig. S4: The number of recruits plotted as a function of IM, splitting the data in negative and

815   positive mismatch years.

816   Fig. S5: The sensitivity of patterns in Fig. 1 to how IM was defined.

817

818

819

820

821

822

823

824

825

826

Table 1. Separating the effects of within-year variation in individual mismatch from

between-year variation in average mismatch on components of great tit fitness in the Hoge

Veluwe study population, Netherlands, from 1973-2010. Each sub-table represents the

minimum adequate models for that fitness component. PM = population mismatch. IM′ =

standardized individual mismatch. The levels for the factor 'Mother age' are abbreviated as:

EXB = experienced breeder, FTB = first-time breeder, U = unknown age. The levels for the

factor BCI (beech crop index) are simply 1, 2 and 3. Intercepts correspond to the level EXB

for mother age and 1 for BCI. Estimates are on the logit scale for models with binomial errors

and on the log scale for models with Poisson errors. ID VC = variance component for random

effect of female identity. Year VC = variance component for random effect of year. n_o =

number of total observations. n_f = number of females. n_y = number of years.

	Estimate	SE	z-value	$\Pr(>|z|)$
(a) Probability of double-brooding (binomial errors, ID VC = 0.544, Year VC = 0.814, n_o = 3472, n_f = 2599, n_y = 37)				
Intercept	1.672	0.763	2.192	0.028
PM	-0.225	0.029	-7.743	< 0.001
IM′	-0.025	0.072	-0.347	0.728
IM′^2	-0.006	0.002	-2.948	0.003
Mother age				
FTB	-0.365	0.129	-2.836	0.005
U	-5.673	1.146	-4.950	< 0.001
PM x IM′	-0.010	0.006	-3.557	< 0.001
Density	-0.027	0.003	-4.537	< 0.001
Density x IM′	-0.001	0.001	2.448	0.014
(b) Clutch size (Poisson errors, ID VC < 0.001, Year VC = 0.004, n_o = 3131, n_f = 2263, n_y = 37)				
Intercept	2.420	0.051	47.44	< 0.001
IM′	-0.009	0.001	-7.61	< 0.001
Mother age				
FTB	-0.021	0.013	-1.59	0.112
U	-0.084	0.022	-3.78	< 0.001
Density	-0.002	> 0.001	-4.07	< 0.001
(c) Probability of producing zero chicks (binomial errors, Year VC = 0.188, n_o = 3469, n_f = 2599, n_y = 37)				
Intercept	-2.563	0.133	-19.284	< 0.001
PM	0.006	0.015	0.394	0.694
IM′	0.049	0.012	4.061	< 0.001
Mother age				
FTB	0.164	0.142	1.154	0.249
U	4.531	0.166	27.375	< 0.001
PM x IM′	0.006	0.002	3.519	< 0.001
(d) Number of fledglings produced (Poisson errors, ID VC = 0.007, Year VC = 0.016, n_o = 2680, n_f = 1896, n_y = 37)				
Intercept	2.555	0.096	26.658	< 0.001

PM	-0.008	0.004	-2.355	0.019
IM′	-0.015	0.002	-9.577	< 0.001
IM′^2	-0.0006	0.0002	-3.011	0.003
Mother age				
FTB	- 0.079	0.015	-5.170	< 0.001
U	-0.105	0.045	-2.317	0.023
Density	-0.004	0.0007	-5.149	< 0.001

(e) Probability of recruitment (binomial errors, ID VC = 0.293, Year VC = 0.150, n_o = 2680, n_f = 1896, n_y = 37)

Intercept	-1.765	0.353	-5.003	< 0.001
PM	-0.017	0.012	-1.436	0.151
IM′	-0.023	0.007	-3.308	< 0.001
Density	-0.011	0.003	-4.247	< 0.001
BCI				
2	0.578	0.184	3.140	0.002
3	0.771	0.217	3.558	< 0.001
PM x IM′	-0.003	0.001	-2.634	< 0.001

(f)   Number of recruits (Poisson errors, IDVC = 0.320, Year VC = 0.161, n_o = 3472, n_f = 2599, n_y = 37)

Intercept	0.650	0.361	-5.003	0.072
PM	-0.025	0.012	-1.436	0.038
IM′	-0.039	0.007	-3.308	< 0.001
IM′^2	-0.002	0.001	-4.247	0.044
Mother age				
FTB	-0.138		0.065	0.032
U	-2.330		0.199	< 0.001
Density	-0.014		0.003	< 0.001
BCI				
2	0.469	0.184	0.187	0.012
3	0.513	0.217	0.221	0.020
PM x IM′	-0.004	0.001	0.001	0.001

(g) Female adult survival (binomial errors, ID VC < 0.001, Year VC = 0.156, n_o = 2912, n_f = 2039, n_y = 37)

Intercept	0.409	0.374	1.095	0.273
Density	-0.008	0.003	-2.955	0.003
BCI				
2	0.411	0.201	2.049	0.041
3	0.535	0.231	2.317	0.021

(h)   Male adult survival (binomial errors, ID VC < 0.001, Year VC = 0.151, n_o = 2912, n_f = 2039, n_y = 37)

Intercept	0.424	0.371	1.146	0.252
Male age				
FTB	-0.016	0.080	-0.195	0.846
U	-1.024	0.466	-2.196	0.028
Density	-0.008	0.003	-2.990	0.003
BCI				
2	0.417	0.199	2.101	0.036
3	0.539	0.228	2.357	0.018

839

840

Figure legends:

Figure 1: Schematic illustration of population/individual-level mismatch. In both panels,

solid black curves show the distribution of laying dates and dashed black curves show the

845   distribution of chick food needs, which peak 30 days after egg-laying. Shaded portions

846   represent female great tits that lay later than the annual average, open portions represent

847   females that lay earlier than the population average. Solid grey curves show the seasonal

848   distribution of caterpillar biomass. *Top panel*: example of a year where there is no population-

849   level mismatch (PM) between the breeding phenology of great tits and the seasonal peak in

850   caterpillar biomass. Late-laying females nonetheless produce broods after the caterpillar peak,

851   and thus exhibit positive values for individual mismatch (IM). Early females exhibit negative

852   values for individual mismatch. *Bottom panel*: example of a year where caterpillar biomass

853   peaks earlier, but there is no change in laying dates, which results in (a positive value for)

854   population-level mismatch. Individual females breeding late relative to the food peak exhibit

855   positive values for individual mismatch in this year, but so too do the earliest females, who

856   are classified as breeding late relative to the food peak.

857

858   **Figure 2:** Individual-level plots of fitness components versus individual mismatch. Data are

859   binned into 10 equally-spaced categories along the individual mismatch axis for ease of

860   illustration (so each data point potentially consists of observations on the same or different

861   females across years) but the statistical analyses are based on the full dataset, with sample

862   sizes given in Table 1. Curves show significant within-year effects of IM, predicted and back-

863   transformed from the GLMMs which also accounted for between-year effects of PM (see

864   Table 1). Error bars are standard errors.

865

866   **Figure 3:** Population-level plots of average fitness components (demographic rates) versus

867   population mismatch. Data points are annual averages. Curves show the predicted, back-

868   transformed fits for the effect of population mismatch from the minimum adequate GLMMs

869   for each fitness component, summarized in Table 1. Error bars are standard errors.
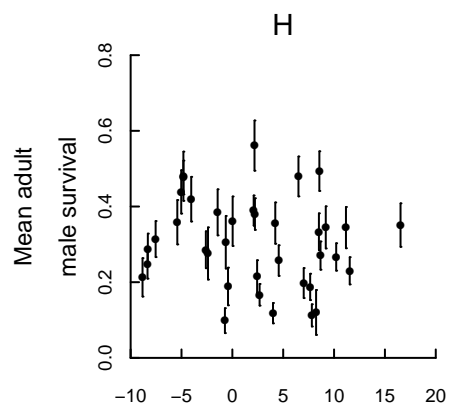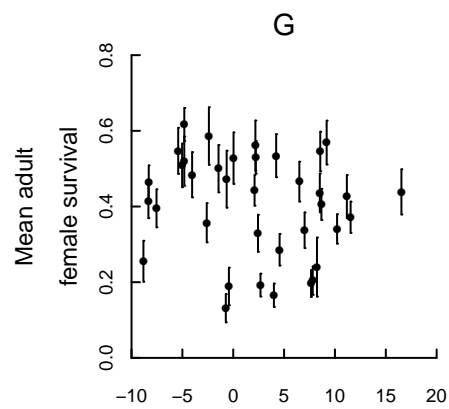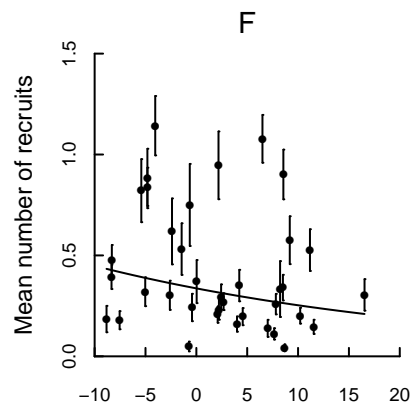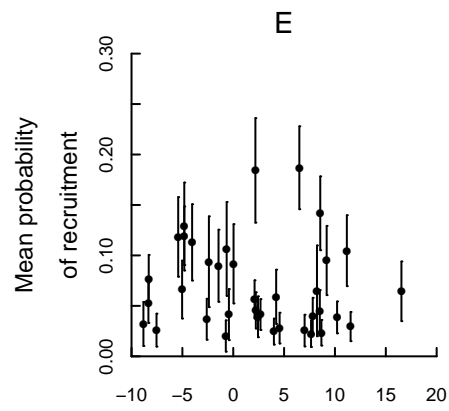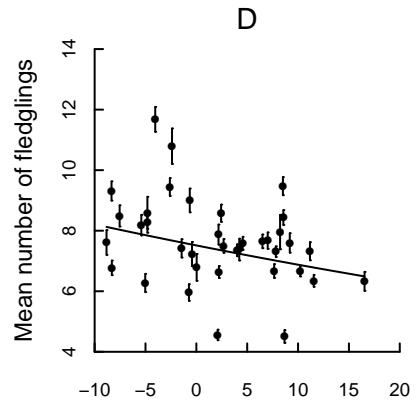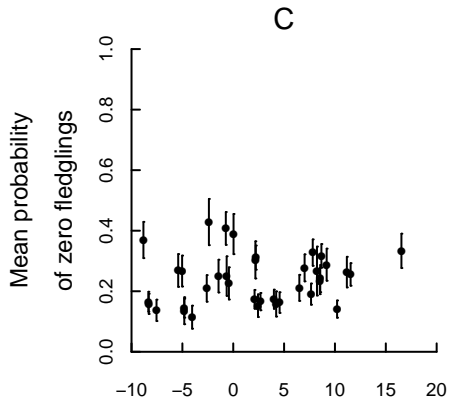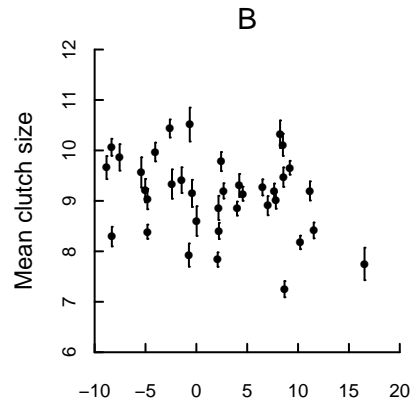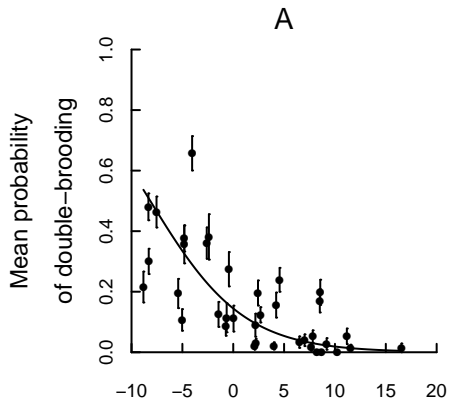
870

871   **Figure 4:** Annual standardized linear selection differentials (β_t) plotted against average

872   population mismatch. Curve shows best-fit from a quadratic model, weighting each data point

873   by $1/[(\text{standard error of } \beta_t)^2]$.
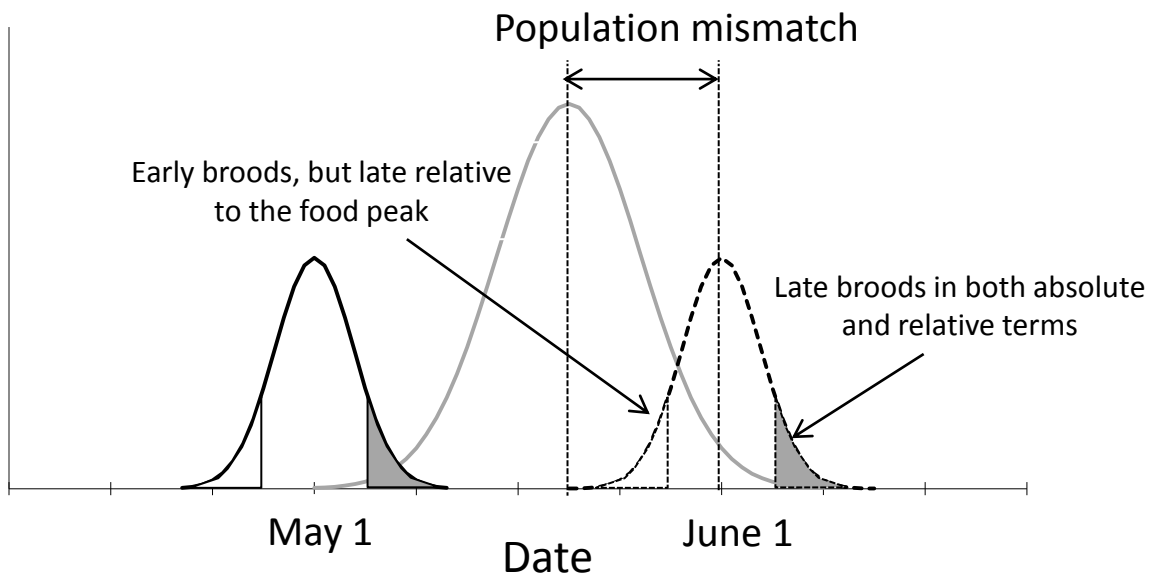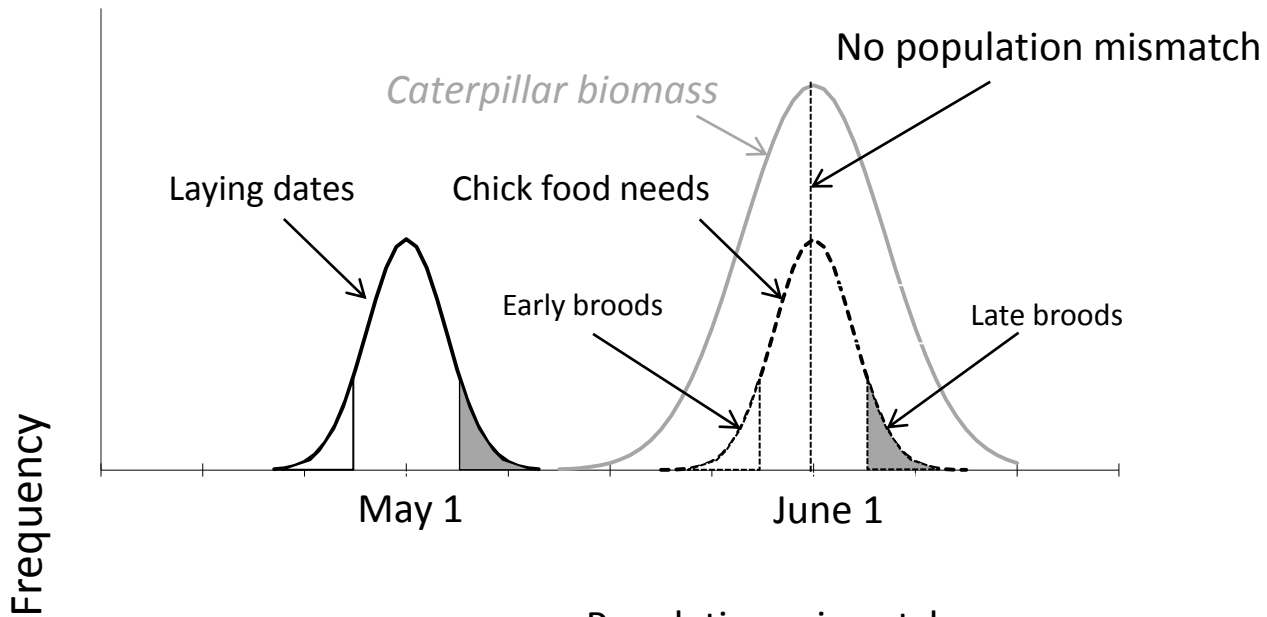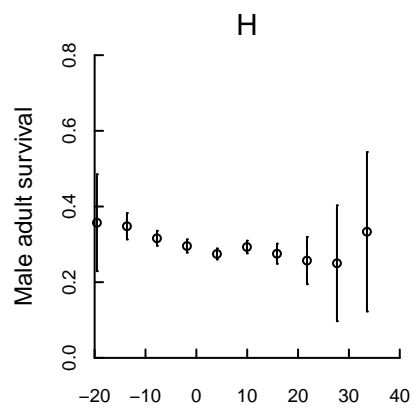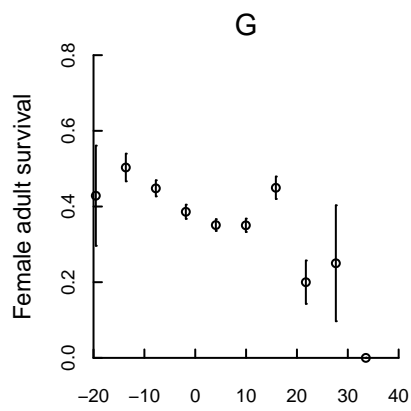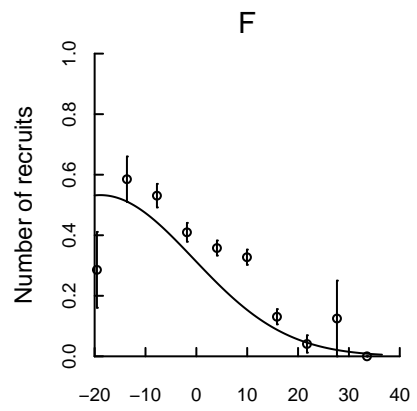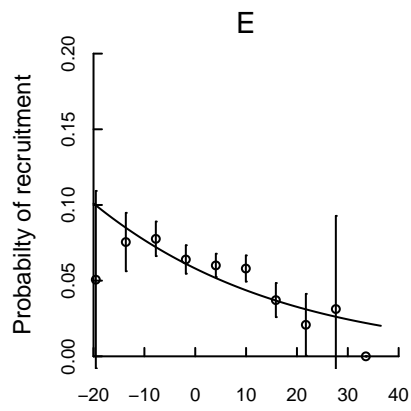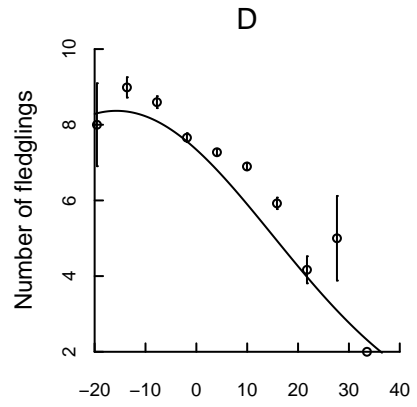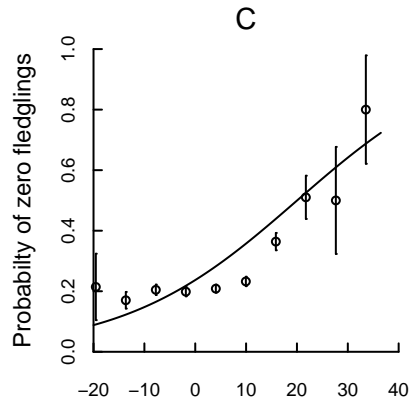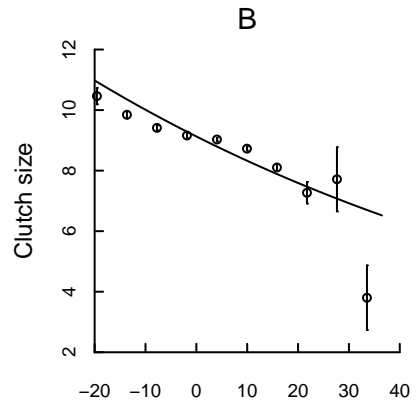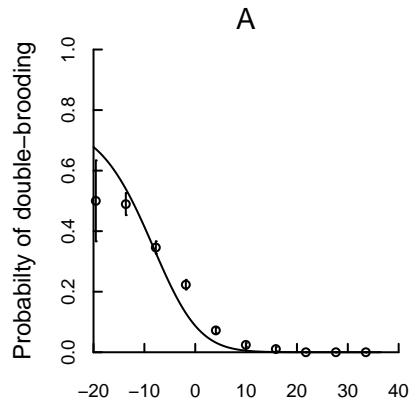
Individual mismatch (days)

Population mismatch (days)