Breeding in a Warming World

Evolution of Avian Breeding Time under Climate Change

RIJKSUNIVERSITEIT GRONINGEN

Breeding in a Warming World

Evolution of Avian Breeding Time under Climate Change

Proefschrift

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
vrijdag 21 oktober 2005
om 14.45 uur

door

Phillip Gienapp
geboren op 17 mei 1971
te Neumünster, Duitsland

Promotor: Prof. dr. M. E. Visser
Beoordelingscommissie: Prof. dr. S. Daan
Prof. dr. B. C. Sheldon
Prof. dr. J. J. M. van Alphen

The research presented in this thesis was carried out at the Netherlands Institute of Ecology (NIOO-KNAW), Centre for Terrestrial Ecology, P.O. Box 40, 6666 ZG Heteren, The Netherlands.

The cover photographs show a male great tit (ringnr. AE27559) bringing food to his brood in a typical NIOO-nest box. He was born in 2002. In 2003, when the photos were taken, he bred for the first (and last) time. In that year, together with his mate who bred in the same nest box already in the previous year, he raised five chicks.
Photos by P. Gienapp

Printed by Ponsen & Looijen BV, Wageningen, The Netherlands

ISBN 90-9019926-8

Contents

Chapter 1

General introduction

Biotic and abiotic environmental conditions change continuously throughout the year and organisms have to fit the different stages of their life-cycle, like mating, breeding and dispersing, into this seasonal pattern. These life-cycle stages are interdependent and to some extent incompatible. Obviously a mate has to be found and courted before it is possible to breed and simultaneously breeding and moulting has a fitness cost (e.g. Hemborg and Lundberg 1998; Jenni and Winkler 1994). For many, if not most, species environmental conditions are favourable for only a limited time and this leaves thus only a short optimal period for the the different stages of their life-cycle. For example, migrating birds should not arrive too early at their breeding grounds because they may suffer from harsh conditions early in season, when it is still too cold and food abundance low (Brown and Brown 2000). On the other hand, they should not arrive too late because late arriving individuals may find all suitable breeding territories taken (Forstmeier 2002) or may not have enough time to moult between breeding and return migration. Consequently, the seasonal timing of many behaviours has important fitness consequences and should be under strong selection.

Understanding how selection and inheritance have determined the way in which species respond to their environment to time their behaviour lies at the heart of evolutionary biology. It requires knowledge of selection pressures and the amount of genetic variation for a given trait. For a full understanding of a species' response to its environment we need to identify and quantify the ecological factors (van Noordwijk *et al*. 1995), for example food abundance, that are responsible for the observed selection pressures . We also need to know whether a species is responding directly to these factors or other environmental cues or internal rhythms. Climate change will likely lead to changed selection on the way species should time their behaviour in response to their environment. These changing selection pressures offer a unique opportunity to evolutionary biologists to study evolution in the wild but, at the same time, they also pose a great threat to global biodiversity.

Effects of climate change on biological systems

During the last 100 years temperatures have risen globally by about 0.5 °C (Houghton *et al*. 2001); Fig. 1.1 shows this temperature increase for the Netherlands. The increase in temperature is not globally uniform; instead the trends differ between regions and within seasons. Winter temperatures have increased more than summer temperatures, and temperatures in northern hemisphere have increased more than in the southern hemisphere (Easterling *et al*. 1997). This spatial and temporal heterogeneity may have important consequences for biological systems (Stenseth and Mysterud 2002; Walther *et al*. 2002).

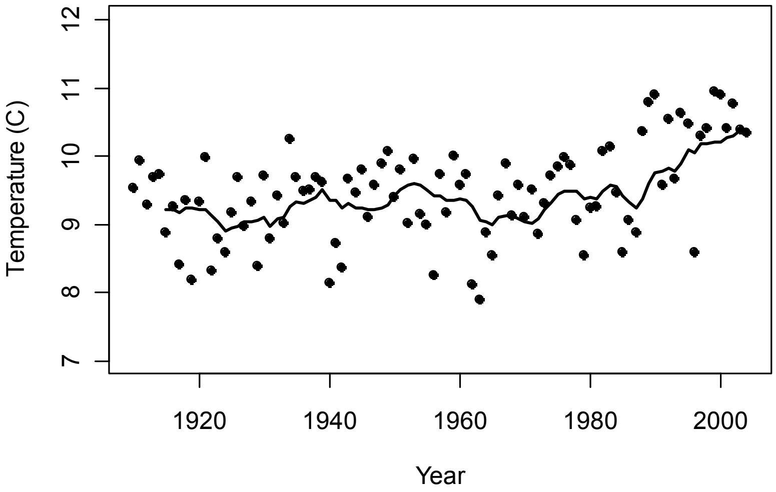

Fig. 1.1 Long-term temperature trends in the Netherlands (1905-2004). Annual mean temperatures (dots) and 10 year-running mean (line) are shown for the Royal Dutch Meteorological Institute (KNMI) weather station at De Bilt. Daily mean temperatures were obtained from the KNMI's website.

Numerous studies covering a wide range of taxa have shown biological responses to global warming and a 'coherent fingerprint of climate change' is visible (Parmesan and Yohe 2003). For example, more than half of 35 studied European non-migratory butterfly species have extended their geographical distribution northwards (Parmesan *et al*. 1999), several Finnish bird species now breed in more northerly areas (Brommer 2004) and North Sea fish species have shifted their range northwards and to deeper water (Perry *et al*. 2005) because the geographic distribution of their 'climate envelopes' shifted. These climate envelopes describe the environmental conditions under which a species can exist under competition and predation. It has been predicted that the shifting of geographical distributions entrained by climate envelopes in combination with a changing climate will have dramatic negative effects on global biodiversity (Thomas *et al*. 2004; Williams *et al*. 2003).

Not only species distributions or assemblages (e.g. Genner *et al*. 2004; Holbrook *et al*. 1997) but also their phenology, the timing of events in the annual cycle, has frequently been affected by climate change. For example, various plant species flower or unfold their leaves earlier nowadays (Bradley *et al*. 1999; Menzel and Fabian 1999). In 21 out of 25 plant species in the Mediterranean region bud burst has advanced (Penuelas *et al*. 2002). Butterflies appear earlier in spring (Penuelas *et al*. 2002; Roy and Sparks 2000; Woiwood 1997) and caterpillar biomass peaks earlier (Visser *et al*. 1998). Amphibians have advanced their date of spawning or appearance in spring (Gibbs and Breisch 2001). Marmots in the Rocky Mountains appear earlier after hibernation (Inouye *et al*. 2000). Many migratory birds arriver earlier in their breeding areas (Ahola *et al*. 2004; Both and Visser 2001; Tryjanowski *et al*. 2002) and migrate earlier (Hüppop and Hüppop 2003). Many bird species also have advanced their laying dates (e.g. Brown *et al*. 1999; Crick *et al*. 1997; Dunn and Winkler 1999).

Although it has generally been found that climate change has led to an advancement of phenology there are differences between species and even populations. As reported by Crick *et al*. (1997), 20 British bird species have advanced their laying dates significantly but 44 have not not and one species has even delayed breeding significantly. While some differences between populations can be explained by regionally differing temperature increases (Both *et al*. 2004) others cannot (Visser *et al*. 2003). However, it is not *a priori* clear whether it is 'good' or 'bad' that populations or species show a response to climate change. Is it 'good' because they seem to be able to respond to changes in their environment or is it 'bad' because it is a sign that climate change is already affecting biological systems with possible negative effects (see Visser *et al*. 2004)?

In many bird species prey abundance is crucial for reproductive success because of the high food demands of fast growing offspring. For example, caterpillars form the main prey species for small woodland passerines during chick rearing (Naef-Daenzer *et al*. 2000; Perrins 1991). Many caterpillar species hatch from their eggs, grow and pupate within a short period in spring and consequently their biomass shows a distinct peak. The synchrony between this biomass peak and the chick rearing period is the main determinant of fledging success in great tits (*Parus major*) (van Balen 1973; van Noordwijk *et al*. 1995; Verboven *et al*. 2001). The same can also be true for marine ecosystems, the reproductive success of kittiwakes (*Rissa tridactyla*) and puffins (*Fratercula artica*) is strongly related to the timing and abundance of their fish prey species (Durant *et al*. 2003; Rindorf *et al*. 2000). Also in species where the young search for their food themselves, as in golden plovers (*Pluvialis apricaria*), reproductive success depends on the synchrony with the phenology of the food (Pearce-Higgins and Yalden 2004). These examples on breeding time illustrate that the phenological synchrony between species can be crucial for individual reproductive success and that a disruption of this synchrony caused by climate change could have negative effects. Consequently, when evaluating the effects of climate change on phenology it is not important whether a species responds or not but that the synchrony with its environment, in this example the phenology of its prey, is maintained.

As already mentioned, the synchrony between caterpillar biomass and chick rearing period is crucial for reproductive success in great tits. Higher spring temperatures cause caterpillars to hatch earlier (Visser and Holleman 2001) and grow faster (Topp and Kirsten 1991). In a Dutch long-term study population of great tits the phenology of caterpillar biomass has been monitored since 1985 and the date of maximum biomass has advanced by about two weeks since then. However, laying dates of great tits have only shifted on average by two days during the same period and the increasing mismatch has led to selection for earlier breeding in this population (Visser *et al*. 1998).

The lack of a response to selection on breeding time seems to be a common phenomenon in birds. In many bird species reproductive success declines during the breeding season (Merilä *et al*. 2001b; Perrins 1970). Since breeding time is heritable (Sheldon *et al*. 2003; van der Jeugd and McCleery 2002; van Noordwijk *et al*. 1981) the directional selection resulting from this seasonal decline in reproductive success should lead to a microevolutionary response. However, in most populations studied – sometimes for several decades – no or at best little evidence has been found that breeding time has advanced evolutionary. Indeed, breeding time in birds has become the classical example of evolutionary stasis (Merilä *et al*. 2001b). However, understanding whether and how populations and species respond to selection has now gained additional importance in the light of climate change. During my PhD project I explored several explanations why selection for earlier breeding has not led to a phenotypic response using great tits and their caterpillar prey as a model system.

Constraints preventing a phenotypic change

In great tits a single egg can weigh up to 10% of the female's body mass and producing eggs is costly in terms of energy, nutrients and future survival (Monaghan *et al*. 1998; Stevenson and Bryant 2000; Visser and Lessells 2001). There is evidence that the temperature increase in spring, which has led to the advancement of caterpillar phenology, has been more pronounced in the later part of spring (Fig. 1.2). This means that in the period when the birds have to produce their eggs environmental conditions have changed relatively little. Producing eggs under low temperatures and consequently low insect abundance may not be possible or incur excessive fitness costs. Hence, low temperatures during the egg production period may hamper an advancement of egg laying. In Chapter 2 I address whether an advancement of egg laying dates is constrained by resources by experimentally manipulating laying dates in two Dutch great tit populations (Hoge Veluwe and Oosterhout).

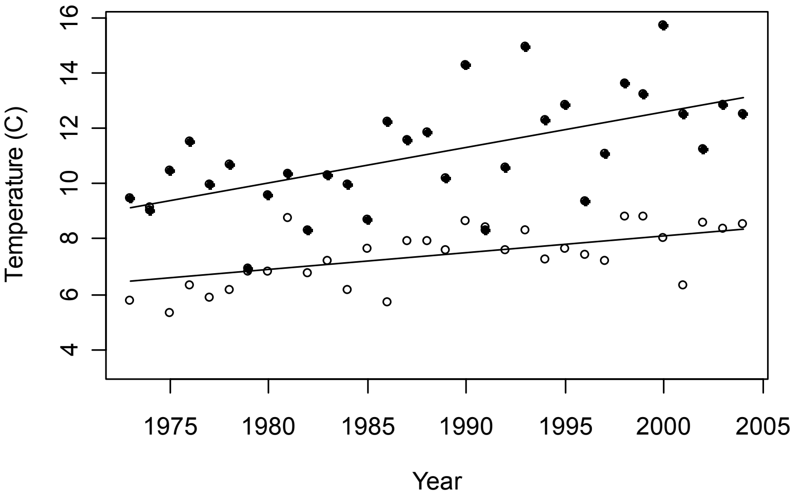

Fig. 1.2 Temperature in spring for the years 1973-2004. Annual mean temperatures during the prelaying period (open dots) and the laying period (filled dots) for Hoge Veluwe great tits increased significantly over this period ($b = 0.13 \pm 0.03$, $p < 0.001$). The temperature increase in the laying period was marginally non-significantly stronger than during the laying period ($p = 0.07$), which means that temperatures during egg production have changed less than later in spring. The prelaying period was defined as the six week period before the grand mean laying date in the Hoge Veluwe (24 April) and the laying period as the four week period after the grand mean laying date (see Visser *et al*. 2003). Temperature data are for the De Bilt weather station (Royal Dutch Meteorological Institute; KNMI) and were obtained from the KNMI's website.

Non-heritable condition biasing selection differentials

Phenotypic selection differentials may not necessarily give a representative picture of selection pressures. If fitness was not related causally to breeding time but a third, unmeasured, non-heritable trait that also affected breeding time, a correlation between fitness and breeding time could arise (Price *et al*. 1988) (Box 1). This has been described as selection on the environmental component of breeding time since females in good health or nutritional condition may be able to breed earlier in relation to their genes than females in bad condition.

Box 1 – *Selection on the environmental component*

The frequently observed seasonal decline in reproductive succes in birds (e.g. Brinkhof *et al*. 1993; Daan *et al*. 1990; Klomp 1970; Perrins 1970; Verhulst *et al*. 1995) leads to directional selection on breeding time but evidence for a microevolutionary response is lacking (Merilä *et al*. 2001b). A hypothesis originally proposed by Fisher (1958) and elaborated by Price *et al*. (1988) explains the lack of a response to selection on breeding time in birds by a correlation between breeding time and fitness caused by a third, unmeasured, environmental variable. Female birds may differ in their health, nutritional state or vigour; in short, in their general condition. Individuals in good condition are likely to reproduce more successfully than individuals in bad condition (e.g. Dyrcz *et al*. 2005; Moreno *et al*. 2002). It is also likely that females in good condition will be able to lay their eggs earlier (e.g. Hasselqvist *et al*. 2001) because producing eggs and incubating them is costly (e.g. Visser and Lessells 2001). Although breeding time is partly determined by a female's genes (Sheldon *et al*. 2003; van der Jeugd and McCleery 2002; van Noordwijk *et al*. 1981) females in good condition will breed on average earlier (assuming no correlation between the genetic component of laying date and condition). Since condition determines reproductive success this will lead to a non-causal correlation between breeding time and reproductive success (Fig. B1).

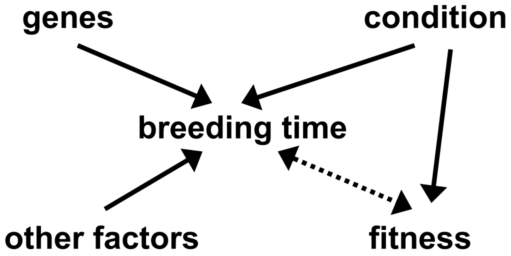

Fig. B1 Schematic illustration of selection on the environmental component. The date when a females starts egg laying, its breeding time, is determined by the female's genes and, in this example, condition. Condition thus determines the environmental component of breeding time. Because condition also determines the female's reproductive success a correlation between breeding time and reproductive success would arise (dashed, double-headed arrow). If condition, or in general the environmental component, remains unmeasured it would seem that fitness is causally related to breeding time and that breeding time should respond to selection, given it is heritable.

The environmental component, or environmental deviation, is the difference between the phenotype expected from an individual's genotype and the observed phenotype. Because condition affects both the environmental component of breeding time and reproductive success this hypothesis has also been described as 'selection on the environmental component', which is however an unlucky term because not the environmental component determines fitness but the third, unmeasured, non-heritable factor (condition in our example).

The seasonal decline of reproductive success has also been adressed in the framework of the 'date' versus the 'quality' hypothesis. While the 'date' hypothesis states that fitness is causally related to breeding time, the 'quality' hypothesis states that individual quality affects both breeding time and fitness thereby creating the observed seasonal decline in reproductive success, which is analogous to selection on the environmental component, as explained above. Many studies have tested these hypotheses by experimentally manipulating breeding time. Breeding time was manipulated, either by removing first clutches, inducing the females to re-nest (e.g. Hatchwell 1991; Verhulst and Tinbergen 1991; Verhulst *et al*. 1995), or by swapping clutches between early and late breeding pairs (e.g. Brinkhof *et al*. 1993; Verboven and Verhulst 1996). In general it was found that delayed females had reduced fitness. This argues against the 'quality' hypothesis and selection on the environmental component. There are however caveats with these kinds of experiments because delayed females either had to lay a complete replacement clutch (removal experiments) or had to incubate longer (swapping experiments). Egg laying and incubation are costly (Visser and Lessells 2001) and might impair the female's ability to successfully raise chicks. Consequently, more elegant tests to manipulate breeding time are desirable (*cf* Christians *et al*. 2001). The ideal experiment would be manipulate laying dates by manipulating the 'decision' of a female when to start egg laying. It has been experimentally shown that blue tits adjust their breeding time to the experienced timing of prey availability (Grieco *et al*. 2002). This learning ability (see also Chapter 2) could offer a means of delaying laying dates without changing the female's reproductive effort, and hence an elegant way to test whether reproductive success is related to breeding time or individual quality.

Carotenoid-based colouration is related to an individual's health or vigour (e.g. Dufva and Allander 1995; Faivre *et al*. 2003), used as a signal in sexual selection (Andersson 1994) and is to a large extent determined by the carotenoid availability in the environment (Olson and Owens 1998). In great tits both sexes display a bright yellow breast. Chapters 3 and 4 address the questions whether and how this carotenoid based colouration is related to the genetic and environmental components of breeding time and fitness and also discuss implications for mutual mate choice.

Counteracting selection on adult survival

Looking only at recruitment in relation to breeding time, as is usually done (e.g. Both and Visser 2001; Svensson 1997; van Noordwijk *et al*. 1995; Visser *et al*. 1998), might give an incomplete picture of selection pressures if breeding time would also affect adult survival. As already explained above, producing and incubating eggs is costly (Visser and Lessells 2001). It seems likely that conditions for egg production become more favourable as spring proceeds and hence that the costs for egg production and incubation decrease during the season. Such a trade-off between reproductive success and adult survival could result in no net selection on breeding time (Brown and Brown 1999). This explanation for a seemingly lacking response to selection is addressed in Chapter 6.

Overestimated heritability

Most heritability estimates are obtained by regressing the phenotype of the offspring against the phenotype of the parent(s). However, such heritability estimates can be upwardly biased by spatial or temporal auto-correlation (van der Jeugd and McCleery 2002) and this would then lead to an overestimation of the expected response. Phenotypes of parents and offspring could resemble each other not only because they share genes but also because they experience similar environmental conditions. In many species natal dispersal is limited with offspring settling in the vicinity of their parents. Additionally, offspring also experience similar environments to their parents bacuse parents and offspring breed in the same or subsequent years and environments tend to be temporally correlated. This effect of shared environments between relatives is strongest between parents and their offspring. So-called ‘animal models’ are an alternative to parent-offspring regression as a method to estimate heritability and allow the inclusion of observations of more distant relatives. As a result, they can reduce the influence of environmental correlation on heritability estimates (Kruuk 2004). In Chapter 6 this methodology is applied to test whether previously used regression methods have overestimated heritability and also the expected response to selection.

Quantifying the expected response to selection

Breeding time in birds shows large year-to-year variation. Annual mean laying dates in the Hoge Veluwe great tit study population can differ by two weeks from one year to the next. Such strong environmental 'noise' may obscure a small expected response to selection and its detection may become simply a matter of statistical power. However, the expected response has first to be quantified and doing this in populations with overlapping generations, possibly temporally varying selection pressures, and different or even counteracting selection via fecundity and adult survival is not straightforward. The solution to this is to predict the response to selection on a year-to-year basis and thereby to quantify the cumulative expected response, as done in Chapter 6. Again, an ‘animal model’ was used to separate the environmental and genetic component of breeding time and the cumulative response to selection quantified at the phenotypic and genetic level. Comparing selection differentials at the phenotypic and genetic level is the most direct way to test whether selection on the environmental component can explain a lacking response to selection (Rausher 1992; Stinchcombe *et al*. 2002).

Predicting the future

As part of the work of the Intergovernmental Panel on Climate Change (IPCC) several climate scenarios for the coming 100 years have been generated. These climate models have already been used to assess possible impacts of climate change (e.g. Halpin 1997; Iverson and Prasad 1998; Thomas *et al*. 2004; Williams *et al*. 2003). As already explained, in great tits, and other small insectivorous passerines, the synchrony between caterpillar biomass and chick rearing period has strong effects on reproductive success and hence possibly also on population dynamics. Both *et al*. (unpublished manuscript) argue that the disapperance of certain pied flycatcher (*Ficedula hypoleuca*) populations is related to increased mismatch between these two periods. Consequently, even small changes in the climate, which do not alter the geographic distribution of climate envelopes (*cf* Thomas *et al*. 2004), can have perceivable influences at population levels. To be able to fully assess the impacts of climate change under the various climate scenarios (Houghton *et al*. 2001) we thus need good predictive models for the phenology of the different trophic levels in the food chain (or web).

Most phenological models in vertebrates have been built by regressing the phenological event (annual mean egg laying date or date of first sighting) against a temperature average measured over a fixed period of the year (e.g. Beebee 1995; Perrins and McCleery 1989; Sparks and Carey 1995). These kinds of descriptive models have however two important drawbacks. First, the average temperature measured over a fixed period, e.g. 1 April to 31 May, can only be a proxy for the factor(s) causally influencing general conditions. It is highly unlikely that temperatures on 31 March or 1 June do not influence egg laying when those at 1 April or 31 May do, or that individuals actually ‘measure’ mean temperatures. Additionally, some individuals will have started egg laying before the end of the period used to predict their laying date and in certain years even the mean laying date can fall within this period. This method, using a proxy instead of the causal factor(s), can be successful in predicting phenology, as long as the relationship between the proxy and the unknown factor(s) remains stable. Unfortunately, this is unlikely to be true under climate change (*cf* Easterling *et al*. 1997; Luterbacher *et al*. 2004). This problem is well illustrated by laying dates of golden plovers. Laying dates for recent years are well predicted by April temperatures but when this model is applied to April temperatures predicted for 2070-2099, the resultant predicted layign dates are in March (Pearce-Higgins *et al*. 2005). The authors tried to circumvent this problem by constructing a model including February temperatures, but these had no explanatory value in the initial model. This obviously limits the use of these models for predicting the effect of climate change on phenology. Second, linear regression models predict only the mean phenology of a population, and not variation between individuals. One solution is to predict spread, measured for example by the standard deviation, from a linear regression of observed variation against a temperature measure. However, extrapolating outside the observed range of temperatures is likely to be misleading.

An alternative approach that adresses these problems is to analyse individual observations rather than annual means and including all temperatures up until the event. Proportional hazards models (Cox 1972) allow exactly this. They were originally developed for medical survival analysis but have already been applied to questions in biology, for example, the foraging behaviour of parasitoids (e.g. Gienapp 1999; Hemerik *et al*. 1993; Wajnberg *et al*. 2003). The models allow properties of the individual, such as age, experience or condition, to be included as fixed variables. Temperature or day length and even their interactions can be included as time-dependent variables. These time-dependent variables are the critical advantage of this type of model because they circumvent the problem described above of measuring temperatures over fixed periods.

In Chapter 7 a proportional hazards model is applied to laying dates of great tits. In Chapter 8 fitness consequences of mistimed breeding are analysed and a descriptive model for caterpillar phenology is developed. These two descriptive models are then used to predict sycnhrony between great tits and their caterpillar food supply under climate change scenarios.

Understanding whether and how populations and species respond to selection is important in evaluating the possible effects of climate change. There are however not many study species that allow such a detailed analysis. Predicting the expected response to selection requires careful quantification of selection acting on the trait under consideration and its genetic variation. Additionally, in order to forecast future evolutionary responses to by climate change we need to have indentified the ecological factors underlying selection. Long-term studies of hole-breeding passerines are an excellent system for this purpose because various traits are easily measured, their fitness consequences can be quantified, the underlying ecological factors are known and pedigrees are easily established. Apart from the obvious need to quantify the ecological consequences of climate change, the resultant directional selection also offers a unique possibility to study how species respond evolutionarily to changing conditions. Thus, by trying to explain why a Dutch great tit population does not seem to have responded to selection for earlier breeding caused by climate change this thesis hopefully adds to our general understanding of evolution in the wild.

Box 2 – *Phenotypic plasticity and reaction norms of breeding time*

Genotypes that express different phenotypes under different environmental conditions are referred to as being phenotypically plastic (Pigliucci 2001). The way in which the phenotypic value of a trait produced by a given genotype varies with the environment is referred to as the genotype's reaction norm (Woltereck 1909) and is generally described by the slope and the intercept of a (linear) regression of the observed phenotypes against the relevant environmental variable(s). The environmental variable is generally scaled in a way that the intercept is equal to the phenotype in the average environment. The slope describes how sensitively a genotypes reacts to the environmental variable. Genotypes may differ in both the intercept and slope of their reaction norms, and the steepness of the reaction norm affects the amount of environmental variation in the trait and consequently its heritability (Postma and van Noordwijk 2005). For a meaningful description of reaction norms it is important to identify the correct environmental variable(s), against which to regress the phenotypes. Furthermore, there is no reason to expect this environmental axis to be one-dimensional.

As expected from the year-to-year variation in food phenology resulting in selection on breeding time (e.g. van Noordwijk *et al*. 1995) breeding time in many bird species is a highly flexible trait (Chapter 5; Brommer *et al*. 2005). Although individuals are selected to breed at the optimal time many species cannot react directly to the phenology of their food but have to rely on indirect cues (Visser and Lambrechts 1999). In birds the duration of laying and incubation means that females must initiate breeding weeks in advance of the food peak to which they are selected to time their breeding. For example, in great tits optimal timing requires females to start egg laying when the caterpillars used to feed the chicks are only themselves just about to hatch and consequently cannot serve as a direct cue. Instead, species are selected to respond to other cues which reliably predict the relevant environmental conditions.

Temperatures in early spring are correlated with temperatures in late spring, which determine caterpillar phenology, and could thus be used by the birds as a cue for caterpillar phenology. Indeed, it has frequently been shown that spring temperatures affect timing of breeding in birds (e.g. Meijer *et al*. 1999; Perrins and McCleery 1989; van Balen 1973) and these temperatures may be used as the environmental variable to describe reaction norms.

However, the different temperature trends in early and late spring (Fig. 1.2) imply that the correlation between cues used by the birds and caterpillar phenology has changed. The same temperature in early spring predicts nowadays a different caterpillar phenology than 30 years ago. Furthermore, since the growing rate of ectotherms is non-linearly related to temperature (Schmidt-Nielsen 1983) the correlation between temperatures in early spring, used as a cue by the birds, and caterpillar phenology may be different for low and high temperatures. While low temperatures still predict the caterpillar peak more or less reliably, the predictive value of high temperatures may have changed. Whether this changed correlation between cues and caterpillar phenology would lead to selection on the intercept or on the slope of the reaction norm would depend on the range of experienced environments (Fig. B2).

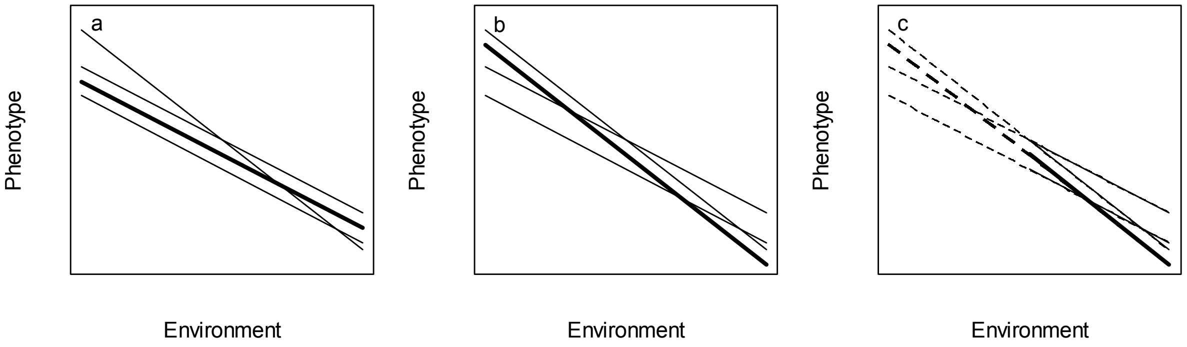

Fig. B2 Schematic illustration of selection on phenotypic plasticity under different environmental regimes. The expressed phenotypes (laying dates) are plotted against the environment (spring temperature). A reaction norm can be described by its elevation in the average environment (the intercept) and its slope. The thick line indicates the optimal reaction norm and the thin lines reaction norm of individuals. In any given environment the distance between the optimal reaction norm and the individual reaction norm represents its fitness (assuming a linear decrease with distance). In (a), representing a hypothetic situation in the past, both reaction norms with the shallower slope have the same fitness and the reaction norm with the steep slope a lower fitness when environments are normally distributed along the axis and the average environment in the middle of the indicated range. In (b) and (c) the optimal reaction norm has become steeper, because the relationship between caterpillar phenology and the cues used by the birds has changed. If the range of experienced environments has not changed (c) the reaction norm with the steepest slope has the highest fitness and the two other do not differ in fitness. Thus there will be only selection on slope. If the range of experienced environments has changed (b) (because cold springs do not occur anymore) the reaction norm with the steepest slope has the highest fitness but also the two shallow reaction norms differ in fitness. Since the one with the lower intercept has a higher fitness there is also selection on the intercept. However, for these two scenarios it was assumed that environments were normally distributed along with the average environment in the middle. If the distribution of environments is skewed and the average environment will shift selection will start acting on the intercept as well since the situation will resemble more scenario (c).

Chapter 2

Possible fitness consequences of experimentally advanced laying dates in great tits: differences between populations in different habitats

Phillip Gienapp and Marcel E. Visser

submitted to Functional Ecology

Summary

In birds early breeding individuals generally reproduce more successfully than late breeding individuals. This seasonal decline in reproductive success leads to directional selection on breeding time in birds. The frequently observed lack of response to this selection could be explained by resource constraints during the egg production period. We made use of the fact that *Parus* species can learn from experienced mismatch between their breeding time and nestling food availability to manipulate laying dates. Creating an artificial food peak in the previous year allowed us to manipulate laying dates without confounding effects of the manipulation in the year of interest and hence to look at possible negative fitness consequences of laying earlier. In two great tit populations we created an artificial food peak by food supplementation. In one population manipulated females advanced their laying dates in response to the artificial food peak while in the other population no response to the treatment was found. This difference could be caused by differences in resource availability in early spring between the two habitats. Low resource availability in early spring could also explain the lacking response to selection for earlier laying in one population caused by climate change.

Introduction

In many bird species early breeding individuals raise more offspring than late breeding individuals (Perrins 1970). This seasonal decline in reproductive success results in directional selection on breeding time. Since breeding time is a heritable trait (Sheldon *et al*. 2003; van der Jeugd and McCleery 2002; van Noordwijk *et al*. 1981) this directional selection should lead to an evolutionary change in breeding time. However, no such change has yet been reported and the lacking response to selection on breeding time has become the classical example for evolutionary stasis (Merilä *et al*. 2001b).

One possible explanation for this paradox is that an advancement of egg laying dates is constrained by resource availability at the time of egg production (Perrins 1970). A single egg can weigh up to 10% of the females body mass and is hence costly in terms of energy, nutrients and future fitness (Monaghan *et al*. 1998; Stevenson and Bryant 2000; Visser and Lessells 2001). Temperatures, and consequently food supply, increase during spring and there might thus be a date before which egg laying is not possible or too costly in terms of fitness. 'Income breeders' (Drent and Daan 1980; Meijer and Drent 1999), like small passerines, do not produce eggs from stored resources but have to acquire them at the time the egg is produced. For these species the benefits of laying earlier may be outweighed by the costs of acquiring enough resources, e.g. increased predation risk, or by a physiological cost of hard work (von Schantz *et al*. 1999). It is also possible that essential resources like certain proteins or calcium are simply not available in early spring because the certain prey species (insects, snails) have not appeared yet.

If breeding time would be constrained by resources this could explain why a Dutch great tit population did not advance their laying dates although caterpillars, the main food supply for their chicks, reach their maximum biomass nowadays earlier (Visser *et al*. 1998). Rising spring temperatures cause the caterpillars to hatch earlier (Visser and Holleman 2001) which has resulted in an advancement of the date when their maximum biomass is reached by about two weeks over the last 15 years (Chapter 8). Since caterpillars form the main food supply for their chicks (Naef-Daenzer *et al*. 2000; Perrins 1991) the birds should follow the changes in caterpillar phenology but annual mean laying dates have advanced by only about 3 days over the same period. The alternative explanation suggested by Visser *et al*. (2004; 1998) for the mismatch between breeding time of the birds and the caterpillar phenology is that the birds are currently responding to the wrong cues to time their breeding. The caterpillars hatch at the beginning of May to feed on the fresh leaves of their host trees. The great tits, however, have to initiate egg laying about a month earlier to match the nestling period with the period when caterpillars are most abundant. The birds can therefore not react directly to the appearance of the caterpillars but have to use other cues, which used to indicate caterpillar phenology reliably. Nowadays this relationship may, however, be disrupted and the cues no longer reliably predict caterpillar phenology.

The ideal experiment to test these two hypotheses is to induce earlier laying than normally found and to look at the fitness consequences. Experiments to manipulate laying date have been carried out frequently: in food supplementation experiments an advancement of laying dates was generally found (Nager *et al*. 1997 and refs therein). However, these experiments are unsuitable to evaluate fitness consequences as in these cases the birds were given additional resources and hence the potential fitness costs of early egg production will be underestimated. In this study we used an experimental approach that did not suffer from this problem.

Great tits use the synchrony between nestling period and food supply that they experience in one year to 'fine-tune' their timing of breeding in the following year (Nager and van Noordwijk 1995). Individuals that bred earlier than optimal delay breeding in the following year, while individuals that breed too late advance their breeding. This ability to learn has also been shown experimentally in blue tits (Grieco *et al*. 2002). Individuals given an artificial food peak late during the nestling period delayed their breeding in the following year. This approach offers an elegant way to manipulate breeding time and to evaluate fitness consequences because the birds are not given any additional resources in the year when their laying dates are manipulated. We used this approach to manipulate laying dates in a Dutch great tit population (Hoge Veluwe), for which selection for earlier laying has been reported (Visser *et al*. 1998), and in another nearby population (Oosterhout) and measured possible fitness costs of advancing laying dates thereby testing whether early breeding is constrained by resource availability.

Methods

Study species and areas

Our study was carried out in 2001-2003 using two Dutch populations of great tits. Great tits (*Parus major* L.) are small (about 20 g), insectivorous, hole-nesting passerines that readily accept nest boxes. Both populations have been followed continuously since 1955 (Hoge Veluwe) and 1956 (Oosterhout). One study area, Oosterhout, is a relatively small woodland of 11 ha, in which 120 nest boxes are placed. The woodland is situated near the river Waal on a river floodplain with rich clay soils. The vegetation consists mainly of large pedunculate oaks (*Quercus robur*), maple (*Acer pseudoplatanus*) and ash (*Fraxinus excelsior*) with rich undergrowth of cherry laurel (*Prunus laurocerasus*), rowan (*Sorbus aucuparia*) and common elderberry (*Sambucucs nigra*). The other study area, the Hoge Veluwe, is part of a larger wooded area on poor sandy soils. The vegetation is a mixture of pedunculate oak (*Quercus robur*), red oak (*Q. rubra*), birch (*Betula pendula*) and scots pine (*Pinus sylvestris*) with sparse undergrowth and interspersed with open heath land. The study area covers 171 ha with about 400 nest boxes.

Standard fieldwork

Nest boxes were checked weekly to determine clutch size and laying date, i.e., the date when the first egg of a clutch was laid. Since successive eggs are generally laid at one-day intervals and clutch size usually exceeds six eggs, weekly checks give accurate laying dates. All nestlings were ringed with aluminium rings at the age of seven days. When the chicks are 15 days old, all chicks were weighed. Adults were caught during the chick feeding period, identified by colour and aluminium rings or ringed if not ringed yet. Additionally, females in the Hoge Veluwe study area were caught in winter when roosting in nest boxes at night and equipped with small (0.5 g) passive integrated transponders (PIT) (Trovan Inc.) being glued to the colour rings. These allowed identification of individuals during the egg laying period when roosting at night in nest boxes by reading the PIT through the nest box. This was necessary to confirm the identity of the food supplemented individuals.

Caterpillars are the main food source for great tit nestlings in these two study areas and relative timing to caterpillar phenology is the main determinant of reproductive success (Chapter 7, van Noordwijk *et al*. 1995; Verboven *et al*. 2001). To determine caterpillar phenology caterpillar droppings (frass) were collected every 3-4 days under oak trees at six locations in Oosterhout and seven locations on the Hoge Veluwe, dried and weighed. Caterpillar biomass was calculated from this dry weight using the equation of Tinbergen and Dietz (1994). The date when the calculated biomass was highest was defined as the caterpillar peakdate and used as a reference point for caterpillar biomass. As a measure for relative timing (synchrony) the difference between the day the chicks were 9 days old and the 'caterpillar peak' date was calculated. When great tit chicks are about 9 days old they grow fastest and food demands are highest (Gebhardt-Henrich 1990; Keller and van Noordwijk 1994).

Food supplementation

To manipulate the birds' experience and thereby its breeding time in following year we provided supplementary food that mimicked an artificial food peak (*cf* Grieco *et al.* 2002) in 2001 in Oosterhout and 2002 in the Hoge Veluwe. Our aim was to experimentally advance egg-laying dates by manipulating the birds' experience in the first year of the experiment and to measure the response and possible consequences in the following year.

As soon as it was discovered that egg-laying had started in a next box during a weekly routine check the nest box was randomly assigned to one of three groups. In one group of birds (manipulated) we wanted to create the perception that they were breeding too late and hence the artificial food peak was earlier than the chick-feeding period. The treatment started on the day at which the 8th egg was laid or at the onset of incubation if clutch size was smaller than eight (4 out of 21 clutches). The food (mealworms, *Tenebrio molitor*) was offered in small trays (approx. 5 x 3 x 4 cm) inside the nest box to assure that no other birds than the manipulated pairs had access to the food. The amount offered rose stepwise from 3 g on day 1 to 20 g on day 7, remained constant until day 9 and then decreased stepwise to 3 g on day 15. Depending on the clutch size the treatment therefore stopped about the time the chicks hatched since incubation normally lasts 13 days. Nest boxes were visited daily; the remaining food was removed and weighed to the nearest 0.1 g and the 'scheduled' amount of fresh food given. Food supplemented birds might be in better in condition and breed earlier in the second year due to carry-over effects and not to their manipulated experience. To control for such effects we provided a second group (control-supplemented) with the same amount of food following the same schedule but starting on the day the chicks hatched. Consequently food supplementation stopped shortly before the chicks fledged. Since this group received the artificial food peak during the chick-feeding period we did not expect any effect of experienced mistiming on breeding time in second year. A third group of birds (unmanipulated) did not receive any food to serve as a natural reference group. Due to small number of breeding pairs this third group had to be omitted in the Oosterhout experiment. In the 2nd year of the experiment, however, a number of first-year breeding females (recruits or immigrants) entered the breeding population and we compared laying dates of the manipulated females to their laying dates. Age affects laying date (e.g. Robertson and Rendell 2001; Wheelwright and Schultz 1994). To correct for this we calculated the long-term difference in laying dates between first-year breeders and older females for the Oosterhout study area (1.46 +/- 0.38 days, $n = 41$ years) and compared this difference with the difference between manipulated and first-year breeding females respectively control-supplemented and first-year breeding females.

Data analysis

Laying dates depend strongly on annual environmental factors, like temperature (e.g. Chapter 7; Perrins and McCleery 1989; van Balen 1973). Since we were interested in individual changes not caused by the environment we corrected individual changes in laying date for the difference in mean laying dates between the first and second year of the experiment ($\Delta ld = ld_{\text{2nd year}} - ld_{\text{1st year}} + \textit{difference in population means}$). All manipulated females were excluded from the calculation of the mean laying date in the second but not in the first year.

Appropriate statistical models were selected using stepwise backward selection procedures starting with a full model including all possible interactions. Non-significant test-statistics for interactions were not reported. All statistical models were GLMs using either normal or binomial error structure and run using R 2.0.1 (R Development Core Team 2004; Venables and Ripley 2002). Means and their standard errors are reported throughout if not stated otherwise. Dates (e.g. laying or hatching) are reported as April days, with 1 being 1st April and 31 being 1st May.

Results

Treatment

In Oosterhout (2001) 28 breeding pairs were food supplemented (14 manipulated and 14 control-supplemented pairs). In the Hoge Veluwe (2002) 42 pairs were food supplemented (21 manipulated and 21 control-supplemented pairs) and 41 pairs were not food supplemented at all (unmanipulated). Laying dates in the first year of the experiment did not differ between treatment groups ($F_{2,107} = 0.85$, $p = 0.43$, corrected for year/area differences) (Tab. 2.1) and this pattern did not change when only females that survived until the second year were analysed ($F_{2,57} = 0.05$, $p = 0.95$). The number of fledged chicks and mean chick weight did not differ between treatment groups (no. fledged chicks: $F_{2,105} = 0.69$. $p = 0.50$; mean chick weight: $F_{2,96} = 0.19$, $p = 0.83$, both corrected for year/area differences) (Tab. 2.1). Again, the same was true when the analysis was restricted to broods of females that survived until the second year for number fledged chicks ($F_{2,57} = 0.21$, $p = 0.81$) and mean chick weight ($F_{2,48} = 0.07$, $p = 0.94$).

When mealworms were offered during the chick feeding period pairs consumed significantly more than when offered during incubation (treatment: $F_{1,67} = 39.2$, $p < 0.001$; area: $F_{1,66} = 1.06$, $p = 0.31$).

Tab. 2.1 Reproductive parameters (mean ± se) for the different experimental groups in the first year of the experiment.

	Oosterhout		Hoge Veluwe		
	manipulated	control-suppl.	manipulated	control-suppl.	unmanip.
Laying date	25.2 ± 1.39	27.5 ± 0.91	22.6 ± 0.82	21.6 ± 0.74	20.6 ± 1.12
No. fledged chicks	6.9 ± 0.71	7.5 ± 0.72	8.1 ± 0.53	7.6 ± 0.73	7.0 ± 0.47
Mean chick weight	17.3 ± 0.27	16.5 ± 0.48	17.3 ± 0.35	17.6 ± 0.25	17.6 ± 0.16

Response to treatment

Survival from the first to the second year of the experiment did not differ between treatment groups or areas (treatment: $X^2 = 0.53$, df = 2, $p = 0.77$; area: $X^2 = 2.21$, $p = 0.14$). Changes in laying dates from the first to second year of the experiment depended on the experienced synchrony with the foodpeak ($F_{1,53} = 45.8$, $p < 0.001$). Females having bred earlier than the natural food peak delayed laying in the second year and vice versa (Fig. 2.1). This response to experienced mistiming was stronger in Oosterhout than in the Hoge Veluwe (area*synchrony: $F_{1,51} = 4.39$, $p = 0.04$). In Oosterhout there was an additional treatment effect (Fig. 2.1). After correcting for the experienced synchrony manipulated females bred on average 5.6 days later than control-supplemented ($F_{1,9} = 5.139$, $p = 0.0496$) while there was no difference between treatment groups in the Hoge Veluwe (treatment: $F_{2,43} = 0.12$, $p = 0.89$) (Fig. 2.2). The difference in the response of the treatment groups between areas was significant (treatment*area: $F_{1,51} = 4.65$, $p = 0.036$).

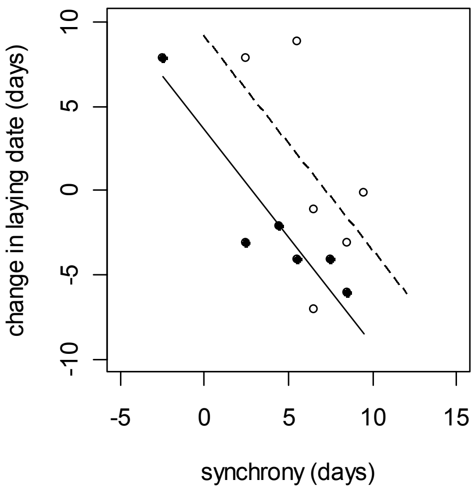

Fig. 2.1 Oosterhout: change in laying date from first to second year versus experienced synchrony with the caterpillar food peak in the first year (closed dots and solid line: manipulated females; open dots and dashed line: control-supplemented females). The change in laying date was significantly affected by experienced synchrony ($F_{1,9} = 11.0$, $p = 0.009$) and experimental treatment ($F_{1,9} = 5.2$, $p = 0.049$). Manipulated females bred on average 5.6 days earlier than control-supplemented females.

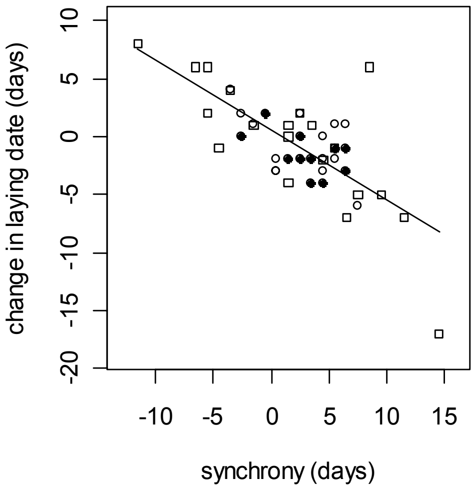

Fig. 2.2 Hoge Veluwe: change in laying date from first to second year versus experienced synchrony with the caterpillar food peak in the first year (closed dots: manipulated females, open dots: control-supplemented females, open squares: unmanipulated females). The change in laying date was significantly affected by experienced synchrony ($F_{1,43} = 44.9$, $p < 0.001$, solid line) but not experimental treatment ($F_{2,43} = 0.12$, $p = 0.89$).

However, we cannot conclude from the experimental treatment that manipulated females advanced their laying date, as expected, because the same pattern would arise if control-supplemented females delayed laying. Due to the small breeding population in Oosterhout it was not possible to have a third group of unmanipulated females, but first-year breeders can serve as an unmanipulated 'outgroup'. Additionally to the surviving experimental females, 10 of these first-year females, either immigrants or recruits, bred in the study area in the second year. The mean laying date of these females was 17.0 +/- 1.68, compared to 10.3 +/- 0.76 for manipulated and 16.0 +/- 2.00 for control-supplemented females (Fig. 2.3). While the difference between first-year breeding and control-supplemented females (1.0 ± 2.03) was not significantly different from the long-term difference between first-year breeding and older females (see Methods) ($t = 0.976$, df = 56, $p = 0.83$) the difference between first-year breeding and manipulated females (6.7 ± 0.85) differed significantly from the long-term difference ($t = -5.736$, df = 56, $p < 0.001$). This indicates that manipulated females advanced their laying date rather than that the control females delayed.

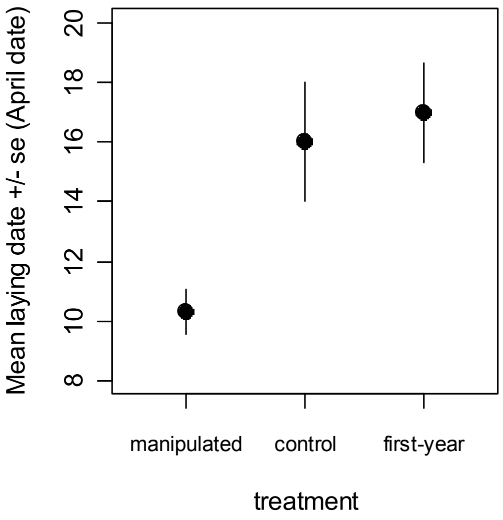

Fig. 2.3 Oosterhout experiment: laying dates of manipulated, control-supplemented and first-year breeding females in 2002 (second year). The difference between control-supplemented females and first-year breeders is not significantly different from the expected difference between older and first-year breeders, while manipulated females bred significantliy earlier than expected for older females (see text for details).

Fitness consequenes

Although in Oosterhout laying dates between manipulated and control-supplemented females differed by about 5 days reproductive success did not differ between groups (mean chick weight: $F_{1,10} = 1.66$, $p = 0.22$; no. fledged chicks: $F_{1,10} = 0.54$, $p = 0.48$) (Tab. 2.2). Survival from the second year of the experiment to the following year did not differ between manipulated and control-supplemented females ($X^2 = 0.028$, df = 1, $p = 0.87$).

Since laying dates in the second year in the Hoge Veluwe did not differ between groups we would not expect any differences in fitness between the groups that were related to breeding time. Reproductive success did not differ between groups (mean chick weight: $F_{2,32} = 1.87$, $p = 0.17$; no. fledged chicks: $F_{2,44} = 0.71$, $p = 0.50$) (Tab. 2.2).

Survival from the second year of the experiment to the following year did not differ between groups (manipulated: 2, control-supplemented: 2, unmanipulated: 1; $X^2 = 1.36$, df = 2, $p = 0.51$, power = 0.52), however, overall survival probability was unusually low (= 0.10) and differed significantly from survival probability from first to second year (= 0.59) ($X^2 = 14.4$, df = 1, $p < 0.001$).

Tab. 2.2 Reproductive success (mean ± se) for the different experimental groups in the second year of the experiment.

	Oosterhout		Hoge Veluwe		
	manipulated	control-suppl.	manipulated	control-suppl.	unmanip.
No. fledged chicks	8.8 ± 0.49	8.0 ± 1.02	7.8 ± 0.95	6.4 ± 1.07	6.3 ± 0.73
Mean chick weight	18.2 ± 0.21	17.6 ± 0.41	16.0 ± 0.54	15.5 ± 0.56	16.9 ± 0.41

Discussion

Tit species can learn from previously experienced synchrony with their nestling food supply to adjust laying dates in the following year (Grieco *et al*. 2002; Nager and van Noordwijk 1995). We made use of this learning ability and manipulated breeding time in two populations of great tits to look at fitness consequences of advanced laying dates. In one study area (Oosterhout) manipulated females bred significantly earlier in the second year of the experiment than control-supplemented females (Fig. 2.1) and thus responded to the treatment as expected. It can be ruled out that this is caused by an improved condition of the females due to the food supplementation in the first year because the control-supplemented females received the same food supplementation. Control-supplemented females also consumed significantly more of the given food and if there were a carry-over effect of condition they should have advanced laying dates. Since the manipulated birds bred earlier than normal and this was achieved without any manipulation to the females in the second year this allowed us to look at possible fitness consequences of advancing laying dates. Possible fitness costs of advancing laying dates were reduced adult survival or reduced adult care, which could result in worse condition of the chicks. Unfortunately, it was not possible to measure differences in recruitment between treatments because overall recruitment was too low to allow statistical testing; in total only seven fledglings from 2002 recruited in 2003. Neither adult survival nor chick condition, measured as fledging weight, differed between manipulated and control-supplemented females. However, sample sizes were small and consequently test power low. We thus cannot conclude that there is no cost of advancing laying dates.

Interestingly, we found different results when repeating the experiment following exactly the same protocol in a different long-term study population (Hoge Veluwe). In this population experimental treatment did not affect change in laying dates from the first to the second year (Fig. 2.2). Since treatment did not affect laying date no effect of treatment of reproductive success or adult survival was expected and found. Although sample sizes were initially twice as large as in the Oosterhout experiment it would have been impossible to detect any differences in adult survival because general survival probability from the second to the following year was unusually low, only five out of 49 females bred in 2004, probably caused by low beech crop (*cf* Perdeck *et al*. 2000). This is also the reason why we did not report any results from two other experiment carried out in 2003. In this year we food supplemented 40 breeding pairs in the Hoge Veluwe and 30 breeding pairs in Warnsborn, another long-term study area. From these 70 food supplemented females only seven survived to the second year in both areas, too few to test for a treatment effect on laying date.

However, the intriguing result that females in the Hoge Veluwe did not advance their laying date, while females in Oosterhout were able to, remains. There was no difference in the amount of mealworms consumed between the areas (treatment*area: $F_{1,65} = 0.06$, $p = 0.81$). The lacking response to the experimental treatment can hence not be explained by the Hoge Veluwe birds ignoring the supplemented food and we should look for other ecological factors possibly explaining this difference.

The two areas differ substantially in their vegetation type, Oosterhout is dominated by large, old oak trees with abundant undergrowth on rich clay soils while the Hoge Veluwe is a mixed oak-pine forest with much less undergrowth and patches of heath on poor sandy soils. This difference in vegetation may translate to difference in insect abundance and/or phenology. In rich, mature woods, like Oosterhout, shrubs and herbs have to make use of the sunlight before it is blocked by the closing canopy and flower and unfold their leaves therefore earlier than the trees. In such rich woods vegetation growth starts thus earlier than in poor woods with less undergrowth and a less closed canopy. Phenology of many insect species is linked to vegetation phenology and this could lead to an earlier availability of insects, a supposed major protein source for egg production in many bird species, in Oosterhout than in the Hoge Veluwe.

Calcium is another important resource during egg production. Small birds cannot rely on stored calcium resources in the bones and have to fulfill the needs for eggshell production on a daily basis (Perrins 1996). Snails form the main supply of this resource. Snail abundance is probably higher in Oosterhout because clay is a more calcareous soil than the sandy soils in the Hoge Veluwe (but see Graveland and Drent 1997). A higher abundance may also be related to an earlier phenology simply because the chance that some individuals appear earlier is higher when abundance is higher. Unfortunately, no data on insect or snail phenology and abundance in early spring are available and it is also not entirely clear which food sources are actually used by great tits during egg production. Results from a radiotracking study of great tits in the Hoge Veluwe indicated that females mainly foraged in coniferous trees during the pre-laying period (own data) but clearly more data on prey choice and its abundance in early spring are needed.

Since producing eggs needs energy, protein and calcium and availability of these resources likely differs between the two areas only the Oosterhout birds were able to advance laying dates in response to our experiment. The manipulated females in the Hoge Veluwe may have refrained from advancing laying dates because either the necessary resources were not available at all or the cost of obtaining them in order to lay earlier would have outweighed the fitness benefits of earlier laying. If this was true it could explain the lacking response to selection in this population. Due to warmer springs the peak date of caterpillar biomass has advanced in the Hoge Veluwe by about two weeks over the last 15 years (Chapter 8) leading to selection for earlier breeding (Visser *et al*. 1998). However, the birds' laying dates have not changed over the same period and two hypotheses have to proposed to explain this (Visser *et al*. 1998). First, the birds are currently responding to the wrong cues to time their breeding. Second, an advancement of laying dates is hampered by low resource availability in early spring and a related high cost of laying early. Our results strongly hint at the latter hypothesis being true. This means that we should not expect an evolutionary response to the observed selection on breeding time because the benefit of laying earlier is outweighed by a cost of laying earlier. This also means that the currently observed mistiming of the great tits may increase which could eventually lead to a population decline as has been found in pied flycatchers in the same area (Both *et al*., unpublished manuscript).

Acknowledgements

J. H. van Balen kept the long-term study on the Hoge Veluwe and Oosterhout going for many years, J. Visser carefully managed the database. R. Raat helped with the food supplementation and L. Holleman and H. Bouwmeester with standard fieldwork. The ‘Stichting Het Nationale Park De Hoge Veluwe’ and Baroness Van Boetzelaer Van Oosterhout kindly allowed us to conduct our research on their property. K. Lessells made helpful comments on the manuscript.

Chapter 3

The relationship between carotenoid colouration and the genetic and environmental components of breeding time in great tits

Phillip Gienapp and Erik Postma

Summary

A seasonal decline in reproductive success is a commonly observed phenomenon in birds. However, to date no microevolutionary response to the resulting directional selection has been observed. One possible explanation is that the correlation between breeding time and reproductive success is caused by a third unmeasured non-heritable factor that is related to quality and which affects breeding time and reproductive success independently (Price *et al*. 1988). Carotenoid colouration provides a measure of quality since relationships with nutritional state, health or survival have been found. To investigate whether condition affects the environmentally determined part of breeding time in two Dutch great tit (*Parus major*) populations we, first, measured carotenoid colouration, and second, we separated the genetic and environmental component of breeding time using an 'animal model'. We found a relationship between carotenoid colouration and the environmental component of laying date as well as the genetic component. We present explanations for these somewhat inconclusive results.

Introduction

In many bird species a seasonal decline in reproductive success is a common phenomenon (Perrins 1970). Since breeding time has been found to be heritable in a range of populations and species (e.g. Sheldon *et al*. 2003; van der Jeugd and McCleery 2002; van Noordwijk *et al*. 1981) the directional selection for breeding earlier that results from the seasonal decline in reproductive success should lead to an evolutionary advancement in breeding time. Yet, to date evidence for an evolutionary change in breeding time has been weak and thereby breeding time has become the classical example of evolutionary stasis (Merilä *et al*. 2001b). One possible explanation for a lacking response to selection is that selection acts on the environmental component of the trait (Fisher 1958; Price *et al*. 1988).

Like many other traits, laying dates are partly determined by the females' genes (the genetic component) and partly by non-heritable, environmental factors (the environmental component). These environmental factors may, for example, be a female's general condition or nutritional state. Since egg production is costly in terms of nutrients and energy (e.g. Monaghan *et al*. 1998; Stevenson and Bryant 2000) individuals in good condition or good nutritional state are able to lay earlier than females in a bad condition. If females in good condition can also raise many offspring because they are in good condition, rather than because they breed early, this would give rise to a correlation between reproductive success and breeding time. If this environmental component, condition in our example, remained unmeasured one may conclude that the correlation between breeding time and reproductive success was causal and thus that breeding time was under directional selection. However, since reproductive success is in fact related to condition, i.e. the environmental component, this was described as selection on the environmental component (Price *et al*. 1988).

Yellow to red carotenoid based colouration is assumed to a condition dependent signal of individual quality in general (Andersson 1994), and especially of immune competence (Hamilton and Zuk 1982). Individuals showing brighter carotenoid colouration have found to be of better health (Dufva and Allander 1995; Hõrak *et al*. 2004; Peters *et al*. 2004), have a higher survival probability (Hill 1991b; Hõrak *et al*. 2001; Saks *et al*. 2003), provide more care for the young (Hill 1991; Senar *et al*. 2002), are relatively heavier (Johnsen *et al*. 2003), are in better nutritional condition (Hill and Montgomerie 1994), and are preferred as mates (Amundsen and Forsgren 2001; Hill 1991; Houde and Endler 1990; Milinski and Bakker 1990). Several experimental studies also showed a causal relationship between carotenoid intake and immune response (Blount *et al*. 2003; McGraw and Ardia 2003; McGraw and Ardia 2004). Besides their function in signaling, carotenoids play an important role in several physiological mechanisms like antioxidant scavenging or immune function (Møller *et al*. 2000). Because carotenoids are produced by plants and cannot be synthesized by animals they have to be acquired via the food. However, carotenoids are relatively scarce in the food of most animal species (Goodwin 1980). Individuals infected by parasites or diseases need more carotenoids for immune response and have to reallocate carotenoids from signalling to this function. E.g., male blackbirds (*Turdus merula*) immunised with sheep red blood cells had less yellow bills (Faivre *et al*. 2003). The important physiological function of carotenoids together with their scarcity result in a trade-off between physiological needs and signaling, and makes carotenoid colouration an honest signal (Olson and Owens 1998). Consequently, if fitness is related to non-heritable condition but not to breeding time per se and condition also affects breeding time a correlation between carotenoid colouration and the environmental component of laying date should be expected. However, no relationship between carotenoid colouration and the genetic component.

Breeding time can be separated into the genetic and environmental components by predicting breeding values for breeding time. Regressing the phenotypic values against breeding values yields then the environmental component or environmental deviation. Recently, mixed model methodology has been used for the estimation of genetic parameters in wild populations (see Kruuk 2004 for a review). By fitting a so-called animal model we can obtain estimates of both genetic and environmental variance componentts, and we can separate an individual’s phenotype into its genetic and environmental components.

In this study we investigated the relationships between carotenoid colouration, breeding time and its genetic and environmental component, and fitness in two Dutch great tit populations and tested whether condition indicated by carotenoid colouration is more closely related to the environmental or the genetic component of breeding time.

Methods

Standard fieldwork and study areas

Our study was carried out in two Dutch long-term study populations of nestbox breeding great tits (*Parus major*). In both study areas (Hoge Veluwe and the island of Vlieland) about 400 nestboxes were put up in the 1950’s. The study area on Vlieland is distributed over five separate woodlands, clustered into to two subareas, the 'West' and the 'East'. However, the breeding population in the 'West’ is considerably smaller and the rate of immigration higher (Postma and van Noordwijk 2005). Therefore we restricted our analysis to the Hoge Veluwe and the Eastern sub-population on Vlieland, hereafter Vlieland.

The vegetation on Vlieland consists mainly of mixed forest on poor sandy soils but also purely coniferous (*Pinus sylvetris*) and decidous stands (*Quercus robur*) are found. The Hoge Veluwe study area is part of a larger wooded area also on poor sandy soils. The vegetation is a mixture of pedunculate oak (*Quercus robur*), red oak (*Q. rubra*), birch (*Betula pendula*) and scots pine (*Pinus sylvestris*) with sparse undergrowth and interspersed with open heath land. Nest boxes were checked weekly to determine clutch size and laying date, i.e., the date when the first egg of a clutch was laid. Since successive eggs are laid at one-day intervals and clutch size usually exceeds six eggs weekly checks give accurate laying dates. All nestlings were ringed with uniquely numbered aluminium rings when they were seven days old. Shortly before fledging, when the chicks are 15 days old, all chicks were weighed and their tarsus length measured. All adults were caught during chick feeding using trap doors inside the nest box and identified on the basis of their aluminium and colour rings or, if not yet ringed, ringed with aluminium and colour rings. At this occasion also their body mass, tarsus length and length of the third primary, hereafter wing length, were measured. Residual body mass was defined as the residual of a regression of body mass against tarsus length, sex, area and the interaction between sex and area (all significant at the 0.05-level).

Colour measurements

In 2004 we collected four yellow breast feathers from each individual taken from a fixed position on the breast when catching the adults for standard field work. The feathers were stored in light tight paper bags in a dry and cool place. For colour measurements four feathers were carefully arranged on top of each other and mounted on black velvet (*cf* Hõrak *et al*. 2001; Saks *et al*. 2003). Five replicate spectrum measurements were taken using an USB2000 Spectrometer with a DH2000 deuterium-halogen lamp as light source (both Avantes) illuminating and measuring perpendicular to the feathers. From the obtained spectra we calculated ‘carotenoid chroma’ as the difference between the reflectance at 450 nm and 600 nm divided by the reflectance at 600 nm (R_{600}-R_{450}/R_{600}) (*cf* Johnsen *et al*. 2003). The yellow to red carotenoid based colouration is produced by the absorbance of blue wavelengths, with a maximum absorbance at 450 nm. The raw data obtained from spectrometer software (Avasoft 5.1 Basic, Avantes) were processed, the chroma calculated and averaged over the five replicates using R 1.9.0 (R Development Core Team 2004).

Quantitative genetic analysis

An animal model approach was used to separate genetic and environmental components of individual laying dates (Kruuk 2004; Lynch and Walsh 1998). These linear mixed models use restricted maximum-likelihood (REML) methodology to estimate genetic and non-genetic variance components, and to simultaneously estimate fixed effects, and predict additive genetic and other random effects. They are well suited for the quantification of genetic parameters of natural populations since they use all available information in the pedigree and are therefore able to accommodate, for example, inbreeding and selection, which are all likely to occur in wild populations. For the quantitative genetic analysis data from the long term data base were used, covering the years 1959 to 2004 for Vlieland and 1957 to 2004 for Hoge Veluwe.

Laying dates show strong year-to-year variation and are influenced by the age (or experience) of the female. To correct for these effects, year and age (first-year breeder or older) were included as fixed effects in the model. To accommodate breeding events of individuals in multiple years, individual identity was fitted as a random effect to account for any permanent environment effects. To summarize, our model separated the remaining phenotypic variance V_P (after taking out the effect of the fixed effects of year and age) into the additive genetic variance V_A, the variance attributed to permanent environment effects V_{PE} and the residual variance V_R. Absolute variance components and ratios with their standard errors were estimated using the VCE package (Groeneveld 1995). Best linear unbiased predictors (BLUPs) for individual breeding values (PBVs) were obtained using the PEST software package (Groeneveld and Kovac 1990; Groeneveld *et al*. 1992).

The pedigree was established by assuming that the parents raising a brood are the biological parents. In some years brood size manipulation experiments were carried out. When brood size was increased by adding eggs, all chicks were assigned unknown parents because it was impossible to track which chick hatched from an added egg. When chicks were cross-fostered or brood size increased by adding chicks, all chicks were individually marked and could thus be assigned the putative genetic parents. Extra-pair paternity or brood parasitism can adulterate the assigned paternity and/or maternity. Although the level of extra-pair paternity in the study population is unknown, it is usually low in great tits (5-9%) (Krokene *et al.* 1998; Lubjuhn *et al.* 1999; Strohbach *et al.* 1998; Verboven and Mateman 1997). Furthermore, no incident of egg-dumping or brood parasitism was ever observed in our study population.

The environmental component of a laying date, or the environmental deviation, was defined as the residual of a regression of predicted breeding value versus phenotype. Phenotypes were corrected for year-to-year differences to allow a combined analysis of all years. The regressions were calculated separatedly for the two areas.

Results

Sex, age and area differerences

We found significant differences in carotenoid chroma between sexes (males: 0.50 ± 0.01; females: 0.42 ± 0.01; $F_{1,255} = 9.8$, $p = 0.002$), age classes (young: 0.46 ± 0.01; old: 0.49 ± 0.01; $F_{1,255} = 16.0$, $p < 0.001$) and also study areas (Hoge Veluwe: 0.53 ± 0.01; Vlieland: 0.44 ± 0.01; $F_{1,255} = 86.5$, $p < 0.001$). Females and young individuals, born the previous year, had significantly a lower carotenoid chroma. Birds in the Hoge Veluwe had a higher carotenoid chroma than birds on Vlieland. Carotenoid chroma was not related to neither residual body mass, weight corrected for tarsus length, tarsus length nor wing length (tarsus length: $F_{1,249} = 0.89$, $p = 0.35$; wing length: $F_{1,248} = 0.007$, $p = 0.80$; residual body mass: $F_{1,250} = 2.4$, $p = 0.12$; samples sizes differ because not all measurements were taken for all individuals).

Laying dates, predicted breeding values and environmental deviations

We found a trend that females with a higher carotenoid chroma laid later ($F_{1,133} = 7.38$, $p = 0.007$) but no differences between age groups ($F_{1,131} = 1.41$, $p = 0.24$) or areas ($F_{1,132} = 1.62$, $p = 0.21$). Neither age groups (age*chroma $F_{1,130} = 0.47$, $p = 0.50$) or areas differed in the relationship between carotenoid chroma and laying date (area*chroma $F_{1,129} = 0.33$, $p = 0.57$).

There was a significant positive relationship between female carotenoid chroma and predicted breeding value ($F_{1,132} = 9.98$, $p = 0.002$) (Fig. 3.1) and a significant difference between areas ($F_{1,132} = 6.42$, $p = 0.01$). The interaction area*carotenoid chroma was not significant ($F_{1,131} = 0.91$, $p = 0.34$). This positive relationship means that genetically late breeding females have a higher carotenoid chroma.

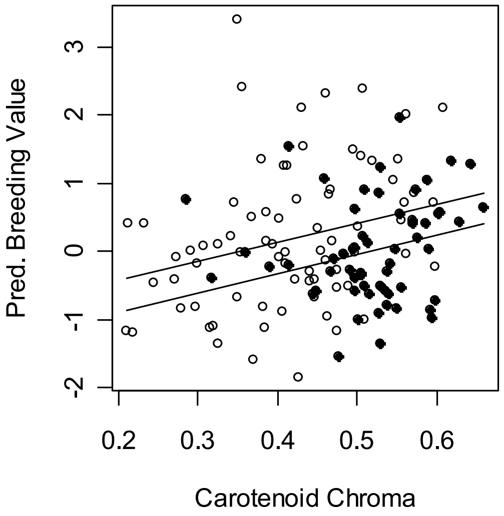

Fig. 3.1 Relationship between predicted breeding value of breeding time and female carotenoid chroma in the Hoge Veluwe (closed dots) and on Vlieland (open dots). Breeding values were significantly related to female carotenoid chroma ($F_{1,132} = 9.98$, $p = 0.002$) and differed between areas ($F_{1,132} = 6.42$, $p = 0.01$). There was no significant difference in this relationship between ares (area*carotenoid chroma: $F_{1,131} = 0.91$, $p = 0.34$).

Environmental deviations were positively related to carotenoid chroma ($F_{1,132} = 5.24$, $p = 0.02$) and negatively with age ($F_{1,132} = 21.5$, $p < 0.001$) (Fig. 3.2). There were no differences between areas ($F_{1,130} = 2.91$, $p = 0.09$). The interaction between area and carotenoid chroma was not significant ($F_{1,129} = 1.62$, $p = 0.21$) and the interaction between age and carotenoid chroma was also not significant ($F_{1,131} = 3.58$, $p = 0.06$).

A negative environmental deviation indicates that an individual breeds earlier than expected from its genotype and a positive environmental deviation that it breeds later than expected. The positive relationship between carotenoid chroma and envrionmental deviation means that females with a higher carotenoid chroma bred later than expected from their genotype. This was opposite to our expectations since females with a higher carotenoid chroma were supposed to be in better condition and therefore able to breed relatively early in relation to their genotypes. However, we found a positve relationship between carotenoid chroma and number of fledged chicks (see also Chapter 4) in old females ($F_{1,21} = 4.57$, $p = 0.04$) but not in young females ($F_{1,33} = 0.28$, $p = 0.60$; Chroma*age: $F_{1,54} = 4.71$, $p = 0.03$) on the Hoge Veluwe. Thus more yellow females fledged more chicks although they bred later than expected from their breeding values. On Vlieland we did not find any relationship between carotenoid chroma and number of fledged chicks ($F_{1,80} = 1.46$, $p = 0.23$).

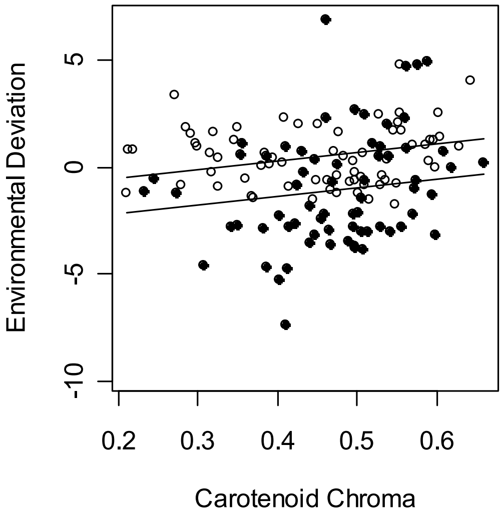

Fig. 3.2 Relationship between environmental deviation of breeding time and female carotenoid chroma for old (filled dots) and young females (open dots). The environmental deviation is defined as the residual of the regression of phenotype versus predicted breeding value. Old females had smaller environmental deviations than young females ($F_{1,132} = 21.5$, $p < 0.001$). Environmental deviations were positively related to carotenoid chroma ($F_{1,132} = 5.24$, $p = 0.02$). Both relationships did not differ significantly between the two areas (area*carotenoid chroma: $F_{1,129} = 1.62$, $p = 0.21$; area*age: $F_{1,128} = 1.19$, $p = 0.28$).

Discussion

It has been argued that the generally observed seasonal decline in reproductive success in birds is not causally related to breeding time but caused by a correlation between breeding time and fitness mediated by individual condition (Price *et al*. 1988). We tested this hypothesis by using female carotenoid based plumage colouration as a measure of inidividual condition and related it to the genetic and environmental component of breeding time and fitness. We used carotenoid colouration as a measure for condition since it is easily measured and related to immune response, survival, body condition (Hill and Montgomerie 1994; Hõrak *et al*. 2001; Saks *et al*. 2003).

We found the expected differences in carotenoid chroma between sexes and age groups: males and older individuals had a higher carotenoid chroma (Hõrak *et al*. 2001; Johnsen *et al*. 2003). We also found a significant difference between the study areas, birds on the island of Vlieland were paler than in the mainland study area Hoge Veluwe. This difference is likely environmentally caused since immigrant birds to Vlieland, which had not moulted on Vlieland, had a higher carotenoid chroma than birds that moulted on Vlieland (Postma and Gienapp, unpubl. manuscript).

Although we could not find a difference in laying dates between old and young females we found that environmental deviations were negatively related to age, older females having a more negative environmental deviation than young females born. Since a negative environmental deviation means that an individual breeds earlier than expected from its genes, this finding corresponds with the generally observed phenomenon that breeding time advances with age (e.g. Robertson and Rendell 2001; Wheelwright and Schultz 1994).

We found a positive correlation between carotenoid chroma and laying date meaning that more yellow individuals bred later. This was unexpected since individuals in good condition being thus more yellow should be able to start egg laying earlier. However, there was a trend that more yellow birds fledged more young, which is in line with the expectation that more yellow individuals are in better condition. This finding is evidence against selection on the environmental component since condition, measured as carotenoid colouration, is related with later laying but higher reproductive success.

Great tits moult their body feathers in August/September (Glutz von Blotzheim 1993). It could be that an individual's condition would change between moult in autumn and breeding in the following spring and would consequently be no longer correlated with plumage colouration. However, if such a time-lag would break up the relationship between plumage colouration and condition we would expect to find no correlation instead of a reversed relationship. Furthermore, several studies found a relationship between feathers moulted in autumn and indvidual properties measured later (Dufva and Allander 1995; Hõrak *et al*. 2001; Saks *et al*. 2003).

When analysing the relationship between carotenoid chroma and breeding values for laying date we found a similar relationship as at the phenotypic level: females with a high carotenoid chroma were genetically late breeding. However, an individual's predicted breeding value can resemble its phenotype if only few additional information from relatives can be obtained. Finding the same relationship as for phenotypes also in breeding values may therefore be a meaningless reflection of the phenotypic relationship. When areas were analysed separatedly the relationship remained significant for the Vlieland population ($F_{1,75} = 8.16$, $p = 0.006$), while it was no longer significant in the Hoge Veluwe ($F_{1,56} = 1.22$, $p = 0.28$). The Vlieland population is more closed than the Hoge Veluwe population and thus the pedigree deeper and wider (E. Postma, unpubl. manuscript). If we should find a spurious relationship with breeding values we should have found it on the Hoge Veluwe. This therefore seems to be an unlikely explanation.

Also for the environmental deviations we found a positive relationship with carotenoid chroma. This means that females with a high carotenoid chroma bred later than expected from their genotype. The question remaining is thus, how can we explain the unexpected relationship between condition or quality, measured as carotenoid colouration and the environmental as well as the genetic component of breeding time.

Since carotenoids have to be acquired from food a higher carotenoid chroma may indicate superior foraging abilities. Individuals that are able to forage more efficiently may also be able to obtain a higher proportion of high quality food containing many carotenoids. Producing eggs and incubating them is costly (Visser and Lessells 2001) and since tempeartures are increasing during spring and consequently also food supply is increasing it seems logical that the costs of egg production and incubation decrease during spring. By delaying breeding time females could reduce these costs but at the same time may breed too late in relation to the phenology of their chicks food supply. Individuals with superior foraging skills may however be able to avoid the costs of laying early because they are able to find enough prey for their chicks even when this food supply is already decreasing. This would then result in a positive relationship between carotenoid chroma and breeding time as well as reproductive success.

Territories could differ in the amount of carotenoids available (Slagsvold and Lifjeld 1985) and carotenoid rich territories do not necessarily have to be the earliest. Although in the Hoge Veluwe generally a decrease in reproductive success over the breeding season is observed (Chapter 7; Visser *et al*. 1998) there are years when this relationship is reversed. It is known that *Parus* species respond to spatial differences in insect phenology by adjusting their breeding time (Grieco *et al*. 2002; Nager and van Noordwijk 1995). If individuals of high quality or condition occupy high quality, carotenoid rich territories, which could be relatively late in a given year, an unexpected positive relationship between carotenoid colouration and environmental deviation could arise.

On the basis of our data we can not distinguish between these possible explanations. It would however be interesting to test whether the unexpected relationship found between carotenoid chroma and breeding time in these two population is caused by spatial variation in carotenoid availability combined with spatial differences in phenology, whether it is caused by a physiological trade-off between early egg production and colouration or by differences in foraging ability and high costs of egg laying. Although it has been repeatedly shown that breeding time is heritable (Sheldon *et al*. 2003; van der Jeugd and McCleery 2002; van Noordwijk *et al*. 1981) it is unclear what the actual underlying factors are. On one hand, individuals may use certain rules (Visser and Lambrechts 1999) to time their breeding time and these rules are inherited. Other factors, like food availability during egg production or basal metabolic rate (BMR), may affect then the environmental component of laying date. On the other hand, breeding time may only be a phenotypic outcome of some underlying trait and not a trait *per se*. This trait may be foraging efficiency, general quality, a physiological property as BMR or even dominance status related to personalities. If this trait is heritable, which has been found in, for example, parental effort (MacColl and Hatchwell 2003), carotenoid colouration (Fitze *et al*. 2003) or personalities (e.g. van Oers *et al*. 2004), it may seem that breeding time is a heritable trait. Addressing the above mentioned explanations chould shed some light on the evolution of breeding time.

Acknowledgements

H. Bouwmeester helped with collecting feather samples. The ‘Stichting Het Nationale Park De Hoge Veluwe’ and 'Staatsbosbeheer Vlieland' kindly allowed us to conduct our research on their property. J. Visser carefully managed the database. K. Lessells kindly allowed me to borrow the spectrophotometer and T. Limburg introduced me into its technical details. M. E. Visser made helpful comments on the manuscript.

Chapter 4

Female carotenoid colouration, female quality and male mate choice in great tits

Phillip Gienapp

Summary

In many bird species not only males but also females bear colourful ornaments. It has repeatedly been shown in males that these ornaments are related to quality and function as an honest signal in female mate choice. However, to date only few studies addressed the question what the role of the same ornaments is in females. They could have evolved as a correlated response to selection for the trait in males or they could signal female quality and act as a signal in male mate choice. In great tits (*Parus major*) males provide a substantial amount of parental care and are thus expected to be choosy in respect to their mate choice. In this study I addressed the question whether the carotenoid based yellow breast colouration in female great tits is related to their quality and whether they are preferred mates. In one population I found a relationship between female colouration and number of fledged young while it was absent in the other population. However, in both populations there was assortative mating with respect to carotenoid colouration. This is evidence that in great tits female carotenoid colouration is an indicator of quality and also used in male mate choice.

Introduction

In many animal species males show a bright carotenoid based colouration, which is thought to have evolved through sexual selection. For such signals to evolve they have to be honest indicators of individual quality, either genetic or phenotypic and there is ample evidence that carotenoid based colouration signals individual quality. For example, brighter individuals are healthier (e.g. Alonso-Alvarez *et al*. 2004; Faivre *et al*. 2003; McGraw and Ardia 2003; Saks *et al*. 2003), survive better (Hill 1991; Hõrak *et al*. 2001), are in better nutritional condition (Hill and Montgomerie 1994) or provide more care for the young (Senar *et al*. 2002). Consequently, these high quality individuals are preferred as mates (Hill 1991; Hill *et al*. 1999; Kodric-Brown and Nicoletto 2001; Milinski and Bakker 1990). While male ornamentation can thus be easily explained it is less clear why also females in many species are ornamented (Amundsen 2000), although generally to a lesser extent. One explanation is that female ornamentation is just a correlated response to selection favouring ornaments in males (Lande and Arnold 1985). Female ornamentation could persist if the selection for this trait in males is stronger than selection against this trait in females. There is, however, no reason why not also female ornamentation could have evolved as a product of male mate choice or female-female competition and there is now comparative evidence that female ornamentation has evolved several times independently in birds (Amundsen 2000) and experimental evidence for male mate preferences for ornamented females (Amundsen and Forsgren 2001; Hill 1993; Hunt *et al*. 1999; Jones and Hunter 1993).

Generally, females are supposed to be the ‘scarce’ sex and therefore can allow being choosy because they have higher costs of mating and reproduction. Under the most extreme scenario, when the costs of mating for males are negligible, males should not be choosy because only access to females limits their reproductive success. However, if the costs of mating increase because of, e.g., sperm depletion, loss of other mating opportunities or bi-parental care males should not select any longer their mates at random (Johnstone 1997). Under the most extreme opposite scenario, when males provide care for the young to the same extent as females, the skew in male mating success is small and extra-pair paternity low, males should be as choosy as females because the reproductive rates for the sexes are equal.

Male mate choice can explain assortative mating between the sexes with respect to the ornamental trait under consideration. If males were not choosy at all they would select randomly from the females being willing to mate with them; males preferred by females only would have a larger set, from which they then selected one female randomly. However, this could not lead to the assortative mating of males and females both being highly ornamented, as observed frequently in birds (Amundsen *et al*. 1997; Andersson *et al*. 1998; Boland *et al*. 2004; Burley and Coopersmith 1987; Jawor *et al*. 2003; MacDougall and Montgomerie 2003).

If female ornamentation evolved indeed as a response to sexual selection, analogous to male ornamentation, the same mechanism has to apply and female ornamentation also has to be an indicator of quality. To date only a limited number of studies looked at the relationship between female ornamentation and quality in birds and found varying evidence (reviewed in Amundsen 2000).

Great tits (*Parus major*) are small monogamous passerines, in which both sexes provide care for their young. Extra pair paternity is generally low in this species (Krokene *et al*. 1998; Lubjuhn *et al*. 1999; Strohbach *et al*. 1998; Verboven and Mateman 1997). Males should thus pay attention to the quality of their mates because it may have an important effect on their reproductive success. In a field study we investigated the relationship between female ornamentation and female quality in great tits. I measured carotenoid colouration in great tit females and related this to their reproductive success, which should be the ultimate currency a male is interested in.

Methods

Standard Fieldwork and Study Areas

Our study was carried out in two Dutch long-term study populations of nest box breeding great tits (*Parus major*). In both study areas ('Hoge Veluwe' and the island 'Vlieland') about 400 nest boxes are placed. The study area on Vlieland is distributed over five separate woodlands, clustered into to two sub areas, the 'West' and the 'East' (*cf* Postma and van Noordwijk 2005). However, the breeding population in the ‘West’ is considerably smaller and the rate of immigration higher (Postma and van Noordwijk 2005). Therefore we restricted our analysis to the Hoge Veluwe and the ‘East’ sub-area on Vlieland, hereafter Vlieland.

The vegetation on Vlieland consists mainly of mixed forest on poor sandy soils but also purely coniferous (*Pinus sylvestris*) and deciduous stands (*Quercus robur*) are found. The Hoge Veluwe study area is part of a larger wooded area also on poor sandy soils. The vegetation is a mixture of pedunculate oak (*Quercus robur*), red oak

(*Q. rubra*), birch (*Betula pendula*) and scots pine (*Pinus sylvestris*) with sparse undergrowth and interspersed with open heath land. Nest boxes were checked weekly to determine clutch size and laying date, i.e., the date when the first egg of a clutch was laid. Since successive eggs are laid at one-day intervals and clutch size usually exceeds six eggs weekly checks give accurate laying dates. All nestlings were ringed with uniquely numbered aluminum rings when they were seven days old. Shortly before fledging, when the chicks are 15 days old, all chicks were weighed. All adults were caught during chick feeding using trap doors inside the nest box and identified on the basis of their aluminum and colour rings or, if not yet ringed, ringed with aluminum and colour rings. At this occasion also their body mass, tarsus length and length of the third primary, hereafter wing length, were measured. Adults were caught only under good weather conditions and traps were placed not longer than 45 min in a nest box and checked at least every 15 min. Handling, and measuring adults took about 10 min per individual and individuals were released immediately afterwards. Adults resumed feeding within one hour at longest. This is not longer than both adults can stay away from the nest under natural conditions (own data). No broods were abandoned because of catching of the adults. Chicks were handled only under warm weather conditions and ringing and measuring took not longer than about 20 min per brood. Adult residual body mass was defined as the residual of a regression of body mass against tarsus length, additional factors in the model were sex, area and the interaction between sex and area (all significant at the 0.05-level). If not stated otherwise all statistical test reported here were GLMs assuming normal error distribution and were performed with R 2.0.1 (R Development Core Team 2004).

Colour Measurements

In 2004 we collected four yellow breast feathers from each individual taken from a standard position on the breast when catching the adults for standard fieldwork. The feathers were stored in light tight paper bags in a dry and cool place. For colour measurements four feathers were carefully arranged on top of each other resembling the natural situation and mounted on black velvet (*cf* Hõrak *et al*. 2001; Saks *et al*. 2003). Five replicate spectrum measurements were taken using an USB2000 Spectrometer with a DH2000 deuterium-halogen lamp as light source (both Avantes) illuminating and measuring perpendicular to the feathers. From the obtained spectra we calculated carotenoid chroma as the difference between the reflectance at 450 nm and 600 nm divided by the reflectance at 600 nm (R_{600}-R_{450}/R_{600}) (*cf* Johnsen *et al*. 2003). The yellow to red carotenoid-based colouration is produced by the absorbance of blue wavelengths, with a maximum absorbance at 450 nm. The chroma used is a measure of the maximum absorbance relative to the maximum reflectance, which is reached at 600 nm, and thus a measure of carotenoid content in the breast feathers. The raw data obtained from spectrometer software (Avasoft 5.1 Basic, Avantes) were processed, the chroma calculated and averaged over the five replicates using R 2.0.1.

Results

Sex, Age and Area Differences

We found significant differences in carotenoid chroma between sexes (males: 0.50 ± 0.01; females: 0.42 ± 0.01; $F_{1,255} = 9.8$, $p = 0.002$), age classes (young: 0.46 ± 0.01; old: 0.49 ± 0.01; $F_{1,255} = 16.0$, $p < 0.001$) and also study areas (Hoge Veluwe: 0.53 ± 0.01; Vlieland: 0.44 ± 0.01; $F_{1,255} = 86.5$, $p < 0.001$). Females and young individuals, born the previous year, had significantly a lower carotenoid chroma. Birds in the Hoge Veluwe had a higher carotenoid chroma than birds on Vlieland. Carotenoid chroma was not related to neither residual body mass, weight corrected for tarsus length, tarsus length nor wing length (tarsus length: $F_{1,249} = 0.89$, $p = 0.35$; wing length: $F_{1,248} = 0.007$, $p = 0.80$; residual body mass: $F_{1,250} = 2.4$, $p = 0.12$; samples sizes differ because not all measurements were taken for all individuals).

Assortative Mating

Carotenoid chroma of mated males and females was significantly correlated (Pearsons correlation: $r = 0.30$, $t_{120} = 3.46$, $p < 0.001$, both areas combined after correcting for area difference) (Fig. 4.1).

Reproductive Success – Hoge Veluwe

In old females a positive relationship between carotenoid chroma and clutch size was found while there was no such relationship in young females, born in the previous year. As usually found, clutch size decreased during the season (laying date: $F_{1,53} = 4.5$, $p = 0.04$, carotenoid chroma*age class: $F_{1,53} = 4.9$, $p = 0.03$). This effect on clutch size was found back in the number of fledged young, in old females there was a positive relationship between carotenoid chroma and number of fledged young while there was no relationship in young females (carotenoid chroma*age class: $F_{1,54} = 4.7$, $p = 0.03$) (Fig. 4.2a). The carotenoid chroma of the male, to which a female was mated, had no effect on both clutch size ($F_{1,52} = 0.16$, $p = 0.69$) and number of fledglings ($F_{1,52} = 0.02$, $p = 0.89$).

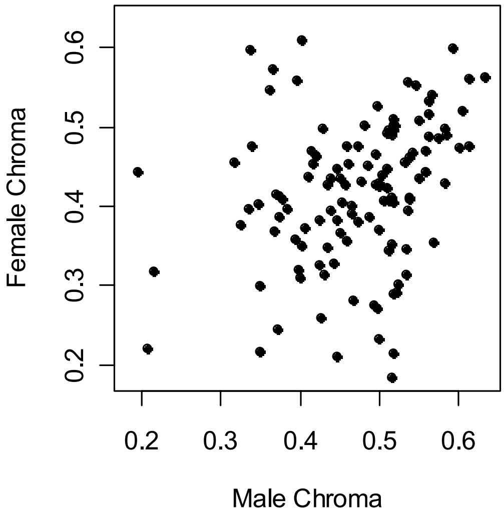

Fig. 4.1 Assortative mating between male and female great tits with respect to carotenoid chroma in two populations (Hoge Veluwe and Vlieland). Female and male carotenoid chroma corrected for area differences were significantly correlated (Pearson correlation: $r = 0.30$, $t_{120} = 3.46$, $p < 0.001$).

Female or male carotenoid chroma had no effect on brood mean chick weight (female: $F_{1,58} = 0.07$, $p = 0.79$; male: $F_{1,60} = 2.97$, $p = 0.09$) nor did age ($F_{1,62} = 2.1$, $p = 0.15$) but it decreased with clutch size ($F_{1,64} = 8.0$, $p = 0.006$). When four exceptionally light broods (all below 14.5 g) were excluded chick weight was negatively affected by hatching date ($F_{1,59} = 11.0$, $p = 0.002$) and clutch size ($F_{1,59} = 5.3$, $p = 0.02$) but not by carotenoid chroma (female: $F_{1,56} = 0.03$, $p = 0.86$; male: $F_{1,58} = 2.48$, $p = 0.12$) or female age ($F_{1,57} = 0.05$, $p = 0.82$).

Reproductive Success – Vlieland

On Vlieland clutch size decreased significantly during the season (laying date: $F_{1,69} = 4.2$, $p = 0.045$) and was positively related to the carotenoid chroma of the male, to which a female was mated ($F_{1,69} = 6.0$, $p = 0.02$). This effect was found back in the number of fledged chicks. While there was no relationship between carotenoid chroma and number of fledged chicks ($F_{1,57}$ = 1.0, p = 0.32) male carotenoid chroma was positively related to number of fledged chicks ($F_{1,69}$ = 5.8, p = 0.02). Older females fledged on average less young ($F_{1,81}$ = 5.0, p = 0.03) (Fig. 4.2b).

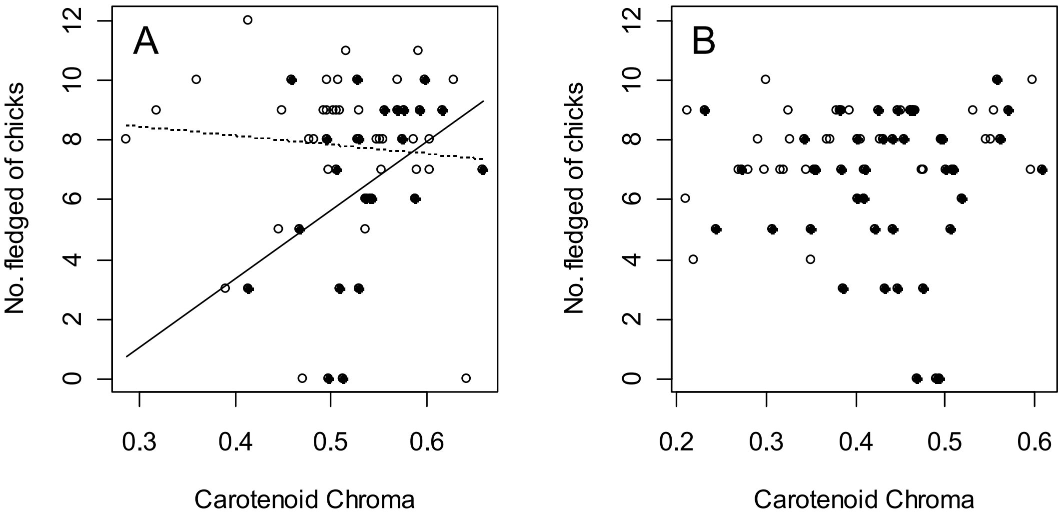

Fig. 4.2 Reproductive success in relation to female carotenoid chroma in two great tit populations. In the Hoge Veluwe (A) number of fledged chicks increased significantly with to carotenoid chroma in old females (filled dots, solid line) but not in young females (open dots, dotted line). On Vlieland (B) the number of fledged chicks was not significantly related to carotenoid chroma in young (open dots) or old females (filled dots) but older females fledged less chicks. See Results for statistical details.

Brood mean chick weight was positively related to hatching date ($F_{1,73}$ = 10.6, p = 0.002) and negatively to age ($F_{1,73}$ = 4.2, p = 0.04). No relationship between female carotenoid chroma ($F_{1,72}$ = 2.2, p = 0.15) or the carotenoid chroma of her mate (F = $_{1,71}$ = 0.5, p = 0.50) with chick weight was found.

Discussion

In many bird species males have brightly carotenoid based plumage colouration, which is generally assumed to signal the bearer's quality. Frequently, also the females are brightly coloured however generally to a lesser extent. To date, surprisingly few studies addressed the question whether female colouration is a signal for quality and play a role in male mate choice (Amundsen 2000). In this study I addressed this question whether carotenoid colouration in great tit females is related to their reproductive success, may therefore signal individual quality and is used as a cue in male mate choice.

In one study population (Hoge Veluwe) clutch size and number of fledged chicks were related to female carotenoid colouration (Fig. 4.2). More yellow females laid larger clutches and, as a consequence of this, fledged more young. However, this relationship was only found in old females; in females born in the previous year this relationship was not found. Out of the 38 young females breeding in this year only three were locally born recruits. Since in great tits natal dispersal takes place after the moult of the body feathers in August/September the vast majority of the young females has not moulted in the study area but elsewhere. Carotenoid availability in the environment has a strong influence on carotenoid based colouration (Olson and Owens 1998). For example, great tits from deciduous and coniferous habitats differed in yellowness (Slagsvold and Lifjeld 1985) and immigrant and resident great tits on the island of Vlieland differed in carotenoid chroma (Postma and Gienapp, unpubl. manuscript). This could mean that high quality individuals having moulted in habitats with low carotenoid availability may have the same carotenoid chroma as low quality individuals from habitats with high carotenoid availability. The lacking relationship between reproductive success and carotenoid colouration in young females may thus rather be an effect of habitat heterogeneity than of an absent relationship between carotenoid colouration and fitness. Brood mean chick weight was not affected by carotenoid chroma but there was a significant decrease over the season (after exclusion of four exceptionally light broods). Caterpillars are the main food source for chicks in great tits (Naef-Daenzer *et al*. 2000; Perrins 1991). In the year of this study the earliest brood hatched on 5 May, mean hatching date was about 12 May and the date when caterpillar biomass reached its maximum was 16 May (Chapter 8). Great tit chicks grow fastest and need most food when they are nine days old (Gebhardt-Henrich 1990; Keller and van Noordwijk 1994). This means that the majority of the Hoge Veluwe population bred too late in relation to their food supply, which led to the decline in, mean chick weight and probably overruled the effect of individual quality.

In contrast to the finding for the Hoge Veluwe population where carotenoid chroma of a female's mate was not related to her reproductive success in the Vlieland population clutch size and consequently number of fledged chicks was positively related to male carotenoid chroma. Clutch sizes of different females mated to the same male were not affected by the male's identity (E. Postma, unpubl. manuscript). If carotenoid chroma is a signal for male quality more yellow males may acquire and defend better territories with, e.g., a higher food abundance, which could be mirrored in the larger clutch sizes of their females.

It has repeatedly been shown that males with brighter carotenoid colouration are of higher quality and are preferred mates (see Introduction). If the costs of mating are low because, e.g., single matings only take a short time, sperm is not limiting, males do not provide parental care and acquiring a new mate is easy, high quality males may mate with every female without discriminating between high and low quality females. With increasing costs of mating when, e.g., males do provide parental care and acquiring a different mate is difficult males should become increasingly choosy. In great tits males take part in feeding the young, sometimes to a larger extent than the female (P. Gienapp, unpubl. data), and stay with the same female, at least for one breeding season (Kluyver 1951). Under these circumstances males mated to a low quality female may suffer a substantial reduction in their fitness and consequently males as well as females should be choosy and compete for high quality mates (Johnstone 1997; Owens and Thompson 1994). This should lead to a positive association between male and female trait. As expected for a species with high male parental effort, there was a positive and significant correlation between male and female carotenoid chroma (Fig. 4.1).

Female ornamentation has evolved as correlated response to selection favouring ornamentation in males (Lande and Arnold 1985) and then no relationship between ornamentation and quality in females should be expected. However, to conclude this from an absence of a relationship between female ornamentation and quality (Hill 1993; Lande and Arnold 1985; Muma and Weatherhead 1989) is difficult. Such absence can be found because the used measure for female quality may be incorrect or incomplete. On Vlieland no effect of carotenoid chroma on clutch size, number of fledged chicks or mean chick weight could be detected while there was a positive relationship with male carotenoid chroma. Although number of fledged chicks and mean chick weight are related to recruitment probability (Verboven and Visser 1998) a substantial proportion of mortality occurs in the first week after fledging (Naef-Daenzer *et al*. 2001). In stable populations lifetime reproductive success is the best feasible measure of fitness (Brommer *et al*. 2002). However, these data are not always available and small effects may not in all years and populations be detectable.

The data I presented here suggest that carotenoid based plumage colouration in great tits is an indicator of individual fitness and is used as a signal in male mate choice. Given the ubiquity of bright colouration in female birds it is surprising how few studies addressed this topic in spite of the large number of studies addressing the male quality and female mate choice in relation to bright colouration in males.

Acknowledgements

E. Postma and H. Bouwmeester helped with collecting feather samples. The ‘Stichting Het Nationale Park De Hoge Veluwe’ and 'Staatsbosbeheer Vlieland' kindly allowed us to conduct our research on their property. K. Lessells kindly allowed me to borrow and use the spectrophotometer and T. Limburg introduced me into its technical details. M. E. Visser made helpful comments on the manuscript. All handlings with adult and nestling birds were performed under license of the Animal Experimental Committee of the Royal Dutch Academy of Science.

Chapter 5

Selection on heritable phenotypic plasticity in a wild bird population

Daniel H. Nussey, Erik Postma, Phillip Gienapp and Marcel E. Visser

submitted to Science

Summary

Theoretical and laboratory research suggests that phenotypic plasticity can evolve under selection. However, evidence for its evolutionary potential from the wild is lacking. We present evidence from a Dutch population of great tits (*Parus major*) for variation in individual plasticity in timing of reproduction, and show that this variation is heritable. Selection favouring highly plastic individuals has intensified over a 32-year period. This temporal trend is concurrent with climate change causing a mismatch between the breeding times of the birds and their caterpillar prey. Continued selection on plasticity can act to alleviate this mismatch.

Introduction

Measuring life history plasticity in wild populations: concepts and past research

Many laboratory and theoretical studies have examined selection and genetic variation in phenotypic plasticity (Pigliucci 2001), most notably in plant systems (e.g. Weis and Gorman 1990; Scheiner 2002; Stinchcombe *et al*. 2004). However, only very recently have researchers begun using data from wild populations of long-lived species to investigate the evolutionary ecology of life history plasticity from a reaction norm perspective. Here, we present the key concepts in the analysis of plasticity in wild vertebrate systems and briefly review their recent applications. We then discuss the patterns of selection on plasticity presented in the main text in more detail. For a more detailed treatment of these issues we refer the reader to Nussey *et al*. (2005) and Brommer *et al*. (2005).

The reaction norm approach to modelling phenotypic plasticity uses linear (or polynomial) regression to estimate changes in a genotype's phenotypic state across a gradient of environmental conditions (Via *et al*. 1995). In the simplest, and most widely applied, form of this approach plasticity is estimated as the slope of a linear regression of a genotype's phenotype on an environmental covariate (e.g. Weis and Gorman 1990). This approach also estimates the elevation of a genotype's slope or plasticity, which – when the environmental covariate is mean-centred – can be taken to represent the genotype's phenotypic state in the average environment (Nussey *et al*. 2005; Brommer *et al*. 2005). In wild populations of longlived species we can treat each individual as a separate genotype and measure components of a reaction norm as phenotypic-environmental regression coefficients estimated across an individual's lifetime.

Early attempts to investigate plasticity in wild vertebrate populations utilising mixed-effects models assumed the slope component of individual reaction norms to be constant, whilst allowing individual elevations to vary (Przybylo *et al*. 2000; Reale *et al*. 2003; Schiegg *et al*. 2002). These studies estimated the average plastic response of individuals to an environmental variable across the population in question. Models that consider between-individual variation in plasticity (or slope) as well as elevation are a straightforward extension of this approach (Brommer *et al*. 2003; Brommer *et al*. 2005; Nussey *et al*. 2005). In particular, the application of a mixed model approach to life-history reaction norms in wild populations has allowed researchers to approximate the patterns of variation in individual elevation and slope for a given trait-environment relationship (Nussey *et al*. 2005), as well as providing estimates (BLUPs, see methods) of individual elevation and slope components for use in selection and quantitative genetics analyses (Brommer *et al*. 2005).

The relationship between elevation and slope components of reaction norms remains a problematic issue in research into phenotypic plasticity. Interdependence between the two is mathematically expected (Pigliucci 2001; Scheiner 1993). However, laboratory research into plasticity suggests that the relationship between elevation and slope components of reaction norms can vary across traits within systems (Newman 1994; Pigliucci *et al*. 2003). Research into plasticity in breeding time in red deer found that these correlations differed between population density phases, and has suggested a relationship between such life history plasticity and the phenotypic quality or condition of individuals (Nussey *et al*. 2005). Furthermore, a recent study of a wild bird population investigating the same trait environment relationship as examined here (laying date-spring temperature) found a negative correlation of lower magnitude between lay date elevation and plasticity to that found in the Hoge Veluwe great tits. The reasons for variation in elevation-slope correlations between and within laboratory systems and natural populations remain an issue for further theoretical and empirical examination.

Here, we have interpreted the high, positive covariance between female great tits' laying date in the average environment and laying date plasticity, as estimated by a random coefficients model, as biologically meaningful. The possibility of the covariance being caused by between-year differences is accounted for in our model by the inclusion of year as a random effect (see methods). We have suggested that a possible reason for such a correlation might be that both components of the reaction norms being correlated to some underlying and unmeasured aspect of condition or quality (see Results and Discussion).

Methods

Phenology data

Each year, nest boxes in the Hoge Veluwe are checked at least weekly during the breeding season and all nestlings ringed and breeding adults caught and identified, so that laying dates and the number of offspring recruiting into the breeding population are known for each female. We used data collected between 1973 and 2004. Laying date refers to the date on which a known female laid the first egg of her first clutch of the year. 'Spring temperature' is the mean daily temperature (in °C) over the period that best predicted mean annual laying date: 16th March – 20th April, measured at de Bilt weather station (http://www.knmi.nl/klimatologie/daggegevens/download.cgi?language=eng). Other data used in the analyses presented here are available on request from MEV (email: m.visser@nioo.knaw.nl).

Dates of the peak in caterpillar biomass were estimated from data collected 1985 – 2004 from frass fall samples in the Hoge Veluwe, from two main species: the winter moth (*Operophtera brumata*) and the oak leaf roller (*Tortrix virirdana*). Mean temperature 8th March – 17th May estimates the caterpillar peak well (r^2 = 0.80), and this relationship was used to predict caterpillar peaks prior to 1985 (see Visser *et al*. 1998 for more details). The synchrony between great tit and caterpillar phenology was estimated, in days, as laying date – peak in caterpillar biomass. We estimated the 'lifetime synchrony' with food availability of each female great tit breeding in more than one year as the average synchrony across her breeding lifetime. As caterpillar phenology has advanced in the last two decades, but great tit phenology has not, females are likely to be breeding too close to the date of the peak in caterpillar biomass to match their peak in food demands (estimated as the date their chicks are around 9 days old). Low lifetime synchrony values therefore represent females that are mistiming their laying dates.

Analysis of phenotypic variation in laying date reaction norms

Variation in females' laying date – spring temperature reaction norms was assessed using a restricted maximum likelihood linear mixed-effects model of laying date, run in GenStat v7.2 (VSN International), including data from 833 females that bred in more than one year. The full random-effects structure of the model included year of breeding (as a factor), female and the female by spring temperature interaction. The model also estimated the covariance between female and female by temperature interaction as an additional random effect (Nussey *et al*. 2005). Fixed-effects for year (as a covariate), spring temperature and its quadratic term (to test for non-linearity in reaction norms), age (as a two-level factor: females breeding in their second year of life, and older females), and their interactions were fitted to the model. All covariates were mean centred. The significance of random effects was assessed using log-likelihood ratio tests comparing models to which random effects were sequentially added (Pinheiro and Bates 2000). Significance of each fixed-effects term was then tested using Wald statistics divided by a term's degrees of freedom. The fixed-effect with the lowest Wald statistic/d.f. value was sequentially dropped until only significant terms ($p < 0.05$) remained (Nussey *et al*. 2005).

Random effects for female and female-temperature interaction generated by such a ‘random coefficients’ model (Brown and Prescott 1999) estimate variation in females’ laying date elevations and laying date-spring temperature slopes respectively. Re-running this model with only females that bred in three or more, or four or more, years did not affect the significance or magnitude of the variance components, suggesting findings were not biased by individuals with just two or three data points. Best linear unbiased predictors (BLUPs) were generated from the random effect for female (elevation estimates) and from the female by temperature interaction (slope or plasticity estimates) (see Nussey *et al*. (2005) for rationale).

Analysis of selection on reaction norms

A female’s total number of male and female offspring that recruited into the population as breeding adults was used as a measure of her lifetime reproductive success (LRS). 25 females still breeding in 2004 were excluded from the selection analysis as they may not have fulfilled their total LRS. Several females experienced one or more experimental manipulations of clutch or brood size during their lifetimes. Inclusion of experience of manipulation as a two-level factor in these selection analyses had no effect on results and was never a significant predictor of LRS. Prior to analysis, LRS was mean standardised and BLUPs were divided by their standard deviations (following Lande and Arnold 1983). Selection gradients were estimated using separate linear regressions of LRS on elevation and slope BLUPs. We checked for stabilising, and correlated selection on reaction norms using a multiple regression of LRS including elevation, slope, their squares and their cross-product (Weis and Gorman 1990). Temporal change in selection on reaction norms was tested by fitting a linear model with LRS as response variable and explanatory variables including both BLUP estimates, the year a female first bred (‘cohort’), and their interactions, and removing the term with the lowest P value sequentially until only significant terms ($p < 0.05$) remained.

To determine the role of the phenological mismatch between birds and caterpillars in driving any observed selection on plasticity we ran linear regressions of lifetime synchrony with females’ cohort, elevation and slope separately. We also investigated the possibility of changes in the relationship between slope and synchrony over time by fitting a linear model of lifetime synchrony with cohort, slope and their interaction.

'Animal model' analysis

A pedigree containing all ringed Hoge Veluwe great tits was constructed based on identified male and female parents of each clutch. For clutches with one or more unidentified parents, a dummy parent identity was assigned to maximise sibship information in the pedigree. Birds not ringed as nestlings were assigned unknown parents. The final Hoge Veluwe great tit pedigree contained 2,946 individuals and had an average lineage length of 3.71 generations. 'Animal models' have the advantage of using all available information on relatives to separate phenotypic variance into additive genetic and residual components (Kruuk 2004). In order to generate additive genetic and residual variance components for each trait as well as estimates of phenotypic and genetic covariances between them (Kruuk 2004) we ran a bivariate animal model of elevation and slope BLUPs in ASReml (VSN International) with cohort of female fitted as a fixed-effect factor. The 'narrow sense' heritability of each trait was calculated as the ratio of additive genetic variance to phenotypic variance (Lynch and Walsh 1998).

Results and Discussion

Phenotypic plasticity – defined as the ability of a single genotype to alter its phenotype in response to environmental conditions – represents an important mechanism by which populations can respond rapidly to changes in ecological conditions (Pigliucci 2001; Scheiner 1993; Via *et al*. 1995). Plasticity in life history traits is ubiquitous in animal populations (Pigliucci 2001) with traits often varying within the lifetimes of individuals depending on the conditions they experience (Nussey *et al*. 2005; Réale *et al*. 2003). It is typically conceptualized and measured using reaction norms: linear functions describing the change in a trait across an environmental gradient (Via *et al*. 1995; Weis and Gorman 1990). Laboratory research has shown that genetic variation for plasticity exists (Newman 1994; Stinchcombe *et al*. 2004) and that heritable plasticity can respond to artificial selection (Scheiner 1993; Scheiner 2002).

Given that many species are currently experiencing long-term anthropogenically driven environmental change (Stenseth *et al*. 2002; Walther *et al*. 2002), a widening of our understanding of how natural selection acts on plasticity under altered levels of environmental variation in the wild is imperative. Detailed analyses of within-population variation in life history plasticity are rarely undertaken in naturally occurring populations because such analyses require data from large numbers of individuals breeding repeatedly across their lifetimes. Recent research using mixed-effects linear models has shown that individuals within two wild vertebrate populations vary in their levels of life history plasticity (Brommer *et al*. 2005; Nussey *et al*. 2005). At present, little is known about the consequences of environmental change for the action of natural selection on plasticity and, ultimately, the ability of populations to continue to respond adaptively to environmental variation. Here we present data from a wild bird population showing temporal trends in natural selection on heritable phenotypic plasticity in the timing of reproduction, which are concurrent with changes in climate and the timing of food availability.

Female passerines often breed earlier following warm springs (McCleery and Perrins 1989; Visser *et al*. 2004) as a result of phenotypic plasticity (Both *et al*. 2004; Visser *et al*. 2004), i.e. the advancement of laying date in response to warm temperatures by individual females. This individual-level response to temperature is considered adaptive as it synchronizes the birds' phenology with the temperature dependent hatching times and growth rates of the caterpillars they rely on to feed their nestlings (Visser *et al*. 1998; Visser and Holleman 2001).

A long-term study of great tits (*Parus major*) in the Hoge Veluwe (the Netherlands) has revealed that, following recent warming of spring temperatures in the region, the timing of growth of their caterpillar prey has advanced whilst the phenology of the birds has not (Visser *et al*. 1998). As a result, over the past three decades the laying dates of female great tits have moved closer to the peak in the caterpillar biomass, so that the peak in demand for food for their offspring no longer coincides with the peak in prey availability (Visser *et al*. 1998; 2004). Selection on the heritable component of great tits' plastic responses to spring temperature could act to reduce this phenological mismatch (Visser *et al*. 2004).

We used information on laying dates for 833 females breeding in more than one year between 1973 and 2004 to examine variation among females in their laying date reaction norms. A random coefficients model of laying date showed that, on average, females began laying earlier following warmer springs (Tab. 5.1). We found significant variation between females in both their estimated laying date at the average spring temperature (Likelihood Ratio Test (LRT) = 226.73, d.f. = 1, $p < 0.001$) and the magnitude of their responses to spring temperature (LRT = 27.07, d.f. = 2, $p < 0.001$). Females in this population lay early following warm springs, but the magnitude of this plastic response varies between females. There was a significant, positive correlation between elevation and slope ($r = 0.40$, LRT = 15.41, d.f. = 1, $p < 0.001$): females that lay early in the average environment also are the most plastic females.

Tab. 5.1. Linear mixed-effects model of 2,195 laying date observations from 833 female great tits that bred in more than one year (1973 – 2004).

Random Effects:			
	Variance	SE	LRT
Year of breeding	9.54	2.55	558.33 ***
Female	8.05	0.76	226.73 ***
Female x Spring temperature	1.05	0.31	27.07 ***
Residual	14.97	0.64	
Fixed Effects:			
	Wald statistic	d.f.	Wald / d.f.
Spring temperature	45.47	1	45.47 ***
Age	116.91	1	116.91 ***
Age x Spring temperature	7.26	1	7.26 **

LRT: Log-likelihood Ratio test statistic; **: $p < 0.01$; ***: $p < 0.001$

Estimated covariance (female, female x spring temperature) = 1.16 ± 0.31

Significant genetic variation in a trait must exist for there to be any response to selection (Lynch and Walsh 1998). We generated predictors for the two components of each female's reaction norm: her laying date in the average environment (elevation) and her change in laying date in response to temperature (plasticity or slope) (Via *et al.* 1995). We used an 'animal model' (Kruuk 2004) to estimate the genetic component of phenotypic variance in predictors of female elevation and slope. We found that significant genetic variation for laying date plasticity exists in the Hoge Veluwe great tit population and laying date plasticity was significantly heritable ($h^2 = 0.30 \pm 0.14$, $z_{(<>0)} = 2.21$, $p < 0.05$, Fig. 5.1A). Genetic variation and heritability estimates for laying date elevation were relatively high but were not significantly greater than zero ($h^2 = 0.24 \pm 0.14$, $z_{(<>0)} = 1.73$, $p > 0.05$, Fig. 5.1B). However, the genetic correlation between slope and elevation was highly positive and not significantly different from one ($r_A = 0.77 \pm 0.18$, $z_{(<1)} = 1.28$, $p > 0.05$).

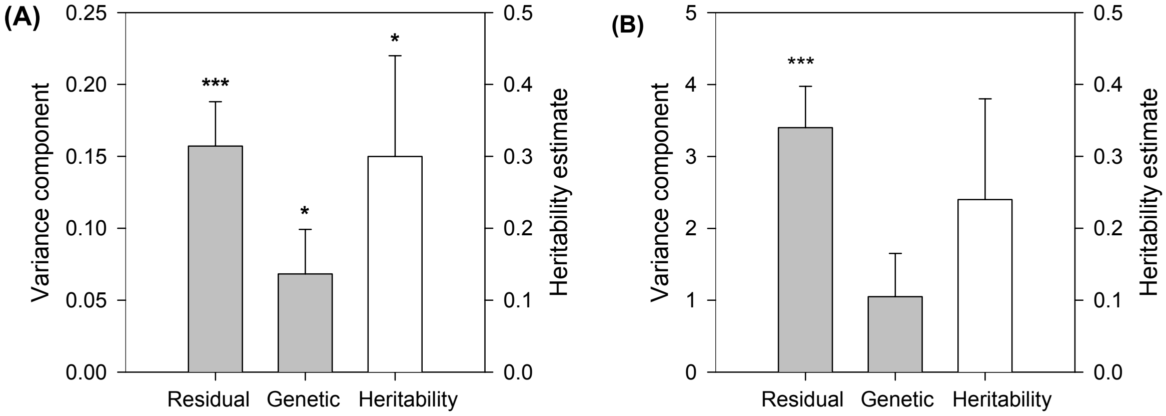

Fig. 5.1 Significant genetic variation exists for laying date reaction norms. Bar plots of 'animal model' estimates of residual and additive genetic variance (grey) and heritability (white) with standard errors bars for (A): laying date – spring temperature slope and (B): laying date elevation. The left y-axis shows variance component values for the grey bars, the right y-axis show predicted heritabilities for the white bars. Stars above bars indicate the estimate is significantly greater than zero (*: $p < 0.05$, ***: $p < 0.001$).

To investigate selection on laying date plasticity across the study period we measured the relationship between a female's lifetime reproductive success (LRS) and predictors of her laying date elevation and slope. There was evidence for directional selection on both reaction norm components, and there was no evidence of stabilizing or correlated selection (Tab. 5.2). Females that laid earlier in the average environment (low elevation) and that responded more strongly to temperature (more negative slope) had significantly more of their offspring recruit into the population as breeding adults (standardized selection gradients for elevation and slope: -0.094 ± 0.039 and -0.085 ± 0.039, respectively).

Tab. 5.2. The table shows the results of a multiple regression of female great tits' lifetime reproductive success including BLUPs for elevation and slope, their quadratics and cross-products. Lifetime reproductive success, and all BLUPs were standardized prior to inclusion in model.

Effect	d.f.	SS	*F*	*p*	*b*
Slope	1	5.82	4.78	0.03	-0.034
Slope^2	1	1.43	1.17	0.28	0.055
Elevation	1	1.71	1.41	0.24	-0.081
Elevation^2	1	0.04	0.04	0.85	0.018
Slope x Elevation	1	1.34	1.10	0.29	-0.063
Residual	802	975.0	1.22		

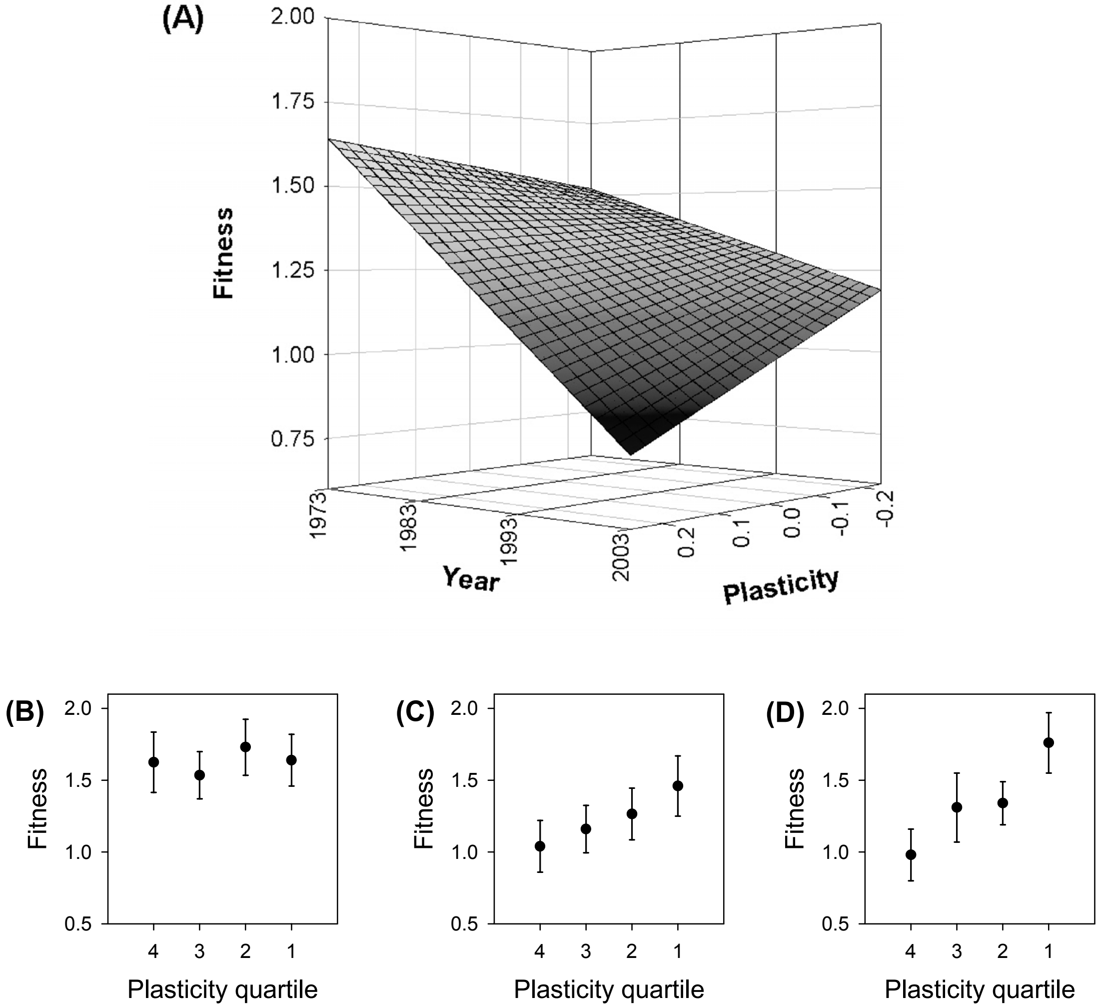

Fig. 5.2 Selection on plasticity is increasing over time. (A): the fitted model for unstandardised lifetime reproductive success ('fitness') of female great tits. 'Plasticity' is the predictor of a female's laying date – spring temperature slope; the predictors are centred on zero, so negative values represent females that advance laying more strongly than average following warm springs. (B, C, D): Plasticity predictor quartiles were estimated across the entire study period and mean LRS for each quartile (with SE bars) are shown for females first breeding between: (B): 1973–1982, (C): 1983–1992, (D): 1993–2002. Quartile 1 contains the most plastic females.

Selection favouring females that advanced laying dates strongly in response to warm springs increased over the study period (Fig. 5.2; slope by cohort interaction: $F_{1,804} = 7.22$, $p < 0.01$). It is clear from both the fitted interaction (Fig. 5.2A) and the data themselves (Fig. 5.2B, C, D) that selection has been strongest in the last two decades of the study, during which time the phenological mismatch with the peak in caterpillar biomass first emerged and then increased (Visser *et al.* 1998). The same pattern of changing selection over time is observed on estimates of females' laying date elevation (Tab. 5.3).

It appears that the strong correlation between females' elevation and slope renders these two components of their reaction norms indistinguishable. Selection favoured those highly plastic females that also lay early on average. How can we explain the correlation between elevation and slope? Both plasticity and elevation in laying date may be correlated to some unmeasured aspect of individual quality or condition (Brommer *et al*. 2005; Nussey *et al*. 2005). If birds differ in their ability to lay early in the year due to variation in some aspect of individual quality, for example because of differences in their ability to gather resources, high-quality birds will lay early in warm years and later in cold years and have steeper slopes and lower elevations, whilst poor-quality birds will usually lay later regardless of temperatures and therefore have shallower slopes. The early / plastic birds would be expected to match their reproductive timing better with the peak in caterpillar biomass, especially as spring temperatures become increasingly warm.

To substantiate the relationship between the great tits' reaction norms and the mismatch with the peak in food availability, we estimated each female's 'lifetime synchrony'. Improved lifetime synchrony was associated with increased laying date plasticity ($F_{1,806} = 189.4$, $p < 0.001$) and earlier breeding in the average environment ($F_{1,806} = 582.4$, $p < 0.001$). Furthermore, the synchrony of highly plastic females increased over time relative to less plastic individuals (slope by cohort interaction: $F_{1,804} = 55.8$, $p < 0.001$). The observed changes in selection on plasticity, as the phenological mismatch has increased over time, appear to be driven by the fact that highly plastic females breed in closer synchrony with the peak in food availability and hence have more resources available for provisioning their young.

Female LRS has decreased across the study period (Fig. 5.1; cohort main effect: $F_{1,804} = 9.92$, $p < 0.01$). If such a decline persists, alongside increased mismatching of phenologies between birds and caterpillars, the population's viability may ultimately be threatened. A phenotypic response to recent selection on laying date reaction norms cannot yet be demonstrated in this population, although the presence of additive genetic variance for plasticity means that a response to selection is predicted (Lynch and Walsh 1998; Kruuk 2004). However, a micro-evolutionary response to selection on the laying date reaction norms towards lower elevations and stronger plasticity would be expected to result in closer synchrony between the great tit's laying dates and the peak in food availability, and ultimately could alleviate the trophic mismatch.

We have shown that selection affects life history plasticity and that it can change with prevailing ecological conditions to potentially alter reaction norms in a wild population. This has wider implications because climate change has the potential to induce mismatches in the timing of breeding between trophic levels across a wide variety of ecosystems (Durant *et al*. 2003; Edwards and Richardson 2004; Visser *et al*. 1998). The capacity for evolutionary change in phenological reaction norms shown here represents a potential means for natural selection to alleviate such mismatches and their ultimately negative consequences for population viability and ecosystem function (Stenseth and Mysterud 2002; Walther *et al*. 2002). However, whether micro-evolutionary change in reaction norm shape can occur fast enough to keep up with the rapid rate of change in ecological conditions remains to be seen.

Tab. 5.3 The table shows maximum (a) and minimum adequate (b) linear models of lifetime reproductive success of female great tits which included BLUP estimates of elevation and slope, the year a female was first recorded breeding (‘cohort’, as a covariate) and their interactions. Lifetime reproductive success, slope and elevation were standardized prior to inclusion in model.

Effect	d.f.	SS	*F*	*p*	*b*
(a) Maximum model of lifetime reproductive success					
Slope	1	5.77	4.82	0.028	12.18
Elevation	1	1.84	1.53	0.216	10.75
Cohort	1	12.08	10.09	0.002	-0.014
Elevation x Slope	1	1.12	0.93	0.334	3.26
Slope x Cohort	1	5.83	4.87	0.028	-0.006
Elevation x Cohort	1	0.67	0.56	0.455	-0.005
Elevation x Slope x Cohort	1	0.30	0.25	0.617	-0.004
Residual	800				
(b) Minimum adequate model					
Slope	1	5.77	4.83	0.028	23.21
Cohort	1	11.84	9.92	0.002	-0.014
Slope x Cohort	1	8.61	7.22	0.007	-0.12
Residual	804	959.16			

Selection on reaction norm components in the Hoge Veluwe great tit population

Here we discuss patterns of selection on BLUP estimates of laying date elevation and slope in our study population in more detail, and in the light of the observed correlation between these two reaction norm components. The correlation implies that females that responded strongly to warm spring temperatures by advancing their laying dates (highly plastic females) also tended to lay earlier in the average environment (females with a low elevation). It should therefore not be surprising that we found little evidence of independent selection on one or other of these components of the birds' reaction norms. Indeed, multiple regressions of the lifetime reproductive success (LRS) of female great tits revealed that selection was not acting independently on BLUPs for elevation and slope. A multiple regression of LRS including both BLUPs, their quadratics and interaction term (Tab. 5.2) revealed no evidence of stabilizing selection on either slope or elevation. The removal of the quadratic and interaction terms from this selection model did not result in any significant change in the model's explanatory power. In a multiple regression containing only the linear terms for slope and elevation, neither effect was significant when assessed using Type III sums of squares (elevation: $F_{1,805} = 1.51$, $p = 0.22$; slope: $F_{1,805} = 0.55$, $p = 0.50$), however either term was significant when fitted singly to a linear regression of LRS (elevation: $F_{1,806} = 5.81$, $p < 0.05$; slope: $F_{1,806} = 4.75$, $p < 0.05$).

When a multiple regression of LRS was fitted with BLUPs for elevation and slope along with the year a female was first recorded breeding ('cohort') and their interactions and simplified to a minimum adequate model, effects for slope, cohort and their interaction were retained as significant (Tab. 5.3). However, fitted alone in a multiple regression, elevation and its interaction with cohort were also significant and showed effects in the same direction and of a similar magnitude to those in the minimum model including slope. Selection favouring both early laying (low elevation) and highly plastic (more negative slope) females increased over time, and the effect of slope is retained in the minimum model because its interaction with cohort explained marginally more residual variance than elevation's interaction with cohort.

As discussed in the main text, the high phenotypic and genetic correlations between elevation and slope, and the non-independent selection on them, lead us to expect that any response to selection on laying date reaction norms will involve a shift towards both lower elevations and higher plasticities in this population. We suspect that, owing to their close phenotypic and genetic correlation, the elevation and slope components of the great tits' lay date reaction norms are inseparable and that selection will effectively act on them as a single trait. The generality and causes of such correlations between reaction norm components in wild or laboratory settings is a subject for further investigation. In this population, selection favouring birds with low elevations and highly negative slopes has emerged alongside the climate change induced advancement of the phenology of the birds' caterpillar prey. We have argued that this is because these early / plastic individuals are breeding in closer synchrony with the caterpillars than the intermediate and late / non-plastic elements of the population.

We examined females' lifetime synchrony in order to substantiate this argument. A female with low lifetime synchrony lays eggs closer to the peak in caterpillar biomass, and is therefore less likely to synchronize her own peak in food demand for her nestlings with the peak in food availability. We found that lifetime synchrony decreased over time ($F_{1,806} = 295.7$, $p < 0.001$), which confirms a widening mismatch between the phenology of the great tits and their caterpillar prey at the individual level. Furthermore, a female's lifetime synchrony is associated with both elevation and slope components of laying date reaction norms. Females that are highly plastic are less likely to mismatch their laying dates with the peak in food availability ($b = -5.24 \pm 0.38$, $F_{1,806} = 189.4$, $p < 0.001$), as are females that lay early in the average environment ($b = -1.72 \pm 0.07$, $F_{1,806} = 582.4$, $p < 0.001$). The presence of a significant interaction in the expected direction between cohort and slope also implies the importance of phenotypic plasticity in matching the phenologies between trophic levels has increased as the caterpillar phenology has advanced following climate change (linear model of lifetime synchrony: $b(\text{slope}) = 500.84 \pm 67.65$, $F_{1,804} = 305.6$, $p < 0.001$; $b(\text{cohort}) = -0.34 \pm 0.02$, $F_{1,804} = 440.8$, $p < 0.001$; $b(\text{slope x cohort}) = -0.25 \pm 0.03$, $F_{1,804} = 55.83$, $p < 0.001$).

The presence of selection on laying date plasticity in this population, coupled with its heritability, leads to the prediction of a micro-evolutionary response to selection (Lynch and Walsh 1998; Kruuk 2004). We found no evidence of a phenotypic response to selection. Significant fixed-effects for year of breeding (as a covariate) or year by spring temperature interaction in our laying date random coefficients model would have indicated temporal trends in elevation or slope, respectively. However, neither was significant and they were dropped from the final model. Additionally, we tested the possibility of a genetic response to selection by averaging the breeding values for females' slope BLUPs generated by our animal model for each cohort of females and assessing temporal trends in these averages (Kruuk 2004). Although the trend in plasticity breeding values was in the expected negative direction, it was not significant ($b = -0.0004 \pm 0.0002$, $F_{1,28} = 2.52$, $p > 0.05$). The lack of evidence of a response to selection is not surprising when one considers the relatively short time period in which selection on reaction norms has been present. However, we would expect any micro-evolutionary response to selection on laying date plasticity and elevation to lead to an overall reduction in the previously documented trophic mismatch in this population.

Acknowledgements

We are grateful to J. H. van Balen and Jan Visser for managing the long-term study for many years, the board of National Park ‘De Hoge Veluwe’ for kindly allowing us to work on their property, and to Loeske Kruuk, Jon Bommer, Kate Lessells and four anonymous referees for comments on earlier drafts of the manuscript.

Chapter 6

Why laying date has not yet evolved in a population of great tits despite climate change induced selection for earlier laying

Phillip Gienapp, Erik Postma and Marcel E. Visser

Summary

Explaining why populations often do not seem to respond to selection is a crucial issue in evolutionary biology and may gain additional importance in the light of climate change. While egg laying dates of a Dutch great tit (*Parus major*) population have changed only marginally over the last decades, the timing of the main food source for the nestlings has advanced substantially, resulting in selection for earlier egg laying. We explored several explanations for the seemingly absent evolution of laying date, amongst which the often cited possibility that selection is only acting upon the environmental component of laying date. We found a small but significant heritability for laying date of 0.17. Fecundity selection via local recruitment on the genetic component was only slightly weaker than on phenotypic values. Selection via adult survival did not counteract fecundity selection and was generally weaker than fecundity selection. These levels of genetic variation and selection, together with the sex limitation of breeding time resulted in a cumulative predicted response to selection of about 1.5 days over a 30-year period. Given this small expected response to strong selection and the observed rates of environmental change, laying dates will evolve at an insufficient rate to restore the synchrony between great tits and their nestling food source, which is expected to have consequences at population level. More in general it emphasizes that the quantitative formulation of the expected response to selection is crucial in the study of microevolution.

Introduction

During the last decades the climate has changed at an unprecedented rate (Stott *et al*. 2000) and this has already affected a variety of biological systems. For example, growing seasons of plants have advanced and have become longer, and many animal species are now breeding earlier or have extended or shifted their distribution (e.g. McCarty 2001; Parmesan and Yohe 2003; Walther *et al*. 2002). There are however exceptions; while twenty British bird species have significantly advanced their laying dates, forty-four did not, and one species delayed egg laying (Crick *et al*. 1997). Additionally, trends can even differ among populations of the same species (Both *et al*. 2004; Dunn and Winkler 1999; Visser *et al*. 2003).

Due to the temporal heterogeneity in the increase in temperature (Easterling *et al*. 1997; Luterbacher *et al*. 2004) different trophic levels are possibly affected differently and thereby the synchrony between the phenology of predators and their prey may be disrupted (Visser *et al*. 2004). Caterpillars of the winter moth (*Opheroptera brumata*) hatch earlier nowadays (Visser and Holleman 2001), grow faster and their biomass peaks earlier (Visser *et al*. 1998). Such caterpillars comprise the main food source for nestlings of many woodland passerine bird species and the synchrony between the timing of breeding and food abundance is considered to be the main selection pressure on breeding time in these species (Dias and Blondel 1996; van Balen 1973; van Noordwijk *et al*. 1995). The date when the biomass of caterpillars is maximal has been measured in one long-term study area of great tits (*Parus major*) since 1985 and has advanced over this period by 13 days (Chapter 8). This change in food phenology has however not yet led to the expected change of laying dates of the birds, they advanced their laying dates by only about two days over the same period. Phenotypic selection differentials became more negative during the study period indicating an increasing selection pressure for earlier laying (Visser *et al*. 1998).

This suggests that the strong phenotypic selection reported by Visser *et al*. (1998) has not led to an evolutionary change in breeding time. There are several possible explanations why no phenotypic response to selection could be observed, especially in a highly plastic trait like laying date (reviewed in Merilä *et al*. 2001b). First, the predicted phenotypic change may be so small that it is obscured by strong year-to-year fluctuations in phenotypes. In our study population, annual mean laying dates can differ by two weeks from one year to the following. Consequently, detecting a time-trend can simply become a matter of statistical power. Second, the heritability estimate can bias the expected response. Heritability estimates obtained by parent-offspring regression can for example be inflated by spatial auto-correlation (van der Jeugd and McCleery 2002). Third, phenotypic selection differentials based on the number of offspring recruited into the population may not reflect all selection pressures acting on the specific trait. If early breeding females survived less well this counteracting survival selection may reduce, cancel or even reverse the selection pressure on breeding time. Fourth, counteracting selection on a genetically correlated trait may reduce the expected response (Lande and Arnold 1983). In passerine birds generally a strong (phenotypic) correlation between laying date and clutch size is found, which may be partly caused by a genetic correlation. Fifth, phenotypic selection differentials can be biased if selection acted not on breeding time itself but on a correlated but unmeasured environmental factor (Fisher 1958; Price *et al*. 1988). If for example female nutritional state determined her reproductive success and also affected the environmental component of breeding time but remained unmeasured one would assume a causal relationship between breeding time and fitness. If this trait was not heritable it would seem that breeding time does not respond to selection despite being heritable. Sixth, a response to selection could be masked if environmental conditions would affect the trait simultaneously in the opposite direction (Merilä *et al*. 2001a). Since producing eggs is costly in terms of nutrients, energy and future survival (Monaghan *et al*. 1998; Stevenson and Bryant 2000; Visser and Lessells 2001) deteriorating food availability during egg-production could constrain an advancement of laying dates, in spite of a genetic change. Accounting for these possible explanations will allow for a reliable and accurate estimation of the expected response to selection.

Recently, mixed model methodology has been used for the estimation of genetic parameters in natural populations (reviewed in Kruuk 2004). By fitting a so-called animal model we can obtain estimates of both genetic and non-genetic variance components, and we can separate an individual’s phenotype into genetic and environmental components. Whereas the variance components provide an accurate and precise estimate of the heritability for laying date, the separation of genetic and environmental components allows us to compare phenotypic and genetic changes over time, and selection differentials based on phenotypic and genotypic values.

In this study we first (re)analyzed data (now covering 31 years) from the same Dutch great tit population as Visser *et al*. (1998). Subsequently, we addressed the question whether the previously reported apparent lack of a phenotypic response to selection can be explained by a (very) small expected phenotypic response, by an inflated heritability estimate, by an incomplete estimation of the selection pressures, by a deteriorating environment obscuring the genetic change, by counteracting selection on a genetically correlated trait or by selection acting mainly on the environmental component of breeding time. Finally, we discuss the implications of our findings in the light of climate change and the potential of natural populations to adapt to environmental change.

Methods

Data Collection and Analysis

The great tit population on the Hoge Veluwe (the Netherlands) has been continuously monitored since 1955. In the winter 1972/73 a major storm devastated a part of the study area, which was subsequently redefined. Here, we thus analyze data for the years 1973 to 2003. The current study area covers a mixed woodland of 171 ha with about 400 nest boxes. The nest boxes are checked weekly to determine laying date and clutch size. Laying dates are calculated back assuming that one egg per day is laid. Adults are caught during chick feeding and identified by their aluminium and colour-bands or banded if not previously caught. All nestlings are banded at an age of seven days.

Only laying dates and clutch sizes of first clutches were used throughout the analyses. In the period 1973 to 2003 on average 125 pairs (sd: 29) of great tits bred in the study area. There was no significant trend in population size over time (linear regression: $b = -0.09 \pm 0.6$, df = 29, $p = 0.81$). Laying dates were independent of population density, whether or not temperature (mean from 1 March until 20 April, (van Balen 1973)) was corrected for (linear regression of mean laying date versus number of breeding pairs: $b = 0.033 \pm 0.030$, df = 29, $p = 0.29$, resp. $b = 0.021 \pm 0.019$, df = 28, $p = 0.26$). Temperature data for 'De Bilt', the main weather station of the Royal Dutch Meteorological Institute (KNMI), were used and obtained from the KNMI’s website (http://www.knmi.nl/voorl/kd/lijsten/daggem/etmgeg_downl.cgi?language=eng). Laying dates are given as Aprildays (1 April is Aprilday 1, 24 May is Aprilday 54).

All statistics except the quantitative genetic analyses were performed using R 2.0.1 (R Development Core Team 2004; Venables and Ripley 2002).

Quantitative Genetic Analysis

An animal model was used to separate the genetic and environmental components of individual laying dates (Kruuk 2004; Lynch and Walsh 1998). These linear mixed models use restricted maximum-likelihood (REML) methodology to estimate genetic and non-genetic variance components, and to simultaneously estimate fixed effects, as well as to predict additive genetic and other random effects. They are well suited for the quantification of genetic parameters of natural populations since they use all available information in the pedigree and are therefore able to accommodate, for example, inbreeding and selection. This provides us with more precise and accurate estimates of genetic parameters than can be obtained with traditional quantitative genetics methodology (Kruuk 2004).

Laying dates show strong year-to-year variation and are influenced by the age (or experience) of the female. To correct for these effects, year and age (first-year breeder or older) were included as fixed effects in the model. To accommodate breeding events of individuals in multiple years, individual identity was fitted as a random effect to account for any permanent environment effects. To summarize, our model separated the remaining phenotypic variance V_P (after taking out the effect of the fixed effects of year and age) into the additive genetic variance V_A, the variance attributed to permanent environment effects V_{PE} and the residual variance V_R. The narrow sense heritability h^2 was calculated as V_A / V_A+V_{PE}+V_R. Absolute variance components and ratios with their standard errors were estimated using the VCE 4 package (Groeneveld 1995).

Best linear unbiased predictors (BLUPs) for individual breeding values were obtained using the PEST software package (Groeneveld and Kovac 1990; Groeneveld *et al*. 1992). The reliability r^2 of the predicted breeding values (PBV) is calculated as the ratio of the variance of the predicted breeding values ($V_{\hat{A}}$) and the additive genetic variance (V_A) and ranges from h^2 when breeding values do not contain additional information on top of the breeding values, to 1 when true breeding values are known (Cameron 1997).

The pedigree was established by assuming that the parents raising a brood are the biological parents. Extra-pair paternity or brood parasitism can adulterate the assigned paternity and/or maternity. Although the level of extra-pair paternity in the study population is unknown, it is usually low in great tits (5-9%) (Krokene *et al*. 1998; Lubjuhn *et al*. 1999; Strohbach *et al*. 1998; Verboven and Mateman 1997). Furthermore, no incident of egg-dumping or brood parasitism was ever observed in our study population. In some years brood size manipulation experiments were carried out. When brood size was increased by adding eggs, all chicks were assigned unknown parents because it was impossible to track which chick hatched from an added egg. When chicks were cross-fostered or brood size increased by adding chicks, all chicks were individually marked and could thus be assigned the putative genetic parents.

Selection Differentials and Expected Responses

All recruits produced by a given breeding pair during a season were included to take into account that early-laying pairs are more likely to produce a second clutch (e.g. Verboven and Verhulst 1996). A female was counted as a survivor if she was caught alive (either breeding or not breeding) in a subsequent year. In some years brood size manipulation experiments were performed likely influencing breeding success or adult survival. All enlarged or reduced broods were excluded from the estimation of selection. Manipulated and unmanipulated broods did not differ in laying date, (unmanipulated) clutch size and age of the adults. Recruits from manipulated broods did not differ in their breeding performance from recruits coming from unmanipulated broods.

We used standard methods to calculate selection differentials and gradients (Arnold and Wade 1984; Brodie *et al.* 1995; Lande and Arnold 1983). Linear and nonlinear standardized selection gradients were calculated separately from regression models of the form:

$$w = c + \beta' * x' \text{ and } w = c + \beta' * x' + \gamma' * x'^2$$

where w is relative fitness, x' the standardized trait value, β' the standardized selection gradient and γ' the standardized non-linear selection gradient. Fitness was converted to relative fitness by dividing each individual's value (number of recruits or survival) by the annual mean fitness (i.e. by the mean number of offspring recruited per individual or by the proportion of surviving females in each year depending on the fitness component under consideration). Standardized laying dates were calculated by subtracting the annual mean and dividing by the annual standard deviation.

Testing whether phenotypic selection differentials are biased because selection is acting on the environmental component is possible by comparing selection differentials for phenotypes and breeding values (Rausher 1992; Stinchcombe *et al.* 2002). If an unmeasured trait is affecting simultaneously breeding time and fitness selection differentials for phenotypes will be larger than selection differentials for breeding values. However, comparing selection differentials and gradients for phenotypes and predicted breeding values is not entirely straightforward (Postma, J. Evol. Biol., in press). Since the variance in predicted breeding values ($\sigma^2(\hat{A})$) is always smaller than the variance in true breeding values, the selection gradient (β) based on PBVs will be upwardly biased, whereas the standardised selection differential (i) or gradient will be downwardly biased. This is however not true for selection differentials (S) since the covariance between relative fitness and predicted breeding values ($\sigma(w,\hat{A})$) is a function of the heritability but not of the reliability. It can therefore be directly compared to the covariance between relative fitness and true breeding values ($\sigma(w,A)$) (Postma, J. Evol. Biol., in press). We therefore based our comparison between selection on phenotypes and the genetic component of laying date on selection differentials (S) only.

In the absence of an environmental bias the selection differential for breeding values (S_{PBV}) is the expected response to selection on phenotypes, which is readily calculated as $R = S_P*h^2$ (Falconer and Mackay 1996). Hence we compared the selection differential for predicted breeding values (S_{PBV}) with the expected response to selection based on the selection differential for phenotypes (S_P) and the heritability. For comparison with other studies we also present linear and non-linear standardized selection gradients for phenotypes, which are equal to standardised selection differentials in univariate analyses. Since non-linear selection differentials for predicted breeding values are possibly biased (Sheldon *et al*. 2003) we restricted our comparative analysis to directional selection.

Predicting a response to selection over a longer period is difficult when selection pressures vary and generations are overlapping. Population density, winter climate or the amount of beech crop (Perdeck *et al*. 2000) could enhance or alleviate selection via recruitment and/or survival. To complicate matters further, selection via recruitment and survival does not need to be in the same direction. We therefore calculated the predicted response to selection on a year-to-year basis as follows:

for phenotypes:

$$\Delta P = S_{P\,\mathrm{recr}}*p_{\mathrm{recruit}}*h^2/2 + S_{P\,\mathrm{surv}}*(1-p_{\mathrm{recruit}})*h^2$$

and for predicted breeding values:

$$\Delta PBV = S_{PBV\,\mathrm{recr}}*p_{\mathrm{recruit}}/2 + S_{PBV\,\mathrm{surv}}*(1-p_{\mathrm{recruit}})$$

where $S_{P\ \mathrm{recr}}$ is the selection differential for phenotypes in year t based on recruitment, $S_{PBV\,\mathrm{recr}}$ is the selection differential for predicted breeding values in year t

based on recruitment, p_{recruit} is the proportion of the breeding population than consists of first year breeders in year *t+1*, $S_{P\,\text{surv}}$ is the selection differential in year *t* based on adult survival for phenotypes, $S_{PBV\,\text{surv}}$ is the selection differential in year *t* based on adult survival for predicted breeding values and h^2 is the heritability for laying dates. Laying date can be considered a sex-limited trait because males do not influence the laying date of the females mated to them (P. Gienapp, unpublished data). The expected response to fecundity selection has therefore to be halved since males are not selected for this trait.

Results

Annual Mean Laying Dates and Breeding Values

Mean laying dates showed considerable year-to-year variation, ranging from 16 April in 1974 to 3 May in 1975. The grand mean laying date for the whole period was 24 April. Annual mean laying dates showed a non-significant advancement of 5.4 days (linear regression of annual mean laying date versus year: $b = -0.182 \pm 0.094$, df = 29, $p = 0.06$) from 1973 until 2003 (Fig. 6.1A). Also annual mean breeding values did not significantly change over the study period ($b = -0.002 \pm 0.002$, df = 29, $p = 0.30$) (Fig. 6.1B).

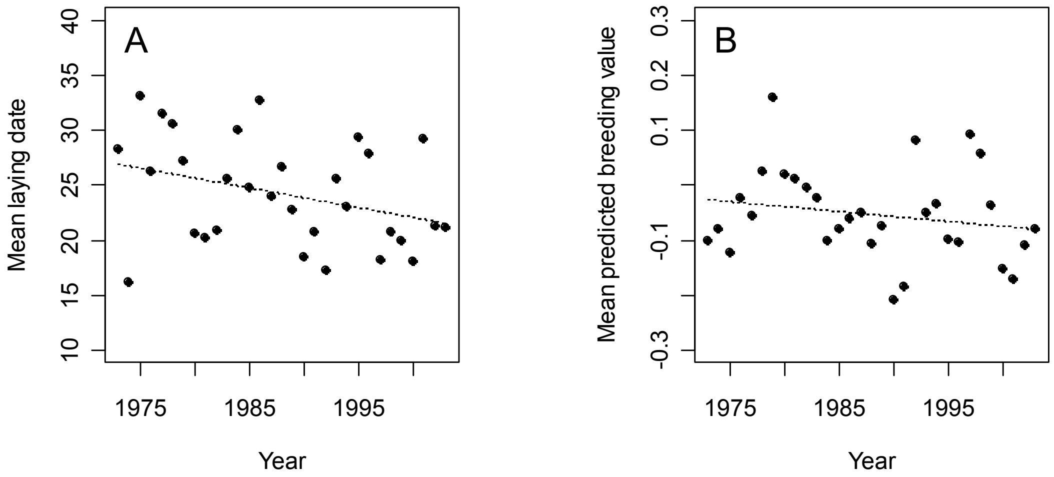

Fig. 6.1 Time trends in laying dates and breeding values for the Hoge Veluwe great tit population 1973 to 2003. Neither mean laying dates (a) (Aprildates, 1 = 1 April) nor mean breeding values (b) advancement significantly during this period (Laying dates: $b = -0.18 \pm 0.09$, df = 29, $p = 0.06$; predicted breeding values: $b = -0.002 \pm 0.002$, df = 29, $p = 0.30$; linear regression versus year).

Climate change has been shown to affect temperatures in spring, and the mean temperature from 1 March until 20 April (van Balen 1973) increased significantly by 1.7 °C (linear regression of mean temperature versus year: $b = 0.057 \pm 0.022$, df = 29, $p = 0.015$). To assess whether the non-significant time-trend in annual mean laying date was caused by phenotypic plasticity, a mixed model ANCOVA with individual as random effect and age (factor with two levels: ‘first year breeder’ and ‘older’) and temperature (mean 1 March until 20 April) as fixed effects was performed. This analysis showed that individuals bred on average 2.8 days earlier per 1°C increase in spring temperature (parameter estimate for temperature: -2.80 ± 0.09, $p < 0.0001$). Based on this and the observed increase in spring temperature, an advancement in mean laying date of 1.7 °C*2.80 days/°C = 4.8 days would be expected. This is very similar to the observed (though non-significant) change in laying dates of 5.4 days.

Heritability and Genetic Correlation

The estimated phenotypic and additive genetic variance for laying date obtained from the animal model were $V_P = 25.17$ and $V_A = 4.28$. The narrow sense heritability for laying date in this population is thus 0.17 ± 0.031 ($z = 5.48$, $p < 0.001$). The reliability ($=V_{\hat{A}}/V_A$) of the predicted breeding values was 0.30, which shows that our predicted breeding values are indeed providing additional information to the phenotypic values (Postma, J. Evol. Biol., in press). The narrow sense heritability for clutch size was 0.25 ± 0.034 ($t = 7.35$, df = 1790, $p < 0.001$). The genetic correlation between laying date and clutch size was small (-0.01 ± 0.12) and not significantly different from zero ($z = -0.08$, $p = 0.53$).

Selection via Adult Survival

There was no evidence for strong directional, stabilizing or disruptive viability selection on phenotypes or predicted breeding values of laying date when all years were averaged ($S_P = -0.08 \pm 0.13$; $i_P = -0.02 \pm 0.024$, $S_{PBV} = -0.03 \pm 0.03$) (Tab. 6.1, Appendix). There was also no evidence that survival selection acted in the opposite direction as fecundity selection, since survival and fecundity selection differentials were not negatively correlated for phenotypes ($r = 0.19$, $p = 0.32$) or predicted breeding values ($r = 0.01$, $p = 0.96$), either.

Selection via Recruitment

– On phenotypes

As reported earlier for this population (Visser *et al*. 1998), we found strong overall selection for early laying when fitness was measured as the number of offspring recruited into the breeding population (i_P = -0.21 ± 0.06, Tab. 6.2, Appendix). Unstandardized as well as standardized selection differentials decreased significantly with year (*b* = -0.093 ± 0.036, df = 28, *p* = 0.014 (Fig. 6.2A), *b* = -0.018 ± 0.006, df = 28, *p* = 0.007, respectively).

– On predicted breeding values

Selection differentials for predicted breeding values cannot be compared directly to selection differentials for phenotypes. However, in the absence of selection on the environmental component the selection differential for predicted breeding values is equal to the selection differential for phenotypes times the heritability, and thus to the expected phenotypic response to selection (see Methods). Hence we compared the expected phenotypic response to selection ($R = S_p{*}h^2$) with the selection differentials for predicted breeding values The average selection differential for predicted breeding values was less negative than the expected phenotypic response, however not significantly so (S_{PBV} = -0.126 ± 0.068; $S_p{*}h^2$ = -0.159 ± 0.57; *t* = 1.95, df = 29, *p* = 0.06, paired *t*-test). The selection differentials for predicted breeding values decreased significantly during the study period (*b* = -0.019 ± 0.007, df = 28, *p* = 0.011, Fig. 6.2B).

Predicted Response to Selection

The predicted change in annual mean breeding values was significantly correlated to the observed difference in mean breeding value from one year to the following (Pearson correlation: *r* = 0.58, df = 28, $p < 0.001$) (Fig. 6.3) This was also true when only fecundity selection was included in the calculation of the predicted change, and when this was compared to the observed change in mean predicted breeding values in first-year breeding females only (Pearson correlation: *r* = 0.38, df = 28, *p* = 0.038).

The cumulative response to selection including both fecundity and survival selection for phenotypes and predicted breeding values were very similar (cum. S_{PBV} = -1.5 days; cum. $S_P{*}h^2$ = -1.6 days) and small compared to the cumulative fecundity selection differential for phenotypes (-34.8 days). Time trends of cumulative responses were not significantly different for phenotypes and predicted breeding values (year^2*type: $F_{1,54} = 3.15$, $p = 0.08$; year*type: $F_{1,55} = 0.29$, $p = 0.59$) (Fig. 6.4). However, cumulative responses for predicted breeding values were less negative than cumulative responses for phenotypes (type: $F_{1,56} = 16.0$, $p < 0.001$). Both the linear and quadratic term of year were significant (year: $b = 10.24 \pm 0.73$, $F_{1,56} = 198.6$, $p < 0.001$; year^2: $b = -0.002 \pm 0.0002$, $F_{1,56} = 200.6$, $p < 0.001$).

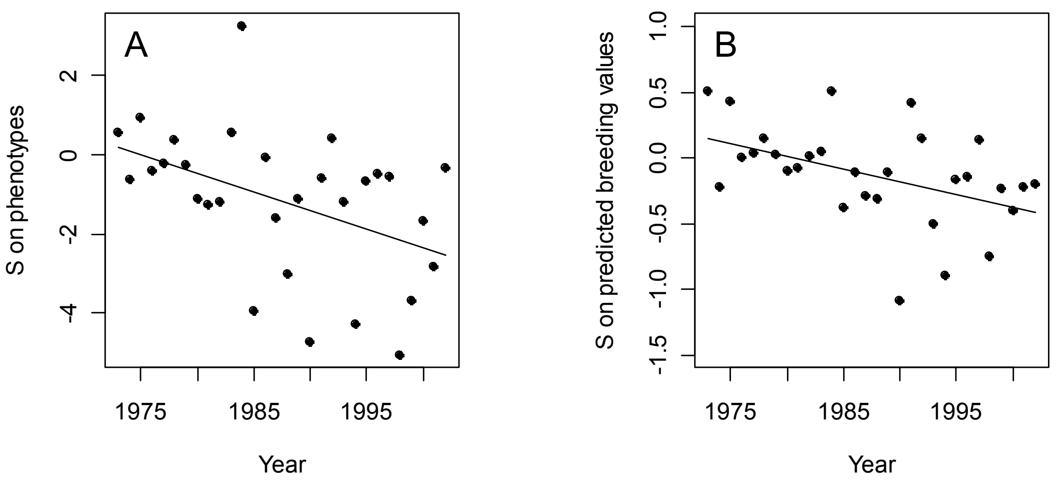

Fig. 6.2 Time-trends in selection differentials for the Hoge Veluwe great tit population 1973-2002. Selection differentials on phenotypes (A) and predicted breeding values (B) decreased significantly during this period (Laying dates: $b = -0.004 \pm 0.001$, df = 28, $p = 0.005$; predicted breeding values: $b = -0.005 \pm 0.002$, df = 28, $p = 0.007$, linear regression versus year).

Discussion

As already reported by Visser *et al.* (1998) we found strong directional selection for earlier breeding time at the phenotypic. The mean phenotypic standardized selection differential was -0.21, which is strong when compared to the median absolute value for the strength of natural selection on life-history/phenology traits (0.08, Kingsolver *et al.* 2001). However, we could not find any indication of a response to selection. Although Population mean laying dates advanced from 1973 to 2003 by 5.4 days (Fig. 6.1a), this non-significant time trend can be explained by a phenotypically plastic response to rising spring temperatures. There are several possible explanation for a lacking or not detectable response to selection (see Merilä *et al.* 2001b), which we will discuss in turn.

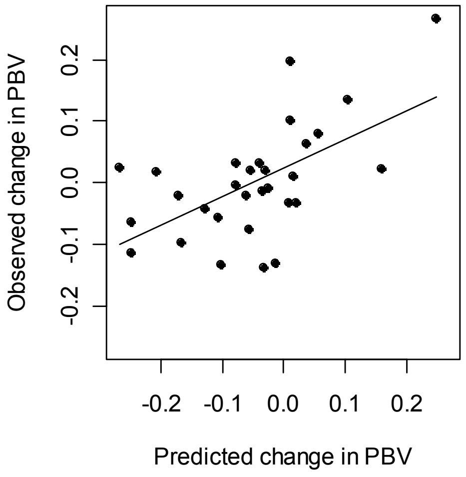

Fig. 6.3 Response to selection on breeding values. Observed and predicted changes in predicted breeding values were positively correlated ($r = 0.58$, df = 28, $p < 0.001$; linear regression observed versus predicted: $b = 0.46 \pm 0.12$). For details on calculation of predicted change see Methods.

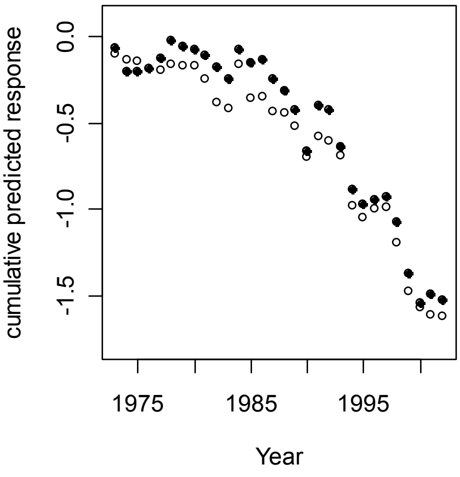

Fig. 6.4 Cumulative responses to fecundity and viability selection for predicted breeding values (filled dots) and phenotypes (open dots). Cumulative responses for both phenotypes and predicted breeding values decreased significantly over the study period (year: $b = 10.2 \pm 0.7$, $F_{1,56} = 198.6$, $p < 0.001$; year^2: $b = 0.003 \pm 0.002$, $F_{1,56} = 200.6$, $p < 0.001$) but there was no difference in time-trends between phenotypes and predicted breeding values (year^2*type: $F_{1,54} = 3.15$, $p = 0.08$; year*type: $F_{1,55} = 0.29$, $p = 0.59$). For details of calculation of predicted responses see Methods.

Heritability estimates based on parent-offspring regression can be upwardly biased due to spatial or temporal auto-correlation (van der Jeugd and McCleery 2002) and this could lead to an overestimation of the expected response. Phenotypes of parents and offspring may resemble each other because they experience similar environmental conditions. Animal model methodology can reduce the influence of environmental correlation on heritability estimates since it makes use of all relatives known from the pedigree (Kruuk 2004). The heritability we obtained from our animal model analysis was however similar to the heritability found for the same population ($h^2 = 0.18 \pm 0.08$, van Noordwijk *et al.* 1981).

Looking only at recruitment in relation to breeding time as is usually done (e.g. Both and Visser 2001; Svensson 1997; van Noordwijk *et al.* 1995; Visser *et al.* 1998) might give an incomplete picture of selection pressures if breeding time would also affect adult survival. A trade-off between reproductive success and adult survival could result in no net selection on breeding time (Brown and Brown 1999). However, in this population we did not find a positive relationship between breeding time and adult survival. Selection on breeding time via adult survival acted on average in the same direction as selection via recruitment, was generally much weaker and fecundity and viability selection were not negatively correlated both for phenotypes and predicted breeding values (Phenotypes: $r = 0.19$, df = 28, $p = 0.32$; predicted breeding values: $r = 0.01$, df = 28, $p = 0.96$).

We also did not find any evidence that selection on a genetically correlated trait affected the (micro)evolution of the focal trait (Lande and Arnold 1983). In this population early breeding individuals laid larger clutches (correlation between breeding time and clutch size: $r = -0.23$, $t = -14.5$, df = 3620, $p < 0.001$), as generally found in many bird species. Despite this strong phenotypic correlation between laying date and clutch size there was no evidence that these two traits are genetically correlated in this population.

One frequently suggested explanation for a lacking response to selection is that selection acted on the environmental component of breeding time (Fisher 1958; Price *et al*. 1988). In this case, selection differentials for predicted breeding values should be smaller than selection differentials for phenotypes times the heritability (Rausher 1992; Stinchcombe *et al*. 2002). Fecundity selection differentials for predicted breeding values were smaller than fecundity selection differentials for phenotypes (after ‘correcting’ for the heritability, see Methods), however not significantly. Under this hypothesis we have a clear expectation about the difference between selection differentials for phenotypes and predicted breeding values in the presence of the selection for earlier laying, namely that selection differentials for breeding values should be smaller than selection differentials for phenotypes 'corrected' for the heritability. However, when phenotypic selection favours later laying the expectation is less clear. Although fecundity selection was in general favouring earlier laying fecundity selection pressures were reversed in single years ($n = 6$). When these years were excluded from the comparison between fecundity selection for phenotypes and predicted breeding values the difference between them became smaller and less significant ($S_P{*}h^2 = -0.29 \pm 0.05$; $S_{PBV} = -0.23 \pm 0.08$; $t = 1.36$, df = 23, $p = 0.19$, paired t-test). Additionally, the slopes of the time-trends in cumulative predicted responses for phenotypes and predicted breeding values did not differ (Fig. 6.4). This is the most powerful way to detect differences in selection since it integrates both selection via recruitment and adult survival over a long time-period. However, in this kind of analysis only the difference in the slope of the time-trends between groups is meaningful since any significant difference in intercept can be caused by a strong difference between selection differentials in one year, which is then ‘pseudo-replicated’ when calculating the cumulative response.

Our finding that selection acted directly on breeding time and not on the environmental component is in accordance with the results of many experiments having manipulated breeding time, either by removing first clutches, inducing the females to re-nest (e.g. Hatchwell 1991; Verhulst and Tinbergen 1991; Verhulst *et al*. 1995), or by swapping clutches between early and late breeding pairs (e.g. Brinkhof *et al*. 1993; Verboven and Verhulst 1996). In general it was found that delayed females had a reduced fitness. There are however caveats with these kinds of experiments because delayed females either had to lay a complete replacement clutch (removal experiments) or had to incubate longer (swapping experiments). Egg laying and incubation are costly (Visser and Lessells 2001) and might impair the female's ability to successfully raise chicks. Consequently, more elegant experiments to manipulate breeding time are desirable (*cf* Christians *et al*. 2001).

To conclude, we did not find evidence that either very low or absent heritability of breeding time, selection on a genetically correlated trait, counteracting selection on adult survival or selection only on the environmental component could explain the seemingly lacking response to selection on breeding time in our study population. The strong phenotypic selection, the cumulative phenotypic selection differential was -34.8 days, however only led to a cumulative expected response of about 1.5 days when selection on adult survival and heritability were taken into account. Although survival selection generally acted in the same direction this difference can be easily explained since survival selection was weaker than fecundity selection and adult survival rates relatively high (mean 1973-2003: 0.54). Since the estimated heritability was relatively low and breeding time is a sex limited trait (thus halving the expected response) the effect of weak adult survival selection reduces the strong phenotypic fecundity selection. This illustrates that drawing conclusions from strong fecundity selection at the phenotypic level about the expected response can be very misleading.

The expected cumulative response of selection on predicted breeding values was about 1.5 days and we should expect to see a corresponding change at the genetic level. However, also predicted breeding values did not show a significant time-trend over the study period (Fig. 6.1b). Predicted breeding values do not show strong year-to-year fluctuations driven by environmental factors, and therefore weak time-trends may be easier to detect in the breeding values than directly in phenotypic values. Additionally, it is possible to detect responses to selection that are being masked when environmental changes affect the phenotype in the opposite direction (Garant *et al*. 2004; Merilä *et al*. 2001a) or to differentiate whether a phenotypic trend is caused by phenotypic plasticity or an underlying genetic change. The absence of a time-trend in mean predicted breeding values can however be misleading. If no information about genetic relationships among individuals were available, an individual's breeding value would simply be the deviation of its phenotype from the population mean times the heritability. In this case the annual mean breeding values will by definition be zero if phenotypes are corrected for among year differences by including year in the animal model. This correction has been applied because otherwise a trend in breeding values might simply reflect a phenotypic trend. How much information about genetic relationships is necessary to obtain reasonably accurate breeding values is an area in need of further research (*cf* Postma, J. Evol. Biol., in press). The fact that we found a correspondence between the expected and observed changes in predicted breeding values from year-to-year (Fig. 6.3) is a clear indication that the predicted breeding values are meaningful. However, the connectivity of the pedigree between subsequent years is higher than over the whole study period as great tits are comparatively short-lived and relatives tend to breed in a limited range of years. Although it was possible to detect meaningful short-term changes in predicted breeding values the low connectivity over the whole study period may limit the power to detect small long-term changes.

In our study population the mismatch between caterpillar phenology and great tit breeding time is currently about two weeks. A microevolutionary response induced by climate change has been reported for phenology in mosquitoes (Bradshaw and Holzapfel 2001) and squirrels (Réale *et al*. 2003). The question however is how fast this will happen in great tits, and whether it can restore the synchrony between caterpillars and birds. Given the small predicted response despite the strong selection at the phenotypic level resulting from the observed mismatch, it will take many decades before an evolutionary response will have restored the synchrony between caterpillars and great tits, assuming that the environment would not change further. Unfortunately, the latter is unlikely to be true: global average temperatures are predicted to increase more or less continuously for the next 100 years by in total 2 to 4.5 °C (Houghton *et al*. 2001) and this will lead to a further advancement of the caterpillar peak. The relatively low heritability of laying dates compared to other life-history traits and the sex-limitation of this trait will slow down the response to selection in such a way that the mismatch between food availability and the birds' breeding time will increase. Consequently, we expect that the increasing mismatch caused by the expected slow evolutionary response will have consequences at the population level in the great tits.

Acknowledgements

J. H. van van Balen kept the long-term study on the Hoge Veluwe going for many years, J. Visser carefully managed the database. Many people helped with fieldwork. The ‘Stichting Het Nationale Park De Hoge Veluwe’ kindly allowed us to conduct our research on their property. EP thanks L. Kruuk for providing invaluable help with VCE and PEST. K. Lessells made helpful comments on earlier versions of the manuscript. EP was financed by the Netherlands Organisation for Scientific Research (ALW-NWO) and a Marie Curie Host Fellowship.

Appendix

Tab. 6.1 Viability selection on laying date

Standardized linear (β'_P) and non-linear selection gradients (β'_P and γ'_P) for phenotypes. In univariate analyses standardized selection gradients are equal to standardized selection differentials.

	linear	non linear	
year	β'_P	β'_P	γ'_P
1973	-0.246	-0.292	0.087
1974	-0.017	-0.085	0.078
1975	-0.049	-0.056	0.110
1976	-0.051	-0.036	-0.027
1977	-0.003	-0.035	0.047
1978	0.050	-0.038	0.108
1979	0.005	0.074	-0.056
1980	0.145	0.022	0.095
1981	-0.065	-0.047	-0.017
1982	-0.179	-0.122	-0.062
1983	-0.188	-0.198	0.026
1984	0.260	0.063	0.180
1985	0.150	0.000	0.145
1986	0.070	0.107	-0.059
1987	-0.036	0.074	-0.055
1988	0.044	0.100	-0.133
1989	-0.057	-0.118	0.163
1990	0.089	0.024	-0.128
1991	0.151	0.145	0.008
1992	-0.083	0.000	-0.136
1993	-0.137	-0.151	-0.042
1994	-0.273	-0.183	-0.080
1995	-0.103	-0.190	0.063
1996	0.215	0.300	-0.075
1997	0.063	0.183	-0.186
1998	-0.069	-0.069	0.001
1999	-0.250	-0.241	0.158
2000	-0.030	-0.032	0.056
2001	0.054	0.086	0.045
2002	0.023	0.026	0.007
mean	-0.017	-0.023	0.011
SE	0.024	0.024	0.018

Tab. 6.2 Fecundity selection on laying date

Standardized linear (β'_P) and non-linear selection gradients (β'_P and γ'_P) for phenotypes. In univariate analyses standardized selection gradients are equal to standardized selection differentials.

	linear	non linear	
year	β'_P	β'_P	γ'_P
1973	0.124	0.226	-0.194
1974	-0.094	0.106	-0.232
1975	0.170	0.178	-0.122
1976	-0.075	-0.060	-0.026
1977	-0.051	0.057	-0.159
1978	0.089	0.256	-0.204
1979	-0.041	-0.083	0.034
1980	-0.258	-0.328	0.053
1981	-0.206	-0.005	-0.193
1982	-0.229	-0.106	-0.133
1983	0.121	0.082	0.097
1984	0.606	0.052	0.509
1985	-0.572	-0.071	0.134
1986	-0.030	-0.071	0.064
1987	-0.337	-0.412	0.038
1988	-0.672	-0.797	0.299
1989	-0.179	-0.267	0.235
1990	-0.676	-0.664	0.024
1991	-0.077	-0.137	0.076
1992	0.076	0.168	-0.149
1993	-0.380	-0.416	-0.111
1994	-0.878	-1.316	0.390
1995	-0.139	-0.012	-0.092
1996	-0.143	-0.275	0.115
1997	-0.096	-0.077	-0.030
1998	-0.872	-0.826	0.426
1999	-0.574	-0.556	0.328
2000	-0.342	-0.340	-0.051
2001	-0.647	-0.567	0.114
2002	-0.058	-0.110	-0.134
mean	-0.215	-0.215	-0.017
SE	0.061	0.061	0.024

Chapter 7

A new statistical tool to predict phenology under climate change scenarios

Phillip Gienapp, Lia Hemerik & Marcel E. Visser

Global Change Biology 11, 600-606 (2005)

Summary

Climate change will likely affect the phenology of trophic levels differently and thereby disrupt the phenological synchrony between predators and prey. To predict this disruption of the synchrony under different climate change scenarios, good descriptive models for the phenology of the different species are necessary. Many phenological models are based on regressing the observed phenological event against temperatures measured over a fixed period. This is problematic, especially when used for future predictions because the paradoxical situation could arise that the phenological event occurs before the period over which temperature is measured. Such models are unable to predict population variation in phenology. Here, we propose a ‘proportional hazards model’ to describe phenology and illustrate it with an example from breeding time in birds. This type of model circumvents the above mentioned problems and is generally applicable for describing phenology, which is essential when assessing the ecological impact of climate change.

Introduction

Climate change has already affected the phenology of various taxa, covering all trophic levels (e.g. Brown *et al*. 1999; Crick *et al*. 1997; Gibbs and Breisch 2001; Parmesan and Yohe 2003; Penuelas *et al*. 2002; Root *et al*. 2003; Roy and Sparks 2000; Tryjanowski *et al*. 2002). However, not all trophic levels are necessarily reacting at the same rate, which can lead to a disruption of the synchrony between consumers and their prey (Visser and Holleman 2001; Visser *et al*. 1998; Voigt *et al*. 2003; Watt and McFarlane 2002). Small, woodland passerine birds, like great tits or flycatchers, feed their young mainly with caterpillars and the synchrony between nestling period and caterpillar abundance is the main determinant of reproductive success in these species (van Balen 1973; van Noordwijk *et al*. 1995). A decline in synchrony should eventually lead to a reduction in overall reproductive success and will thereby affect population numbers negatively (Visser *et al*. 2004). Consequently, even small changes in climate, which do not shift the ‘climate envelope’ (*cf* Thomas *et al*. 2004), can already have perceivable influences at population levels. To be able to fully assess the impacts of climate change under the various climate scenarios (Houghton *et al*. 2001) we thus need good predictive models for the phenology of the different trophic levels in the ‘food chain’.

Several studies already established relationships between spring phenology and spring temperatures in a variety of taxa, ranging from trees to birds (e.g. Beebee 1995; Cannell and Smith 1983; Meijer *et al*. 1999; Nizinski and Saugier 1988; Perrins and McCleery 1989; Sparks and Carey 1995; Visser and Holleman 2001). In these studies, two different general methods have been used to describe phenology. The mean dates of phenological events (e.g. bud burst, date of appearance, breeding time) has been described either by regressing these dates against a temperature measured over a certain fixed period or by a threshold, i.e. the phenological event occurs when certain temperature conditions are fulfilled. There are however differences between taxa in which models are applied. In almost all studies on vertebrates (e.g. Beebee 1995; Meijer *et al*. 1999; Perrins and McCleery 1989; Sparks and Carey 1995) regression of the phenological event (mostly breeding time) against a temperature measured over a fixed period was used, while the phenology of plants and insects is usually described by temperature-thresholds (e.g. when a certain number of degree-days is reached) (Cannell and Smith 1983; Nizinski and Saugier 1988; Visser and Holleman 2001).

The two ways of describing phenology also differ in their suitability for future predictions of phenology. If the timing of the phenological event is regressed against a certain temperature measured over a fixed period (e.g. average egg laying date in a bird species versus average daily mean temperature in April and May) this temperature-measure can only be a proxy for the factor(s) causally influencing general conditions. It is highly unlikely that temperatures on 31 March or 1 June do not influence egg laying or that individuals ‘measure’ mean temperatures. Additionally, some individuals will have started egg laying before the end of the period predicting their laying date. This method, using a proxy instead of the causal factor(s), can explain among year variation in phenology, as long as the relationship between the proxy and the unknown factor(s) remains stable. Unfortunately, this is unlikely to be true under climate change, which limits its use for predicting phenological changes induced by climate change. Temporal temperature patterns have already changed and will continue to do so (e.g. late winter and spring have warmed more strongly than summer) (Easterling *et al*. 1997). In that respect, threshold-models, which predict a given phenological event when certain conditions are fulfilled, may be more flexible and better suited. However, they will only perform better if the true causal factor has been identified and incorporated into the model, which will be difficult to test. Another problem, which applies to both types of models, is that a model describing the (annual mean) date of the given phenological event is per se not able to predict the within-year variation of this phenological event. It is, however, important to be able to predict this, because, even if the ‘average’ individual suffers from mistiming, the earliest individuals may still be well synchronized and breed successfully. These problems can be solved by analyzing individual observations rather than annual means with a statistical model that is based on all temperatures up until the event rather than temperatures measured over a fixed period. Moreover, this can allow us to predict the within-year variation.

Proportional hazards models (‘Cox regression model’ Cox 1972) were originally developed for medical survival analysis but have already been applied to questions in biology (e.g. for analyzing foraging behavior in parasitoids or development time in mosquitoes) (Haccou *et al*. 1991; Hemerik *et al*. 1993; Koenraadt *et al*. 2004; Wajnberg *et al*. 2003). Any 'time to event' can be analyzed with survival analysis (Kleinbaum 1996). The event can for instance be leaving a patch, pupation or laying the first egg of a clutch. The model describes the probability per time unit that the event occurs, as a function of a basic probability (the baseline hazard) and a set of explanatory variables. The baseline hazard is modified by the variables, which can include fixed properties of the individual (e.g. age, experience or condition) and time dependent variables (e.g. temperature or photoperiod). The model is non parametric and no assumptions about distributions of the ‘time to event’ have to be made. The probabilistic nature of the model allows us to predict the variation in timing of the phenological event (here egg laying).

Our aim is thus, to develop a descriptive model for laying dates of great tits based on a proportional hazards model. Great tits were used as a model organism because their biology and the biology of their main prey species are well studied, the consequences of the synchrony between prey phenology and nestling period are known and accurate and detailed long-term data sets are available.

Materials and Methods

Data collection

We used 31 years of data (1973 to 2003) from a long-term study of a population of great tits (*Parus major*) on the Hoge Veluwe (The Netherlands). Great tits are small, hole-nesting passerines. The study area covers a mixed woodland of 171ha (1ha = 10^4 m^2) with about 400 nest boxes, which are readily accepted by great tits. Nest boxes were checked weekly to determine clutch size and laying date, i.e., the date when the first egg of a clutch was laid. Since successive eggs are laid at one-day intervals and clutch size usually exceeds six eggs weekly checks give accurate laying dates. All nestlings were ringed with uniquely numbered aluminum rings when they were seven days old. Almost all adults were caught during chick feeding and ringed with aluminum and color rings which allowed their identification. From 1973 to 2003, 3410 first clutches of females of known identity could be included in the analysis. Laying dates of the first egg of a clutch, times from an arbitrary starting point until a female lays her first egg, are 'time to event' data and were analyzed with a proportional hazards model.

Data analysis – Proportional hazards model

The proportional hazards model summarizes the ‘times to event’ as a baseline hazard multiplied by the effects of a set of variables (Cox 1972). The hazard rate at time *t* is described by

$$h(t,z_1,\ldots,z_n) = h_0(t) * \exp(z_1(t)\beta_1+, \ldots, +z_n(t)\beta_n)$$

where $h_0(t)$ is an arbitrary, unspecified baseline hazard rate, $z_1(t)$, …, $z_n(t)$ a set of n independent variables, that can be fixed or time-dependent, and β_1, …, β_2 the corresponding regression coefficients. Using a partial maximum likelihood method allows to estimate regression coefficients without knowing the baseline hazard function (*cf* Cox 1972).

The scale on which the times to egg laying are measured can be set arbitrarily, as long as no failures occur before the zero-point of the scale. In our standard protocol, laying dates were reported in Aprildays with 1st April being Aprilday 1 and 1st May correspondingly Aprilday 31. For reasons of convenience, we therefore measured the times to egg-laying on the same time scale.

It is known that laying date is influenced by age (Blums *et al*. 1997; Perdeck and Cave 1992; Robertson and Rendell 2001), temperature (Meijer *et al*. 1999; Perrins and McCleery 1989; van Balen 1973) and also day length (Lambrechts *et al*. 1997), and, therefore, these parameters were included into our model. Age was included as a categorical variable separating first-year breeders and older birds. An individual's age was either known because it was ringed as nestling in the study area, or by the coloration of the primary coverts (Glutz von Blotzheim 1993). Daily mean temperature and day length were obtained from the weather station of the Royal Netherlands Meteorological Institute at De Bilt (the Netherlands) from the institute's website (http://www.knmi.nl/ voorl/ kd/ lijsten/ daggem/ etmgeg_downl.cgi? language =eng). Both were included into the model as a non-categorical, time-dependent variable. Daily minimum or maximum temperatures are closely correlated with daily mean temperature and using them did not give quantitative or qualitative different results (not presented).

To avoid the problems related to measuring temperature over a fixed period, we used an iterative procedure based on a 'linear predictor' (Bush and Mosteller 1955) to calculate a temperature variable (λ) that included current temperature as well as earlier temperatures according to the following formula:

$$\lambda(t) = \text{daily mean temperature}(t)*\alpha + \lambda(t\text{-}1)*(\alpha)$$

where λ is the temperature variable and α is the ‘weighing factor’ and indicates the weighing of the current temperature relative to previous temperatures. A small α means that all but the most recent temperatures are weighed almost equally high, while a large α places a high weighing on the most recent temperature. Fig. 7.1 shows daily mean temperatures for two different years and the values of temperature variables illustrating the effect of differing α’s. For an α of 0.5 the values are similar to the daily mean temperature while the values based on an α of 0.04 are considerably smoothed. To avoid weighing an unusual warm or cold starting-day too heavily, when α is small, the average winter temperature (December until February) for the period 1972/73 to 2002/3 was used as $\lambda(0)$ for all α's and years. We made a series of models with α ranging from 0.01 to 0.05 in 0.01 steps and from 0.05 to 0.95 in 0.05 steps. For our model, the specific hazard rate could thus be described as:

$$h(t,A,\lambda,D) = h_0(t)\exp(\beta_A * A + \beta_\lambda * \lambda(t) + \beta_{\lambda D} * \lambda(t) * D(t))$$

where t is Aprilday, h_0 the baseline hazard, A age (a categorical variable indicating first-year (=0) or older (=1)), λ the iteratively calculated temperature measure and D the day length measured in hours. Day length was only fitted in interaction with temperature because it is (almost) linearly correlated with date.

Data analysis – model selection procedure and validation

The 'coxph' procedure in R 1.9.0 (R Development Core Team 2004; Venables and Ripley 2002) was used to calculate parameter estimates and Likelihood ratios using Efron-approximation for handling ties (Efron 1977). The overall fit and significance of the model was judged using the Likelihood ratio test and significance of variables using *t*-scores.

To compare the results with the observed laying dates and thereby validate the model internally we predicted individual laying dates of all females and then calculated the annual mean laying date and its standard deviation. To calculate the baseline survivor function the 'survfit' procedure in R 1.9.0 was used with all variable values set to 0. The survivor function describes the proportions of individuals having survived until time t. Assuming a discrete proportional hazards model the baseline hazard rate $h_0(t)$ was calculated from the baseline survivor function $F_0(t)$ as the ratio of the number of 'failures' (i.e. number of females started laying) during day t over the number of individuals having 'survived' (i.e. did not start laying) until day t, according to the formula

$$h_0(t) = \frac{F_0(t) - F_0(t+1)}{F_0(t)}$$

For every single female, the hazard rate for all days was calculated based on the value of the baseline hazard rate, multiplied by the effects of her specific values for the significant fixed variables, the temperatures of the year in which she bred and the interaction between temperature and day length. Then a laying date was simulated, as described in the following. Starting at day 1, a random number from a uniform distribution ranging from 0 to 1 was drawn (using the R random number generator). If the random probability was larger than the hazard rate for that given day, the female did not start laying on that day and the procedure repeated by using the hazard rate for the following day and drawing a new random probability. This was repeated until the random number was smaller than the hazard rate and that day was taken as the simulated laying date. This simulation was repeated 1000 times for every female. Annual means and standard deviations based on the first simulation run of the females breeding in the specific year were calculated and this was repeated for all other simulation runs. The obtained means and standard deviations, one per simulation run and year, were then averaged giving the predicted annual means and standard deviations. The simulation was performed using R 1.9.0.

Results

The best model describing the laying dates of great tits in our population included a $\lambda(t)$ based on $\alpha = 0.04$ (Tab. 7.1). Furthermore, the regression coefficients indicate that older females lay earlier, that females lay earlier under higher temperatures and that the influence of temperature decreases with increasing day length (Tab. 7.1).

Using the coefficients obtained and the baseline hazard individual survivor curves were calculated. These were used to simulate individual laying dates and predict annual population mean laying dates and standard deviations (Fig. 7.2). Predicted mean laying dates correlated well with observed mean laying dates ($r = 0.85$, df = 29, $p < 0.001$) (Fig. 7.2a) and also the predicted standard deviations were correlated with the observed standard deviations ($r = 0.57$, df = 29, $p < 0.001$) (Fig. 7.2b).

Discussion

As expected, egg laying dates of great tits were statistically significantly influenced by spring temperature and age of the female. First-year females lay later than older females (*cf* Blums *et al*. 1997; Robertson and Rendell 2001). Temperatures had a strong effect on laying dates. Higher temperatures increased the hazard rate and led thus to earlier laying (*cf* Meijer *et al*. 1999; Perrins and McCleery 1989; van Balen 1973). The α of 0.04 indicates that current temperatures play a less important role

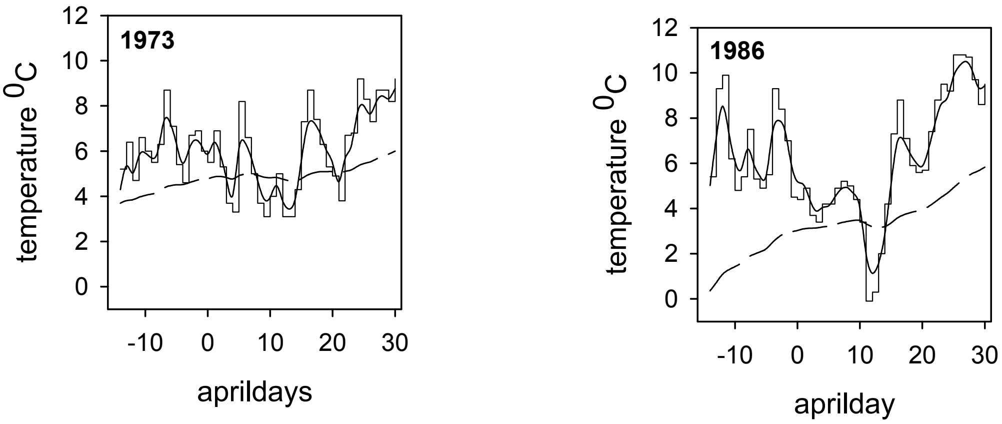

Fig. 7.1 Daily mean temperatures (step line), and two temperature measures (λ) used in the analysis (α = 0.04 dashed line; α =0.5 solid line) at De Bilt (KNMI) from 15 March (Aprilday -16) until 30 April (Aprilday 30). For details about calculation of λ see Methods. Temperature data from Royal Dutch Meteorological Institute (KNMI).

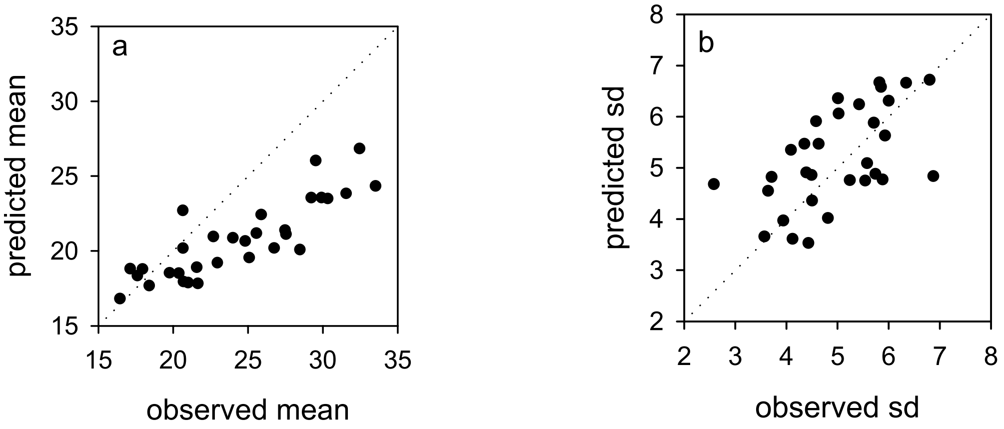

Fig. 7.2 Comparison between observed and predicted annual mean (a) and standard deviation (b) of laying dates. Dotted lines indicate the 1-to-1 relationship. Predicted means (a) corresponded well with observed means (r = 0.85, p < 0.0001) and the same was true for predicted (b) and observed standard deviations (r = 0.57, p = 0.01). For further details of predicting means and standard deviations see Methods.

than the long-term temperature trend. Fig. 7.1 shows, however, that short-term temperature signals are still visible. Interestingly, we also found a significant effect of day length in interaction with temperature. It is known that day length influences laying dates (Lambrechts *et al*. 1997) by inducing slow gonadal growth (Wingfield *et al*. 1996). However, this acts in a more general way by setting the 'window of reproduction' (Visser and Lambrechts 1999). The effect of temperature was influenced by day length in a way that high temperatures under short day length result in a higher hazard rate than the same temperatures under long day length, i.e. late in the season. This means that a period of warm weather early in spring will on average induce more individuals to start egg laying that warm weather in late spring. On the other hand, cold weather in late spring reduces the hazard less than cold weather in early spring as day length becomes an overriding cue late in the season (Lambrechts and Perret 2000). It was not possible to include day length as a main effect into the model since in our data set day length was almost perfectly correlated with date and did not show any variation between years. As a result, the effect of day length is contained in the baseline hazard. When we added a small random noise to our day length data to introduce variation between years it was possible to include the main effect of day length and the baseline hazard became smaller while the other coefficients changed only marginally (analysis not shown). Earlier descriptive models of laying dates in birds were not able to show any effect of day length on egg laying dates because of the limits of the statistical models used.

The phenological event analyzed in this paper is laying date, and we have used the date of first egg laying as the event analyzed. However, the ‘decision’ to lay has been taken when egg formation started. The duration of egg formation in great tits is short but may be considerably longer in larger bird species. In mammals gestation period lasts weeks or months. As it is likely that environmental conditions during egg formation or gestation play a different role than the conditions before the ‘decision’ is made to start reproducing this time-lag should be taken into account when describing breeding times by analyzing the period up to the start of egg formation or of embryonic development in mammals.

Our proportional hazards model gives the probability per day that egg laying occurs, given that it has not occurred yet. Using these rates, it is possible to predict the number/proportion of females starting egg laying on a given day and thereby to model annual distributions of laying dates respectively the spread in laying dates. This spread, expressed as standard deviations, correlated well with the observed standard deviations (Fig. 7.2b). This means that the probabilistic nature of the model was able incorporate the effects of other factors, like individual condition (Hasselqvist *et al*. 2001), food availability (e.g. Nager *et al*. 1997; Nilsson 1994) and a genetic component (Sheldon *et al*. 2003; van Noordwijk *et al*. 1981), that also influence laying date. Although it would be possible to include them in the model, using the distribution of the characters as we have observed them over the past years, this would only be of limited use for future predictions. While accurate enough predictions of temperatures of the next hundred years (Houghton *et al*. 2001) are available predictions of condition or genetic composition of populations are absent. But as the predicted standard deviations are correlated with the observed spread in laying dates our model allows quantitative predictions whether the distribution of laying dates will become in general more narrow or wider.

The proposed descriptive model is not based on temperatures measured over a fixed period and was able to predict the spread of laying dates of great tits. Flexible models such as this, which allow several fixed and time-dependent variables and even their interactions to be included and describes the probability of individual phenological events to occur, will prove to be indispensable tools in making valuable predictions about future changes in the phenology of a variety of species.

Tab. 7.1 The log Likelihood and regression coefficients (β) for the five best models. λ's (temperature measures) were calculated using the given α's, see text for further details. The hazard is calculated by multiplying the baseline hazard with the variables times their regression coefficients. A coefficient larger than 1 consequently means that a variable has a positive effect on the hazard, i.e., it increases the probability to start egg laying. As can be seen older females (age=1) lay earlier and higher temperature also lead to earlier laying. The effect of day length is that per experienced unit of $\lambda(t)$ and per hour lengthening of the day the probability per day to start laying an egg is almost halved (multiplication factor 0.52). However, every experienced unit of $\lambda(t)$ alone increases this probability much higher (multiplication factor 24100.79). *** $p<0.001$, ** $p<0.01$, * $p<0.05$

α	log Likelihood	coefficients Exp(β_{age})	 Exp(β_{λ})	 Exp($\beta_{\lambda * D}$)
Null model	- 23543			
0.02	- 24091 ***	1.35***	52260.70***	0.49***
0.03	- 24253 ***	1.38***	50161.35***	0.49***
0.04	**- 24325 *****	**1.41*****	**24100.79*****	**0.52*****
0.05	- 24320 ***	1.43***	8038.52***	0.56***
0.10	- 24086 ***	1.44***	315.45***	0.69***

Acknowledgments

J. H. van van Balen kept the long-term study on the Hoge Veluwe going for many years. J. Visser carefully managed the database and many other people helped with fieldwork. The ‘Stichting Het Nationale Park De Hoge Veluwe’ kindly allowed us to conduct our research on their property. K. Lessells and two anonymous referees made helpful comments on the manuscript.

Chapter 8

Shifts in caterpillar biomass phenology due to climate change and its impact on the breeding biology of an insectivorous bird

Marcel E. Visser, Leonard J. M. Holleman & Phillip Gienapp

Oecologia, *accepted*

Summary

Timing of reproduction has major fitness consequences, which can only be understood when the phenology of the food for the offspring is quantified. For insectivorous birds, like great tits (*Parus major*), synchronisation of their offspring needs and abundance of caterpillars is the main selection pressure. We measured caterpillar biomass over a 20-year period and show that annual peak date is correlated with temperatures from 8 March to 17 May. We use this descriptive model to predict shifts in caterpillar peak for 2005-2100, using an IPCC climate scenario. Laying dates also correlate with temperatures, but over an earlier period (16 March - 20 April). However, as we would predict from a reliable cue used by birds to time their reproduction, also the food peak correlates with these temperatures. Moreover, the slopes of the phenology of the birds and caterpillar biomass, when regressed against the temperatures in this earlier period, do not differ. The major difference is that due to climate change, the relationship between the timing of the food peak and the temperatures over the 16 March – 20 April period is changing, while this is not so for great tit laying dates. As a consequence, the synchrony between offspring needs and the caterpillar biomass has been disrupted in the recent warm decades. This may have severe consequences as we show that both the number of fledglings as well as their fledging weight is affected by this synchrony. We use a predictive model for the great tit laying dates to calculate the bird phenology for 2005-2100, using the same IPCC scenario. The birds will also start breeding earlier and this advancement is predicted to be at the same rate as the advancement of the food peak, and hence they will not reduce the amount of the current mistiming of about 10 days.

Introduction

Timing of reproduction is a life-history trait with important fitness consequences (e.g. Daan *et al*. 1990; Klomp 1970; Perrins 1970; Verhulst *et al*. 1995) because for many species there is only a short time-window in the annual cycle where conditions are sufficiently good so that they can successfully reproduce. This time-window is often set by the abundance of food necessary to raise offspring (Lack 1950; Martin 1987). As the seasonal pattern in the environmental conditions varies from year to year, also the optimal time-window for reproduction varies annually. In accordance with this, timing of reproduction shows corresponding within individual year-to-year variation caused by phenotypic plasticity. To understand the selection acting on the timing of reproduction we need to identify the environmental factors, e.g. prey phenology, which determine the optimal time window for reproduction. Only then we can understand why, in an ultimate sense, the animals reproduce early in one year and late in another, and why within a year late reproducing animals do better or worse than early reproducing conspecifics. What is more, it is only the only avenue to a more predictive study on the impact of climate change on the reproductive output of animals.

There is ample evidence that climate change has led to changes in the phenology of many species (e.g. Beebee 1995; Brown *et al*. 1999; Crick *et al*. 1997; Parmesan *et al*. 1999; Penuelas *et al*. 2002). Parmesan and Yohe (2003) summarized the literature and concluded that there is '… 'very high confidence' that climate change is already affecting living systems.'. It is however unclear how we should interpret this. If a species has shifted 10 days in 20 years, is that positive (adapting to changing conditions) or negative (climate change is affecting living systems)? What we need is a yardstick: how much should a species shift given the change in temperatures and other climate variables (Visser *et al*. 2004)? The yardstick is of course the shift in the moment in the year the environment is most suitable for reproduction; if this moment has shifted 10 days as well, the focus species is doing very well, if this moment has shifted 20 days, or not at all, then obviously climate change has a negative effect.

In many taxa the main variable determining reproductive success is the abundance of prey items (e.g. Dias and Blondel 1996; Durant *et al*. 2003; Pearce-Higgins and Yalden 2004; van Noordwijk *et al*. 1995; Verboven *et al*. 2001) affecting also the workload of the adults (Thomas *et al*. 2001). Yet, there are only a limited number of systems where data is collected on the annual cycle of a species and on the temporal variation of its prey abundance. We will focus on one of the more well know systems of great tits (*Parus major*) relying on caterpillars in oaks as the main food source for their nestlings (e.g. Naef-Daenzer *et al*. 2000; Perrins 1991). In this system, the caterpillar prey appears after the trees have their bud burst and disappears when the caterpillars are fully-grown and pupate in the soil and are thus no longer available. Hence, there is only a short period of ample food, and these insectivorous birds need to synchronise their breeding in such a way that the time of the maximal need of their offspring for feed coincides with the time of maximal food abundance.

The phenology of caterpillars in trees is generally measured indirectly by collecting caterpillar droppings with so-called frass-nets placed beneath a tress (e.g. Fischbacher *et al*. 1998; Verboven *et al*. 2001; Visser *et al*. 1998). The term *frass* stems from a mistranslation of the German word *Fraß* meaning 'amount eaten away (from a certain food) rather than dropping. The biomass of the caterpillars is calculated from the mass of the droppings collect in the nets corrected for temperature (Tinbergen and Dietz 1994). Although this method has been validated (Fischbacher *et al*. 1998; Liebhold and Elkinton 1988a; Liebhold and Elkinton 1988b; Zandt 1994) it may be possible that weather conditions (wind, rain) and properties of the tree (height, size) introduce systematic errors. To validate the established relationship further and possibly to correct for the mentioned effects we compared caterpillar biomass calculated from collected droppings and from direct branch samples from trees.

We have reported on the impact of climate change on the synchronisation of great tit laying dates and caterpillar abundance in previous publications (Visser *et al*. 2004; Visser *et al*. 1998) where we showed that caterpillar peak biomass date has shifted forward over the past decades but that there has been no advancement of laying dates of great tits. Here, we will take our studies on this system a step further and predict how forecasted climate change will impact the phenology of both caterpillar biomass and bird reproduction in the next 100 years.

Our objectives are to describe the caterpillar biomass changes throughout a season, to construct a descriptive model of the peak date and width of the biomass curve and to extrapolate this model under IPCC scenarios to predict future caterpillar biomass distributions. We will also compare the temperature sensitivity of the great tit laying dates and of the food peak to asses the impact of changing temperatures on the synchrony of the birds with their main food. Furthermore, we correlate a number of great tit fitness parameters to the biomass distribution to demonstrate the importance of this synchronisation. Finally, we use a descriptive model to predict the laying dates and from these predicted caterpillar and bird phenologies estimate the impact of the changing food phenology on the birds' breeding success.

Methods

Fieldwork

Fieldwork was carried out in the Hoge Veluwe, one of our long-term (1955-2004) study areas. This study area consists of 171 ha mixed woodland on poor sandy soils, dominated by oak and pine with about 400 nest boxes. Some additional comparisons are made from data from another long-term study area, Oosterhout, a smaller, isolated woodland (11 ha) on rich clay soils near the river Waal, in which 120 nest boxes are placed.

Frass nets (a cheese cloth of 0.25 m^2 in a metal frame, with a weight hung from the centre of the net, see Fig. 3a in Tinbergen, 1960) were put up under oak trees (*Quercus robur*). Two of these nets are placed under a tree (about 1-1.5 meter from the stem) and every 3-4 days all caterpillar droppings are collected, dried at 60° C for 24 hours, sorted (i.e. all debris is removed), weighted and from this the caterpillar biomass is calculated (see below).

From 1993 – 2004 frass nets were placed under a varying number of trees on the Hoge Veluwe but there is fixed set of 7 oak trees, which were sampled throughout these years. Furthermore, from 1985 to 1992 frass nets were used in the same area by J.M. Tinbergen, and we use these data on the phenology of the caterpillar biomass (see Verboven *et al*. 2001b) to obtain a 20 year time series. For Oosterhout frass net data are available for 1957-1967 (van Balen 1973), 1986-1988 (Zandt 1994), 1996 & 2001-2004 (own data).

During three years (1994-1996) we took branch samples using a sky lift at the Hoge Veluwe. Branches of about 2 meter were first put into a large plastic bag, then cut from the trees and later (within 36 hours) sorted for insects. The length of the branches were measured to convert the total biomass to biomass / meter branch. We also measured height and stem diameter of these trees.

To analyse the spatial variation in the time of maximal caterpillar biomass we correlated the peak date of caterpillar biomass with bud burst date of the tree the nets were place under. Tree leave phenology was scored twice a week and we use the date at which the leaves protruded from their buds in the canopy of the tree as their bud burst date (Visser and Holleman 2001).

Data on laying dates and clutch sizes of great tits in the various populations was collected by weekly checks of the nest boxes. Laying dates are calculated back assuming that one egg per day is laid. All analyses of laying dates and clutch sizes were restricted to first clutches. Adults are caught during chick feeding and identified by their aluminium and colour-or ringed if not previously caught. All nestlings are banded at an age of seven days and weighed at 15 days, which is a good approximation of their fledging weight. Laying dates are given as April-days (1 April is April-day 1, 24 May is April-day 54).

Temperature and other climatic data for 'De Bilt', the main weather station of the Royal Dutch Meteorological Institute (KNMI), were used and obtained from the KNMI's website.

Caterpillar biomass distribution

Caterpillar biomass was calculated from the raw frass data using the formula of Tinbergen & Dietz (1994), which essentially corrects the amount of frass produced per unit biomass caterpillar for the temperature. Biomass is expressed as gram caterpillar biomass per m^2 per day, and thus ignores the height of the tree under which the frass nets are placed. We therefore include height in the subsequent analysis where we correlate biomass as obtained form branch samples with the biomass as calculated from the frass nets. Furthermore, we analyse only the data from our set of standard trees (see above) and hence the there is no between year variation in tree height for the samples analysed.

The distribution of the biomass of caterpillars over the season cannot be fitted with a normal or any other distribution and hence we use three parameters to describe the food peak: the peak date is the date at which the largest biomass is recorded (also expressed in April-days), the peak height is the largest biomass recorded in a season (g biomass m^{-2} day^{-1}) and the peak width is the number of days where the biomass is above 1 g m^{-2} day^{-1}. The date at which the biomass rises above and falls below this threshold value is estimated by fitting a cubic polynomial through the data points. In three years sampling either started after or ended before this threshold was reached and hence for these years we cannot estimate the peak width. Peak height and width depend on the trees sampled and hence were only calculated using the restricted set of trees that were sampled every year (1993-

2004). The area peak date is a much less tree dependent measure and hence we used also the 1985-1992 period from Verboven *et al*. (2001).

When correlating the biomass as recorded from the branch samples with the biomass as calculated from the frass samples we used a log transformation to normalize the data. We also used a repeated measures design with sample date repeated over trees to take into account that the same trees were sampled at different sampling dates (MIXED procedure SAS). In the analysis on the spatial variation in caterpillar peak date we used ANOVA and ANCOVA models (GLM procedure SAS) with peak caterpillar biomass date as dependent variable.

We constructed a descriptive model for the peak biomass date (n = 19 years, excl. 1991 as this was a year with a late frost which damaged the leaves of the oak trees, *cf* Visser *et al*., 1998) by regressing this peak dates against the mean temperature in all periods of at least 10 day long between the 1 January and the 31 May (in total 10153 different periods) and selected the period which had the highest correlation with peak date. Finally, we use the descriptive model and the predicted temperatures from an IPCC-SRES model, with an intermediate increase in temperature (SRES-B2) (Esch 2005), and use these to predict the caterpillar biomass distributions for 2005-2100.

Great tit fitness parameters

– Synchrony with food peak

The synchronisation between the birds breeding time and the timing of the food peak was defined as the difference between hatching date plus nine days and peak date (synchrony = hatching date + 9 - peak date) because great tit chicks grow fastest at an age of nine days and food demands are then highest (Gebhardt-Henrich 1990; Keller and van Noordwijk 1994).

– Number and weight of fledglings

As part of our standard field protocol great tit chicks are weighed to the nearest 0.1 g when they are 15 days old and the number of fledged chicks is recorded by checking the nest for dead chicks after fledging. For all broods in the years 1985-2004 we calculated mean chick weight per brood and the number of fledged chicks. In some years clutch or brood size manipulations were carried out. Since these experiments possibly affected chick weight or number of fledged of a brood manipulated broods were excluded from this analysis. The relationship between mean chick weight respectively number of fledged with synchrony with the caterpillar peak was analysed with GLMs in R 2.0.1 (R Development Core Team 2004). Full models included the following explanatory variables and all 2-way interactions: synchrony, synchrony^2, peak height (maximum amount of caterpillar biomass recorded), peak width (number of days with caterpillar biomass above 1 g m^{-2} day^{-1}) and brood size (number of hatched chicks). Identity link and normal error structure was used for analysing mean chick weight and log link and Poisson error structure for analysing number of fledged chicks. Minimum adequate models were selected using stepwise backward deletion beginning with interactions.

– Selection differentials

Selection differentials are defined as the covariance between trait and relative fitness and quantify the strength of selection on the trait (e.g. Lande & Arnold 1983). We used number of offspring found breeding in our population (termed recruits) as a fitness measure. Fitness was converted to relative fitness by dividing the number of recruits by the mean number of recruits in a given year. To analyse whether the strength of selection is related to synchrony of the population with the timing, the height or the width of the caterpillar peak we regressed the annual selection differentials against these variables and all 2-way interactions.

Predictive model of great tit laying dates

We used a Cox's proportional hazards model (Cox 1972) to build a descriptive model, based on the results of Chapter 7. This kind of model describes the probability that an individual female will start with egg laying as a function of an unspecified baseline hazard and a set of explanatory variables. The explanatory variables included in the model were female age (first year breeder or older), temperature (calculated via a 'linear predictor'), day length and the interaction between temperature and day length. To avoid very small values of the baseline hazard and possible arithmetic inaccuracies temperature and day length were rescaled by subtracting the corresponding minimum value.

– Temperature and day length data

We used an IPCC-SRES model, with an intermediate increase in temperature (SRES-B2) (Esch 2005) as temperature data for our prediction. The predictions of this climate model are available at a temporal resolution of one day and a spatial resolution of 50*50 km, in our case around Arnhem (the Netherlands). Day length data for every fifth year were obtained from the US Naval Office via its website (http://aa.usno.navy.mil/data/docs/RS_OneYear.html).

– Simulation

The baseline hazard is only specified for days on which a laying date was observed. To be able to predict also laying dates outside this 'time window' we fitted an exponential function to the baseline hazard estimated by the proportional hazards model ($r^2 = 0.86$, $p < 0.0001$) and used it for extrapolating the baseline hazard. We calculated the hazard for all days from 1 Jan to 31 May for all years from 2005 to 2100 using the temperature and day length data, assuming that 43% of the females are older than one year (which is the average percentage for 1973-2004) and the (extrapolated) baseline hazard. Using this hazard we simulated egg-laying dates of 100 females per year (see Chapter 7 for details). From the obtained distribution of laying dates the annual mean simulated laying dates were calculated.

Results

Calibration of frass measurements

In three years (1994 to 1996) branches were sampled from trees that also had caterpillar frass nets ($n = 60$ tree-samples days). The main species of caterpillars found were winter moth (*Operophtera brumata*) and oak leaf roller (*Tortrix virirdana*). The two measures of caterpillar biomass were highly correlated (log biomass from branch samples versus log biomass calculated from frass in a repeated measures analysis of sample dates over trees: $F_{1,13} = 19.35$, $p < 0.001$). There were no additional significant explanatory variables, neither of tree characteristics (height and width at breast height) nor of sample date characteristics (daily precipitation amount (*cf* Zandt 1994), sunshine duration, daily mean wind speed and daily mean temperature (note that the biomass as calculated from frass was already corrected for temperature according to Tinbergen & Dietz, 1994); all p-values > 0.30 in a step wise analysis). We therefore used the biomass as calculated from the frass samples (using the Tinbergen & Dietz 1994 formula) without any further corrections as a measure for the biomass of caterpillars in the trees.

Describing the caterpillar biomass changes throughout a season

In general, the peak in caterpillar biomass is very narrow, the number of days with a biomass above 1 g m^{-2} day^{-1} is on average 24.3 days (n = 9, range 19-33.5). There is a correlation between the height of the biomass peak and the timing: in late years the peak is higher (Pearson's $r = 0.76$, $p = 0.004$). There is no correlation between the peak width and either the height or the timing of the peak (both $p > 0.13$).

Peak dates vary from year-to-year with a range of about 3 weeks, roughly equal to the width of the peak. There is a significant advancement of the peak date over the years 1985-2004 of 0.74 days a year ($F_{1,17} = 13.15$, $p = 0.002$ excl. 1991; $F_{1,18} = 10.88$, $p = 0.004$ incl. 1991; Fig. 8.1a, *cf* Visser *et al*, 1998; Visser *et al.*, 2005). There is no significant change in the width or the height of the biomass peak over the years (both $p > 0.40$).

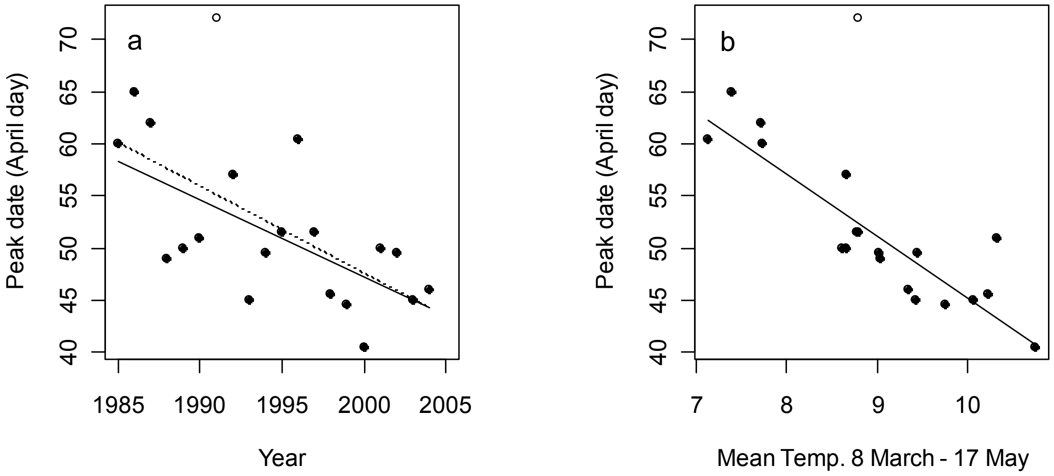

Fig. 8.1 Phenology of caterpillar biomass at the Hoge Veluwe (1985-2004), in 1991 (open dot) a late frost damaged the fresh oak leaves resulting in an extremely late peak date. (a) Advancement of caterpillar peak date over time (broken line for all 20 years, solid line with 1991 is excluded), and (b) caterpillar peak date (excl. 1991) versus the mean temperature from 8 March to 17 May (the period that correlates best with the caterpillar phenology).

There is clear spatial variation is the timing of the biomass peak of oaks: some sites within the Hoge Veluwe area are consistently early while others are late (site: $F_{6,65} = 17.52$, $p < 0.001$; year: $F_{11,65} = 22.35$, $p < 0.001$; *cf* Grieco *et al.* 2002). When site is replaced by characteristics of the sampled trees (retaining year as a factor in the model, $F_{11,70} = 12.26$, $p < 0.001$), then height and width of the tree are not significant (both $p > 0.28$) but there is a clear effect of the date of bud burst

($F_{1,70}$ = 57.73, $p < 0.001$): for every day the bud burst is later, the peak biomass date is 0.47 days later.

Construct descriptive models of food phenology

Of all periods tested, the correlation between peak biomass data and temperature was the highest for the period of the 8th March to the 17th of May (r^2 = 0.78; Fig. 8.1b; *cf* Visser *et al.* 1998, but note different temperature period). The width of the peak (i.e. the no. of days between caterpillar biomass above/below 1 g m^{-2} day^{-1}) for which we have 9 years of data (1994, 1996, 1998-2004) also correlates with temperature. Here, the period with the highest correlation is 16th of March to the 17th of May (r^2 = 0.66).

There is also no interaction between site and temperature (over the best fitting period; $F_{6,69}$ = 0.42, p = 0.86), but sites differ in timing ($F_{6,75}$ = 12.14, $p < 0.001$, with temperature in the analysis: $F_{1,75}$ = 141.44, $p < 0.001$), the earliest and the latest site differing by 9 days.

Predict future caterpillar biomass distributions

Over the period 2005 to 2100, caterpillar peak dates are predicted from the linear regression against spring temperature and an IPCC-SRES temperature scenario. Peak dates will advance by 0.20 days per year, which will add up to a total advancement of 18 days (linear regression: $F_{1,94}$ = 50.0, $p < 0.001$; Fig. 8.2b). The predicted width of the caterpillar peak will significantly decrease (linear regression: $b = -0.13 \pm 0.020$, $F_{1,94}$ = 40.0, $p < 0.001$).

Correlation between great tit fitness parameters and biomass distribution

– Laying date

When we analyse the relationship between the timing of the food peak and of the start of egg laying by the birds within an area we find a correlation: laying is earlier in years with an early food peak ($F_{1,17}$ = 4.52, p = 0.048). Laying date however only advances with 0.3 days for every day the food peak is earlier.

Verboven *et al.* (2001) have argued that the laying dates of great tits in the Hoge Veluwe and the Oosterhout areas differ in the way they correlate with the caterpillar biomass peak dates but we do not find this (effect of peak date on laying date: $F_{1,36}$ = 16.91, $p < 0.001$ (combining the Oosterhout and Hoge Veluwe data), no effect of area*peak date interaction ($F_{1,34}$ = 0.81, p = 0.37) nor of area ($F_{1,35}$ = 0.01,

$p = 0.98$); note that the food peak data for Oosterhout partly come from much earlier years (1957-1967) than for the Hoge Veluwe).

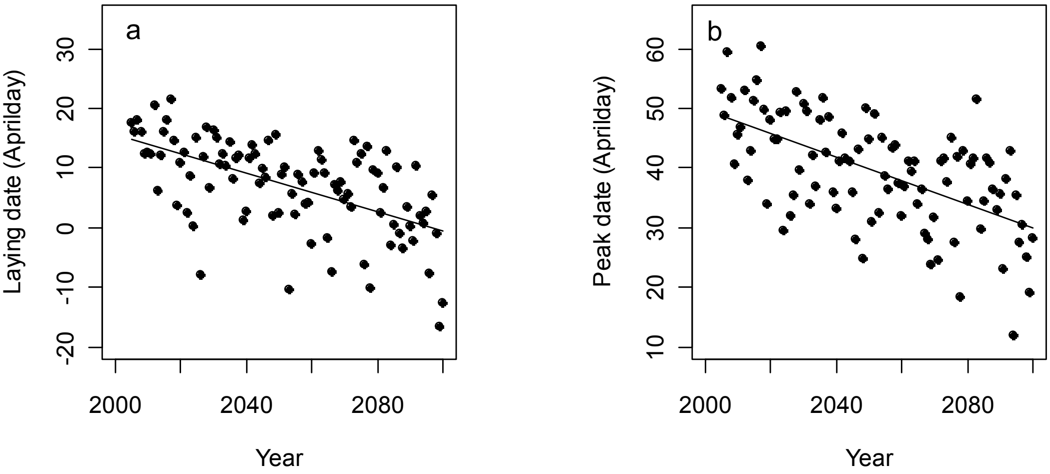

Fig. 8.2 Predicted phenology based on an IPCC-SRES scenario for the Hoge Veluwe area. (a) Annual means of predicted laying dates, and (b) predicted dates when the caterpillar biomass reaches its maximum (peak date).

– Synchrony with food peak

The phenology of both the birds, their laying date, and of the food peak depend on temperature, but temperatures of different periods correlate best with these phenological events. However, if the temperatures used by the birds as cues to time their reproduction we would expect these temperatures to predict the time of optimal conditions to raise their offspring, i.e. caterpillar peak (Visser *et al.* 2004). We therefore used the temperatures for the best fitting period for the laying dates (16 March – 20 April, $r^2 = 0.61$ for 1973-2004) and determined how well these temperatures predict the food peak. When we restricted the analysis of the relationship between mean temperature and annual mean laying dates to the years, for which we have measured caterpillar peak dates, this relationship still holds ($r^2 = 0.42$, $F_{1,17} = 12.22$, $p = 0.003$, no effect of year $F_{1,16} = 0.26$, $p = 0.62$ or year*temperature interaction: $F_{1,15} = 0.37$, $p = 0.55$). This mean temperature over 16 March – 20 April also correlates with the timing of the food peak ($r^2 = 0.66$, $F_{1,16} = 10.38$, $p = 0.005$) but there is an additional effect of year ($F_{1,16} = 11.18$, $p = 0.004$, no significant year*temperature interaction $F_{1,15} = 1.26$, $p = 0.28$). The sensitivity of the timing of laying and of the food peak to the temperatures over the period 16 March-20 April is not significantly different (laying date: -3.34 ± 0.97 day $°C^{-1}$, food peak: -4.01 ± 1.25 day $°C^{-1}$, $F_{1,32} = 0.23$, $p = 0.63$). This means that the temperatures used by the birds as a cue are true predictors of the timing of the food peak and it is therefore meaningful for the birds to use them (Fig. 8.3a). However, the year effect on the food peak shows that the timing of the food peak is advancing: for the same mean temperature over the period 16 March-20 April the food peaks 0.57 days a year earlier.

We also have 19 years of data on the timing of the food peak of Oosterhout, but 11 of these are from an earlier period (1957-1967). We do find however also for this area that the laying dates and the food peak correlate well with the temperatures from 16 March to 20 April (Fig. 8.3b, temperature effect on laying date: $F_{1,17} = 107.30$, $p < 0.001$; temperature effect on food peak: $F_{1,17} = 16.17$, $p < 0.001$) and also here there is no significant difference in the temperature sensitivity of the two phenologial variables (laying date: -4.14 ± 0.40 day $°C^{-1}$, food peak: -4.22 ± 1.05 day $°C^{-1}$, $F_{1,32} = 0.01$, $p = 0.91$). But in contrast to the Hoge Veluwe, there is no significant year effect for the food peak ($F_{1,16} = 0.76$, $p = 0.40$). This may reflect actual differences between these areas but it is perhaps more likely that this is due to the large number of years from the earlier period for Oosterhout.

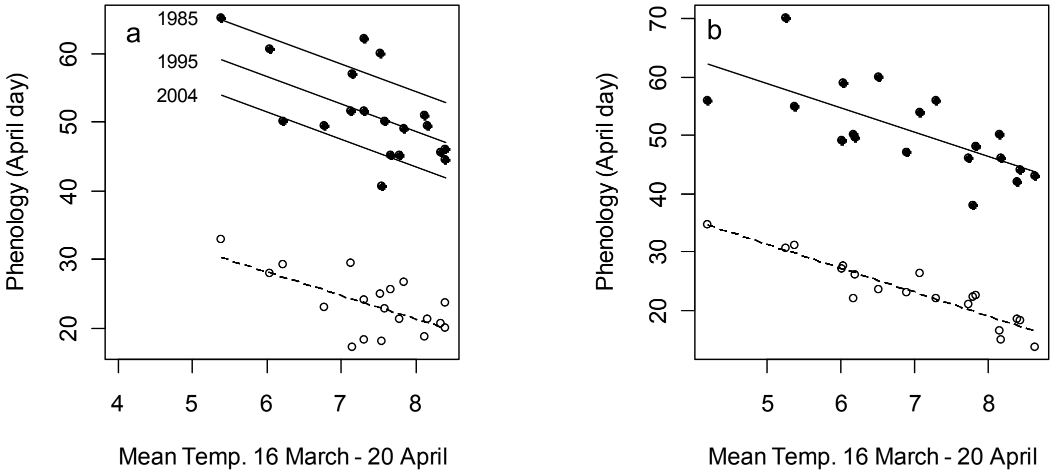

Fig. 8.3 Laying date (open symbols) and peak biomass date (closed symbols) for (a) Hoge Veluwe (1985-2004, excl. 1991) and (b) Oosterhout (1957-1967, 1986-1988, 1996, 2001-2004) against the mean temperature for 16 March – 20 April (period that best correlated with mean annual laying date). The dashed lines are the regression lines for the laying dates, the solid lines for the food peak. For the Hoge Veluwe there is a significant year effect on the food peak phenology, which is indicated with regression lines for 1985, 1995 and 2004. The difference in elevation for the 1985 and 2004 line is 10.9 days.

– *Clutch size*

There was no effect of the height of the peak caterpillar biomass on the mean annual clutch size ($F_{1,10} = 0.33$, $p = 0.58$) for the 12 years of the Hoge Veluwe (1993-2004) where we sampled the same trees for a number of years, making the peak height data comparable over the years. The relationship between laying date and clutch size varies significantly between years (interaction between laying date and year: $F_{18,2062} = 3.99$, $p < 0.001$ in a model with year and laying date, no significant quadratic terms) but the estimates for the these annual slopes do not correlate with the biomass peak date ($F_{1,17} = 3.46$, $p = 0.08$).

– *Reproductive success*

The number of fledged chicks was strongly determined by brood size (number hatched: $X^2{}_1 = 646.0$, $p < 0.001$) and by synchrony with the caterpillar peak (Fig. 8.4a): when corrected for brood size, fewer chicks fledged from broods raised before or after the food peak (synchrony^2: $X^2{}_1 = 97.0$, $p < 0.001$) and this relationship did not reach its maximum at a synchrony of 0 days but 1.7 days later, i.e. when the chicks are 10.7 days old (synchrony: $X^2{}_1 = 10.9$, $p < 0.001$).

Mean chick weight is significantly influenced by the synchrony with the caterpillar peak (Fig. 8.4b): chicks raised before or after the caterpillar peak are lighter (*cf* Verboven *et al.*, 2001). This effect is influenced by brood size and the height of the caterpillar peak (synchrony^2*number hatched: $F_{1,529} = 18.3$, $p < 0.001$; synchrony^2*peak height: $F_{1,529} = 5.3$, $p = 0.02$): the effect of synchrony on chick weight becomes weakened with an increasing maximum caterpillar biomass but becomes stronger with increasing brood size. The linear synchrony term of the model means that maximum chick weight is not reached at 0 days synchrony (i.e. chicks are 9 days at the date of the biomass peak) but 1.1 days later (when the chicks are 10.1 day old). This relationship is also affected by the width of the caterpillar peak, the wider the peak is the earlier becomes the date relative to the caterpillar peak when chick weight is maximal (synchrony*peak width: $F_{1,529} = 7.1$, $p = 0.007$; correcting for the brood size: $F_{1,529} = 66.0$, $p < 0.001$).

– *Selection differentials*

We could not find any relationship between width or height of the caterpillar peak or population mean synchrony and selection differentials (peak width: $F_{1,3} = 0.02$, $p = 0.90$; peak height: $F_{1,4} = 0.0008$, $p = 0.98$; mean synchrony: $F_{1,15} = 1.27$, $p = 0.28$).

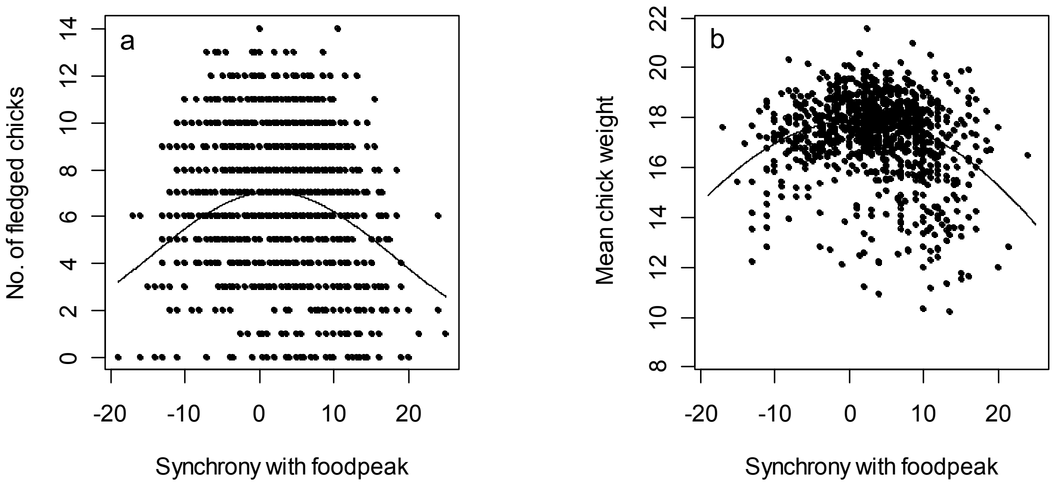

Fig. 8.4 Mean chick weight (a) and number of fledged chicks (b) of great tits on the Hoge Veluwe in relation to synchrony with caterpillar peak (synchrony = hatching date + 9 – peak date). In (a) fitted lines show relationship between number of fledged chicks and synchrony for different brood sizes, in (b) the fitted line shows the relationship between chick weight and synchrony for the average number of hatched chicks, peak width and peak height.

Estimate impact of the changing food phenology on breeding success

– Proportional hazards model

Our 'Cox's regression' model' explained a significant amount of variation in egg laying dates (Likelihood-ratio = 1549, df = 4, $p < 0.001$, see also Chapter 7). All included explanatory variables were highly significant (age: $b = 0.35$, $p < 0.001$; temperature: $b = 1.60$, $p < 0.001$, day length: $b = 5.03$, $p < 0.001$, temperature*day length: $b = -1.03$, $p < 0.001$). As found in various other studies (e.g. Perdeck & Cave 1992, Wheelwright & Schultz 1994, Robertson & Rendell 2001) older females laid earlier than young females as indicated by the positive regression coefficient. Both temperature, calculated as 'linear predictor', and day length had a positive effect on the probability that a females starts with egg laying (both regression coefficients are positive). The negative regression coefficient of the interaction between temperature and day length means that the effect of temperature decreases with increasing day length. Late in the season thus lower temperatures already trigger laying than in early in the season.

– Simulation

From 2005 until 2100 mean laying dates are predicted to advance by 0.16 days (± 0.02) per year (linear regression: $F_{1,94} = 47.1$, $p < 0.001$) (Fig. 8.2a). This amounts to a total advance of 15 days over the whole period. Simulated standard deviations show no significant time-trend (linear regression vs year: $F_{1,94} = 2.28$, $p = 0.13$). The predicted time trends for laying dates in birds and caterpillar phenology are however not significantly different (linear regression: year*species interaction: $F_{1,188} = 0.99$, $p = 0.32$).

Discussion

The optimal time for reproduction in the great tit is clearly set by the time the biomass of caterpillars peak: birds that have 10-day old chicks in their nest at the annual peak date in biomass fledge the most chicks (for their clutch size, Fig. 8.4a) and these are also the heaviest (Fig. 8.4b), increasing the chance that they will survive and breed (Verboven and Visser 1998). There is both spatial and annual variation in the time of maximal caterpillar biomass and hence the optimal time of breeding for the birds varies in space and time. The annual variation in the date of the biomass peak is as large as the width of the peak (about 3 weeks) emphasising the need for phenotypic plasticity.

One way the birds can cope with the spatial variation is to learn when best to breed at the place they have established themselves (Grieco *et al*. 2002). They cope with the temporal variation by phenotypic plasticity of their laying dates: the same individual lays at different times in different years. Obviously, the birds need to start reproduction quite some time before their chicks are 10 days, i.e. when they need to be synchronized with the food peak. They therefore use cues from their environment at the time of laying. We have shown that laying dates correlate very well with temperature (from 16 March – 20 April), that the food peak also correlates with these temperatures and that the phenology of the birds and the caterpillar biomass respond very similar to these temperatures. Hence, the birds seem to be able to respond adequately to the temporal variation in biomass peak date.

The observed mismatch between bird and caterpillar biomass phenology at the Hoge Veluwe seems to be at variance with the observation that both phenologies respond in the same way to temperature. The reason for this is that the timing of the food peak is also affected by temperatures after the 20 April (note that the best fitting period for the food peak is 8 March to the 17 of May, see above) and that these temperatures have also increased (Visser *et al*. 1998). It turns out that this is more or less additive so that the change of 1 °C in the earlier period still leads to a food peak that is 0.3 days earlier, but that the food peak has advanced because the temperatures in the later period have increased (see Fig. 8.3a). As the phenology of the food is now earlier for the same temperatures over the early period, but the phenology of the birds is not, the consequence is that the interval between laying and the food peak has become shorter. For a mean temperature of 7 °C over the period 16 March – 20 April, this interval was 37.5 days in 1985, but is only 26.6 days in 2004 (a shift of 10.9 days). Given a clutch of 10 eggs (= 10 days), 12 days of incubation and chicks of 10 days when they should synchronise with the food peak (=32 days), the result is that many of the birds have their offspring in the nest too late to profit from the short peak in caterpillar biomass. In the statistical model this leads to a significant year (as a continuous variable) effect in the analysis of the food peak.

The changed relationship between early spring temperature (16 March – 20 April) and caterpillar peak date will lead to selection for a changed reaction norm. However, there is no need for the birds to become more sensitive to temperature (the slope of the reaction norm) but just that they have to lay earlier over the whole range of temperatures (the intercept of the reaction norm). We have indeed detected selection on the reaction norm for the Hoge Veluwe population, but surprisingly the selection was stronger on the temperature sensitivity, i.e. the slope of the reaction norm, than on the laying date in the average environment, i.e. the intercept (Chapter 5). At present, we cannot explain this from the changes in the caterpillar biomass phenology as this has not become more temperature sensitive (no significant year*temperature interaction). We can only propose several different explanations. A time series of 20 years may not be long enough to detect such an interaction. For a meaningful description of reaction norms it is crucial to identify the correct explanatory variable, against which phenotypes are regressed. Mean temperatures are only a proxy for the real cues used by the birds (*cf* Chapter 7) and describing reaction norms using a different explanatory variable may give different results. Another explanation could be that there are other selection pressures that select for a steeper reaction norm, such as earlier settlements in warmer years, and hence an earlier competition for territories among the fledged offspring.

Great tits are facultative multi-brooders at the Hoge Veluwe but over the past 20 years the proportion of birds producing a second brood has strongly declined (Visser *et al*. 2003). A potential explanation for this decline would be a reduction in the width of the food peak; a narrow food peak means a short time-window for reproduction and hence fewer broods. However, we cannot demonstrate a reduction in the width of the food peak. Moreover, it is also likely that for second brood offspring caterpillars in oak are not the main food source (Verboven *et al*. 2001).

Annual mean laying dates are correlated to annual mean food peak dates, both on Oosterhout and the Hoge Veluwe. This is what is expected if the phenotypic plasticity in laying date is adaptive. Our results are in contrast to those of Verboven *et al*. (2001) who claim that there is such a relationship in Oosterhout and Marley Wood (UK), where birds are generally not muli-brooded, but not on the Hoge Veluwe and Vlieland (NL), where part of the birds use to be double brooded. A major problem with the analysis of Verboven *et al*. (2001) is however that they compare two areas for which they have data from the 50ties and 60ties (Oosterhout and Marley Wood) with two areas for which they have data from the 80ties and 90ties (Hoge Veluwe and Vlieland), and find a difference. Given the disrupted synchrony in the recent two decades this could well explain their results, rather than the incidence in second broods in these two pairs of areas. In our analysis, with more recent years for both areas, there is no longer a statistical difference between Oosterhout and the Hoge Veluwe.

We have shown that the current reaction norm of great tit laying date against temperature is no longer adaptive. Given that laying dates are heritable (Sheldon *et al*. 2003; van der Jeugd and McCleery 2002; van Noordwijk *et al*. 1981), that selection act on the genotypes rather than on an environmental variable which correlates both with fitness and laying date (Chapter 6), we expect a response to this selection and thereby a change in the reaction norm of the birds. We predicted laying dates and food phenology for 2005-2100 and found that both advance over the next 100 years (Fig. 8.2) and are predicted to do so at the same rate. However, this prediction is for a great tit population in which no micro-evolution occurs. Obviously, what is needed is a predicted rate of change in laying dates due to selection. Next, we should then compare the rate of change in ecological conditions with the rate of micro-evolution. In the absence of a change in reaction norm we predict that the current mistiming will not increase, which might mean that selection may occur fast enough for the synchrony between the birds and the caterpillar biomass phonologies to be restored, and thereby reducing the negative impact of global climate change.

Acknowledgements

We thank Jan Visser for maintaining the database, Ruben Smit for the measurements of the trees, Arie van Noordwijk and many students for their help with the branch sampling. We are grateful to the board of the National Park de Hoge Veluwe, to Barones van Boetzelaer van Oosterhout and the State Forestry Service in Vlieland for the permission to work in their woodlands.

Chapter 9

General Discussion

Fitness consequences of phenological synchrony between great tits and their prey

In many species, the exact timing of many life-history events, like migration or breeding, has strong fitness effects because there is often only a short time-window with optimal conditions. Caterpillars, mainly wintermoths (*Opheroptera brumata*) form the main food supply for nestlings in great tits and other small passerine birds. Great tit chicks raised before or after the period when caterpillar biomass fledge at lighter weight (Chapter 7, Verboven *et al*. 2001) and subsequent survival and recruitment probability is related to fledging weight (Naef-Daenzer *et al*. 2001; Verboven and Visser 1998). In the Hoge Veluwe long term-study area the phenology of caterpillar biomass has been monitored since 1985. Warmer springs have led to an advancement of caterpillar phenology by about two weeks while the birds have shifted their breeding time by only about three days over the same period. This has consequently led to selection for earlier breeding (Visser *et al*. 1998). The aim of this thesis was to explain why this great tit population was not able to track the advancement in the phenology of caterpillars, and in doing so enhancing our understanding of evolution in the wild.

Phenotypic changes in laying dates

Breeding time in birds is a highly plastic trait (Chapter 5; Brommer *et al*. 2005). Optimal timing of breeding can vary strongly from year-to-year and consequently adaptive phenotypic plasticity has evolved. However, the birds cannot respond directly to their caterpillar food supply. In order to have their chicks in the nest when caterpillar biomass peaks the birds have to lay their first eggs at a time when the caterpillars are only just hatching (inside the leaf buds). Consequently, the birds must use other cues that reliably predict caterpillar phenology. Although photoperiod plays a role in the timing of breeding in birds (Chapter 7; Lambrechts *et al*. 1997; Wingfield *et al*. 1996) it cannot account for year-to-year differences. In contrast, it has frequently been found that laying dates are related to temperature (e.g. Chapter 7; Meijer *et al*. 1999; Perrins and McCleery 1989). The caterpillar peak has shifted forward but this does not necessarily mean that the cues used by the birds have changed at a corresponding rate (see Fig. 1.2). Mean temperatures from 16 March to 20 April, which are correlated to mean laying dates, have increased significantly by about 0.05 °C per year since 1973. This temperature increase is expected to have led to advancement of only about 5.5 days caused by phenotypic plasticity. This is very similar to the observed non-significant advancement of annual mean laying dates of 5.4 days (Fig. 6.1a). Since caterpillar phenology has shifted at a faster rate (Chapter 8) this means that the relationship between caterpillar phenology and the cues the birds are using has changed. In the great tit population in Wytham Wood near Oxford (UK) a similar phenomenon has been observed, only in the opposite direction. In this population mean laying dates have advanced faster than the caterpillar phenology and this has led to reduced selection on breeding time, increased fledging success and a prolonged incubation period (Cresswell and McCleery 2003).

Selection on and heritability of phenotypic plasticity

From the changed relationship between caterpillar phenology and cues used by the birds it might be expected that there is not only selection for earlier laying but also selection for a different reaction norm of breeding time. Which will be stronger will however depend on whether the range of experienced environments (spring temperatures) has changed (Chapter 1 – Box 2). In Chapter 8 we showed that at the same temperature (from 16 March to 20 April) the caterpillar peak is nowadays 10 days earlier than 20 years ago but that the slope of the relationship between caterpillar peak and this temperature has not changed (Fig. 8.2a). This means that there is selection on the intercept of the reaction norm but not on the slope. In contrast to this it is shown in Chapter 5 that selection on the slope of the reaction norm increased over time; individuals that reacted more strongly to temperatures thus had higher reproductive success. However, when describing reaction norms it is crucial against which environmental variable phenotypes are regressed. As discussed in Chapter 7 using mean temperatures measured over a fixed period can be problematic. This is especially true if these temperatures have changed systematically with time, which is unfortunately true in this case.

There was also a strong phenotypic and genetic correlation between slope and intercept, which means that individuals breeding early in the average environment were also more responsive to temperatures (Chapter 5). Intercept and slope of the breeding time reaction norm may not be traits *per se* but simply an emerging property of genetic quality. Producing eggs and incubating them is costly in terms of nutrients (Blount *et al*. 2004) and energy (Reid *et al*. 2000; Stevenson and Bryant 2000). Individuals of low quality are poor in obtaining the necessary resources and thus need longer before being able to lay the first egg. There is however (probably) a date after which successful reproduction is no longer possible. For example, simultaneously rearing chicks and moulting has a fitness cost (Hemborg and Lundberg 1998; Jenni and Winkler 1994) and more time for moulting leads to heavier and better plumage (Dawson *et al*. 2000; Serra 2001), which in turn is important for winter survival (Nilsson and Svensson 1996). Because producing eggs is costly, low quality birds may always wait until the last moment to start reproducing. This would minimise their costs for egg production, which would however be higher in years with a late spring. Another explanation could be that in such years conditions become favourable more quickly after a delay. If individuals can start egg laying at different levels of food availability because of their quality this would lead to less variation in laying dates in late years which in turn would lead to the observed correlation between intercept and slope of the reaction norms. However, we cannot conclude that breeding time is an outcome of heritable quality or a heritable trait *per se* but this may deserve further investigation.

No response to selection?

Since we found selection on, and heritability of, either breeding time per se (Chapter 6) or of its reaction norms (Chapter 5) we should expect an evolutionary change but the small observed shift in laying dates can be explained by phenotypic plasticity. The lack of a response to selection seems to be a common phenomenon, especially in breeding time of birds, and several possible explanations have been put forward (see Merilä *et al*. 2001b).

Selection differentials for breeding time are generally based on the number of offspring recruited into the breeding population (Both and Visser 2001; Cresswell and McCleery 2003; Svensson 1997; van Noordwijk *et al*. 1981; van Noordwijk *et al*. 1995; Visser *et al*. 1998). Producing eggs and incubating them is costly in terms of fitness (Monaghan *et al*. 1998; Visser and Lessells 2001). A trade-off between reproductive success and adult survival could result in no net selection on breeding time (Brown and Brown 1999). In Chapter 6 we however showed that in our study population there was only weak selection on breeding time via adult survival. We also found no genetic correlation of laying date with clutch size, which could hamper evolution of breeding time if clutch size was under independent selection (Lande and Arnold 1983).

Deteriorating environmental conditions hampering earlier laying?

A genetic change may be masked by simultaneously deteriorating environmental conditions. Decreasing food supply or increasing population density caused a reduction in fledging weight (corrected for tarsus length) of collared flycatchers and great tits despite selection for heavier fledging weight. Although individuals became phenotypically lighter they became genetically heavier as a result of selection (Garant *et al*. 2004; Merilä *et al*. 2001a). If temperatures are not only a cue but also a constraint (either directly or via closely related insect phenology) unchanged temperatures in early spring could hamper an evolutionary advancement of laying dates. We manipulated the experienced synchrony with the food peak in great tits in two study populations and expected the birds to advance their laying dates if egg production was not constrained by resource availability (Chapter 2). While the manipulated individuals in one population (Oosterhout, a mature oak forest on rich clay soils, Fig. 2.1) advanced their laying dates the individuals in the other population (Hoge Veluwe, a mixed forest on poor sandy soils, Fig. 2.2) did not. The difference in vegetation between the two areas and likely correlated differences in insect abundance and phenology could explain this outcome of the experiment but also hint at a possible resource constraint in the Hoge Veluwe, which could prevent earlier breeding.

Selection on the environmental component?

Price *et al*. (1988) elaborated on an explanation originally suggested by Darwin (1859) and Fisher (1958) how a trait could be heritable, under selection and yet not evolve, which has subsequently been described as 'selection on the environmental component'. If female 'condition' determines her reproductive success, also affects her breeding time but remains unmeasured it would seem that the arising correlation between breeding time and reproductive success was causal and breeding time was under selection (Chapter 1 - Box 1).

Comparing selection on phenotypes and the genetic component is the most direct way to test for selection on the environmental component since selection on the environmental component will lead to weaker selection on the genetic component (Rausher 1992; Stinchcombe *et al*. 2002). Using a so-called animal model we separated laying dates into their genetic and environmental component, quantified and compared selection on phenotypes and breeding values (Chapter 6). Selection differentials based on recruitment for breeding values, our measure of the genetic component, were slightly (but non-significantly) smaller than the corresponding selection differentials for phenotypes.

Selection differentials for a single fitness component may not give a complete picture of selection pressures on a given trait. Although survival selection was overall weak and non-significant it could still counteract (or enhance) fecundity selection. Quantifying and comparing time-trends in predicted cumulative responses at the genetic and phenotypic should be the sensitive test for possible differences in selection because it integrates viability and fecundity selection. However, quantifying cumulative responses to selection in populations with overlapping generations is not straightforward. Selection pressures via reproductive success and adult survival may differ and vary among years. Recruitment and survival rates may also vary and may even be correlated, negatively or positively, with selection pressures. We therefore calculated predicted responses on a year-to-year basis integrating both fecundity and viability selection as well as recruitment and adult survival rates. Predicted responses and observed changes in breeding values corresponded on a year-to-year basis (Fig. 6.3), which should give us confidence in the reliability of our predicted breeding values. From these predicted year-to-year responses we calculated the cumulative response to selection over the whole study period. Time-trends in cumulative responses for phenotypes and breeding values did not differ and we could thus find no evidence for selection on the environmental component.

This result was corroborated by the findings of the Chapters 3 and 4 since carotenoid colouration, as measure of non-genetic quality, was associated with both the genetic and the environmental components of laying date. Carotenoid colouration is an indicator of an individual's health or vigour (e.g. Dufva and Allander 1995; Faivre *et al*. 2003; Peters *et al*. 2004) and to large extent determined by the carotenoid availability in the environment (Olson and Owens 1998). If selection acted on the environmental component and carotenoid colouration was a suitable measure of non-genetic quality or condition we would expect a correlation between carotenoid colouration, the environmental component of laying date, and fitness. There was evidence that carotenoid colouration is an indicator of female quality since brighter yellow females raised more offspring and were preferred mates (Chapter 4). Carotenoid colouration was correlated with the genetic and the environmental components of laying date, but in the opposite way to that expected (Fig. 3.1, Fig. 3.2). Brighter yellow females laid both phenotypically and genetically later. This seems paradoxical. Why should high quality individuals lay later if early breeding is better and they should be able to lay early because of their high quality? One possible explanation is that high quality individuals are able to raise a brood successfully even after the optimal period for breeding, the caterpillar peak, because of their good foraging skills. Consequently, they could avoid any costs of early egg laying and survive better (Hill 1991; Horak *et al*. 2001), while individuals of lower quality would have to pay to breed in synchrony with the caterpillar peak. Carotenoid availability can depend on the environment (Olson and Owens 1998) and small-scale spatial differences could lead to the observed pattern if good quality habitats with high carotenoid availability sometimes have a late caterpillar phenology. Clearly, the results are not conclusive yet and data from more years and other population are necessary.

Small expected response to selection

The cumulative expected response to selection was small (-1.5 days) (Fig. 6.4) compared with the cumulative phenotypic selection differential (-34.8 days) since survival selection was much weaker than fecundity selection and adult survival rates relatively high (mean 1973-2003: 0.54). The estimated heritability was moderate to low and the sex-limitation of breeding time (thus halving the expected response) further reduced the expected response to strong phenotypic fecundity selection. Detecting such small expected response to selection and separating it from phenotypic trends may simply be a matter of statistical power.

However, when calculating the expected response to selection we assumed that immigrant individuals had undergone the same selection regime as locally born individuals. Immigration into a population can counteract selection if immigrating individuals had undergone different selection regimes. Great tits disperse before their first breeding attempt and females disperse more than males (Greenwood *et al*. 1979). The proportion of non-locally born first year-breeders varied in our study population from 0.20 to 0.45 and was on average 0.35. The correlation between observed changes in mean breeding values and expected year-to-year responses to selection (see Fig. 6.3) was weaker when it was based only on fecundity selection and annual mean breeding values of first-year breeders ($r = 0.38$, $p = 0.04$) than when adult survival selection and annual mean breeding values of all birds were included ($r = 0.58$, $p < 0.001$). Consequently, our assumption that immigrant females were selected similarly to locally-born females may not be entirely justified. Since we still found a positive correlation between expected response to fecundity selection and observed change in annual mean breeding values this effect cannot be strong enough to counteract selection acting in our study population.

To conclude, we found no evidence of evolutionary stasis of breeding time in this population. Inferring evolutionary stasis from an absence of a phenotypic trend despite (strong) phenotypic selection can be misleading if the expected response to selection is not carefully quantified.

Predicting the future

As demonstrated above, the evolutionary change in breeding time is and will be small compared to phenotypic changes. Using a phenotypic descriptive model (Chapter 7) to predict changes under climate change scenarios should thus at least give an indication how the relationship between breeding time in great tits and caterpillar phenology will develop in the future. Since the currently observed mismatch between chick rearing period and caterpillar phenology (Visser *et al*. 1998) has been produced by differential warming of periods early and late in spring it could be possible that changing temperature patterns in the future will restore the synchrony between birds and caterpillars.

We used daily mean temperatures generated by the SRES-B2 climate model (Esch 2005) to predict mean laying dates in great tits and dates when caterpillar biomass reached its maximum. Both laying dates and caterpillar biomass will advance at more or less the same rates (Fig. 8.4). This means forecasted temperatures will not alleviate the current mismatch, which might have negative effects at the population level in the future. Both *et al*. (unpublished manuscript) argue that the disapperance of certain pied flycatcher (*Ficedula hypoleuca*) populations in the Netherlands is related to increased mistiming.

General conclusion

In this thesis several possible explanations why breeding time of great tits in the Hoge Veluwe population did not respond to selection were explored. It could be ruled out that selection acted on the environmental component of laying date, survival selection counteracted fecundity selection or that heritability was very low or absent. There was some evidence that resource availability in early spring may constrain earlier laying. However, the moderate heritability found, the sex limitation of breeding time and overlapping generations reduced the expected response to selection to about 1.5 days in 30 years, which is not distinguishable from the simultaneously observed phenotypic change. This phenotypic trend was however too small to track the changes in the environment, especially caterpillar phenology, resulting in the currently observed mistiming.

In contrast to this population phenological changes related to climate change have been reported for a number of species covering a range of taxa (e.g. Beebee 1995; Inouye *et al*. 2000; McCleery and Perrins 1998; Weishampel *et al*. 2004; Woiwood 1997). Although phenological changes from year-to-year are assumed to be a product of phenotypic plasticity (e.g. Przybylo *et al*. 2000), rather than a fast response to selection, in most cases it can only be speculated whether the observed changes are a result of phenotypic plasticity or adaptation. Phenotypic plasticity will only be a sufficient response to climate change if the whole environment shifts at the same rate, which is highly unlikely to be true (see Stenseth and Mysterud 2002). For example, the developmental time of endotherm and ectotherm species depends very differently on ambient temperatures. Insect eggs will develop faster under warmer temperatures while the developmental time of bird eggs or mammalian embryos is independent of ambient temperature. Most studies reporting a phenological change as a response to climate change do not report whether selection on that trait changed, i.e., whether the phenotypic change was sufficient to fully track shifts in the species' environment (but see Cresswell and McCleery 2003).

How fast a species will be able respond to fecundity selection depends obviously on the heritabilty of the trait and on its generation time. Heritabilities of life-history or behavioural traits tend to be low to moderate and smaller than heritabilities for morphological traits (Stirling *et al*. 2002). Furthermore, the sex-limitation of certain life-history traits, like clutch respectively litter size and breeding time, will halve the expected response to selection. While morphological traits can respond quickly to selection (e.g. Grant and Grant 1995) the same is unlikely to be true for life-history or behavioural traits. To date only two studies reported a microevolutionary change of phenology in wild populations as a response to climate change. Bradshaw and Holzapfel (2001) showed a change in heritable critical photoperiod inducing diapause in mosquitos and argue that this change is a response to selection induced by climate change. Unfortunately, they did not attempt to quantify selection. Parturition date in Canadian squirrels has advanced strongly over only a couple of decades and this can be attributed partly to phenotypic plasticity and partly to a genetic change (Reale *et al*. 2003). However, the sex-limitation of parturition date was not included in their calculation of the expected response and the correct expected response would have been smaller than the observed change in predicted breeding values. This means that the breeding values or the selection differential must have been incorrectly quantified.

To conclude, the evidence of microevolutionary responses to climate change is scarce and the majority of observed changes are most likely a product of phenotypic plasticity. However, whether phenotypic plasticity will enable a species to fully track changes in its environment is questionable. This study showed that breeding time in great tits has not and will not evolve fast enough to track the phenological changes in its environment. Given the observed and expected rate of environmental change due to climate change and the slow response of life-history or behavioural traits it seems very unlikely that many species will be able adapt fast enough to track shifts in their environment caused by climate change (Etterson and Shaw 2001; Hendry and Kinnison 1999; Janzen 1994).

Summary

One central question in ecology and evolutionary biology is understanding how species adapt to their environment. Matching the timing of seasonal activities with favourable environmental conditions may have important fitness consequences. Increasing temperatures and other effects of climate change are likely to shift the timing of these environmental conditions and species should respond to this. It has previously been shown that increasing spring temperatures affected caterpillar phenology. Caterpillars are the main chick food supply for many passerine birds. Their changed phenology has led to selection for earlier breeding in the great tit population in the Hoge Veluwe (NL) but no corresponding change in the birds' egg laying dates has been observed (Visser *et al*. 1998). The aim of this thesis was to explain why this bird population did not seem to respond to the observed selection.

Temperature increases have not been uniform in time and space (Easterling *et al*. 1997; Luterbacher *et al*. 2004). Birds have to produce and incubate their eggs before the growth period of the caterpillars, which form the main food source for their chicks. There is an indication that temperatures during the period of egg production and incubation have warmed less than later during the growth period of caterpillars. Producing eggs and incubating them is costly (e.g. Visser and Lessells 2001) and the unchanged conditions in early spring may hamper the advancement of laying dates. In Chapter 2 this explanation was addressed by manipulating laying dates and measuring possible fitness cost of advanced egg laying. It has frequently been shown that food supplementation results in earlier laying (reviewed in Nager *et al*. 1997) but such experiments are unsuitable to measure fitness consequences of earlier breeding. There is experimental evidence that blue tits learn from experienced synchrony with their caterpillar prey (Grieco *et al*. 2002). We used this learning ability to induce earlier egg laying in two nest box populations of great tits. In one study area (Oosterhout, NL) females responded to the treatment by laying earlier but due to low sample sizes it was not possible to quantify the fitness of earlier laying. In the Hoge Veluwe study area females did not lay earlier in response to the experiment. This might indicate that fitness costs of advanced laying in this area are too high and consequently females refrained from it. The two study areas differ in vegetation and a correlated difference in insect prey abundance or phenology at the time of egg production could explain this difference between areas.

If fitness was not related to breeding time but an unmeasured, non-heritable trait, which affects both breeding time and fitness, a non-causal correlation between fitness and breeding time could arise (Price *et al*. 1988). Carotenoid colouration is related to an individual's health or vigour (e.g. Dufva and Allander 1995; Faivre *et al*.

2003) and consequently a possible indicator for this unmeasured trait. In Chapter 3 and Chapter 4 the relationships between carotenoid colouration, breeding time and fitness were measured in two great tit populations (Hoge Veluwe and Vlieland, NL). While there was evidence that females with a brighter yellow breast colouration had a higher reproductive success an unexpected relationship with breeding time was found: yellower females bred later. This relationship was found both for the genetic and environmental component of laying date and explanations for this could only be proposed but not tested.

Breeding time in birds is a highly plastic trait (e.g. Brommer *et al*. 2005; Przybylo *et al*. 2000). Differing temperature trends in early and late spring may lead to selection on the reaction norms underlying the observed phenotypic plasticity. In Chapter 5 we addressed whether individual breeding time reaction norms differ in their slope and intercept. Individual reaction norms were described in relation to average temperature from 16 March to 20 April. There was evidence that birds differ in the slope of their reaction norm, i.e. the sensitivity of their response to temperature, and their intercept, and that both slope and intercept are heritable. During the study period selection on the slope of the reaction norm increased, favouring more plastic females. There was also a strong genetic correlation of slope and intercept possibly indicating that both are emerging properties of the same underlying trait.

Comparing selection differentials for phenotypes and the genetic component of a trait is the most direct way to test whether an unmeasured factor, like health or condition, creates a correlation between the trait and fitness, which could lead to the wrong impression that selection acted on the trait itself (Rausher 1992; Stinchcombe *et al*. 2002). In Chapter 6 breeding values for laying dates obtained from an ‘animal model’ were used as a measure of the genetic component. Subsequently, we calculated and compared selection differentials for phenotypes and breeding values. These selection differentials did not differ significantly and hence there was no evidence that an unmeasured factor, like the female’s health or condition, would create a correlation between breeding time and fitness. The estimate of the additive genetic variance of laying date obtained from the ‘animal model’ gave a low (0.17) but significant heritability. Selection on adult survival was weak overall and did not counteract fecundity selection. However, the cumulative response to selection was small (1.5 days) because of the low heritability, the sex-limitation of breeding time and relatively high adult survival rates. During the study period laying date advanced non-significantly by 5.4 days. It was not possible to separate the response to selection from the expected change caused by phenotypic plasticity (5.5 days).

Although linear regressions of mean laying dates against mean temperatures measured over a fixed period have been widely used to describe phenology (e.g. Both *et al*. 2004; Meijer *et al*. 1999; Perrins and McCleery 1989) this is problematic when the temperature measure shows a systematic time-trend. Analysing individual observations and using a time-dependent explanatory variable can solve this problem. In Chapter 7 we applied a proportional hazards model (Cox 1972) to describe laying dates of great tits. This model successfully predicted annual mean laying dates and also annual standard deviations. Laying dates were affected by female age (*cf* Blums *et al*. 1997; Robertson and Rendell 2001), temperature, photoperiod and an interaction between photoperiod and temperature.

In great tits, the synchrony between caterpillars and chick rearing period has important consequences for reproductive success, as shown in Chapter 8 (see also van Noordwijk *et al*. 1995; Verboven *et al*. 2001). Great tits breeding too early or too late in relation to caterpillar phenology fledged significantly less chicks, which were also significantly lighter reducing their survival probability (Verboven and Visser 1998). Good descriptive models are necessary to assess possible effects of climate change on reproductive success and consequently population dynamics. In Chapter 8 a descriptive model for food caterpillar phenology was developed. In combination with the descriptive model for great tit laying dates (Chapter 7) it was used to forecast the synchrony between birds and caterpillars for the coming 100 years based on an IPCC scenario (Esch 2005). Both laying dates and caterpillar phenology will advance at the same rate. Since the birds are currently breeding too late in relation to the caterpillars this means that we should not expect a restoration of the synchrony without an evolutionary change in the birds' breeding time. However, as shown in Chapter 6 this is most likely to happen at a (too) slow rate.

To date, phenological changes as a response to climate change have been reported in several species covering a range of taxa (e.g. Beebee 1995; Crick *et al*. 1997; Inouye *et al*. 2000; Weishampel *et al*. 2004; Woiwood 1997), which were rather a result of phenotypic plasticity than a response to selection. Given the general low heritability of life-history or behavioural traits (Stirling *et al*. 2002) and the rate of the predicted environmental change (Houghton *et al*. 2001) it seems unlikely that species' life-history or behaviour can evolve fast enough to track the predicted changes in their environment (Chapter 6; Etterson and Shaw 2001; Janzen 1994).

Samenvatting

De vraag hoe soorten zich aanpassen aan hun leefomgeving staat centraal binnen zowel de ecologie als de evolutionaire biologie. Vanuit het oogpunt van reproductief succes is het van groot belang om die seizoensgebonden activiteiten die de grootste inspanning vergen samen te laten vallen met de periode waarin de omstandigheden het gunstigst zijn. Klimaatsverandering leidt tot, onder andere, steeds hogere temperaturen, en daarmee tot een verschuiving van de periode waarin condities optimaal zijn. Soorten zullen hier op de een of de andere manier op moeten reageren. Eerder onderzoek heeft laten zien dat een stijging van de temperatuur een effect heeft gehad op de fenologie van rupsen, die een belangrijke voedselbron voor koolmezen zijn tijdens de jongen periode. De vervroeging van de rupsen fenologie heeft geleid tot selectie voor vroeger leggen in een koolmezen populatie op de Hoge Veluwe. Er is echter geen verschuiving in gemiddelde legdatum van de vogels opgetreden (Visser *et al.* 1998). Het doel van dit proefschrift was om een verklaring te vinden voor de afwezigheid van een dergelijke verschuiving.

De temperatuur stijgt niet overal in dezelfde mate, er is variatie in zowel ruimte als tijd (Easterling *et al.* 1997; Luterbacher *et al.* 2004). Vogels moeten hun eieren produceren en bebroeden voordat het groeiseizoen van de rupsen, het hoofdbestanddeel van het dieet van hun jongen, is aangebroken. In Nederland is de temperatuur tijdens de periode van leggen en bebroeden van de eieren minder sterk gestegen dan de temperatuur tijdens de groei van de rupsen en het voeren van de jongen. Nu kost het produceren en bebroeden van eieren veel energie (zie bijvoorbeeld Visser en Lessels 2001), en de onveranderende condities in het vroege voorjaar zouden het de koolmezen dus onmogelijk kunnen maken om hun eieren vroeger te leggen. Dat zou kunnen verklaren waarom de koolmezen op de Hoge Veluwe niet vroeger zijn gaan leggen. Deze mogelijke verklaring hebben we getoetst in een experiment dat beschreven is in Hoofdstuk 2. In dit experiment hebben we geprobeerd om door het manipuleren van legdata in twee populaties (Oosterhout en de Hoge Veluwe) de fitness consequenties van vroeger leggen te meten. Hoewel men verschillende keren heeft laten zien dat het verstrekken van extra voedsel leidt tot vroeger leggen (zie bijvoorbeeld Nager *et al.* 1997), zijn zulke experimenten helaas ongeschikt om de fitness consequenties van vroeger leggen te bepalen. In een eerdere studie heeft men echter laten zien dat pimpelmezen kunnen leren van de door hun ervaren synchronie tussen de voedselbehoefte van hun jongen en het moment van de rupsenpiek: de vogels passen hun legdatum in het volgende jaar aan wanneer ze het jaar ervoor te laat waren (Grieco *et al.* 2002). We hebben deze capaciteit om te leren uitgebuit om koolmeesvrouwtjes eerder te laten leggen zonder ze meer voedsel te geven in het jaar waarin we de fitness consequenties van de legdatum hebben bepaald. Hoewel in Oosterhout de vrouwtjes inderdaad reageerden op een experimentele vervroeging van de door hun waargenomen voedselpiek en zij het jaar daarna eerder legden, was de steekproefgrootte helaas te klein om de fitness consequenties hiervan te kunnen bepalen. In tegenstelling tot de vrouwtjes in Oosterhout, reageerden vrouwtjes op de Hoge Veluwe niet op de behandeling. Een mogelijke verklaring hiervoor is dat de kosten van vroeger leggen, in termen van fitness, te hoog zijn op de Hoge Veluwe, en dat vrouwtjes het daarom niet doen. Daarnaast verschillen de twee studieterreinen qua samenstelling van de vegetatie. Een daarmee samenhangend verschil in de beschikbaarheid van insecten of in de fenologie tijdens de eileg zou het verschil tussen de twee terreinen eveneens kunnen verklaren.

Het is mogelijk dat het reproductief succes van een vrouwtje niet direct bepaald wordt door het moment waarop ze haar eieren legt, maar door een derde, niet gemeten en niet erfelijke eigenschap. Als deze eigenschap naast haar legdatum ook haar reproductieve succes beïnvloedt, dan leidt dit tot een niet causaal verband tussen fitness en legdatum (Price *et al.* 1988). Carotenoïden, onder andere aanwezig in de gele borstveren van koolmezen, houden verband met een individu’s conditie of gezondheid (zie bijvoorbeeld Dufva en Allander 1995; Faivre *et al.* 2003). Hiermee vormen ze een mogelijke indicator voor de aanwezigheid van een niet erfelijke derde factor die er toe leidt dat vrouwtjes zowel vroeg broeden als veel jongen groot kunnen brengen. In Hoofdstuk 3 en 4 wordt de relatie tussen carotenoïde kleur, legdatum en fitness beschreven in twee koolmezen populaties (Hoge Veluwe en Vlieland). Hoewel we bewijs vonden dat vrouwtjes met een gelere borst inderdaad meer jongen groot brachten, vonden we tegen onze verwachting in dat deze vrouwtjes juist later broedden. Dit verband bestond voor zowel de genetische als de omgevingscomponent van legdatum. Hoewel er verschillende verklaringen zijn voor het gevonden verband, was het helaas niet mogelijk om deze te toetsen.

Legdatum is een buitengewoon flexibele eigenschap in vogels (bijv. Przybylo *et al.* 2000; Brommer *et al.* 2005). Het eerder genoemde verschil in hoe de temperatuur is veranderd in het vroege en late voorjaar zou kunnen leiden tot selectie op de manier waarop vrouwtjes hun legdatum aanpassen aan de temperatuur, of te wel hun reactienorm. In Hoofdstuk 5 hebben we voor alle vrouwtjes die meer dan een keer hebben gebroed op de Hoge Veluwe hun reactienorm voor legdatum opgesteld door hun legdatum te koppelen aan de gemiddelde temperatuur van 16 maart tot 20 april. Vervolgens hebben we getoetst of vrouwtjes verschillen in hun reactienorm, zowel wat betreft de helling als de hoogte. We vonden dat vogels inderdaad verschillen in hoogte van hun reactienorm, of te wel in hun legdatum in het gemiddelde voorjaar. Ze bleken echter ook te verschillen in de helling, of met andere woorden in hun gevoeligheid voor variatie in voorjaarstemperatuur. Beide eigenschappen bleken bovendien een erfelijke basis te hebben. Ook bestond er een sterke genetische correlatie tussen de hoogte en de helling van een reactienorm, hetgeen suggereert dat beide het resultaat zijn van dezelfde onderliggende eigenschap. Tijdens de studieperiode was er selectie voor een steilere helling van de reactie norm, en deze selectie is met de tijd sterker geworden.

De meest directe manier om te testen of een derde factor, zoals gezondheid of conditie, de oorzaak is van het waargenomen verband tussen legdatum en reproductief succes, is een vergelijking van selectiedifferentialen gebaseerd op legdatum zelf met selectiedifferentialen die gebaseerd zijn op de genetische component van legdatum (Rausher 1992; Stinchcombe *et al.* 2002). In Hoofdstuk 6 hebben we daarom voor ieder vrouwtje de genetische component van legdatum geschat op basis van de legdatum van haar familieleden. Vervolgens hebben we selectiedifferentialen berekend voor zowel legdatum als de genetische component van legdatum. Deze selectie differentialen bleken echter onderling niet significant van elkaar te verschillen, hetgeen er op wijst dat de gezondheid of conditie van een vrouwtje niet ten grondslag ligt aan het waargenomen negatieve verband tussen legdatum en fitness. Een vergelijking met de legdatum van verwante vrouwtjes gaf een significante maar relatief lage erfelijkheidsgraad voor legdatum van 0,17. Selectie via adulte overleving was over het algemeen zwak en werkte niet tegen selectie via de productie van rekruten in. De cumulatieve verwachtte respons op selectie was echter slechts klein, namelijk 1,5 dag gedurende de gehele studieperiode. Dit is het gevolg van de lage erfelijkheidsgraad, het feit dat de genen voor legdatum alleen tot uitdrukking komen in vrouwtjes, en tot slot de relatief hoge adulte overleving. Gedurende de studieperiode is legdatum (niet significant) met 5,4 dagen vervroegd. Het was niet mogelijk om de respons op selectie van de verwachte verandering als gevolg van de fenotypische plasticiteit, die 5,5 dagen bedraagt, te scheiden.

Bij het beschrijven van de fenologie van een soort maakt men vaak gebruik van lineaire regressies van legdatum tegen de temperatuur zoals die gemeten is over een vaste periode (bijv. Both *et al.* 2004; Meijer *et al.* 1999; Perrins en McCleery 1989). Dit is echter problematisch als de temperatuur tijdens deze vaste periode een systematische trend vertoont in de tijd. In Hoofdstuk 7 maken we daarom gebruik van een zogenaamd ‘proportional hazards model’ (Cox 1972) om legdatum in koolmezen te beschrijven. Op deze manier was het mogelijk om zowel de gemiddelde legdatum per jaar als de jaarlijkse standaarddeviaties te voorspellen. Verder konden we aantonen dat legdata beïnvloed worden door de leeftijd van een vrouwtje (zie Blums *et al.* 1997; Robertson en Rendell 2001), maar ook door de temperatuur, de daglengte en de interactie tussen beide.

In Hoofdstuk 8 hebben we laten zien dat in koolmezen de synchronie tussen de rupsen en de periode van het opvoeden van de jongen belangrijke consequenties heeft voor hun reproductieve succes (zie ook van Noordwijk *et al.* 1995; Verboven *et al.* 2001). Individuen, die te vroeg of laat ten opzichte van de rupsen phenologie broeden, brengen minder jongen groot, die bovendien ook lichter zijn. Goede beschrijvende modellen zijn noodzakelijk om de mogelijke consequenties van klimaatsverandering op reproductief succes op zowel individueel als op populatie niveau te kunnen overzien. In Hoofdstuk 8 hebben we daarom, naast het model dat de legdatum van de koolmezen beschrijft (Hoofdstuk 7), ook een model opgesteld waarmee de fenologie van de rupsen beschreven kon worden. Aan de hand van een IPCC scenario (Esch 2005) en deze twee modellen hebben we vervolgens de synchronie tussen vogels en rupsen voorspeld voor de komende 100 jaar. Dit liet zien dat zowel legdatum als rups fenologie in de zelfde mate zullen vervroegen. Aangezien de vogels momenteel te laat broeden ten opzichte van de rupsen betekent dit dus dat we zonder een evolutionaire verandering in de legdatum van de vogels geen herstel van de synchronie hoeven te verwachten.

Tot nu toe heeft men fenologische veranderingen op klimaatsverandering laten zien in groot aantal soorten behorende tot een wijde range van taxa (bijv. Beebee 1995; Crick *et al.* 1997; Inouye *et al.* 2000; Weishampel *et al.* 2004; Woiwood, 1997). Deze veranderingen zijn eerder het resultaat zijn van fenotypische plasticiteit dan van een respons op selectie. Gegeven de lage erfelijkheidsgraad van gedragskenmerken, inclusief legdatum (Stirling *et al.* 2002), en de voorspelde snelheid waarmee de omgeving zal veranderen (Houghton *et al.* 2001), lijkt het onwaarschijnlijk dat soorten snel genoeg kunnen evolueren om de verwachtte veranderingen bij te houden met waarschijnlijke, negatieve effecten op het voortbestaan van populaties of zelfs soorten (Hoofdstuk 6; Etterson en Shaw 2001; Janzen 1994).

Zu verstehen wie Arten sich an ihre Umgebung anpassen, ist eine zentrale Fragestellung in Ökologie und Evolutionsbiologie. Die zeitliche Übereinstimmung von Fortpflanzung und entsprechenden, günstigen Umweltbedingungen kann wichtige Fitnesskonsequenzen haben. Steigende Temperaturen und andere Auswirkungen globaler Klimaveränderungen, wie z. B. mehr oder weniger Niederschlag, haben diese Umweltbedingen bereits beeinflusst und wir können erwarten, dass Arten sich an diese veränderten Umweltbedingungen anpassen. Wärmere Frühlingstemperaturen haben zu einer früheren Phänologie bestimmter Schmetterlingsraupen geführt. Diese Raupen sind jedoch eine wichtige Nahrungsquelle für viele insektivore Singvögel und deren veränderte Phänologie hat zu zunehmender Selektion für einen früheren Brutbeginn bei Kohlmeisen im "Hoge Veluwe" (NL) geführt (Visser *et al*. 1998). Allerdings haben die Meisen dieser Population ihren Brutbeginn noch nicht, wie erwartet, verändert und das Ziel dieser Doktorarbeit war es zu erklären, warum diese Population nicht auf Selektion reagiert zu haben scheint.

Die Umgebungstemperatur hat nicht in allen Regionen und Jahreszeiten gleichermaßen zugenommen, so vor allem Winter und Frühling wärmer geworden während sich Sommer- und Herbsttemperaturen wenig geändert haben (Easterling *et al*. 1997; Luterbacher *et al*. 2004). Zwischen dem Zeitpunkt, an dem das erste Ei eines Geleges gelegt wird, und dem Zeitpunkt, an dem die Jungen schlüpfen, liegen ungefähr drei Wochen. Folglich müssen Kohlmeisen ihre Eier legen und bebrüten bevor die Raupen, die die Hauptnahrungsquelle für die Nestlinge sind, überhaupt schon geschlüpft sind. Die Temperatur während der Eilege- und Brutperiode scheint in den letzten Jahrzehnten weniger stark gestiegen zu sein als die Temperatur während der Periode, in der sich die Raupen entwickeln. Eier zu produzieren und zu bebrüten braucht Nährstoffe und Energie und hat dies kann die Überlebenswahrscheinlichkeit eines Weibchens beeinflussen (Visser und Lessells 2001). Die unveränderten Temperaturen im frühen Frühjahr, wenn die Vögel ihre Eier produzieren und bebrüten müssen, können einen früheren Brutbeginn verhindern. Diese Erklärung wurde durch experimentelle Manipulation des Legebeginns und die Quantifizierung möglicher Fitnesskosten getestet, wie in Kapitel 2 beschrieben. Es ist häufig gezeigt worden, dass ein künstliches zusätzliches Futterangebot zu einem früheren Legebeginn führt (z.B. Nager *et al*. 1997). Leider sind solche Experimente ungeeignet, die Fitnesskonsequenzen eines früheren Legebeginns zu quantifizieren, denn die manipulierten Individuen erhalten ja zusätzliche Nahrung. Es ist auch gezeigt worden, dass Blaumeisen aus der zeitlichen Abstimmung ihres Brutzeitpunkts mit dem Nahrungsangebot lernen können (Grieco *et al*. 2002). Wir machten uns diese Lernfähigkeit zunutze, um den Brutzeitpunkt bei Kohlmeisen durch einen künstliches Futterangebot experimentell zu beeinflussen. In einen Untersuchungsgebiet ("Oosterhout", NL) reagierten Weibchen auf diese Manipulation und legten im darauffolgenden Jahr ihre Eier früher. Durch eine (zu) kleine Stichprobengröße war es jedoch nicht möglich, Fitnesseffekte dieses manipulierten Legebeginns zu messen. In einem zweiten Untersuchungsgebiet ("Hoge Veluwe", NL) reagierten die Weibchen nicht auf die experimentelle Manipulation. Möglicherweise wären die Kosten eines früheren Legebeginns zu groß gewesen, und die Weibchen haben schlicht davon abgesehen, ihre Eier früher zu legen. Die beiden Untersuchungsgebiete unterscheiden sich in ihrer Vegetation und ein entsprechender Unterschied in der Nahrungsverfügbarkeit, z. B. von Insekten, während der Eilegeperiode könnte die unterschiedlichen Ergebnisse in den zwei Untersuchungsgebieten erklären.

Wenn Fitness nicht vom Brutzeitpunkt *per se* sondern von einer anderen, nicht gemessenen und nicht-erblichen Eigenschaft, z. B. Qualität des Territoriums oder Vitalität, die außerdem auch den Brutzeitpunkt beeinflusste, abhinge, würde eine nicht-kausale Korrelation zwischen Brutzeitpunkt und Fitness entstehen (Price *et al*. 1988). Bei Vögeln korreliert bunte Carotenoid-basierte Gefiederfärbung positiv mit der Gesundheit oder Vitalität (z. B. Dufva und Allander 1995; Faivre *et al*. 2003). Carotenoid-basierte Gefiederfärbung könnte deshalb ein Indikator für diese Eigenschaft sein, denn Weibchen in generell guter Kondition sollten in der Lage, früh zu brüten und auch viele Nachkommen aufzuziehen. In Kapitel 3 und 4 wurde diese Verhältnis zwischen carotenoid-basierter Färbung der Brustfedern, Brutzeitpunkt und Fitness in zwei Populationen ("Hoge Veluwe" und "Vlieland", NL) analysiert. Während stärker gefärbte Weibchen mehr Eier legen und mehr Jungen aufziehen konnten, brüteten sie im Gegensatz zu unseren Erwartungen später als ihre weniger stark gefärbten Artgenossinnen. Dieses Verhältnis zwischen Brutzeitpunkt und Carotenoid-basierter Gefiederfärbung wurde sowohl für die genetische Komponente als auch für die Umgebungskomponente des Brutzeitpunktes gefunden. Mögliche Erklärungen konnten leider nur vorgeschlagen aber nicht getestet werden.

Der Legebeginn bei Vögeln kann beträchtliche phänotypische Plastizität zeigen (z. B. Brommer *et al*. 2005; Przybylo *et al*. 2000). Unterschiedliche Temperaturtrends im frühen und späten Frühjahr könnten zu Selektion auf die Reaktionsnormen, die dieser phänotypischen Plastizität zugrunde liegen, führen. In Kapitel 5 wurde untersucht, ob sich individuelle Reaktionsnormen in ihrer Steigung und "Höhe" (Y-Achsenabschnitt) unterscheiden. Reaktionsnormen wurden als das

Verhältnis von Legebeginn zur Durchschnittstemperatur vom 16. März bis 20. April beschrieben. Individuelle Reaktionsnormen unterschieden sich sowohl in ihrer Steigung, d. h., wie sensitiv die Weibchen auf ihre Umgebungstemperatur reagieren, als auch in der "Höhe", d. h. wie früh oder spät ein Weibchen unter der durchschnittlichen Temperatur beginnt, Eier zu legen. Diese Unterschiede haben eine genetische Komponente, sind also erblich. Allerdings existiert auch eine starke genetische Korrelation zwischen der Steigung einer Reaktionsnorm und ihrer "Höhe", was darauf hinweist, dass Steigung und "Höhe" dieser Reaktionsnormen Effekte einer gemeinsamen zugrunde liegenden genetischen Eigenschaft sind.

Die Stärke von Selektion auf Phänotypen und die genetische Komponente einer Eigenschaft zu vergleichen, ist der direkteste Test, ob ein "dritter Faktor", wie z. B. Vitalität, eine Korrelation zwischen der Eigenschaft, in diesem Falle natürlich Legebeginn, und Fitness hervorruft, die dann fälschlicherweise als Selektion auf die Eigenschaft selbst wahrgenommen wird (Rausher 1992: Stinchcombe *et al*. 2002). In Kapitel 6 wurde ein sogenanntes "animal model" benutzt, um "breeding values" als Maß für die genetische Komponente zu schätzen. Anschließend wurden Selektionsdifferentiale für Phänotypen und "breeding values" berechnet. Diese Selektionsdifferentiale unterschieden sich nicht signifikant und es konnte also kein Hinweis darauf gefunden werden, dass ein "dritter Faktor" zu scheinbarer Selektion auf Legebeginn geführt hätte. Die durch das "animal model" berechnete Erblichkeit von Legebeginn war klein (0,17) aber signifikant größer als Null. Da sich früh und spät brütende Individuen nur wenig in ihrer Überlebenswahrscheinlichkeit unterschieden, gibt es in dieser Population keine Selektion *via* Überleben der Adulten, die der Selektion via reproduktivem Erfolg entgegenwirken könnte. Die kumulative erwartete genetische Veränderung des Legebeginns aufgrund der beobachteten Selektion war jedoch nur klein (1,5 Tage), da die Erblichkeit von Legebeginn relativ klein ist, Legebeginn eine geschlechts-limitierte Eigenschaft ist und die Generationen überlappen. Während des Untersuchungszeitraumes begannen die Kohlmeisen 5,4 Tage früher ihre Eier zu legen. Diese Veränderung ist jedoch nicht von der erwarteten Veränderung des Legebeginnes von 5,4 Tagen, die man aufgrund von phänotypischer Plastizität erwarten würde.

Obwohl lineare Regressionen von jährlichen gemitteltelten Legebeginn gegen Frühlings-Durchschnittstemperaturen üblich sind, um die Phänologie der Fortpflanzung zu beschreiben (z. B. Both *et al*. 2004; Meijer *et al*. 1999; Perrins und McCleery 1989) ist dies problematisch, wenn ebendiese Temperaturen einen zeitlichen Trend zeigen. Einzelne Eilegedaten zu analysieren und hierfür ein statistisches Modell mit zeitabhängigen Variablen zu benutzen kann dieses Problem umgehen. In Kapitel 7 wurde ein "Proportional Hazards Model" (Cox 1972) benutzt, um den Legebeginn bei Kohlmeisen zu beschreiben. Dieses Modell war in der Lage, sowohl jährliche Mittelwerte als auch Standardabweichungen zu beschreiben. Nach diesen Ergebnissen wird Legebeginn bei Kohlmeisen durch Alter (siehe auch Blums *et al*. 1997; Robertson und Rendell 2001), Umgebungstemperatur und Tageslänge beeinflusst.

Bei Kohlmeisen hat die zeitliche Übereinstimmung zwischen der Jungenaufzuchtperiode und der Phänologie der Raupenarten, die die Hauptnahrungsquelle darstellen, wichtige Konsequenzen für den Fortpflanzungserfolg, wie in Kapitel 8 gezeigt wird (siehe auch van Noordwijk *et al*. 1995; Verboven *et al*. 2001). Kohlmeisen, die zu früh oder zu spät im Verhältnis zur Phänologie der Raupen brüteten, hatten signifikant weniger Flügglinge, die außerdem auch noch signifikant leichter waren, was negative Effekte auf ihre Überlebenswahrscheinlichkeit hat (Verboven und Visser 1998). Um den Effekt von Klimaveränderung auf den Fortpflanzungserfolg und darausfolgend die Populationsdynamik abschätzen zu können, sind gute Modelle für sowohl die Phänologie der Kohlmeisen wie die ihrer Nahrung notwendig. In Kapitel 8 haben wir ein solches Modell für die Raupen-Phänologie entwickelt. Basierend auf den Temperaturen eines IPCC-Szenarios (Esch 2005) wurden dieses Modell und das in Kapitel 7 entwickelte Modell für Kohlmeisen-Phänologie benutzt, um die zeitliche Übereinstimmung zwischen Kohlmeisen und ihrer Raupennahrung für die nächsten 100 Jahre vorherzusagen: sowohl die Phänologie der Kohlmeisen wie ihrer Hauptbeute wird sich mit derselben Rate ändern. Da die Vögel gegenwärtig zu spät im Verhältnis zu ihrer Beute brüten, können wir darum nicht erwarten, dass diese zeitliche Übereinstimmung ohne eine evolutionäre Veränderung des Legebeginns wieder herstellt werden kann. Allerdings wird eine evolutionäre Veränderung des Legebeginns sehr wahrscheinlich (zu) langsam vor sich gehen (siehe Kapitel 6).

Bis jetzt sind phänologische Veränderungen bei mehreren Arten aus verschiedensten Taxa beschrieben worden (z. B. Beebee 1995; Crick *et al*. 1997; Inouye *et al*. 2000; Weishampel *et al*. 2004; Woiwood 1997), die jedoch eher ein Produkt von phänotypischer Plastizität als von evolutionärer Veränderung sind. Berücksichtigt man die generell geringen Erblichkeiten von "life-history" Eigenschaften oder Verhalten (Stirling *et al*.2002) und die Rate der vorhergesagten Klimaveränderungen (Houghton *et al*. 2001) scheint es unwahrscheinlich zu sein, dass die "life-history" oder das Verhalten von Arten schnell genug evolvieren kann, um mit den vorhergesagten Umweltveränderungen Schritt halten zu können, was zu negativen Konsequenzen für den Fortbestand von Population oder sogar Arten führen kann (Kapitel 6; Etterson und Shaw 2001; Janzen 1994).

References

References

Ahola, M., T. Laaksonen , K. Sippola, T. Eeva, K. Rainio, and E. Lehikoinen. 2004. Variation in climate warming along the migration route uncouples arrival and breeding dates. Global Change Biology 10:1610–1617.

Alonso-Alvarez, C., S. Bertrand, G. Devevey, M. Gaillard, J. Prost, B. Faivre, and G. Sorci. 2004. An Experimental Test of the Dose-Dependent Effect of Carotenoids and Immune Activation on Sexual Signals and Antioxidant Activity. American Naturalist 164:651-659.

Amundsen, T. 2000. Why are female birds ornamented? Trends in Ecology & Evolution 15:149-155.

Amundsen, T., and E. Forsgren. 2001. Male mate choice selects for female coloration in a fish. Proceedings of the National Academy of Sciences of the United States of America 98:13155-13160.

Amundsen, T., E. Forsgren, and L. T. T. Hansen. 1997. On the function of female ornaments: male bluethroats prefer colourful females. Proceedings of the Royal Society of London Series B-Biological Scienes 264:1579-1586.

Andersson, M. 1994, Sexual selection. Princeton, Princeton University Press.

Andersson, S., J. Ornborg, and M. Andersson. 1998. Ultraviolet sexual dimorphism and assortative mating in blue tits. Proceedings of the Royal Society of London Series B-Biological Sciences 265:445-450.

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

Beebee, T. J. C. 1995. Amphibian Breeding and Climate. Nature 374:219-220.

Blount, J. D., D. C. Houston, P. F. Surai, and A. P. Møller. 2004. Egg-laying capacity is limited by carotenoid pigment availability in wild gulls *Larus fuscus*. Proceedings of the Royal Society of London Series B-Biological Sciences (Suppl.) 271: S79-S81.

Blount, J. D., N. B. Metcalfe, T. R. Birkhead, and P. F. Surai. 2003. Carotenoid modulation of immune function and sexual attractiveness in zebra finches. Science 300:125-127.

Blums, P., G. R. Hepp, and A. Mednis. 1997. Age-specific reproduction in three species of European ducks. Auk 114:737-747.

Boland, C. R. J., M. C. Double, and G. B. Baker. 2004. Assortative mating by tail streamer length in Red-tailed Tropicbirds *Phaethon rubricauda* breeding in the Coral Sea. Ibis 146:687-690.

Both, C., A. C. Artemyev, B. Blaauw, R. J. Cowie, A. J. Dekhuijzen, T. Eeva, A. Enemar et al. 2004. Large-scale geographical variation confirms that climate change causes birds to lay earlier. Proceedings of the Royal Society of London Series B-Biological Sciences 271:1657-1662.

Both, C., and M. E. Visser. 2001. Adjustment to climate change is constrained by arrival date in a long-distance migrant bird. Nature 411:296-298.

Bradley, N. L., A. C. Leopold, J. Ross, and W. Huffaker. 1999. Phenological changes reflect climate change in Wisconsin. Proceedings of the National Academy of Sciences of the United States of America 96:9701-9704.

Bradshaw, W. E., and C. M. Holzapfel. 2001. Genetic shift in photoperiodic response correlated with global warming. Proceedings of the National Academy of Sciences of the United States of America 98: 14509-14511.

Brinkhof, M. W. G., A. J. Cave, F. J. Hage, and S. Verhulst. 1993. Timing of reproduction and fledging success in the coot *Fulica atra*: evidence for a causal relationship. Journal of Animal Ecology 62:577-587.

Brodie, E. D., A. J. Moore, and F. J. Janzen. 1995. Visualizing and quantifying natural-selection. Trends in Ecology & Evolution 10:313-318.

Brommer, J. E. 2004. The range margins of northern bird species shift polewards. Annales Zoologici Fennici 41:391-397.

Brommer, J. E., J. Merilä, and H. Kokko. 2002. Reproductive timing and individual fitness. Ecology Letters 5:802-810.

Brommer, J. E., J. Merilä, B. Sheldon, and L. Gustafsson. 2005. Natural Selection and genetic variation for reproductive reaction norms in a wild bird population, Evolution 59: 1362-1371.

Brown, C. R., and M. B. Brown. 1999. Fitness components associated with laying date in the cliff swallow. Condor 101:230-245.

—. 2000. Weather-mediated natural selection on arrival time in cliff swallows (*Petrochelidon pyrrhonota*). Behavioural Ecology and Sociobiology 47:339-345.

Brown, H., and R. Prescott. 1999, Applied mixed models in medicine. Chichester, John Wiley & Sons.

Brown, J. L., S. H. Li, and N. Bhagabati. 1999. Long-term trend toward earlier breeding in an American bird: A response to global warming? Proceedings of the National Academy of Sciences of the United States of America 96:5565-5569.

Burley, N., and C. B. Coopersmith. 1987. Bill color preferences of zebra finches. Ethology 76:133-151.

Bush, and Mosteller. 1955, Stochastic models of learning. New York, Wiley & Sons.

Cameron, N. D. 1997, Selection indices and prediction of genetic merit in animal breeding. Wallingford, CAB International.

Cannell, M. G. R., and R. I. Smith. 1983. Thermal time, chill days and prediction of budburst in *Picea sitchensis*. Journal of Applied Ecology 20:951-963.

Chastel, O., A. Lacroix, and M. Kersten. 2003. Pre-breeding energy requirements: thyroid hormone, metabolism and the timing of reproduction in house sparrows *Passer domesticus*. Journal of Avian Biology 34:298-306.

Christians, J. K., M. Evanson, and J. J. Aiken. 2001. Seasonal decline in clutch size in European starlings: a novel randomization test to distinguish between the timing and quality hypotheses. Journal of Animal Ecology 70:1080-1087.

Cox, D. R. 1972. Regression models and life tables. Biometrics 39:67-77.

Cresswell, W., and R. McCleery. 2003. How great tits maintain synchronization of their hatch date with food supply in response to long-term variability in temperature. Journal of Animal Ecology 72:356-366.

Crick, H. Q. P., C. Dudley, D. E. Glue, and D. L. Thomson. 1997. UK birds are laying eggs earlier. Nature 388:526-526.

Daan, S., C. Dijkstra, and J. M. Tinbergen. 1990. Family planning in the kestrel (*Falco tinnunculus*) - the ultimate control of covariation of laying date and clutch size. Behaviour 114:83-116.

Darwin, C. R. 1859. On the Origin of Species. London, John Murray.

Dawson, A., S. A. Hinsley, N. Ferns, R. H. C. Bonser, and L. Eccleston. 2000. Rate of moult affects feather quality: a mechanism linking current reproductive effort to future survival. Proceedings of the Royal Society of London Series B-Biological Sciences 267: 2093-2098.

Dias, P. C., and J. Blondel. 1996. Breeding time, food supply and fitness components of Blue Tits *Parus caeruleus* in Mediterranean habitats. Ibis 138:644-649.

Drent, R., and S. Daan. 1980. The prudent parent: energetic adjustments in avian breeding. Ardea 68:225-252.

Dufva, R., and K. Allander. 1995. Intraspecific variation in plumage coloration reflects immune response in Great Tit (*Parus major*) males. Functional Ecology 9: 785-789.

Dunn, P. O., and D. W. Winkler. 1999. Climate change has affected the breeding date of tree swallows throughout North America. Proceedings of the Royal Society of London Series B-Biological Sciences 266:2487-2490.

Durant, J. M., T. Anker-Nilssen, and N. C. Stenseth. 2003. Trophic interactions under climate fluctuations: the Atlantic puffin as an example. Proceedings of the Royal Society of London Series B-Biological Sciences 270:1461-1466.

Dyrcz, A., M. Wink, A. Kruszewicz, and B. Leisler. 2005. Male reproductive success is correlated with blood parasite levels and body condition in the promiscuous aquatic warbler (*Acrocephalus paludicola*). Auk 122:558-565.

Easterling, D. R., B. Horton, P. D. Jones, T. C. Peterson, T. R. Karl, D. E. Parker, M. J. Salinger et al. 1997. Maximum and minimum temperature trends for the globe. Science 277:364-367.

Edwards, M., and A. J. Richardson. 2004. Impact of climate change on marine pelagic phenology and trophic mismatch. Nature 430:881-884.

Efron, B. 1977. The efficiency of Cox's likelihood function for censored data. Journal of the American Statistical Association 72:557-565.

Esch, M. 2005. ECHAM4_OPYC_SRES_B2: 110 years coupled B2 run 6H values, DOI: 10.1594/WDCC/EH4_OPYC_SRES_B2.

Etterson, J. R., and R. G. Shaw. 2001. Constraint to adaptive evolution in response to global warming. Science 294: 151-154.

Faivre, B., A. Grégoire, M. Préault, F. Cézilly, and G. Sorci. 2003 Immune activation rapidly mirrored in a secondary sexual trait. Science 300:103.

Falconer, D. S., and T. F. Mackay. 1996, Introduction to Quantitative Genetics. New York, Longman.

Fischbacher, M., B. Naef-Daenzer, and L. Naef-Daenzer. 1998. Estimating caterpillar density on trees by collection of frass droppings. Ardea 86:121-129.

Fisher, R. A. 1958, The genetical theory of natural selection. London, Clarendon.

Fitze, P. S., M. Kölliker, and H. Richner. 2003. Effects of common origin and common environment on nestling plumage coloration in the great tit (Parus major). Evolution 57:144-150.

Forstmeier, W. 2002. Benefits of early arrival at breeding grounds vary between males. Journal of Animal Ecology 71:1-9.

Garant, D., L. E. B. Kruuk, R. H. McCleery, and B. C. Sheldon. 2004. Evolution in a Changing Environment: A Case Study with Great Tit Fledging. American Naturalist 164: E115-E129.

Gebhardt-Henrich, S. G. 1990. Temporal and spatial variation in food availability and its effect on fledgling size in the great tit. *in* J. Blondel, A. Gosler, J.-D. Lebreton, and R. McCleery, eds. Population biology of passerine birds. Berlin, Springer. pp. 175-186.

Genner, M. J., D. W. Sims, V. J. Wearmouth, E. J. Southall, A. J. Southward, P. A. Henderson, and S. J. Hawkins. 2004. Regional climatic warming drives long-term community changes of British marine fish. Proceedings of the Royal Society of London Series B-Biological Sciences 271:655-661.

Gibbs, J. P., and A. P. Breisch. 2001. Climate warming and calling phenology of frogs near Ithaca, New York, 1900-1999. Conservation Biology 15:1175-1178.

Gienapp, P. 1999. Suchverhalten von *Myoleja lucida* (Diptera, Tephritidae) und ihrem Parasitoiden *Halticoptera laevigata* (Hymenoptera, Pteromalidae). Diplomarbeit thesis, Christian Albrechts Universtät, Kiel.

Glutz von Blotzheim, U. N. 1993. Handbuch der Vögel Mitteleuropas. Wiesbaden, Aula Verlag.

Goodwin, T. W. 1980, The Biochemistry of the Carotenoids. London, Chapman & Hall.

Grant, P. R., and B. R. Grant. 1995. Predicting microevolutionary responses to directional selection on heritable variation. Evolution 49: 241-251.

Graveland, J., and R. Drent. 1997. Calcium availability limits breeding success of passerines on poor soils. Journal of Animal Ecology 66:279-288.

Greenwood, P. J., P. H. Harvey, and C. M. Perrins. 1979. The role of dispersal in the great tit (*Parus major*): the causes, consequences and heritability of natal dispersal. Journal of Animal Ecology 48: 123-142.

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

Groeneveld, E. 1995. REML VCE: a multivariate multi model restricted maximum likelihood (co)variance component estimation package, version 3.2.

Groeneveld, E., and M. Kovac. 1990. A generalized computing procedure for setting up and solving mixed linear-Models. Journal of Dairy Science 73:513-531.

Groeneveld, E., M. Kovac, T. L. Wang, and R. L. Fernando. 1992. Computing algorithms in a general purpose BLUP package for multivariate prediction and estimation. Archiv für Tierzucht-Archives of Animal Breeding.

Haccou, P., S. J. Devlas, J. J. M. van Alphen, and M. E. Visser. 1991. Information-processing by foragers - effects of Intra-patch experience on the leaving tendency of *Leptopilina heterotoma*. Journal of Animal Ecology 60:93-106.

Halpin, P. N. 1997. Global climate change and natural-area protection: Management responses and research directions. Ecological Applications 7:828-843.

Hamilton, W. D., and M. Zuk. 1982. Heritable true fitness and bright birds: a role for parasites? Science 218:384-387.

Hasselqvist, D., M. F. Wasson, and D. W. Winkler. 2001. Humoral immunocempetence correlates with date of egg-laying and reflects work load in female tree swallows. Behavioral Ecology 12:93-97.

Hatchwell, B. J. 1991. An experimental-study of the effects of timing of breeding on the reproductive success of common guillemots (*Uria aalge*). Journal of Animal Ecology 60:721-736.

Hemborg, C., and A. Lundberg. 1998. Costs of overlapping reproduction and moult in passerine birds: an experiment with the pied flycatcher. Behavioral Ecology and Sociobiology 43:19-23.

Hemerik, L., G. Driessen, and P. Haccou. 1993. Effects of intra-patch experiences on patch time, search time and searching efficiency of the parasitoid *Leptopilina clavipes*. Journal of Animal Ecology 62:33-44.

Hendry, A. P., and M. T. Kinnison. 1999. The pace of modern life: Measuring rates of contemporary microevolution. Evolution 53:1637-1653.

Hill, G. E. 1991. Plumage coloration is a sexually selected indicator of male quality. Nature 350:337-339.

—. 1993. Male mate choice and the evolution fo female plumage coloration in the House Finch. Evolution 47:1515-1525.

Hill, G. E., and R. Montgomerie. 1994. Plumage color signals nutritional condition in the house finch. Proceedings of the Royal Society of London Series B-Biological Scienes 258:47-52.

Hill, G. E., P. M. Nolan, and A. M. Stoehr. 1999. Pairing success relative to male plumage redness and pigment symmetry in the house finch: temporal and geographic constancy. Behavioral Ecology 10:48-53.

Holbrook, S. J., R. J. Schmitt, and J. S. Stephens. 1997. Changes in an assemblage of temperate reef fishes associated with a climate shift. Ecological Applications 7:1299-1310.

Hõrak, P., I. Ots, H. Vellau, C. Spottiswoode, and A. P. Møller. 2001. Carotenoid-based plumage coloration reflects hemoparasite infection and local survival in breeding great tits. Oecologia 126:166-173.

Hõrak, P., L. Saks, U. Karu, I. Ots, P. F. Surai, and K. J. McGraw. 2004. How coccidian parasites affect health and appearance of greenfinches. Journal of Animal Ecology 73:935-947.

Houde, A. E., and J. A. Endler. 1990. Correlated evolution of female mating preferences and male colour patterns in the guppy (*Poecilia reticulata*). Science 248:1405-1408.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell et al. 2001. Climate change 2001. The scientific basis. Cambridge, Cambridge University Press.

Hunt, S., I. C. Cuthill, A. T. D. Bennett, and R. Griffiths. 1999. Preferences for ultraviolet partners in the blue tit. Animal Behaviour 58:809-815.

Hüppop, O., and K. Hüppop. 2003. North Atlantic Oscillation and timing of spring migration in birds. Proceedings of the Royal Society of London Series B-Biological Sciences 270:233-240.

Inouye, D. W., B. Barr, K. B. Armitage, and B. D. Inouye. 2000. Climate change is affecting altitudinal migrants and hibernating species.

Iverson, L. R., and A. M. Prasad. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. Ecological Monographs 68:465-485.

Janzen, F. J. 1994. Climate Change and Temperature-Dependent Sex Determination in Reptiles. Proceedings of the National Academy of Sciences of the United States of America 91:7487-7490.

Jawor, J. M., S. U. Linville, S. M. Beall, and R. Breitwisch. 2003. Assortative mating by multiple ornaments in northern cardinals (*Cardinalis cardinalis*). Behavioral Ecology 14:515-520.

Jenni, L., and R. Winkler. 1994. Moult and ageing of European passerines. London, Academic Press.

Johnsen, A., K. Delhey, S. Andersson, and B. Kempenaers. 2003. Plumage colour in nestling blue tits: sexual dichromatism, condition dependence and genetic effects. Proceedings of the Royal Society of London Series B-Biological Sciences 270:1263-1270.

Johnstone, R. A. 1997. The tatics of mutual mate choice and competitive search. Behavioral Ecology and Sociobiology 40:51-59.

Jones, I. L., and F. M. Hunter. 1993. Mutual sexual selection in a monogamous seabird. Nature 362:238-239.

Keller, L. F., and A. J. van Noordwijk. 1994. Effects of local environmental conditions on nestling growth in the great tit (*Parus major* L.). Ardea 82:349-362.

Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang et al. 2001. The strength of phenotypic selection in natural populations. American Naturalist 157:245-261.

Kleinbaum, D. G. 1996, Survival analysis - a self-learning text. New York, Springer.

Klomp, H. 1970. Determination of clutch-size in birds - a review. Ardea 58:1-147.

Kluyver, H. N. 1951. The population ecology of the great tit Parus m. major. Ardea 39:1-135.

Kodric-Brown, A., and P. F. Nicoletto. 2001. Female choice in the guppy (*Poecilia reticulata*): the interaction between male color and display. Behavioral Ecology and Sociobiology 50:346-351.

Koenraadt, C. J. M., S. Majambere, W. Takken, and L. Hemerik. 2004. The effects of food and space on the occurrence of cannibalism and predation among larvae of *Anopheles gambiae s.l.* Entomologia Experimentalis et Applicata 112:125-134.

Krokene, C., K. Rigstad, M. Dale, and J. T. Lifjeld. 1998. The function of extrapair paternity in blue tits and great tits: good genes or fertility insurance? Behavioral Ecology 9:649-656.

Kruuk, L. E. B. 2004. Estimating genetic parameters in natural populations using the 'animal model'. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 359:873-890.

Lack, D. 1950. The breeding seasons of European birds. Ibis 92:288-316.

Lambrechts, M. M., J. Blondel, M. Maistre, and P. Perret. 1997. A single response mechanism is responsible for evolutionary adaptive variation in a bird's laying date, Pages 5153-5155, Proceedings of the National Academy of Sciences of the United States of America.

Lambrechts, M. M., and P. Perret. 2000. A long photoperiod overrides non-photoperiodic factors in blue tits' timing of reproduction. Proceedings of the Royal Society of London Series B-Biological Sciences 264:585-588.

Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. Evolution 37:1210-1226.

—. 1985. Evolution of mating preference and sexual dimorphism. Journal of Theoretical Biology 117:651-664.

Liebhold, A. M., and J. S. Elkinton. 1988a. Techniques for estimating the density of late-instar gypsy-moth, *Lymantria dispar* (Lepidoptera, Lymantriidae), populations using frass drop and frass production measurements. Environmental Entomology 17:381-384.

—. 1988b. Estimating the density of larval gypsy-moth, *Lymantria dispar* (Lepidoptera, Lymantriidae), populations using frass drop and frass production measurements - sources of variation and sample size. Environmental Entomology 17:385-390.

Lubjuhn, T., S. Strohbach, J. Brun, T. Gerken, and J. T. Epplen. 1999. Extra-pair paternity in great tits (*Parus major*) - A long term study. Behaviour 136:1157-1172.

Luterbacher, J., D. Dietrich, E. Xoplaki, M. Grosjean, and H. Wanner. 2004. European seasonal and annual temperature variability, trends, and extremes since 1500. Science 303:1499-1503.

Lynch, M., and B. Walsh. 1998, Genetics and analysis of quantitative traits. Sunderland, Massachusetts, Sinauer.

MacColl, A. D. C., and B. J. Hatchwell. 2003. Heritability of parental effort in a passerine bird. Evolution 57:2191-2195.

MacDougall, A. K., and R. Montgomerie. 2003. Assortative mating by carotenoid-based plumage colour: a quality indicator in American goldfinches, *Carduelis tristis*. Naturwissenschaften 90:464-467.

Martin, T. E. 1987. Food as a limit on breeding birds - a life-history perspective. Annual Review of Ecology and Systematics 18:453-487.

McCarty, J. P. 2001. Ecological consequences of recent climate change. Conservation Biology 15:320-331.

McCleery, R., and C. Perrins. 1998. ... temperature and egg-laying trends. Nature 391:30-31.

McGraw, K. J., and D. R. Ardia. 2003. Carotenoids, immunocompetence, and the information content of sexual colors: An experimental test. American Naturalist 162:704-712.

—. 2004. Immunoregulatory activity of different dietary catotenoids in male zebra finches. Chemoecology 14:25-29.

Meijer, T., and R. Drent. 1999. Re-examination of the capital and income dichotomy in breeding birds. Ibis 141:399-414.

Meijer, T., U. Nienaber, U. Langer, and F. Trillmich. 1999. Temperature and timing of egg-laying of European Starlings, Pages 124-132, Condor.

Menzel, A., and P. Fabian. 1999. Growing season extended in Europe. Nature 397:659-659.

Merilä, J., L. E. B. Kruuk, and B. C. Sheldon. 2001a. Cryptic evolution in a wild bird population. Nature 412:76-79.

Merilä, J., B. C. Sheldon, and L. E. B. Kruuk. 2001b. Explaining stasis: microevolutionary studies in natural populations. Genetica 112:199-222.

Milinski, M., and T. C. M. Bakker. 1990. Female sticklebacks use male colouration in mate choice and hence avoid parasitised males. Nature 344:330-333.

Møller, A. P., C. Biard, J. D. Blount, D. C. Houston, P. Ninni, N. Saino, and P. F. Surai. 2000. Carotenoid-dependent signals: Indicators of foraging efficiency, immunocompetence or detoxification ability? Avian and Poultry Biology Reviews 11:137-159.

Monaghan, P., R. G. Nager, and D. C. Houston. 1998. The price of eggs: increased investment in egg production reduces the offspring rearing capacity of parents. Proceedings of the Royal Society of London Series B-Biological Sciences 265:1731-1735.

Moreno, J., P. Yorio, P. Garcia-Borboroglu, J. Potti, and S. Villar. 2002. Health state and reproductive output in Magellanic penguins (*Spheniscus magellanicus*). Ethology Ecology & Evolution 14:19-28.

Muma, K. E., and P. J. Weatherhead. 1989. Male traits expressed in females: direct or indirect selection? Behavioral Ecology and Sociobiology 25:23-31.

Naef-Daenzer, B., F. Widmer, and M. Nuber. 2001. Differential post-fledging survival of great and coal tits in relation to their condition and fledging date. Journal of Animal Ecology 70:730-738.

Naef-Daenzer, L., B. Naef-Daenzer, and R. G. Nager. 2000. Prey selection and foraging performance of breeding Great Tits *Parus major* in relation to food availability. Journal of Avian Biology 31:206-214.

Nager, R. G., C. Rüegger, and A. J. van Noordwijk. 1997. Nutrient or energy limitation on egg formation: A feeding experiment in great tits. Journal of Animal Ecology 66:495-507.

Nager, R. G., and A. J. van Noordwijk. 1995. Proximate and Ultimate Aspects of Phenotypic Plasticity in Timing of Great Tit Breeding in a Heterogeneous Environment. American Naturalist 146:454-474.

Newman, R. A. 1994. Genetic variation for phenotypic plasticity in the larval life history of spadefoot toads (*Scaphiopus couchii*). Evolution 48:1773-1785.

Nilsson, J.-Å. 1994. Energetic bottle-necks during breeding and the reproductive cost of being too early. Journal of Animal Ecology 63:200-208.

Nilsson, J.- Å., and E. Svensson. 1996. The cost of reproduction: a new link between current reproductive effort and future reproductive success. Proceedings of the Royal Society of London Series B-Biological Sciences 263:711-714.

Nizinski, J. J., and B. Saugier. 1988. A model of leaf budding and development for a mature *Quercus* forest. Journal of Applied Ecology 25:643-652.

Nussey, D. H., T. H. Clutton-Brock, D. A. Elston, S. D. Albon, and L. E. B. Kruuk. 2005. Phenotypic plasticity in a maternal trait in red deer. Journal of Animal Ecology 74:387-396.

Olson, V. A., and I. P. F. Owens. 1998. Costly sexual signals: are carotenoids rare, risky or required? Trends in Ecology & Evolution 13:510-514.

Owens, I. P. F., and D. B. A. Thompson. 1994. Sex differences, sex ratios and sex roles. Proceedings of the Royal Society of London Series B-Biological Scienes 258:93-99.

Parmesan, C., N. Ryrholm, C. Stefanescu, J. K. Hill, C. D. Thomas, H. Descimon, B. Huntley et al. 1999. Poleward shifts in geographical ranges of butterfly species associated with regional warming. Nature 399:579-583.

Parmesan, C., and G. Yohe. 2003. A globally coherent fingerprint of climate change impacts across natural systems. Nature 421:37-42.

Pearce-Higgins, J. W., and D. W. Yalden. 2004. Habitat selection, diet, arthropod availability and growth of a moorland wader: the ecology of European Golden Plover Pluvialis apricaria chicks. Ibis 146:335-346.

Pearce-Higgins, J. W., D. W. Yalden, and M. J. Whittingham. 2005. Warmer springs advance the breeding phenology of golden plovers *Pluvialis apricaria* and their prey (Tipulidae). Oecologia 143:470-476.

Peñuelas, J., I. Filella, and P. Comas. 2002. Changed plant and animal life cycles from 1952 to 2000 in the Mediterranean region. Global Change Biology 8:531-544.

Perdeck, A. C., and A. J. Cave. 1992. Laying date in the coot - effects of age and mate choice. Journal of Animal Ecology 61:13-19.

Perdeck, A. C., M. E. Visser, and J. H. Van Balen. 2000. Great Tit Parus major survival, and the beech-crop cycle. Ardea 88:99-108.

Perrins, C. M. 1970. The timing of bird's breeding seasons. Ibis 112:242-255.

—. 1991. Tits and their caterpillar food supply. Ibis 133 suppl.:49-54.

—. 1996. Eggs, egg formation and the timing of breeding. Ibis 138:2-15.

Perrins, C. M., and R. McCleery. 1989. Laying Dates and Clutch Size in the Great Tit. Wilson Bulletin 101:236-253.

Perry, A. L., P. J. Low, J. R. Ellis, and J. D. Reynolds. 2005. Climate Change and Distribution Shifts in Marine Fishes. Science 308:1912-1915.

Peters, A., A. G. Denk, K. Delhey, and B. Kempenaers. 2004. Carotenoid-based bill colour as an indicator of immunocompetence and sperm performance in male mallards. Journal of Evolutionary Biology 17:1111-1120.

Pigliucci, M. 2001, Phenotypic plasticity: Synthesis in Ecology and Evolution. Baltimore, The John Hopkins University Press.

Pigliucci, M., H. Pollard, and M. Cruzan. 2003. Comparative studies of evolutionary responses to light environments in *Arabidopsis*. American Naturalist 161:68-82.

Pinheiro, J. C., and D. M. Bates. 2000, Mixed-Effects Models in S and S-Plus. New York, Springer.

Postma, E., and A. J. van Noordwijk. 2005a. Gene flow maintains a large genetic difference in clutch size at a small spatial scale. Nature 433:65-68.

—. 2005b. Genetic variation for clutch size in natural populations of birds from a reaction norm perspective. Ecology, in press.

Price, T., M. Kirkpatrick, and S. J. Arnold. 1988. Directional Selection and the Evolution of Breeding Date in Birds, Pages 798-799, Science.

Przybylo, R., B. C. Sheldon, and J. Merilä. 2000. Climatic effects on breeding and morphology: evidence for phenotypic plasticity. Journal of Animal Ecology 69:395-403.

R Development Core Team. 2004. R: A language and environment for statistical computing. Version 2.1.0. R Foundation for Statistical Computing, Vienna.

Rausher, M. D. 1992. The Measurement of Selection on Quantitative Traits - Biases Due to Environmental Covariances between Traits and Fitness. Evolution 46:616-626.

Réale, D., A. G. McAdam, S. Boutin, and D. Berteaux. 2003. Genetic and plastic responses of a northern mammal to climate change. Proceedings of the Royal Society of London Series B-Biological Sciences 270:591-596.

Reid, J. M., P. Monaghan, and G. D. Ruxton. 2000. Resource allocation between reproductive phases: the importance of thermal conditions in determining the cost of incubation. Proceedings of the Royal Society of London Series B-Biological Sciences 267:37-41.

Rindorf, A., S. Wanless, and M. P. Harris. 2000. Effects of changes in sandeel availability on the reproductive output of seabirds. Marine Ecology-Progress Series 202:241-252.

Robertson, R. J., and W. B. Rendell. 2001. A long-term study of reproductive performance in tree swallows: the influence of age and senescence on output. Journal of Animal Ecology 70:1014-1031.

Root, T. L., J. T. Price, K. R. Hall, S. H. Schneider, C. Rosenzweig, and J. A. Pounds. 2003. Fingerprints of global warming on wild animals and plants. Nature 421:57-60.

Roy, D. B., and T. H. Sparks. 2000. Phenology of British butterflies and climate change. Global Change Biology 6:407-416.

Saks, L., I. Ots, and P. Hõrak. 2003. Carotenoid-based plumage coloration of male greenfinches reflects health and immunocompetence. Oecologia 134:301-307.

Scheiner, S. M. 1993. Genetics and evolution of phenotypic plasticity. Annual Review of Ecology and Systematics 24:35-68.

—. 2002. Selection experiments and the study of phenotypic plasticity. Journal of Evolutionary Biology 15:889-898.

Schiegg, K., G. Pasinelli, J. R. Walters, and S. J. Daniels. 2002. Inbreeding and experience affect response to climate change by endangered woodpeckers. Proceedings of the Royal Society of London Series B-Biological Sciences 269:1153-1159.

Schmidt-Nielsen, K. 1983, Animal physiology: adaptation and environment. Cambridge, Cambridge University Press.

Senar, J. C., J. Figuerola, and J. Pascual. 2002. Brighter yellow blue tits make better parents. Proceedings of the Royal Society of London Series B-Biological Sciences 269:257-261.

Serra, L. 2001. Duration of primary moult affects primary quality in Grey Plovers *Pluvialis squatarola*. Journal of Avian Biology 32:377-380.

Sheldon, B. C., L. E. B. Kruuk, and J. Merilä. 2003. Natural selection and inheritance of breeding time and clutch size in the collared flycatcher. Evolution 57:406-420.

Slagsvold, T., and J. T. Lifjeld. 1985. Variation in plumage colour of Great tit *Parus major* in relation to habitat, season and food. Journal of Zoology 206:321-328.

Sparks, T. H., and P. D. Carey. 1995. The responses of species to climate over 2 centuries - an analysis of the Marsham Phenological Record, 1736-1947. Journal of Ecology 83:321-329.

Stenseth, N. C., and A. Mysterud. 2002. Climate, changing phenology, and other life history and traits: Nonlinearity and match-mismatch to the environment. Proceedings of the National Academy of Sciences of the United States of America 99:13379–13381.

Stevenson, I. R., and D. M. Bryant. 2000. Climate change and constraints on breeding. Nature 406:366-367.

Stinchcombe, J. R., L. A. Dorn, and J. Schmitt. 2004. Flowering time plasticity in *Arabidopsis thaliana*: a reanalysis of Westerman & Lawrence (1970). Journal of Evolutionary Biology 17:197-207.

Stinchcombe, J. R., M. T. Rutter, D. S. Burdick, P. Tiffin, M. D. Rausher, and R. Mauricio. 2002. Testing for environmentally induced bias in phenotypic estimates of natural selection: Theory and practice. American Naturalist 160:511-523.

Stirling, D. G., D. Réale, and D. A. Roff. 2002. Selection, structure and the heritability of behaviour. Journal of Evolutionary Biology 15:277-289.

Stott, P. A., S. F. B. Tett, G. S. Jones, M. R. Allen, J. F. B. Mitchell, and G. J. Jenkins. 2000. External control of 20th century temperature by natural and anthropogenic forcings. Science 290:2133-2137.

Strohbach, S., E. Curio, A. Bathen, J. T. Epplen, and T. Lubjuhn. 1998. Extrapair paternity in the great tit (*Parus major*): a test of the "good genes" hypothesis. Behavioral Ecology 9:388-396.

Svensson, E. 1997. Natural selection on avian breeding time: Causality, fecundity- dependent, and fecundity-independent selection. Evolution 51:1276-1283.

Thomas, C. D., A. Cameron, R. E. Green, M. Bakkenes, L. J. Beaumont, Y. C. Collingham, B. F. N. Erasmus et al. 2004. Extinction risk from climate change. Nature 427:145-148.

Thomas, D. W., J. Blondel, P. Perret, M. M. Lambrechts, and J. R. Speakman. 2001. Energetic and fitness costs of mismatching resource supply and demand in seasonally breeding birds. Science 291:2598-2600.

Tinbergen, J. M., and M. W. Dietz. 1994. Parental energy-expenditure during brood rearing in the great tit (*Parus major*) in relation to body-mass, temperature, food availability and clutch size. Functional Ecology 8:563-572.

Topp, W., and K. Kirsten. 1991. Synchronisation of pre-imaginal development and reproductive success in the Winter Moth, *Operophtera brumata* L. Journal of Applied Entomology 111:137-146.

Tryjanowski, P., S. Kuzniak, and T. Sparks. 2002. Earlier arrival of some farmland migrants in western Poland. Ibis 144:62-68.

van Balen, J. H. 1973. A comparative study of the breeding ecology of the Great Tit (Parus major) in different habitats. Ardea 61:1-93.

van der Jeugd, H. P., and R. McCleery. 2002. Effects of spatial autocorrelation, natal philopatry and phenotypic plasticity on the heritability of laying date. Journal of Evolutionary Biology 15:380-387.

van Noordwijk, A. J., R. McCleery, and C. Perrins. 1995. Selection for the Timing of Great Tit Breeding in Relation to Caterpillar Growth and Temperature. Journal of Animal Ecology 64:451-458.

van Noordwijk, A. J., J. H. van Balen, and W. Scharloo. 1981. Genetic variation in the timing of reproduction in the great tit. Oecologia 49:158-166.

van Oers, K., P. J. Drent, G. de Jong, and A. J. van Noordwijk. 2004. Additive and nonadditive genetic variation in avian personality traits. Heredity 93:496-503.

Venables, W. N., and B. D. Ripley. 2002, Modern Applied Statistics with S: Statistics and Computing. New York, Springer.

Verboven, N., and A. C. Mateman. 1997. Low frequency of extra-pair fertilizations in the Great Tit *Parus major* revealed by DNA fingerprinting. Journal of Avian Biology 28:231-239.

Verboven, N., J. M. Tinbergen, and S. Verhulst. 2001. Food, reproductive success and multiple breeding in the Great Tit *Parus major*. Ardea 89:387-406.

Verboven, N., and S. Verhulst. 1996. Seasonal variation in the incidence of double broods: The date hypothesis fits better than the quality hypothesis. Journal of Animal Ecology 65:264-273.

Verboven, N., and M. E. Visser. 1998. Seasonal variation in local recruitment of great tits: the importance of being early. Oikos 81:511-524.

Verhulst, S., and J. M. Tinbergen. 1991. Experimental-evidence for a causal relationship between timing and success of reproduction in the great tit *Parus major*. Journal of Animal Ecology 60:269-282.

Verhulst, S., J. H. van Balen, and J. M. Tinbergen. 1995. Seasonal decline in reproductive success of the great tit - variation in time or quality. Ecology 76:2392-2403.

Via, S., R. Gomulkiewicz, G. De Jong, S. M. Scheiner, C. D. Schlichting, and P. H. van Tienderen. 1995. Adaptive phenotypic plasticity - consensus and controversy. Trends in Ecology & Evolution 10:212-217.

Visser, M. E., F. Adriaensen, J. H. van Balen, J. Blondel, A. A. Dhondt, S. van Dongen, C. du Feu et al. 2003. Variable responses to large-scale climate change in European *Parus* populations. Proceedings of the Royal Society of London Series B-Biological Sciences 270:367-372.

Visser, M. E., C. Both, and M. M. Lambrechts. 2004. Global climate change leads to mistimed avian reproduction. Advances in Ecological Research 35:89-110.

Visser, M. E., and L. J. M. Holleman. 2001. Warmer springs disrupt the synchrony of oak and winter moth phenology. Proceedings of the Royal Society of London Series B-Biological Sciences 268:289-294.

Visser, M. E., and M. M. Lambrechts. 1999. Information constraints in the timing of reproduction in temperate zone birds: Great and Blue Tits. *in* N. Adams, and R. Slotow, eds. 22 International Ornithological Congress.

Visser, M. E., and C. M. Lessells. 2001. The costs of egg production and incubation in great tits (*Parus major*). Proceedings of the Royal Society of London Series B-Biological Sciences 268:1271-1277.

Visser, M. E., A. J. van Noordwijk, J. M. Tinbergen, and C. M. Lessells. 1998. Warmer springs lead to mistimed reproduction in great tits (*Parus major*). Proceedings of the Royal Society of London Series B-Biological Sciences 265:1867-1870.

Voigt, W., J. Perner, A. J. Davis, T. Eggers, J. Schumacher, R. Bahrmann, B. Fabian et al. 2003. Trophic levels are differentially sensitive to climate. Ecology 84:2444-2453.

von Schantz, T., S. Bensch, M. Grahn, D. Hasselquist, and H. Wittzell. 1999. Good genes, oxidative stress and condition-dependent sexual signals. Proceedings of the Royal Society of London Series B-Biological Sciences 266:1-12.

Wajnberg, E., P. A. Gonsard, E. Tabone, C. Curty, N. Lezcano, and S. Colazza. 2003. A comparative analysis of patch-leaving decision rules in a parasitoid family. Journal of Animal Ecology 72:618-626.

Walther, G. R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J. M. Fromentin et al. 2002. Ecological responses to recent climate change. Nature 416:389-395.

Watt, A. D., and A. M. McFarlane. 2002. Will climate change have a different impact on different trophic levels? Phenological development of winter moth *Operophtera brumata* and its host plants. Ecological Entomology 27:254-256.

Weis, A. E., and W. L. Gorman. 1990. Measuring selection on reaction norms: an exploration of the *Eurosta-Solidago* system. Evolution 44:820-831.

Weishampel, J. F., D. A. Bagley, and L. M. Ehrhart. 2004. Earlier nesting by loggerhead sea turtles following sea surface warming. Global Change Biology 10:1424.

Wheelwright, N. T., and C. B. Schultz. 1994. Age and Reproduction in Savanna Sparrows and Tree Swallows. Journal of Animal Ecology 63:686-702.

Williams, E. S., E. E. Bolitho, and S. Fox. 2003. Climate change in Australian tropical rainforests: an impending environmental catastrophe. Proceedings of the Royal Society of London Series B-Biological Sciences 270:1887-1892.

Wingfield, J. C., T. P. Hahn, M. Wada, L. B. Astheimer, and S. J. Schoech. 1996. Interrelationship of day length and temperature on the control of gonadal development, body mass, and fat score in white-crowned sparrows, *Zonotrichia leucophrys gambelii*. General and Comparative Endocrinology 101:242-255.

Woiwood, I. P. 1997. Detecting the effects of climate change on Lepidoptera. Journal of Insect Conservation 1:149-158.

Woltereck, R. 1909. Weitere experimentelle Untersuchungen über Artveränderung, speziell über das Wesen quantitativer Artunterschiede bei Daphniden. Verhandlungen der Deutsch-Tschechischen Zoologischen Gesellschaft 1909:110-172.

Zandt, H. S. 1994. A comparison of three sampling techniques to estimate the population-size of caterpillars in trees. Oecologia 97:399-406.

Acknowledgements

A PhD project is a big undertaking and one that is not easy to accomplish without the help, assistance, and support of others.

First of all I would like to thank Marcel Visser for his supervision. Marcel always made himself accessible to questions and discussions. Working with him was really enjoyable since he left me enough freedom to follow my own ideas and to work the way I liked, but kept a close enough watch to press things when necessary!

There were, of course, many other people in Heteren whom I would like to thank. Everybody in the Department was friendly, helpful and open for questions. Kate Lessells was great for discussion and provided excellent comments and suggestions on many manuscripts. Arie van Noordwijk, Piet Drent, Christiaan Both, Rüdiger Cordts and David Thomsen took part in many scientific discussions in journal clubs, science lunches and werkbesprekingen and made insightful suggestions. Jan Visser introduced me to the database and always had an open ear for questions about it. Henry Bouwmeester collected many data, which contributed to this thesis by catching, ringing and measuring great tits on the Hoge Veluwe. Leonard Holleman did the same in Oosterhout and, much to his credit, collected and sorted caterpillar droppings too! Rogier, Barry, Joep and Robert also helped with fieldwork as part of their various student projects. Many others besides the aforementioned also helped with the roosting inspections. By the way, many thanks to Kate for providing food afterwards. Christa Mateman could find almost everything in the stock of the mo-lab if necessary. Ab and Keultjes made various things, like 400 nest box bottoms.

Of course, there were also the other PhD students - Christel, Kees, Niels, Tobi, Erik and Margriet (in 'chronological' order) - who all contributed immensely to scientific discussions and also helped to create a nice athmosphere within which to work. Tobi measured the occasional brood for me and was generally great company during field work. Erik helped with the animal model, made many comments on manuscripts and we had many good discussions about science (and other things).

People outside the NIOO also played an important part. Dan Nussey did the phenotypic plasticity analysis and agreed to include the manuscript into my thesis. Lia Hemerik provided help with the Proportional Hazards Model. My parents

supported me during my study in Kiel and never questioned whether becoming a biologist is a wise thing to do (or not!). My friends helped to remember that there is a life besides science and doing a PhD. Last but not least, I have to thank Katja for various things, too numerous to mention.

Curriculum Vitae

Phillip Gienapp was born on 17 May 1971 in Neumünster (Germany). From 1977 to 1981 he went to primary school in Hamburg and from 1981 to 1990 to secondary school in Glückstadt. From December 1990 to February 1992 he did alternative civilian service and worked for a local nature protection society on the island of Sylt. In October 1992 he started studying biology at the Christian-Albrechts-Universität, Kiel, specialising in zoology, marine biology and oceanography. During this time, he worked as a research assistant on projects concerning the feeding ecology of eider ducks at the Forschungs- und Technologiezentrum Westküste of Kiel University, and arthropod diversity at Ökologie-Zentrum Kiel. He conducted a project on the feeding ecology of lapwing chicks in intensively and extensively managed farmland at the NABU Institut für Wiesen und Feuchtgebiete. He also took part in aerial surveys of ducks in the Waddensea and the Baltic Sea. In 1999 he finished his study with a thesis (Diplom) about foraging behaviour and chemical communication of a tephritid fly and its parasitoid at the University of Kiel. Afterwards he worked freelance in various nature conservation and monitoring projects before starting his PhD study at the Netherlands Institue of Ecology in March 2001, which resulted in this thesis. In March/April 2005 he worked at the University of Helsinki with a grant from the Spatial Ecology Visitor's Programme.

Publications

Gienapp, P., L. Hemerik & M. E. Visser (2005) A new statistical tool to predict phenology under climate change scenarios. *Global Change Biology* 11, 600-606.

Hoffmeister, T. S. & P. Gienapp (2001) Discrimination against previously searched, host-free patches by a parasitoid foraging for concealed hosts. *Ecological Entomology* 26, 487-494.

Gienapp, P. & T. S. Hoffmeister (2000) Risky communication – the exploitation of marking pheromone by a parasitoid. *Mitt. der Deutsch. Gesellsch. für Allg. und Angew. Entomologie* 12, 451-454.

Hoffmeister, T. S. & P. Gienapp (1999) Exploitation of the host's chemical communication in a parasitoid searching for concealed host larvae. *Ethology* 105, 223-232.

Nehls, G. & P. Gienapp (1997) Prey selection and foraging behaviour of Cormorants *Phalacrocorax carbo* in the Wadden Sea. *Vogelwelt* 118, 33-40.

accepted

Visser M. E., C. Both & P. Gienapp (2005) Mistimed reproduction due to global climate change. *Acta Zoologica Sinica* In press

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

submitted

Gienapp, P. & M. E. Visser. Possible fitness consequences of experimentally advanced laying dates in great tits: differences between populations in different habitats. *Functional Ecology*

Nussey, D. H., E. Postma, P. Gienapp & M. E. Visser. Selection on heritable phenotypic plasticity in a wild bird population. *Science*

Addresses of co-authors

Lia Hemerik
Biometris
Department of Mathematical and Statistical Methods
Wageningen University
P.O. Box 100
6700 AC Wageningen, The Netherlands

Leonard J. M. Holleman
Netherlands Institute of Ecology (NIOO-KNAW)
Centre for Terrestrial Ecology
P.O. Box 40
6666 ZG Heteren, The Netherlands

Daniel H. Nussey
Institute of Evolutionary Biology
University of Edinburgh
Edinburgh EH9 3JT, UK

Erik Postma
Netherlands Institute of Ecology (NIOO-KNAW)
Centre for Terrestrial Ecology
P.O. Box 40
6666 ZG Heteren, The Netherlands

Marcel E. Visser
Netherlands Institute of Ecology (NIOO-KNAW)
Centre for Terrestrial Ecology
P.O. Box 40
6666 ZG Heteren, The Netherlands