**Appendix:**

POTENTIAL LIMITATIONS OF THE MISMATCH MEASURE

While our phenological mismatch measure provides a useful ‘yardstick’ of how well birds manage to track temporal shifts in their food supply ([Visser & Both 2005](#_ENREF_75)), it assumes that reproductive success is largely determined by the temporal overlap between the period when chicks require most food and the caterpillar peak. Asides from the timing of peak food availability, the length of the period over which caterpillars are available (width of the food peak) and the maximal amount of caterpillar biomass available (height of the food peak) could also be important ([Durant *et al.* 2005](#_ENREF_16); [Visser 2008](#_ENREF_73)). For example, a good match with the food peak could be crucial in years where the peak is lower or narrower, but relatively unimportant in years where caterpillars are plentiful for extended periods ([Perrins 1991](#_ENREF_50)). In our study population, there is no correlation between the width and either the height or the timing of the peak, although the peak tends to be higher in late years (Visser *et al*. 2006). Moreover, the main effect of climate change to date has been to shift caterpillar peak dates, whereas no significant changes in the width or height of the peak have been observed (Visser *et al.* 2006). While there is clear spatial variation in caterpillar peak dates among sampled oaks in our study area ([Grieco, van Noordwijk & Visser 2002; Visser et al. 2006](#_ENREF_20)), relative differences are persistent across years – an early year is an early year for all trees. We lacked sufficient spatial resolution of caterpillar phenology and information on great tit foraging ranges to be able to define individual mismatch at the scale of breeding territories. Thus, breeding time relative to the estimated caterpillar peak date for the entire study area remains our best estimate of phenological mismatch at both the individual and population levels. This measure was also highly correlated with our ‘annual phenology overlap’ measure, which quantified more directly the extent of temporal overlap between caterpillar availability and chick food needs (see Materials and Methods). While the latter explains more of the variation in great tit performance (Reed & Visser, unpublished), the former is more straightforward to estimate and interpret, while it clearly also captures much of the variation in individual reproductive success (Fig. 2, Table 1).

We did not account for individual variation in clutch size when calculating IM, which could introduce a potential bias into our estimation of the relationships between IM and fledging/recruitment success, given that late breeders tend to lay smaller clutches. However, the patterns remain largely unchanged when clutch size variation was taken into account (see Fig. S5 and its legend for full explanation).

Cited papers:

Durant, J.M., Hjermann, D.O., Anker-Nilssen, T., Beaugrand, G., Mysterud, A., Pettorelli, N. & Stenseth, N.C. (2005) Timing and abundance as key mechanisms affecting trophic interactions in variable environments. *Ecology Letters,* **8,** 952-958.

Grieco, F., van Noordwijk, A.J. & Visser, M.E. (2002) Evidence for the effect of learning on timing of reproduction in blue tits. *Science,* **296,** 136-138.

Perrins, C.M. (1991) Tits and their caterpillar food supply. *Ibis,* **133,** 49-54.

Visser, M.E. & Both, C. (2005) Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings of the Royal Society B-Biological Sciences,* **272,** 2561-2569.

Visser, M.E., Holleman, L.J.M. & Gienapp, P. (2006) Shifts in caterpillar biomass phenology due to climate change and its impact on the breeding biology of an insectivorous bird. *Oecologia,* **147,** 164-172.

Visser, M.E. (2008) Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings of the Royal Society B-Biological Sciences,* **275,** 649-659.

**Figure S1:** Population-level plots of average fitness components versus year. Data points are annual averages. Error bars are standard errors.

**Figure S2:** Annual recapture probability as a function of (A) year and (B) annual population-level mismatch. Open triangles = females; closed triangles = males.

**Figure S.3.** Panel A: Annual standardized linear selection differentials (*βt*) plotted against year. Curve shows best-fit from a quadratic model, weighting each data point by 1/[(standard error of *βt* )2]. Panel B: average population mismatch plotted against year. Line shows best-fit from a linear model. Panel C: Annual standardized nonlinear selection differentials (*γt*) plotted against year. Curve shows best-fit from a quadratic model, weighting each data point by 1/[(standard error of *λt* )2]. Panel D: *γt* plotted against annual average population mismatch (no significant linear or quadratic relationship).

**Figure S.4.** The number of recruits plotted as a function of IM, splitting the data into negative mismatch years (left panel, years where many individuals bred too early relative to the food peak) and positive mismatch years (right panel). Data are binned into 10 equally-spaced categories along the individual mismatch axis for ease of illustration When the data are split this way, the bell-shaped relationship between fitness and IM becomes more obvious, but when all years are combined (as in Fig. 2 in the manuscript), the curves come down less on the left and directional selection is the overall net pattern.

**Figure S.5.** Testing whether the relationships between fitness components and IM were sensitive to how IM was defined. As in Fig.2 in the main manuscript, data are binned into 10 equally-spaced categories along the individual mismatch axis for ease of illustration. Blue symbols represents the way IM is calculated in the main manuscript, i.e. IM = lay date + 10 + 12 + 9 – food peak date. Red symbols take into account individual variation in clutch size, i.e. IM = lay date + observed clutch size + 12 + 9 – food peak date. Green symbols represent IM calculated based on observed hatch dates (i.e. IM = hatch date + 9 – food peak date), and therefore taking into account individual variation in clutch size and the length of the period between clutch completion and hatch date. This figure shows that the patterns are largely robust to how IM is defined:

1. The relationship between probability of double-brooding and IM is largely unaffected.
2. When clutch size variation is taken into account (red data points), the relationship between clutch size and IM becomes flatter, or even slightly hump-shaped. This makes sense, because early layers tend to lay larger clutches than late layers, and are shifted to the right along the IM axis when clutch size variation is taken into account.
3. The elevation of the relationship between prob. of producing 0 fledglings and IM become lower when clutch size variation is taken into account, and even lower when variation in the interval between clutch completion and hatching (hereafter simply ‘interval’) is accounted for. This is because some nests fail completely at the egg stage, which drop out of the analysis when clutch size variation and interval variation is accounted for.
4. The relationship between no. of fledglings and IM becomes more hump-shaped, with more of a drop on the left of the peak, when CS variation is accounted for. The slope of the decline to the right of the peak becomes weaker. Additionally accounting for interval variation seems to make the relationship similar again to the original blue one.
5. The relationship for recruitment probability is little affected by the method of calculating IM. It is slightly more hump-shaped when interval variation is accounted for.
6. The relationships for no. of recruits is largely unaffected by the chosen IM measure, except that the elevations go up when CS variation and interval variation are accounted for. This is because nests that fail completely at the egg stage are omitted with the red and green data points.
7. The relationship for female survival is largely unaffected.
8. The relationship for male survival is largely unaffected.

Figure S1

FF

E



Figure S2



B

A

Figure S3



A

B

C

D

Figure S4

**

Figure S5

