THE UNIVERSITY
of EDINBURGH

Breeding phenology and its effects on reproductive success in seabirds

Katharine Keogan

Thesis submitted for the degree of Doctor of Philosophy

University of Edinburgh

February 2019

For my dad

i. Abstract

The timing of reproduction is important for fitness, and has been used to measure the effect of widespread environmental change across ecosystems globally. Across trophic levels, species occupying higher levels of a food web are generally adjusting their timing of breeding in response to environmental change at a slower rate than their prey (Poloczanska et al., 2013; Thackeray et al., 2010). This may lead to a trophic mismatch between the energy requirements of consumers and the timing of peak availability of resources during the crucial reproductive period, potentially negatively impacting on fitness. However, the effects of environmental change have not been uniform across populations, species, or regions of the world. This makes it difficult to predict how different populations will adjust their response to environmental change and the consequences of this for fitness. Marine species are generally underrepresented in studies of environmental change, and seabirds are a group of marine organisms that may be particularly at risk. They generally occupy higher trophic levels, are long-lived, and reproduce slowly, meaning they may lack the evolutionary capacity to adapt if the timing of key resources shifts rapidly under climate change. However, the disconnected nature of previous studies of the trends and drivers of seabird breeding phenology and the effects of trophic mismatch on seabird fitness has precluded a global understanding of the extent to which seabirds will respond to climate-mediated environmental change.

In this thesis, I make use of resources contributed by a global network of collaborators to first establish the global average trends in seabird breeding phenology over time and in response to sea surface temperature. I then identify which seabird populations may be at higher risk of mismatch with prey by characterising sources of variance around these phenological trends (e.g. due to differences in phylogeny, biogeographic region, or life history traits). I go on to explore the scales at which phenology is correlated across breeding North Atlantic seabird populations, to understand whether it is likely that phenology is driven by conditions experienced by populations at the breeding grounds, overwintering locations, or across multiple spatial scales. Finally, I examine the fitness consequences of trophic mismatch between the resource and consumer in two ways. I first use 30+ years of data from the long-term monitored population of European shags *Phalacrocorax aristotelis* on the Isle of May, Scotland, to identify the impact of trophic mismatch on population- and individual-level fitness

over time and in relation to changes in SST and diet. My final data chapter expands the focus on the effects of mismatch on population level breeding success back to the global scale. In the absence of detailed information on prey availability and phenology, I develop on an existing framework that allows us to predict when phenological change may impact on population level fitness to identify whether trophic mismatch is both present in a population and getting worse over time. I use these criteria to compare relationships across populations, regions and life history traits to identify the prevalence of trophic mismatch across populations on a global scale.

ii. Lay Summary

Globally, climate is changing at alarming rates, which is negatively affecting organisms that live on land, in fresh water and in the sea. One of the key ways in which these negative effects are observed are through changes in the time of year at which organisms breed. Breeding at the right time of year often relies on organisms correctly responding to seasonal changes in their environment, like rising spring temperatures, which indicate that conditions are now suitable to raise young, or will be in the near future. Conditions such as prey availability may only be available at sufficient quantities for a few weeks in a year. Reproduction is therefore very important to time correctly, because parents need to ensure there are sufficient prey available to feed their young as they grow. It is therefore energetically demanding for both parents and their young. However, as climate changes, some organisms may not be able to respond to changing temperatures as accurately as others.

In general, animals occupying higher levels of a food chain, like birds and mammals, are responding to environmental change at a slower rate than those at lower levels, like plants and insects. This means that the timing at which organisms high up in the food chain (i.e. predators) breed may become asynchronous with the timing of breeding of those occupying lower levels (i.e. prey). As climate change increases, there may therefore be less prey available for predators to feed their young when they need it most, and as a result, each breeding season may be less successful in terms of the number of offspring produced. However, while we know that this lack of synchrony between predators and prey is occurring in some food chains, sometimes with negative consequences for successfully producing young, we currently do not know how prevalent it is on the global scale.

In marine environments our understanding of how climate change impacts organisms is limited, because most studies focus on what is happening on land. Marine birds are a group of highly threatened species, yet we know much less about how they are responding to environmental change than other avian groups. Furthermore, their timing of breeding is likely to become asynchronous with the time that prey are available because they generally occupy high levels of a food chain. In my thesis I investigated how climate change will impact when seabirds lay their eggs, and whether they are likely to be able to track the time at which prey are available for foraging parents by laying earlier in more recent or warmer years. I also examined

whether the number of chicks to fledge each nest has declined over time, or as temperatures rise. If climate change really is impacting the time at which seabirds and their prey breed at different rates, then we expect birds to fledge fewer chicks as climate change progresses.

Overall, I found no evidence that seabirds lay eggs earlier or later than they did in the past, or in years where the sea is warmer. This means that seabirds may be at risk of breeding at a time when there are not enough prey available to feed their young. However, my results showed no evidence that the number of chicks to fledge each nest has declined over time or as temperatures rise, meaning that if seabirds really are becoming asynchronous with their prey, the consequences may be evident in other ways, such as impacting adult survival or the likelihood that chicks will survive after fledging.

iii. Acknowledgements

Sue and Ally, you have been the best double act of main supervisors I could ever have hoped for. Thank you for investing so much of your time and sharing your wealth of expertise and knowledge, and for encouraging me to constantly do my very best – even when I didn't think I could. But most importantly thank you both for being a source of friendship that I value so highly, and without which I would not have made it to the end. Sue, I will never see a cardamom bun without thinking of you, and Ally, I will never see a speeding ticket without you crossing my mind.

To the rest of the gang, Francis, Sarah and Richard, I have felt completely privileged to be supervised by each of you. Thank you for contributing your time and sharing your vast quantities of knowledge, for all of your encouragement and involvement, and for travelling from near and far to have so many chats about seabirds and science throughout my PhD. What a wonderful group of people to be lucky enough to call friends.

Thank you to all the countless researchers and field assistants from long-term monitoring projects around the world. My PhD would have been impossible without your collective decades of effort and I am hugely grateful to you all. Thank you to everyone who took the time to answer my questions (often successive) over email or Skype, read numerous drafts of numerous manuscripts, and those who emailed little words of encouragement from all corners of the world as the projects progressed.

Tim. Thank you for putting up with all my shite for the last 3.5 years (sure I've been doing it a lot longer than that, says you), and for being my most supportive best friend. Thanks for listening to me talk relentlessly about seabirds and their silly breeding habits and for dropping everything when I needed your support, encouragement and big bear hugs the most. Thank you for putting things in perspective, particularly whenever I was acting like a big dope, but mostly, thank you for continuing to pay more than half the rent so I could spend my limited finances on wine.

Thanks to Jess. What a silly goose to become pals with and it has been a pleasure. Thanks for introducing me to the world of fancy wine (and for sharing so many bottles with me) and for being my partner in crime in just about everything since we started. You are an inspirational scientist and a wonderful friend.

Thanks also to Kirsty and Jack for being top drawer drinking buddies and fantastic pals; The Phenoweb and Team Shrub lab groups and the inhabitants of office 1.33 for the craic and chats; All of the girls at home, especially the Aoifes and Colleen; The Meadows bbq and beer enthusiasts, Nathan, Rich, Lydia, Ed and Sam; The CEH seabird group; The UCC seabird research group for taking me under your wing at every conference we've been to together and letting me help with the gannets; Willie, Laura, Martin, Sophie, Francis and the rest of the Lothian Ringing and Forth Seabird

Groups for providing great banter during the much needed real life bird times. Thanks to Deirdre for all the encouragement in the last few months, and to my sister Chloe for sending encouraging texts from across the Atlantic.

Finally, thank you to my mam. Your infectious love of the natural world has been the driver of everything I've done. Thank you for being eternally proud of me, which you may not realise has been the thing that has kept me going the most.

iv. Declaration

All of the data chapters have been highly collaborative, but for each chapter I collated the datasets, coordinated the project, designed and executed the analyses and wrote the first draft. I detail the additional contributions of others below.

Chapter two: the version that is presented here merges my own work with the expertise and suggestions of 88 additional co-authors and three reviewers. However, I had the final approval on the submitted draft, which is now published in Nature Climate Change.

Chapter three: the version presented in this thesis merges my own work with the suggestions of 34 additional collaborators, three reviewers and the Editor of Global Change Biology. The simulations in Appendix 3.4 were conducted by one of my supervisors (Ally Phillimore). This chapter is currently under revision for Global Change Biology.

Chapter four: I received substantial guidance on the statistical models from Ally Phillimore and Francis Daunt contributed substantial rewriting in preparation for submission to a peer reviewed journal. Additional input on biology of sandeels was given by Richard Howells.

Chapter five: the analyses in this chapter were designed by myself with guidance from my supervisors and with additional input from Nina McLean and Jarrod Hadfield. I conducted all analyses and wrote the first draft. I received comments from my supervisors. This chapter has not been reviewed by the collaborators that contributed data.

This work has not been submitted for any other degree or professional qualification.

Contents

1. General Introduction

1.1 Global environmental change

In recent decades, global temperature increases have had a detectable effect on organisms across terrestrial and aquatic biomes (Walther, 2010; Stocker *et al.*, 2013). Climate change has been linked to species decline (Hughes, 2000; Sydeman *et al.*, 2015; Spooner *et al.*, 2018), shifts in species distributions (Poloczanska *et al.*, 2013) and altered ecological processes across numerous taxa and trophic levels (Parmesan & Yohe, 2003; Parmesan, 2006; Thackeray *et al.*, 2010; Poloczanska *et al.*, 2013). As climate change is projected to continue (Stocker *et al.*, 2013), it is particularly important to identify the species most at risk or sensitive to changing environments. However, as data are often deficient and because the effects of warming have not been uniform across regions of the world (Burrows *et al.*, 2011; Stocker *et al.*, 2013), predicting exactly how specific species or populations will respond to changing climates presents a challenge. With so many current pressures on biodiversity, it is important to understand how climate change affects the processes of organisms throughout their range, and the mechanisms by which they respond to environmental variation on a global scale (Visser, 2008).

1.2 Phenology

One of the most widely used indicators to track organismal response to climate change is that of phenology, the study of the timing of seasonally recurring biological events such as breeding or migration (Sparks & Smithers, 2002; Parmesan & Yohe, 2003; Parmesan, 2006; Janetos *et al.*, 2012). Reproductive phenology is particularly important to time correctly, as reproduction plays a leading role in population dynamics (Rasmussen *et al.*, 2014) and should therefore intersect a time when both biotic and abiotic conditions are at their most favourable to maximise fitness (Poethke

et al., 2016). Organisms generally rely on cues such as photoperiod or temperature to initiate breeding (Bradshaw & Holzapfel, 2010). However, the availability of favourable conditions may only last a few weeks and can fluctuate between years and regions (Visser *et al.*, 2006; Grémillet & Boulinier, 2009; Thackeray, 2016), and organisms therefore face additional challenges in maximising fitness when adjusting breeding phenology from year to year (Bradshaw & Holzapfel, 2010). Recent global analyses have indicated that both spring temperatures and phenological events are occurring earlier now than in previous decades (Parmesan & Yohe, 2003; Thackeray *et al.*, 2010, 2016; Poloczanska *et al.*, 2013), yet it is unclear whether these phenological changes are sufficient to keep up with environmental change, and to what extent this will impact fitness.

1.3 Phenology changing at different rates

When considering how environmental and phenological change will impact absolute (i.e. the actual number of offspring) or relative (i.e. the number of offspring an individual has relative to conspecifics) fitness, it is important to understand that organisms in a community do not always respond to the same cues, in the same direction, or with similar magnitude (Visser & Both, 2005; Thackeray *et al.*, 2016). For example, lower trophic level organisms are generally adjusting their phenology at a faster rate than organisms higher in the food chain (Thackeray *et al.*, 2010; Poloczanska *et al.*, 2013). This may be due to the fact that at lower trophic levels organisms generally have shorter generation times, more plasticity in their responses to the environment, or a combination of both (Visser *et al.*, 2004). This may leave organisms at higher trophic levels more vulnerable to rapid climate change and be at risk of becoming asynchronous or mismatched with their resources during critical periods such as breeding if they are less capable of adjusting their phenology to coincide with suitable conditions (Visser *et al.*, 2004).

1.4 Consequences of phenological asynchrony

The terms asynchrony and mismatch are often used interchangeably to describe year to year differences in phenological overlap between trophic levels. However, while there is clear evidence that climate (especially temperature) leads to variation in asynchrony (Thackeray *et al.*, 2010, 2016; Poloczanska *et al.*, 2013), evidence for this leading to fitness consequences (i.e. mismatch) is somewhat lacking. For mismatch to occur, there should be evidence of a) a difference in plasticity between resource and consumer (i.e. a differential response to environmental cues) (Gienapp *et al.*, 2014), b) a decoupling between the peaks of resource availability and consumer energy requirements (Thackeray *et al.*, 2010), which may lead to negative fitness consequences (reduced survival or fecundity) experienced either by individuals or the entire population of the consumer (Thackeray, 2012; Reed *et al.*, 2013a; McLean *et al.*, 2016). Yet, often several of these criteria are not met (Thackeray, 2012), precluding evidence for mismatch across a food chain in terrestrial and marine systems. Trophic mismatch may lead to reduced fitness for consumers, both between and within species groups, and alter predator-prey dynamics, competition and ecosystem function (Forrest & Miller-rushing, 2010; Rasmussen *et al.*, 2014). However, as most of our evidence comes from a limited number of well-studied systems, in particular great tits (Visser & Both, 2005; Charmantier *et al.*, 2008; Reed *et al.*, 2013b, 2013a), we have little idea of the scale and magnitude of a problem it may pose.

1.5 Phenology in the marine environment

The Intergovernmental Panel on Climate Change has reported increased frequency of marine storms and wind speed; global rise of average sea level and surface temperature (SST), and altered patterns of ocean circulation, upwelling, and ultimately, primary productivity in the marine environment since 1970 (Brierley &

Kingsford, 2009; Aoki *et al.*, 2013). Yet, the effects of climate change on biotic systems have been understudied in marine communities in comparison with terrestrial systems, despite the fact that marine systems are thought to be extremely vulnerable to climate change (Richardson & Poloczanska, 2008; Aoki *et al.*, 2013). Of the relatively few marine studies that exist, evidence suggests that phenological shifts are advancing at different rates across trophic levels in marine habitats (Edwards & Richardson, 2004; Frederiksen *et al.*, 2006; Thackeray *et al.*, 2010; Burthe *et al.*, 2012), indicating that the mechanisms which drive phenology differ throughout marine food webs. Poloczanska *et al.*, (2013) identify that marine primary producers may be advancing spring phenological events by ~6 days/decade, zooplankton by ~12 days/decade, larval bony fish by ~6 days/decade, and seabirds may be delaying their spring phenology by ~1 day/decade, although in this study the sample sizes were in single figures for most taxonomic groups. This suggests that organisms occupying higher trophic levels may become asynchronous with key resources as they advance phenology at a slower rate than that of their prey.

However, when phenological trends that occur in summer were considered in the same study (Poloczanska *et al.*, 2013), these rates of change were much more synchronous, with phytoplankton, zooplankton and seabirds (the only three levels for which summer phenology was included) all advancing at the same rate (~4 days/decade). The large margins of error surrounding each trend suggest that there may in fact be variation around the global average for each trophic level, although this may also be a result of measurement error due to small sample sizes. Potential sources of variation in phenological slopes (i.e. within taxonomic groups and regions) were explored in Poloczanska *et al.*, (2016), although the data were visually presented rather than formally analysed. This makes it difficult to arrive at any robust conclusions about which taxonomic groups and trophic levels are truly adjusting phenology over

time, as sources of non-independence in the data (i.e. whether the time series were collected from different species, locations, or time periods), and measurement error were not taken into account. It is therefore unclear what the overall trends in phenology is at any trophic level, whether there are any patterns of variation around each trend.

1.6 Seabirds as a study system

The ecology of seabirds is thought to make them effective bio-indicators (Piatt & Sydeman, 2007; Croxall *et al.*, 2012). As a long lived group of animals with low fecundity and delayed sexual maturity, even small increases in mortality can lead to population decline (Grémillet & Boulinier, 2009; Sato *et al.*, 2016). While their behaviour varies between wide ranging pelagic and locally resident, an individual seabird may experience a broad range of environmental conditions throughout its life (Frederiksen *et al.*, 2004), and populations may evolve to use cues to adjust migration or breeding phenology accordingly (Grémillet & Boulinier, 2009). However, there are physiological limits to this ability to adjust (Grémillet & Boulinier, 2009) and the degree to which a plastic response to a phenological cue is expected to evolve depends on the reliability of the cue as a predictor of later conditions (Reed *et al.*, 2010). Alternatively, the relationship between the environment and phenology may be reflective of environmental constraints (such as snow-covered breeding sites, Watanuki *et al.*, 2009), or time taken to reach breeding condition after the winter season (Daunt *et al.*, 2014), rather than cues that forecast future environmental conditions. The relatively late recruitment age and long generation times of seabirds may make them less capable of adjusting to rapid environmental change than organisms with accelerated life histories. Furthermore, high levels of philopatry, strong pair bonds and specialised diets may mean that seabirds exhibit geographic and dietary constraints in their response to climate (Grémillet & Boulinier, 2009). As

such, seabirds have been widely used to assess ecosystem health in response to a suite of global threats, including warming climates (Reviewed in Piatt & Sydeman, 2007; Mallory *et al.*, 2010). However, the potential for seabirds to exhibit some levels of behavioural plasticity may make them less effective as ecological indicators (Grémillet & Charmantier, 2010), therefore, one of the aims of this thesis is to revisit whether or not seabirds are good indicators of marine health.

Seabirds generally forage above the primary consumer level, though vary from planktonivorous to scavenger, with diets often favouring key prey species (Grémillet & Boulinier, 2009; Thackeray *et al.*, 2016). This foraging ecology makes seabirds a particularly useful system for studying the presence of mismatch with lower trophic level prey and the effects of mismatch on fitness. Furthermore, avian systems are amenable to such work because individuals can be monitored relatively easily, particularly during the breeding season (Crick, 2004). In the marine environment, where most other organisms are concealed under water, seabirds are highly visible throughout their range. This makes them more easy to study than marine mammals, fish, or low trophic level invertebrates (Piatt & Sydeman, 2007). They are generally philopatric, found throughout the world's oceans, and return to established colonies to breed year after year (Grémillet & Boulinier, 2009; Wanless *et al.*, 2009). This makes them especially accessible as a study system, with one fifth of all seabird colonies in the world currently being monitored (Paleczny *et al.*, 2015). As a result, many long-term datasets on breeding phenology at individual colonies are available.

1.7 Trends in seabird breeding phenology

Studies across populations, species and regions of the globe have found evidence that seabirds are advancing (Bertram *et al.*, 2001), delaying (Barbraud & Weimerskirch, 2006) or showing no trend (Bond *et al.*, 2011) in their timing of breeding over time or with changing environmental conditions. However, some species may be

more sensitive or exposed to climate change than others (Williams *et al.*, 2008; Dawson *et al.*, 2011). For example, back-legged kittiwakes are heavily restricted to feeding on the surface of the water, and therefore may be more vulnerable to reduced survival or breeding success during periods of food shortage than diving species which can seek out prey, such as shags and cormorants (Furness & Tasker, 2000). Other species may have less ecological or evolutionary capacity to adjust or adapt, like many albatross species which do not breed until they are more than ten years old, and take almost a year to raise a chick (Schreiber & Burger, 2002), giving them a lower rate of evolutionary rescue than those which reproduce at a faster pace. Furthermore, populations of the same species will be exposed to different climate conditions across their range (Dawson *et al.*, 2011), particularly if they are widely distributed globally.

1.8 The drivers of seabird breeding phenology

To predict how populations will respond to climate change in the future, it is important to identify the cues and drivers to which organisms respond, and the scale at which these drivers act. Previous studies have linked breeding phenology in seabirds to a range of climate drivers including local SST and upwelling indices (Wolf *et al.*, 2009), and large scale drivers like the El Niño Southern Oscillation (ENSO) or North Atlantic Oscillation (NAO) (Frederiksen *et al.*, 2004; Surman & Nicholson, 2009). However, varying methods and results have impeded a complete understanding of the drivers of phenology across populations. When considering which drivers are most influential to breeding phenology, both the scale of the driver and the life history traits of the seabirds are important (Frederiksen *et al.*, 2004; Afán *et al.*, 2015). Species which spend the non-breeding season far from the colony may be less likely to follow local environmental cues to initiate breeding than those who are year round residents (Frederiksen *et al.*, 2004). Long-distance migrants might instead rely on circannual

rhythms or body condition to initiate migration (Visser *et al.*, 2004), with carry-over effects on their breeding phenology. This difference may even be observed between closely related species, like sympatric *Pygoscelis* penguins which show noticeably different breeding phenologies due to their differing overwintering strategies (Lynch *et al.*, 2012). The response observed by an individual population may not be indicative of the full range of phenotypic plasticity of the species, and therefore to understand the range of responses a species is capable of there is value to considering multiple populations (Thackeray, 2016).

1.9 Phenological asynchrony and trophic mismatch in seabirds

Once trends and drivers of phenology in the context of environmental change have been identified, the next step is to understand the consequences for the birds. Previous studies have investigated the impacts of potentially mistimed phenology on the demography of individual seabird populations with contrasting results (Frederiksen *et al.*, 2004; Gaston *et al.*, 2009; Shultz *et al.*, 2009; Sorensen *et al.*, 2009; Watanuki *et al.*, 2009; Bond *et al.*, 2011; Burthe *et al.*, 2012; Ramirez *et al.*, 2016; Youngflesh *et al.*, 2017), and there are still clear gaps in our knowledge of what this means for seabird populations globally. It is plausible that seabirds may become asynchronous with their prey; short-lived organisms and those occupying lower trophic levels tend to adjust their phenology at a faster rate than long-lived, higher trophic level organisms like seabirds (Visser *et al.*, 2004; Poloczanska *et al.*, 2013). However, direct tests of increasing asynchrony between higher predators and their prey are challenging, as it is often difficult to obtain time series' on marine prey that are both long enough and at a sufficiently fine scale to address questions about mismatch (but see Youngflesh et al., 2017 for a good example in seabird populations). Few studies directly link changes in phenology of predators with that of their specific prey to detect asynchrony (discussed in a marine context in David Grémillet et al.,

2008), and even fewer go on to test the resulting consequences for breeding success (but for a good example in a terrestrial system see Reed et al., 2013). Furthermore, the disconnection between methods, the disparity in the quality of data used across studies and the lack of clearly laid out predictions to facilitate interpretation of results has precluded an understanding of the extent to which deleterious effects of mismatch are observed on a global scale. As such, our understanding of the presence and effects of trophic mismatch in marine higher predators remains poorly understood (Richardson & Poloczanska, 2008; Thackeray *et al.*, 2010).

1.10 Insights from individual studies

Much of what we know about how organisms respond to environmental change comes from detailed studies of individual populations. These studies are invaluable, because they have allowed us to answer questions about how certain aspects of environmental change may be impacting each population in a unique way. For example, many high-latitude seabird populations are constrained in their reproductive phenology by the presence of ice and snow during the breeding season (Gaston *et al.*, 2009; Moe *et al.*, 2009; Watanuki *et al.*, 2009). However, Alaskan black guillemots *Cepphus grylle mandtii* rely on summer pack ice to forage during the chick rearing period, and reduced ice concentrations during the breeding season negatively impact reproductive success (Divoky *et al.*, 2015). Understanding their ecology has allowed the correct variables to be considered in an individual analysis, and more accurate predictions to be made about what we expect to find.

Individual studies can also be beneficial where data at the level of a sample of individuals within a population is available. Population level responses, even estimated within a single study, may overlook the mechanisms underpinning variation in phenological plasticity and breeding success between individuals (Reed *et al.*, 2009, 2013a; Szostek *et al.*, 2015). Considering the relationships and potential

differences between population- and individual-level phenological responses and reproductive success are key to understanding how organisms and populations may respond to environmental change (van de Pol & Wright, 2009).

1.11 Linking individual studies for a broad-scale perspective

An effective method for understanding species' responses on a global scale and identifying variation around the average trend is to conduct a cross-species comparative analysis of long-term time series. Scientific generalisations are helpful because they increase our confidence in the evidence for climate change impacts, more so than when studies are considered independently (Oro, 2014). They can be useful for considering differences observed between species and populations due to spatial, trait based and taxonomic variation, and indeed, this approach is recommended by global change biologists (Thackeray, 2016). Comparisons of individual studies have been made within and between taxa (Møller *et al.*, 2008; Dalleau *et al.*, 2012), and across trophic levels (Thackeray *et al.*, 2016), habitats (Thackeray *et al.*, 2010) and regions of the globe (Poloczanska *et al.*, 2016) to highlight the influences of climate change across taxonomic groups. Life history traits, such as range distance, body size, and foraging strategy can also affect a species' ability to respond to climate, potentially resulting in differential responses across populations of the same species (Stevenson & Bryant, 2000; Sandvik & Erikstad, 2008; Sabarros *et al.*, 2012; Paleczny *et al.*, 2015). Comparative analyses can therefore reveal the drivers of inter-, intra-, and trait-specific variation in phenological trends and their consequences for fitness in seabirds, for which more information is currently needed at all scales (Sandvik & Erikstad, 2008; Thackeray, 2016).

Meta-analyses of previously published time series of seabird breeding phenology have been carried out on populations from both the Northern (Sydeman *et al.*, 2012) and Southern Hemispheres (Chambers *et al.*, 2013) and at the global scale

(Poloczanska *et al.*, 2013, 2016). These analyses have merits, however, caution should be taken when comparing previously published time series, as bias of publication towards positive or interesting results may lead to the trends being over estimated (Thackeray *et al.*, 2010). An alternative method which had not yet been attempted in meta-analyses of seabird phenology is to collect raw data from known researchers in the field of interest to include unpublished time series. This ensures that species with positive, negative and neutral responses are all reported and is the approach taken by Thackeray *et al.*, (2010, 2016). Furthermore, when comparing traits or levels of response among species, it is important to statistically consider the measurement error associated with each study (Ives *et al.*, 2007). The standard error associated with each individual analysis can account for various sources of measurement error and can be included in a meta-analysis (Ives *et al.*, 2007; Hadfield & Nakagawa, 2010), to give a more accuate estimate of the mean trend, and identify whether there is true variation around it.

1.12 Thesis outline and aims

In this thesis I assess the extent to which seabirds adjust their breeding phenology between years, investigate phenological drivers, and test for fitness consequences of phenological asynchrony with lower trophic levels in the form of reduced breeding success. Overall, my thesis will include a global analysis of trends in seabird breeding phenology (**Chapter 2**); an investigation into the spatial scale at which seabirds respond to environmental cues (**Chapter 3**); and look at the consequences of potential trophic mismatch within a single population (**Chapter 4)** and across multiple populations on a global scale (**Chapter 5**).

The first step in understanding how seabirds will respond to environmental change is to identify the phenological responses of seabirds to environmental change. In **Chapter 2** I will do this by combining 209 raw phenological time series from 61

species to undertake a global phylogenetic meta-analysis of seabird breeding phenology. I will look at trends over time and with spring sea surface temperature (sSST). In addition to estimating average trends I will assess whether slopes show any patterns across species, regions or with different life histories. Identifying trends across species groups will allow us to further understand which species may be more vulnerable to changing environments. However, it is important to remember that sSST is just one of many potential environmental variables that seabirds may respond to, although the only one that can be readily compared across populations on a global scale. To identify which species may be more responsive to another unmeasured aspect of their environment, I will also identify which species have the highest levels of variance in breeding phenology between years.

Some seabird species, such as the shag, are known to vary substantially in their breeding time from year to year, suggesting that one or more aspect of the environment may impact on the timing of breeding. However, understanding what drives phenology in seabirds is challenging, as they potentially encounter a wide range of conditions both within and outside of the breeding season. It is currently unclear whether phenology is driven by conditions experienced on a local scale, over winter, or a combination of both. Therefore, in **Chapter 3** I will investigate the scale at which phenology is driven in the North Atlantic by estimating the extent to which phenology correlates across 51 populations from 9 species in this ocean basin. If all populations at a breeding site, in a larger region, sharing the same wintering location or of the same species share early or late years, then this may indicate that they are responding to a similar environmental driver. Understanding the scale at which phenology is driven is a useful step in allowing us to predict how species will respond to environmental change.

Given my finding in **Chapter 2** that seabirds are generally not shifting their phenology in a consistent way in response to warming temperatures, there is potential for them to suffer reductions in fitness if prey at lower trophic levels are adjusting their phenology at a faster rate. In **Chapter 4** I will test for evidence of increasing trophic mismatch at the population and individual levels in a population of European shags *Phalacrocorax aristotelis*, by combining data on individual and population level phenology and breeding success with three proxies of trophic mismatch (time, SST and diet composition). I will present a set of clear predictions that must be met if trophic mismatch is present and increasing at both the individual and population levels. In **Chapter 5** I will test for evidence for trophic mismatch across seabird populations globally by combining data on temperature change, annual phenology and population level breeding success. Although this chapter lacks information on phenology of prey, I will present a set of clear predictions and criteria that must be met for mismatch with prey to be evident in populations. I will also test whether mismatch is observed at similar magnitudes across populations occupying different latitudes and with a range of life history traits.

2. Global phenological insensitivity to shifting ocean temperatures among seabirds

This Chapter is published as:

Keogan, K., Daunt, F., Wanless, S., Phillips, R. A., Walling, C. A., Agnew, P., Ainley, D. G., Anker-Nilssen, T., Ballard, G., Barrett, R. T., Barton, K. J., Bech, C., Becker, P., Berglund, P.-A., Bollache, L., Bond, A. L., Bouwhuis, S., Bradley, R. W., Burr, Z. M., Camphuysen, K., Catry, P., Chiaradia, A., Christensen-Dalsgaard, S., Cuthbert, R., Dehnhard, N., Descamps, S., Diamond, T., Divoky, G., Drummond, H., Dugger, K. M., Dunn, M. J., Emmerson, L., Erikstad, K. E., Fort, J., Fraser, W., Genovart, M., Gilg, O., González-Solís, J., Granadeiro, J. P., Grémillet, D., Hansen, J., Hanssen, S. A., Harris, M., Hedd, A., Hinke, J., Igual, J. M., Jahncke, J., Jones, I., Kappes, P. J., Lang, J., Langset, M., Lescroël, A., Lorentsen, S.-H., Lyver, P. O. B., Mallory, M., Moe, B., Montevecchi, W. A., Monticelli, D., Mostello, C., Newell, M., Nicholson, L., Nisbet, I., Olsson, O., Oro, D., Pattison, V., Poisbleau, M., Pyk, T., Quintana, F., Ramos, J. A., Ramos, R., Reiertsen, T. K., Rodríguez, C., Ryan, P., Sanz-Aguilar, A., Schmidt, N. M., Shannon, P., Sittler, B., Southwell, C., Surman, C., Svagelj, W. S., Trivelpiece, W., Warzybok, P., Watanuki, Y., Weimerskirch, H., Wilson, P. R., Wood, A. G., Phillimore, A. B. & Lewis, S. **Global phenological insensitivity to shifting ocean temperatures among seabirds**. *Nature Climate Change* 8, 313 (2018).

All supplementary tables and meta-data are published online at:
https://doi.org/10.1038/s41558-018-0115-z

2.1 Abstract

Reproductive timing in many taxa plays a key role in determining breeding productivity (Visser & Both, 2005) and is often sensitive to climatic conditions (Thackeray *et al.*, 2016). Current climate change may alter timing of breeding at different rates across trophic levels, potentially resulting in temporal mismatch between the resource requirements of predators, and their prey (Thackeray *et al.*, 2010). This is of particular concern for higher trophic-level organisms, whose longer generation times confer a lower rate of evolutionary rescue than primary producers or consumers (Visser *et al.*, 2004). However, the disconnection between studies of ecological change in marine systems makes it difficult to detect general patterns of timing of reproduction (Richardson & Poloczanska, 2008). Here, we use a comprehensive meta-analysis of

209 phenological time series from 145 breeding populations to show that on average, seabird populations worldwide have not adjusted their breeding seasons over time (-0.020 days yr^{-1}) or in response to sea surface temperature (SST) (-0.272 days $°C^{-1}$) between 1952 and 2015. However, marked between-year variation in timing observed in resident species and some Pelecaniformes and Suliformes (cormorants, gannets and boobies), may imply that timing, in some cases, is affected by unmeasured environmental conditions. This limited temperature-mediated plasticity of reproductive timing in seabirds potentially makes these top predators highly vulnerable to future mismatch with lower trophic-level resources (Thackeray *et al.*, 2016).

2.2 Introduction and Results

The effects of rising global temperatures are having a profound impact on terrestrial and aquatic biota, including species abundance, distributions, behaviours, and interactions (Walther *et al.*, 2002). Changes in phenology - the timing of seasonally recurring life-history events - are one of the most apparent responses to rising global temperatures; at higher latitudes many spring and early summer events are advancing over time across a suite of terrestrial, freshwater and marine ecosystems (Thackeray *et al.*, 2010, 2016). As timing of breeding affects the abiotic conditions and biotic interactions to which parents and their offspring are exposed (Miller-Rushing *et al.*, 2010), breeding phenology is expected to play a key role in mediating the relationship between environmental temperature and fitness (Visser & Both, 2005).

Globally, many species at higher trophic levels have poor conservation status (Purvis *et al.*, 2000). Current evidence indicates that the phenology of species occupying higher trophic levels is less responsive to environmental change than that of primary producers and consumers (Visser *et al.*, 2004; Thackeray *et al.*, 2010,

2016), making them particularly susceptible to trophic mismatch and the associated negative demographic consequences (Thackeray *et al.*, 2010; Poloczanska *et al.*, 2013). However, previous studies which have combined estimates of phenological sensitivity (i.e. phenological change over time or in response to temperature) of multiple high trophic-level species to global change (Thackeray *et al.*, 2010, 2016; Sydeman *et al.*, 2012, 2015; Chambers *et al.*, 2013; Poloczanska *et al.*, 2013, 2016) have typically included few species or focused primarily on mean responses within taxa, trophic levels, or regions. Moreover, most earlier multi-species analyses have ignored sampling error in estimates of phenological sensitivity (Parmesan & Yohe, 2003; Sydeman *et al.*, 2012, 2015; Poloczanska *et al.*, 2013, 2016) (but see (Thackeray *et al.*, 2016) for an alternative approach) or sources of statistical non-independence, such as phylogeny (but see (Dunn & Møller, 2014)). As such, it is not clear whether the variation in rates of phenological sensitivity reported in the literature is simply the result of the sampling error variance that is characteristic of regression using short time series (Hadfield & Nakagawa, 2010; Youngflesh *et al.*, 2017), or represents true variation. If true variation in phenological sensitivity exists, this may arise where the strength of plasticity covaries with attributes of particular species (e.g. body size, feeding ecology, migration strategy), biogeography (e.g. upwelling, latitude, hemisphere or ocean basin), or an interaction between two or more of these effects. Testing the influence of these variables on variation in phenological sensitivity at a global scale across multiple populations will help to ascertain general patterns and highlight those taxa and regions most likely to be vulnerable to climate change.

Seabirds are one of the best-studied groups of higher trophic level organisms, and are considered here to include species from the orders Sphenisciformes, Procellariiformes, Suliformes, Pelecaniformes and Charadriiformes. Found throughout the world's oceans, they range in size from ~20g to ~30kg, and generally

exhibit long generation times and slow, inflexible life histories. They are more threatened than any other comparable avian group, with the conservation status of many species rapidly deteriorating (Croxall *et al.*, 2012). Seabirds exhibit considerable interspecific variation in feeding strategies, with breeding season foraging ranges varying from <10 to >1000 km and foraging depths from <1 m to 100s of metres deep. Outside the breeding season, some species remain close to their colony while others undertake the longest migrations known in the animal kingdom (Schreiber & Burger, 2002).

Studies of seabird breeding phenology have reported a variety of different trends over time (Chambers *et al.*, 2014). Among the local environmental drivers of phenology that have been identified, sea surface temperature (SST) is widely reported to correlate with the distribution, abundance and phenology of both local and migratory prey populations (Cheung *et al.*, 2013), of which the effects on higher trophic level organisms can be compared at global scales. Therefore, changes in temperature driven by climate change could be critical, generating a mismatch with prey availability (see further discussion below) (Ainley & Boekelheide, 1990). Directional SST changes and fluctuations have been recorded in the waters surrounding many seabird breeding sites (Figure 2.1a, b, Supplementary Figure 1 (in published version)), with both metrics of change varying geographically. Large-scale climatic variables, such as the North Atlantic Oscillation and the Southern Oscillation Index may also explain annual variation in reproductive phenology (reviewed in (Sydeman *et al.*, 2012)). However, using large-scale proxies instead of data on specific climate drivers (e.g. SST) may lead to spurious and simplistic assumptions of climate-ecology dynamics (Mesquita *et al.*, 2015). Furthermore, proxies at this scale are not amenable to global analyses, since regional proxies are not equivalent or comparable in a single analysis (Mesquita *et al.*, 2015). Thus, variation in the

sensitivity in timing of breeding across species and regions remains unclear (but see (Youngflesh *et al.*, 2017)). Due to their trophic position, global distribution and the numerous long-term studies available, seabirds constitute a tractable and powerful group for a global meta-analysis of breeding phenology. Such an analysis allows us to not only make general inferences about the degree to which breeding phenology has changed both over time and in relation to SST, but also about the life history traits underpinning variation in phenological responsiveness (Table 2.1). Finally, it allows us to examine predictors of between-year phenological variation, with high variance potentially indicative of phenological sensitivity to one or more unspecified environmental drivers. We applied a phylogenetic mixed model meta-analysis to a global dataset comprising 209 phenological time series of breeding dates obtained from 145 seabird populations (Figure 2.1c. Median number of years/time-series = 18; min = 5; max = 48. Median sample size/year /time-series = 72; min = 6; max = 936), covering 61 species from five main orders. These taxonomic groups exhibit a wide variety of life-history, migration and foraging strategies, and are distributed from equator to poles across all principal oceanographic regimes. Meta-analyses provide a robust approach for identifying average effect sizes across studies, and for identifying predictors of variation around the average (Nakagawa & Santos, 2012). Here, we (i) characterised latitudinal trends in the mean and between-year variance of seabird breeding phenology (laying and hatching dates), (ii) estimated the mean sensitivity of breeding phenology over time and in relation to SST in the waters around the sampled colonies, and (iii) identified predictors (body size, biogeography, phylogeny) of inter- and intra-specific variation around the mean response (mean, variance and both sensitivity measures) of each species/population (for specific predictions see Table 2.1 & Methods).

With increasing latitude, we found that breeding occurred later in the calendar year and that between-year variance in phenology decreased (Supplementary Table 1, Figure 2.2a, b), which concurs with earlier results obtained from regional studies (Wanless *et al.*, 2008; Burr *et al.*, 2016). The low variance at high latitudes may arise due to the shortened period of favourable conditions and the strong seasonal cue that photoperiod provides, whereas the much greater variance at lower latitudes may relate to the reduction of seasonality and the relatively weaker cue from day length (Bradshaw & Holzapfel, 2010).

Overall, the between-year variance in lay date among populations in our dataset ranged from < 1 in the black-browed albatross (*Thalassarche melanophris*) at New Island, Falklands, consistent with 95% of annual means occurring within a three-day period, to 1573 in the blue-footed booby (*Sula nebouxii*) at Isla Isabel, Mexico, consistent with 95% of annual means occurring within a five-month period. Examination of life history traits potentially explaining this variation (Supplementary meta-data) indicated that resident species were more variable than migrants (Supplementary Table 2, Figure 2.3b). This result is in accordance with results for terrestrial birds(Moussus *et al.*, 2011) and may arise if the laying dates of resident species are more sensitive to local foraging conditions as a cue to initiate breeding in anticipation of the timing of future resources. Controlling for biogeographic trends, we find that between-year variance in laying date was highly phylogenetically conserved (H^2 = 0.84, 95% Credible Interval [CI]: 0.508 – 1, n = 208, Supplementary Table 2). From inspection of the best linear unbiased predictors (BLUPs) for the phylogenetic effects, the most threatened order(Croxall *et al.*, 2012), Procellariiformes, particularly giant petrels and fulmars (Procellariidae), and albatrosses (Diomedeidae), stood out as least variable in timing of breeding. This response is consistent with a strong reliance on photoperiod as a cue (Gwinner, 1996). In contrast, we find that

Pelecaniformes and Suliformes (cormorants, gannets and boobies) vary substantially among years in timing of breeding, suggesting that these species may adjust egg laying in relation to some aspect of the local environment (weather, oceanographic conditions or food availability) in the lead-up to the breeding season (Dawson, 2008).

Table 2.1. Predictions of the effect of life history and environmental variables on phenology from the four key models. Predictions in bold indicate they are supported by the model. Proximate ([p]) versus ultimate ([u]) cues ([q]) and constraints ([c]) are specified in superscript.

Prediction		Reason
Mean Phenology		
Phenology will be later:	**at high latitudes**	**due to stronger photoperiodic cues at high latitudes[u,c]** (Wanless *et al.*, 2008; Burr *et al.*, 2016)**.**
Between-year variance		
Higher between-year variance will be observed in:	smaller birds	as they are more sensitive to environmental change[u,c] (Stevenson & Bryant, 2000)
	residents & short-distance migrants	**because they may be more sensitive to conditions at the breeding site[u,q]** (Moussus *et al.*, 2011)**.**
	surface feeders	which are more constrained in the water column, meaning that they can only exploit prey near the water surface[p,c] (Furness & Tasker, 2000).
	populations in upwelling zones	**due to high between-year variation in productivity in these areas[p,q]** (Chavez & Messié, 2009; Reed *et al.*, 2009)**.**
Temporal trends		
A steeper negative slope will be observed:	**in birds with smaller body size**	**to avoid incurring fitness costs of thermoregulation when breeding at higher temperatures[p,q]** (Stevenson & Bryant, 2000)**.**
	in birds which feed at the surface	as they may be more sensitive to the timing at which lower trophic level prey are available[p,c] (Furness & Tasker, 2000).
	at high latitudes	because polar systems are experiencing warming faster than other areas[u,c] (Stocker *et al.*, 2013).
Sea Surface Temperature trends		
A steeper negative slope will be observed:	in birds with smaller body size	to avoid incurring fitness costs of thermoregulation when breeding at higher temperatures[p,q] (Stevenson & Bryant, 2000).
	in residents & short-distance migrants	as they are likely to respond to conditions at the breeding site more readily than species which overwinter in different basins[u,q] (Moussus *et al.*, 2011).
	in birds which feed at the surface	as they are predicted to be more sensitive to the timing at which lower trophic level prey are available[p,c] (Furness & Tasker, 2000).
	at high latitudes	as polar systems are experiencing warming faster than other areas[u,c] (Stocker *et al.*, 2013).

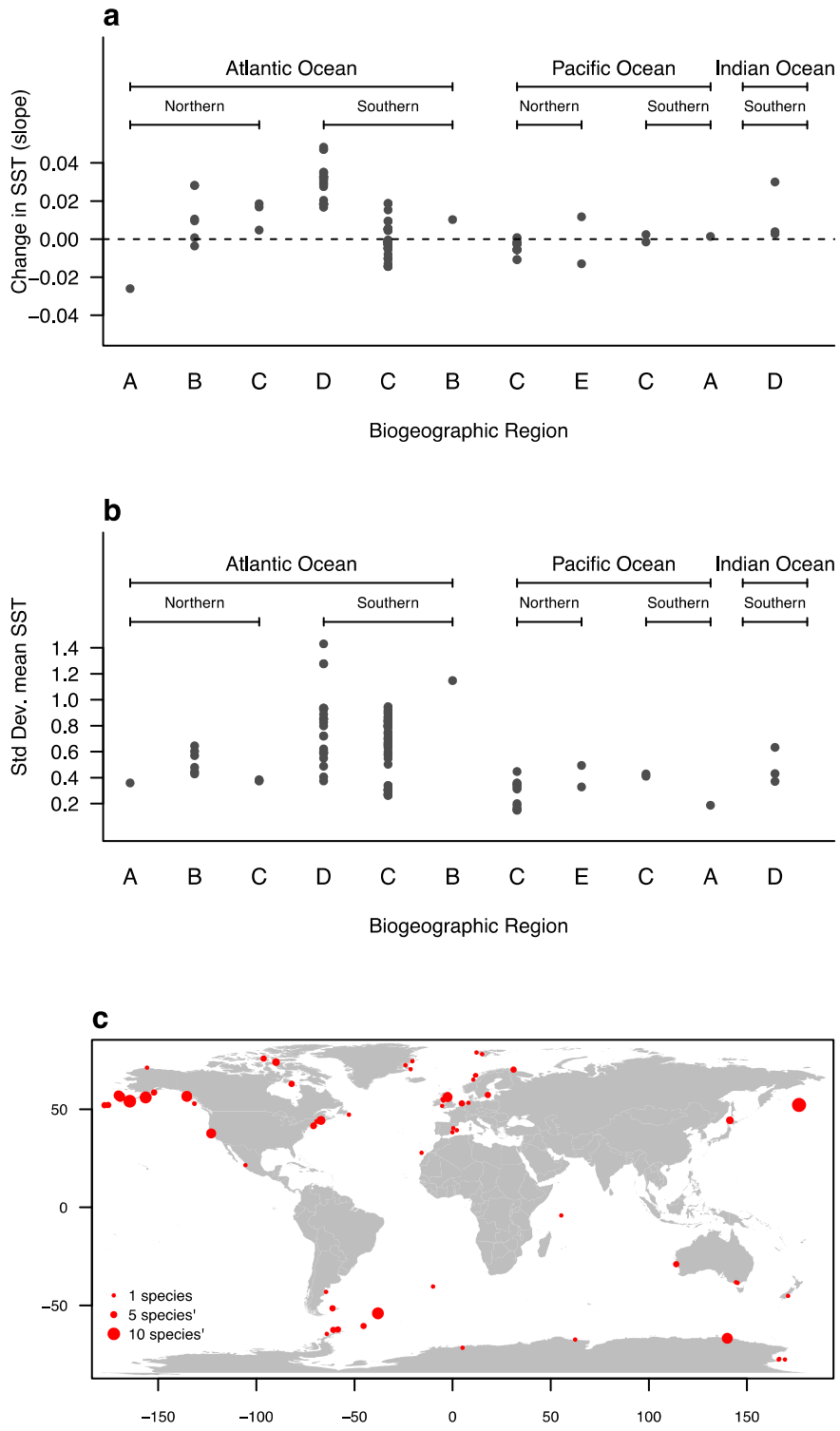


Figure 2.1 SST trends and map of study sites included in the analyses. a) Across year temporal changes in mean Sea Surface Temperature (SST) in the three months prior to breeding across all biogeographic regions represented by slopes between 1982 (when SST time series' began) and 2015 for each site. Each point represents a slope, with positive slopes indicating warming and negative slopes indicating cooling. b) Standard deviation from

the mean SST at each site during the same study period. A = Polar, B = Subpolar, C = Temperate, D = Subtropical, E = Tropical. c) The full dataset comprises 209 time series from 61 seabird species and across 64 locations, collected between 1952 and 2015. The data include slopes for 32 genera, 9 families, and 5 orders (Sphenisciformes (6), Procellariiformes (15), Suliformes (3), Pelecaniformes (5), Charadriiformes (32)) and spans all seven continents. The underrepresentation of tropical time series is due to a combination of a paucity of long-term data for these regions and the asynchronous nature of breeding in many tropical species, which diminishes the informativeness of measuring annual phenological central tendency.

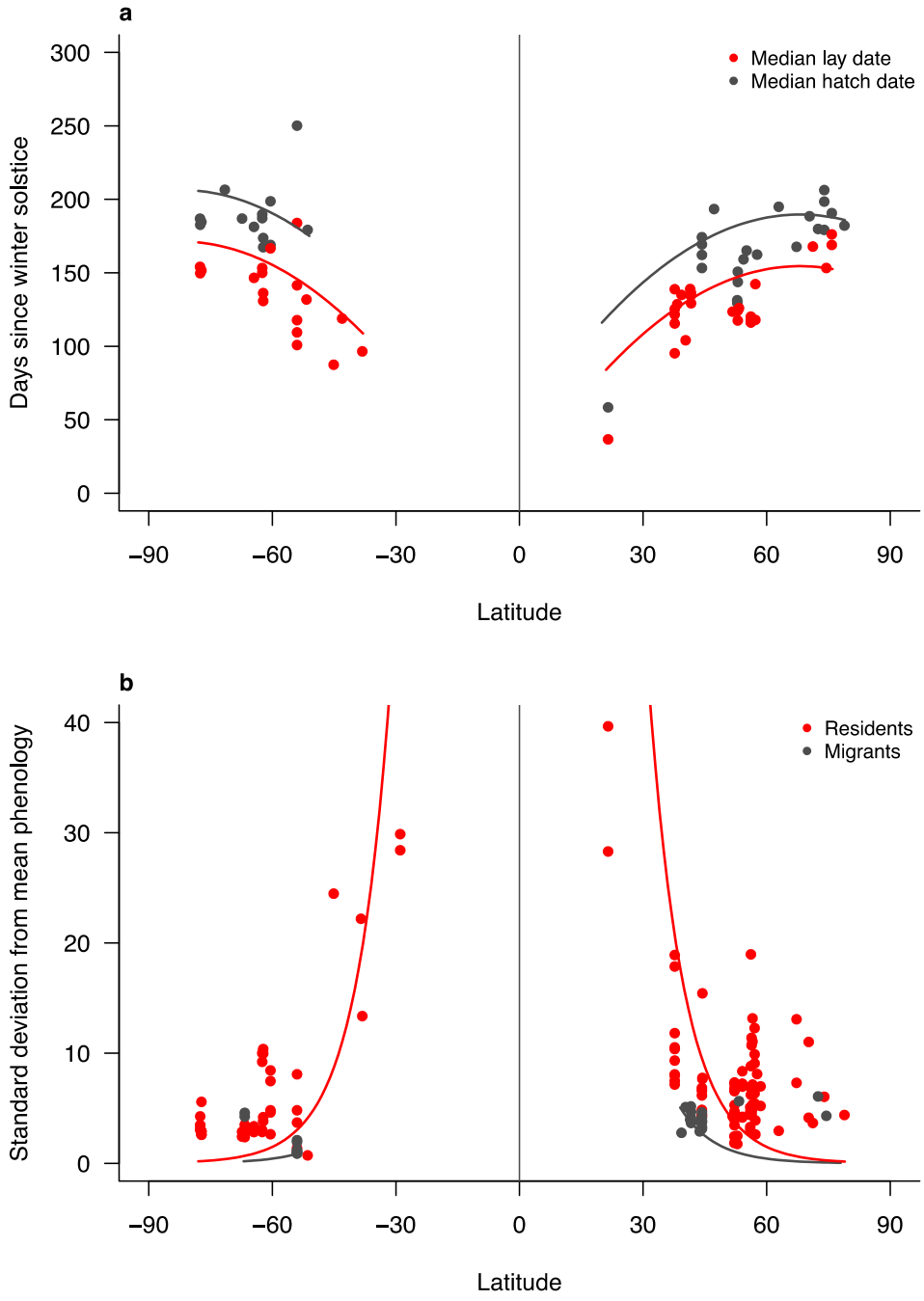


Figure 2.2. Mean and between-year variance in phenology separated by hemisphere. a) represents the differences in latitudinal gradient between Northern and Southern Hemispheres, where each data point (grey or red) represents the median timing of breeding of a population. Lines (grey = lay date, red = hatch date) represent the delay in phenology approaching the poles in days lat $^{-1}$, and were estimated using values from Supplementary Table 1. **b)** represents the between-year standard deviation in mean timing for residents (represented by red dots) and migrants (grey dots). Lines are plotted from the ecological model and represent the median lay date in the mean year of study of an average surface feeding resident bird, weighing 800g, in a region where there is no major upwelling system. The non-linearity in the plot is due to back calculation from the log scale.

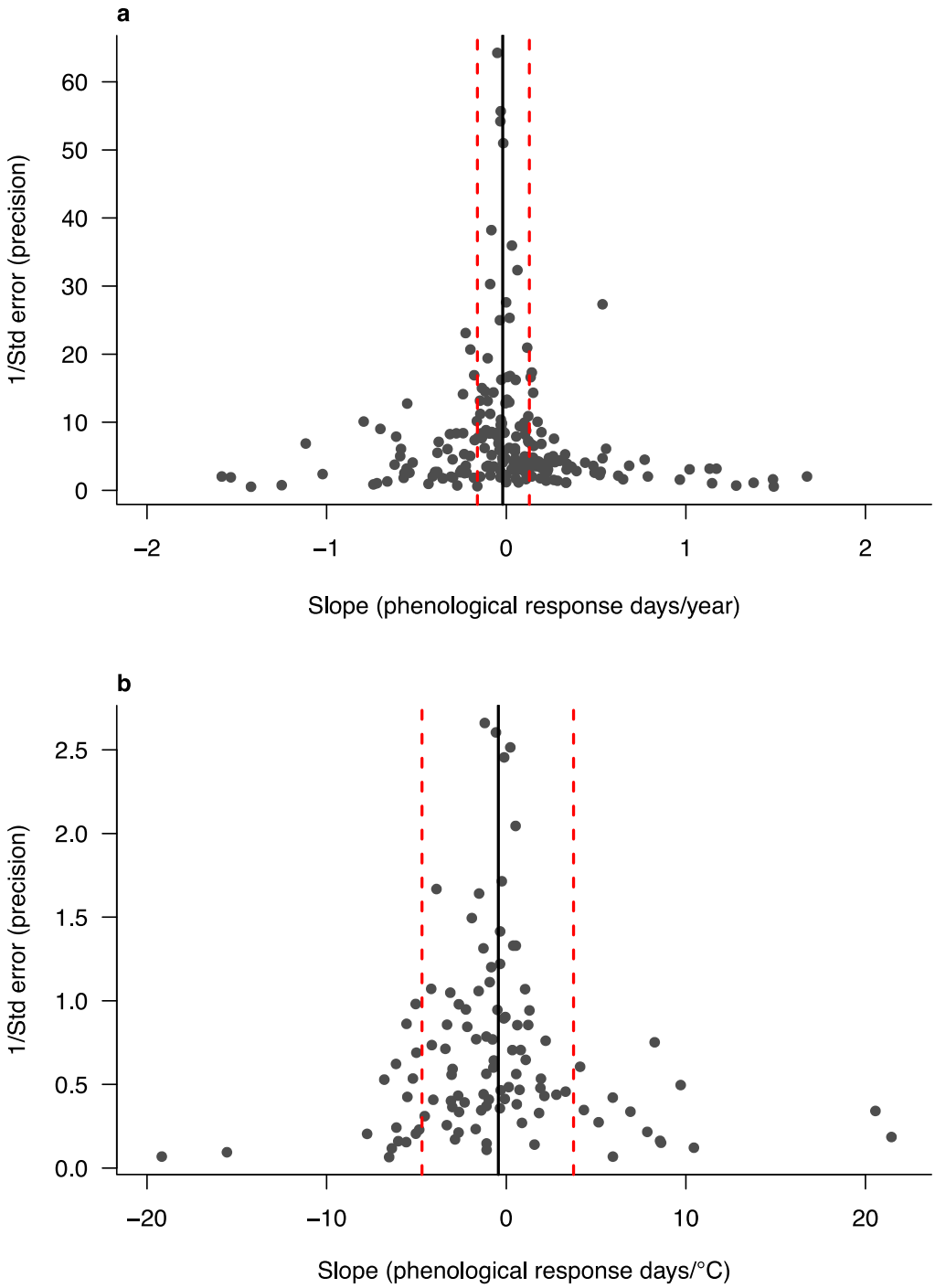


Figure 2.3. Funnel plots of phenological trends in relation to year and sea surface temperature. a) represents year and **b)** represents sea surface temperature. Each point represents a slope estimate from the meta-analysis, with negative slopes indicating an advance and positive slopes indicating a delay, in phenological trends. Positioning of each point on the y-axis indicates the precision (1/S.E) of the estimate. Thus, points with higher precision are expected to converge on the true average response. Lines represent the posterior for the average response or intercept (black) and its 95% credible intervals (dashed red) from the basic model (Tables S3a, S5a).

On average, seabirds showed no tendency to advance or delay breeding phenology over time (-0.020 days yr^{-1}, 95% CI: -0.160 – 0.129, *n* = 209, Figure 2.3a). This is in agreement with previous studies of this species group (Poloczanska *et al.*, 2013; Chambers *et al.*, 2014), but the overall slope was much less steep than those from similar analyses of UK birds (Thackeray *et al.*, 2010) (mean = -0.19 days yr^{-1}), terrestrial and marine vertebrates (Thackeray *et al.*, 2010) (terrestrial mean = ~-0.25 days yr^{-1}, marine mean = ~-0.35 days yr^{-1}) or global estimates of marine species in general (Poloczanska *et al.*, 2013) (mean = ~-0.4 days yr^{-1}). We found limited evidence for true variation around the mean response (Supplementary Table 3), with 83% of the variation in raw slope estimates of phenology over time attributable to sampling error arising from linear regressions based on small datasets (Supplementary Table 4). Of the remaining true variation, we found that the mean slope estimates did not differ significantly among oceans (Supplementary Table 3). This result runs counter to previous studies of seabird breeding phenology, which have reported variation in long-term trends among biogeographic realms (Chambers *et al.*, 2014; Poloczanska *et al.*, 2016). However, we found some evidence that temporal response may vary among species at shared breeding sites (Supplementary Table 3), although sampling covariance between the different phenological measures is likely to inflate this variance estimate. Among-population variation makes it difficult to predict which species and sites will be most phenologically responsive to changing environments, as it implies that the degree of environmental sensitivity in seabird breeding may be determined by a combination of intrinsic and extrinsic factors (Daunt *et al.*, 2014). Of the environmental or life history variables we considered, body mass was the only significant positive predictor of the temporal trend (Supplementary Table 3), with larger-bodied species responding at a slower rate over time than smaller species, in accordance with our predictions (Table 2.1).

Globally we found no evidence that seabirds as a group have shifted their laying date in relation to SST in waters around the breeding site in the three months preceding egg laying (mean = -0.272 days $^{\circ}C^{-1}$, 95% CI: -4.896 – 4.482, *n* = 108, Figure 2.3b, Supplementary Table 5). The average response is much shallower than the average response of lay date to air temperature reported for 27 UK terrestrial birds (mean = -3.8 days $^{o}C^{-1}$ (air temperature)) (McLean *et al.*, 2016). In broad agreement with the temporal analysis we found no evidence that true variation in the slope of the covariation with SST is predicted by phylogeny, species, biogeographic region, or life-history traits. We did, however, find significant variation in slopes among sites, and the lowest BLUP was -2.96 days $^{\circ}C^{-1}$ (95% CI: -6.00 – 0.13) at Skomer Island, Wales, where SST in the focal time period has increased significantly by 0.6°C decade^{-1} since 1982 (Supplementary meta-data 1). In contrast, the most positive BLUP was 7.32 days $^{\circ}C^{-1}$ (95% CI: 4.96 – 9.73) at Southeast Farallon Island, California, which is located in a highly variable upwelling zone, where inter-annual variance in SST is higher than average (Figure 2.1b, Supplementary meta-data), a condition that might select for plasticity. So, although on average, seabirds appear to be unresponsive to SST, we cannot rule out the possibility some populations are temperature-sensitive in either direction.

That we could detect no trend in seabird phenology over time or in relation to SST (Supplementary meta-data), suggests that if lower trophic levels are shifting in parallel with changing SST, seabirds, in general, may be at risk from increasing levels of trophic mismatch (Durant *et al.*, 2007). To date, there are very few studies that have reported the slope of the phenology of poikilothermic seabird prey and lower trophic levels in relation to SST (but see (Ainley & Boekelheide, 1990)). Differing rates of phenological response between seabirds and their food resources (Poloczanska *et al.*, 2013) may leave them short of critical prey during the breeding season under

future climate regimes. However, there is limited and mixed evidence on the frequency of climate-induced mismatch (Ainley & Boekelheide, 1990; Youngflesh *et al.*, 2017), and whether it has an impact on breeding success (Burthe *et al.*, 2012) or population dynamics (Reed *et al.*, 2013b). Alternatively, any negative fitness consequences of trophic asynchrony may be ameliorated by the ability of some species to alter their behaviour, for example by switching prey or adjusting foraging effort (Ainley & Boekelheide, 1990; Howells *et al.*, 2017).

Our study represents the most statistically rigorous and spatially representative meta-analysis to date of the reproductive phenology of a group of upper trophic-level predators, seabirds. Contrary to previous assertions, we find that once sampling error has been taken into account, in most cases the phenology of seabirds shows no trend over time and appears to be largely insensitive to changing SST. While certain populations may be responding, most of the among-species variation in estimates of phenological sensitivity can be attributed to sampling error. Overall, this inflexibility in breeding phenology in relation to temperature may leave seabirds vulnerable to trophic mismatch arising from shifts in timing of their prey.

2.3 Methods

2.3.1 Data collection

To prevent an effect of publication bias and to ensure that positive, negative and neutral phenological trends were included, we used only raw time series (see PRISMA checklist). For each time series we used consistent methods to calculate slopes (i.e. rate of phenological change), between-year variance and crucially, standard error. Raw phenological data were compiled from a variety of sources between October 2015 and October 2016. We contacted 120+ known seabird researchers and owners of time series to request annual data on seabird breeding

phenology and life history. Furthermore, requests were made via Twitter and at the World Seabird Conference in Cape Town (October 2015); the Pacific Seabird Group Annual Meeting in Oahu (February 2016); The Seabird Group conference in Edinburgh (September 2016); and the International Albatross and Petrel Conference in Barcelona (September 2016).

2.3.2 Data

Annual data on breeding phenology during the period 1952 and 2015 were the median or mean date of laying or the median, mean or first date of hatching of the study population, in units of ordinal days. Population was defined as an individual species at a breeding site. We only considered populations that breed seasonally during spring and summer (austral and boreal) months, as measures of phenological central tendency are not informative for species which breed asynchronously or subannually (i.e. many tropical species (Schreiber & Burger, 2002)). Time series' were required to be a minimum of five years for the temporal analysis and ten years for the analysis of SST, although the years did not need to be consecutive. Details of criteria used to choose suitability of time series' are given in Supplementary Table 9, and the field methods used to collect each time series are outlined in the Supplementary Methods.

Monthly means of NOAA Optimum Interpolation (OI) Sea Surface Temperature (SST) V2 for the period 1982 – 2015 were *obtained from the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA,* a resource which provides interpolated *in situ* and satellite SST data on a one-degree grid(Reynolds *et al.*, 2002).

For each time series we characterised the biogeography at the colony it was located. We collated information on the location (latitude and longitude) and hemisphere of each population, and for our primary fixed effects model we assigned each location to one of the three main oceans: Atlantic, Pacific or Indian. Global climate zones (Equatorial, Tropical, Subtropical, Temperate, Subpolar or Polar) were

identified using the classification from Trujillo & Thurman (2014). These zones correspond to latitudinal bands of similar sea surface temperature and are categorized by levels of precipitation, wind and water temperature (Trujillo & Thurman, 2014). We combined hemisphere, ocean and global climate zone to identify 15 Biogeographic Regions (e.g. North Atlantic Temperate; South Pacific Subpolar etc.). Finally, we used the Longhurst Biogeographical Provinces to determine whether each location was situated within an Eastern Boundary (upwelling) zone (Longhurst, 2006). These are areas of high productivity within the marine environment, and are also highly variable across seasons, years and decades (Stenseth *et al.*, 2003; Chavez & Messié, 2009).

We collated data on several aspects of the ecology and life history of each species that may affect the phenological slope (with year or temperature), mean or between-year variance. These data were provided by authors and supplemented using online resources: www.audubon.org, www.birdlife.org, nzbirdsonline.org.nz, www.bird-research.jp and www.npolar.no (Supplementary meta-data). Feeding strategy was categorised either as surface feeder (feeding <1 metre below the surface), diver (feeding >1 metre below the surface), or kleptoparasite/predator (part-time marine foragers). Species which seek out prey by diving under water may be able to exploit a wider range of prey than those constrained to feeding on the surface (<1 metre depth), thus reducing the necessity to adjust breeding phenology to buffer mismatch (Sabarros *et al.*, 2012; Cabot & Nisbet, 2013; Passuni *et al.*, 2015). We also compiled data on average body mass of every species (Supplementary meta-data), as small-bodied seabird species are predicted to be more sensitive to temperature change due to the higher cost of thermoregulation (Reiss, 1989; Stevenson & Bryant, 2000). Furthermore, body mass can be used as a proxy for trophic level, which is difficult to classify explicitly in seabirds (Romero-Romero *et al.*,

2016). We used log body mass in analyses. The migration strategy of individuals from each population was assigned based on the behaviour of the majority (>80%) of individuals. Long distance trans-equatorial migrants, and species which spend the winter outside the sector in which they breed were categorised together as “migrants”, and those which remain in the same ocean sector throughout the year were classified as “residents”. Sectors were defined as North Atlantic, Mediterranean, South Atlantic, Southern Ocean-Atlantic sector, North Pacific, South Pacific, Southern Ocean-Pacific sector, Indian, Southern Ocean-Indian sector.

We took into account phylogenetic relationships among species using 100 samples of the pseudo-posterior species tree (Jetz *et al.*, 2012) using the Hackett *et al.* (2008) backbone (Hackett *et al.*, 2008).

2.3.3 Statistics

We used the *MCMCglmm* package(Hadfield, 2010) in R (v 3.2.2; R Core Team 2015), to fit Bayesian generalised linear mixed-effects models (GLMMs). We adopted a random effects meta-analytic (REMA) approach, estimating both fixed and random effects, while taking sampling error characteristic or regression using short time series into account (Hadfield & Nakagawa, 2010; Nakagawa & Santos, 2012).

We included cross-classified random effects to account for and estimate sources of variance, though not every random variable was included in each model (see Tables S1-S5). The model was of the form

$$y_i = \mu + \beta x_i + \alpha_{f[i]} + s_{f[i]} + b_{g[i]} + l_{h[i]} + p_{j[i]} + e_i + m_i \qquad \text{eq. 1.}$$

where y is the phenological response variable of each time series i, μ represents the global mean response (intercept), and βx_i the fixed effects. For each response variable we also included a null model with the intercept as the sole fixed effect, as this allowed us to infer which random terms captured most of the variance.

$\alpha_{f[i]}$ is the effect of phylogenetic non-independence due to shared evolutionary history (Hadfield & Nakagawa, 2010) for the *f*th species. $s_{f[i]}$ is the non-phylogenetic species-specific effect for the *f*th species. Spatial variation was accounted for via two terms, *g*th biogeographic region ($b_{g[i]}$) (see Supplementary meta-data) and *h*th site ($l_{h[i]}$). In certain analyses we included multiple measures/traits for a time series and in these cases we could fit the interaction between site and species (population) ($p_{j[i]}$), which provided us with an estimate of intraspecific geographic variation that is unique to each (*j*th) population. In these cases the residual term (e_i) captures variation within a site and species (population), and we allowed this variance to be heterogeneous across different phenophases (i.e. median lay date, mean lay date, first hatch date, median hatch date, mean hatch date). In other analyses only a single measure/trait was included and in such instances $p_{j[i]}$ was not estimable. In this case the residual term captured variance both due to intraspecific geographic variation that is unique to each species and differences among phenological measures/traits. Our response variables were themselves estimates that have error associated with them and we incorporated sampling error variances as m_i, which means that the analyses were weighted. For the sampling error term, the among-observation variance was set to 1, and for all other random terms the variance was estimated. The specification of these models assumed that random effects for different measures were perfectly correlated. To test whether this impacted on our estimation of phylogenetic signal we then relaxed this assumption and estimated the covariance between random effects for measures of laying and hatching phenology (Variance Structure of Models section, below).

We calculated phylogenetic signal (Housworth *et al.*, 2004; Hadfield & Nakagawa, 2010) in our response variables (H^2), i.e. the tendency of closely related

species to resemble each other more than distantly related species, from σ_a^2 (the phylogenetic variance), and σ_s^2 (the species variance)

$$H^2 = \frac{\sigma_a^2}{(\sigma_a^2 + \sigma_s^2)} \quad \text{eq. 2.}$$

We considered the following four response variables and clearly identify where analyses are *post hoc* rather than *a priori*:

(1) Multi-year mean phenology: we estimated the mean phenology (e.g. average laying date overall) across all years for each time series. Measurement variance in the mean was quantified as the squared standard error. To examine latitudinal trends in mean date we included both absolute latitude and its quadratic term (to test both linear and non-linear effects); hemisphere; and the interaction between latitude and hemisphere as fixed effects. Additional fixed effects were trait (laying and hatching date) and phenological measurement (mean, median, first date). See Table 1 for predictions.

Post hoc tests: mean phenology is delayed as latitude increases in both hemispheres, with a significant quadratic term, such that the slope appears to reach an asymptote toward the poles (Figure 2.2, Supplementary Table 1). However, seabirds at low latitudes are underrepresented in this study. When we removed three low latitude data points, there was no support for the quadratic relationship (Supplementary Table 1) but the positive linear relationship between latitude and breeding phenology remained (posterior mean = 0.81 days.lat^{-1}, 95% CI: 0.33 – 1.29, *n* = 206, Supplementary Table 1). The intercepts of each measure of phenology (i.e. mean laying date, first hatching date) differed significantly, although a test including the interaction between latitude and phenological measure revealed no difference in their latitudinal slopes (Supplementary Table 1).

(2) Between-year variance in phenology: the response variable (eq. 3) was based on the natural log of the between-year standard deviation (*s*) of each population ($\ln \sigma$), taking into account the number of years (*n*). The sampling variance of this measure was quantified as ($s^2_{\ln \sigma}$) as in eq. 4 (Nakagawa *et al.*, 2015):

$$\ln \hat{\sigma} = \ln s + \frac{1}{2(n-1)} \qquad \text{eq. 3.}$$

$$s^2_{\ln \hat{\sigma}} = \frac{1}{2(n-1)} \qquad \text{eq. 4.}$$

The model included phenological trait and measure, latitude and its quadratic term, hemisphere, presence or absence of upwelling and, to test for decadal patterns, the mean year of each time series as fixed effects. We included body mass, foraging and migration strategies in the same model to investigate the effects of life history traits on between-year variance. See Table 2.1 for predictions.

(3) Temporal trend in phenology: we estimated the linear slope (and standard error) of phenological change over time for each measure (median, mean, first date) and trait (laying or hatching date) of a population using Generalised Least Squares (GLS) in nlme (Pinheiro & Bates, 2000), fitting an autoregressive model of order 1, AR(1) (Box *et al.*, 1990), to take into account temporal autocorrelation in each individual time series. We used these slope estimates in a meta-analysis, and included the squared standard error of the slope to weight the analysis. We included three types of fixed effects: methodology (trait, measure, mean year of time series), life history and ecology (body mass and foraging strategies), and biogeography (ocean basin, hemisphere, latitude). See Table 2.1 for predictions. We did not make predictions about which ocean basins or hemisphere might show the steepest slopes, but allowed the response to differ among ocean basis and hemispheres in our model.

Post hoc test: our primary ecological fixed effects model categorised locations into one of the three main ocean basins (Atlantic, Indian, Pacific), and included the interaction between latitude and hemisphere as an additional parameter. This approach considered the life histories of wide-ranging polar species which may have large foraging ranges. Yet many species forage near to the colony, or may have evolved alongside the unique oceanographic features of polar systems (Brierley & Kingsford, 2009). To consider these species we re-categorised ocean basins into five discrete water bodies (Arctic, Atlantic, Indian, Pacific, Southern) and ran our ecological model again, replacing the three ocean variable with five oceans, and removing the interaction between latitude and hemisphere.

(4) Phenological response to SST: for each time series we averaged monthly temperature data from the local grid cell for the pre-breeding period (three, two and one month prior to laying, including the month in which laying began) each year. In some cases sea ice cover meant that an average temperature was not estimable and affected time series' were excluded from this analysis. We restricted this analysis to laying dates only, representing each population with a single time series in declining order of preference of measurements: median, mean and first date. In populations for which we only had data on timing of hatching, we back-calculated lay dates using information on the duration of incubation period and average number of eggs. These data were provided by authors and supplemented using online resources: www.audubon.org, www.birdlife.org, nzbirdsonline.org.nz, www.bird-research.jp and www.npolar.no (Supplementary meta-data). Where incubation period was reported as a range, we calculated the central value; this method was used for 70 time series (Supplementary meta-data).

For each colony we calculated the reaction norms and associated standard errors of phenological response to SST (days $^{o}C^{-1}$) using the GLS methods as

described for the temporal trends, but retaining year as an additional predictor, in order to de-trend the data and allow us to consider the effects of SST independently of time (Supplementary meta-data). We compared among pre-breeding on the basis of AIC and found very little difference, as expected given the overlap between time periods and month-to-month temporal autocorrelation is SST. Across time series the three-month period had the lowest mean AIC (2 month mean ΔAIC = 0.02, 1 month mean ΔAIC = 0.50) and for consistency we used this time period in subsequent analyses.

We then passed the slopes of phenology regressed on three-month SST into a meta-analysis, with the squared standard error of the slope included for weighting. We tested similar predictions as in (3) above, predicting that timing of laying would be more sensitive to pre-breeding SST in species with smaller body mass, which feed on the surface, or that remain in the same ocean basin over winter. Measure, trait and mean year of study were also included as fixed effects.

All models were run for 30,000 iterations on each phylogenetic tree sample, discarding the first 10,000 as burn-in and sampling every 10^{th} iteration. We repeated this process over 100 phylogenetic trees and the pooled posterior distributions take into account both model and phylogenetic uncertainties (Pagel & Lutzoni, 2002). Parameter-expanded priors were used for all random effects except the residual, which followed an inverse Wishart distribution. Plots of the mean and variance of the posterior distribution were examined to assess autocorrelation in the posterior samples. Statistical significance of fixed effects was inferred where 95% credible intervals did not span zero.

2.3.4 Variance Structure of Models:

Our dataset contains five phenophases: median lay date (1), mean lay date (2), first hatch date (3), median hatch date (4) and mean hatch date (5). The core models (with the exception of temperature) run under the assumption that within the residual term (e_i) the variance would be heterogeneous, with each phenophase varying independently of the other four (eq. S1). We used the idh() variance structure function in the MCMCglmm package(Hadfield, 2010). This is consistent with phenophases being uncorrelated at the residual level (i.e. covariance = 0) but at the other random effects the correlation between phenophases is assumed to be 1.

$$\boldsymbol{V}_{e_i} = \begin{bmatrix} V_{1,1} & 0 & 0 & 0 & 0 \\ 0 & V_{2,2} & 0 & 0 & 0 \\ 0 & 0 & V_{3,3} & 0 & 0 \\ 0 & 0 & 0 & V_{4,4} & 0 \\ 0 & 0 & 0 & 0 & V_{5,5} \end{bmatrix} \quad \text{eq. S1}$$

These assumptions can be relaxed for each random effect and the covariance between phenophase can be estimated. We used the us() variance structure function (eq. S2), where *V* = variance, *C* = covariance and e_i = random effect.

$$\boldsymbol{V}_{e_i} = \begin{bmatrix} V_{1,1} & C_{1,2} & C_{1,3} & C_{1,4} & C_{1,5} \\ C_{1,2} & V_{2,2} & C_{2,3} & C_{2,4} & C_{2,5} \\ C_{1,3} & C_{2,3} & V_{3,3} & C_{3,4} & C_{3,5} \\ C_{1,4} & C_{2,4} & C_{3,4} & V_{4,4} & C_{4,5} \\ C_{1,5} & C_{2,5} & C_{3,5} & C_{4,5} & V_{5,5} \end{bmatrix} \quad \text{eq. S2}$$

Allowing slopes of phenophases to covary for every random effect may result in a more informative estimate of phylogenetic signal (i.e. perhaps signal is observed at one stage of reproduction but not another), but requires a large amount of data at each level to confidently estimate multiple (co)variances. As our dataset was not large enough to run models with fully unstructured (co)variance, we only estimate the covariance between lay and hatch dates. We restructured the covariance matrix for each random effect (eq. S2) into a 2 x 2 grid (eq. S3).

$$\boldsymbol{V_{R.E}} = \begin{bmatrix} V_{lay,lay} & C_{lay,hatch} \\ C_{lay,hatch} & V_{hatch,hatch} \end{bmatrix} \quad \text{eq. S3}$$

Thus, three slopes (lay date, hatch date and the covariance between the two) were estimated for each random effect (phylogeny; species; biogeographic region; location and species:location). We ran the three key models (between year variance, temporal and SST) using this error structure to assess whether any of our key insights were sensitive to the assumption that lay and hatch dates are perfectly correlated.

When the assumption of perfect correlation between the two measures was relaxed, we found that phylogenetic signal remained significant for the variance and SST models (Supplementary Tables 6, 8). We also found some evidence for phylogenetic signal in the temporal model (Supplementary Table 7). These results are in agreement with the key findings of our core models.

3. Variation and correlation in the timing of breeding of North Atlantic seabirds across multiple spatial scales

Katharine Keogan, Sarah Wanless, Francis Daunt, Richard A. Phillips, David Alvarez, Tycho Anker-Nilssen, Robert T. Barrett, Claus Bech, Peter H. Becker, Per-Arvid Berglund, Sandra Bouwhuis, Zofia M. Burr, Olivier Chastel, Signe Christensen-Dalsgaard, Sebastien Descamps, Tony Diamond, Kyle Elliott, Kjell-Einar Erikstad, Mike Harris, Jonas Hentati-Sundberg, Martin Heubeck, Magdalene Langset, Svein-Håkon Lorentsen, Mark Mallory, Mick Mellor, Will T. S. Miles, Børge Moe, Carolyn Mostello, Mark Newell, Ian Nisbet, Tone Kirstin Reiertsen, Jennifer Rock, Øystein Varpe, Sue Lewis, Albert B. Phillimore

3.1 Abstract

Environmental conditions can be highly variable in both space and time, causing fluctuations in the timing of resource availability. These conditions may affect breeding phenology of organisms, as individuals may use cues to ensure optimal timing, or be constrained by resource availability during the pre-breeding period. Contrary to trends observed in terrestrial habitats, breeding phenology of marine birds has shown no general trend over time or with rising spring temperatures. However, for many seabird populations there is substantial among-year variance in timing of breeding, suggesting that they may respond to one or more aspects of their environments. At present, it is unclear whether the observed variance is driven by a single environmental driver with a consistent effect across populations on a large spatial scale, conditions acting at a local scale, or species-specific responses to large-scale or local drivers. We combined 51 long-term datasets on breeding phenology spanning 50 years from nine seabird species across 29 North Atlantic sites, and assessed the extent to which reproductive phenology positively covaried among groupings of populations. We found no covariance in phenology between years on a large spatial scale. Instead, the timing of reproduction showed greater positive covariance when we considered multiple species breeding at the same site or groups of sites, showing

the importance of local terrestrial and/or marine conditions. In addition, breeding phenology in one species, black-legged kittiwake (*Rissa tridactyla*) covaried positively across 16 populations within the North Atlantic breeding range, but the degree of covariance decayed with distance. By applying variance partitioning in a new context, we show that in the absence of information on relevant environmental variables, the spatial and taxonomic scale at which phenology is determined can be established by identifying the extent to which timing is correlated between time-series.

3.2 Introduction

Predicting how organisms will respond to changing climate presents one of the greatest global challenges for ecologists. Some of the key responses that have been observed are changes in timing of seasonally recurring events (Parmesan & Yohe, 2003), which are often sensitive to environmental conditions such as temperature (Thackeray *et al.*, 2016). Timing of reproduction in relation to resource availability is an important trait affecting fitness (Visser & Both, 2005; Varpe, 2017). In order to respond to fluctuating environments, organisms should adjust timing of breeding to coincide with suitable conditions, and may respond to drivers which indicate the future arrival of a favourable environment (McNamara *et al.*, 2011), or be limited by environmental constraints (Perrins, 1970). Breeding phenology may therefore be adjusted in response to one or multiple cues and drivers, such as temperature (Chambers *et al.*, 2009), photoperiod (Dawson *et al.*, 2001), wintering conditions (Dobson *et al.*, 2017) or resource availability, potentially mediated by body condition in the pre-breeding season (Love *et al.*, 2010; Daunt *et al.*, 2014). The extent to which these different drivers combine or interact to elicit a phenological response may differ between species and regions, hampering our ability to make general predictions regarding population responses to environmental change (Thackeray, 2016; van de Pol *et al.*, 2016; Cohen *et al.*, 2018).

Determining the conditions that drive phenological responses and the spatiotemporal scales at which they act requires both long-term data on phenology and fine-scale data on candidate environmental variables (van de Pol *et al.*, 2016). However, identifying these conditions and scales is more straightforward in some species than others. For instance, those species occupying lower trophic levels may respond directly to temperature (Visser & Both, 2005) or increasing sunlight (Townsend *et al.*, 1994), and species that are rooted/sessile or have small year-round ranges should respond to conditions at a relatively small spatial scale (Frederiksen *et al.*, 2004; Lindestad *et al.*, 2018). In contrast, it may be more difficult to identify the correct drivers for wide-ranging or migratory species, or those occupying higher trophic levels. These species may respond to cues or conditions in the area where they breed (Frederiksen *et al.*, 2004), at their wintering areas (Szostek *et al.*, 2015; Dobson *et al.*, 2017), or both (Harrison *et al.*, 2011). Determining how the environment affects phenology also requires that the time period during which the cue or constraint acts is correctly identified, though rather few studies conduct the necessary comparisons among different time-windows (van de Pol *et al.*, 2016).

The challenges involved in identifying the drivers of phenology are well-illustrated by seabirds. Globally, seabirds on average show no phenological trend over time or with spring sea surface temperature (Keogan *et al.*, 2018), yet some analyses of particular species groups have found high levels of variance in timing of breeding between years (Burr *et al.*, 2016; Keogan *et al.*, 2018; Youngflesh *et al.*, 2018). This suggests that individuals are responding to their environment, but the specific environmental drivers (and when they occur, e.g. just before egg laying or as carry-over effects from several months before the breeding period), and associated population responses remain to be established. For the most part seabirds occupy higher trophic levels, and the breeding ranges of many species span large spatial

gradients in environmental conditions. They can forage long distances from the breeding site during the breeding season, and have some of the longest migrations known in the animal kingdom (Egevang *et al.*, 2010). Although many seabird species winter far from their colonies, many also spend time at the breeding site before egg laying commences, such that conditions at both breeding (Frederiksen *et al.*, 2004; Love *et al.*, 2010) and wintering grounds (Szostek *et al.*, 2015; Dobson *et al.*, 2017) may affect breeding phenology. Furthermore, seabird breeding sites may be home to multiple species, among which phenology may be positively correlated (i.e. relative differences in timing are similar) or even synchronous (i.e. timing is the same) (Lahoz-Monfort *et al.*, 2013; Samplonius & Both, 2017).

To overcome the difficulties of identifying the precise timing and nature of specific drivers and the spatiotemporal scales at which they act, several previous studies have regressed the breeding phenology of seabirds on aspects of the environment (Haest *et al.*, 2018a). Examples include average local or regional sea surface temperatures in the pre-breeding period (Reed *et al.*, 2009; Keogan *et al.*, 2018), abundance or availability of species two or more levels lower in the food chain (discussed in Grémillet *et al.*, 2008), or large-scale climate indices, such as the North Atlantic Oscillation (Frederiksen *et al.*, 2004), which attribute a single value to the diverse conditions encountered throughout a season (Stenseth *et al.*, 2002). However, using such indices may conceal the underlying mechanisms driving the interaction between organisms and their environment (Mesquita *et al.*, 2015; Haest *et al.*, 2018b), or misidentify the points in space or time at which environmental conditions are crucial (van de Pol *et al.*, 2016).

An alternative approach is to identify the scales at which populations respond to their surroundings. This can be achieved by examining the extent to which multiple populations within multiple species covary in their breeding phenology across years.

There is a clear parallel between this approach and those that study the Moran effect as it applies to population synchrony (Bjørnstad *et al.*, 1999). If interannual variation in timing of breeding positively covaries among populations (Figure 3.1), this may indicate shared environmental conditions. This deductive approach can then be used to identify the groupings of populations that share a correlated response, such as (i) taxonomic relatedness, (ii) breeding at the same site or in the same region, or (iii) wintering in the same areas. Furthermore, a spatial forecast horizon approach may be used to assess the distance beyond which populations may cease to share a correlated response (Petchey *et al.*, 2015), which may be used to inform conservation management decisions (Oppel *et al.*, 2018). In lieu of identifying the environmental conditions themselves, it is therefore possible to identify the scales at which different populations are responding similarly to interannual variation in an environmental driver. Similar studies have applied these techniques to examine inter- and intra-specific covariance in survival (Pyper *et al.*, 2004), productivity (Bond *et al.*, 2011), phenology (Raimondo *et al.*, 2004) and fluctuating population dynamics (Jarillo *et al.*, 2018), with a view to further understanding the impacts of harvesting, trophic interactions and climate change on their demography.

In this study, we aimed to identify the degree to which 51 populations (defined as a species breeding at a particular site) of nine seabird species breeding in the North Atlantic show correlated breeding phenology responses. In particular, we determined the degree to which interannual breeding phenology positively covaries (i.e phenology varies in tandem) between different groupings of these birds. We examine the evidence to support five specific hypotheses: 1) if there is a large-scale driver of phenology in the North Atlantic ocean basin that varies in a similar way from year to year across space and all populations respond similarly to that driver, then between-year variation in breeding phenology will positively covary among all

populations; 2) if drivers of phenology vary between regions, or if a large-scale driver elicits different responses between regions, then the phenology of populations that share similar breeding or wintering regions will positively covary among years; 3) if large-, regional-, or local-scale conditions drive phenology, but responses to these drivers are species-specific, then the phenology of populations of the same species will positively covary at a large-, regional-, or local-scale; 4) if local conditions drive phenology, then populations of different species at an individual site or of the same or different species at sites in close proximity will respond in similar ways across years, and; 5) if local conditions drive phenology, then covariance of lay date will decrease as the distance between populations of a species increases.

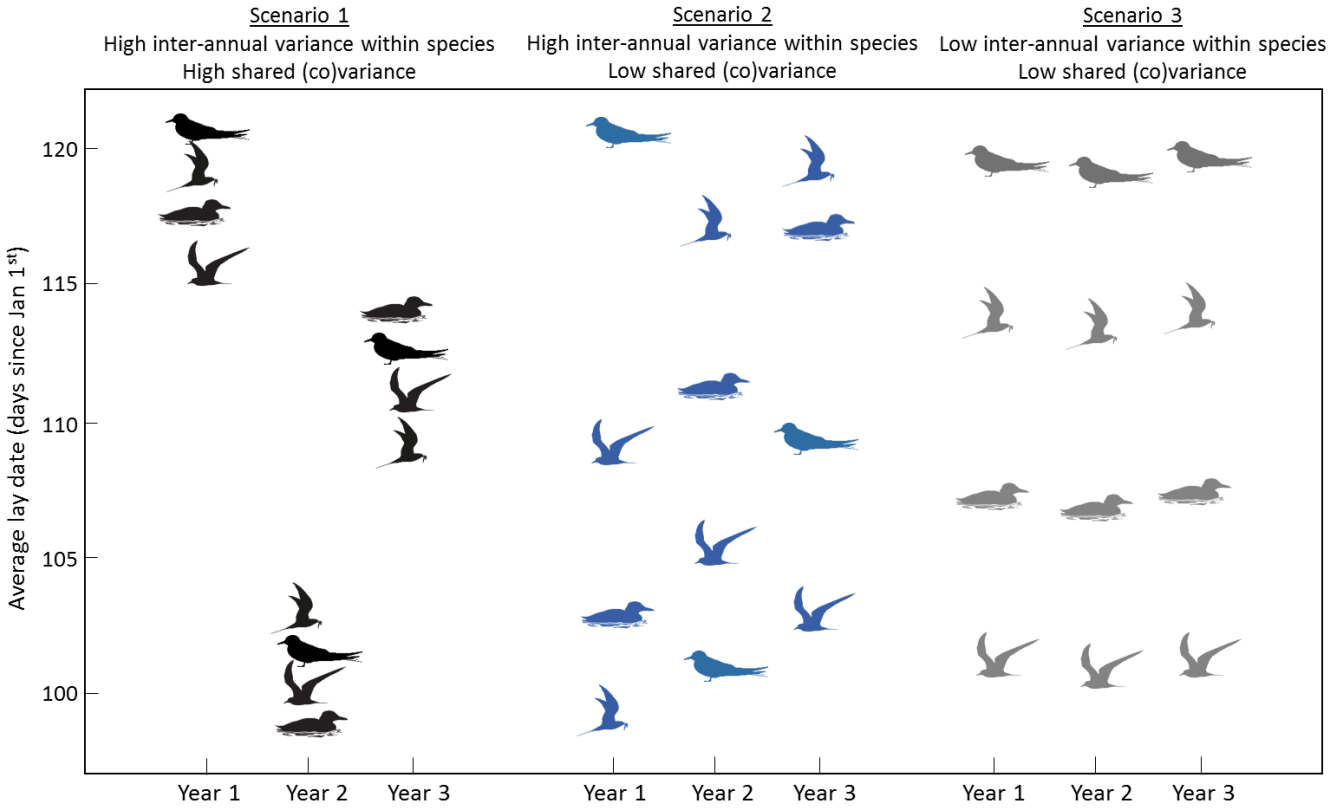


Figure 3.1. Schematic representation of covariance in phenology across hypothetical populations of four seabird species. Each image represents a population, and these populations may be grouped together by common attributes to assess the scale at which timing of breeding covaries. In this case, we represent populations of different species occupying the same breeding site. In Scenarios 1 & 2, populations of each species vary in laying date across years, while in Scenario 3 variance is negligible. Populations in Scenario 1 (black) have high positive co-variance in lay date across years, indicating a shared response to one or more environmental drivers. Populations in Scenario 2 (blue) also vary across years, but do not positively co-vary, and it is unlikely that they share a similar response to one or more environmental drivers. Populations in Scenario 3 (grey) show low variance and co-variance across years, so they are unlikely to be responding to the environment at the same scale.

3.3 Methods

3.3.1 Data collection

We compiled annual average phenological data on nine North Atlantic seabird species (black-legged kittiwake (*Rissa tridactyla*), common tern (*Sterna hirundo*), roseate tern (*Sterna dougallii*), Arctic tern (*Sterna paradisaea*), European shag (*Phalacrocorax aristotelis*), razorbill (*Alca torda*), Atlantic puffin (*Fratercula arctica*), common guillemot (*Uria aalge*) and Brünnich's guillemot (*Uria lomvia*)). Annual data on breeding phenology during the period from 1968 to 2017 were (in order of preference): median lay date; mean lay date; median hatch date; mean hatch date; first hatch date of the study population (defined as a species breeding at a particular site), in units of ordinal days. We used only one measure of phenology for each population, and where only hatch date was available, we back-calculated lay date using information on the average incubation period (Sources in Table A.7). All time series were a minimum of eight years to allow the estimation of covariance between all populations, although the years did not need to be consecutive. The minimum number of individuals on average per population per year was > 10, which we set to limit the contribution of measurement error to our measures of annual central tendency, although in most cases the sample size was considerably higher than this. Low sample sizes may inflate the interannual variation within a population, and to ensure that our sample size was sufficient we ran an additional analysis with this threshold increased to > 50 (the key results remained the same: see Appendix A.2). Field methods used to collect each time series are summarised in Table A.7.

For each time series we collated information on the latitude and longitude of the breeding site, and categorised sites as being either east (< 35° W) or west (> 35° W) coast to assess whether covariance could be attributed to the similarities of conditions observed on either side of the North Atlantic. Sites were also assigned to

one of eleven breeding regions using the Large Marine Ecosystem (LME) classification to assess covariance at a smaller spatial scale (Figure 3.2a, Table 3.1). The wintering region of individuals from each population was determined from available published tracking data, with eleven potential wintering grounds identified in total (Figure 3.2b, Table 3.1, see Table A.7 for sources). Individuals from the same population often use multiple wintering locations, but as this was a population level analysis, wintering regions could not be split within a single time series. We accounted for this in different ways depending on the tracking data available for a population. For most populations, information came from published papers (cited in Table A.7), which identified the most common locations used to overwinter for each species. For 11 Norwegian populations and two which breed in Shetland, we used information from seatrack.seapop.no/map/, which presents wintering distributions from 2 - 4 years in kernel distribution maps. Based on visual inspection of the maps we assigned a wintering distribution as the region where >80% of individuals within a population spent the winter across all years available. Within a wintering region that contained populations of more than one species, we acknowledged the potential for conspecifics to covary more than heterospecifics by grouping populations of the same species into a new variable with four potential "wintering species subgroups" (Table A.7).

Breeding sites were distributed across the North Atlantic, which could introduce spatial autocorrelation in the degree to which phenology of different populations covaried among years. For this reason, if breeding sites were < 120 km apart, a new variable ("site group") was generated that grouped them together. We chose 120 km based on average foraging ranges during the breeding season of the study species which are generally markedly less than this value (Thaxter *et al.*, 2012). 120 km also represents a natural discontinuity in the data i.e. 23/51 (45%) of populations were clustered within groups where all sites were < 120 km apart and the

remaining 28 populations bred at sites which were more than 120 km from another site (Figure 3.2a). This classification allowed us to estimate the average positive covariance between populations within a site group. If a group of sites held more than one population of the same species, we took this into account by creating a new intraspecific variable ("breeding species subgroup") to group these populations together.

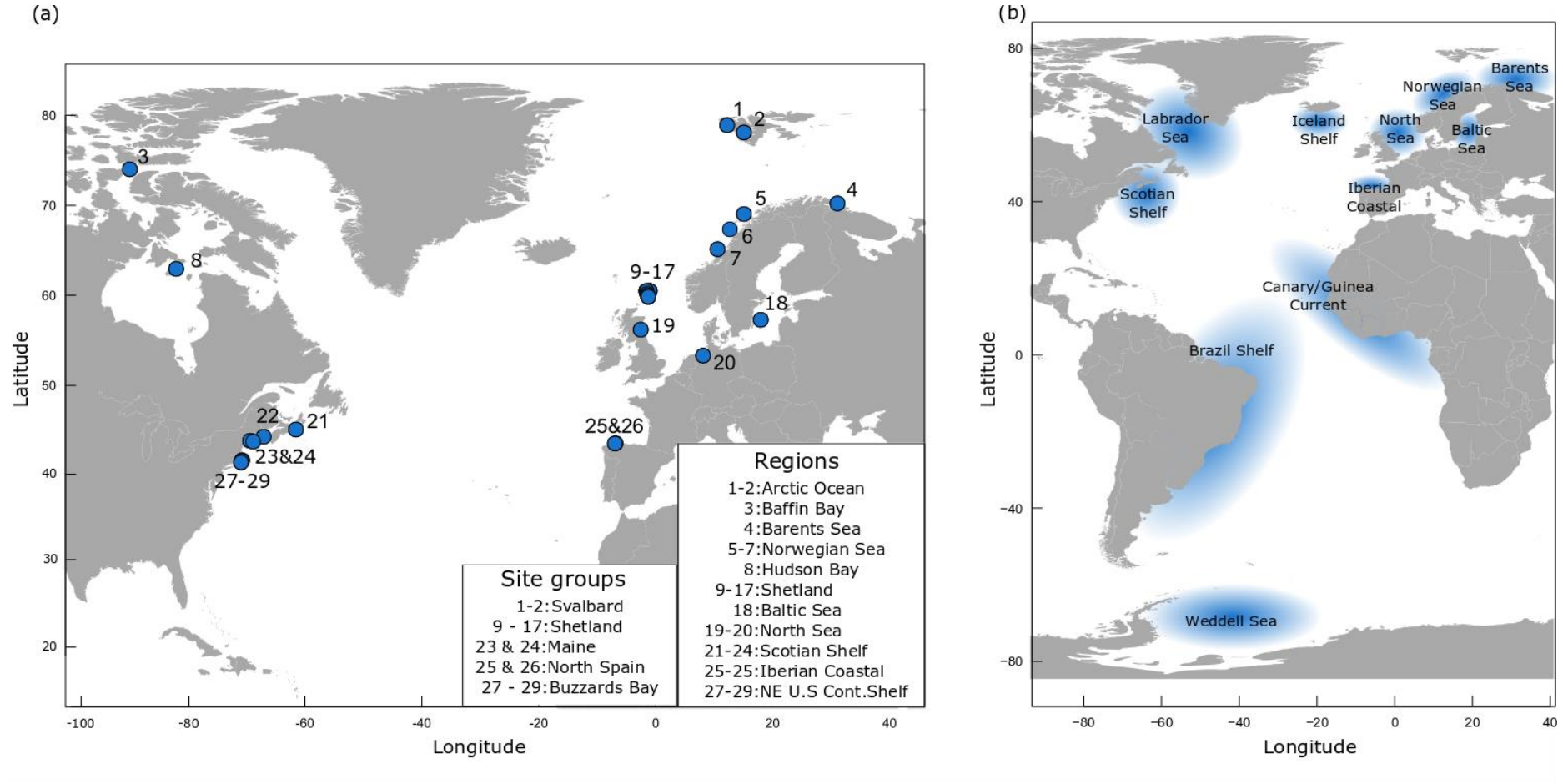


Figure 3.2 Map of sites in the North Atlantic included in the analyses. a) during the breeding season. Site numbers correspond to the breeding sites named in Table 3.1. Sites are numbered according to the region in which they occur. Only regions in which data for more than one site were available were included in the analysis of the annual covariance across regions. b) during winter. Wintering areas represent the general area used by each population over winter. For further information, sources and site coordinates see Table A.7

Table 3.1. List of breeding sites and species included in the analyses in order of decreasing latitude, with breeding and wintering regions indicated. Site numbers correspond to those in Figure 3.2a. Species are as follows: KI = black-legged kittiwake, CT = common tern, RT = roseate tern, AT = Arctic tern, SH = European shag, RA = razorbill, AP = Atlantic puffin, CG = common guillemot, BG = Brünnich's guillemot, with numbers in parenthesis indicating the number of populations of each species included in the analyses. A region was only included in the analysis of annual covariance when data for more than one population were available.

	Site	AP (6)	RA (3)	CG (4)	BG (2)	SH (6)	KI (16)	AT (3)	CT (7)	RT (4)	Breeding region	Winter region
1	Kongsfjorden						x				Arctic Ocean[d] (as Svalbard site group)	Labrador Sea[c]
2	Grumantbyen						x					
3	Prince Leopold Island[a]				x		x				Baffin Bay	Labrador Sea[c]
4	Hornøya[a]	x	x	x			x				Barents Sea	Barents Sea[c] / Norwegian Sea / Barents Sea[c] / Labrador Sea[c]
5	Anda[a]	x					x				Norwegian Sea[b]	Iceland Shelf[c] / Labrador Sea[c]
6	Røst[a]	x				x	x				Norwegian Sea[b]	Iceland Shelf[c] / Norwegian Sea[c] / Labrador Sea[c]
7	Sklinna[a]	x				x					Norwegian Sea[b]	Iceland Shelf[c] / Norwegian Sea[c]
8	Coats Island				x						Hudson Bay	Labrador Sea[c]
9	Burravoe						x				North Sea[b] (also as Shetland site group)	Labrador Sea[c]
10	Esha Ness						x					Labrador Sea[c]
11	Westerwick						x					Labrador Sea[c]
12	Ramna Geo						x					Labrador Sea[c]
13	Kettla Ness						x					Labrador Sea[c]
14	No Ness						x					Labrador Sea[c]
15	Troswick Ness						x					Labrador Sea[c]
16	Compass Head						x					Labrador Sea[c]
17	Sumburgh Head[a]			x		x	x					North Sea[c] / North Sea[c] / Labrador Sea[c]
18	Stora Karlsö			x							Baltic Sea	Baltic Sea

19	Isle of May[a]	x	x	x		x	x			North Sea[b]	North Sea[c] / North Sea[c] / North Sea[c] / North Sea[c] / Labrador Sea[c]
20	Banter See							x		North Sea[b]	Canary or Guinea Current
21	Country Island[a]						x	x	x	Scotian Shelf[b]	Weddell Sea[d] / Brazil Shelf [c] / Brazil Shelf [c]
22	Machias Seal Island[a]	x	x				x	x		Scotian Shelf[b]	Gulf of Maine[c] / Unknown / Weddell Sea[d] / Brazil Shelf [c]
23	Eastern Egg Rock							x		Scotian Shelf[b] (also Maine site group)	Brazil Shelf [c]
24	Matinicus Rock						x				Weddell Sea[d]
25	A Forcada				x					Iberian Coastal[b,d] (as North Spain site group)	Iberian Coastal[c]
26	As Pantorgas				x						
27	Bird Island[a]							x	x	North East U.S Continental Shelf[b,d] (as Buzzards Bay site group)	Brazil Shelf[c]
28	Ram Island[a]							x	x		
29	Penikese Island[a]							x	x		

[a] represents sites over which (co)variance between years was estimated, [b] represents breeding regions for which among-year (co)variances were estimated, [c] represents wintering regions for which among-year (co)variances were estimated. [d] represents effects which are confounded because the same combination of populations is grouped into another term, see main text for details. Confounded terms were not included in the model unless specified in the main text.

Table 3.2. Syntax and description of random terms used in the analyses. b indicates terms included in the breeding model, w indicates terms included in the wintering model.

MCMCglmm Random Term Syntax	Number of Parameters	Description	Hypothesis addressed
Year[b,w]	1	Partitions the among-year variance in annual phenology means across all populations breeding in the North Atlantic. Provides an estimate of the magnitude of a shared response to a trans-North Atlantic driver.	Hypothesis 1
idh(breeding region):year[b,w]	3	Partitions among year variance in the annual mean phenology of all populations in the same breeding region. This accounts for populations in a region sharing a phenological response to a common regional driver during the summer.	Hypothesis 2 (Table 3.1, Figure 3.2)
idh(wintering region):year[w]	8	Partitions among year variance in the annual mean phenology of all populations that share the same wintering region. This accounts for populations in a region sharing a phenological response to a common regional driver during the winter.	Hypothesis 2 (Table 3.1, Figure 3.2)
idh(species):year[b,w]	9	Partitions among year variance in the annual mean phenology of all populations that belong to the same species. This takes into account the potential for species to share a phenological response to a spatially consistent driver	Hypothesis 3 (Table 3.1)
idh(site group):year[b]	5	Partitions among year variance in the annual mean phenology of all populations found within 120km of each other (“site groups”). This accounts for a shared phenological response to conditions at nearby sites.	Hypothesis 4 (Table A.7)
idh(breeding species subgroup):year[b]	3	Partitions among year variance in the annual mean phenology of all populations within a site group (above) that belong to the same species. This takes into account the potential for a shared	Hypothesis 3 (Table A.7)

		phenological response to conditions at nearby sites, but allows this response to differ across species.	
idh(wintering species subgroup):year^{w}	4	Partitions among year variance in the annual mean phenology of all populations of the same species that share the same wintering region. This accounts for populations in a region sharing a phenological response to a common regional driver during winter, but allows this response to differ across species.	Hypothesis 3 (Table A.7)
idh(Site):year^{b}	12	Partitions among year variance in the annual mean phenology of all populations found at the same site. This accounts for a shared phenological response to local conditions.	Hypothesis 4 (Table 3.1, Figure 3.2)
rcov=~idh(Population):year^{b,w}	51	Allows for the residual among year variance to be heterogeneous across all populations. High residual variance implies that phenology is largely determined by a driver that is idiosyncratic to the population.	None (Table A.7)

3.3.2 Statistical Analyses

We used the *MCMCglmm* package (Hadfield, 2010) in R (v 3.2.2; R Core Team 2015), to fit linear mixed-effect models (LMMs) in a Bayesian framework. In these models, the yearly average breeding phenology of each population is a Gaussian response and random terms are used to control for differences in mean timing among populations and, more centrally to our aims, to identify the sources of positive covariance in phenology among populations (see Table 3.2 for full list of terms used). We used separate models to distinguish the positive covariance among populations that share breeding grounds (core model) versus wintering grounds (wintering model). The core model included latitude and the continental coast of the breeding site (east or west Atlantic Ocean) as fixed effects to account for geographic variation in the multi-year mean phenology of populations.

We used random terms in two ways. First, we controlled for variation among populations in the mean phenology of each time series, by including breeding region, species, site group (groups of sites that are < 120km apart), breeding species subgroup (site group:species), site, and population (site:species) as random terms. Our further random terms were all focused on partitioning the among year variance so that we could examine the extent to which populations covary (Table 3.2). The year random term estimates the degree to which the phenology of all populations covaries from year to year. This term does not allow for the potential for some populations to positively covary more than others. To accommodate this, where there was replication of time-series' within a level of a random term we allowed year variances to be heterogeneous (using the idh structure in MCMCglmm, Appendix A.3) across levels of breeding region, species, site group, breeding species subgroup and site. These variance estimates allow us to infer the degree to which populations in each level of the random term (e.g., a breeding region) positively co-vary in addition to the overall

year variance. For each of the random terms which included heterogeneous year effects, high values within a level of the random term indicated positive covariance among associated time series, indicating similar patterns of early or late breeding (Figure 3.1a). Conversely, low covariance indicated no uniformly early or late breeding events among the time series (Figure 3.1b,c). We allowed for heterogeneity in variance among years only where data were available for two or more populations in each region, species, site, site group, breeding species subgroup, or wintering region. We also allowed the residual year variance to be heterogeneous across populations. For all random terms, effects were drawn from a normal distribution with mean = 0, and with the variance estimated.

Given the five alternative random terms in the core model (i.e. by region, species, site group, breeding species subgroup, and site), the combination of populations was sometimes the same for more than one level across them. For example, both populations of European shag in North Spain were located < 120 km apart and were therefore included in the same site group, and this same combination was found in the breeding region, Iberian coastal. It was therefore not possible to estimate year variance across both of these groupings, and only the grouping in terms of close proximity (breeding sites < 120 km apart) was included. In such cases the effect of local conditions versus regional conditions are confounded and we highlight such cases in the results.

In the wintering model, we tested for positive covariance among populations based on the wintering region. Year, species, population, and the interaction between year:species and year:population were retained from the core model. We also retained the effect of breeding species subgroup and the interaction with year to control for similar responses of adjacent populations of the same species, which may travel to the same wintering region. This was particularly important for kittiwake

populations in Shetland (comprising nine populations). In addition to estimating positive covariance in phenology among all populations wintering in the same region, we also estimated the species-specific positive covariance among populations across years by including (wintering species subgroup):year as an effect.

All models were run for 250,000 iterations, discarding the first 3,000 as burn-in and sampling every 10th iteration. For the residual priors we used an inverse-Wishart distribution. For the remaining variance terms to improve mixing we adopted parameter-expanded priors (Gelman *et al.*, 2008), which give a scaled F distribution with numerator and denominator degrees of freedom = 1 and scale parameter = 1000 (Gelman, 2006). Trace plots of posterior distributions were examined to assess autocorrelation and model convergence. Statistical significance of fixed effects was inferred where 95% credible intervals (CIs) did not span zero. As variance estimates are bounded at zero, we infer that a random term is significant where the 2.5% CI was removed from zero.

To examine how properties of the data (effect size, replication, number of overlapping years etc.) may affect the accuracy and power of our approach for estimating (co)variances we conducted simulations (Appendix A.4). Simulations of phenology for six populations over 30 years revealed that the method returns unbiased covariance estimates, but that residual variances can be biased if the true covariance between populations within a grouping is negative. The method generally has good power to detect a variance of 40 and moderate power to detect a variance of 20. Power is greatest for parameters where the among year signal (among year variance in that parameter) to noise ratio (among year variance of other parameters that affect the contributing populations) is high and power is only slightly reduced if years are not overlapping (for the number of time series per year see Figure A.2). Power to detect a non-zero covariance is reduced when time series are reduced to

15 years duration and we suggest care should be taken in interpreting covariance estimates with broad credible intervals, as this may reflect low power rather than a true absence.

3.3.3 Correlation in timing of breeding between populations

Using the posterior distribution from the core model, we can derive the predicted among-year correlation coefficient r in phenology between pairs of populations. We calculated this for each possible pairwise intraspecific combination of the four species for which *n* populations > 5 (*n* kittiwake populations = 16, *n* combinations = 120; *n* tern populations = 7, *n* combinations = 21, *n* puffin or shag populations = 6, *n* combinations = 15) using the equation

$$r = \frac{\sigma(X,Y)}{\sigma X \sigma Y}$$

where $\sigma(X,Y)$ is the covariance in interannual phenology between populations X and Y, σX is the standard deviation of phenology in population X and σY is the standard deviation of phenology in population Y. To obtain a model based posterior for $\sigma(X,Y)$, we summed the posterior distributions of all among year (co)variance components that are common to the two populations (i.e. the year variance, the species specific year variance etc.). We obtained the model based posterior of σX (and σY), as the posterior of the square root of $\sigma(X,Y)$ plus the sum of the posterior distributions of the square roots of the among year variances unique to that population, i.e. site-specific year variance, residual year variance. For further details regarding this method refer to Appendix A.3.

`The mixed-model method that we used to estimate phenological correlations among populations has three properties that may introduce bias. First, all among year covariances are constrained to be ≥ 0 and, whilst it is plausible that phenology might covary negatively between populations, we lack a biological explanation for this

phenomenon. Second, in each random term, the mean among year covariance is estimated across all populations within a grouping, such that some pairwise covariances may be over- or under-estimated. Third, within a random term, between-level covariances are treated as 0. To assess whether such biases impact our inferences, we tested whether the pairwise phenological correlations derived from model estimates agreed with pairwise phenological correlations between populations (with ≥ 3 overlapping years), estimated from the raw data using Pearson's correlation (all populations $n = 51$, combinations overlapping ≥ 3 years $n = 1181$). Both estimates of pairwise correlations were plotted against pairwise great circle distance between populations (km) and visually inspected to identify any spatial dependency in the tendency for the model derived correlations to over- or under-estimate the Pearson's correlations.

We also used a spatial forecast horizon approach to test the effect of distance on temporal correlation between populations of the same species. For each of the four species, we used a Mantel test in the *vegan* package version 2.5-1 (Oksanen *et al.*, 2018) to estimate the effect of great circle geographic distance on 1 – the among-years breeding time Pearson correlation between pairs of populations.

As a final model diagnostic for our core model we examined the relationship between the quantiles of observed among year pairwise phenological correlations between all populations and those that were predicted based on the model. To generate the expected relationship we conducted 1000 *a posteriori* simulations under the core model, then calculated the pairwise phenological correlations in the resulting data and summarise the expected correlations as the mean quantile values. Using this approach provided us with a quantile-quantile plot for observed versus expected pairwise correlations, which allows us to examine whether our model had any tendency to over or under-predict pairwise correlations.

3.4 Results

The full dataset of 1043 phenological observations (annual means) spanned 50 years of breeding phenology averages from 51 populations across nine species and 29 breeding sites, with more recent years represented by more time-series than earlier years (Figure A.2). Time-series were assigned to 16 site groups (sites < 120 km apart), 11 breeding regions and 11 wintering regions (Figures 3.2a, b, Tables 3.1, 3.2, A.7). All parameter estimates correspond to those obtained from the core model unless the wintering model is specified. Average lay date increased with latitude (b = 1.763 days lat^{-1}, 95% CI = 1.003, 2.555), and, controlling for latitude, laying occurred 38 days later (95% CI = 20.540, 56.141) in the west Atlantic compared to the east Atlantic.

To test whether the phenology of populations in the North Atlantic Ocean basin varies in a similar way from year to year (hypothesis 1) we tested for shared variance in timing between years across all time series'. The shared variance (in units of $days^2$) was very low (σ^2 = 0.181, 95% CI = 0.000, 1.135, years = 49) in comparison to the average amount of variance in lay date shown by each population, indicating that for North Atlantic seabirds in general, early and late years were not shared across all of the populations.

To test whether drivers of phenology vary between regions, or if a large-scale driver elicits different responses between regions (hypothesis 2), we estimated among year phenological covariance between populations sharing similar breeding or wintering regions. We detected no statistically significant covariance within breeding regions (Figure 3.3e, Tables A.1, A.2), although in the Norwegian Sea and the North Sea regions the credible intervals were wide. In the wintering region model, significant variance was shared only among populations in the North Sea (σ^2 =17.879, 95% CI = 9.910, 28.852, time series = 6, Figure 3.4b, Table A.3), with the estimated variance

corresponding to the shared response being in the range of ± 8.3 days in 95% of years. The posteriors for inter-year variance in timing of breeding for populations that wintered in three additional regions (Gulf of Maine, Iceland Shelf and Barents Sea) were somewhat removed from zero, although the 2.5 CI was approximately 0. In the case of breeding populations that are found on the Iceland Shelf the among year variance was potentially large, but there was high uncertainty in these variance estimates (Figure 3.4b).

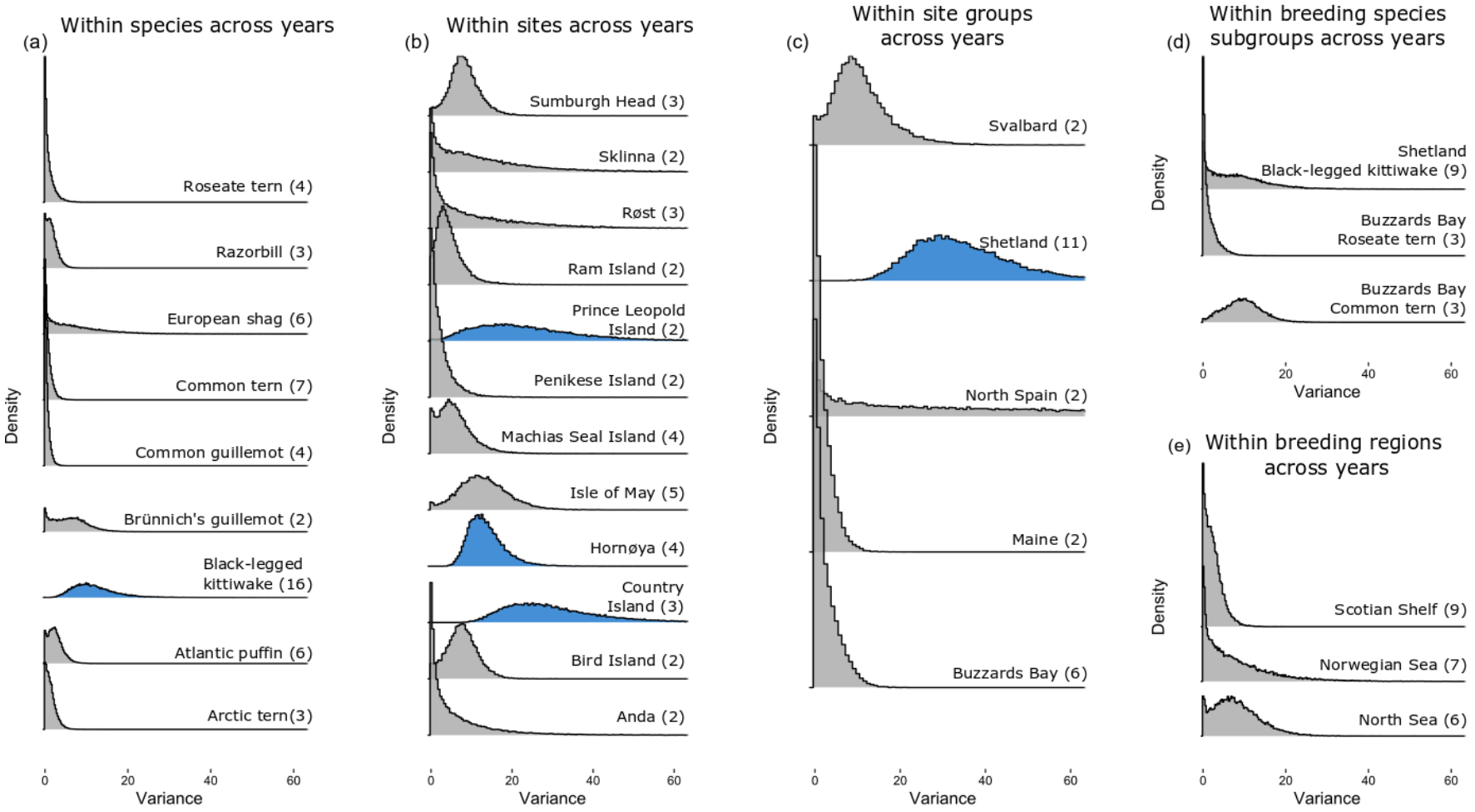

(a)
Within species across years
Roseate tern (4)
Razorbill (3)
European shag (6)
Common tern (7)
Common guillemot (4)
Brünnich's guillemot (2)
Black-legged kittiwake (16)
Atlantic puffin (6)
Arctic tern(3)
Density
0
20
40
60
Variance
(b)
Within sites across years
Sumburgh Head (3)
Sklinna (2)
Røst (3)
Ram Island (2)
Prince Leopold Island (2)
Penikese Island (2)
Machias Seal Island (4)
Isle of May (5)
Hornøya (4)
Country Island (3)
Bird Island (2)
Anda (2)
Density
0
20
40
60
Variance
(c)
Within site groups across years
Svalbard (2)
Shetland (11)
North Spain (2)
Maine (2)
Buzzards Bay (6)
Density
0
20
40
60
Variance
(d)
Within breeding species subgroups across years
Shetland Black-legged kittiwake (9)
Buzzards Bay Roseate tern (3)
Buzzards Bay Common tern (3)
Density
0
20
40
60
Variance
(e)
Within breeding regions across years
Scotian Shelf (9)
Norwegian Sea (7)
North Sea (6)
Density
0
20
40
60
Variance

Figure 3.3. (Co)variance in timing of breeding of seabird populations across years during the breeding season. Plotted from the posterior distribution of the core random-effects model, representing shared variance across years according to (a) species, (b) site, (c) site groups (< 120 km apart), (d) breeding species subgroups (i.e. populations of the same species within a group of nearby sites), and (e) breeding regions. On the y axes labels, values in parenthesis indicate the number of populations associated with each term. For interpretation, narrower histograms indicate a posterior distribution that has been estimated with higher precision (i.e. a smaller credible interval), and histograms with a centre of mass further removed from zero represent more posterior support for a positive (co)variance. Groups for which significant positive covariance was estimated (i.e. where 2.5% credible interval was not ~ 0) are shaded in blue.

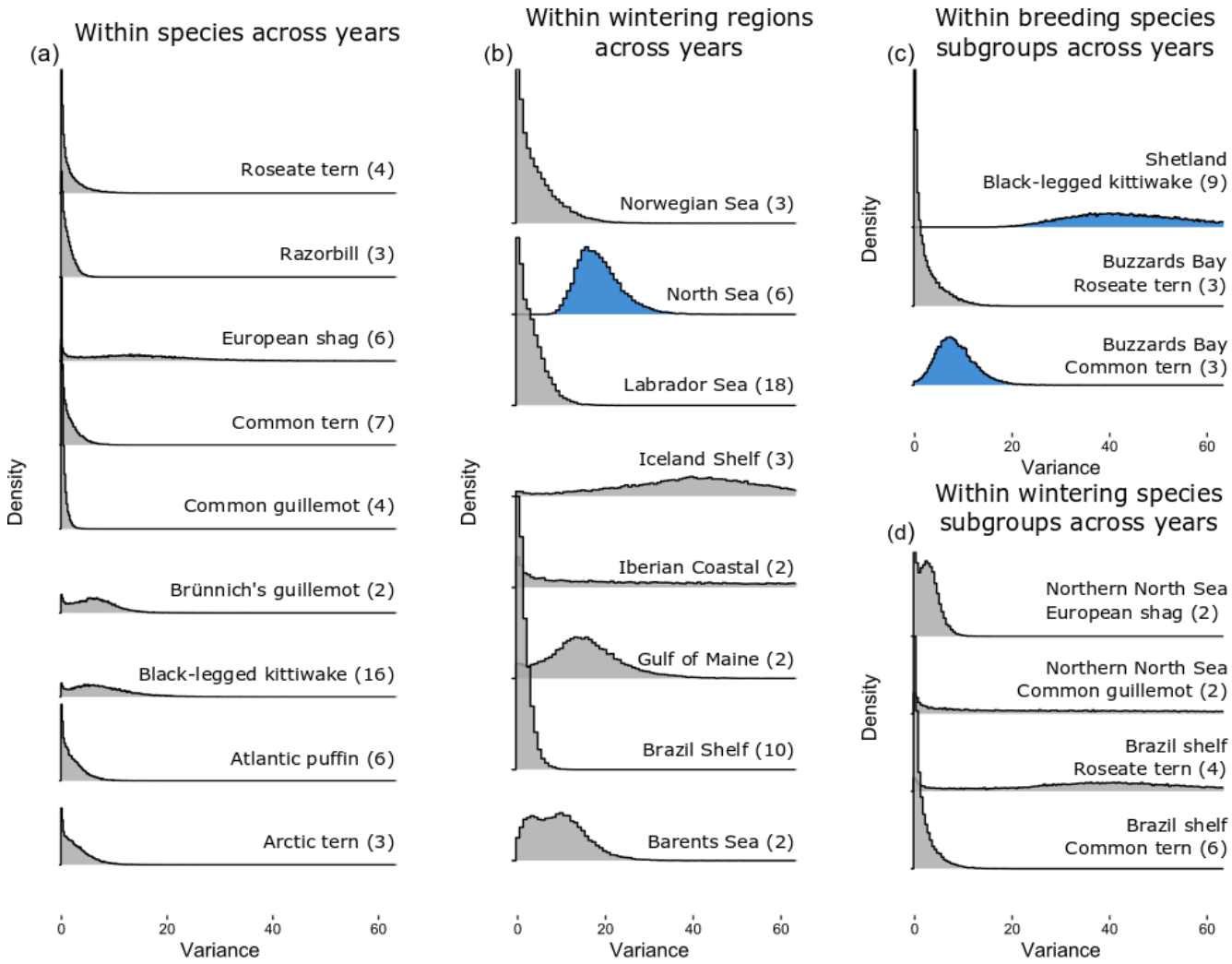
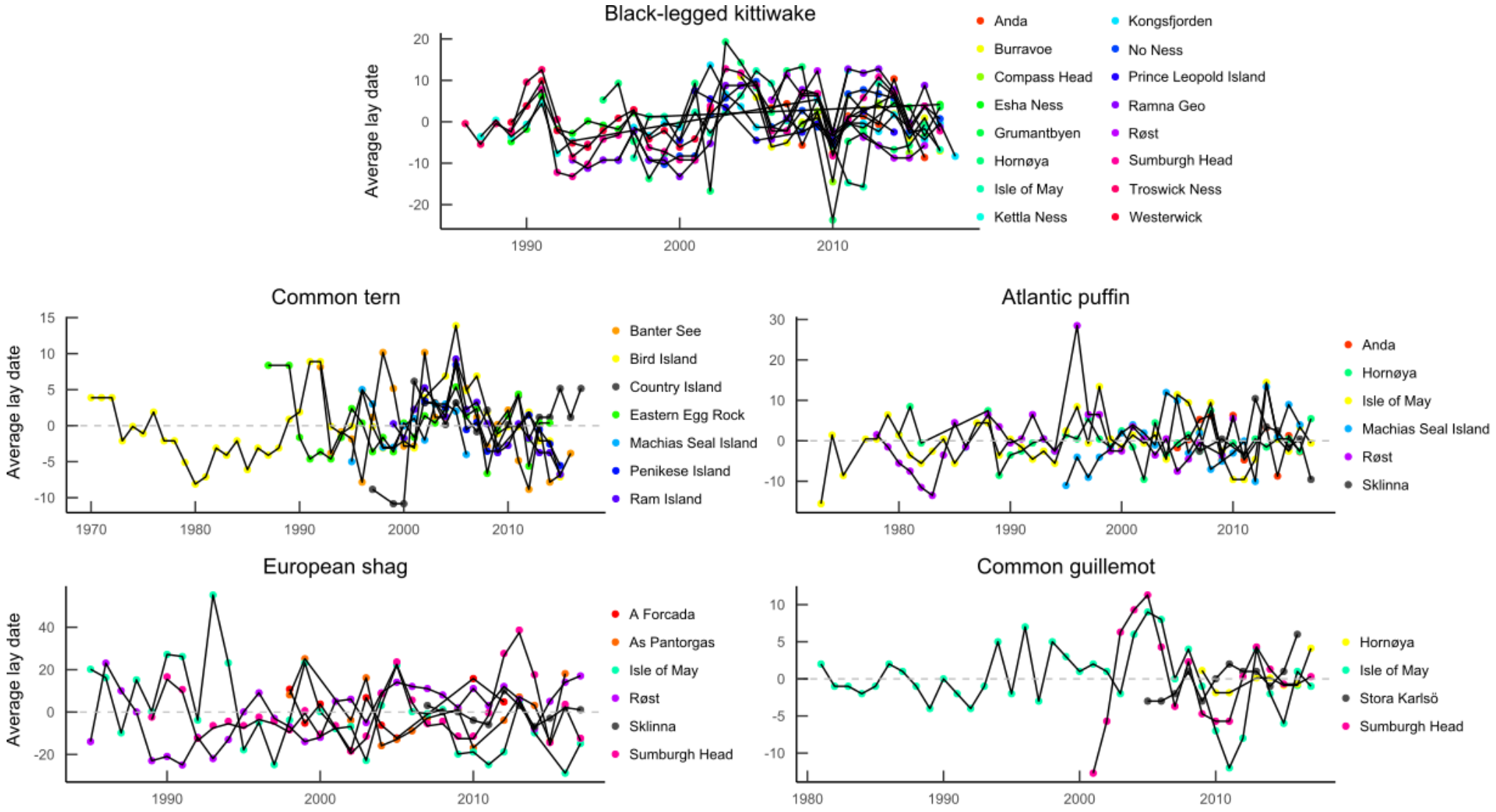

(a)
Within species across years
Roseate tern (4)
Razorbill (3)
European shag (6)
Common tern (7)
Common guillemot (4)
Brünnich's guillemot (2)
Black-legged kittiwake (16)
Atlantic puffin (6)
Arctic tern (3)
Density
0
20
40
60
Variance
(b)
Within wintering regions across years
Norwegian Sea (3)
North Sea (6)
Labrador Sea (18)
Iceland Shelf (3)
Iberian Coastal (2)
Gulf of Maine (2)
Brazil Shelf (10)
Barents Sea (2)
Density
0
20
40
60
Variance
(c)
Within breeding species subgroups across years
Shetland
Black-legged kittiwake (9)
Buzzards Bay
Roseate tern (3)
Buzzards Bay
Common tern (3)
Density
0
20
40
60
Variance
(d)
Within wintering species subgroups across years
Northern North Sea
European shag (2)
Northern North Sea
Common guillemot (2)
Brazil shelf
Roseate tern (4)
Brazil shelf
Common tern (6)
Density
0
20
40
60
Variance

Figure 3.4. (Co)variance in timing of breeding across years of seabird populations wintering in different regions. Plotted from the posterior distribution of the winter random-effects model, representing shared variance across years with in (a) species, (b) winter region, (c) breeding species subgroup (i.e. populations of the same species within a group of nearby sites), and (d) wintering species subgroup (i.e. populations of the same species within a wintering region). On the y axes labels, values in parenthesis indicate the number of populations associated with each term. For interpretation, narrower histograms indicate a posterior distribution that has been estimated with higher precision (i.e. a smaller credible interval), and histograms with a centre of mass further removed from zero represent more posterior support for a positive (co)variance. Groups for which significant positive co-variance was estimated (i.e. where 2.5% credible interval was not ~ 0) are shaded in blue. In (b) note that populations in both the Gulf of Maine and the Barents Sea come from the same breeding site, and thus the effects of breeding and wintering conditions cannot be separated in these cases.

We tested for covariance between populations of the same species to address hypothesis 3 - that if large-, regional- or local-scale conditions drive phenology, the conditions may be species-specific. Our models allowed for the year variance to be estimated within nine species, which estimates the covariance in annual means between populations of the same species (Figure 3.5). This among-year variance was only significant for black-legged kittiwakes (σ^2 = 11.038, 95% CI = 2.761, 22.639, time series = 16, Figure 3.3a, Table A.1). Under a normal distribution with mean = 0 and variance = 11.038, the shared deviations in timings are expected to lie in the range ± 6.5 days in 95% of years. Between populations of the same species within a site group, the posterior suggested positive covariance for common terns, although the 2.5% CI was approximately zero (Figure 3.3d, Tables A.1, A.2). We found no evidence that breeding phenology of populations of the same species within a wintering region covaried (Figure 3.4d, Table A.3).


Black-legged kittiwake
Average lay date
20
10
0
-10
-20
1990
2000
2010
Anda
Burravoe
Compass Head
Esha Ness
Grumantbyen
Hornøya
Isle of May
Kettla Ness
Kongsfjorden
No Ness
Prince Leopold Island
Ramna Geo
Røst
Sumburgh Head
Troswick Ness
Westerwick
Common tern
Average lay date
15
10
5
0
-5
-10
1970
1980
1990
2000
2010
Banter See
Bird Island
Country Island
Eastern Egg Rock
Machias Seal Island
Penikese Island
Ram Island
Atlantic puffin
30
20
10
0
-10
1980
1990
2000
2010
Anda
Hornøya
Isle of May
Machias Seal Island
Røst
Sklinna
European shag
Average lay date
40
20
0
-20
1990
2000
2010
A Forcada
As Pantorgas
Isle of May
Røst
Sklinna
Sumburgh Head
Common guillemot
10
5
0
-5
-10
1980
1990
2000
2010
Hornøya
Isle of May
Stora Karlsö
Sumburgh Head

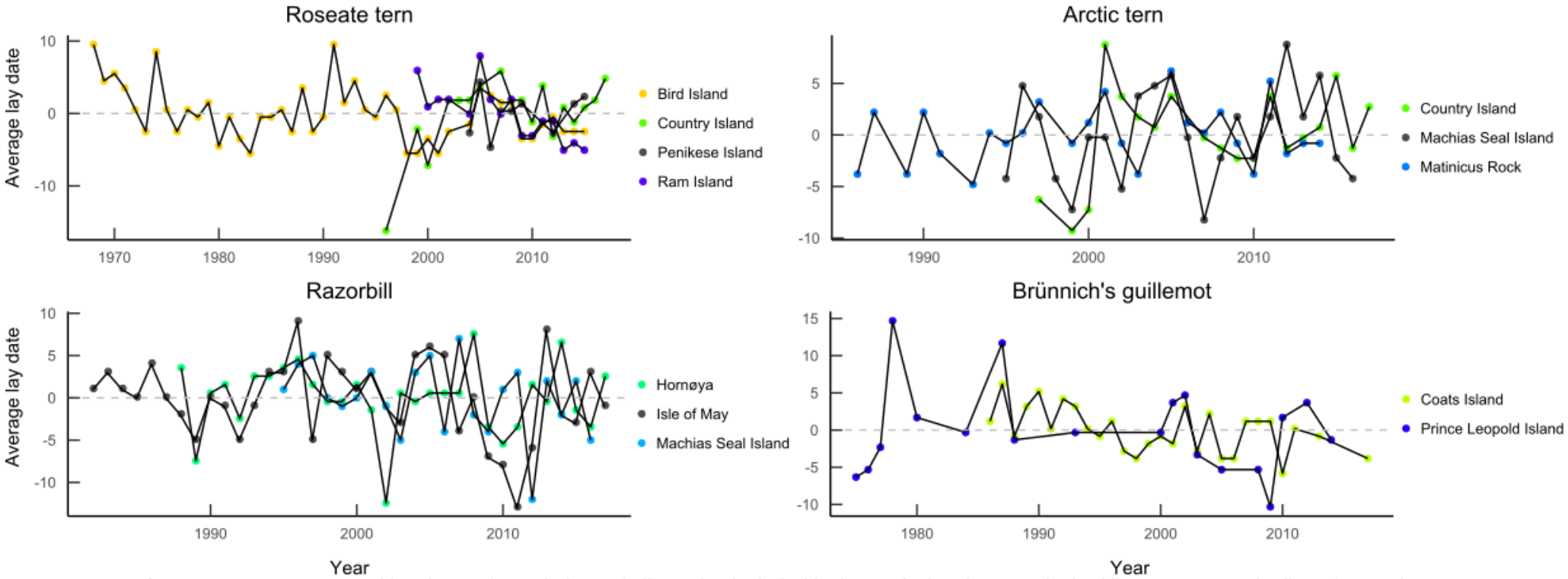


Figure 3.5. Average annual lay dates of populations of all species included in the analysis. The grey dashed line represents the line of central tendency of laying for each species.

We tested whether local conditions drive phenology (hypothesis 4) by estimating whether populations of the same or different species at an individual breeding site or site groups covary across years. Of the 29 breeding sites, 12 held more than one species, allowing interspecific covariance across years to be estimated (Figures 3.3b, 3.6). The site-level among-year variance (i.e. individual species at a site sharing early versus late years) was only significant for three sites: Country Island, Hornøya, and Prince Leopold Island, and the values corresponding to the 95% quantiles of shared phenological deviations (calculated for a normal distribution with mean = 0 and variance specified) were in the range ± 10.4 days, ± 7.0 days and ± 9.1 days, respectively. Whilst the peaks of the posterior distribution for inter-year variance for five additional sites (Bird Island, Isle of May, Machias Seal Island, Ram Island and Sumburgh Head) were removed from 0, the 2.5 CI was approximately 0 (Figure 3.3b). Among-year variance was estimated for five site groups and this tended to be quite high in North Spain, Shetland and Svalbard, but was estimated well only for Shetland (σ^2 = 33.106, 95% CI = 14.253, 59.942, time series = 11, Figure 3.3c, Table A.1).

Residual variance was significant for 18 of the 51 populations (35%, Table A.1). Average residual variance varied substantially among species (Table 3.3), being most pronounced in European shags.

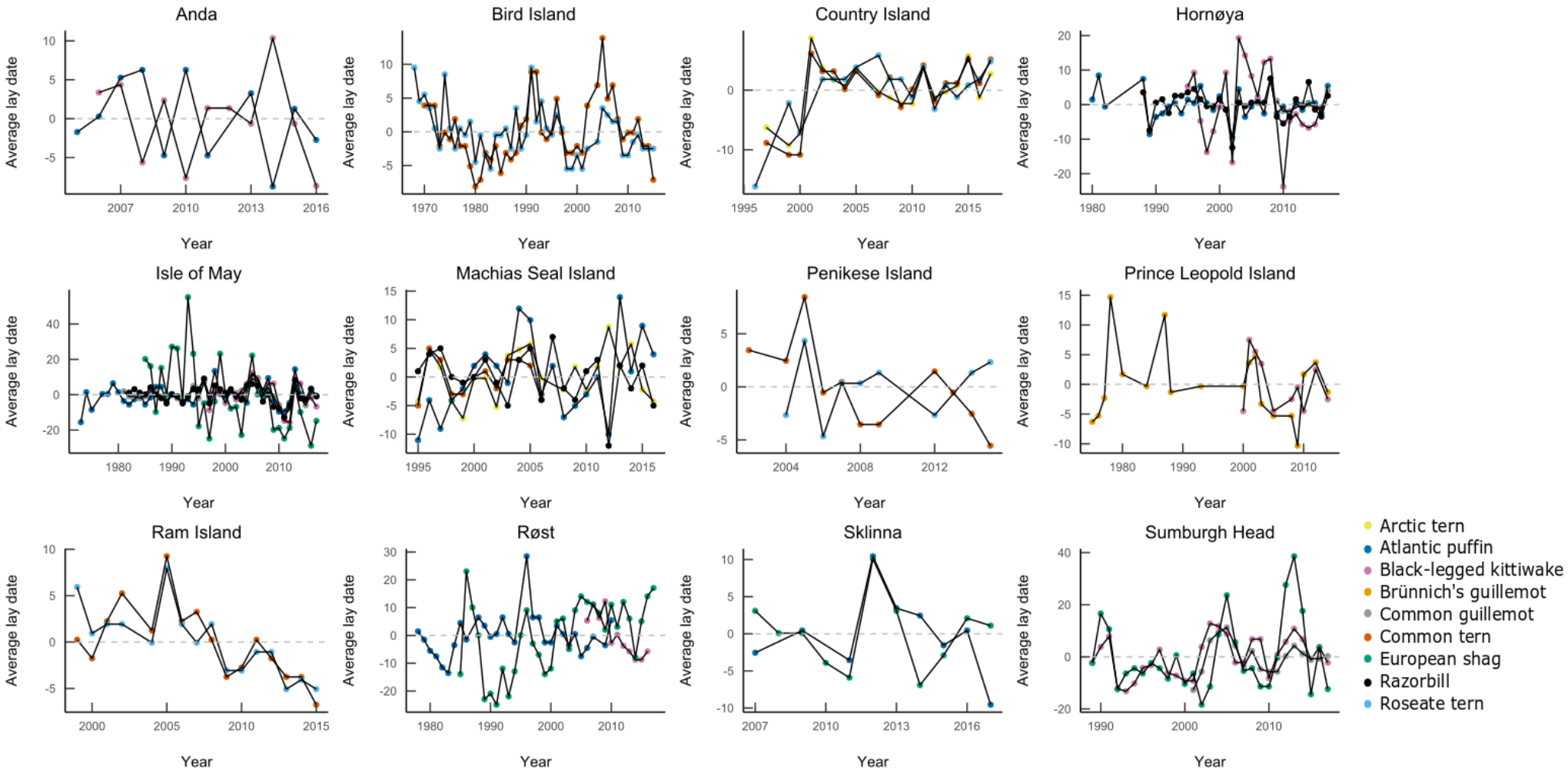


Figure 3.6. Average annual lay dates of populations at twelve sites for which more than one time series was available for analysis. The grey dashed line represents the central tendency of laying at each site.

Table 3.3. Median residual variance for the nine species included in the analysis in order of decreasing variance. Residual variance is calculated from the core random effects model, and species are placed in order from highest to lowest values. Numbers in brackets indicate 95% credible intervals for the species means. 95% range in days corresponds to the 0.025 and 0.975 quantiles of a normal distribution of mean = 0 and σ calculated from the residual variance.

Species	Median among-year residual variance	95% range in days
European shag	138.33 (55.85 - 309.83)	± 23.05 days
Atlantic puffin	21.86 (4.92 - 56.39)	± 9.16 days
Black-legged kittiwake	9.31 (3.59 - 41.94)	± 5.98 days
Roseate tern	6.38 (0.58 - 18.59)	± 4.95 days
Brünnich's guillemot	5.32 (0.00 - 21.06)	± 4.52 days
Razorbill	5.3 (1.2.0 - 15.25)	± 4.51 days
Arctic tern	3.66 (1.36 - 13.86)	± 3.75 days
Common tern	2.24 (2.05 - 15.06)	± 2.93 days
Common guillemot	1.76 (0.68 - 12.46)	± 2.60 days

We tested whether covariance of lay date decreased as the distance between populations increased (hypothesis 5). Intraspecific pairwise Pearson's correlations of annual phenology between populations of black-legged kittiwakes, Atlantic puffins, and European shags decreased with increased distance (black-legged kittiwake: Mantel statistic [between distance and 1-correlation] $r = 0.572$, $p = 0.002$; Atlantic puffin: $r = 0.750$, $p = 0.025$; European shag: $r = 0.847$, $p = 0.008$. Figure S3.4, Table 3.4). Correlations estimated from the raw data were compared against those from the posterior distribution of the core random effects model as a diagnostic of the performance of the mixed-model approach for the four species. The model-based estimates corresponded well with estimates from pairwise correlations using the raw data, and captured the spatial decay in pairwise correlations (Figure A.4).

The *a posteriori* quantile-quantile plot for pairwise population correlations revealed an excellent correspondence between empirical and model based quantiles (Figure A.5). The model yielded a similar frequency of negative pairwise correlation between populations, such that the observed frequency of negative phenological correlations was consistent with that expected by chance in the absence of true negative covariances.

Table 3.4. Results of Mantel tests of the effect of distance on correlation of lay date, and minimum, maximum and mean correlations for each species using all pairwise correlations from raw data, and correlations where each pairwise correlation was based on 10 or more years. Higher values indicate more tightly correlated phenology between closely positioned populations.

	Black-legged kittiwake	Common tern	Atlantic puffin	European shag
Mantel statistic r (all data)	**0.412****	0.216	**0.586***	**0.535*****
Min. correlation	-0.830	-0.436	-0.428	-0.613
Max. correlation	0.976	0.867	0.777	0.607
Mean correlation	0.368	0.347	0.028	-0.025
Mantel statistic r (>10 years overlap)	**0.572****	0.218	**0.750***	**0.847****
Min. correlation	-0.385	-0.436	-0.428	-0.613
Max. correlation	0.917	0.867	0.387	0.607
Mean correlation	0.298	0.342	0.037	0.030

* < 0.05, ** < 0.01, ***< 0.005

3.5 Discussion

Timing of breeding is often used as an indicator of response to environmental change, yet for many species the drivers of phenology and the spatiotemporal scale at which they operate remain unclear. We examined phenology in a group of North Atlantic seabird populations with diverse breeding and migration strategies, wide breeding ranges and high between-year variance in breeding phenology, and asked to what extent populations share early versus late breeding seasons between sites, species, breeding and wintering regions. Using this approach, we established that in general, laying dates of populations tended to covary at the site or small-scale regional level. Additionally, we found that black-legged kittiwakes positively covary in laying dates with conspecifics at ocean basin scales, and heterospecifics at local scales, suggesting that phenology in this species is sensitive to environmental conditions in both the wintering and pre-breeding seasons. Overall, we shed light on the potential

factors to drive phenology in our study species thus furthering our understanding of the variation in the scales at which different seabirds interact with a variable environment, and their capacity to respond to future environmental change.

We found no evidence that populations in this study collectively breed early or late (hypothesis 1), suggesting that if there is a common driver of phenology in the North Atlantic, such as the North Atlantic Oscillation, it does not act in the same way across regions, or elicit a consistent response across populations. However, we did identify a pronounced difference in the mean timings between the east and west Atlantic, with phenology more than a month later in the west. This is potentially due to the differences in temperature of the currents passing each coast (southward flowing Labrador Current being cold in comparison with the warmer and northward flowing Gulf Stream) which leads to more pronounced seasonality in water temperature in the west for a given latitude (Mackas *et al.*, 2012).

We found no shared variance between populations occupying the same breeding region (hypothesis 2), which may originate from a number of mechanisms. Populations of seabirds may partition feeding areas so as to reduce competition (Wakefield *et al.*, 2013), such that although they occupy the same general region, the scale, magnitude and direction of any adjustment in timing of breeding in response to the environment may differ across sites within it. Indeed, primary productivity (Behrenfeld *et al.*, 2006) and abundance of prey (Frederiksen *et al.*, 2005) vary in their temporal availability at spatial scales smaller than the regional basin categorisation used in this study, which potentially explains why such an effect was not observed. Bathymetry, tides and currents are all important for prey distributions and aggregations, and thereby for seabird foraging (Amélineau *et al.*, 2016; Christensen-Dalsgaard *et al.*, 2018; Vihtakari *et al.*, 2018), and may vary considerably within small areas (Sankaranarayanan, 2007). An absence of synchrony on a regional scale may

benefit species via a portfolio effect, such that if extreme weather negatively impacts a population at one stage of the breeding season, a population at a different stage of reproduction may experience less severe effects, thereby promoting population stability at higher aggregate levels such as groupings of species at the regional or meta-population level (Schindler *et al.*, 2015).

We detected correlated responses across populations of only one species, the black-legged kittiwake (hypothesis 3), with timing of breeding in populations from both sides of the Atlantic and spanning almost all of the breeding range tending to vary in tandem by ±6 days. In the North Atlantic, the majority of kittiwakes from most populations winter in the Labrador Sea, and it is likely that they experience similar conditions during this period (Frederiksen *et al.*, 2012; Bogdanova *et al.*, 2017). It is plausible that water temperature over the winter, via its effect on resources, may determine when kittiwakes return to waters around their colonies, with carry-over effects on timing of breeding. Although there was high among-year covariance in laying dates of kittiwakes across breeding sites, this only explained an average of 27.2% of the total among-year variance experienced by each population (min. = 12.11% [Hornøya], max. = 79.6% [Prince Leopold Island]), and correlations in lay date decreased with distance between sites. As kittiwakes are poor divers and hence restricted to foraging on the water's surface, they may be more responsive to local conditions than other species that can dive (Furness & Tasker, 2000). It is evident that kittiwakes may therefore be sensitive to environmental conditions across multiple spatial scales.

With the exception of the black-legged kittiwake, we found no shared variance across populations of the same species, which implies that they do not respond similarly to a spatially consistent driver. We found that residual variance for European shags (i.e. between-year variance in lay date within a population, after all other terms

have been taken into account) greatly exceeded the levels estimated for other species in the analysis (Table 3.3). European shags have a shorter-distance migration strategy than the other study species (Grist *et al.*, 2014), so may be more sensitive to local conditions near the colony, such as abundance of forage fish (Lorentsen *et al.*, 2015a) and have an unusually high capacity to adjust laying dates accordingly. While auk populations in our analysis do remain in the North Atlantic over winter and spring, they migrate to a variety of different areas (Frederiksen *et al.*, 2016; Fayet *et al.*, 2017), suggesting that the conditions driving auk phenology are unlikely to be consistent for all populations. Finally, the tern species included in this analysis (common, roseate and Arctic) are all long-distance migrants, and individuals from the same or different breeding sites may take alternative migration routes, at different times, and to different destinations (Egevang *et al.*, 2010; Mostello *et al.*, 2014; Becker *et al.*, 2016; Nisbet *et al.*, 2017), potentially experiencing different conditions. Further research comparing laying dates of tracked individuals known to have similar migration strategies would therefore elucidate the extent to which phenology covaries between individuals within and across colonies (Grecian *et al.*, 2016).

We found that laying dates of the sympatric populations positively covaried at two sites (hypothesis 4), and at five additional sites the posterior distribution was somewhat removed from 0 (Figure 3.5b), suggesting that local effects do play a role in driving phenology. Positive covariance in phenology may be driven by several factors, such as local habitat or weather conditions (Porlier *et al.*, 2012); abundance and phenology of prey (Frederiksen *et al.*, 2005); inter- (Schoener, 1974) and intra-specific competition for food (Lewis *et al.*, 2001a), or a combination of effects. Small-scale physical features potentially cause subtle differences in conditions at each site despite site proximity, resulting in the observed differences in covariance between sites. This may be beneficial from a resilience perspective, as phenological

asynchrony observed across wide distances may allow meta-populations (such as those formed by many North Atlantic seabird species (Bogdanova *et al.*, 2017; Fayet *et al.*, 2017)) to increase resilience in the face of wide-scale perturbations (Schindler *et al.*, 2015) expected under future climate scenarios (Stocker *et al.*, 2013).

Although our analysis included datasets of eight or more years in duration, occasionally the time series overlap was low, reducing our ability to infer precise covariances. As our simulations show (Appendix A.4), longer datasets would rectify this and we hope this approach will be repeated in the future on seabirds and other taxa. While our model structure did not allow for negative covariance between phenological time series, when we compared pairwise estimates of phenological correlations expected under our model to those obtained from raw data we found a good correspondence between the two (Figure A.4). We therefore suggest that the observed negative covariances are consistent with what one would expect to observe by chance when sample sizes are small and the true covariance is close to zero. Finally, our analysis considered the effects of conditions at the breeding and main wintering grounds, but did not take into account pre-breeding, post-breeding, staging and migration routes. More detailed tracking information would allow future analyses to take this into account.

Phenology is widely used as a measure of species' response to environmental change, yet for higher trophic level species particularly those that are highly mobile, its specific drivers are often poorly understood. We estimated covariance of average lay date across multiple populations of seabirds, to identify the scale at which drivers of phenology operate in this group of highly mobile top predators. For many populations, we identified covariance at the breeding site level, highlighting the importance of local conditions in driving phenology for some species in this taxonomic group. However, we conclude that the near absence of regional covariance, with the

exception of black-legged kittiwakes, may allow for increased resilience at the meta-population scale via portfolio effects, should broad-scale perturbations cause conditions to deteriorate rapidly across a large region. Further research combining individual tracking and phenology data could reveal drivers operating at additional spatial, temporal and biological scales, for example conditions experienced by individuals or populations on migration routes, stop-overs, or during autumn or spring periods. Identifying the multiple scales at which phenology is driven will allow us to further understand how organisms respond to fluctuating conditions, and how they may continue to do so in the future.

4. No evidence for fitness signatures of increasing trophic mismatch over 30 years in a population of European shag *Phalacrocorax aristotelis*

Katharine Keogan, Sue Lewis, Richard J Howells, Sarah Wanless, Sarah Burthe, Richard A Phillips, Ally B Phillimore*, Francis Daunt*

*joint last authors

4.1 Abstract

Timing of reproduction is changing at different rates across trophic levels, potentially resulting in mismatch between consumers and their resources. Trophic mismatch is widely considered to have negative impacts on average productivity of consumers, because the timing of breeding of the population as a whole is out of synchrony with resource availability. Furthermore, selection on timing of breeding is predicted to strengthen with increasing mismatch. Quantifying changes in average productivity and strength of selection therefore constitutes a test of the presence of increasing mismatch. However, the limited number of long-term data sets on breeding timing and success from wild animal populations has meant that this analysis has rarely been undertaken. Here, using a 30 year individual-level data set of breeding phenology and success from a population of European shags (*Phalacrocorax aristotelis*), we quantified changes in average breeding success and strength of selection on timing of breeding over time and in relation to sea surface temperature (SST) and diet composition to test for fitness signatures of trophic mismatch both over time and in relation to climate. Annual average (population level) breeding success was negatively correlated with average lay date, such that years in which breeding was earlier were more successful, yet showed no trend over time, with increasing SST or in relation to the proportion of the principal prey in the diet, as would be expected if

mismatch with prey was increasing. At the individual level, we found evidence for stabilising selection on earlier timing of breeding, yet this slope did not become steeper over time, with SST or in relation to the slope of seasonal shift in diet from principal to secondary prey. Our results indicate that average performance is affected by population timing of breeding, and there is stabilising selection on earlier laying at the individual level, but no fitness signatures of climate-induced trophic mismatch in this population. These results highlight the importance of testing population and individual level variation in productivity in relation to breeding phenology when assessing evidence of mismatch.

4.2 Introduction

In recent decades, surface temperatures around the globe have risen (Stocker *et al.*, 2013), causing the timing of seasonally recurring life history events, such as reproduction, to change at different rates across trophic levels (Visser & Both, 2005; Thackeray *et al.*, 2010). The plastic phenological responses of higher trophic level organisms to changing temperatures often appear to be weaker than those of organisms lower in the food web (Thackeray *et al.*, 2016, Keogan *et al*., 2018). Studies have shown that this difference in responsiveness could potentially lead to trophic mismatch, whereby the timing of peak demands of consumers and availability of their resources are asynchronous (Figure 4.1a, 4.1b; Visser *et al.*, 2004; Thackeray *et al.*, 2016). Trophic mismatch is predicted to have negative consequences on fitness, with important implications for population dynamics (Miller-Rushing *et al.*, 2010; Reed *et al.*, 2013a). A negative effect on average demography may arise because the population as a whole is less well matched with the availability of resources (Durant *et al.*, 2007; Thackeray *et al.*, 2010, 2016; Reed *et al.*, 2013a; McLean *et al.*, 2016). Furthermore, selection on timing of breeding is predicted to strengthen with increasing trophic mismatch as the among-individual variation in fitness increases (Reed *et al.*,

2013a). The extent to which climate-mediated trophic mismatch has led to negative effects on population demography and stronger selection on timing of breeding is a central question in understanding the population and evolutionary consequences of trophic mismatch (Reed *et al.*, 2013a; McLean *et al.*, 2016).

A widespread finding in higher trophic-level species breeding in seasonal environments is a negative relationship between timing of breeding and reproductive success (Clutton-Brock, 1988; Newton, 1989). This relationship holds both within and among years, whereby years in which breeding is earlier are more successful than later years (Figure 4.1c), and individuals breeding relatively early within a year are more successful than those breeding relatively late (Figure 4.1d; Verhulst & Nilsson, 2008a). The extent of matching of peak demands and availability of resources may underpin these relationships. In earlier years, and among earlier breeders within years, timing may be more closely matched with resource availability, with positive consequences on fitness (Visser *et al.*, 2004). Crucially, if trophic mismatch is increasing under climate change, average breeding success is expected to decline and directional selection on earlier breeding to strengthen (Reed et al. 2013). However, there is a critical shortage of long-term data sets on phenology of breeding of higher trophic level species and timing of availability of their prey, hampering the ability to test the consequences of trophic mismatch on fitness directly.

An alternative approach is to quantify the change in average breeding success and strength and direction of selection on timing of breeding in relation to proxies of change in trophic mismatch. In seasonally breeding species, two measures that are often used in this context because they are important indicators of change in trophic mismatch are time (i.e. years) and annual average temperature (Keogan et al. 2018). Diet proportion offers an additional proxy of trophic mismatch between consumers and their prey (Watanuki *et al.*, 2009), since predators are in general constrained to

feed themselves or developing young on prey types that are available during the breeding season. Consumer diet composition is therefore determined by the timing of key life history events in the annual cycle of the two trophic levels (Miller-Rushing *et al.*, 2010). Annual mean diet composition can therefore be considered in the same way as year and temperature variables (Figure 4.1e). Furthermore, diet data may be available spanning a range of dates within seasons, and the opportunity then arises to quantify within-season shifts between different prey types (Figure 4.1f). Signatures in the fitness data consistent with trophic mismatch at the population level include average breeding success becoming lower over time, with increasing temperature and when the principal average diet proportions is reduced (Figure 4.1g). At the individual level, fitness signatures consistent with mismatch include stronger directional selection on timing of breeding is predicted over time, with increasing temperature and when average change in diet within a season is steeper (Figure 4.1h).

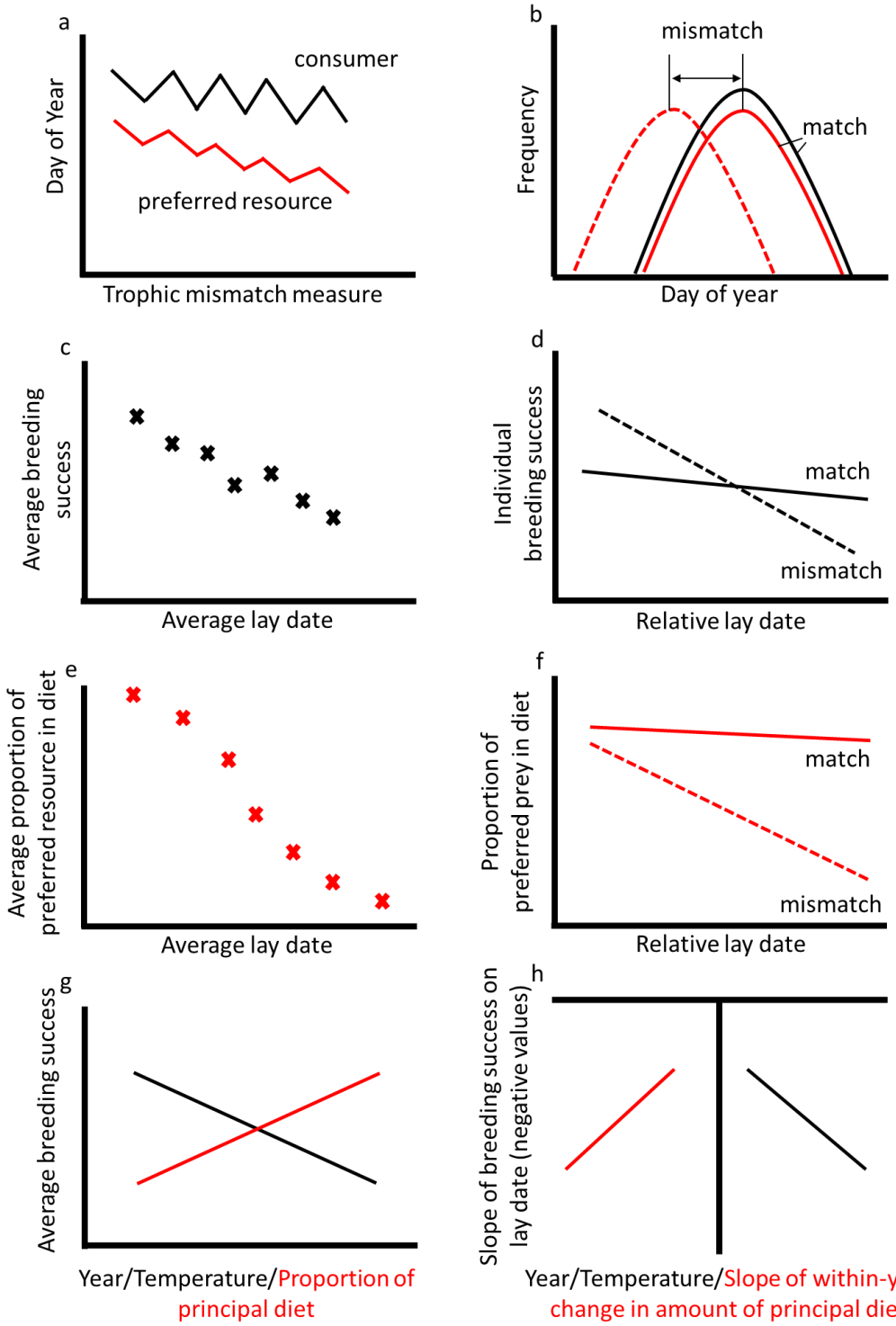


Figure 4.1. Schematic of requirements for mismatch (a, b), consequences for breeding success (c,d) and diet (e,f) if mismatch is present and expectations if the impacts of mismatch are worsening (g,h) in relation to year, temperature or proportion of principal diet. Left-hand plots (a, c, e, and g) show expected outcomes at the population (between-year) level and right-hand plots (b, d, f, and h) show expected outcomes at the individual (within-year) level. Red lines are representative of prey, and black lines are representative of predator, except in (g and h) where red lines are indicative of the relationship between prey and predator. In (b,d, and f), solid lines indicate matched years, and dashed lines indicate mismatched years. In (c) and (e), x's are used to represent average breeding success rather than a

continuous line because these are detrended for year. In (h), a more negative y-value of slope represents stronger selection on laying date.

The Match-Mismatch Hypothesis was first proposed to explain changes in marine fisheries productivity (Cushing, 1990). However current research has mainly focussed on its prevalence in terrestrial systems (Thackeray *et al.*, 2010). Our understanding of the effects of trophic decoupling on fitness in marine systems is therefore less well understood than in terrestrial systems (Richardson & Poloczanska, 2008; Thackeray *et al.*, 2010). In the marine environment, rising sea surface temperatures (SSTs) have been correlated with advances in the timing of plankton blooms (Edwards & Richardson, 2004) and fish spawning events (Asch, 2015), and there is increasing evidence that trophic mismatch may be occurring (Burthe *et al.*, 2012; Régnier *et al.*, 2017). On average the phenology of higher trophic level marine groups such as seabirds has been unresponsive over time (Poloczanska *et al.*, 2013; Keogan *et al.*, 2018), and to rising SST (Keogan *et al.*, 2018). In general the rate of change in seabirds is much slower than that of fish or plankton (Poloczanska *et al.*, 2013), potentially leaving many seabird species at risk of asynchrony with their food resources.

In this paper, we used a long-term data set of a marine higher predator, the European shag *Phalacrocorax aristotelis*, to test whether average breeding success and the strength of selection on timing of breeding has changed in relation to time, temperature and diet composition. European shags breed and overwinter throughout temperate latitudes in the North Atlantic, and are highly variable in their annual mean phenology within and among breeding seasons. The study population on the Isle of May National Nature Reserve, south-east Scotland, has advanced its laying date by approximately 1 day per year and by about 5 days per every 1°C rise in SST (Chapter 2). The principal prey of shags is adult ('1+ group') lesser sandeels *Ammodytes*

marinus, but they show a seasonal shift in diet to the second most common prey, juvenile ('0 group') sandeels (Howells *et al.*, 2017). Furthermore, there is marked variation in the structure of the seasonal shift among years (Howells et al. 2017). This dataset therefore presents an excellent opportunity to test whether average breeding success and selection on timing and breeding success has changed in relation to year, temperature and average and within-season changes in diet.

This study has two principal aims. First, we estimate the effect of lay date on breeding success at the population-level (between year means) and individual-level (within year slopes), addressing average population breeding success and the presence, direction and form of selection respectively. Second, we test whether annual mean breeding success and strength of selection on timing of breeding have changed in relation to three proxies of mismatch. If mismatch has increased over time, the fitness signatures that would be consistent with this trend are that mean annual breeding success will decline and strength of selection on relative lay date within a season will increase over the course of the study. Similarly, the predicted fitness signatures consistent with mismatch having increased with rising temperatures are that mean annual breeding success will correlate negatively with Sea Surface Temperature (hereafter SST) and strength of selection should increase as SST rises. Finally, using diet composition as a proxy for resource availability we test for fitness signatures consistent with mismatch by estimating the between-year and within-year correlations between breeding success and the proportion of sandeels in the diet samples. If European shags are becoming mismatched with this resource, the fitness signatures we expect to observe are a positive correlation between resource and breeding success both between and within years. However, it is important to consider that there may be alternative mechanisms driving these patterns, and the presence

or absence of such fitness signatures do not demonstrate unequivocally the presence or absence of trophic mismatch.

4.3 Methods

4.3.1 Data Collection

Breeding phenology and success data Breeding phenology and success were recorded for a sample of nests every year between 1987 and 2016 (range = 35 - 266; no data available for 1993 and 2003). Nests were monitored in 18 plots spread throughout the colony. Nests were checked every seven days from before laying started until after the last chick had fledged. Lay date was taken to be three days prior to the first date that incubation was recorded, unless the number of eggs in the nest were counted on the first occasion that laying was confirmed, in which case lay date could be estimated with greater accuracy based on standard laying intervals of three days in this species (Potts *et al.*, 1980). While the maximum error in lay date is an overestimate by six days (for a nest where laying occurred just after the previous visit), variation in accuracy across nests should be consistent within and between years, and measurement error variance is therefore much smaller than the within-year variance on lay date. In cases where lay date could not be determined by start of incubation or number of eggs found in the nest, it was back-calculated from chick wing length or hatch date based on previously derived relationships between wing length and hatch date (Daunt, 2000) and an incubation period of 36 days (Potts *et al.*, 1980). Breeding success was recorded as the number of chicks fledged per nest (range 0 - 4). In cases where a breeding pair failed and laid a second clutch, the new attempt was not included in the core analyses because lay date was not independent of the laying and failure date of the first clutch. However, because overall breeding success of a nest (i.e. from all breeding attempts) may be impacted by the timing of the first breeding attempt, an additional analysis was included to test the effect of lay date of the first clutch on overall breeding success. We found no qualitative difference in the

results between these two models and therefore our subsequent models only included first laying attempts (Table B1.a).

Population counts of breeding pairs were available for each year.

Inshore temperature data Following Frederiksen et al. (2004), sea surface temperature (SST) data were extracted for February and March in each year from http://www.bsh.de, for an area surrounding the Isle of May that overlapped shag foraging distribution in the breeding season (Bogdanova *et al.*, 2014; bounded by ca. 56° 0' to 56° 4' N, and 2° 7' to 2° 3' W). We averaged the monthly records to obtain a mean late winter temperature for each year. This period was selected because late winter temperature is a key driver of sandeel somatic investment and recruitment (the number of sandeels that successfully transition from 0 group to 1 group) (Wright & Bailey, 1996; Arnott & Ruxton, 2002; Van Deurs *et al.*, 2009; Wright *et al.*, 2017a, 2017b).

Diet Chicks and adult shags sometimes regurgitate food during routine fieldwork, and the biomass proportions of each prey type can be estimated using standardised methods (full details in Harris & Wanless, 1991; Howells *et al.*, 2017). Regurgitates were collected on the Isle of May during the chick-rearing period (April – July) between 1985 and 2014 (n = 863; median 29 samples per year; range 4-69; Howells et al. 2017). Collection dates showed a strong positive correlation with the timing of the shag breeding season (median collection date versus median laying date, $r = 0.86$, $n = 25$ years, 95% CI = 0.70, 0.94, $p < 0.0001$). The two most important diet types are 1+ group sandeels (70% of biomass) and 0 group sandeels (12%; Howells et al. 2017). 1+ group sandeels are replaced by 0 group sandeels over the course of the breeding season, from a predicted relative proportion of 1.00 in April to 0.24 in August (Howells et al. 2017). This shift is in line with the seasonal life history of the two age classes of sandeels. 1+ group sandeels are active in the water column in the early

spring (April/May), before burying in sandy sediments for the remainder of the year, next entering the water column to spawn in midwinter. In contrast, 0 group sandeels become available from June onwards following metamorphosis from the larval stage (Winslade, 1974; Wright & Bailey, 1996; Boulcott & Wright, 2008; Régnier *et al.*, 2018). This seasonal diet shift has been recorded in other seabirds breeding on the Isle of May (Lewis *et al.*, 2001b; Wanless *et al.*, 2004; Daunt *et al.*, 2008), and so it would appear that the seabird community is responding to the changes in availability of different sandeel age classes over the course of the breeding season. 1+ group sandeels are markedly larger than 0 group sandeels, and have a higher energy density (Wanless *et al.*, 2004; Harris *et al.*, 2008). We therefore predict that average breeding success will be higher when the proportion of 1+ group relative to 0 group sandeels in the diet is higher. Further, there is marked variation among years in the relative proportions of 1+ to 0 group sandeels in the diet (Howells et al 2017), which is likely to largely be determined by the timing and strength of the within-season change in proportion of the two age classes of sandeel relative to the timing of the shag breeding season. We quantify this link here, and predict that years in which the slope of change in proportion of the two age classes is steeper show stronger directional selection on timing of breeding. Crucially, since the among and within-year diet proportion measures are likely to be strongly determined by relative timing of key life history events among shags, sandeels and their prey, they are useful proxies of trophic mismatch (Wright & Bailey, 1996; Frederiksen *et al.*, 2011; Burthe *et al.*, 2012; Régnier *et al.*, 2017; Wright *et al.*, 2017a, 2017b).

4.3.2 Statistical Analysis

Environmental variation and temporal trends We estimated the linear slope of SST change and phenological trend over time to assess overall patterns within the study

system. We used Generalised Least Squares (GLS) in nlme (Pinheiro & Bates, 2000), and fitted an autoregressive model of order 1, AR(1) (Box *et al.*, 1990), to take into account temporal autocorrelation in both time series'. Howells et al. (2017) have previously demonstrated that there is no trend in the proportion of 1+ group relative to 0 group over the study period.

Phenology and breeding success We used Generalised Linear Mixed-Effects Models (GLMMs) in a Bayesian framework in the *MCMCglmm* package (Hadfield, 2010) in R (v 3.2.2; R Core Team 2015) as a framework for examining the relationship between lay date and breeding success (Aim 1), which we defined as the proportion of chicks fledged to the maximum number of potential fledglings (following Burthe *et al.*, 2012), with the maximum brood size of shags taken to be four (Newell *et al.*, 2015) (4 nests out of 2746). Breeding success was recorded at each nest, and therefore the nest was the unit of measure, not the individual bird, since shags are long-lived biparental species and both males and females may move nests between breeding attempts (Barlow *et al.*, 2013). All nests were pooled in the analysis. We assumed binomial family errors because breeding success was under-dispersed as compared with the expectation under a Poisson process. However, see supplementary methods and results for outputs of the same models assuming Poisson family errors, which made no qualitative difference to the results. For all models coefficients are presented as the mean of the posterior distribution and uncertainty is presented as the 95% credible intervals (CIs).

We included three key fixed effects in all models (Table 4.1): i) annual mean lay date, which allowed us to test whether there was any linear decline in the average breeding success with later population mean laying date. ii) within-year centred relative lay date as a linear effect. The within year-centring removed the effect of between-year variation in laying date (van de Pol & Verhulst, 2006), and the slope of

breeding success on relative lay date estimates the direction and strength of selection on lay date within each year. iii) relative lay date as a quadratic effect, as finding evidence of a significant quadratic term and peak within the range of data is consistent with stabilising selection with an optimum that lies within the data range. We calculated the vertex of the quadratic curve as $-b/2a$, where b = linear slope and a = quadratic slope, to estimate the date within an average year when individual fitness (in relation to lay date) was maximised. Log transformed annual population size was included in the initial version of our core model, but produced no qualitative differences in the results and so was not included in the final version (see Table B1.b for summary statistics from this model).

We allowed the relative slopes to vary across years by fitting the regression lines as a random effect.

Equation 1 below corresponds to our core model.

$$z_{ij} = \hat{\mu} + \hat{\mu}_i + \hat{\beta}_A(\bar{x}_j) + \hat{\beta}_R(\bar{x}_{ij} - \bar{x}_j) + \beta_{R_i}(\bar{x}_{ij} - \bar{x}_j) + e_{ij} \quad \text{eq. 1}$$

Where z is the number of expected offspring in year i for population j. μ represents the overall grand mean and μ_i represents the deviation of the mean population breeding success in year i. $\bar{x}$ represents mean lay date, and $\hat{\beta}_A$ estimates the among year slope; $\hat{\beta}_R$ represents the average relative (within year) slope, and $\hat{\beta}_{R_i}$ represents the deviation of the relative slope in year i. e_{ij} represents the random intercept with fixed variance. Random effects were assumed to come from a normal distribution with mean = 0 and a variance that was estimated from the data.

We used the core model to estimate the slope of breeding success on lay date at the population (annual mean lay dates) and individual (within year relative lay dates) levels. We then considered three additional models that test the hypothesised effects

of mismatch on breeding success (Table 4.1), each of which build upon the core model (Aims 2 and 3).

1) Year model: we included year as a mean-centred continuous variable, and the interaction between year and relative lay date. If mismatch has increased over time we predict that population mean breeding success will decline and the within-year relative slope will become steeper (Figure 4.1g, h).

2a) SST model 1: we included SST in the current year, the interaction between SST and relative lay date, and year as a mean-centred continuous variable to detrend the analysis (Iler *et al.*, 2016). Timing of 0-group presence in the water column may be dependent on temperature in the current year (Eerkes-Medrano *et al.*, 2017; Wright *et al.*, 2017b). If mismatch is occurring, we predict that there should be a negative relationship between SST and shag annual population mean breeding success and the slope in relation to within-year relative lay date should be steeper in warmer years (Figure 4.1g, h).

2b) SST model 2: we included SST in the previous year (SST-1), the interaction between SST-1 and relative lay date, and current year as a mean-centred continuous variable to detrend the analysis (Iler *et al.*, 2016). High sandeel recruitment is strongly dependent on cool temperatures in the previous year (Boulcott & Wright, 2008; Van Deurs *et al.*, 2009), potentially influencing timing of key life history events of 1+ group sandeels in the current year. We therefore predict that if the previous spring was warm, shag population mean fitness in the current year should be reduced and the within-year slope of the relationship between breeding success and relative lay date should be steeper (Figure 4.1g, h).

3) Diet model: we used the proportion of 1+ group to 0 group sandeels in each diet sample to test for the seasonal shift in diet between the two age groups. Only samples

which contained sandeel prey were used in this model (n = 745). We tested whether years where the mean date of sample collection was later had lower than average proportions of 1+ to 0 group sandeels (Figure 4.1e) and lower average breeding success (Figure 4.1g). We then estimated the slope of the proportion of 1+ group relative to 0 group sandeels in the diet regressed on collection date in each year (Figure 4.1f) and tested whether there was be a positive relationship between the within-year slopes of 1+ to 0 group diet proportions and breeding success (Figure 4.1h). First we considered the diet in isolation and tested for between year and within year trends. We included two fixed effects: 1) annual mean date of sample collection, which allowed us to test whether there was any linear increase or decrease in the average proportion of 1+ group relative to 0 group sandeels in the diet in relation to the mean date of sample collection (as in Howells *et al.*, 2017). 2) relative date of sample collection, which was within-year centred (van de Pol & Wright, 2009) to remove the effect of between-year variation (van de Pol & Verhulst, 2006), as a linear effect which allowed us to consider the direction and magnitude of seasonal shifts in diet between the two age classes of sandeel within each year. The random effect allowed relative slopes to vary across years in the same way as eq. 1.

We tested whether the proportion of 1+ group relative to 0 group sandeels in the diet showed a linear trend towards increasing or decreasing over time by including mean-centred year as a continuous variable. To test whether the strength or direction of seasonal shifts in diet changed over time, we included the interaction between year (as above) and relative date of sample collection.

The response variable (proportion of 1+:0 group sandeels) was logit transformed, with 0.01 added to both the numerator and denominator of the logit function to avoid -∞ and ∞ values for proportions of 0 and 1 respectively (Collett, 2002; Warton & Hui, 2011). Random regression models allowed us to test whether the intra-

annual slopes of seasonal change in proportion of 1+ group relative to 0 group sandeels in the diet varied among years (Hadfield, 2010).

We considered shag breeding success and proportions of 1+ group relative to 0 group sandeels in the diet as a bivariate response, with a binomial family error for breeding success and Gaussian for sandeel diet.

The model terms were as outlined in the core shag and sandeel models with the following differences. First, the effect of within year timing was centred around the mean lay date of shags in each year for both shags and sandeels. Second, for each random term (the among-year variation in intercepts and the among-year variation in the relative timing slope) we also estimated covariance (σ) between shag (*Sh*) and sandeel (*Sa*). This gave a 4 x 4 matrix

$$\begin{bmatrix} {\sigma^2}_{M_{sh}} & \sigma_{M_{Sa}M_{Sh}} & \sigma_{R_{Sh}M_{Sh}} & \sigma_{R_{Sa}M_{Sh}} \\ \sigma_{M_{Sh}M_{Sa}} & {\sigma^2}_{M_{Sa}} & \sigma_{R_{Sh}M_{Sa}} & \sigma_{R_{Sa}M_{Sa}} \\ \sigma_{M_{Sh}R_{Sh}} & \sigma_{M_{Sa}R_{Sh}} & {\sigma^2}_{R_{Sh}} & \sigma_{R_{Sa}R_{Sh}} \\ \sigma_{M_{Sa}R_{Sh}} & \sigma_{M_{Sa}R_{Sa}} & \sigma_{R_{Sa}R_{Sh}} & {\sigma^2}_{R_{Sa}} \end{bmatrix}$$

where M_{Sh} represents mean shag breeding success, M_{Sa} represents mean sandeel proportion, R_{Sh} represents relative shag breeding success, and R_{Sa} represents relative sandeel proportion.

If mismatch with sandeels impacts at the population level we predict $\sigma_{M_{Sa}M_{Sh}} > 0$**,** and if it impacts at the individual (within year level) we predict $\sigma_{R_{Sa}R_{Sh}} > 0$.

Model structure

All models were run for 100,000 iterations to allow effective sample sizes for focal parameters to reach >1000, sampling every 10th iteration and with the first 10,000 iterations discarded as burn-in. Parameter-expanded priors were used for all random terms except the residual variance, which was fixed at 1. Plots of the mean and

variance of the posterior distribution were examined to assess autocorrelation in the posterior samples. Statistical significance of fixed effects was inferred where 95% credible intervals did not span zero.

Table 4.1. The effects included in each model in this analysis

Model name	Response variable	Fixed effects	Random effects	MCMCglmm syntax
Core	chicks fledged:all potential chicks	mean lay date; relative lay date; relative lay date (quadratic)	random regression allowing intercept and relative lay date slope slope to (co)vary across years	`MCMCglmm(cbind(chicks.fledged,failedeggs)~ mean.lay+relative+I(relative^2), random=~us(1+relative):year, prior = prior_pa, family="multinomial2", data=shags, pr=TRUE, nitt = 100000, burnin = 10000)`
Core (with population size)	chicks fledged:all potential chicks	mean lay date; relative lay date; relative lay date (quadratic); population size (log)	random regression allowing intercept and relative lay date slope slope to (co)vary across years	`MCMCglmm(cbind(chicks.fledged,failedeggs)~ mean.lay+relative+I(relative^2)+log(pop.size), random=~us(1+relative):year, prior = prior_pa, family="multinomial2", data=shags, pr=TRUE, nitt = 100000, burnin = 10000)`
Year	chicks fledged:all potential chicks	mean lay date; relative lay date; relative lay date (quadratic); year (centred & continuous); relative:year (centred & continuous)	random regression allowing intercept and relative lay date slope slope to (co)vary across years	`MCMCglmm(cbind(chicks.fledged,failedeggs)~ mean.lay+relative+yearcentre +relative:yearcentre +I(relative^2), random=~us(1+relative):year, prior = prior_pa, family="multinomial2", data=shags, nitt = 100000, burnin = 10000, pr = TRUE)`

SST-1 / SST	chicks fledged:all potential chicks	mean lay date; relative lay date; relative lay date (quadratic); year (centred and continuous) Inshore mean; relative:Inshore mean	random regression allowing intercept and relative lay date slope slope to (co)vary across years	`MCMCglmm(cbind(chicks.fledged,failedeggs)~ mean.lay+relative+yearcentre+Inshore.mean.present+relative:Inshore.mean.present+I(relative^2), random=~us(1+relative):year, prior = prior_pa, family="multinomial2", data=shags, nitt = 100000, burnin = 10000, pr = TRUE)`
Sandeel core	Proportion 1+ : 0 group sandeels (logit transformed)	mean sample date; relative sample date	random regression allowing intercept and linear slope to (co)vary across years	`MCMCglmm(logit~meandate+datediff, random=~us(1+datediff):Year, prior = prior_pa, data=shagdiet, pr=TRUE, nitt = 100000, burnin = 10000)`
Sandeel year	Proportion 1+ : 0 group sandeels (logit transformed)	year (centred & continuous); relative year (centred & continuous)	random regression allowing intercept and linear slope to (co)vary across years	`MCMCglmm(logit~meandate+datediff+yearcentre+yearcentre:datediff, random=~us(1+datediff):Year, prior = prior_pa, data=shagdiet, pr=TRUE, nitt = 100000, burnin = 10000)`
Bivariate core	chicks fledged:all potential chicks *and* Proportion 1+ : 0 group sandeels (logit transformed)	Mean lay date (for both); relative lay date (for both); relative lay date (quadratic) (breeding success only)	random regression allowing intercept and slope of both breeding success and sandeel diet to (co)vary across years	`MCMCglmm(cbind(chicks.fledged,failedeggs, sandeelprop)~trait-1+trait:Mean.Lay+trait:relative+at.level(trait,1):I(relative^2), random=~us(trait-1+trait:relative):year,rcov=~idh(trait):units, family=c("multinomial2","gaussian"),data=ov`

```
eralldata,prior=prior_pa, pr = TRUE,
nitt=100000,burnin=10000)
```

4.4 Results

4.4.1 Temporal trends

SST became warmer between 1987 and 2016 (mean temperature = 5.94°C , min – max = 5.08°C – 6.78°C, *b* = 0.02°C yr^{-1}, SE = 0.007, *p* = 0.0085, Phi = 0.27, Figure 4.2a). However, upon visual inspection of Figure 4.2a, it is important to consider that this trend may be driven by the lack of cooler SST in more recent years rather than by increasing temperatures overall. Mean lay date became earlier during the study period (mean lay date (day of year) = 127, min – max = day 71 – 217, *b* = -0.94 days yr^{-1}, SE = 0.33, *p* = 0.0087, Phi = 0.15, Figure 4.2b).

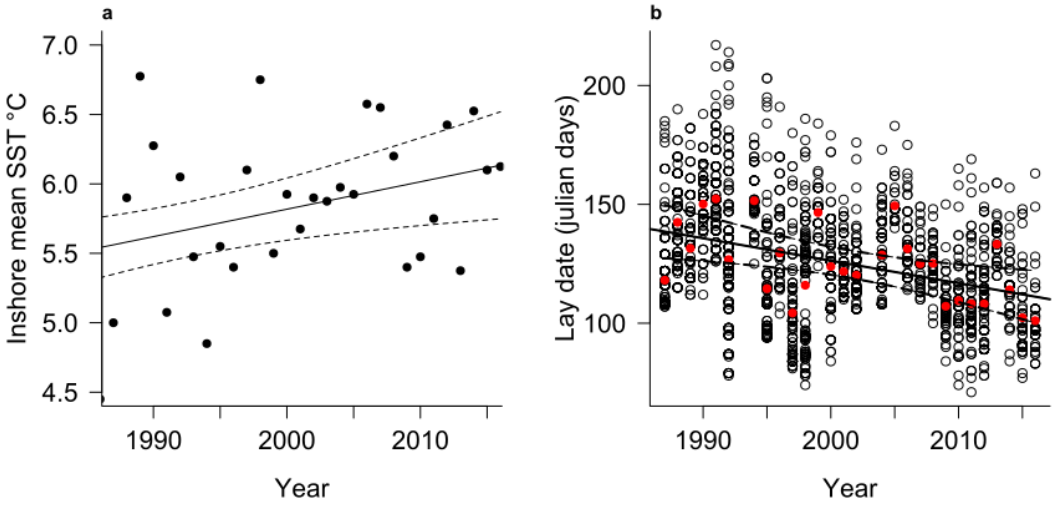


Figure 4.2. Environmental and demographic inter-annual variation in a) Sea Surface Temperature, and b) lay date. Red dots depict annual mean lay dates. Black lines indicate average trends in SST (a) and lay date (b) over time. Ordinal day refers to number of days after Jan 1st, allowing for leap years. Dotted lines represent 95% confidence intervals around the slope estimate.

4.4.2 Aim 1: Phenology and breeding success

We found that between years, mean breeding success declined significantly with mean lay date (Slope = -0.035, 95% Credible Interval [CI]: -0.053, -0.017. Figure 4.3a.

Table B.1). Within years, there was a negative relationship between relative lay date and breeding success (Slope = -0.026, 95% CI: -0.034, -0.019. Figure 4.3b. Table B.1) and a significant negative quadratic term (Slope = -0.0007, 95% CI: -0.0009, -0.0005, Figure 4.3b. Table B.1), such that breeding success was highest in birds breeding early in the year but not the earliest. Within a year, fitness was estimated to be maximised for birds that lay 19.34 (95% CI: -29.19, -11.31) days prior to the annual mean laying date. However, overall we found no significant difference in fitness slope between years (core model variance = 0.0001, 95% CI: 0.0000, 0.0003. Figure 4.3b), indicating that the shape and strength of directional selection remains similar across years.

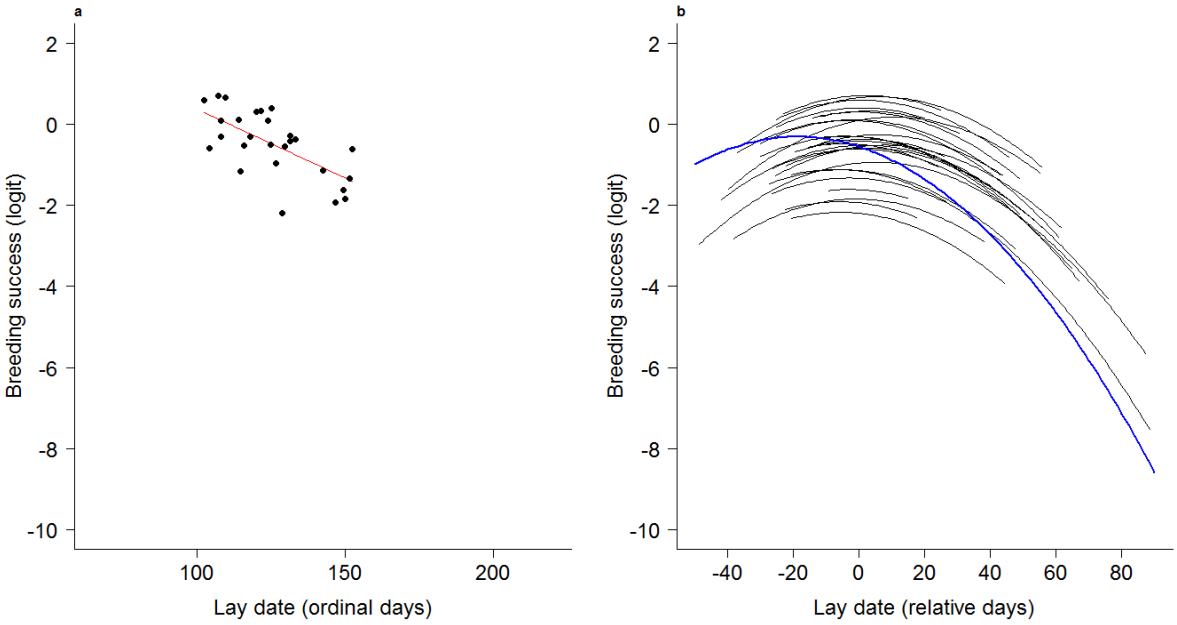


Figure 4.3. The relationship between lay date and breeding success (logit scale) (a) at the between-year level and (b) at the within year level. Points in (a) are mean values from the data, red line corresponds to the slope across annual means estimated from the core model and estimates the change in mean fitness. Ordinal day refers to number of days after Jan 1st, allowing for leap years. Black lines in (b) correspond to best linear unbiased predictors of the within-year slopes estimated in different years and the blue line is the average within-year slope, with all coefficients taken from the core model. See Fig S1 for a projection of these slopes onto the proportion scale.

4.4.3 Aim 2: Mismatch and Breeding Success

The time model (Table B.2) revealed no significant temporal decline in annual mean breeding success (slope = 0.029, 95% CI = -0.008, 0.064. Figure 4.4a), nor steepening of the within-year slope (relative slope:year interaction = 0.00007, 95% CI: -0.0008, 0.0009. Figure 4.4b). SST in the previous year had no effect on population-level fitness (breeding success $^{o}C^{-1}$ = -0.203, 95% CI = -0.750, 0.328. Figure B.2, Table B.3), nor on the strength of stabilising selection (change in strength of selection $^{o}C^{-1}$ = -0.005, 95% CI: -0.009, 0.019). Spring SST in the current year had no significant effect on population-level fitness (population breeding success $^{o}C^{-1}$ = 0.482, 95% CI = -0.074, 1.053. Figure 4.4c, Table B.4), nor on the strength of stabilising selection (change in strength of selection $^{o}C^{-1}$ = -0.007, 95% CI: -0.027, 0.013. Figure 4.4d). As such, warmer years neither impacted population average breeding success, nor the relative fitness of individuals breeding later or earlier than the average.

The relative proportion of 1+ to 0 group sandeels in the diet varied significantly among years (variance = 6.86, 95% CI = 3.35, 11.24, Figure 4.5a, Tables B.6, B.7). However, this proportion was not correlated with annual mean date of sample collection as predicted if mismatch were present (slope = -0.026, 95% CI = -0.084, 0.030, Figures 4.1e, 4.5a, Tables B.5, B.6). Within a year, the proportion of 1+ group relative to 0 group sandeels in the diet declined significantly throughout the season (relative slope = -0.097, 95% CI = -0.14, -0.049, Figure 4.5b, Tables B.5, B.6), and the within-year slope varied significantly among years (variance between slopes = 0.016, 95% CI = 0.0076, 0.027, Figure 4.5b, Tables B.5, B.6). In an expanded model that included year as a continuous fixed effect and the interaction between year and relative timing, there was no change in the proportion of 1+ group relative to 0 group sandeels across years (slope = 0.007, 95% CI = -0.121, 0.147, Tables B.5, B.6) or change in the within-year

slope over time (interaction coefficient = 0.004, 95% CI = -0.002, 0.009, Tables B.5, B.6).

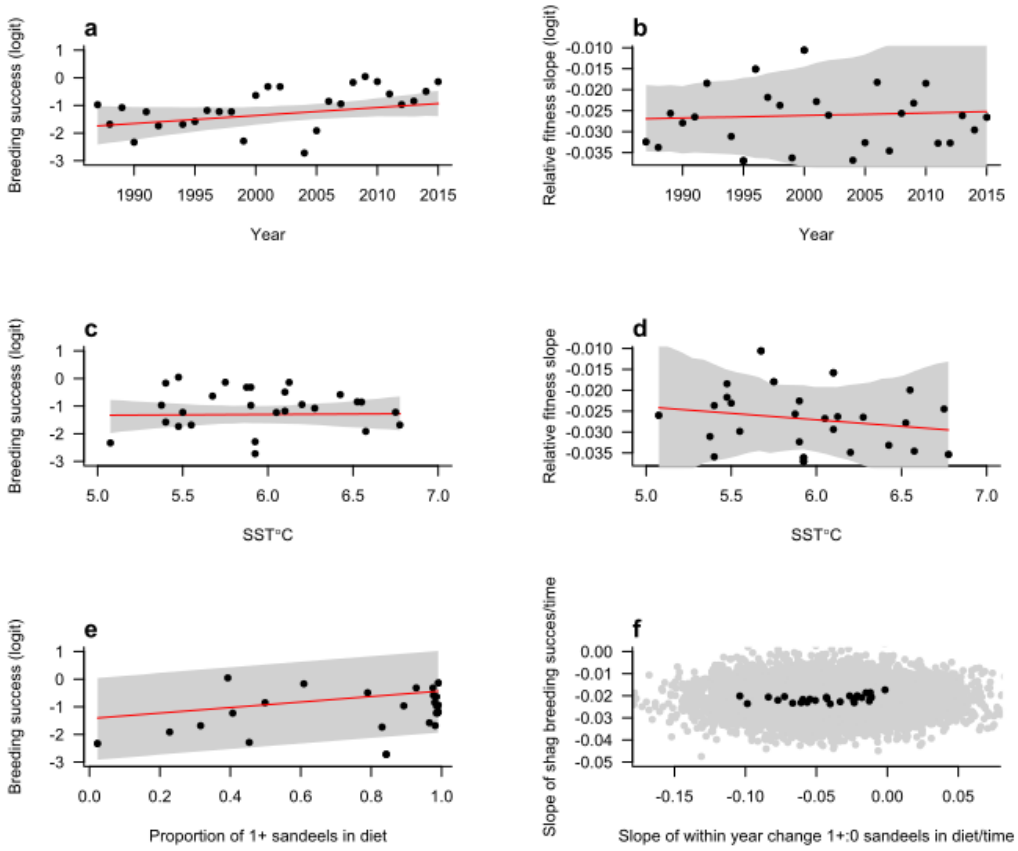


Figure 4.4. The effect of time (a), SST in the current year (c), and increased proportion of principal prey in diet (e) on breeding success (logit transformed) at the population level. Changes in strength of selection over time (b), with SST in the current year (d) and with the within year change in the proportion of principal diet (f). Red lines indicate average response and grey areas represent 95% credible intervals. In (f), each point (black) represents the mean value from the posterior distribution in a given year. Grey points represent the full posterior distribution of covariance between slopes for each year, taken from the bivariate model for each measure of diet change.

There was no evidence for positive covariance between mean shag breeding success and mean proportion of 1+ group relative to 0 group sandeels in the diet ($\sigma_{M_{Sa}M_{Sh}}$ = -0.078, 95% CI = -0.561, 0.297. Figure 4.4e, Table B.7), nor did the slopes of relative breeding success and relative proportion of the two age classes of sandeels in diet

covary positively among years ($\sigma_{R_{Sh}R_{Sa}}$ = 0.00004, 95% CI = -0.0002, 0.0003, Figure 4.4f, Table B.7).

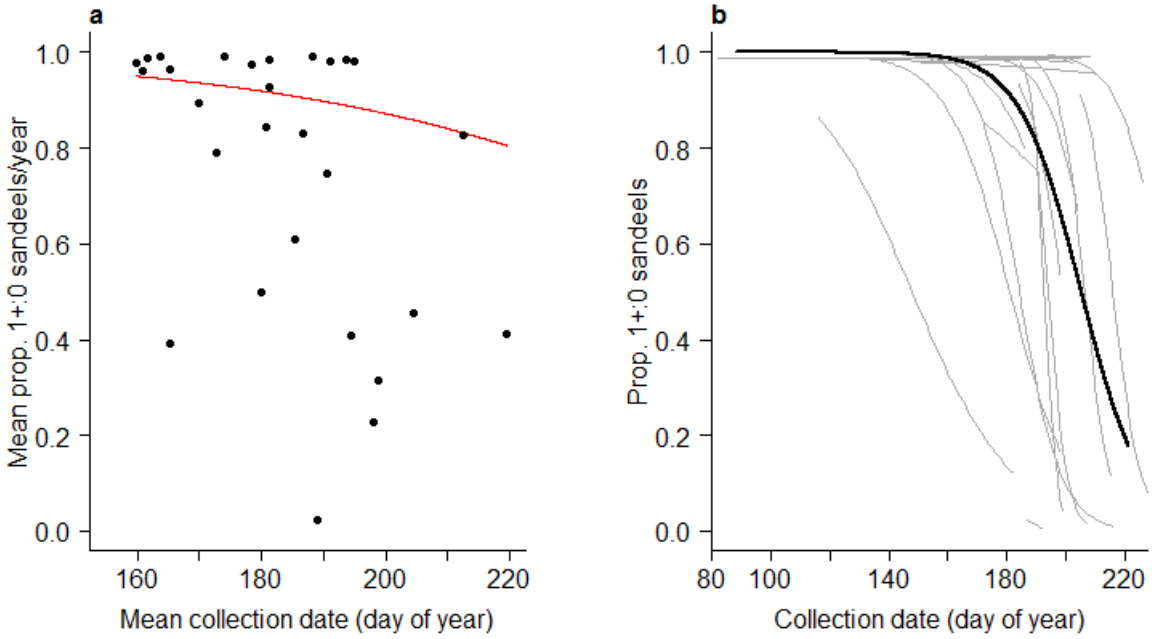


Figure 4.5. Between-year (a) and within-year (b) proportions of sandeels in the diet of shags during the chick-rearing period. **a)** Each point represents a yearly mean of the proportion of sandeels (1+ group to 0 group) in the diet and mean date of sample collection in that year (days after 1 Jan). The red line is the estimate from the core model of the change in diet proportion with mean lay date, without controlling for the effect of year, and back-transformed from the logit scale. **b)** Within-year changes in diet proportions for each year (grey lines) and the average within-year slope across all years (black line), all from the core model without controlling for year.

4.5 Discussion

We examined the effect of lay date on breeding success in a population of European shag, at both the population and individual levels. We found clear fitness benefits at the population level from breeding earlier in the year, and evidence that selection favours earlier (but not earliest) lay date. We then tested whether annual mean breeding success and strength of selection on timing of breeding have changed in relation to three proxies of mismatch (time, temperature and principal prey proportion in the diet). There was no trend towards decreasing population mean fitness over time, in warmer years, or in years where 1+ group sandeels formed a greater proportion of the diet. Moreover, strength of selection did not vary among years and showed no trend over time or with SST, and did not covary with the slope of the proportion of 1+ group relative to 0 group sandeels in relation to date through the season. We therefore conclude that while timing of breeding appears to be inherently important for reproductive success, there are no fitness signatures consistent with climate-induced trophic mismatch in this population.

4.5.1 Effect of lay date on population-level breeding success

At the population level, we found timing of breeding to be a negative correlate of fitness, which corresponds with previous studies (Clutton-Brock, 1988; Newton, 1989; Frederiksen *et al.*, 2004). There may be several reasons for this. Shags lay fewer eggs in later years at some breeding sites (Sklinna and Røst, Norway; Lorentsen *et al.*, 2015), potentially as a result of lower body conditions due to poor weather over winter, or a lower abundance of principal prey in the pre-laying period. Shags may also lay fewer eggs later in the season as a response to a photoperiodic cue, although this remains to be tested. If this were the case, shags may lay fewer eggs to avoid the energetic costs of losing the brood. Alternatively, years in which breeding is later may experience poorer prevailing conditions (food availability, parasite loads, weather

conditions), resulting in greater losses during incubation or chick-rearing in comparison to early years (Frederiksen *et al.*, 2004; Daunt *et al.*, 2007). These patterns could arise because individuals are constrained by these conditions, and breeding success may be as good as can be achieved, or they may show restraint to safeguard future reproductive potential (Williams, 1996; Frederiksen *et al.*, 2004).

4.5.2 Effect of lay date on individual-level breeding success (selection)

At the individual level, those breeding towards the end of the season were less successful than earlier conspecifics in all years, a result echoed in many other studies of breeding phenology (Verhulst & Nilsson, 2008; Sorensen *et al.*, 2009; Ramirez *et al.*, 2016; Smiley & Emmerson, 2016). The observed relationship was consistent with stabilising selection around an optimum laying date that is around 19 days earlier than the annual population mean. This means that in all years, it was disadvantageous to be among the earliest breeders, with fitness reaching a peak before declining again for the remainder of the season. The observed non-linear trend in breeding success may have been due to the energetic consequences of breeding very early within a year, before both environmental and body conditions have become good enough to ameliorate the fitness costs of producing eggs (Perrins, 1970; Stevenson & Bryant, 2000). Alternatively, very early breeders may be more vulnerable to factors such as increased predation risk or poor weather conditions. In this population, physiological constraints during winter affect breeding phenology, whereby those individuals with a lower foraging effort breed earlier (Daunt *et al.*, 2014). It may be the case that only these higher condition birds are able to initiate breeding when conditions are optimal (Verhulst & Nilsson, 2008), which may confer significant advantages since shags are income breeders that require a constant supply of food for successful breeding. In contrast, lower quality individuals may only reach a condition threshold later in the season (Daunt *et al.*, 1999). As such, rather than following cues to initiate breeding to

coincide with optimal extrinsic conditions, later breeders may be constrained to breed late by reduced body condition due to parasite burden (Hicks *et al.*, 2018), or the carry over effects of increased foraging effort over winter (Daunt *et al.*, 2014).

It is therefore important to consider that this within year relationship between timing and breeding success may arise via a third variable. In addition to individual quality (Verhulst & Nilsson, 2008), several other factors may influence the differential breeding success observed within a year. Later breeders are generally younger and less experienced, with intrinsically lower breeding success than more experienced breeders (Potts *et al.*, 1980; Daunt *et al.*, 2007). Additionally, nest site quality (Aebischer, 1985; Newell *et al.*, 2015), parasite burden (Reed *et al.*, 2008) and differential migration strategies within this population (Grist *et al.*, 2017) affect both the timing and reproductive success of breeding pairs (Potts *et al.*, 1980; Daunt *et al.*, 2007; Grist *et al.*, 2017). Although alternative mechanisms may underpin the link between breeding phenology and success in different years, the key result is that this relationship is remarkably consistent across years.

4.5.3 Evidence for trophic mismatch at the population level

Despite inter-annual variation in breeding success, there was no evidence that it had declined linearly over time, with SST, or with a reduced proportion of principal prey in the diet. The fact that breeding success has not declined over time suggests that perhaps sandeels are not adjusting their phenology at a faster rate over time than shags. This is contrary to what is suggested by multi-trophic-level phenological studies on other marine systems (Poloczanska *et al.*, 2013). Sandeel phenology is likely to respond to fluctuating conditions in the North Sea (Boulcott & Wright, 2008; Wright *et al.*, 2017a). Some of these conditions have shown a systematic trend over the course of the study, notably SST. However, although it is a known predictor of sandeel phenology in some regions (Boulcott & Wright, 2008; Frederiksen *et al.*, 2011;

Burthe *et al.*, 2012; Wright *et al.*, 2017a), it did not correlate with shag population level breeding success, as shown in previous studies based on a shorter study period (Frederiksen *et al.*, 2004). Thus, it may be that phenology of the very localised sandeel populations that shags forage on, inshore of the Isle of May, is driven by factors that we could not quantify in this study but which have shown no trend over time, or are uncorrelated with temperature. However, we cannot discount the possibility that sandeel phenology has changed but its effects have been overridden by other factors with positive effects on breeding success, such as diversification of diet which has been observed in this population (Howells *et al.*, 2017, 2018).

Alternatively, given the absence of any temporal or environmental trend, one clear possibility is that potential mismatch is not climate induced. If mismatch is present, it may instead be driven by inter-annual variability in environmental conditions that are largely independent of the directional trend of anthropogenic climate change (Youngflesh *et al.*, 2017). Alternatively, reduced breeding success in later years may be unrelated to asynchrony with prey. In at least one seabird species, Adélie penguin *Pygoscelis adelie*, timing of breeding at the population level has been found to exhibit patterns that are consistent with inherent stochasticity unrelated to measured environmental conditions, instead being embedded in the species' breeding behaviour (Youngflesh *et al.*, 2018). Youngflesh *et al* suggest that stochastic phenology exhibited by Adélie penguins may be reinforced by their synchronous breeding behaviour, as these birds use cues from conspecifics as an indicator of when to lay. It would be interesting to test whether this is also true in the case of the European shag. This species and other members of Phalacrocoracidae show very high levels of inter-annual variability in the mean laying date (Keogan *et al.*, 2018; this chapter), of which the drivers are not currently fully understood. Although shags are are much more variable in timing of breeding within a year than Adélie penguins and

other synchronous breeders (Reed *et al.*, 2006), some populations do form large foraging rafts during which time information transfer is thought to take place (Evans *et al.*, 2016). If timing of breeding is impacted even partially by transfer of information between conspecifics, this might explain why we found no effect of SST on breeding success in the Isle of May population, despite warmer springs being linked to earlier sandeel spawning phenology (Arnott & Ruxton, 2002; Van Deurs *et al.*, 2009).

While we cannot rule out that sandeels are changing their spawning phenology at a faster rate than shags are adjusting their timing of breeding, we found no evidence of positive covariance between the proportion of 1+ group sandeels in diet samples and breeding success throughout the season. Our study therefore suggests that shags do not rely on timing their breeding with the peak of availability of a single prey species, or that abundance of prey may be sufficient enough to be of low concern (Durant *et al.*, 2005) i.e. there is no clear food peak. In fact, while Isle of May shag adults do feed their chicks largely on 1+ group sandeels during the chick rearing period (Howells *et al.*, 2017), they have adopted a more generalist diet in recent years (Fortin *et al.*, 2013; Howells *et al.*, 2017, 2018). This may serve to buffer them from experiencing a reduction in population level breeding success if they are mismatched with respect to 1+ group sandeels. Currently phenology and abundance data do not exist at the scale required to fully examine whether trophic mismatch across multiple prey species impacts on breeding success of this population of shags.

4.5.4 Evidence for mismatch at the individual level

We found no evidence that the strength of stabilising selection on lay date varied between early and late seasons; that it has changed over time or between relatively cool versus relatively warm years; or that it is correlated with within-season changes in prey availability. This is contrary to other studies of selection on lay date, where changes in strength of selection across a variety of groups have been observed both

over time, and attributed to climate-induced changes in environmental conditions (Reed *et al.*, 2009, 2013a; Visser *et al.*, 2015; Smiley & Emmerson, 2016; Marrot *et al.*, 2018). However, in this study population, shags have actually advanced their lay date over time, suggesting that they may be keeping up with the general trend towards earlier spawning of principal prey, i.e. mismatch may not be present in this population.

However, there may be other reasons why we did not find evidence that potential mismatch is increasing. In this population, experienced breeders are able to adjust foraging effort and deliver food to young at the same rate throughout the season (Daunt *et al.*, 2007), suggesting they may be able to ameliorate the effects of potential mismatch with peaks in prey availability. Furthermore, developmental plasticity in this species allows offspring to prioritise structural growth during periods of food shortage, enabling them to fledge at comparable sizes to conspecifics (Moe *et al.*, 2004). As a result, any negative consequences of trophic mismatch may not be observed until a later point in life. Increased foraging effort on the part of the adults, or differential growth rates in offspring during mismatched years may also have negative fitness consequences at a later stage. Studies of the potential effect of mismatch on post-fledging and adult survival would require detailed information on the survival and reproduction of individuals throughout their lives. However, these results provide no evidence that the strength of selection on timing of breeding has changed in relation to our three proxies of mismatch.

4.6 Conclusion

It is evident that timing of breeding correlates with both population and individual fitness, which is why it is extensively used to quantify the extent to which organisms respond to environmental change. However, the assumption is often that differential rates of phenological change across trophic levels will result in peak energy availability of the resource and the requirements of the consumer becoming

asynchronous (Thackeray *et al.*, 2010), with negative consequences for consumer fitness (Visser & Both, 2005; Durant *et al.*, 2007). Yet, there are several factors which are often not considered, and may have resulted in the prevalence of climate induced phenological mismatch being overstated. Firstly, the abundance of prey may outweigh the importance of being aligned with the resource peak (Durant *et al.*, 2005), or alternatively it may be the case that no peak is present. Secondly, many species at higher latitudes are trophic generalists, and generalist consumers such as shags may simply shift prey or adopt a broader diet if they miss the peak of preferred prey (Howells *et al.*, 2017). Finally, the fitness consequences associated with mismatch may be more nuanced than simply impacting upon the number of chicks to fledge the nest. To our knowledge, no marine study (and only one terrestrial system (Charmantier *et al.*, 2008; Reed *et al.*, 2013a)) has used timing and abundance of a full suite of potential prey species coupled with information on both adult and offspring consumer phenology, growth, survival and recruitment, and environmental variables that drive their interactions. Such analyses are urgently needed for us to fully understand the causes and consequences of changes in food web dynamics, and predict how such systems would respond to future environmental change.

5. A global analysis of the signatures of temperature-mediated trophic mismatch and its effects on breeding success in seabird populations

Katharine Keogan, Francis Daunt, Nina McLean, Sarah Wanless, Richard Phillips, Sue Lewis*, Ally Phillimore*

(With the additional co-authorship of all contributors of data to the analysis)

*Joint last authors.

5.1 Abstract

Rising global temperatures have been linked to changes in phenology, with species at different trophic levels exhibiting changes of different magnitudes. This will potentially result in the timing of peak demand of consumers becoming desynchronised from the timing of peak availability of prey resources. Such temperature-mediated trophic mismatch (T-MTM) may impact negatively on demographic rates (i.e. breeding success or survival). In the marine environment, seabirds, a group of top consumers, are more threatened than any other comparable avian group. On average, they have not advanced their breeding phenology over time, nor with rising spring sea surface temperatures. However, information on prey phenology at appropriate temporal and spatial scales is lacking for most species, hindering efforts to test directly for trophic mismatch and assess its consequences for seabird breeding success at the global scale. In the absence of direct information on prey, a proxy may be used to infer the timing of peak prey abundance, coupled with a set of clear criteria that are likely to be observed if mismatch is present in a population. We combined time series' on spring sea surface temperature, annual average breeding phenology and reproductive success from 62 seabird populations and 34 species with a range of life history traits, from both hemispheres, and breeding from the equator to the poles. In the absence of information on prey biomass and

phenology, we identified criteria that would constitute evidence for consequences of T-MTM. Two of the key criteria are that breeding success should decline with increasing temperature and over time, but we found no significant evidence for this, either at the global scale or in individual populations. In 50% of populations, average breeding success declined significantly in years where lay date was later than average and we attribute this to deteriorating conditions throughout the season that may be linked to another unmeasured aspect of trophic mismatch or environmental change. We suggest that if seabird populations have become mismatched with peaks in prey availability, they may have the ability to ameliorate its effects by altering other aspects of their behaviour, or switching to another resource. Whilst seabirds are often proposed as potential casualties of T-MTM, we find limited evidence that any mismatch has led to a decline in breeding success. However, in order to fully understand how climate impacts on marine trophic interactions between consumers, it is crucial to obtain detailed information on the abundance and phenology of lower trophic level prey that are spatiotemporally relevant during the breeding season.

5.2 Introduction

The effects of ongoing climate change can be observed through changes in trophic level dynamics across both terrestrial and aquatic food webs (Thackeray *et al.*, 2010). Of particular concern is the way in which rising temperatures influence the relative timings of peak resource availability and the requirement of food by consumers during periods of peak energy requirement, for example during reproduction (Visser *et al.*, 2004; Durant *et al.*, 2007; Reed *et al.*, 2013a). In general, higher trophic level organisms are advancing their timing of breeding at a slower rate than the resources on which they feed, potentially resulting in trophic mismatch (Thackeray *et al.*, 2010; Poloczanska *et al.*, 2013). However, in order for this to be of conservation concern, trophic mismatch should have deleterious consequences for the consumer at the

population level e.g., population-level breeding success or survival should be reduced (Reed *et al.*, 2013a; McLean *et al.*, 2016). Yet, despite growing evidence for differential rates of change across trophic levels in both terrestrial and aquatic systems globally (Thackeray *et al.*, 2010, 2016; Poloczanska *et al.*, 2013), phenological trends of consumers and specific prey species have rarely been measured simultaneously. Few studies have therefore explicitly linked trophic asynchrony with changes in consumer demographic rates (but see Reed et al., 2013; Youngflesh et al., 2017).

While it is unclear exactly how prevalent temperature mediated trophic mismatch is in consumers, some species and populations may be more at risk of suffering demographic costs as a result of mismatched phenology than others (McLean *et al.*, 2016). For instance, species which breed in highly seasonal environments, such as those experienced at high latitudes, may experience higher reductions in breeding success as a consequence of missing the peak in resource availability than populations which breed year round (Moe *et al.*, 2009). Furthermore, species with inflexible foraging strategies, such as dietary specialists or those restricted to certain areas or modes of foraging (i.e. provisioning young with one or multiple prey at a time) may be more vulnerable to mismatch as they may be unable to compensate for mistimed phenology by shifting to alternative prey or foraging areas (Reed *et al.*, 2013b; Gaglio *et al.*, 2018a). Additionally, income breeding species (i.e. those which rely on resources obtained throughout the breeding season to fuel reproduction (Stephens *et al.*, 2009)) may be more at risk of mismatch than capital breeders (i.e. those species which rely on stored fat reserves to fuel reproduction (Stephens *et al.*, 2009)) (Kerby & Post, 2013), mediated by a reduced ability of adults to maintain provisioning rates to young. Consequently, the deleterious effects of

mismatch may not be spread evenly across populations of higher consumers on a global scale.

The extent to which mismatch is present in populations of higher consumers is particularly unclear in the marine environment. Here, detailed information on phenological time series across multiple trophic levels is limited (Croxall *et al.*, 2002), because both reproduction and trophic interactions between predators and their prey are hard to monitor (Schreiber & Burger, 2002). Seabirds are a diverse group of long-lived marine consumers with an array of feeding and breeding strategies and with ranges spanning from the equator to the poles (Schreiber & Burger, 2002). Although these birds forage at sea (with the exception of some kleptoparasitic species), they return to land to breed. This makes detailed studies of their reproductive phenology and success more straightforward than in other marine consumers. Available studies suggest that species occupying lower trophic levels, e.g. phytoplankton, zooplankton and larval fish, may be advancing their phenology at a faster rate than seabirds (Poloczanska *et al.*, 2013) – which in general, have neither adjusted their reproductive phenology over time (Poloczanska *et al.*, 2013; Keogan *et al.*, 2018), nor with rising sea surface temperatures (Keogan *et al.*, 2018). This may leave seabirds short of preferred prey at critical points in the season, although the diversity of seabirds as a group means that some populations may be more at risk of mismatch than others (Oro, 2014).

Previous studies have looked for evidence of mismatch and reduced breeding success in individual seabird populations (Frederiksen *et al.*, 2004; Gaston *et al.*, 2009; Shultz *et al.*, 2009; Sorensen *et al.*, 2009; Watanuki *et al.*, 2009; Bond *et al.*, 2011; Burthe *et al.*, 2012; Regular *et al.*, 2014; Ramirez *et al.*, 2016; Youngflesh *et al.*, 2017). However, the variety of methods used have made direct comparisons across studies difficult. Furthermore, the paucity of phenological time series available

for lower trophic level organisms still remains a problem. While cooler water is generally thought to indicate later peaks in abundance of lower trophic levels (plankton: Edwards & Richardson, 2004; squid: Sims *et al.*, 2001; fish: Davoren *et al.*, 2012), few studies have published time series' of the phenology of specific seabird prey over time or in response to changes in their environment (but see: Bertram *et al.*, 2001; Watanuki *et al.*, 2009; Davoren *et al.*, 2012). Evidence for the frequency of temperature mediated mismatch is therefore currently limited and mixed in seabirds, as is whether it has an impact on population level breeding success (Burthe *et al.*, 2012). Given the range in risk of mismatch in seabirds (Oro, 2014), we currently lack a complete understanding of the presence and impact of mismatch globally. With seabirds currently more widely threatened than any other comparable avian group (Croxall *et al.*, 2012), undertaking a global analysis of the impacts of climate change on mismatch is crucial to identify the species groups, regions and life history traits that expose seabirds to detrimental effects of climate on demographic rates.

In the absence of annual time series on peaks of prey abundance, identifying populations in which mismatch and its consequences are evident presents a challenge. To address the question of temperature-mediated trophic mismatch (T-MTM), one solution is to use a proxy to infer the timing of peak prey abundance, coupled with a set of clear criteria that are likely to be observed if mismatch consequences are present in a population. Sea surface temperature (SST) is generally thought to be a good proxy of the phenology, distribution and abundance of prey (Cheung *et al.*, 2013), with warmer years often indicating earlier peaks of availability of lower trophic levels (but see Reed *et al.*, 2009). In Figure 1a, we use a framework presented by Mclean *et al*., (2016) to identify five causal relationships that should exist if fitness signatures of T-MTM are present in seabird populations and is becoming more detrimental over time. However, the presence or absence of any of

these trends do not provide unequivocal evidence that mismatch is present and increasing, as other processes could be buffering or driving the observed patterns. Mismatch requires that the timing of an assumed peak biomass of seabird prey (typically plankton or forage fish) is more responsive to warming temperatures than the timing of peak demand by seabird consumers, so that the peaks in availability of and demand for nutrients become desynchronised (Durant et al., 2005; McLean et al., 2016; Thackeray, 2012) (Figures 1b, c). Ultimately this should lead to reduced overall fitness of the seabird population in mismatched years, a consequence of lower breeding success due to lower offspring provisioning rates (Figure 1d). Therefore, if T-MTM is indeed evident and becoming compounded (Thackeray *et al.*, 2016), temperature should be warming over time (criterion 1: Figure 5.1a, arrow 1), consumer breeding success should be lower in warmer years (criterion 2: Figures 5.1a, d, arrow 4), and this pathway should be negative overall. However, in order for this to be attributable to mismatch, it must be mediated by breeding phenology. Thus, timing of breeding should also show weak or neutral trends over time and with temperature (criteria 3 & 4: Figure 5.1a arrows 2 & 3), and breeding success should decline as annual average timing of breeding delays relative to an assumed food peak (criterion 5: Figure 5.1a arrow 5).

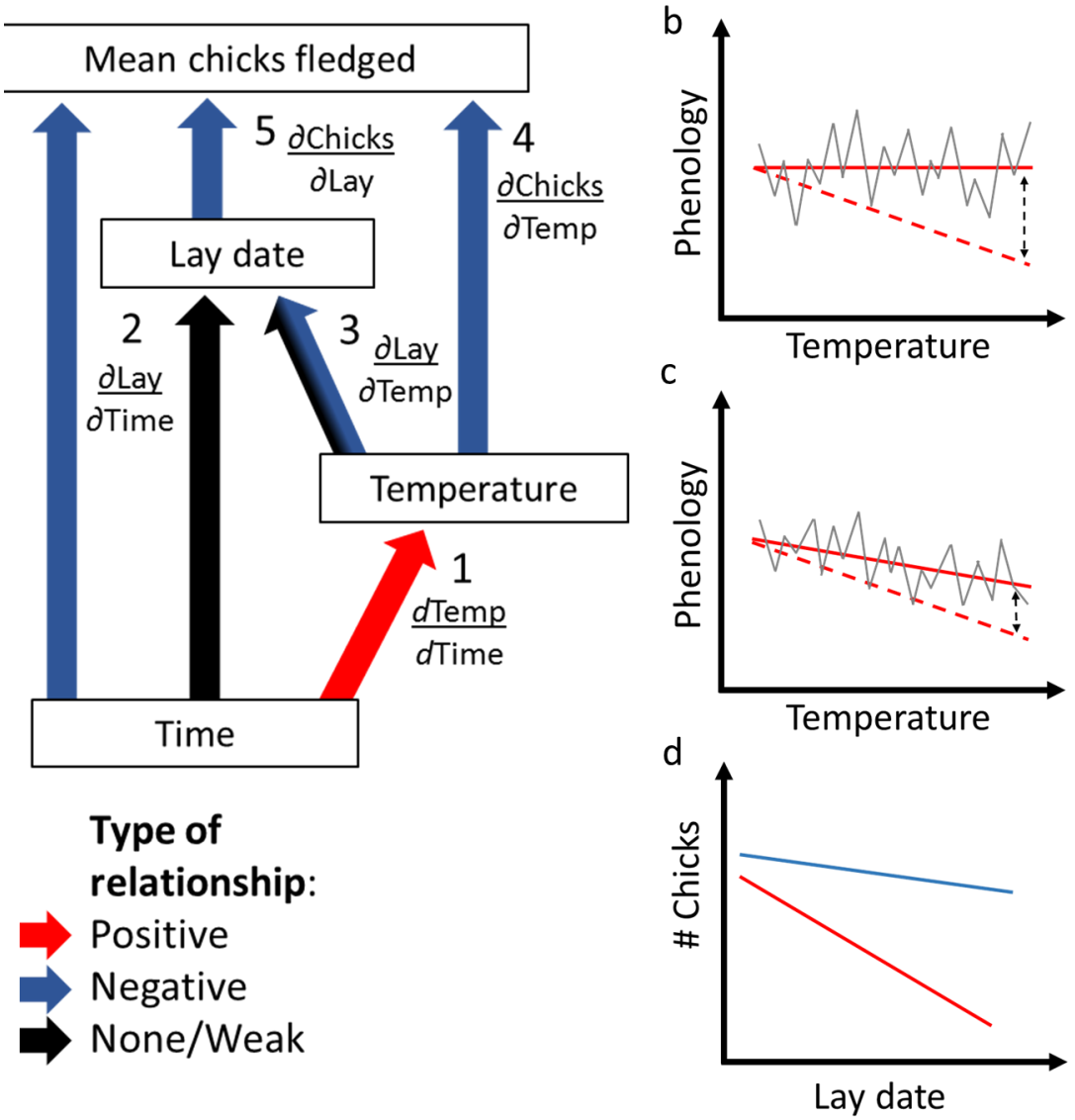


Figure 5.1. Criteria that must be observed if trophic mismatch is present in a population. **(a)** Pathways (1-5) included in the analysis and the direction of the relationships required if mismatch or its consequences are present in seabird populations globally. Red arrows indicate that a positive coefficient is required to show consistency with temperature mediated trophic mismatch, and blue arrows represent the requirement of a negative coefficient. The regression coefficients of pathway 2 (black arrow) should be shallow (Keogan *et al.*, 2018), and we predict that for T-MTM to be present, regression coefficients of pathway 3 should not be strongly negative (black/blue arrow). **(b – d)** show potential relationships between bird reproductive phenology (**(b, c)** red solid line = average response, grey solid line = interannual variation (Keogan *et al.*, 2018)), timing of peak fish abundance (red dashed line) and mismatch (black dashed arrow and **(d)**) under the assumption that in the past birds were trophically matched with their resources. In **(b)** birds are not responding to rising temperatures, and $\frac{\partial Chicks}{\partial Temp}$ and $\frac{\partial Chicks}{\partial Lay}$ **(a)** are predicted to show negative slopes. In **(c)** birds are responding to rising temperatures, and $\frac{\partial Chicks}{\partial Temp}$ and $\frac{\partial Chicks}{\partial Lay}$ **(a)** are negative slopes, and $\frac{\partial Lay}{\partial Temp}$ is weakly negative. $\frac{\partial Chicks}{\partial Temp}$ will be less steep in **(c)** than in **(b)**. **(d)** represents a negative interaction between

population average lay date and temperature, where mismatch is stronger in warm years (red line) than cold years (blue line). On x and y axes, direction of arrows indicate increasing lay date, number of chicks and temperature, respectively.

In this study, we assess the evidence that T-MTM leads to reduced breeding success in global seabird populations with a variety of feeding and breeding strategies and from a range of latitudes, using surface temperature as a proxy for peak prey availability and with the five criteria necessary for mismatch explained in Figure 1. Specifically, we aim to answer three key questions: (i) Does mismatch result in reduced breeding success at the population level? We predict that if T-MTM results in reduced breeding success at the population level, breeding success should be lower when SST is warmer and lay date is later. (ii) Is T-MTM increasing over time? If mismatches are intensifying over time, SST should be increasing and the pathway time-temperature-chicks fledged should be negative. (iii) What (if any) factors influence the impact of mismatch? If mismatch is present, we predict it should be stronger in populations breeding at higher latitudes, in single-prey loaders and in income breeders.

5.3 Methods

5.3.1 Data

Datasets were collated between March and October 2018 by contacting researchers directly and by requesting data at the 4th World Seabird Twitter Conference (April 2018) and the 14th International Seabird Conference (September 2018).

Each time series was required to be 15 or more years in duration and contain information on annual mean lay date (and associated standard error) and annual mean number of chicks fledged per nest (including failed nests). Mean number of chicks fledged per nest was log transformed for all analyses after the data had been collected, and associated standard errors could therefore not be accurately estimated

for this term. If data were collected from multiple plots in a population, the data were pooled. Datasets which did not include information for each of these three terms were still used in certain sections of the analysis, but were omitted from a model if data for each included term were not available every year, e.g. if there were missing standard errors. The number of breeding phenology records measured per year were ten or more, so that the annual standard error of the mean could be calculated. In cases where standard error could not be calculated, we ran the analysis both with these time series (without weighting for standard error) and without them (weighting response variables for standard error). Where breeding phenology was measured as date of hatching, we back-calculated to lay date using average incubation period (details in Keogan et al., 2018). To focus analyses on within population variation in predictors we centred all predictor variables with respect to population means (van de Pol & Wright, 2009).

Mean monthly temperature interpolated on a 100km^2 grid - centred on the breeding site – was downloaded for each population for the three months prior to breeding from https://www.esrl.noaa.gov/psd/data/gridded/data.noaa.oisst.v2.html (Keogan *et al.*, 2018). This allowed us to determine whether years were warm or cold relative to the long term trend, and thus whether peak prey biomass abundance is likely to occur early (relatively warm year) or late (relatively cold year). For each time series, we collated additional information on latitude and longitude of the breeding site, and whether the species feeds its chick using single or multiple prey loading (multiple prey loaders included those that regurgitate food for the chick) or an income (i.e. those which rely on taking resources throughout the breeding season to fuel reproduction) or capital breeder (i.e. those species which may use stored fat reserves to fuel reproduction in the case where food resources were limited) (for sources, see Table C.14). To account for the effect of hemisphere, day 1 was taken to be 1st of Jan

in the Northern hemisphere and 1st of July in the Southern hemisphere, with leap years taken into account.

Prior to running analyses we submitted a pre-registration (https://osf.io/4cwnt/) and we identify below where our methods departed from those set out in the pre-registration using post hoc tests.

5.3.2 Questions

Question 1: Does temperature-mediated trophic mismatch result in reduced breeding success at the population level?

Prediction: ***If T-MTM results in reduced breeding success at the population level, breeding success should be lower when SST is warmer and lay date is later.***

i) A strongly negative relationship between annual mean lay date and time ($\frac{\partial \text{Lay}}{\partial \text{Time}} < 0$, Figure 5.1a - arrow 2), or temperature ($\frac{\partial \text{Lay}}{\partial \text{Temp}} < 0$, Figure 5.1a - arrow 3) would be inconsistent with the hypothesis that seabird breeding phenology is failing to keep pace with temperature-mediated changes to resource phenology. If mismatch is present, a relationship between $\frac{\partial \text{Lay}}{\partial \text{Time}}$ and $\frac{\partial \text{Lay}}{\partial \text{Temp}}$ should therefore be weak or absent.

ii) If climate change-mediated mismatches impact negatively on reproduction, then T-MTM should be greater in the years when the SST is warmer ($\frac{\partial \text{Chicks}}{\partial \text{Temp}} < 0$, Figure 5.1a - arrow 4). This should arise due to birds failing to adjust their lay date at a sufficient rate to keep up with the prey resource timing, and thus decreasing offspring provisioning. There should also be a negative regression

coefficient between lay date and the number of chicks fledged ($\frac{\partial \text{Chicks}}{\partial \text{Lay}} < 0$, Figure 5.1a - arrow 5).

iii) On their own, neither $\frac{\partial \text{Chicks}}{\partial \text{Temp}} < 0$ nor $\frac{\partial \text{Chicks}}{\partial \text{Lay}} < 0$ constitutes strong evidence that trophic mismatch impacts negatively on breeding success. If $\frac{\partial \text{Chicks}}{\partial \text{Temp}} < 0$ (i.e. fewer chicks fledged at warmer temperatures), but $\frac{\partial \text{Chicks}}{\partial \text{Lay}}$ is ≥ 0 (i.e. equal or more chicks fledged when lay date is later), this will indicate that climate is important for breeding success, but not as a result of trophic mismatch, i.e. timing (lay date) is a key component of mismatch so must have a significantly negative effect for it to be present.

iv) If seabird peak resource demand is earlier than peak resource availability in the coldest years but after peak resource in the warmest years it is possible that a negative interaction between lay date and temperature exists, such that the negative slope between lay date and chicks fledged is most negative in warmer years (as shown in Figure 5.1c) and potentially positive in cold years.

Question 2: is temperature-mediated trophic mismatch increasing over time?

Prediction: ***If mismatches are intensifying over time, SST should be increasing and the pathway time-temperature-chicks fledged should be negative.***

For T-MTM to be worsening under climate change, temperatures must be rising over time ($\frac{d\text{Temp}}{d\text{Time}} > 0$, Figure 5.1a - arrow 1). This is a necessary condition but not sufficient alone as there must also be evidence that mismatch negatively impacts reproduction. If mismatch is intensifying over time we predict that the $\frac{d\text{Temp}}{d\text{Time}} \times \frac{\partial \text{Chicks}}{\partial \text{Temp}}$ pathway will be

negative (pathway: arrow 1 – arrow 4), but we can only infer mismatch if timing is important i.e a negative value of $\frac{\partial \text{Chicks}}{\partial \text{Lay}}$ (arrow 5).

Question 3: what (if any) factors influence the impact of mismatch?

Prediction: ***If mismatch is present, it should be stronger in populations breeding at higher latitudes; in single-prey loaders and in income breeders.***

We will use three regional and life history variables to explain among population variance surrounding the average response of $\frac{\partial \text{Chicks}}{\partial \text{Temp}}$ and $\frac{\partial \text{Chicks}}{\partial \text{Lay}}$.

i) If mismatch is present, we predict the consequences are likely to be more negative in populations breeding at high latitudes. Regions at higher latitudes are characterised by strongly seasonal environments and shorter peaks in food resources, increasing the chances that a trophic mismatch between seabirds and peak availability of their food will have a negative impact on breeding success (Moe *et al.*, 2009).

ii) We predict that single-prey loaders (i.e. those can only deliver one prey item at a time) will be at a higher risk of reduced breeding success as a result of trophic asynchrony than those which can deliver multiple prey or regurgitate. Increased foraging effort required during mismatched years is therefore likely to impact negatively on breeding success more strongly in single-prey loaders.

iii) We predict that income breeding species will be more at risk of mismatch than capital breeders (Kerby & Post, 2013), mediated by a reduced ability of adults to maintain provisioning rates to young.

5.3.3 Statistical analyses

We ran three core statistical models (Table 5.1) that together estimated all of the effects in Figure 5.1a, and allowed the intercepts and slopes to vary across populations. Models were run in a Bayesian generalised linear mixed model framework in R (Hadfield, 2010), using default diffuse priors for the fixed effects and parameter expanded priors for the random effects, with the exception of the residual which followed an inverse-Wishart. All model posteriors were checked for adequate mixing and ran for 250,000 iterations.

The fixed effects allowed us to interpret the main effect for each predictor (Figure 5.1a). To infer individual population estimates, we extracted the posterior mode of the random effects (equivalent to BLUPs) of the intercept and slope for each predictor at the population level. Any observed variance in intercept (i.e. average temperature, average lay date & average breeding success) across populations may additionally be impacted by the location at which they breed. To account for this, we summed the posterior modes of the breeding site intercept and the population intercept to estimate the average response for each population. To calculate the % decline in breeding success for each of the main effects (models 3 - 4f), we subtracted the exponent of the posterior mean from 1, and multiplied by 100.

Where there is measurement error in a causal predictor variable this will downwardly bias the estimated effects of this predictor and has the potential to upwardly bias the effect of any correlated predictor included in the model. We did not account for the effect of phylogenetic relatedness in these analyses in order to avoid over-parameterising the models. This may upwardly bias the estimated effects of correlated predictor variables, such as the effect of species on breeding phenology or success.

Table 5.1. Terms used in each model, which together examined all of the arrows in Figure 5.1. Fixed effects estimate the average slope for each term and the us() variance structure as random effects allow the slopes and intercepts to (co)vary across populations. All response variables were normally distributed.

Question/ Prediction	Model	Response	Fixed effects*	Random effect syntax*
Questions 1 & 2	1	SST	Year $\left(\frac{\partial \text{Temp}}{\partial \text{Time}}\right)$	yearF + site + site:yearF + us(1 + year):population
Question 1	2	Lay date	Year $\left(\frac{\partial \text{Lay}}{\partial \text{Time}}\right)$ + SST $\left(\frac{\partial \text{Lay}}{\partial \text{Temp}}\right)$	yearF + site + site:yearF + us(1 + year + temperature):population
Questions 1 & 2	3	Chicks fledged	year $\left(\frac{\partial \text{Chicks}}{\partial \text{Time}}\right)$ + SST $\left(\frac{\partial \text{Chicks}}{\partial \text{Temp}}\right)$ + lay date $\left(\frac{\partial \text{Chicks}}{\partial \text{Lay}}\right)$	yearF + site + site:yearF + us(1 + year + temperature + lay date):population
Question 1	3a	Chicks fledged	year $\left(\frac{\partial \text{Chicks}}{\partial \text{Time}}\right)$ + SST $\left(\frac{\partial \text{Chicks}}{\partial \text{Temp}}\right)$ + lay date $\left(\frac{\partial \text{Chicks}}{\partial \text{Lay}}\right)$ + SST:lay date	yearF + site + site:yearF + us(1 + year + temperature + lay date):population
Question 3	4a, 4b	Chicks fledged	year $\left(\frac{\partial \text{Chicks}}{\partial \text{Time}}\right)$ + SST $\left(\frac{\partial \text{Chicks}}{\partial \text{Temp}}\right)$ + lay date $\left(\frac{\partial \text{Chicks}}{\partial \text{Lay}}\right)$ + latitude + SST:latitude **OR** lay date:latitude	yearF + site + site:yearF + us(1 + year + temperature + lay date):population
Question 3	4c, 4d	Chicks fledged	year $\left(\frac{\partial \text{Chicks}}{\partial \text{Time}}\right)$ + SST $\left(\frac{\partial \text{Chicks}}{\partial \text{Temp}}\right)$ + lay date $\left(\frac{\partial \text{Chicks}}{\partial \text{Lay}}\right)$ + income/capital + SST:income/capital **OR** lay date:income/capital	yearF + site + site:yearF + us(1 + year + temperature + lay date):population

Question 3	4e, 4f	Chicks fledged	year $\left(\frac{\partial \text{Chicks}}{\partial \text{Time}}\right)$ + SST $\left(\frac{\partial \text{Chicks}}{\partial \text{Temp}}\right)$ + lay date $\left(\frac{\partial \text{Chicks}}{\partial \text{Lay}}\right)$ + feeding strategy + SST:feeding strategy **OR** lay date:feeding strategy	yearF + site + site:yearF + us(1 + year + temperature + lay date):population

* Note that all models include year as a categorical random effect (denoted yearF) and as a continuous fixed covariate (denoted year). The motivation for including year as a categorical random effect is to take into account the potential for some years (or site:yearF, due to multiple species at some sites) to be associated with high (or low) values of the response variable due to the effects of other variables that we have not included. The motivation for including year as a covariate is to de-trend the analysis between temperature and response variables (Iler *et al.*, 2016; Keogan *et al.*, 2018), thereby reducing the risk of attributing to temperature any effects caused by a third variable (e.g., impact of fisheries, increase in plastic pollution) that has changed over time. In model 1, the response variable was weighted by its annual standard error. However, in model 2, standard error of lay date was unavailable for three time series', and so we analysed the effects of time and SST on lay date both with (population n = 59) and without (population n = 62) weighting for standard error. In models 3 - 4f, standard errors of the log transformed breeding success were not available and these models are therefore not weighted by error.

5.3.4 Path Analysis

In order for temperature mediated trophic mismatch to be increasing over time, breeding success should be decreasing over time and mediated by increasing temperature (Figure 5.1a, arrow 1 – arrow 4). We estimated the strength of the pathway 1 - 4 for each population in the analysis using the posteriors of the regression coefficients from model 1 ($\frac{d\text{Temp}}{d\text{Time}}$, Table 5.1) and model 3 ($\frac{\partial \text{Chicks}}{\partial \text{Temp}}$, Table 5.1). For temperature mediated trophic mismatch to be getting stronger over time, this pathway must be negative.

$\frac{d\text{Temp}}{d\text{Time}}$ x $\frac{\partial \text{Chicks}}{\partial \text{Temp}}$ < 0 (fewer chicks over time), $\frac{\partial \text{Chicks}}{\partial \text{Lay}}$ < 0 (fewer chicks as lay date increases)

(arrow 1 x arrow 4 < 0, contingent on arrow 5 < 0)

5.4 Results

We collated 62 time series from 34 species and 24 breeding sites spanning 21.52° – 74.50° latitude and 0.07° – 176.56° longitude between 1975 and 2018 (Figure 5.2, Table C.14). In total there were 46 populations for which income or capital energy storage strategies could be defined (income breeders = 42 populations, 5 species; capital breeders = 4 populations, 3 species), and 56 populations for which chick provisioning strategy could be defined (single prey loaders = 15 populations, 5 species; multiple prey loaders = 14 populations, 9 species; regurgitators = 27 populations, 16 species).

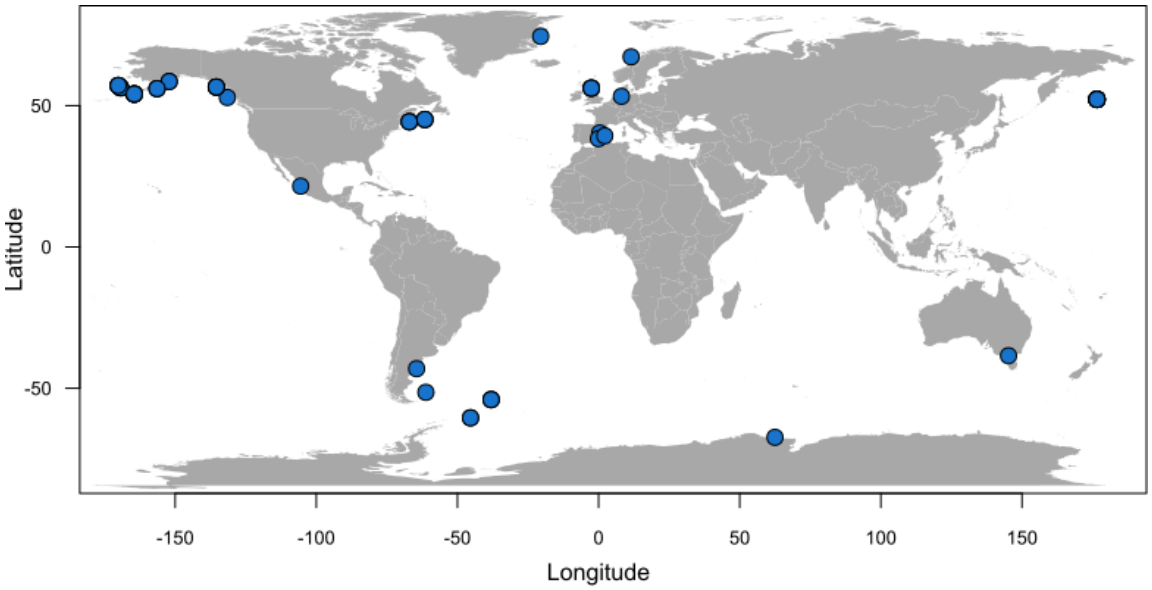


Figure 5.2. Map of all sites included in the analysis.

5.4.1 Question 1: does mismatch result in reduced breeding success at the population level?

We examined trends in average lay date over time ($\frac{\partial \text{Lay}}{\partial \text{Time}}$, Figure 5.1a - arrow 2, Table 5.1) and with sea surface temperature ($\frac{\partial \text{Lay}}{\partial \text{Temp}}$, Figure 5.1a - arrow 3), and explored

variation in slopes between populations. All estimates are therefore reported from the weighted model which takes standard error of the response into account. On average, we found no trend in lay date over time (-0.67 days decade^{-1}, 95% CIs = -1.76, 0.41. Figure 5.3b, Table C.3) or with SST (-0.60 days °C^{-1}, 95% CIs = -1.97, 0.79. Figure 5.3c, Table C.3). However, there was significant variation in the responses of populations, with slopes of lay date both over time (variance in year slopes: 0.08, 95% CIs = 0.04, 0.13. Table C.3) and with temperature (variance in temperature slopes: 16.58, 95% CIs = 9.34, 24.66. Table C.3) varying substantially. Extreme year slopes ranged from -7.23 days decade^{-1} (95% CIs = -9.86, -4.63. Table C.15) for European shags *Phalacrocorax aristotelis* on the Isle of May, southeast Scotland to 6.9 days decade^{-1} (95% CIs = 3.227, 10.744. Table C.15) for common guillemots *Uria aalge* on East Amatuli Island, Alaska. Extreme temperature slopes ranged from -7.93 days °C^{-1} (95% CIs = -14.75, -0.85. Table C.15) for little penguins *Eudyptula minor* on Phillip Island, southeast Australia to 18.91 days °C^{-1} (95% CIs = 16.67, 21.20. Table C.15) for blue-footed boobies *Sula nebouxii* on Isla Isabel, Mexico.

We examined trends in breeding success over time, and with increasing sea surface temperature ($\frac{\partial \text{Chicks}}{\partial \text{Temp}}$, Figure 5.1a - arrow 4) and lay date ($\frac{\partial \text{Chicks}}{\partial \text{Lay}}$, Figure 5.1a - arrow 5, Table 5.1). We found no overall decline in mean number of chicks fledged per nest (our measure of breeding success), over time (2% decline decade^{-1}, slope = -0.02, 95% CIs = -0.03, 0.0003. Figure 5.4a, Table C.4), or with temperature (1.4% decline °C^{-1}, slope = -0.014, 95% CIs = -0.03, 0.01. Figure 5.4b, Table C.4). However, breeding success became significantly lower as average lay date became later (0.7% decline day^{-1}, slope = -0.007, 95% CIs = -0.008, -0.005. Figure 5.4c, Table C.4), equivalent to a 10.0% reduction in breeding success for every two week delay in laying date. The populations in this analysis had between-year variation in lay date ranging from 2 days (black-browed albatross *Thalassarche melanophris*, New Island,

Falklands/Malvinas) to 112 days (blue footed booby *Sula nebouxii*, Isla Isabel, Mexico) (mean = 25 days) . Therefore, some populations may experience a decline in breeding success as lay date delays because they are biologically more likely to adjust their lay date from year to year. In a *post hoc* test, we tested the prediction that populations with higher between-year variation in average lay date (i.e. a greater number of days between earliest and latest lay dates during the study period) would exhibit steeper slopes than those which breed on a similar day each year. We found no evidence that populations with a wider range of lay dates showed a steeper decline in breeding success as lay date delayed (minimum range = 2 days, black-browed albatross *Thalassarche melanophris*; maximum range = 112 days, blue-footed booby *Sula nebouxii*, Table C.15).

To estimate the baseline change in breeding success over time $\frac{\partial \text{Chicks}}{\partial \text{Time}}$ (Figure 5.1a) without taking other predictors into account, we ran a *post hoc* analysis using model 3 (Table 5.1) with year as the only fixed effect. There was no significant decline in breeding success over time (Table C.13) when year was the only term included in the model (Figure 5.5).

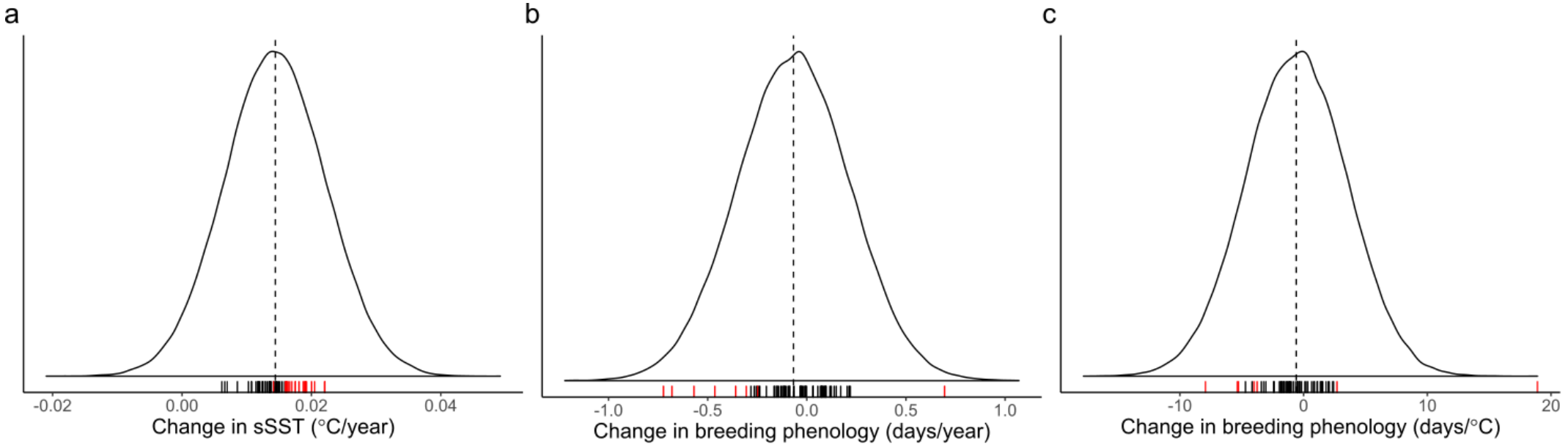


Figure 5.3. Smoothed histogram of the distribution of trends in temperature change (a) and change in lay date (b, c). To make each histogram, 100000 samples were drawn from a normal distribution with the mean = posterior mean and the variance = variance in slopes for each population. Dashed line represents the average response slope across the study period (a), temperature range (b) and range of average lay dates (c). The dashed lines below each histogram represent the individual population level slope estimates (black = non-significant slopes, red = significant slopes), which were estimated using the posterior mode of the random effect for each term and with the site effect taken into account. See Table C.15 for regression coefficients for each population.

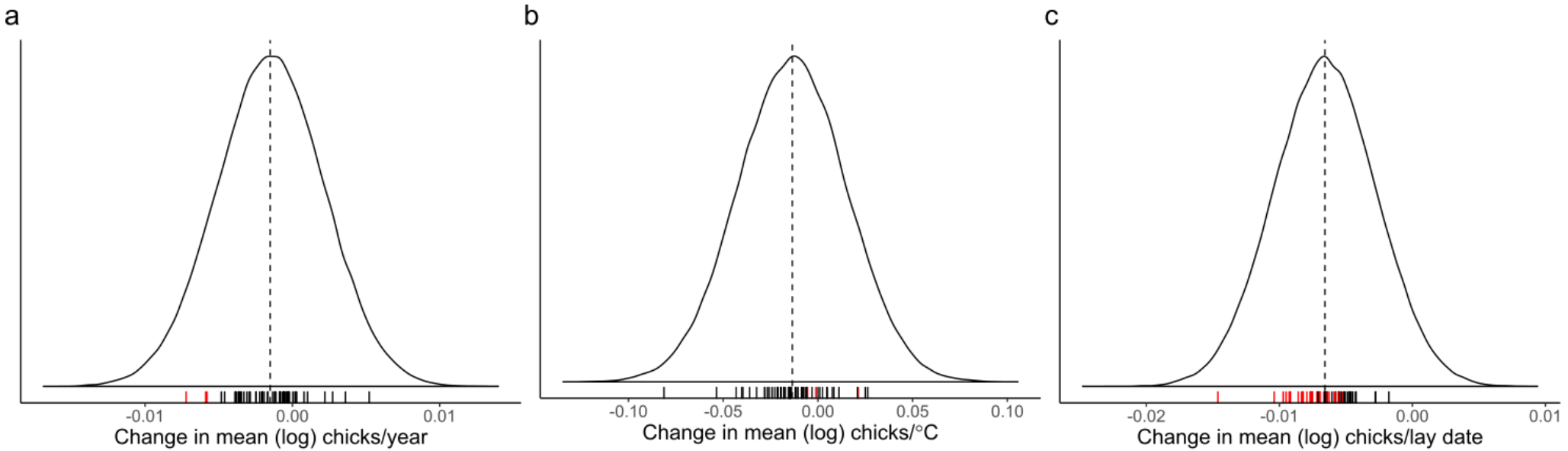


Figure 5.4. Smoothed histogram of the distribution of trends in breeding success (log transformed average number of chicks fledged per nest) over time (a), temperature (b) and lay date (c). To make each histogram, 100000 samples were drawn from a normal distribution with the mean = posterior mean and the variance = variance in slopes for each population. Dashed line represents the average response slope across the study period (a), temperature range (b) and range of average lay dates (c). The dashed lines below each histogram represent the individual population level slope estimates (black = non-significant slopes, red = significant slopes), which were estimated using the posterior mode of the random effect for each term and with the site effect taken into account. See Table C.15 for regression coefficients for each population.

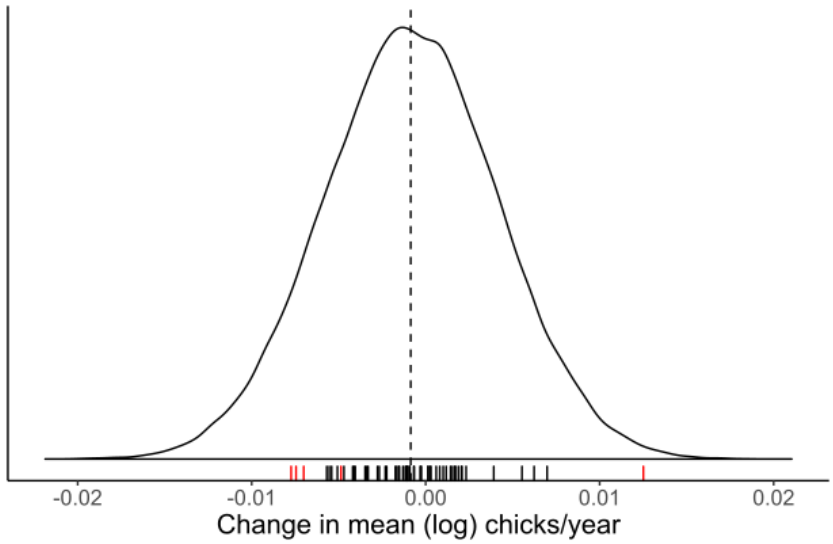


Figure 5.5. Smoothed histogram of the distribution of trends in breeding success (log transformed average number of chicks fledged per nest) over time, using the breeding success model (model 3) with year as the only fixed effect. To make the histogram, 100000 samples were drawn from a normal distribution with the mean = posterior mean and the variance = variance in slopes for each population. Dashed line represents the average response slope across the study period. The dashed lines below the histogram represent the individual population level slope estimates (black = non-significant slopes, red = significant slopes), which were estimated using the posterior mode of the random effect for the year term and with the site effect taken into account. See Table C.15 for regression coefficients for each population.

We tested the prediction that years when mean lay date was later and temperatures were high would result in a steeper decline in breeding success than the effect of breeding later in a cool year (Figure 5.1b-d, Table 5.1). We found no significant positive/negative interaction (slope = -0.0012, 95% CIs = -0.0023, 0.00003), although the effect of temperature was marginally more detrimental to breeding success in years where lay dates were earlier. This is the equivalent of a 1.4% reduction in breeding success (95% CIs = -0.032, 0.003) for every 1°C increase in sea surface temperature when average lay date was day 125 (the average lay date across all populations), and a 3.0% reduction in breeding success (95% CIs = -0.054, -0.005) for every 1°C increase in temperature if average lay date occurred two weeks later.

In addition to overall trends, we also explored variation (i.e. in breeding success over time, temperature and lay date) across populations. Only 3/57 populations (~5%) showed significant declines in breeding success over time, with extreme year slopes ranging from a 7% decline decade^{-1} in Atlantic puffin, *Fratercula arctica* on Hernyken, Norway (95% CIs = -0.122, -0.026. Table C.15), to a 5% increase in breeding success decade^{-1} (-0.0009, 0.0113. Table C.15) for the European shag *Phalacrocorax aristotelis* on the Isle of May, Scotland. Only one population (1.75% of total (n = 57)) showed a significant decline in breeding success in warmer years (red-faced cormorant *Phalacrocorax urile* on St. Paul island, Alaska, 8% decline °C^{-1} (95% CIs = -0.1408, -0.0258. Table C.15). The black-legged kittiwake *Rissa tridactyla* population on Chowiet Island, Alaska, was the most extreme positive slope, and showed a non-significant increase in breeding success of 3% °C^{-1} (95% CIs = -0.0184, 0.0737. Table C.15). Finally, 28/57 (49%) populations showed significant declines in breeding success for every day later in the season average lay date occurred. Extreme slopes for lay date ranged from 1% decline day^{-1} (95% CIs = -0.0214, -0.0071. Table C.15) in the red-faced cormorant *Phalacrocorax urile* on St. Paul Island, Alaska, to 0.2% decline day^{-1} (95% CIs = -0.0078, 0.0043) in the Adélie penguin *Pygoscelis adelie* population on Signy Island, South Orkney.

5.4.2 Question 2: is temperature-mediated trophic mismatch increasing over time?

We examined trends in sea surface temperature over time ($\frac{\partial \mathrm{Temp}}{\partial \mathrm{Time}}$, Figure 5.1a - arrow 1, Table 5.1) and explored variation in slopes between populations. Overall, average sea surface temperatures surrounding the breeding site during the pre-breeding period increased over time during the study period (0.144°C decade $^{-1}$, 95% Credible Intervals [CIs] = 0.048, 0.247. Figure 5.3a, Table C.1), which satisfies the criterion for

prediction 1a, though the increase was only significant for 34% of time series (Table C.15). Extreme temperature slopes ranged from 0.06°C decade $^{-1}$ (95% CIs = -0.074, 0.19. Table C.15) in the population of black-browed albatross *Thalassarche melanophris*, on Bird Island (South Georgia) to 0.221°C decade $^{-1}$ in two populations: blue-footed booby *Sula nebouxii* on Isla Isabel (Mexico) (95% CIs = 0.046, 0.43. Table C.15) and roseate tern *Sterna dougallii* on Country Island, Nova Scotia, Canada (95% CIs = 0.018, 0.46. Table C.15). Variation in the rate of temperature change over time across populations was non-significant (variance in temperature slopes: 0.00006, 95% CIs = 0, 0.0002).

We estimated the strength of pathway 1 - 4 ($\frac{d\text{Temp}}{d\text{Time}}$ x $\frac{\partial\text{Chicks}}{\partial\text{Temp}}$, Figure 5.1a) using the regression coefficients from models 1 $\frac{d\text{Temp}}{d\text{Time}}$ and 3 $\frac{\partial\text{Chicks}}{\partial\text{Temp}}$. For consequences of mismatch to be present, this pathway should be negative, contingent on a negative regression coefficient of $\frac{\partial\text{Chicks}}{\partial\text{Lay}}$ (model 3). Overall the change in breeding success over time mediated by temperature ($\frac{d\text{Temp}}{d\text{Time}}$ x $\frac{\partial\text{Chicks}}{\partial\text{Temp}}$) was non-significant (0.002% decline in breeding success days decade^{-1}, 95% CIs = -0.005, 0.0007). While $\frac{\partial\text{Chicks}}{\partial\text{Lay}}$ was significantly negative on average (0.007% decline in breeding success day^{-1}, 95% CIs = -0.008, -0.005), the criteria required to provide evidence that mismatch is increasing over time were not all met. We also found no evidence that T-MTM is increasing over time in individual populations (Table C.15).

5.4.3 Question 3: what (if any) factors influence the impact of mismatch?

With breeding success as the response variable, we found no significant interaction between temperature and latitude (Table C.6), income/capital breeding strategy (Table C.8) or chick provisioning method (Table C.10), or between lay date and

income/capital breeding strategy (Table C.9), or chick provisioning method (Table C.11). However, we found a very small but significant interaction between lay date and absolute latitude (interaction slope = -0.00006, 95% CIs = (-0.0001, -0.00001), Table C.7). This is equivalent to a 9% reduction in breeding success if lay date occurred 14 days later than average at 40° latitude (slope = -0.09, 95% CIs = -0.12, -0.07), but a reduction of 11% if lay date occurred the same number of days later at 65° latitude (slope = -0.11, 95% CIs = -0.14, -0.09).

5.5 Discussion

Across 62 populations of seabirds distributed around the world's oceans we found little evidence that temperature-mediated mismatch is impacting negatively on breeding success. Out of five criteria for T-MTM, we found that four were met overall: temperature has risen over time across the study sites on average (criterion 1), lay date showed no trend over time or with temperature overall (criteria 3 & 4), and breeding success was reduced in later years (criterion 5). However, only nine individual populations met all of these four criteria, suggesting that these pathways are not particularly strong. Furthermore, the key criterion that would suggest that signatures of T-MTM are present, i.e. that breeding success must be lower in warmer years due to an earlier peak of prey availability (criterion 2) was met neither at the global average nor individual population levels. We found little evidence that signatures of mismatch are more pronounced in populations breeding at higher latitudes, or in populations of species that are income breeders or single prey loaders. Population declines in seabirds (Paleczny *et al.*, 2015) are likely to be driven by factors other than T-MTM, such as adult survival. We infer that if mismatch is present it is unlikely to be driven by temperature, and it is not increasing over time.

By analysing environmental change, phenological insensitivity and breeding success data from many independent studies we were able to directly compare trends in mismatch and breeding success across populations occupying different regions and with diverse life history traits. Previous large scale comparative studies have suggested that the differential rates of change over time that have been observed across trophic levels may lead to trophic mismatch in the form of reduced breeding success. We found little evidence for this, and while some individual studies are in accordance with our findings (Frederiksen *et al.*, 2004; Shultz *et al.*, 2009; Bond *et al.*, 2011; Burthe *et al.*, 2012), others find contrasting results (Gaston *et al.*, 2009; Sorensen *et al.*, 2009; Watanuki *et al.*, 2009; Ramirez *et al.*, 2016; Youngflesh *et al.*, 2017). This may be due to differences in methods used between individual studies, e.g. differential measures of the environment or of breeding success, which makes direct comparisons difficult. For example, Gaston *et al.*, (2009) found evidence for mismatch of median egg laying of common guillemots with peak in food availability at lower trophic levels, but using sea ice extent and peak colony attendance, not sea surface temperature as proxies for resource availability. Watanuki et al (2009) showed that in warmer years, the date at which adult rhinoceros auklets switch to key prey is mismatched with hatching date of their chicks. As information on year specific diet data were not available for each population, we did not use this approach in our analysis. Ramírez et al (2016) and Youngflesh et al (2017) use detrended analyses and found evidence for mismatch in some years, suggesting that if mismatch is present, it may be driven by interannual variability in environmental conditions, and not temperature-driven or linear trends. In this analysis, we made clear predictions about the trends expected under T-MTM, and the criteria that should be met if it were increasing over time. A further global analysis of the magnitude of interannual

variation in phenology and breeding success across populations would be a useful and interesting next step.

Previous studies suggest that marine species occupying lower trophic levels are adjusting their phenology at a faster rate over time than seabirds (Edwards & Richardson, 2004; Poloczanska *et al.*, 2013). If prey are getting earlier in their phenological peaks over time at a faster rate than that of the peak in seabird resource requirements, there may be several reasons why we found no evidence for reduced breeding success. Seabirds typically have long incubation periods and periods until chicks reach peak energy requirements, hence predicting future conditions may be less relevant to seabird species than those which have shorter reproductive events. In mismatched years or those where weather or climate conditions are poor, adults within a population may skip breeding all together to avoid incurring fitness costs that might be deleterious to their own survival (Stearns, 1989; Erikstad *et al.*, 1998). Skipped breeding occurs in numerous seabird species including European shags (Aebischer, 1985), black-legged kittiwakes (Goutte *et al.*, 2010), common guillemots (Reed *et al.*, 2015), red footed boobies (Cubaynes *et al.*, 2011) and procelariiformes (Jenouvrier *et al.*, 2005), and at varying frequencies within a population (high: Chastel *et al.*, 1993; low: Reed *et al.*, 2015). It is plausible that in years when a lower proportion of the population breed, a decrease in intraspecific competition allows breeding individuals to be more successful (Reed *et al.*, 2013a). Furthermore, the effects of mismatch may be evident not only in offspring quantity, but also in their quality, i.e. growth rate and mass at fledging (Watanuki *et al.*, 2009; Doiron *et al.*, 2015) or post fledging (Divoky *et al.*, 2015). Additionally, offspring survival may not only be impacted prior to, but also after fledging (Furness & Tasker, 2000). Merely measuring the number of chicks to leave the nest therefore makes it difficult to detect alternative or additional effects of trophic mismatch.

The main determinant of breeding success in our analyses was laying date, with a later than average year resulting in a lower number of chicks fledged per nest in a population. One explanation for this may be that prey respond to a driver (other than spring sea surface temperature) which causes spring to arrive earlier than the peak of consumer requirements in later years, such as solar irradiation (Townsend *et al.*, 1994) or ocean stratification (Behrenfeld *et al.*, 2006). While it is unclear exactly what drives breeding phenology in seabirds, in some cases it is influenced by carry over effects of conditions experienced over winter (Sorensen *et al.*, 2009; Daunt *et al.*, 2014). After a winter when conditions have been unfavourable, individuals may take longer to reach breeding condition (Daunt *et al.*, 2014). In such years, the whole population may be mismatched with resources that peaked in their availability earlier in the season, or be forced to abandon breeding if resources are limited (Regular *et al.*, 2014). If breeding occurs later in the year, seabirds may also need to allocate finite resources towards self-maintenance (Stearns, 1989), preparing for energetically expensive periods of post-breeding migration and moulting (Schreiber & Burger, 2002) at the expense of rearing chicks.

Surprisingly, we found no overall trend towards decreasing breeding success over time, despite the fact that between 1950 and 2010, the globally monitored seabird population (approx. 20% of the total global seabird population) declined by 70% (Paleczny *et al.*, 2015). This suggests that population sizes are driven by factors other than breeding success. In seabirds and other long-lived species with slow life histories, even a small reduction in adult survival can have drastic implications for population growth rate (Furness & Tasker, 2000; McLean *et al.*, 2016). Survival of adult seabirds is currently threatened by a suite of drivers, including pollutants (Provencher *et al.*, 2019), competition with fisheries (Grémillet *et al.*, 2018), extreme climate events resulting in mass mortalities (Jones *et al.*, 2018), and interference from

invasive mammalian predators (Schreiber & Burger, 2002). Therefore, while adults may still be able to raise chicks to fledging age, perhaps a result of their plastic foraging strategies and a diversification of diets, this is clearly not sufficient to keep population sizes constant.

While seabirds may be able to mitigate the effects of missing the peak in resources, an alternative viewpoint may be that in arriving at general predictions we have failed to capture the complex ways in which match-mismatch operates. Timing reproduction in order to provide offspring with sufficient resources may be much more complicated than simply intersecting the period at which prey availability peaks. In complex food webs it is likely that predator reproductive success is also correlated with prey abundance (Durant *et al.*, 2007), quality (Österblom *et al.*, 2008), and growth rate (Burthe *et al.*, 2012) and species composition (Howells *et al.*, 2018). Furthermore, although long-lived organisms like seabirds are unlikely to increase parental effort as environmental conditions deteriorate (Drent & Daan, 1980), the signs of mismatch may be ameliorated by an increase in variety of prey composition during the chick provisioning stage. Diversification of diet compositions have been observed in Southern rockhopper penguins (Dehnhard *et al.*, 2016), crested terns (Gaglio *et al.*, 2018b), and European shag (Howells *et al.*, 2017), and may allow adults to cope with a mismatch with the availability of previously preferred prey items. However, poor foraging conditions during the breeding season have been found to negatively impact the post breeding body condition of adult seabirds (Harding *et al.*, 2011), potentially resulting in declines in adult survival over winter. In order to understand how climate change impacts upon marine food webs, more complete information on the availability of prey biomass throughout the year would be of enormous value.

There are several ways in which our methods may have caused us to underestimate the presence and consequences of trophic mismatch in global seabird

populations. We used average sea surface temperature in the 100km^2 grid surrounding each colony averaged across three month prior to the breeding period to ask whether mismatch is likely to occur in years that are warmer. However, while this allowed us to question whether generally warmer years result in more severe mismatch, without information on prey phenology it is difficult to conclude whether this proxy of the environment is accurate or even adequate. A different environmental condition (or conditions) may drive peak availability of resources, or the driver(s) may influence prey phenology in location(s) other than the waters surrounding the breeding site. The effect of the environment on trophic interactions may be much more complicated and involve a range of conditions, such as ocean stratification (Carroll *et al.*, 2015), or increasing wind speeds (Lewis *et al.*, 2015). Furthermore, although we found no evidence for signatures of T-MTM in the form of reduced breeding success, it may instead be evident through reductions in adult survival or offspring recruitment to the breeding population. However, both age at recruitment and the extent to which natal philopatry is prevalent vary within and between species and populations (Coulson, 2016; Johnston *et al.*, 2019), making it difficult to ascertain when and where individuals recruit. Future analyses that could fill the crucial knowledge gaps on prey phenology, recruitment, and adult and offspring survival and quality would be hugely beneficial.

We expanded on aspects of an existing framework for identifying the extent to which climate driven phenotypic change impacts demographic rates (McLean *et al.*, 2016) by generating predictions for the relationships between temperature, lay date, and breeding success that would be expected if trophic mismatch was both present and mediated by temperature, and in the absence of information about prey. We then used this framework to estimate these relationships across 62 seabird populations and found little evidence for T-MTM. Assuming prey are responding to temperature at

a faster rate than seabirds, this result suggests that evidence for the effects of mismatch may have been previously overstated, or alternatively, that mismatch manifests in ways other than simply a reduction in the number of chicks to fledge the nest. However, in order to fully understand how climate impacts trophic interactions and subsequent demographic rates, this crucial knowledge gap – lack of information on prey – must be addressed.

6. General Discussion

The key aims of my thesis were to identify trends and patterns in timing of breeding of seabird populations globally, to further understand the scales at which seabird breeding phenology is driven, and to identify the impacts of reproductive timing on seabird breeding success at individual, population and global levels.

In **chapter two** I collated raw data from 209 phenological time series from 61 species and all ocean basins to estimate overall global trends and patterns in breeding phenology, over time and with spring Sea Surface Temperature (sSST). I found no temporal trend in timing of breeding on average, and little true variation around this trend across populations. This implies that at least some of the significant temporal trends that have been reported may have been due to estimation error. However, in some cases, significant trends may be due to real causal effects of changes to the environment, such as Alaskan black guillemots, *Cepphus grylle mandtii* which show temporal trends in advancing their breeding phenology, a result of sea ice melting earlier now than in previous decades (Divoky *et al.*, 2015). Breeding phenology was not correlated with local SST, and again I found little evidence of true variation around the overall trend. This result implies that overall, seabirds do not adjust breeding phenology (at least not consistently in one direction) in years where sea surface temperatures are warmer or cooler than average. The results of this chapter were in broad agreement with previous studies of temporal trends in phenology that have been done on smaller spatial scales or with fewer species, and provided a more complete understanding of phenological patterns in this group of higher trophic level marine species. A *post hoc* test revealed that between-year variance in phenology is lower in populations breeding at higher latitudes and higher in species that are resident to the breeding site over winter. Between year phenological variance was phylogenetically conserved, from which I infer that species with particular life-history

traits or those that breed in more seasonal environments may be more likely to respond to (unmeasured) environmental cues.

Chapter three was influenced by my findings from chapter two that seabirds are not responding to SST but that some species are highly variable in timing from year to year. This led me to ask whether different populations may be responding to a similar environmental driver. To do this I combined data from 61 populations of 9 species in the North Atlantic, and tested the scale at which timing of breeding covaries between populations across years, an approach which has not been used to identify drivers of phenology in seabirds (or other taxa) before. I found no evidence for large scale covariance across time series', suggesting that either there is no common driver of phenology, or that populations respond to a common driver in different ways. However, there was some evidence that local conditions drive phenology, with covariance of lay dates of populations breeding at the same site or sites positioned near to one another observed in some cases. This finding suggests that future studies should focus on local scale environmental variables when identifying cues that drive breeding phenology. I found no wintering or species-specific effects, with the exception of black-legged kittiwakes, which covaried across populations from both sides of the North Atlantic and almost their whole North Atlantic breeding range. Although phenology tended to covary across years at the site level, I also found that residual variance was quite high in some populations – most notably the European shags. From this I infer that phenology is most likely driven by a suite of intrinsic (age, individual quality) and extrinsic (climate, weather, site quality) effects.

Chapters four and five built on the results from chapter two, where I found no general phenological response over time or with SST. If resources during the breeding season are advancing their phenology, as suggested by other work (Poloczanska *et al.*, 2013) this may lead seabirds to become mismatched with their

prey. However, in general the implications of this for fitness have been understudied in individual populations, making a global understanding of the extent to which climate-mediated trophic mismatch is present limited. In **chapter four** I adapted an approach presented by Reed *et al.*, (2013) to identify the extent to which trophic mismatch impacts fitness at both individual and population levels. With a 30-year time series of individual level breeding phenology, reproductive success and diet from the Isle of May population of European shags and a random regression approach, I identified that fitness implications of trophic mismatch have not worsened over time, with rising SST or with a proxy of diet change, either at the population-level or between individuals. Therefore in a species that has been proposed as potentially at risk from climate-mediated mismatch (Burthe *et al.*, 2012), I find no evidence consistent with this hypothesis

In **chapter five** I expanded on a framework for identifying the extent to which climate driven phenotypic change impacts population level demographic rates presented by McLean *et al.*, (2016) and identified five clear criteria that can be used to test for evidence that SST induced mismatch between consumers and their resources impacts negatively on breeding success. This approach will be a particularly useful first step in the cases where there is no information on lower trophic level prey, which is a common knowledge gap for studies of trophic mismatch, and is the case with seabirds. I applied these criteria to 62 populations using path analyses and mixed model approaches. I found that globally, temperatures have risen at all breeding sites (although not all of them significantly), but there is no effect of rising temperature on breeding success. Furthermore, breeding success has not declined over time for the focal populations.

Taken together the results from **chapters four and five** suggest that if trophic mismatch is present, there is no evidence that it is temperature-induced or that it is

getting worse over time. However, in years where the whole population bred early, breeding success was higher than in late years. I conclude that this may be a result of deteriorating conditions throughout the season that are linked to another unmeasured aspect of seasonal environmental change. Alternatively, in years where birds breed late they may choose to invest less in their offspring due to a trade-off between offspring growth and their own survival. By combining time series from multiple populations, this chapter allowed us to further understand that lay date is associated with breeding success, and that breeding success has not declined over time across multiple species. Furthermore, assuming prey are responding to temperature at a faster rate than seabirds, the results from these chapters suggest that evidence for mismatch effects in seabirds may have been previously overstated, or alternatively, that mismatch effects manifest in ways other than simply a reduction in the number of chicks to fledge the nest, for example by impacting chick survival or adult survival and recruitment.

6.1 Global versus individual studies

The data used in chapters two, three and five were generously contributed by researchers from numerous long-term seabird studies, and incorporated many decades worth of field seasons and effort from scientists all over the world. The highly collaborative nature of each of the chapters presented here has greatly enriched the conclusions drawn from each analysis; a result of the huge wealth of pooled knowledge from experts of each population. In addition to the benefits of collaborations, large-scale approaches to answering ecological questions have some clear strengths. They allow us to estimate overall trends and observe similarities across populations or individuals, and to identify patterns that may allow us to predict which populations or species will exhibit a certain response. Linking individual time series' in previous large-scale analyses has, for example, facilitated global

understanding of at-risk groups and regions where emerging infectious diseases are likely to rise (Jones *et al.*, 2008), and identification of species and traits which are likely to be most sensitive to habitat fragmentation (Keinath *et al.*, 2017).

Even at smaller spatial scales, combining information from individual studies can be incredibly useful. For example, by bringing together hive-specific honeybee microbiome data from across the UK, Regan *et al.*, (2018) identified characteristics of individual honeybee populations that may make them more sensitive to disease burden, and created a platform to monitor the health of British honeybees more effectively. In their key papers, Thackeray *et al.*, (2010, 2016) brought together phenological time series' from all over the UK to provide key insight as to how phenology has changed at different rates across trophic levels – a key motivation for the work in this thesis. In this thesis, combining phenological time series' across large spatial scales provided a key insight, that seabird populations generally show idiosyncratic phenological responses (chapters 2 and 3), even when sources of non-independence (e.g. when time series' come from different species', sites or time periods) and measurement error have been taken into account. It suggests that using individual studies to make general predictions about how seabirds will respond to environmental change as a whole may not be the best approach (McLean *et al.*, 2018), and highlights the need to continue long-term and individual studies to understand further what drives these idiosyncrasies across populations.

However, in generalising across species, populations or individuals, large-scale studies may overlook some of the detail and nuance specific to individual systems (Sandvik & Erikstad, 2008). A key example in this thesis comes from using general environmental predictors to explain patterns across multiple species and from many different regions. By using a very general measure of temperature in chapters 2 and 5 I was able to ask whether breeding phenology and success differ when spring

is warmer or cooler than average. However I was not able to answer questions about cues or environmental drivers that may be specific to each population. The results from Chapters 2 and 3 provide a motivation for researchers to consider local drivers across multiple species at a colony.

Another example of the merits of individual studies comes from comparison of the methods I used in chapters 4 and 5 to examine the fitness consequences of potential climate-induced trophic mismatch. By using individual level data in chapter 4 I was able to characterise the relationship between timing of breeding and fitness within a season to further understand the underlying mechanisms driving the overall population-level response. My finding that the strength and direction of selection does not vary between years allowed me to conclude that while timing of breeding is inherently important for individuals, there is no evidence that climate-mediated trophic mismatch is increasing over time, with SST or diet. By contrast, in chapter 5 I was not able to characterise the relationship between phenology and breeding success at the individual level. Therefore, while in this chapter I found no evidence that trophic mismatch is present or increasing at the population level, I cannot rule out that some individuals in these populations may suffer fitness consequences of mistimed phenology within a year. Both individual studies and large-scale generalisations therefore have their merits and both approaches should be used together to understand mechanisms of how organisms respond to environmental change across populations.

6.2 Implications and future directions

6.2.1 Information on climate drivers

In Chapters 2 and 5 I used an interpolated measure of average sea surface temperature in the three months prior to breeding as a proxy of environmental change.

This approach provided a comparable measure of the environment across populations that could be replicated in future studies, and a general measure of warm versus cool years (see further discussion of the strengths and weaknesses of global studies, above). However, in order to explain what drives phenology, a greater number of climate variables collected at a finer spatial and temporal scale are needed (Bailey & van de Pol, 2016). This presents a challenge in wide-ranging and long-lived species such as seabirds, because there are potentially a large number of environmental drivers of phenology and breeding success, the effects of which may carry over from previous seasons (Daunt *et al.*, 2006, 2014) and have substantial time lags (Thompson & Ollason, 2001). For this reason, the approach I took in chapter 3, to first address the scale at which phenology is correlated was a particularly useful first step in identifying suitable additional environmental conditions that may be included in further analyses.

Understanding what drives phenology in seabirds will ultimately benefit from the marriage of both individual and global studies. In chapter 3 I found evidence to suggest that local cues drive phenology by combining time series' from many populations, which would have been difficult when considering studies individually. However, this work was limited by the relatively small number of time series available for each species and region, and the lack of comparably detailed information across all time series' about wintering areas, migration routes and stop over locations. If further work could expand time series' to include these missing data and identify the spatiotemporal scales at which drivers of phenology act, individual studies could incorporate this information to test specific cues and drivers at the correct scale for each population. However, if each individual study identifies a different environmental driver of phenology then it will be impossible to do a future formal meta-analysis to identify trends across populations.

6.2.2 Data on lower trophic levels

One of the key elements limiting the progression of research on pelagic trophic mismatch is the shortage of data on the phenology of mid-trophic-level marine prey. I found this to be a limiting factor throughout my thesis which presents an obstacle to making robust conclusions about either the drivers of phenology (e.g., with resources acting as a determinant of breeding conditions) or the consequences of environmental change on breeding success. Because few time series' on forage fish phenology exist, in many cases we do not know when or to what extent prey availability peaks or acts as a cue, whether seabirds are sampling the complete availability of prey throughout the breeding season, or in many cases what they are bringing back to feed their chicks.

To address the lack of information on prey phenology, some studies investigating the presence of mismatch in fish-eating seabird species have used innovative proxies of prey availability, such as monitoring the date of peak attendance by individual seabirds at the breeding colony to indicate peak food availability (Gaston *et al.*, 2009); estimating the date at which forage fish in the waters around the colony reach a threshold length deemed nutritionally suitable (Burthe *et al.*, 2012); using fisheries data to test for correlation between local fish abundance and timing of breeding (Reed *et al.*, 2009); or measuring the date at which seabirds switch to a nutritionally richer prey species (Watanuki *et al.*, 2009).

In chapter 4 I used a direct measure of diet throughout the season: regurgitates from adult shags as they returned to the breeding site to provision their young. This provided a reasonable estimate of the timing and rate of change of the shift from 1+ group (principal prey) to 0 group sandeels, and therefore allowed me to identify to some extent whether the shag population had synchronised breeding with principal prey availability. In this chapter I assumed that shags were sampling from

the entire availability of prey in their environment. This assumption is reasonable because shags can feed both pelagically and benthically, disturb the substrate to look for prey, and there are no depths within their range that they cannot attain. However, I did not have information on the amount of energy expended by the adults to obtain their prey. This is poorly understood and potentially extremely relevant. Therefore, while these studies and my own work have shed some light on what may be occurring in individual populations, quantifying the phenology, abundance and energy composition of total available prey throughout the breeding season and between years remains a significant challenge.

The ideal resource dataset would incorporate information on the distribution (both vertically within the water column and the patterns of migration), abundance, and breeding phenology of prey, as well as data on species diversity throughout the season, and information on how the environment influences growing period, energy content and size of forage fish. This would allow us to not only understand the extent to which prey availability during the breeding season impacts breeding success, but also shed light on potential cues that fish might provide earlier in the season. In woodland systems, where mismatch and its impacts have been well-studied, the ability to monitor food availability over winter (Reed *et al.*, 2013a) and throughout the onset of spring by means of directly sampling insect phenology (Franks *et al.*, 2017), caterpillar frass (Reed *et al.*, 2013a; Burgess *et al.*, 2018), plant phenology and adult and chick passerine diet (Shutt, 2018) has provided researchers with a better understanding of the shape of peaks in food availability, and the cues these prey may provide.

However, in the absence of appropriate techniques to collect this information in the marine environment, an alternative approach would be to sample the diet of chicks throughout the season and measure the size, energy content, age and species

diversity of the prey (combining techniques used in Wanless *et al.*, 2005; Burthe *et al.*, 2012; Howells *et al.*, 2017, 2018). This would allow researchers to reconstruct the peaks in energy composition of the diet in a very useful proxy of energy availability, although this would rely on the assumption that parents are sampling their prey from all available biomass in the environment, which I made in Chapter 4. Linking this with information on adult foraging effort, i.e. how far they travel and how much energy they expend to catch prey, and body condition before, during and after the breeding season would be invaluable in further understanding the patterns of availability of prey across years, although this would be challenging data to collect across many species.

6.2.3 Trophic mismatch

The difficulties encountered when testing for evidence of asynchrony across trophic levels in the absence of information on prey and drivers of phenology are evident. However, under the assumption that suggested asynchrony is present between seabirds and their prey (Thackeray *et al.*, 2010; Poloczanska *et al.*, 2013, 2016), the final aim of my thesis was to test for evidence that temperature-mediated trophic mismatch has consequences for breeding success. There may be several biological explanations, apart from the coarse scale of the environmental variable, as to why I did not observe any evidence for T-MTM. Firstly, the chick rearing period is generally quite long in many seabird species (Schreiber & Burger, 2002). This suggests that timing reproduction to coincide with a period of seasonal resources may not be as important for some seabird species, which may be able to respond to poor foraging conditions in alternative ways (Wojczulanis-Jakubas *et al.*, 2018). This could partially explain the differences in observed mismatch and declines in breeding success between seabirds and well-studied terrestrial systems involving temperate woodland passerines (Charmantier *et al.*, 2008; Reed *et al.*, 2013a; Franks *et al.*, 2017), which generally have nestling stages of 2-3 weeks. The breeding success of temperate

passerines may be much more reliant on seasonal food peaks and therefore more susceptible to trophic mismatch.

Further explanation as to why mismatch was not observed could be explained by the fact that many species of seabird have the capacity to store lipids during the nestling phase for use when resources are not as plentiful (Riou & Hamer, 2010). Several hypotheses have been suggested as to why this may be the case, including as a mechanism to cope with periods of prolonged fasting during poor foraging conditions experienced by seabird parents (Lack, 1968), and to increase post-fledging survival during a period when chicks are learning to forage alone (Phillips & Hamer, 1999). This ability to store resources may mean that chicks can ameliorate the negative implications of periods of food shortage experienced during prolonged nestling phases, making them less susceptible to mismatch than species with faster life histories, such as passerines. However, an ability to store resources may also mean that mismatch does not simply manifest in the number of chicks to successfully leave the nest, but in the condition the chicks are in when they fledge. Chicks may be able to buffer mismatch conditions, but fledge at a lighter mass than if conditions have been good throughout the nestling phase (Perrins *et al.*, 1973). The post-fledging period is critical for seabirds, with high levels of mortality occurring in the first year (Schreiber & Burger, 2002), particularly in lighter birds. It is therefore plausible that it is post-fledging survival, perhaps mediated by fledgling mass that is impacted during periods of mismatch with lower trophic level prey.

The demographic rate I considered in chapters four and five was the average number of chicks to successfully fledge the nest (including failed nests) per year. This determinant of success was useful as collaborators across multiple studies could readily provide it (and it was of a standard quality), and informative as it gives us a proximate indication of how good or bad the breeding season was. However, if trophic

mismatch is apparent in seabirds, it is clear that the effects may be more wide-ranging than simply influencing the average number of chicks to fledge the nest. Future studies of the consequences of potential mismatch would benefit from the inclusion of information on body condition of both chicks (e.g. wing length or mass at fledging) and adults (e.g. stress hormones or mass), and adult and offspring survival, which could be measured by mark-recapture studies or population growth rate. It will be difficult to include the latter in future studies, as population sizes are not always counted each year, and may only include breeding pairs, which precludes any analysis or comparison of the number of breeders versus non-breeding adults, which may change in proportion between matched and mismatched years. Furthermore, population decline mediated by stress experienced by adults during a difficult breeding year would be difficult to disentangle from other drivers, such as decline from predation or fisheries (Oro, 2014). To assess the level at which mismatch impacts offspring recruitment also presents a challenge. It takes many years for juveniles to reach recruitment age (Schreiber & Burger, 2002), and the degree of natal philopatry exhibited by offspring remains unclear in many species (Coulson, 2016). More detailed tracking data in the years before recruitment to identify where juveniles settle, and population counts across multiple breeding sites would allow us to quantify the extent to which potential mismatch impacts on population growth rates.

6.3 Concluding remarks

In such a complex system as marine food webs, it is likely that a spectrum of conditions, constraints and cues drive timing of breeding in these long lived, socially complex (e.g. with high breeding site fidelity and strong pair bonds), and wide ranging marine higher predators. Without long term information on the potential intrinsic and extrinsic drivers across a multitude of spatiotemporal scales, successfully identifying specific drivers of phenology in seabirds presents an enormous challenge. The

consequences of (mis)timing may not become evident until many years after the fact, and may manifest in many subtle ways other than declines in breeding success or survival. Whilst the analyses in this thesis have furthered our understanding of the sensitivity of seabird breeding phenology to the environment, the scale at which phenology is driven and the consequences of potential mismatch for breeding success, the extent to which climate-mediated trophic mismatch between seabirds and their prey is present is still unclear. Deeper knowledge of a combination of the extent to which natal philopatry is present in populations and ages at which individuals reach maturity would allow us to understand the far reaching effects of phenology and mismatch. The effect of climate on breeding phenology and its consequences will be difficult to disentangle from other more transparent issues. Observed decreases in survival and variation in breeding success may indeed be linked to mismatch with prey, but also with the decline in abundance of resources, the increase in contaminants, and the continuing development of the marine environment for human use. A substantial amount of further research is required to identify the impacts of environmental change on trophic interactions within complex food webs.

Bibliography

Aebischer NJ (1985) *Aspects of the biology of the shag: (Phalacrocorax aristotelis)*. Durham University.

Afán I, Chiaradia A, Forero MG, Dann P, Ramírez F (2015) A novel spatio-temporal scale based on ocean currents unravels environmental drivers of reproductive timing in a marine predator. *Proceedings. Biological sciences / The Royal Society*, **282**, 20150721-.

Ainley D., Boekelheide R. (1990) *Seabirds of the Farallon Islands: ecology, dynamics and structure of an upwelling-system community*. Stanford University Press, 306–338 pp.

Amélineau F, Grémillet D, Bonnet D, Bot T Le, Fort J (2016) Where to Forage in the Absence of Sea Ice? Bathymetry As a Key Factor for an Arctic Seabird. *PLoS ONE*, **11**, 1–19.

Aoki S, Campos E, Chambers D et al. (2013) Climate Change 2013: The Physical Science Basis (Chapter 3). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 62.

Arnott SA, Ruxton GD (2002) Sandeel recruitment in the North Sea: Demographic, climatic and trophic effects. *Marine Ecology Progress Series*, **238**, 199–210.

Asch RG (2015) Climate change and decadal shifts in the phenology of larval fishes in the California Current ecosystem. *Proceedings of the National Academy of Sciences*, 4065–4074.

Bailey LD, van de Pol M (2016) Climwin: An R Toolbox for Climate Window Analysis. *PLoS ONE*, **11**, 1–27.

Barbraud C, Weimerskirch H (2006) Antarctic birds breed later in response to climate change. *Proceedings of the National Academy of Sciences of the United States of America*, **103**, 6248–6251.

Barlow EJ, Daunt F, Wanless S, Reid JM (2013) Estimating dispersal distributions at multiple scales: Within-colony and among-colony dispersal rates, distances and directions in European Shags Phalacrocorax aristotelis. *Ibis*, **155**, 762–778.

Becker PH, Schmaljohann H, Riechert J, Wagenknecht G, Zajková Z, González-Solís J (2016) Common Terns on the East Atlantic Flyway: temporal–spatial distribution during the non-breeding period. *Journal of Ornithology*, **157**, 927–940.

Behrenfeld MJ, O'Malley RT, Siegel DA et al. (2006) Climate-driven trends in contemporary ocean productivity. *Nature*, **444**, 752–755.

Bertram DF, Mackas DL, McKinnell SM (2001) The seasonal cycle revisited: interannual variation and ecosystem consequences. *Progress In Oceanography*, **49**, 283–307.

Bjørnstad ON, Ims RA, Lambin X (1999) Spatial population dynamics: analyzing patterns and processes of population synchrony. *Trends in Ecology and Evolution*, **14**, 427–432.

Bogdanova MI, Wanless S, Harris MP et al. (2014) Among-year and within-population variation in foraging distribution of European shags Phalacrocorax aristotelis over two decades: Implications for marine spatial planning. *Biological Conservation*, **170**, 292–299.

Bogdanova MI, Butler A, Wanless S et al. (2017) Multi-colony tracking reveals spatio-temporal variation in carry-over effects between breeding success and winter movements in a pelagic seabird. *Marine Ecology Progress Series*, **578**, 167–181.

Bond AL, Jones IL, Sydeman WJ, Major HL, Minobe S, Williams JC, Byrd GV (2011) Reproductive success of planktivorous seabirds in the North Pacific is related to ocean climate on decadal scales. *Marine Ecology Progress Series*, **424**, 205–218.

Boulcott P, Wright PJ (2008) Critical timing for reproductive allocation in a capital breeder: Evidence from sandeels. *Aquatic Biology*, **3**, 31–40.

Box G, Jenkins G, Reinsel G (1990) *Time Series Analysis, Forecasting and Control*, Fifth Edit edn. Wiley.

Bradshaw WE, Holzapfel CM (2010) Light, time, and the physiology of biotic response to rapid climate change in animals. *Annual Review of Physiology*, **72**, 147–166.

Brierley AS, Kingsford MJ (2009) Impacts of climate change on marine organisms and ecosystems. *Current Biology*, **19**, R602–R614.

Burgess MD, Smith KW, Evans KL et al. (2018) Tritrophic phenological match–mismatch in space and time. *Nature Ecology & Evolution*, **2**, 970–975.

Burr ZM, Varpe Ø, Anker-Nilssen T et al. (2016) Later at higher latitudes: large-scale variability in seabird breeding timing and synchronicity. *Ecosphere*, **7**, 1–12.

Burrows MT, Schoeman DS, Buckley LB et al. (2011) The pace of shifting climate in marine and terrestrial ecosystems. *Science*, **334**, 652–655.

Burthe S, Daunt F, Butler A et al. (2012) Phenological trends and trophic mismatch across multiple levels of a North Sea pelagic food web. *Marine Ecology Progress Series*, **454**, 119–133.

Cabot D, Nisbet I (2013) *Terns*. HarperCollins, London.

Carroll M, Butler A, Owen E et al. (2015) Effects of sea temperature and stratification changes on seabird breeding success. *Climate Research*, **66**, 75–89.

Chambers LE, Cullen JM, Coutin PC, Dann P (2009) Predicting onset and success of breeding in little penguins Eudyptula minor from ocean temperatures. *Marine Ecology Progress Series*, **378**, 269–278.

Chambers LE, Altwegg R, Barbraud C et al. (2013) Phenological changes in the Southern hemisphere. *PLoS ONE*, **8**.

Chambers LE, Dann P, Cannell B, Woehler EJ (2014) Climate as a driver of phenological change in southern seabirds. *International Journal of Biometeorology*, **58**, 603–612.

Charmantier A, McCleery RH, Cole LR, Perrins C, Kruuk LEB, Sheldon BC (2008) Adaptive phenotypic plasticity in response to climate change in a wild bird population. *Science*, **320**, 800–3.

Chastel O, Weimerskirch H, Jouventin P (1993) High annual variability in reproductive success and survival of an Antarctic seabird, the snow petrel Pagodroma nivea - A 27-year study. *Oecologia*, **94**, 278–285.

Chavez FP, Messié M (2009) A comparison of Eastern Boundary Upwelling Ecosystems. *Progress in Oceanography*, **83**, 80–96.

Cheung WWL, Watson R, Pauly D (2013) Signature of ocean warming in global fisheries catch. *Nature*, **497**, 365–368.

Christensen-Dalsgaard S, May R, Lorentsen S-H (2018) Taking a trip to the shelf : Behavioral decisions are mediated by the proximity to foraging habitats in the black-legged kittiwake. *Ecology and Evolution*, **8**, 866–878.

Clutton-Brock TH (1988) *Reproductive success: studies of individual variation in contrasting breeding systems*. The University of Chicago Press.

Cohen JM, Lajeunesse MJ, Rohr JR (2018) A global synthesis of animal phenological responses to climate change. *Nature Climate Change*, **8**, 224–228.

Collett D (2002) *Modelling Binary Data*, Second Edi edn (ed Hall/CRC C and).

Coulson JC (2016) A Review of Philopatry in Seabirds and Comparisons with Other Waterbird Species. *Journal of the Waterbird Society*, **39**, 229–240.

Crick HQP (2004) The impact of climate change on birds. *Ibis*, **146**, 48–56.

Croxall JP, Trathan PN, Murphy EJ (2002) Environmental change and Antarctic seabird populations. *Science*, **297**, 1510–1514.

Croxall JP, Butchart SHM, Lascelles B, Stattersfield AJ, Sullivan B, Symes A, Taylor P (2012) Seabird conservation status, threats and priority actions: a global assessment. *Bird Conservation International*, **22**, 1–34.

Cubaynes S, Jr PFD, Schreiber EA, Gimenez O (2011) To breed or not to breed : a seabird ' s response to extreme climatic events. *Biology Letters*, **7**, 303–306.

Cushing DH (1990) *Plankton production and year-class strength in fish populations: An update of the match/mismatch hypothesis*, Vol. 26. Advances in Marine Biology, Academic Press, 249–293 pp.

Dalleau M, Ciccione S, Mortimer JA, Garnier J, Benhamou S, Bourjea J (2012) Nesting Phenology of Marine Turtles: Insights from a Regional Comparative Analysis on Green Turtle (Chelonia mydas). *PLoS ONE*, **7**.

Daunt F (2000) *Age specific reproductive performance in the shag (Phalacrocorax aristotelis)*. University of Glasgow.

Daunt F, Harris MP, Monaghan P (1999) Experimental evidence that age-specic reproductive success is independent of environmental effects. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **266**, 1489–1493.

Daunt F, Afanasyev V, Silk JRD, Wanless S (2006) Extrinsic and intrinsic determinants of winter foraging and breeding phenology in a temperate

seabird. *Behavioral Ecology and Sociobiology*, **59**, 381–388.

Daunt F, Wanless S, Harris MP, Money L, Monaghan P (2007) Older and wiser: improvements in breeding success are linkled to better foraging performance in European shags. *Functional Ecology*, **21**, 561–567.

Daunt F, Wanless S, Greenstreet SPR, Jensen H, Hamer KC, Harris MP (2008) The impact of the sandeel fishery closure on seabird food consumption, distribution, and productivity in the northwestern North Sea. *Canadian Journal of Fisheries and Aquatic Sciences*, **65**, 362–381.

Daunt F, Reed TE, Newell M, Burthe S, Phillips RA, Lewis S, Wanless S (2014) Longitudinal bio-logging reveals interplay between extrinsic and intrinsic carry-over effects in a long-lived vertebrate. *Ecology*, **95**, 2077–2083.

Davoren GK, Penton P, Burke C, Montevecchi WA (2012) Water temperature and timing of capelin spawning determine seabird diets. *ICES Journal of Marine Science*, **69**, 1234–1241.

Dawson A (2008) Control of the annual cycle in birds: endocrine constraints and plasticity in response to ecological variability. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 1621–1633.

Dawson A, King VM, Bentley GE, Ball GF (2001) Photoperiodic control of seasonality in birds. *Journal of Biological Rhythms*, **16**, 365–380.

Dawson TP, Jackson ST, House JI, Prentice IC, Mace GM (2011) Beyond predictions: biodiversity conservation in a changing climate. *Science (New York, N.Y.)*, **332**, 53–58.

Dehnhard N, Ludynia K, Masello JF, Voigt CC, McGill RAR, Quillfeldt P (2016) Plasticity in foraging behaviour and diet buffers effects of inter-annual environmental differences on chick growth and survival in southern rockhopper penguins Eudyptes chrysocome chrysocome. *Polar Biology*, **39**, 1627–1641.

Van Deurs M, Van Hal R, Tomczak MT, Jónasdóttir SH (2009) Recruitment of lesser sandeel Ammodytes marinus in relation to density dependence and zooplankton composition. *Marine Ecology Progress Series*, **381**, 249–258.

Divoky GJ, Lukacs PM, Druckenmiller ML (2015) Effects of recent decreases in arctic sea ice on an ice-associated marine bird. *Progress in Oceanography*, **136**, 151–161.

Dobson FS, Becker PH, Arnaud CM, Bouwhuis S, Charmantier A (2017) Plasticity results in delayed breeding in a long-distant migrant seabird. *Ecology and Evolution*, **7**, 3100–3109.

Doiron M, Gauthier G, Lévesque E (2015) Trophic mismatch and its effects on the growth of young in an Arctic herbivore. *Global Change Biology*, **21**, 4364–4376.

Drent RH, Daan S (1980) The Prudent Parent: Energetic Adjustments in Avian Breeding. *Ardea*, **55**, 225–252.

Dunn PO, Møller AP (2014) Changes in breeding phenology and population size of birds. *Journal of Animal Ecology*, **83**, 729–739.

Durant JM, Hjermann D, Anker-Nilssen T, Beaugrand G, Mysterud A, Pettorelli N, Stenseth NC (2005) Timing and abundance as key mechanisms affecting

trophic interactions in variable environments. *Ecology Letters*, **8**, 952–958.

Durant JM, Hjermann DO, Ottersen G, Stenseth NC (2007) Climate and the match or mismatch between predator requirements and resource availability. *Climate Research*, **33**, 271–283.

Edwards M, Richardson AJ (2004) Impact of climate change on marine pelagic phenology and trophic mismatch. *Nature*, **430**, 881–884.

Eerkes-Medrano D, Fryer RJ, Cook KB, Wright PJ (2017) Are simple environmental indicators of food web dynamics reliable: Exploring the kittiwake-temperature relationship. *Ecological Indicators*, **75**, 36–47.

Egevang C, Stenhouse IJ, Phillips RA, Petersen A, Fox JW, Silk JRD (2010) Tracking of Arctic terns Sterna paradisaea reveals longest animal migration. *Proceedings of the National Academy of Sciences*, **107**, 2078–2081.

Erikstad KE, Fauchald P, Tveraa T, Steen H (1998) On the Cost of Reproduction in Long-Lived Birds: The Influence of Environmental Variability. *Ecology*, **79**, 1781–1788.

Evans JC, Dall SRX, Bolton M, Owen E, Votier SC (2016) Social foraging European shags: GPS tracking reveals birds from neighbouring colonies have shared foraging grounds. *Journal of Ornithology*, **157**, 23–32.

Fayet AL, Freeman R, Anker-Nilssen T et al. (2017) Ocean-wide Drivers of Migration Strategies and Their Influence on Population Breeding Performance in a Declining Seabird. *Current Biology*, **27**, 1–8.

Forrest J, Miller-rushing AJ (2010) Toward a synthetic understanding of the role of phenology in ecology and evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **365**, 3101–3112.

Fortin M, Bost C, Maes P, Barbraud C (2013) The demography and ecology of the European shag Phalacrocorax aristotelis in Mor Braz, France. *Aquatic Living Resources*, **185**, 179–185.

Franks SE, Pearce-Higgins JW, Atkinson S et al. (2017) The sensitivity of breeding songbirds to changes in seasonal timing is linked to population change but cannot be directly attributed to the effects of trophic asynchrony on productivity. *Global Change Biology*, 1–15.

Frederiksen M, Harris MP, Daunt F, Rothery P, Wanless S (2004) Scale-dependent climate signals drive breeding phenology of three seabird species. *Global Change Biology*, **10**, 1214–1221.

Frederiksen M, Wright PJ, Harris MP, Mavor RA, Heubeck M, Wanless S (2005) Regional patterns of kittiwake Rissa tridactyla breeding success are related to variability in sandeel recruitment. *Marine Ecology Progress Series*, **300**, 201–211.

Frederiksen M, Edwards M, Richardson AJ, Halliday NC, Wanless S (2006) From plankton to top predators: Bottom-up control of a marine food web across four trophic levels. *Journal of Animal Ecology*, **75**, 1259–1268.

Frederiksen M, Elston D a., Edwards M, Mann AD, Wanless S (2011) Mechanisms of long-term decline in size of lesser sandeels in the North Sea explored using a growth and phenology model. *Marine Ecology Progress Series*, **432**, 137–

147.

Frederiksen M, Moe B, Daunt F et al. (2012) Multicolony tracking reveals the winter distribution of a pelagic seabird on an ocean basin scale. *Diversity and Distributions*, **18**, 530–542.

Frederiksen M, Descamps S, Erikstad KE et al. (2016) Migration and wintering of a declining seabird, the thick-billed murre Uria lomvia, on an ocean basin scale: Conservation implications. *Biological Conservation*, **200**, 26–35.

Furness RW, Tasker ML (2000) Seabird-fishery interactions : quantifying the sensitivity of seabirds to reductions in sandeel abundance, and identification of key areas for sensitive seabirds in the North Sea. *Marine Ecology Progress Series*, **202**, 253–264.

Gaglio D, Cook TR, Sherley RB, Ryan PG (2018a) How many can you catch? Factors influencing the occurrence of multi-prey loading in provisioning Greater Crested Terns§. *Ostrich*, **89**, 145–149.

Gaglio D, Cook TR, McInnes A, Sherley RB, Ryan PG (2018b) Foraging plasticity in seabirds: A non-invasive study of the diet of greater crested terns breeding in the Benguela region. *PLoS ONE*, **13**, 1–20.

Gaston AJ, Gilchrist HG, Mallory ML, Smith P a. (2009) Changes in Seasonal Events, Peak Food Availability, and Consequent Breeding Adjustment in a Marine Bird: A Case of Progressive Mismatching. *The Condor*, **111**, 111–119.

Gelman A (2006) Prior distributions for variance parameters in hierarchical models (Comment on Article by Browne and Draper). *Bayesian Analysis*, **1**, 515–534.

Gelman A, Van Dyk DA, Huang Z, Boscardin WJ (2008) Using redundant parameterizations to fit hierarchical models. *Journal of Computational and Graphical Statistics*, **17**, 95–122.

Gienapp P, Reed TE, Visser ME (2014) Why climate change will invariably alter selection pressures on phenology. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20141611–20141611.

Goutte A, Angelier F, Chastel CC et al. (2010) Stress and the timing of breeding: glucocorticoid-luteinizing hormones relationships in an arctic seabird. *General and comparative endocrinology*, **169**, 108–116.

Grecian WJ, Witt MJ, Attrill MJ et al. (2016) Seabird diversity hotspot linked to ocean productivity in the Canary Current Large Marine Ecosystem. *Biology Letters*, **12**.

Grémillet D, Boulinier T (2009) Spatial ecology and conservation of seabirds facing global climate change: a review. *Marine Ecology Progress Series*, **391**, 121–137.

Grémillet D, Charmantier A (2010) Shifts in phenotypic plasticity constrain the value of seabirds as ecological indicators of marine ecosystems. *Ecological Applications*, **20**, 1498–1503.

Grémillet D, Lewis S, Drapeau L et al. (2008) Spatial match-mismatch in the Benguela upwelling zone: Should we expect chlorophyll and sea-surface temperature to predict marine predator distributions? *Journal of Applied Ecology*, **45**, 610–621.

Grémillet D, Ponchon A, Paleczny M, Palomares M-LD, Karpouzi V, Pauly D (2018) Persisting Worldwide Seabird-Fishery Competition Despite Seabird Community Decline. *Current Biology*, 1–5.

Grist H, Daunt F, Wanless S et al. (2014) Site fidelity and individual variation in winter location in partially migratory European shags. *PLoS One*, **9**, 98562.

Grist H, Daunt F, Wanless S, Burthe SJ, Newell MA, Harris MP, Reid JM (2017) Reproductive performance of resident and migrant males, females and pairs in a partially migratory bird. *Journal of Animal Ecology*, **86**, 1010–1021.

Gwinner E (1996) Circannual clocks in avian reproduction and migration. *Ibis*, **138**, 47–63.

Hackett SJ, Kimball RT, Reddy S et al. (2008) A Phylogenomic Study of Birds Reveals Their Evolutionary History. *Science*, **320**, 1763–1768.

Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, **33**, 1–22.

Hadfield JD, Nakagawa S (2010) General quantitative genetic methods for comparative biology: Phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *Journal of Evolutionary Biology*, **23**, 494–508.

Haest B, Hüppop O, Bairlein F (2018a) The influence of weather on avian spring migration phenology: What, where, and when? *Global Change Biology*, **24**, 5769–5788.

Haest B, Hüppop O, Bairlein F (2018b) Challenging a 15-year-old claim: The North Atlantic Oscillation index as a predictor of spring migration phenology of birds. *Global Change Biology*, **24**, 1523–1537.

Harding AMA, Welcker J, Steen H et al. (2011) Adverse foraging conditions may impact body mass and survival of a high Arctic seabird. *Oecologia*, **167**, 49–59.

Harris MP, Wanless S (1991) The importance of the lesser sandeel Ammodytes marinus in the diet of the shag Phalacrocorax aristotelis. *Ornis Scandinavica*, **22**, 375–382.

Harris MP, Newell M, Daunt F, Speakman JR, Wanless S (2008) Snake Pipefish Entelurus aequoreus are poor food for seabirds. *Ibis*, **150**, 413–415.

Harrison XA, Blount JD, Inger R, Norris DR, Bearhop S (2011) Carry-over effects as drivers of fitness differences in animals. *Journal of Animal Ecology*, **80**, 4–18.

Hicks O, Burthe SJ, Daunt F et al. (2018) The energetic cost of parasitism in a wild population. *Proceedings of the Royal Society B: Biological Sciences*, **285**, 20180489.

Housworth EA, Martins EP, Lynch M (2004) The phylogenetic mixed model. *The American Naturalist*, **163**, 84–96.

Howells RJ, Burthe SJ, Green J et al. (2017) From days to decades: short- and long-term variation in environmental conditions affect diet composition of a marine top-predator. *Marine Ecology Progress Series*, **583**, 227–242.

Howells RJ, Burthe SJ, Green JA et al. (2018) Pronounced long‐term trends in year

-round diet composition of the European shag Phalacrocorax aristotelis. *Marine Biology*, **165**, 1–15.

Hughes L (2000) Biological consequences of global warming: is the signal already apparent? *Trends in Ecology & Evolution*, **15**, 56–61.

Iler AM, Inouye DW, Schmidt NM, Høye TT (2016) Detrending phenological time series improves climate-phenology analyses and reveals evidence of plasticity. *Ecology*, **98**, 647–655.

Ives AR, Midford PE, Garland T (2007) Within-species variation and measurement error in phylogenetic comparative methods. *Systematic biology*, **56**, 252–270.

Janetos AC, Chen RS, Arndt D, Kenney MA (2012) *National Climate Assessment Indicators: Background, Development, & Examples*. 59 pp.

Jarillo J, Sæther BE, Engen S, Cao FJ (2018) Spatial scales of population synchrony of two competing species: effects of harvesting and strength of competition. *Oikos*, **127**, 1459–1470.

Jenouvrier S, Barbraud, Christophe, Weimerskirch H (2005) Long-term contrasted responses to climate of two Antarctic seabird species. *Ecology*, **86**, 2889–2903.

Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO (2012) The global diversity of birds in space and time. *Nature*, **491**, 444–448.

Johnston DT, Furness RW, Robbins AMC, Taggart MA, Tyler G, Masden EA (2019) Variations in Black Guillemot *Cepphus grylle* natal dispersal and philopatry across Britain and Ireland. *Bird Study*, **3657**, 1–6.

Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P (2008) Global trends in emerging infectious diseases. *Nature*, **451**, 990–993.

Jones T, Parrish JK, Peterson WT et al. (2018) Massive Mortality of a Planktivorous Seabird in Response to a Marine Heatwave. *Geophysical Research Letters*, **45**, 3193–3202.

Keinath DA, Doak DF, Hodges KE et al. (2017) A global analysis of traits predicting species sensitivity to habitat fragmentation. *Global Ecology and Biogeography*, **26**, 115–127.

Keogan K, Daunt F, Wanless S et al. (2018) Global phenological insensitivity to shifting ocean temperatures among seabirds. *Nature Climate Change*, **8**, 313–318.

Kerby JT, Post E (2013) Capital and income breeding traits differentiate trophic match−mismatch dynamics in large herbivores. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **368**, 20120484.

Lack D (1968) *Ecological adaptations for breeding in birds*. Methuen London.

Lahoz-Monfort JJ, Morgan BJT, Harris MP, Daunt F, Wanless S, Freeman SN (2013) Breeding together: modeling synchrony in productivity in a seabird community. *Ecology*, **94**, 3–10.

Lewis S, Sherratt T., Hamer K., Wanless S (2001a) Evidence of intra-specific competition for food in a pelagic seabird. *Nature*, **412**, 816–819.

Lewis S, Wanless S, Wright PJ, Harris MP, Bull J, Elston DA (2001b) Diet and

breeding performance of black-legged kittiwakes Rissa tridactyla at a North Sea colony. *Marine Ecology Progress Series*, **221**, 277–284.

Lewis S, Phillips RA, Burthe SJ, Wanless S, Daunt F (2015) Contrasting responses of male and female foraging effort to year-round wind conditions. *Journal of Animal Ecology*, **84**, 1490–1496.

Lindestad O, Wheat CW, Nylin S, Gotthard K (2018) Local adaptation of photoperiodic plasticity maintains life cycle variation within latitudes in a butterfly. *Ecology*, **100**, 1–14.

Longhurst A (2006) *Ecological Geography of the Sea.*, 2nd Editio edn. Academic Press, San Diego, 560 pp.

Lorentsen SH, Anker-Nilssen T, Erikstad KE, Røv N (2015a) Forage fish abundance is a predictor of timing of breeding and hatching brood size in a coastal seabird. *Marine Ecology Progress Series*, **519**, 209–220.

Lorentsen S, Anker-Nilssen T, Erikstad K, Røv N (2015b) Forage fish abundance is a predictor of timing of breeding and hatching brood size in a coastal seabird. *Marine Ecology Progress Series*, **519**, 209–220.

Love OP, Gilchrist HG, Descamps S, Semeniuk CAD, Bêty J (2010) Pre-laying climatic cues can time reproduction to optimally match offspring hatching and ice conditions in an Arctic marine bird. *Oecologia*, **164**, 277–286.

Lynch HJ, Fagan WF, Naveen R, Trivelpiece SG, Trivelpiece WZ (2012) Differential advancement of breeding phenology in response to climate may alter staggered breeding among sympatric pygoscelid penguins. *Marine Ecology Progress Series*, **454**, 135–145.

Mackas DL, Greve W, Edwards M et al. (2012) Changing zooplankton seasonality in a changing ocean: Comparing time series of zooplankton phenology. *Progress in Oceanography*, **97**–**100**, 31–62.

Mallory ML, Robinson SA, Hebert CE, Forbes MR (2010) Seabirds as indicators of aquatic ecosystem conditions: A case for gathering multiple proxies of seabird health. *Marine Pollution Bulletin*, **60**, 7–12.

Marrot P, Charmantier A, Blondel J, Garant D (2018) Current Spring Warming as a Driver of Selection on Reproductive Timing in a Wild Passerine. *Journal of Animal Ecology*, **87**, 754–764.

McLean N, Lawson CR, Leech DI, van de Pol M (2016) Predicting when climate-driven phenotypic change affects population dynamics. *Ecology Letters*, **19**, 595–608.

McLean N, van der Jeugd HP, van de Pol M (2018) High intra-specific variation in avian body condition responses to climate limits generalisation across species. *Plos One*, **13**, e0192401.

McNamara JM, Barta Z, Klaassen M, Bauer S (2011) Cues and the optimal timing of activities under environmental changes. *Ecology Letters*, **14**, 1183–1190.

Mesquita MDS, Erikstad KE, Sandvik H et al. (2015) There is more to climate than the North Atlantic Oscillation: a new perspective from climate dynamics to explain the variability in population growth rates of a long-lived seabird. *Frontiers in Ecology and Evolution*, **3**, 1–14.

Miller-Rushing AJ, Høye TT, Inouye DW, Post E (2010) The effects of phenological mismatches on demography. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **365**, 3177–3186.

Moe B, Brunvoll S, Mork D, Brobakk T., Bech C (2004) Developmental plasticity of physiology and morphology in diet-restricted European shag nestlings (Phalacrocorax aristotelis). *Journal of Experimental Biology*, **207**, 4067–4076.

Moe B, Stempniewicz L, Jakubas D et al. (2009) Climate change and phenological responses of two seabird species breeding in the high-Arctic. *Marine Ecology Progress Series*, **393**, 235–246.

Møller AP, Rubolini D, Lehikoinen E (2008) Populations of migratory bird species that did not show a phenological response to climate change are declining. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 16195–16200.

Mostello CS, Nisbet ICT, Oswald SA, Fox JW (2014) Non-breeding season movements of six North American Roseate Terns Sterna dougallii tracked with geolocators. *Seabird*, **27**, 1–21.

Moussus J-P, Clavel J, Jiguet F, Julliard R (2011) Which are the phenologically flexible species? A case study with common passerine birds. *Oikos*, **120**, 991–998.

Nakagawa S, Santos ESA (2012) Methodological issues and advances in biological meta-analysis. *Evolutionary Ecology*, **26**, 1253–1274.

Nakagawa S, Poulin R, Mengersen K, Reinhold K, Engqvist L, Lagisz M, Senior AM (2015) Meta-analysis of variation: ecological and evolutionary applications and beyond. *Methods in Ecology and Evolution*, **6**, 143–152.

Newell M, Wanless S, Harris MP, Daunt F (2015) Effects of an extreme weather event on seabird breeding success at a North Sea colony. *Marine Ecology Progress Series*, **532**, 257–268.

Newton I (1989) *Lifetime reproduction in birds*. London Academic Press.

Nisbet ICT, Arnold JM, Oswald SA, Pyle P, Pattern MA (2017) Common tern Sterna hirundo. No. 618 in The Birds of North America (P.G. Rodewald, Ed.). *Ithaca: Cornell Lab of Ornithology; Retrieved from the Birds of North America: https://birdsna-org.bnaproxy.birds.cornell.edu/Species-Account/bna/species/comter*.

Oksanen J, Blanchet FG, Friendly M et al. (2018) vegan: Community Ecology Package. *R package version 2.5-2.*

Oppel S, Bolton M, Carneiro APB et al. (2018) Spatial scales of marine conservation management for breeding seabirds. *Marine Policy*, **98**, 37–46.

Oro D (2014) Seabirds and climate: knowledge, pitfalls, and opportunities. *Frontiers in Ecology and Evolution*, **2**, 1–12.

Österblom H, Olsson O, Blenckner T, Furness RW (2008) Junk-food in marine ecosystems. *Oikos*, **117**, 967–977.

Pagel M, Lutzoni F (2002) *Accounting for phylogenetic uncertainty in comparative studies of evolution and adaptation*. Springer Berlin Heidelberg, 148–162 pp.

Paleczny M, Hammill E, Karpouzi V, Pauly D (2015) Population Trend of the World's Monitored Seabirds, 1950-2010. *Plos One*, **10**, e0129342.

Parmesan C (2006) Ecological and Evolutionary Responses to Recent Climate Change. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 637–669.

Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**, 37–42.

Passuni G, Barbraud C, Chaigneau A et al. (2015) Seasonality in marine ecosystems: Peruvian seabirds, anchovy and oceanographic conditions. *Ecology*, **97**, 182–193.

Perrins CM (1970) The timing of birds' breeding seasons. *Ibis*, **112**, 242–255.

Perrins CM, Harris MP, Britton CK (1973) Survival of Manx shearwater. *Ibis*, **115**, 535–548.

Petchey OL, Pontarp M, Massie TM et al. (2015) The ecological forecast horizon, and examples of its uses and determinants. *Ecology Letters*, **18**, 597–611.

Phillips RA, Hamer KC (1999) Lipid reserves, fasting capability and the evolution of nestling obesity in procellariiform seabirds. *Proceedings of the Royal Society B: Biological Sciences*, **266**, 1329–1334.

Piatt I, Sydeman W (2007) Seabirds as indicators of marine ecosystems. *Marine Ecology Progress Series*, **352**, 199–204.

Pinheiro J, Bates D (2000) *Mixed-Effects Models in S and S-PLUS*.

Poethke HJ, Hovestadt T, Mitesser O (2016) The evolution of optimal emergence times: Bet hedging and the quest for an ideal free temporal distribution of individuals. *Oikos*, **125**, 1647–1656.

van de Pol M, Verhulst S (2006) Age-Dependent Traits : A New Statistical Model to Separate Within- and Between-Individual Effects. **167**, 766–773.

van de Pol M, Wright J (2009) A simple method for distinguishing within- versus between-subject effects using mixed models. *Animal Behaviour*, **77**, 753–758.

van de Pol M, Bailey LD, McLean N, Rijsdijk L, Lawson CR, Brouwer L (2016) Identifying the best climatic predictors in ecology and evolution. *Methods in Ecology and Evolution*, **7**, 1246–1257.

Poloczanska ES, Brown CJ, Sydeman WJ et al. (2013) Global imprint of climate change on marine life. *Nature Climate Change*, **3**, 919–925.

Poloczanska ES, Burrows MT, Brown CJ et al. (2016) Responses of marine organisms to climate change across oceans. *Frontiers in Marine Science*, **3**, 1–21.

Porlier M, Charmantier A, Bourgault P, Perret P, Blondel J (2012) Variation in phenotypic plasticity and selection patterns in blue tit breeding time: between- and within-population comparisons. *Journal of Animal Ecology*, **81**, 1041–1051.

Potts GR, Coulson JC, Deans IR (1980) Population dynamics and breeding success of the shag, Phalacrocorax aristotelis, on the Farne Islands, Northumberland. *Journal of Animal Ecology*, **49**, 465–484.

Provencher JF, Borrelle S, Sherley RB et al. (2019) Seabirds. In: *World Seas: an Environmental Evaluation. Volume III: Ecological Issues and Environmental Impacts*, Second Edi edn, pp. 133–162. Elsevier Ltd.

Purvis A, Gittleman JL, Cowlishaw G, Mace GM (2000) Predicting extinction risk in declining species. *Proceedings of the Royal Society of London B*, **267**, 1947–1952.

Pyper BJ, Mueter FJ, Peterman RM (2004) Across-species comparisons of spatial scales of environmental effects on survival rates of Northeast Pacific salmon. *Transactions of the American Fisheries Society*, **134**, 86–104.

Raimondo S, Liebhold AM, Strazanac JS, Butler L (2004) Population synchrony within and among Lepidoptera species in relation to weather, phylogeny, and larval phenology. *Ecological Entomology*, **29**, 96–105.

Ramirez F, Afan I, Tavecchia G, Catala IA, Oro D, Sanz-aguilar A (2016) Oceanographic drivers and mistiming processes shape breeding success in a seabird. *Proceedings of the Royal Society B: Biological Sciences*.

Rasmussen NL, Van Allen BG, Rudolf VHW (2014) Linking phenological shifts to species interactions through size-mediated priority effects. *Journal of Animal Ecology*, **83**, 1206–1215.

Reed TE, Wanless S, Harris MP, Frederiksen M, Kruuk LEB, Cunningham EJA (2006) Responding to environmental change: plastic responses vary little in a synchronous breeder. *Proceedings. Biological sciences / The Royal Society*, **273**, 2713–9.

Reed TE, Daunt F, Hall ME, Phillips RA, Wanless S, Cunningham EJA (2008) Parasite Treatment Affects Investment in Sons. *Science*, **321**, 2007–2008.

Reed TE, Warzybok P, Wilson AJ, Bradley RW, Wanless S, Sydeman WJ (2009) Timing is everything: flexible phenology and shifting selection in a colonial seabird. *Journal of Animal Ecology*, **78**, 376–387.

Reed TE, Waples RS, Schindler DE, Hard JJ, Kinnison MT (2010) Phenotypic plasticity and population viability: the importance of environmental predictability. *Proceedings of the Royal Society B: Biological Sciences*, **277**, 3391–400.

Reed TE, Jenouvrier S, Visser ME (2013a) Phenological mismatch strongly affects individual fitness but not population demography in a woodland passerine. *Journal of Animal Ecology*, **82**, 131–144.

Reed TE, Grotan V, Jenouvrier S, Saether B-E, Visser ME (2013b) Population growth in a wild bird is buffered against phenological mismatch. *Science*, **53**, 1689–1699.

Reed TE, Harris MP, Wanless S (2015) Skipped breeding in common guillemots in a changing climate: restraint or constraint? *Frontiers in Ecology and Evolution*, **3**, 1–13.

Regan T, Barnett MW, Laetsch DR et al. (2018) Characterisation of the British honey bee metagenome. *Nature Communications*, **9**, 4995.

Régnier T, Gibb FM, Wright PJ (2017) Importance of trophic mismatch in a winterhatching species: Evidence from lesser sandeel. *Marine Ecology*

Progress Series, **567**, 185–197.

Régnier T, Gibb FM, Wright PJ (2018) Temperature effects on egg development and larval condition in the lesser sandeel, Ammodytes marinus. *Journal of Sea Research*, **134**, 34–41.

Regular PM, Hedd A, Montevecchi W a., Robertson GJ, Storey AE, Walsh CJ (2014) Why timing is everything: Energetic costs and reproductive consequences of resource mismatch for a chick-rearing seabird. *Ecosphere*, **5**, 1–13.

Reiss MJ (1989) *The scaling of average daily metabolic rate and energy intake*. Cambridge University Press, 7–22 pp.

Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W (2002) An improved in situ and satellite SST analysis for climate. *Journal of Climate*, **15**, 1609–1625.

Richardson AJ, Poloczanska ES (2008) Under-resourced, under threat. *Science*, **320**, 1294–1295.

Riou S, Hamer KC (2010) Lipid metabolism, begging behaviour and nestling obesity in a pelagic seabird. *Functional Ecology*, **24**, 340–346.

Romero-Romero S, Molina-Ramírez A, Höfer J, Acuña JL (2016) Body size- based trophic structure of a deep marine ecosystem. *Ecology*, **97**, 171–181.

Sabarros PS, Durant JM, Grémillet D, Crawford RJM, Stenseth NC (2012) Differential responses of three sympatric seabirds to spatio-temporal variability in shared resources. *Marine Ecology Progress Series*, **468**, 291–301.

Samplonius JM, Both C (2017) Competitor phenology as a social cue in breeding site selection. *Journal of Animal Ecology*, **86**, 615–623.

Sandvik H, Erikstad KE (2008) Seabird life histories and climatic fluctuations: a phylogenetic-comparative time series analysis of North Atlantic seabirds. *Ecography*, **31**, 73–83.

Sankaranarayanan S (2007) Modeling the tide and wind-induced circulation in Buzzards Bay. *Estuarine, Coastal and Shelf Science*, **73**, 467–480.

Sato M, Tarzia M, Wanless RM (2016) Chapter 38 – Seabirds. *United Nations: First World Ocean Assessment*, **2258**, 1–17.

Schindler DE, Armstrong JB, Reed TE (2015) The portfolio concept in ecology and evolution. *Frontiers in Ecology and the Environment*, **13**, 257–263.

Schoener TW (1974) Resource partitioning in ecological communities. *Science*, **185**, 27–39.

Schreiber E, Burger J (2002) *Biology of Marine Birds*. 722 pp.

Shultz MT, Piatt JF, Harding AMA, Kettle AB, Van Pelt TI (2009) Timing of breeding and reproductive performance in murres and kittiwakes reflect mismatched seasonal prey dynamics. *Marine Ecology Progress Series*, **393**, 247–258.

Shutt JD (2018) *Expanding a classic woodland food chain into a geographically variable food web*. University of Edinburgh.

Sims DW, Genner MJ, Southward AJ, Hawkins SJ (2001) Timing of squid migration

reflects North Atlantic climate variability. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 2607–2611.

Smiley K, Emmerson LM (2016) A matter of timing: Adélie penguin reproductive success in a seasonally varying environment. *Marine Ecology Progress Series*, **542**, 235–249.

Sorensen MC, Hipfner JM, Kyser TK, Norris DR (2009) Carry-over effects in a Pacific seabird: Stable isotope evidence that pre-breeding diet quality influences reproductive success. *Journal of Animal Ecology*, **78**, 460–467.

Sparks TH, Smithers RJ (2002) Is spring getting earlier? *Weather*, **57**, 157–166.

Spooner FEB, Pearson RG, Freeman R (2018) Rapid warming is associated with population decline among terrestrial birds and mammals globally. *Global Change Biology*, **24**, 4521–4531.

Stearns S. (1989) Trade-Offs in Life-History Evolution. *Functional Ecology*, **3**, 259–268.

Stenseth NC, Mysterud A, Ottersen G, Hurrell JW, Chan K-S, Lima M (2002) Ecological Effects of Climate Fluctuations. *Science*, **297**, 1292–1297.

Stenseth NC, Ottersen G, Hurrell JW et al. (2003) Studying climate effects on ecology through the use of climate indices: the North Atlantic Oscillation, El Niño Southern Oscillation and beyond. *Proceedings. Biological sciences / The Royal Society*, **270**, 2087–96.

Stephens PA, Boyd IL, McNamara JM, Houston AI (2009) Capital breeding and income breeding:their meaning, measurement, and worth. *Ecology*, **90**, 2057–2067.

Stevenson IR, Bryant DM (2000) Climate change and constraints on breeding. *Nature*, **406**, 366–367.

Stocker TF, D. Qin G-K, Plattner M et al. (2013) IPCC: Summary for Policymakers. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 1–29.

Surman CA, Nicholson LW (2009) The good, the bad and the ugly: ENSO-driven oceanographic variability and its influence on seabird diet and reproductive performance at the Houtman Abrolhos, Eastern Indian Ocean. *Marine Ornithology*, **37**, 129–138.

Sydeman WJ, Thompson SA, Kitaysky A (2012) Seabirds and climate change: roadmap for the future. *Marine Ecology Progress Series*, **454**, 107–117.

Sydeman WJ, Poloczanska ES, Reed TE, Thompson SA (2015) Climate change and marine vertebrates. *Science*, **350**, 772–777.

Szostek KL, Bouwhuis S, Becker PH (2015) Are arrival date and body mass after spring migration influenced by large-scale environmental factors in a migratory seabird? *Frontiers in Ecology and Evolution*, **3**, 1–12.

Thackeray SJ (2012) Mismatch revisited: What is trophic mismatching from the perspective of the plankton? *Journal of Plankton Research*, **34**, 1001–1010.

Thackeray SJ (2016) Casting your network wide: a plea to scale-up phenological research. *Biology Letters*, **12**, 20160181.

Thackeray SJ, Sparks TH, Frederiksen M et al. (2010) Trophic level asynchrony in rates of phenological change for marine, freshwater and terrestrial environments. *Global Change Biology*, **16**, 3304–3313.

Thackeray SJ, Henrys PA, Hemming D et al. (2016) Phenological sensitivity to climate across taxa and trophic levels. *Nature*, **535**, 241–245.

Thaxter CB, Lascelles B, Sugar K et al. (2012) Seabird foraging ranges as a preliminary tool for identifying candidate Marine Protected Areas. *Biological Conservation*, **156**, 53–61.

Thompson PM, Ollason JC (2001) Lagged effects of ocean climate change on fulmar population dynamics. *Nature*, **413**, 417–420.

Townsend DW, Cammen LM, Holligan PM, Campbell DE, Pettigrew NR (1994) Causes and consequences of variability in the timing of phytoplankton blooms. *Deep-Sea Research Part I*, **41**, 747–765.

Trujillo A, Thurman H (2014) *Essentials of Oceanography (11th Edition)*, 11th edn. Pearson Prentice Hall, Chapter 6 p.

Varpe Ø (2017) Life history adaptations to seasonality. *Integrative and Comparative Biology*, **57**, 943–960.

Verhulst S, Nilsson J. (2008) The timing of birds' breeding seasons: a review of experiments that manipulated timing of breeding. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 399–410.

Vihtakari M, Welcker J, Moe B et al. (2018) Black-legged kittiwakes as messengers of Atlantification in the Arctic. *Scientific Reports*, **8**, 1–11.

Visser ME (2008) Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings. Biological sciences / The Royal Society*, **275**, 649–59.

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

Visser ME, Both C, Lambrechts MM (2004) Global Climate Change Leads to Mistimed Avian Reproduction. *Advances in Ecological Research*, **35**, 89–110.

Visser ME, Holleman LJM, Gienapp P (2006) Shifts in caterpillar biomass phenology due to climate change and its impact on the breeding biology of an insectivorous bird. *Oecologia*, **147**, 164–72.

Visser ME, Gienapp P, Husby A, Morrisey M, de la Hera I, Pulido F, Both C (2015) Effects of Spring Temperatures on the Strength of Selection on Timing of Reproduction in a Long-Distance Migratory Bird. *PLoS Biology*, **13**, 1–17.

Wakefield ED, Bodey TW, Bearhop S et al. (2013) Space partitioning without territoriality in gannets. *Science*, **341**, 68–70.

Walther G. (2010) Community and ecosystem responses to recent climate change. *Philosophical transactions of the Royal Society of London. Series B, Biological*

sciences, **365**, 2019–24.

Walther G, Post E, Convey P et al. (2002) Ecological responses to recent climate change. *Nature*, **416**, 389–395.

Wanless S, Wright PJ, Harris MP, Elston DA (2004) Evidence for decrease in size of lesser sandeels Ammodytes marinus in a North Sea aggregation over a 30-yr period. *Marine Ecology Progress Series*, **279**, 237–246.

Wanless S, Harris MP, Redman P, Speakman JR (2005) Low energy values of fish as a probable cause of a major seabird breeding failure in the North Sea. *Marine Ecology Progress Series*, **294**, 1–8.

Wanless S, Harris MP, Lewis S, Frederiksen M, Murray S (2008) Later breeding in northern gannets in the eastern Atlantic. *Marine Ecology Progress Series*, **370**, 263–269.

Wanless S, Frederiksen M, Walton J, Harris MP (2009) Long-term changes in breeding phenology at two seabird colonies in the western North Sea. *Ibis*, **151**, 274–285.

Warton D, Hui F (2011) The arcsine is asinine: the analysis of proportions in ecology. *Ecology*, **92**, 3–10.

Watanuki Y, Ito M, Deguchi T, Minobe S (2009) Climate-forced seasonal mismatch between the hatching of rhinoceros auklets and the availability of anchovy. *Marine Ecology Progress Series*, **393**, 259–271.

Williams J. (1996) Energetics of avian incubation. In: *Avian energetics and nutritional ecology*, pp. 375–415. Chapman and Hall, New York.

Williams SE, Moritz C, Shoo LP, Isaac JL, Hoffmann AA, Langham G (2008) Towards an Integrated Framework for Assessing the Vulnerability of Species to Climate Change. *PLoS Biol*, **6**, e325.

Winslade P (1974) Behavioural studies on the lesser sandeel Ammodytes marinus (Raitt) III. The effect of temperature on activity and the environmental control of the annual cycle of activity. *Journal of Fish Biology*, **6**, 587–599.

Wojczulanis-Jakubas K, Araya-Salas M, Jakubas D (2018) Seabird parents provision their chick in a coordinated manner. *PLoS ONE*, **13**, 1–13.

Wolf SG, Sydeman WJ, Hipfner JM, Abraham CL, Tershy BR, Croll D a. (2009) Range-wide reproductive consequences of ocean climate variability for the seabird Cassin's Auklet. *Ecology*, **90**, 742–753.

Wright PJ, Bailey MC (1996) Timing of hatching in Ammodytes marinus from Shetland waters and its significance to early growth and survivorship. *Marine Biology*, **126**, 143–152.

Wright PJ, Orpwood JE, Boulcott P (2017a) Warming delays ovarian development in a capital breeder. *Marine Biology*, **164**, 1–9.

Wright PJ, Orpwood JE, Scott BE (2017b) Impact of rising temperature on reproductive investment in a capital breeder: The lesser sandeel. *Journal of Experimental Marine Biology and Ecology*, **486**, 52–58.

Youngflesh C, Jenouvrier S, Li Y et al. (2017) Circumpolar analysis of the Adelie

penguin reveals the importance of environmental variability in phenological mismatch. *Ecology*, **98**, 940–951.

Youngflesh C, Jenouvrier S, Hinke JT et al. (2018) Rethinking “normal”: The role of stochasticity in the phenology of a synchronously breeding seabird. *Journal of Animal Ecology*, **87**, 682–690.

A. Chapter 3 Appendix

A.1 Supplementary Tables

Table A.1. Coefficients (and CIs) from core breeding model. Intercept represents the median lay date across all populations in the analysis. Random effects estimate covariance in timing of breeding across populations: across whole North Atlantic (Year effect); of the same species (Species.Year effects); breeding at the same site (Site.Year effects); breeding in the same group of sites (<120km) (Site group.Year effects); of the same species breeding in the same group of sites (Breeding region.Year effects) and breeding in the same Large Marine Ecosystem (Sea.Year effects). Residual terms are the remaining unexplained variance within individual populations (residual Population.Year effects). [l] represents populations for which phenology was measured using lay date, and [h] represents populations for which phenology was measured using hatch date.

Breeding model - random effects	Coefficient/Variance (Median and CI)	Effective sample size
Fixed Terms		
Intercept	143.767 (125.577 - 160.953)	24700
Random Terms		
Year	0.181 (0 - 1.135)	661
Species	43.62 (0 - 238.377)	122
Colony	16.475 (0 - 62.779)	163
Colony_Species	26.975 (2.279 - 81.237)	297
Site	43.578 (0 - 381.196)	217
Breeding region	193.921 (47.889 - 433.827)	418
Sea	438.515 (0.005 - 1520.431)	240
Species: Arctic tern.Year	1.127 (0 - 3.736)	4400
Species: Atlantic puffin.Year	2.245 (0 - 5.705)	4462
Species: Black-legged kittiwake.Year	11.038 (2.761 - 22.639)	1055
Species: Brünnich's guillemot.Year	5.68 (0 - 12.905)	2458
Species: Common guillemot.Year	0.29 (0 - 1.73)	2766
Species: Common tern.Year	0.498 (0 - 2.566)	1997
Species: European shag.Year	5.503 (0 - 21.568)	5183
Species: Razorbill.Year	1.398 (0 - 4.056)	8546
Species: Roseate tern.Year	0.538 (0 - 3.513)	5062
Site: Anda.Year	3.091 (0 - 24.658)	9969
Site: Bird Island.Year	7.506 (0 - 14.104)	1104
Site: Country Island.Year	28.275 (10.177 - 57.824)	3107
Site: Hornøya.Year	12.88 (5.858 - 22.379)	6216
Site: Isle of May.Year	12.499 (0 - 23.795)	1423
Site: Machias Seal Island.Year	5.208 (0 - 13.617)	6977
Site: Penikese Island.Year	1.608 (0 - 8.507)	6464

Site: Prince Leopold Island.Year	21.664 (1.873 - 53.279)	3012
Site: Ram Island.Year	3.927 (0 - 11.189)	2829
Site: Røst.Year	6.408 (0 - 36.054)	4447
Site: Sklinna.Year	9.334 (0 - 43.325)	5283
Site: Sumburgh Head.Year	7.997 (0 - 15.025)	4878
Site group: Buzzards Bay.Year	1.548 (0 - 8.521)	572
Site group: Maine.Year	1.217 (0 - 6.363)	4596
Site group: North Spain.Year	50.921 (0 - 223.875)	3688
Site group: Shetland.Year	33.106 (14.253 - 59.942)	1736
Site group: Svalbard.Year	9.741 (0 - 23.179)	4576
Breeding region:species: Buzzards Bay, common tern.Year	9.477 (0.768 - 18.085)	891
Breeding region:species: Buzzards Bay, roseate tern.Year	0.834 (0 - 5.248)	1892
Breeding region:species: Shetland, black-legged kittiwake.Year	7.049 (0 - 21.638)	729
Sea: North Sea.Year	7.448 (0 - 18.625)	1204
Sea: Norwegian Sea.Year	5.413 (0 - 24.508)	2642
Sea: Scotian Shelf.Year	1.981 (0 - 6.389)	1915
Residual Terms		
Population: aforcada_europeanshag.Year[l]	116.587 (0.002 - 447.077)	3131
Population: anda_atlanticpuffin.Year[h]	13.65 (0.002 - 51.629)	5706
Population: anda_blackleggedkittiwake.Year[h]	36.956 (0.005 - 110.161)	6845
Population: aspantorgas_europeanshag.Year[l]	145.448 (0.002 - 359.973)	3814
Population: bantersee_commontern.Year[l]	25.034 (9.132 - 48.595)	10497
Population: birdisland_commontern.Year[l]	1.425 (0.002 - 8.579)	950
Population: birdisland_roseatetern.Year[h]	2.119 (0.002 - 8.083)	1669
Population: burravoe_blackleggedkittiwake.Year[l]	3.159 (0.001 - 9.415)	13525
Population: coatsisland_thickbilledmurre.Year[h]	3.753 (0.002 - 11.401)	1817
Population: compasshead_blackleggedkittiwake.Year[l]	13.547 (2.218 - 41.986)	23663
Population: countryisland_arctictern.Year[h]	0.179 (0.002 - 2.17)	6509
Population: countryisland_commontern.Year[h]	1.982 (0.003 - 5.381)	7264

Population: countryisland_roseatetern.Year[h]	9.238 (2.312 - 22.163)	18245
Population: easterneggrock_commontern.Year[h]	12.676 (5.189 - 23.923)	11835
Population: eschaness_blackleggedkittiwake.Year[l]	15.707 (4.576 - 38.779)	24700
Population: grumantbyen_blackleggedkittiwake.Year[h]	10.18 (0.003 - 29.335)	16639
Population: hornøya_atlanticpuffin.Year[h]	1.06 (0.002 - 5.254)	3220
Population: hornøya_blackleggedkittiwake.Year[h]	80.026 (35.633 - 150.205)	21858
Population: hornøya_commonguillemot.Year[h]	0.424 (0.002 - 5.183)	11606
Population: hornøya_razorbill.Year[h]	4.356 (0.004 - 8.768)	10371
Population: isleofmay_atlanticpuffin.Year[l]	16.251 (8.692 - 26.78)	17395
Population: isleofmay_blackleggedkittiwake.Year[l]	18.67 (6.687 - 37.055)	18155
Population: isleofmay_commonguillemot.Year[l]	0.292 (0.001 - 2.578)	2727
Population: isleofmay_europeanshag.Year[l]	339.382 (192.04 - 552.553)	25149
Population: isleofmay_razorbill.Year[l]	3.699 (0.003 - 7.033)	4754
Population: kettlaness_blackleggedkittiwake.Year[l]	6.633 (0.006 - 19.114)	17154
Population: kongsfjorden_blackleggedkittiwake.Year[h]	0.54 (0.002 - 10.337)	3324
Population: machiassealisland_arctictern.Year[l]	14.131 (4.066 - 30.048)	13544
Population: machiassealisland_atlanticpuffin.Year[l]	45.255 (20.843 - 85.551)	22641
Population: machiassealisland_commontern.Year[l]	1.578 (0.002 - 11.171)	7690
Population: machiassealisland_razorbill.Year[l]	14.109 (3.587 - 29.958)	14835
Population: matinicusrock_arctictern.Year[h]	3.593 (0.003 - 9.358)	4279
Population: noness_blackleggedkittiwake.Year[l]	6.082 (2.012 - 12.195)	16619
Population: penikeseisland_commontern.Year[l]	0.465 (0.002 - 5.425)	9312
Population: penikeseisland_roseatetern.Year[l]	9.015 (0.003 - 29.46)	5091

Population: princeleopoldisland_blackleggedkittiwake.Year[h]	0.539 (0.002 - 12.695)	7844
Population: princeleopoldisland_thickbilledmurre.Year[h]	8.177 (0.002 - 30.715)	2905
Population: ramisland_commontern.Year[l]	0.169 (0.002 - 2.321)	8932
Population: ramisland_roseatetern.Year[l]	5.514 (0.002 - 14.662)	1710
Population: ramnageo_blackleggedkittiwake.Year[l]	9.56 (3.677 - 18.456)	18278
Population: røst_atlanticpuffin.Year[h]	42.267 (0.003 - 78.494)	6037
Population: røst_blackleggedkittiwake.Year[h]	35.372 (0.003 - 107.955)	10022
Population: røst_europeanshag.Year[h]	157.851 (79.693 - 264.812)	15611
Population: sklinna_atlanticpuffin.Year[h]	29.688 (0.002 - 90.644)	15499
Population: sklinna_europeanshag.Year[h]	0.808 (0.002 - 21.694)	6291
Population: storakarlso_commonguillemot.Year[l]	8.496 (2.708 - 20.853)	13900
Population: sumburghhead_blackleggedkittiwake.Year[l]	0.653 (0.001 - 6.975)	3230
Population: sumburghhead_commonguillemot.Year[l]	5.094 (0.001 - 21.211)	1522
Population: sumburghhead_europeanshag.Year[l]	124.912 (63.362 - 212.875)	18851
Population: troswickness_blackleggedkittiwake.Year[l]	16.4 (2.58 - 62.883)	16398
Population: westerwick_blackleggedkittiwake.Year[l]	0.287 (0.001 - 3.54)	8007

Table A.2. Coefficients (and CIs) from fixed effects breeding model. Intercept represents the median lay date for a population found at 0 degrees latitude on the East coast of the North Atlantic. Random effects estimate covariance in timing of breeding across populations: across the whole North Atlantic (Year effect); of the same species (Species.Year effects); breeding at the same site (Site.Year effects); breeding in the same group of sites (<120km) (Site group.Year effects); of the same species breeding in the same group of sites (Breeding region.Year effects) and breeding in the same Large Marine Ecosystem (Sea.Year effects). Residual terms are the remaining unexplained variance within individual populations (residual Population.Year effects). [l] represents populations for which phenology was measured using lay date, and [h] represents populations for which phenology was measured using hatch date.

Breeding model - fixed effects	Coefficient/Variance (Median and CI)	Effective sample size
Fixed Terms		
Intercept	27.212 (-22.384 - 75.554)	3691
Latitude	1.763 (1.003 - 2.555)	2321
West Atlantic	37.873 (20.54 - 56.141)	24700
Random Terms		
Year	0.197 (0 - 1.205)	463
Species	37.533 (0 - 232.431)	110
Colony	13.819 (0 - 58.798)	136
Colony_Species	28.438 (2.011 - 78.331)	335
Site	12.738 (0 - 104.13)	284
Breeding region	170.342 (60.116 - 321.96)	799
Sea	10.978 (0 - 127.56)	243
Species: Arctic tern.Year	1.163 (0 - 3.759)	4781
Species: Atlantic puffin.Year	2.248 (0 - 5.671)	4832
Species: Black-legged kittiwake.Year	10.493 (2.998 - 22.029)	1396
Species: Brünnich's guillemot.Year	5.308 (0 - 12.585)	2430
Species: Common guillemot.Year	0.283 (0 - 1.721)	3244
Species: Common tern.Year	0.479 (0 - 2.615)	1987
Species: European shag.Year	5.529 (0 - 22.187)	5341
Species: Razorbill.Year	1.402 (0 - 4.001)	10034
Species: Roseate tern.Year	0.575 (0 - 3.689)	5171
Site: Anda.Year	3.115 (0 - 25.266)	7703
Site: Bird Island.Year	7.289 (0 - 13.865)	1228
Site: Country Island.Year	27.675 (10.189 - 56.213)	2912
Site: Hornøya.Year	12.875 (5.908 - 22.387)	6165
Site: Isle of May.Year	12.717 (1.27 - 25.118)	1456
Site: Machias Seal Island.Year	5.104 (0 - 13.553)	6172
Site: Penikese Island.Year	1.653 (0 - 8.667)	7565
Site: Prince Leopold Island.Year	21.983 (1.689 - 54.07)	2829
Site: Ram Island.Year	3.966 (0 - 11.339)	2735

Site: Røst.Year	6.165 (0 - 36.114)	3654
Site: Sklinna.Year	9.032 (0 - 41.963)	5391
Site: Sumburgh Head.Year	7.976 (0 - 15.157)	4278
Site group: Buzzards Bay.Year	1.853 (0 - 8.711)	595
Site group: Maine.Year	1.21 (0 - 6.345)	4189
Site group: North Spain.Year	54.42 (0 - 228.425)	3207
Site group: Shetland.Year	33.875 (14.478 - 61.162)	1803
Site group: Svalbard.Year	9.756 (0 - 23.374)	4120
Breeding region:species: Buzzards Bay, common tern.Year	9.214 (0 - 16.974)	854
Breeding region:species: Buzzards Bay, roseate tern.Year	0.775 (0 - 5.084)	2818
Breeding region:species: Shetland, black-legged kittiwake.Year	6.827 (0 - 21.645)	827
Sea: North Sea.Year	7.278 (0 - 17.91)	1571
Sea: Norwegian Sea.Year	5.375 (0 - 24.201)	3285
Sea: Scotian Shelf.Year	2.134 (0 - 6.565)	1864
Residual Terms		
Population: aforcada_europeanshag.Year[l]	115.26 (0.002 - 441.847)	3671
Population: anda_atlanticpuffin.Year[h]	13.613 (0.002 - 51.704)	5700
Population: anda_blackleggedkittiwake.Year[h]	36.813 (0.003 - 111.723)	7699
Population: aspantorgas_europeanshag.Year[l]	147.301 (0.003 - 354.32)	3507
Population: bantersee_commontern.Year[l]	25.033 (9.551 - 47.24)	14126
Population: birdisland_commontern.Year[l]	1.423 (0.002 - 8.348)	1127
Population: birdisland_roseatetern.Year[h]	2.338 (0.002 - 8.248)	1615
Population: burravoe_blackleggedkittiwake.Year[l]	3.185 (0.002 - 9.549)	13364
Population: coatsisland_thickbilledmurre.Year[h]	4.176 (0.002 - 11.501)	2076
Population: compasshead_blackleggedkittiwake.Year[l]	13.621 (2.438 - 43.257)	20086
Population: countryisland_arctictern.Year[h]	0.179 (0.002 - 2.152)	6019
Population: countryisland_commontern.Year[h]	1.984 (0.002 - 5.331)	6765
Population: countryisland_roseatetern.Year[h]	9.239 (2.038 - 22.167)	19198
Population: easterneggrock_commontern.Year[h]	12.669 (5.174 - 24.063)	11881

Population: eschaness_blackleggedkittiwake.Year[l]	15.61 (4.825 - 38.589)	24700
Population: grumantbyen_blackleggedkittiwake.Year[h]	10.114 (0.002 - 29.401)	16888
Population: hornøya_atlanticpuffin.Year[h]	0.96 (0.002 - 5.134)	3143
Population: hornøya_blackleggedkittiwake.Year[h]	79.85 (35.916 - 149.142)	23538
Population: hornøya_commonguillemot.Year[h]	0.411 (0.001 - 5.204)	14358
Population: hornøya_razorbill.Year[h]	4.356 (0.005 - 8.821)	10713
Population: isleofmay_atlanticpuffin.Year[l]	16.196 (8.524 - 26.591)	21852
Population: isleofmay_blackleggedkittiwake.Year[l]	18.553 (6.982 - 37.323)	19821
Population: isleofmay_commonguillemot.Year[l]	0.266 (0.002 - 2.534)	2565
Population: isleofmay_europeanshag.Year[l]	339.154 (193.247 - 553.974)	24700
Population: isleofmay_razorbill.Year[l]	3.745 (0.003 - 7.102)	4787
Population: kettlaness_blackleggedkittiwake.Year[l]	6.594 (0.01 - 19.242)	19903
Population: kongsfjorden_blackleggedkittiwake.Year[h]	0.581 (0.002 - 11.242)	3048
Population: machiassealisland_arctictern.Year[l]	14.181 (4.055 - 29.964)	15707
Population: machiassealisland_atlanticpuffin.Year[l]	45.477 (20.829 - 86.544)	21325
Population: machiassealisland_commontern.Year[l]	1.632 (0.002 - 11.305)	7670
Population: machiassealisland_razorbill.Year[l]	14.024 (4.281 - 30.328)	17717
Population: matinicusrock_arctictern.Year[h]	3.52 (0.002 - 9.164)	4848
Population: noness_blackleggedkittiwake.Year[l]	6.099 (2.099 - 12.229)	16741
Population: penikeseisland_commontern.Year[l]	0.452 (0.001 - 5.389)	9641
Population: penikeseisland_roseatetern.Year[l]	9.231 (0.003 - 30.759)	6481
Population: princeleopoldisland_blackleggedkittiwake.Year[h]	0.553 (0.001 - 12.889)	7170
Population: princeleopoldisland_thickbilledmurre.Year[h]	7.785 (0.002 - 30.119)	2590
Population: ramisland_commontern.Year[l]	0.152 (0.001 - 2.19)	8100
Population: ramisland_roseatetern.Year[l]	5.296 (0.002 - 14.336)	1795

Population: ramnageo_blackleggedkittiwake.Year[l]	9.475 (3.859 - 18.207)	18785
Population: røst_atlanticpuffin.Year[h]	41.844 (0.002 - 78.463)	5381
Population: røst_blackleggedkittiwake.Year[h]	35.613 (0.002 - 107.393)	11942
Population: røst_europeanshag.Year[h]	157.874 (82.139 - 266.449)	13279
Population: sklinna_atlanticpuffin.Year[h]	29.512 (0.002 - 91.134)	14019
Population: sklinna_europeanshag.Year[h]	0.862 (0.001 - 22.938)	5204
Population: storakarlso_commonguillemot.Year[l]	8.467 (2.513 - 20.833)	16520
Population: sumburghhead _blackleggedkittiwake.Year[l]	0.708 (0.002 - 7.106)	3107
Population: sumburghhead _commonguillemot.Year[l]	5.147 (0.002 - 21.45)	2115
Population: sumburghhead _europeanshag.Year[l]	123.624 (62.236 - 210.359)	17050
Population: troswickness_blackleggedkittiwake.Year[l]	16.537 (2.583 - 69.766)	13827
Population: westerwick_blackleggedkittiwake.Year[l]	0.277 (0.001 - 3.539)	7938

Table A.3. Coefficients (and CIs) from core wintering model. Intercept represents the median lay date across all populations in the analysis. Random effects estimate covariance in timing of breeding across populations: across whole North Atlantic (Year effect); of the same species (Species.Year effects); breeding in the same group of sites (<120km) (Site group.Year effects), which takes spatial autocorrelation of populations at closely positioned breeding sites into account; wintering in the same Large Marine Ecosystem (Sea.Year effects) and of the same species wintering in the same group of sites (Wintering region.Year effects). Residual terms are the remaining unexplained variance within individual populations (residual Population.Year effects). [l] represents populations for which phenology was measured using lay date, and [h] represents populations for which phenology was measured using hatch date.

Wintering model - random effects	Coefficient/Variance (Median and CI)	Effective sample size
Fixed Terms		
Intercept	141.056 (121.598 - 159.645)	8900
Random Terms		
Year	0.654 (0 - 2.113)	768
Species	89.147 (0 - 592.466)	71
Colony_Species	45.53 (11.547 - 124.955)	766
Breeding Region	150.39 (0 - 319.625)	1410
Winter	377.983 (0 - 1481.614)	51
Wintering Region	78.722 (0 - 545.699)	70
Species: Arctic tern.Year	2.153 (0 - 7.809)	6317
Species: Atlantic puffin.Year	1.536 (0 - 6.404)	2160
Species: Black-legged kittiwake.Year	6.89 (0 - 18.3)	2137
Species: Brünnich's guillemot.Year	5.945 (0 - 13.131)	2923
Species: Common guillemot.Year	0.201 (0 - 1.542)	4253
Species: Common tern.Year	0.987 (0 - 5.185)	7262
Species: European shag.Year	13.989 (0 - 36.918)	3415
Species: Razorbill.Year	0.822 (0 - 3.328)	2825
Species: Roseate tern.Year	0.853 (0 - 6.419)	3789
Breeding region:species: Buzzards Bay, common tern.Year	7.963 (0.862 - 16.3)	8956
Breeding region:species: Buzzards Bay, roseate tern.Year	1.303 (0 - 8.718)	2486
Breeding region:species: Shetland, black-legged kittiwake.Year	44.586 (21.346 - 78.584)	3697
Winter: Barents Sea.Year	8.937 (0 - 19.923)	2562
Winter: East or South Brazil shelf.Year	1.02 (0 - 4.553)	4376
Winter: Gulf of Maine.Year	14.772 (0 - 30.604)	2594
Winter: Iberian coastal.Year	62.352 (0 - 244.257)	3957
Winter: Iceland shelf.Year	41 (0 - 70.898)	2637

Winter: Labrador Sea.Year	2.372 (0 - 8.888)	1511
Winter: North Sea.Year	17.879 (9.91 - 28.852)	4548
Winter: Norwegian Sea.Year	3.007 (0 - 13.731)	3643
Wintering region:species: East or South Brazil shelf, common tern.Year	0.882 (0 - 6.448)	4713
Wintering region:species: East or South Brazil shelf, roseate tern.Year	40.32 (0 - 81.627)	3386
Wintering region:species: North Sea, common guillemot.Year	52.535 (0 - 189.297)	1892
Wintering region:species: North Sea, European shag.Year	2.565 (0 - 6.422)	2288
Residual Terms		
Population: aforcada_europeanshag.Year[l]	107.046 (0.002 - 435.052)	3558
Population: anda_atlanticpuffin.Year[h]	11.236 (0.002 - 44.998)	5731
Population: anda_blackleggedkittiwake.Year[h]	17.934 (4.13 - 45.57)	23052
Population: aspantorgas_europeanshag.Year[l]	157.445 (0.002 - 377.412)	3514
Population: bantersee_commontern.Year[l]	30.634 (14.801 - 53.87)	23920
Population: birdisland_commontern.Year[l]	8.066 (3.036 - 15.334)	17493
Population: birdisland_roseatetern.Year[h]	6.717 (0.002 - 13.792)	2328
Population: burravoe_blackleggedkittiwake.Year[l]	3.017 (0.002 - 9.165)	13306
Population: coatsisland_thickbilledmurre.Year[h]	0.585 (0.001 - 7.288)	2811
Population: compasshead_blackleggedkittiwake.Year[l]	13.858 (2.651 - 42.51)	22127
Population: countryisland_arctictern.Year[h]	17.214 (6.684 - 33.564)	24700
Population: countryisland_commontern.Year[h]	24.445 (10.279 - 47.523)	24700
Population: countryisland_roseatetern.Year[h]	25.885 (10.157 - 52.356)	21843
Population: easterneggrock_commontern.Year[h]	14.371 (6.326 - 26.124)	22286
Population: eschaness_blackleggedkittiwake.Year[l]	15.81 (4.66 - 38.509)	24828
Population: grumantbyen_blackleggedkittiwake.Year[h]	5.809 (0.003 - 20.946)	16976
Population: hornøya_atlanticpuffin.Year[h]	4.196 (0.002 - 16.085)	2314
Population: hornøya_blackleggedkittiwake.Year[h]	113.123 (55.426 - 208.817)	23762
Population: hornøya_commonguillemot.Year[h]	0.227 (0.002 - 3.786)	11221

Population: hornøya_razorbill.Year[h]	11.867 (0.003 - 21.851)	5386
Population: isleofmay_atlanticpuffin.Year[l]	16.534 (7.816 - 27.715)	12757
Population: isleofmay_blackleggedkittiwake.Year[l]	58.993 (27.478 - 111.722)	23430
Population: isleofmay_commonguillemot.Year[l]	1.47 (0.002 - 4.428)	3407
Population: isleofmay_europeanshag.Year[l]	333.779 (166.502 - 584.826)	6720
Population: isleofmay_razorbill.Year[l]	0.383 (0.002 - 4.291)	2722
Population: kettlaness_blackleggedkittiwake.Year[l]	6.812 (0.009 - 20.008)	18192
Population: kongsfjorden_blackleggedkittiwake.Year[h]	9.422 (0.004 - 21.78)	7928
Population: machiassealisland_arctictern.Year[l]	16.267 (6.899 - 31.396)	18266
Population: machiassealisland_atlanticpuffin.Year[l]	1.39 (0.002 - 48.53)	2529
Population: machiassealisland_commontern.Year[l]	9.409 (2.123 - 24.88)	23488
Population: machiassealisland_razorbill.Year[l]	1.797 (0.002 - 17.52)	2499
Population: matinicusrock_arctictern.Year[h]	5.691 (0.003 - 11.438)	8228
Population: noness_blackleggedkittiwake.Year[l]	5.96 (1.771 - 12.052)	14516
Population: penikeseisland_commontern.Year[l]	2.564 (0.002 - 9.83)	6054
Population: penikeseisland_roseatetern.Year[l]	9.858 (0.004 - 31.831)	11168
Population: princeleopoldisland_blackleggedkittiwake.Year[h]	12.638 (0.95 - 39.121)	19555
Population: princeleopoldisland_thickbilledmurre.Year[h]	30.559 (12.214 - 61.112)	24700
Population: ramisland_commontern.Year[l]	2.825 (0.002 - 8.798)	5654
Population: ramisland_roseatetern.Year[l]	11.755 (3.108 - 26.462)	12962
Population: ramnageo_blackleggedkittiwake.Year[l]	9.851 (3.779 - 18.603)	19573
Population: røst_atlanticpuffin.Year[h]	6.998 (0.002 - 47.539)	1897
Population: røst_blackleggedkittiwake.Year[h]	60.692 (18.062 - 158.054)	22681
Population: røst_europeanshag.Year[h]	153.749 (79.841 - 258.878)	19743
Population: sklinna_atlanticpuffin.Year[h]	20.04 (0.002 - 86.747)	7094
Population: sklinna_europeanshag.Year[h]	2.71 (0.002 - 34.581)	5097
Population: storakarlso_commonguillemot.Year[l]	9.39 (2.978 - 23.648)	11791

Population: sumburghhead _blackleggedkittiwake.Year[I]	8.952 (4.324 - 15.872)	23338
Population: sumburghhead _commonguillemot.Year[I]	31.654 (13.169 - 63.709)	20599
Population: sumburghhead _europeanshag.Year[I]	117.438 (0.003 - 234.806)	2053
Population: troswickness_blackleggedkittiwake.Year[I]	16.728 (2.62 - 72.582)	12047
Population: westerwick_blackleggedkittiwake.Year[I]	0.3 (0.002 - 3.535)	7992

Table A.4. Median correlations between populations breeding at the same site. Pearson's pairwise correlations for each combination of populations were calculated from the raw data.

	Median correlation	p-value	Number of populations
Anda	-0.529	0.116	2
Bird Island	0.513	0.000	2
Country Island	0.704	0.002	3
Hornøya	0.524	0.079	4
Isle of May	0.655	0.000	5
Machias Seal Island	0.370	0.176	4
Penikese Island	0.070	0.858	2
Prince Leopold Island	0.464	0.176	2
Ram Island	0.831	0.000	2
Røst	0.040	0.544	3
Sklinna	0.471	0.200	2
Sumburgh Head	0.516	0.004	3

A.2 Testing for sufficient sample size

In the full model, we used populations for which average sample size across all years was >10 breeding pairs and where sample size within a year was >5. However, low sample size may lead to imprecise (co)variance estimates. We therefore tested that our key results were still valid by running our core breeding model using a reduced dataset, in which we only used years where sample size was >50 breeding pairs. In some populations, only some years

had the required sample size, and removing them caused the total number of years in the time series to fall below 8 (which was our stipulated cut-off for a time series to qualify for the analysis), and in these cases we removed the full time series from the analysis. This reduced dataset therefore comprised 675 annual phenological means from 33 populations, across 20 breeding sites, and 10 breeding regions (Table A.5).

Table A.5. List of breeding sites and species included in the analyses in order of decreasing latitude, with breeding and wintering regions indicated. Site numbers correspond with those in Figure 2a. Species are as follows: AP = Atlantic puffin, RA = razorbill, CG = common guillemot, BG = Brünnich's guillemot, SH = European shag, KI = black-legged kittiwake, AT = Arctic tern, CT = common tern, RT = roseate tern, with numbers in parenthesis indicating the number of populations of each species included in the analyses. A region was only included in the analysis of annual covariance when data for more than one population were available.

Site	AP (3)	RA (1)	CG (3)	BG (2)	SH (3)	KI (12)	AT (2)	CT (5)	RT (2)	Breeding region
1 Kongsfjorden						x				Arctic Ocean
2 Prince Leopold Island[a]				x		x				Baffin Bay
3 Hornøya[a]						x				Barents Sea
4 Røst[a]	x				x	x				Norwegian Sea
5 Coats Island				x						Hudson Bay
6 Burravoe						x				North sea[b] (also as Shetland site group)
7 Esha Ness						x				
8 Westerwick						x				
9 Ramna Geo						x				
10 Kettla Ness						x				
11 Troswick Ness						x				
12 Sumburgh Head[a]			x		x	x				
13 Stora Karlsö			x							Baltic Sea
14 Isle of May[a]	x	x	x		x	x				North Sea[b]
15 Banter See								x		North Sea[b]
16 Country Island[a]							x	x		Scotian Shelf[b]
17 Machias Seal Island[a]	x						x			Scotian Shelf[b]
18 Bird Island[a]								x	x	North East U.S Continental Shelf[b,c] (as Buzzards Bay site group)
19 Ram Island[a]								x	x	
20 Penikese Island[a]								x		

[a] represents sites over which variance between years was estimated, [b] represents breeding regions for which among-year variances were estimated. [c] represents effects which are confounded because the same combination of populations is grouped into another term, see main text for details. Confounded terms were not included in the model unless specified in the main text.

Results

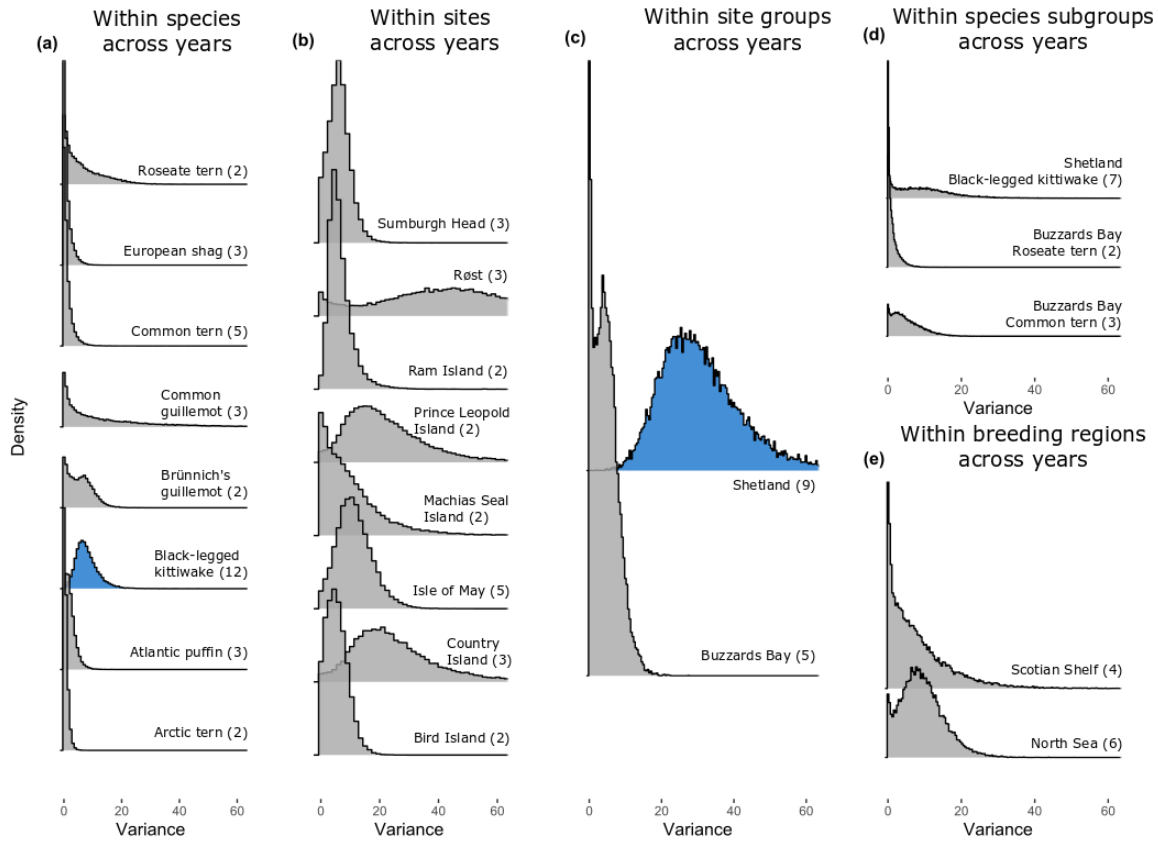


Figure A.1. (Co)variance in timing of breeding of seabird populations across years during the breeding season. Plotted from the posterior distribution of the random-effects model using the reduced dataset limited to years where sample size was > 50 breeding pairs. Represents shared variance across years according to (a) species, (b) site, (c) site groups (< 120 km apart), (d) species subgroups (i.e. populations of the same species within a group of nearby sites), and (e) breeding regions. On the y axes, values in parenthesis indicate the number of populations associated with each group. For interpretation, narrower histograms indicate a posterior distribution that has been estimated with higher certainty (i.e. a smaller credible interval), and histograms further removed from zero indicate higher likelihood of significance. Groups for which significant positive covariance was estimated (i.e. where 2.5% credible intervals did not overlap zero) are shaded in blue.

Using the reduced dataset, we found that the key results of the core breeding analysis remained the same: There was no shared covariance across large spatial scales, with covariance instead being observed between populations at the same site or sites within 120 km of one another. The significant covariance observed in the original model between populations of black-legged kittiwake remained in the new model when 4 out of 16 populations were removed (Figure A.1, Table A.6).

Table A.6. Coefficients (and CIs) from breeding model using the reduced dataset. Intercept represents the average lay date across all populations in the analysis. Random effects estimate covariance in timing of breeding across populations: across whole North Atlantic (Year effect); of the same species (Species.Year effects); breeding at the same site (Site.Year effects); breeding in the same group of sites (<120km) (Site group.Year effects); of the same species breeding in the same group of sites (Breeding region.Year effects) and breeding in the same Large Marine Ecosystem (Sea.Year effects). Residual terms are the remaining unexplained variance within individual populations (residual Population.Year effects).

	Coefficient/Variance (Mean and CI)	**Effective sample size**
Fixed Terms		
Intercept	145.561 (132.021 - 159.748)	15776
Random Terms		
Year	0.925 (0 - 2.733)	477
Species	53.627 (0 - 207.052)	149
Colony	51.42 (0 - 168.133)	65
Colony_Species	56.516 (1.296 - 161.247)	171
Site	143.266 (0 - 476.934)	92
Breeding region	177.95 (0 - 374.666)	399
Sea	123.878 (0 - 446.919)	252
Species: Arctic tern.Year	1.207 (0 - 4.361)	5089
Species: Atlantic puffin.Year	6.705 (0 - 19.955)	1338
Species: Black-legged kittiwake.Year	7.973 (1.105 - 15.787)	1487
Species: Brünnich's guillemot.Year	5.367 (0 - 12.582)	2599
Species: Common guillemot.Year	0.673 (0 - 2.163)	3229
Species: Common tern.Year	2.092 (0 - 5.628)	1896
Species: European shag.Year	15.108 (0 - 53.512)	16892

Species: Roseate tern.Year	1.176 (0 - 4.336)	2957
Site: Bird Island.Year	5.35 (0 - 11.851)	1116
Site: Country Island.Year	26.841 (0 - 59.176)	1457
Site: Isle of May.Year	11.19 (0 - 21.733)	1198
Site: Machias Seal Island.Year	10.808 (0 - 32.518)	3799
Site: Prince Leopold Island.Year	22.67 (0 - 49.814)	4323
Site: Ram Island.Year	6.357 (0 - 13.547)	3422
Site: Røst.Year	43.575 (0 - 86.724)	2767
Site: Sumburgh Head.Year	6.35 (0 - 12.875)	3216
Site group: Buzzards Bay.Year	4.531 (0 - 10.452)	621
Site group: Shetland.Year	31.056 (10.978 - 55.771)	707
Breeding region: Buzzards Bay, common tern.Year	5.211 (0 - 13.137)	648
Breeding region: Buzzards Bay, roseate tern.Year	1.178 (0 - 4.31)	2578
Breeding region: Shetland, black-legged kittiwake.Year	10.474 (0 - 25.672)	387
Sea: North Sea.Year	9.655 (0 - 20.003)	976
Sea: Scotian Shelf.Year	8.566 (0 - 26.127)	741
Residual Terms		
Population: bantersee_commontern.csv.Year	24.472 (7.021 - 46.899)	10333
Population: birdisland_commontern.csv.Year	4.443 (0.002 - 10.479)	1236
Population: birdisland_roseatetern.csv.Year	4.223 (0.002 - 10.012)	1277
Population: burravoe_blackleggedkittiwake.csv.Year	0.841 (0.002 - 3.42)	10908
Population: coatsisland_thickbilledmurre.csv.Year	4.756 (0.002 - 11.864)	2416
Population: countryisland_arctictern.csv.Year	0.743 (0.001 - 3.241)	9051
Population: countryisland_commontern.csv.Year	1.679 (0.001 - 5.986)	8115
Population: eschaness_blackleggedkittiwake.csv.Year	23.812 (4.784 - 55.67)	24000
Population: hornoya_blackleggedkittiwake.csv.Year	143.036 (33.056 - 321.912)	22127
Population: isleofmay_atlanticpuffin.csv.Year	13.118 (0.003 - 25.03)	1340
Population: isleofmay_blackleggedkittiwake.csv.Year	19.313 (7.674 - 34.419)	23186
Population: isleofmay_commonguillemot.csv.Year	0.483 (0.002 - 1.95)	3858

Population: isleofmay_europeanshag.csv.Year	348.49 (180.596 - 538.863)	23215
Population: isleofmay_razorbill.csv.Year	5.396 (2.481 - 8.926)	14057
Population: kettlaness_blackleggedkittiwake.csv.Year	0.843 (0.001 - 2.854)	16104
Population: kongsfjorden_blackleggedkittiwake.csv.Year	1.578 (0.002 - 5.839)	5628
Population: machiassealisland_arctictern.csv.Year	5.34 (0.002 - 22.41)	4188
Population: machiassealisland_atlanticpuffin.csv.Year	46.824 (0.003 - 95.056)	4503
Population: penikeseisland_commontern.csv.Year	1.672 (0.001 - 5.927)	9347
Population: princeleopoldisland_blackleggedkittiwake.csv.Year	5.835 (0.001 - 24.258)	4786
Population: princeleopoldisland_thickbilledmurre.csv.Year	15.499 (0.002 - 38.598)	3840
Population: ramisland_commontern.csv.Year	0.509 (0.002 - 2.151)	8411
Population: ramisland_roseatetern.csv.Year	2.958 (0.002 - 10.561)	860
Population: ramnageo_blackleggedkittiwake.csv.Year	7.137 (1.851 - 14.411)	17027
Population: rost_atlanticpuffin.csv.Year	35.689 (0.002 - 97.391)	2915
Population: rost_blackleggedkittiwake.csv.Year	21.022 (0.002 - 89.204)	6513
Population: rost_europeanshag.csv.Year	170.378 (73.764 - 285.079)	11500
Population: storakarlso_commonguillemot.csv.Year	8.558 (1.346 - 19.515)	18865
Population: sumburgh_blackleggedkittiwake.csv.Year	3.003 (0.002 - 9.313)	2572
Population: sumburgh_commonguillemot.csv.Year	7.344 (0.002 - 22.824)	1612
Population: sumburgh_europeanshag.csv.Year	127.454 (47.53 - 221.698)	19849
Population: troswickness_blackleggedkittiwake.csv.Year	27.434 (4.655 - 63.961)	15584
Population: westerwick_blackleggedkittiwake.csv.Year	0.644 (0.001 - 2.274)	12293

Table A.7. Meta data used in this thesis are available at

https://github.com/katkeogan/PhDthesis
Chapter 3 Metadata

A.3 Further details regarding the method for estimating among year (co)variance

One approach to inferring the covariance among populations would be to simply estimate the unstructured covariance between all pairs of populations. However, there are too few years in our dataset for this to be feasible across the number of populations that we have. Also, this approach does not readily lend itself to inference of the factors that determine among population covariance.

Our approach estimating the among year variance for populations belonging to a group, but allows for heterogeneous variance across different groups (Hadfield, 2010). Below we present an explanation of how the (co)variances estimated via different random terms combine.

Below we present an example involving six populations and present the MCMCglmm syntax to fit each random term in parentheses.

By including a year term (year) the estimated (co)variance among populations is

$$\begin{bmatrix} Vy & Vy & Vy & Vy & Vy & Vy \\ Vy & Vy & Vy & Vy & Vy & Vy \\ Vy & Vy & Vy & Vy & Vy & Vy \\ Vy & Vy & Vy & Vy & Vy & Vy \\ Vy & Vy & Vy & Vy & Vy & Vy \\ Vy & Vy & Vy & Vy & Vy & Vy \end{bmatrix}$$

Populations 1, 2 and 3 represent one species and populations 4, 5, and 6 represent another. Allowing for heterogeneous year variance (idh(species):year) across species, *s*, describes the following pattern of among population covariance

$$\begin{bmatrix} Vys_1 & Vys_1 & Vys_1 & 0 & 0 & 0 \\ Vys_1 & Vys_1 & Vys_1 & 0 & 0 & 0 \\ Vys_1 & Vys_1 & Vys_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & Vys_2 & Vys_2 & Vys_2 \\ 0 & 0 & 0 & Vys_2 & Vys_2 & Vys_2 \\ 0 & 0 & 0 & Vys_2 & Vys_2 & Vys_2 \end{bmatrix}$$

Populations 1, 3 and 5 come from one region and 2, 4 and 6 come from another. Allowing for the year variance to be heterogeneous (idh(region):year) across regions, *r*, describes the following pattern of among population covariance

$$\begin{bmatrix} Vyr_1 & 0 & Vys_1 & 0 & Vyr_1 & 0 \\ 0 & Vyr_2 & 0 & Vyr_2 & 0 & Vyr_2 \\ Vyr_1 & 0 & Vyr_1 & 0 & Vyr_1 & 0 \\ 0 & Vyr_2 & 0 & Vyr_2 & 0 & Vyr_2 \\ Vyr_1 & 0 & Vyr_1 & 0 & Vyr_1 & 0 \\ 0 & Vyr_2 & 0 & Vyr_2 & 0 & Vyr_2 \end{bmatrix}$$

Populations 1 and 3 come from one colony and populations 4 and 6 come from another. Allowing for the year variance to be heterogeneous (idh(colony):year) across colonies, *c*, describes the following pattern of among population covariance

$$\begin{bmatrix} Vyc_1 & 0 & Vyc_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ Vyc_1 & 0 & Vyc_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & Vyc_2 & 0 & Vyc_2 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & Vyc_2 & 0 & Vyc_2 \end{bmatrix}$$

In addition the residual variance, e, is allowed to be heterogeneous across populations (rcov=idh(population):year). Combining these sources of (co)variance gives us the following

$$\begin{bmatrix} Vy + Vys_1 + Vyr_1 + Vyc_1 + e_1 & Vy + Vys_1 & Vy + Vys_1 + Vyr_1 + Vyc_1 & Vy & Vy + Vyr_1 & Vy \\ Vy + Vys_1 & Vy + Vys_1 + Vyr_2 + e_2 & Vy + Vys_1 & Vy + Vyr_2 & Vy & Vy + Vyr_2 \\ Vy + Vys_1 + Vyr_1 + Vyc_1 & Vy + Vys_1 & Vy + Vys_1 + Vyr_1 + Vyc_1 + e_3 & Vy & Vy + Vyr_1 & Vy \\ Vy & Vy + Vyr_2 & Vy & Vy + Vys_2 + Vyr_2 + Vyc_2 + e_4 & Vy + Vys_2 & Vy + Vys_2 + Vyr_2 + Vyc_2 \\ Vy + Vyr_1 & Vy & Vy + Vyr_1 & Vy + Vys_2 & Vy + Vys_2 + Vyr_1 + e_5 & Vy + Vys_2 \\ Vy & Vy + Vyr_2 & Vy & Vy + Vys_2 + Vyr_2 + Vyc_2 & Vy + Vys_2 & Vy + Vys_2 + Vyr_2 + Vyc_2 + e_6 \end{bmatrix}$$

We can use these (co)variances to derive model based estimates of among population correlation. For instance, the correlation between populations 1 and 3 is given by

$$cor_{1,2} = \frac{(Vy + Vys_1 + Vyr_1 + Vyc_1)}{\sqrt{(Vy + Vys_1 + Vyr_1 + Vyc_1 + e_1)}\sqrt{(Vy + Vys_1 + Vyr_1 + Vyc_1 + e_3)}}$$

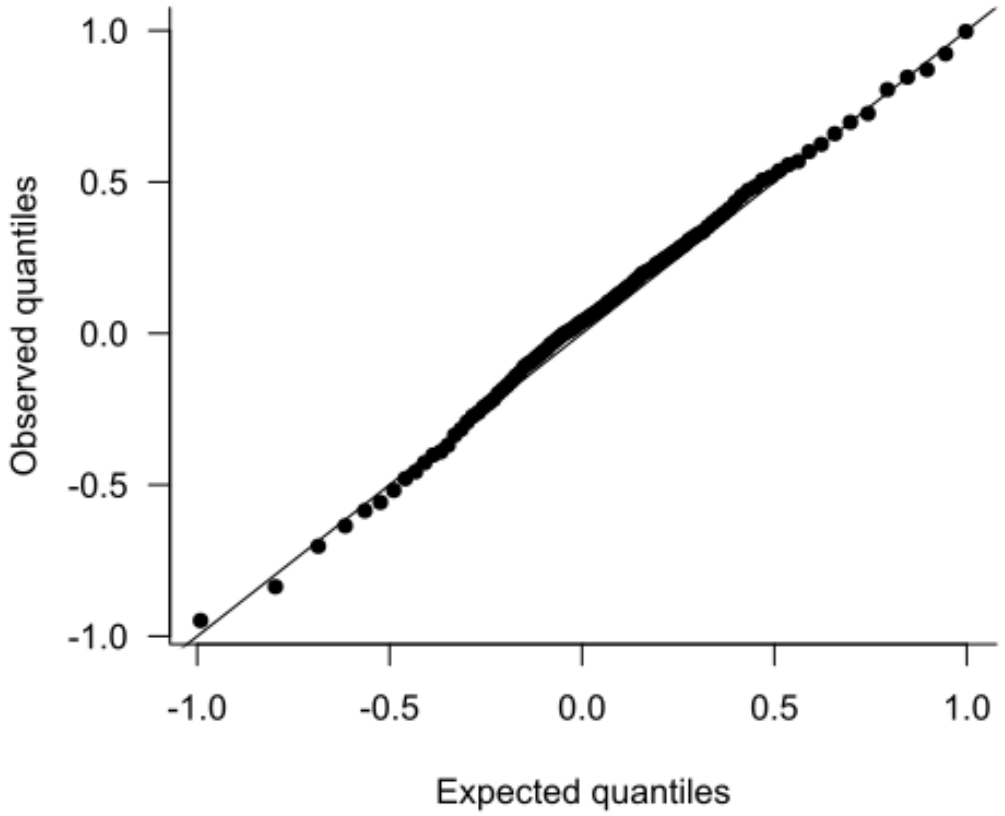


Figure A.2. Q-Q plot of observed versus expected pairwise correlations. The expected relationship was generated under the core model using 1000 a posteriori simulations, with pairwise correlations in the resulting data calculated, and the expected correlations summarised as the mean quantile values. Each point represents a quantile, and the black line represents a 1:1 relationship.

B. Chapter 4 Appendix

Table B.1. Coefficients (and CIs) from core model that tests mismatch hypotheses as predictors of variation in breeding success from the first laying attempt between annual population means (superscript p) and within years among individuals (superscript i). Significant terms highlighted in bold.

Shags - Core Model	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	**3.161 (0.914 - 5.473)**	9000
Mean lay date^{p}	**-0.035 (-0.053 - -0.017)**	9000
Relative lay date^{i}	**-0.026 (-0.034 - -0.019)**	4702
Relative lay date (quadratic)^{i}	**-0.0007 (-0.0009 - -0.0005)**	1388
Random Terms		
(Intercept):(Intercept).year	**0.486 (0.21 - 0.831)**	1882
relative:(Intercept).year	0.002 (-0.003 - 0.008)	3380
relative:relative.year	1e-04 (0 - 4e-04)	2307

Table B.2.a Coefficients (and CIs) from core model that tests mismatch hypotheses as predictors of variation in breeding success from all attempts between annual population means (superscript p) and within years among individuals (superscript i). Significant terms highlighted in bold.

Shags - Core Model	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	**3.258 (0.926 - 5.413)**	8489
Mean lay date^{p}	**-0.035 (-0.053 - -0.017)**	8399
Relative lay date^{i}	**-0.034 (-0.041 - -0.029)**	4555
Relative lay date (quadratic)^{i}	**-0.0004 (-0.0007 - -0.0003)**	1542
Random Terms		
(Intercept):(Intercept).year	**0.481 (0.23 - 0.819)**	4096
relative:(Intercept).year	0.002 (-0.002 - 0.008)	2239
relative:relative.year	7e-05 (0 – 0.0002)	2407

Table B.3.b Coefficients (and CIs) from core model that tests mismatch hypotheses as predictors of variation in breeding success from all attempts between annual population means (superscript p) and within years among individuals (superscript i). Significant terms highlighted in bold.

Shags - Core Model	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	**3.774 (0.001 – 7.060)**	8581
Mean lay date^{p}	**-0.034 (-0.053 - -0.016)**	8215
Relative lay date^{i}	**-0.026 (-0.034 - -0.019)**	5097
Relative lay date (quadratic)^{i}	**-0.0007 (-0.0009 - -0.0005)**	1416
Population size (log transformed)	-0.112 (-0.607 – 0.362)	8410
Random Terms		
(Intercept):(Intercept).year	**0.497 (0.22 - 0.857)**	2465
relative:(Intercept).year	0.003 (-0.003 - 0.009)	3122
relative:relative.year	0.0001 (0 – 0.0004)	2223

Table B.4. Coefficients (and CIs) from temporal model that tests mismatch hypotheses as predictors of variation in breeding success between annual population means (superscript p) and within years among individuals (superscript i). Significant terms highlighted in bold.

Shags - Time Model	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	1.951 (-0.673 - 4.546)	7754
Mean lay date^{p}	**-0.026 (-0.047 - -0.007)**	7738
Relative lay date^{i}	**-0.027 (-0.035 - -0.02)**	4798
Relative lay date (quadratic)^{i}	**-7e-04 (-9e-04 - -5e-04)**	1369
Year (mean centred)^{p}	0.029 (-0.008 - 0.064)	8518
Relative lay date:Year (mean centred)^{i}	7e-05 (-0.0008 - 0.0009)	5699
Random Terms		
(Intercept):(Intercept).year	**0.456 (0.196 - 0.794)**	2664
relative:(Intercept).year	0.003 (-0.002 - 0.01)	2839
relative:relative.year	2e-04 (0 - 4e-04)	2374

Table B.5. Coefficients (and CIs) from past sea surface temperature model that tests mismatch hypotheses as predictors of variation in breeding success between annual population means (superscript p) and within years among individuals (superscript i). Significant terms highlighted in bold.

Shags - SST -1 Model	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	3.301 (-1.204 - 7.54)	8508
Mean lay date^{p}	**-0.027 (-0.047 - -0.006)**	8116
Relative lay date^{i}	-0.055 (-0.141 - 0.024)	5200
Relative lay date (quadratic)^{i}	**-0.0006 (-0.0009 - -0.0005)**	1310
Year (mean centred)	0.03 (-0.005 - 0.065)	7987
Inshore SST (past)^{p}	-0.203 (-0.75 - 0.328)	9000
Inshore SST (past): Relative lay date^{i}	0.005 (-0.009 - 0.019)	5389
Random Terms		
(Intercept):(Intercept).year	**0.462 (0.203 - 0.809)**	2998
relative:(Intercept).year	0.003 (-0.003 - 0.01)	2436
relative:relative.year	2e-04 (0 - 4e-04)	2179

Table B.6. Coefficients (and CIs) from present sea surface temperature model that tests mismatch hypotheses as predictors of variation in breeding success between annual population means (superscript p) and within years among individuals (superscript i). Significant terms highlighted in bold.

Shags - SST Model	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	1.811 (-2.421 - 6.168)	8406
Mean lay date^{p}	**-0.026 (-0.047 - -0.006)**	8443
Relative lay date^{i}	-0.008 (-0.108 - 0.091)	3991
Relative lay date (quadratic)^{i}	**-0.0007 (-0.0009 - -0.0005)**	1456
Year (mean centred)	0.028 (-0.008 - 0.063)	8384
Inshore SST (present)^{p}	0.029 (-0.592 - 0.652)	9000
Inshore SST (present): Relative lay date^{i}	-0.003 (-0.021 - 0.013)	4084
Random Terms		
(Intercept):(Intercept).year	**0.479 (0.209 - 0.841)**	2594
relative:(Intercept).year	0.003 (-0.003 - 0.01)	3357
relative:relative.year	2e-04 (0 - 4e-04)	2020

Table B.7. Coefficients (and CIs) from core sandeel model (proportion of 1+: 0 group sandeels) that tests mismatch hypotheses as changes in proportion of 1+

sandeels between annual sample means (superscript p) and within years among samples (superscript i). Significant terms highlighted in bold.

Sandeels - Core 1+:0 group	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	7.04 (-2.921 - 17.744)	5241
Mean collection date^{p}	-0.026 (-0.084 - 0.028)	5051
Relative collection date^{i}	**-0.096 (-0.144 - -0.049)**	9110
Random Terms		
(Intercept):(Intercept).year	**6.86 (3.351 - 11.242)**	1757
relative:(Intercept).year	**0.189 (0.036 - 0.369)**	6006
relative:relative.year	**0.0161 (0.0076 - 0.0272)**	2872

Table B.8. Coefficients (and CIs) from expanded sandeel model (proportion of 1+: 0 group sandeels) that tests mismatch hypotheses as changes in proportion of 1+ sandeels between annual sample means (superscript p) and within years among samples (superscript i). Significant terms highlighted in bold.

Sandeels - Expanded 1+:0 group	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	9.324 (-2.384 - 21.366)	8327
Mean collection date^{p}	-0.038 (-0.106 - 0.024)	8322
Relative collection date^{i}	**-0.092 (-0.14 - -0.043)**	9000
Year (mean centred)^{p}	0.007 (-0.126 - 0.138)	9000
Relative collection date: Year (centred)^{i}	0.004 (-0.002 - 0.01)	9000
Random Terms		
(Intercept):(Intercept).year	**6.855 (3.461 - 11.189)**	2077
relative:(Intercept).year	**0.194 (0.039 - 0.375)**	8718
relative:relative.year	**0.0161 (0.0071 - 0.0269)**	2810

Table B.9. Coefficients (and CIs) from bivariate model (chicks fledged : 1+ : 0 group sandeels) that tests mismatch hypotheses as the covariance between changes in proportion of chicks fledged and changes in proportion of 1+ sandeels, between annual population and sample means (superscript p) and within years among populations and samples (superscript i). Significant terms highlighted in bold.

Bivariate - Chicks fledged: SE1:SE0	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept - Chicks fledged	1.32 (-0.018 - 2.72)	5287
Intercept - Proportion of Sandeels	**14.999 (7.487 - 22.207)**	1292
Chicks fledged:Mean lay^{p}	**-0.018 (-0.029 - -0.007)**	5412
Proportion of sandeels:mean lay^{p}	**-0.0213 (-0.0281 - -0.0141)**	1041
Chicks fledged:relative lay^{i}	**-0.021 (-0.028 - -0.014)**	5321

Proportion of sandeels:relative lay[i]	**-0.042 (-0.067 - -0.02)**	1862
Relative lay date (quadratic)[i]	-0.001 (-0.001 - 0)	494
Random Terms		
Mean shag lay:mean sandeel sample[p]	-0.078 (-0.561 - 0.297)	823
relative shag lay:relative sandeel sample[i]	0 (0 - 0)	1456
Residual variance		
Residual - chicks fledged	**0.949 (0.676 - 1.228)**	1207
Residual - sandeel sample	**4.929 (4.195 - 5.633)**	7038

B.1 Models with Poisson error structure

All models were estimated using a binomial family error structure, because breeding success was underdispersed as compared with the expectation under a Poisson process. However, to ensure there was no qualitative difference in the results, we reran each shag and co-variance (shag & sandeel) model assuming a Poisson error. In this case, the response variable was the number of chicks fledged per nest, with all fixed and random effects the same as described by equations 2 & 6 (main text) and in Table 4.1. Parameter expanded priors were used for all models, and the residual variance was not fixed at 1.

Supplementary results

There was no qualitative difference in the results under a Poisson error structure for shag (Table B.8) or co-variance (Table B.9) models.

Table B.10. Key results from core models, assuming a Poisson family error structure.

	As average lay date increases	over time	SST	SST -1
Change in mean breeding success	**-0.020 (-0.032, -0.009)**	0.020 (-0.007, 0.045)	0.057 (-0.404, 0.504)	-0.129 (-0.524, 0.260)
Change in relative breeding success (linear)	**-0.017 (-0.022, -0.012)**	0.0002 (-0.0004, 0.0007)	-0.002 (-0.014, 0.009)	-0.004 (-0.007, 0.015)
Change in relative breeding success (quadratic)	**-0.0005 (-0.0007, -0.0004)**	-	-	-

Table B.11. Key results from co-variance model, assuming a Poisson family error structure.

	1+ : 0 group
Between-year proportions	
Covariance with mean shag breeding success	-0.040 (-0.276, 0.141)
Within-year proportions	
Covariance with slope of relative shag fitness	0.000009 (-0.0002, 0.0002)

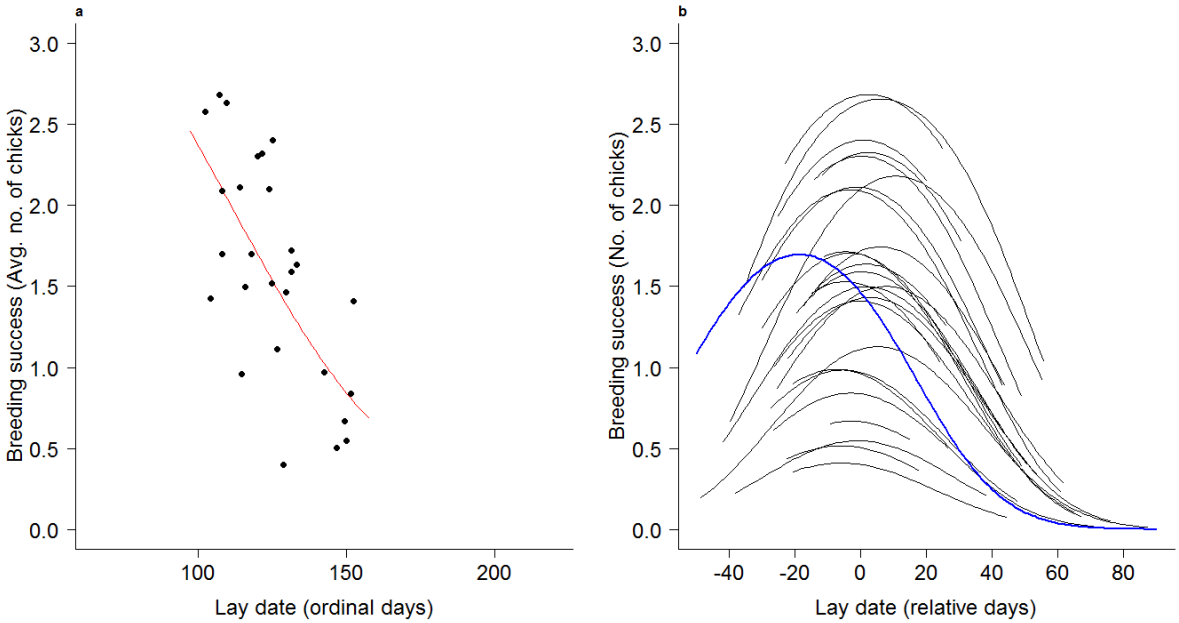


Figure B.1. Comparisons of mean (a) and relative (b) fitness on the normal scale. **a**) slope in red (from the basic model) and annual means as black points (from the raw data) of mean lay date in relation to fitness. **b)** average relative fitness (blue) slope and annual relative fitness slopes from the random regression (black), with lay date on the relative scale for comparison of slopes across years.

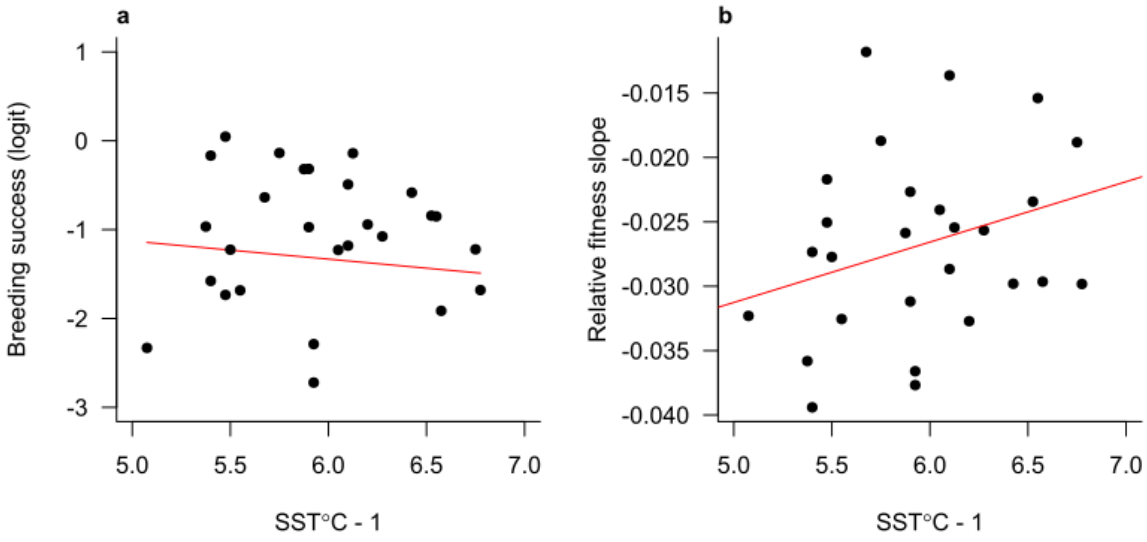


Figure B.2. The effect of SST in the previous year (a) on breeding success (logit transformed) at the population level. Changes in strength of selection with SST in the previous year (b). Red lines indicate average response.

C. Chapter 5 Appendix

Table C.1. Coefficients (and Credible Intervals) from the model (1) which tests the effect of time on temperature change. Year is centred on 2003.

Response = Temperature	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	8.174 (5.83, 10.559)	25159
Year (centred)	0.014 (0.005, 0.025)	13719
Random Terms		
Year (factor)	0.044 (0.015, 0.078)	567
Site	34.009 (15.834, 55.908)	24700
Year (factor):Site	0.118 (0.084, 0.153)	5988
(Intercept):(Intercept).population	1.458 (0.808, 2.25)	22393
Year (centred):(Intercept).population	0.004 (-0.003, 0.014)	1478
(Intercept):Year (centred).population	0.004 (-0.003, 0.014)	1478
Year (centred):Year (centred).population	0 (0, 0)	267
Standard Error temperature.units	1 (1, 1)	0
Units	0.001 (0, 0.001)	514

Table C.2. Coefficients (and Credible Intervals) from the model (2) which tests the effect of time and temperature on annual population average lay date, but where annual phenology is not weighted by standard error. Year is centred on 2003. Temperature is centred on 6°C.

Response = Lay date (unweighted)	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	141.854 (129.942, 153.112)	24554
Year (centred)	-0.054 (-0.163, 0.056)	24194
Temperature (centred)	-0.311 (-1.751, 1.099)	24700
Random Terms		
Year (factor)	0.714 (0, 2.148)	14024
Site	651.337 (168.718, 1246.881)	2652
Year (factor):Site	22.527 (16.732, 28.661)	22598
(Intercept):(Intercept).population	288.77 (164.549, 437.358)	3136
Year (centred):(Intercept).population	2.053 (0.245, 4.13)	25062
Temperature (centred):(Intercept).population	8.205 (-23.021, 40.795)	9567
(Intercept):Year (centred).population	2.053 (0.245, 4.13)	25062
Year (centred):Year (centred).population	0.089 (0.043, 0.141)	15543
Temperature (centred):Year (centred).population	-0.138 (-0.6, 0.296)	23525
(Intercept):Temperature (centred).population	8.205 (-23.021, 40.795)	9567
Year (centred):Temperature (centred).population	-0.138 (-0.6, 0.296)	23525
Temperature (centred):Temperature (centred).population	17.642 (9.805, 26.292)	11448
Units	27.126 (24.308, 30.061)	23659

Table C.3. Coefficients (and Credible Intervals) from the model (2a) which tests the effect of time and temperature on annual population average lay date, where annual phenology is weighted by standard error. Year is centred on 2003. Temperature is centred on 6°C.

Model 2a - Lay date (weighted)	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	158.801 (141.131, 176.04)	24700
Year (centred)	-0.067 (-0.176, 0.041)	23248
Temperature (centred)	-0.596 (-1.968, 0.79)	24034
Random Terms		
Year (factor)	0.964 (0, 2.662)	11180
Site	1553.271 (596.425, 2755.237)	15525
Year (factor):Site	21.995 (16.373, 28.208)	21544
(Intercept):(Intercept).population	285.945 (160.662, 433.113)	4468
Year (centred):(Intercept).population	1.679 (-0.131, 3.582)	23275
Temperature (centred):(Intercept).population	16.457 (-19.161, 54.005)	8937
(Intercept):Year (centred).population	1.679 (-0.131, 3.582)	23275
Year (centred):Year (centred).population	0.081 (0.04, 0.128)	16158
Temperature (centred):Year (centred).population	-0.152 (-0.571, 0.262)	24700
(Intercept):Temperature (centred).population	16.457 (-19.161, 54.005)	8937
Year (centred):Temperature (centred).population	-0.152 (-0.571, 0.262)	24700
Temperature (centred):Temperature (centred).population	16.581 (9.343, 24.656)	12377
Standard error phenology.units	1 (1, 1)	0
Units	22.945 (20.173, 25.631)	22736

Table C.4. Coefficients (and Credible Intervals) from the model (3) which tests the effect of time, temperature and average lay date on annual population average breeding success (log transformed mean number of chicks to fledge/nest). Year is centred on 2003. Temperature is centred on 6°C. Mean lay date is centred on day 125 after Jan 1st for Northern hemisphere and July 1st for Southern hemisphere.

Model 3 - Breeding success	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	0.451 (0.387, 0.516)	14215
Year (centred)	-0.002 (-0.003, 0.0003)	23533
Temperature (centred)	-0.014 (-0.032, 0.003)	24040
Mean lay date (centred)	-0.007 (-0.008, -0.005)	23546
Random Terms		
Year (factor)	0.00049 (0, 0.0012)	15277
Site	0.01011 (0, 0.02653)	1112
Year (factor):Site	0.00626 (0.00415, 0.00854)	21349
(Intercept):(Intercept).population	0.02687 (0.01493, 0.04064)	2347
Year (centred):(Intercept).population	1e-04 (-0.00016, 0.00038)	10554
Temperature (centred):(Intercept).population	-0.00306 (-0.00595, -0.00048)	4318
Mean lay date (centred):(Intercept).population	-9e-05 (-0.00038, 2e-04)	11497
(Intercept):Year (centred).population	1e-04 (-0.00016, 0.00038)	10554
Year (centred):Year (centred).population	1e-05 (0, 3e-05)	13037
Temperature (centred):Year (centred).population	-3e-05 (-1e-04, 3e-05)	13663
Mean lay date (centred):Year (centred).population	0 (-1e-05, 1e-05)	19359
(Intercept):Temperature (centred).population	-0.00306 (-0.00595, -0.00048)	4318
Year (centred):Temperature (centred).population	-3e-05 (-1e-04, 3e-05)	13663
Temperature (centred):Temperature (centred).population	0.00081 (5e-05, 0.00172)	20719
Mean lay date (centred):Temperature (centred).population	2e-05 (-4e-05, 9e-05)	15891
(Intercept):Mean lay date (centred).population	-9e-05 (-0.00038, 2e-04)	11497
Year (centred):Mean lay date (centred).population	0 (-1e-05, 1e-05)	19359
Temperature (centred):Mean lay date (centred).population	2e-05 (-4e-05, 9e-05)	15891
Mean lay date (centred):Mean lay date (centred).population	1e-05 (0, 3e-05)	14652
units	0.01305 (0.01155, 0.01461)	21059

Table C.5. Coefficients (and Credible Intervals) from the model (3a) which tests the effect of time, temperature and lay date on annual population average breeding success, and the interaction between temperature and lay date. Year is centred on 2003. Temperature is centred on 6°C. Mean lay date is centred on day 125 after Jan 1st for Northern hemisphere and July 1st for Southern hemisphere.

Model 3a - Breeding success (temperature:lay date interaction)	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	0.452 (0.387, 0.516)	15805
Year (centred)	-0.001 (-0.003, 0)	23517
Temperature (centred)	-0.014 (-0.032, 0.003)	23933
Mean lay date (centred)	-0.006 (-0.008, -0.005)	24083
Temperature:lay date	-0.001 (-0.002, 0)	24700
Random Terms		
Year (factor)	0.00044 (0, 0.00113)	15516
Site	0.01038 (0, 0.02703)	1198
Year (factor):Site	0.00618 (0.00405, 0.00845)	20429
(Intercept):(Intercept).population	0.02624 (0.01488, 0.03965)	3370
Year (centred):(Intercept).population	9e-05 (-0.00017, 0.00035)	9524
Temperature (centred):(Intercept).population	-0.00289 (-0.00565, -0.00031)	3982
Mean lay date (centred):(Intercept).population	-9e-05 (-0.00039, 0.00018)	16044
(Intercept):Year (centred).population	9e-05 (-0.00017, 0.00035)	9524
Year (centred):Year (centred).population	1e-05 (0, 3e-05)	13852
Temperature (centred):Year (centred).population	-2e-05 (-9e-05, 3e-05)	13581
Mean lay date (centred):Year (centred).population	0 (-1e-05, 1e-05)	18558
(Intercept):Temperature (centred).population	-0.00289 (-0.00565, -0.00031)	3982
Year (centred):Temperature (centred).population	-2e-05 (-9e-05, 3e-05)	13581
Temperature (centred):Temperature (centred).population	0.00074 (0, 0.00159)	21478
Mean lay date (centred):Temperature (centred).population	2e-05 (-4e-05, 9e-05)	16134
(Intercept):Mean lay date (centred).population	-9e-05 (-0.00039, 0.00018)	16044
Year (centred):Mean lay date (centred).population	0 (-1e-05, 1e-05)	18558
Temperature (centred):Mean lay date (centred).population	2e-05 (-4e-05, 9e-05)	16134
Mean lay date (centred):Mean lay date (centred).population	1e-05 (0, 3e-05)	15305
units	0.01311 (0.01159, 0.01465)	19902

Table C.6. Coefficients (and Credible Intervals) from the model (4b) which tests the effect of time, temperature and lay date on annual population average breeding success, and the interaction between temperature and latitude. Year is centred on 2003. Temperature is centred on 6°C. Mean lay date is centred on day 125 after Jan 1st for Northern hemisphere and July 1st for Southern hemisphere.

Model 4b - Breeding success (temperature:latitude interaction)	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	0.491 (0.416, 0.572)	22839
Year (centred)	-0.002 (-0.003, 0)	23145
Temperature (centred)	-0.029 (-0.063, 0.003)	24700
Mean lay date (centred)	-0.007 (-0.008, -0.005)	23015
Latitude	-0.001 (-0.003, 0)	23309
Temperature:latitude	0 (0, 0.001)	23820
Random Terms		
Year (factor)	0.00048 (0, 0.00119)	16699
Site	0.00995 (0, 0.02616)	1102
Year (factor):Site	0.00626 (0.00409, 0.00843)	18805
(Intercept):(Intercept).population	0.026 (0.01441, 0.03878)	3251
Year (centred):(Intercept).population	7e-05 (-0.00019, 0.00034)	8779
Temperature (centred):(Intercept).population	-0.00291 (-0.00571, -0.00039)	5089
Mean lay date (centred):(Intercept).population	-0.00014 (-0.00045, 0.00015)	18661
(Intercept):Year (centred).population	7e-05 (-0.00019, 0.00034)	8779
Year (centred):Year (centred).population	1e-05 (0, 3e-05)	14346
Temperature (centred):Year (centred).population	-2e-05 (-9e-05, 4e-05)	14249
Mean lay date (centred):Year (centred).population	0 (-1e-05, 1e-05)	18880
(Intercept):Temperature (centred).population	-0.00291 (-0.00571, -0.00039)	5089
Year (centred):Temperature (centred).population	-2e-05 (-9e-05, 4e-05)	14249
Temperature (centred):Temperature (centred).population	0.00079 (0, 0.00166)	18571
Mean lay date (centred):Temperature (centred).population	3e-05 (-3e-05, 0.00011)	15708
(Intercept):Mean lay date (centred).population	-0.00014 (-0.00045, 0.00015)	18661
Year (centred):Mean lay date (centred).population	0 (-1e-05, 1e-05)	18880
Temperature (centred):Mean lay date (centred).population	3e-05 (-3e-05, 0.00011)	15708
Mean lay date (centred):Mean lay date (centred).population	1e-05 (0, 3e-05)	16211
units	0.01305 (0.01162, 0.01462)	21613

Table C.7. Coefficients (and Credible Intervals) from the model (4b) which tests the effect of time, temperature and lay date on annual population average breeding success, and the interaction between lay date and latitude. Year is centred on 2003. Temperature is centred on 6°C. Mean lay date is centred on day 125 after Jan 1st for Northern hemisphere and July 1st for Southern hemisphere.

Model 4b - Breeding success (lay date:latitude interaction)	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	0.481 (0.407, 0.558)	24700
Year (centred)	-0.002 (-0.003, 0)	22018
Temperature (centred)	-0.013 (-0.031, 0.004)	24700
Mean lay date (centred)	-0.005 (-0.007, -0.002)	24700
Latitude	-0.001 (-0.002, 0)	24106
Lay date:latitude	-0.00006 (-0.0001, -0.00001)	23860
Random Terms		
Year (factor)	0.00045 (0, 0.00115)	14413
Site	0.00921 (0, 0.02411)	1277
Year (factor):Site	0.00621 (0.00402, 0.00843)	20346
(Intercept):(Intercept).population	0.02619 (0.01465, 0.03924)	4036
Year (centred):(Intercept).population	7e-05 (-0.00019, 0.00034)	9960
Temperature (centred):(Intercept).population	-0.00309 (-0.00593, -0.00056)	5562
Mean lay date (centred):(Intercept).population	-2e-04 (-0.00049, 7e-05)	18065
(Intercept):Year (centred).population	7e-05 (-0.00019, 0.00034)	9960
Year (centred):Year (centred).population	1e-05 (0, 3e-05)	13460
Temperature (centred):Year (centred).population	-2e-05 (-9e-05, 4e-05)	14940
Mean lay date (centred):Year (centred).population	0 (-1e-05, 1e-05)	18062
(Intercept):Temperature (centred).population	-0.00309 (-0.00593, -0.00056)	5562
Year (centred):Temperature (centred).population	-2e-05 (-9e-05, 4e-05)	14940
Temperature (centred):Temperature (centred).population	0.00083 (6e-05, 0.00179)	20945
Mean lay date (centred):Temperature (centred).population	4e-05 (-2e-05, 0.00011)	16464
(Intercept):Mean lay date (centred).population	-2e-04 (-0.00049, 7e-05)	18065
Year (centred):Mean lay date (centred).population	0 (-1e-05, 1e-05)	18062
Temperature (centred):Mean lay date (centred).population	4e-05 (-2e-05, 0.00011)	16464
Mean lay date (centred):Mean lay date (centred).population	1e-05 (0, 2e-05)	12619
units	0.01311 (0.01161, 0.01463)	19757

Table C.8. Coefficients (and Credible Intervals) from the model (4c) which tests the effect of time, temperature and lay date on annual population average breeding success, and the interaction between temperature and income/capital breeders. Year is centred on 2003. Temperature is centred on 6°C. Mean lay date is centred on day 125 after Jan 1st for Northern hemisphere and July 1st for Southern hemisphere.

Model 4c - Breeding success (temperature:income/capital interaction)	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	0.479 (0.265, 0.688)	2353
Year (centred)	-0.002 (-0.004, 0.001)	2957
Temperature (centred)	-0.025 (-0.144, 0.083)	2534
Mean lay date (centred)	-0.007 (-0.009, -0.006)	2700
Income/Capital-income	-0.058 (-0.282, 0.164)	2264
Temperature:Income/Capital-income	0.011 (-0.103, 0.128)	2700
Random Terms		
Year (factor)	0.00068 (0, 0.00159)	2209
Site	0.00968 (0, 0.02765)	166
Year (factor):Site	0.00766 (0.00505, 0.0104)	2311
(Intercept):(Intercept).population	0.03002 (0.01545, 0.04661)	827
Year (centred):(Intercept).population	0.00019 (-0.00014, 0.00056)	1116
Temperature (centred):(Intercept).population	-0.00407 (-0.00797, -0.00079)	1283
Mean lay date (centred):(Intercept).population	-1e-04 (-0.00041, 0.00024)	2246
(Intercept):Year (centred).population	0.00019 (-0.00014, 0.00056)	1116
Year (centred):Year (centred).population	2e-05 (0, 3e-05)	1483
Temperature (centred):Year (centred).population	-4e-05 (-0.00014, 4e-05)	1493
Mean lay date (centred):Year (centred).population	0 (-1e-05, 1e-05)	2564
(Intercept):Temperature (centred).population	-0.00407 (-0.00797, -0.00079)	1283
Year (centred):Temperature (centred).population	-4e-05 (-0.00014, 4e-05)	1493
Temperature (centred):Temperature (centred).population	0.00109 (0, 0.00235)	2171
Mean lay date (centred):Temperature (centred).population	2e-05 (-4e-05, 0.00011)	1849
(Intercept):Mean lay date (centred).population	-1e-04 (-0.00041, 0.00024)	2246
Year (centred):Mean lay date (centred).population	0 (-1e-05, 1e-05)	2564
Temperature (centred):Mean lay date (centred).population	2e-05 (-4e-05, 0.00011)	1849
Mean lay date (centred):Mean lay date (centred).population	1e-05 (0, 3e-05)	1443

units	0.01389 (0.0121, 0.01584)	2308

Table C.9. Coefficients (and Credible Intervals) from the model (4d) which tests the effect of time, temperature and lay date on annual population average breeding success, and the interaction between lay date and income/capital breeders. Year is centred on 2003. Temperature is centred on 6°C. Mean lay date is centred on day 125 after Jan 1st for Northern hemisphere and July 1st for Southern hemisphere.

Model 4d - Breeding success (lay date:income/capital interaction)	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	0.478 (0.268, 0.675)	22675
Year (centred)	-0.002 (-0.004, 0)	24700
Temperature (centred)	-0.014 (-0.035, 0.005)	24700
Mean lay date (centred)	-0.002 (-0.008, 0.004)	24700
Income/Capital-income	-0.058 (-0.283, 0.155)	21651
Lay date:Income/Capital-income	-0.006 (-0.012, 0)	24700
Random Terms		
Year (factor)	0.00062 (0, 0.00152)	16572
Site	0.00803 (0, 0.02452)	1259
Year (factor):Site	0.00774 (0.00498, 0.01043)	20884
(Intercept):(Intercept).population	0.03099 (0.01565, 0.04748)	3867
Year (centred):(Intercept).population	2e-04 (-0.00013, 0.00054)	6801
Temperature (centred):(Intercept).population	-0.00393 (-0.00765, -0.00058)	5219
Mean lay date (centred):(Intercept).population	-1e-04 (-0.00041, 0.00017)	20009
(Intercept):Year (centred).population	2e-04 (-0.00013, 0.00054)	6801
Year (centred):Year (centred).population	2e-05 (0, 3e-05)	13966
Temperature (centred):Year (centred).population	-4e-05 (-0.00013, 4e-05)	13231
Mean lay date (centred):Year (centred).population	0 (-1e-05, 1e-05)	20743
(Intercept):Temperature (centred).population	-0.00393 (-0.00765, -0.00058)	5219
Year (centred):Temperature (centred).population	-4e-05 (-0.00013, 4e-05)	13231
Temperature (centred):Temperature (centred).population	0.00102 (5e-05, 0.00221)	20001
Mean lay date (centred):Temperature (centred).population	2e-05 (-5e-05, 9e-05)	19962
(Intercept):Mean lay date (centred).population	-1e-04 (-0.00041, 0.00017)	20009
Year (centred):Mean lay date (centred).population	0 (-1e-05, 1e-05)	20743
Temperature (centred):Mean lay date (centred).population	2e-05 (-5e-05, 9e-05)	19962
Mean lay date (centred):Mean lay date (centred).population	1e-05 (0, 3e-05)	13110
units	0.01387 (0.01205, 0.01574)	20514

Table C.10. Coefficients (and Credible Intervals) from the model (4e) which tests the effect of time, temperature and lay date on annual population average breeding success, and the interaction between temperature and feeding strategy. Year is centred on 2003. Temperature is centred on 6°C. Mean lay date is centred on day 125 after Jan 1st for Northern hemisphere and July 1st for Southern hemisphere.

Model 4e - Breeding success (temperature:feeding interaction)	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Feeding method - single prey	0.476 (0.373, 0.574)	15274
Feeding method - multiple prey	0.505 (0.393, 0.619)	5474
Feeding method - regurgitation	0.404 (0.329, 0.479)	24700
Year (centred)	-0.001 (-0.003, 0.001)	22642
Temperature (centred)	-0.028 (-0.058, 0.005)	23343
Mean lay date (centred)	-0.007 (-0.008, -0.005)	24700
Temperature:Feeding method - single	0.011 (-0.024, 0.046)	23787
Temperature:Feeding method - regurgitation	-0.013 (-0.062, 0.031)	23293
Random Terms		
Year (factor)	0.00052 (0, 0.00126)	15025
Site	0.01023 (0, 0.02711)	1221
Year (factor):Site	0.00639 (0.00426, 0.00867)	21227
(Intercept):(Intercept).population	0.02441 (0.01295, 0.03755)	2813
Year (centred):(Intercept).population	0.00018 (-7e-05, 0.00046)	6968
Temperature (centred):(Intercept).population	-0.00263 (-0.00538, 6e-05)	5783
Mean lay date (centred):(Intercept).population	-9e-05 (-0.00038, 0.00018)	7941
(Intercept):Year (centred).population	0.00018 (-7e-05, 0.00046)	6968
Year (centred):Year (centred).population	1e-05 (0, 2e-05)	14237
Temperature (centred):Year (centred).population	-3e-05 (-1e-04, 2e-05)	13588
Mean lay date (centred):Year (centred).population	0 (-1e-05, 1e-05)	18486
(Intercept):Temperature (centred).population	-0.00263 (-0.00538, 6e-05)	5783
Year (centred):Temperature (centred).population	-3e-05 (-1e-04, 2e-05)	13588
Temperature (centred):Temperature (centred).population	0.00072 (0, 0.00166)	18239
Mean lay date (centred):Temperature (centred).population	2e-05 (-4e-05, 9e-05)	16218
(Intercept):Mean lay date (centred).population	-9e-05 (-0.00038, 0.00018)	7941
Year (centred):Mean lay date (centred).population	0 (-1e-05, 1e-05)	18486

Temperature (centred):Mean lay date (centred).population	2e-05 (-4e-05, 9e-05)	16218
Mean lay date (centred):Mean lay date (centred).population	1e-05 (0, 3e-05)	16664
units	0.01312 (0.01159, 0.01466)	21610

Table C.11. Coefficients (and Credible Intervals) from the model (4f) which tests the effect of time, temperature and lay date on annual population average breeding success, and the interaction between lay date and feeding strategy. Year is centred on 2003. Temperature is centred on 6°C. Mean lay date is centred on day 125 after Jan 1st for Northern hemisphere and July 1st for Southern hemisphere.

Model 4f - Breeding success (lay date:feeding interaction)	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Feeding method - single prey	0.472 (0.373, 0.568)	14186
Feeding method - multiple prey	0.486 (0.377, 0.595)	7988
Feeding method - regurgitation	0.41 (0.339, 0.486)	24700
Year (centred)	-0.001 (-0.003, 0.001)	20292
Temperature (centred)	-0.014 (-0.032, 0.004)	23627
Mean lay date (centred)	-0.005 (-0.008, -0.002)	23583
Lay date:Feeding method - single	-0.004 (-0.009, 0.001)	24700
Lay date:Feeding method - regurgitation	-0.012 (-0.063, 0.029)	23032
Random Terms		
Year (factor)	0.00051 (0, 0.00125)	16250
Site	0.00896 (0, 0.02483)	1171
Year (factor):Site	0.00635 (0.00425, 0.00869)	22050
(Intercept):(Intercept).population	0.02487 (0.01345, 0.03809)	2810
Year (centred):(Intercept).population	0.00019 (-7e-05, 0.00047)	6895
Temperature (centred):(Intercept).population	-0.00291 (-0.00577, -0.00028)	5077
Mean lay date (centred):(Intercept).population	-9e-05 (-0.00039, 0.00018)	13841
(Intercept):Year (centred).population	0.00019 (-7e-05, 0.00047)	6895
Year (centred):Year (centred).population	1e-05 (0, 3e-05)	12408
Temperature (centred):Year (centred).population	-4e-05 (-0.00011, 2e-05)	12668
Mean lay date (centred):Year (centred).population	0 (-1e-05, 1e-05)	17358
(Intercept):Temperature (centred).population	-0.00291 (-0.00577, -0.00028)	5077
Year (centred):Temperature (centred).population	-4e-05 (-0.00011, 2e-05)	12668
Temperature (centred):Temperature (centred).population	0.00083 (0, 0.00177)	17609
Mean lay date (centred):Temperature (centred).population	2e-05 (-5e-05, 9e-05)	16557
(Intercept):Mean lay date (centred).population	-9e-05 (-0.00039, 0.00018)	13841
Year (centred):Mean lay date (centred).population	0 (-1e-05, 1e-05)	17358
Temperature (centred):Mean lay date (centred).population	2e-05 (-5e-05, 9e-05)	16557

Mean lay date (centred):Mean lay date (centred).population	2e-05 (0, 3e-05)	16138
units	0.01308 (0.01156, 0.01457)	21669

Table C.12. Coefficients (and Credible Intervals) from the post hoc model which tests the effect of time, temperature and lay date on annual population average breeding success, and the interaction between lay date and between-year variation in average lay date. Year is centred on 2003. Temperature is centred on 6°C. Mean lay date is centred on day 125 after Jan 1st for Northern hemisphere and July 1st for Southern hemisphere.

***Post hoc* test- Breeding success (lay date:between-year variation lay date (days))**	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	0.443 (0.345, 0.546)	20135
Year (centred)	-0.002 (-0.003, 0)	22934
Temperature (centred)	-0.013 (-0.031, 0.005)	23865
Mean lay date (centred)	-0.007 (-0.01, -0.003)	24000
Between-year variation in lay date (days)	0 (-0.003, 0.003)	21520
Lay date:Between-year variation in lay date (days)	1.4e-06 (-8.43e-05, 8.27e-05)	23263
Random Terms		
Year (factor)	0.00049 (0, 0.00121)	15973
Site	0.01035 (0, 0.02801)	1221
Year (factor):Site	0.00627 (0.0041, 0.00848)	21202
(Intercept):(Intercept).population	0.02724 (0.01535, 0.04132)	2746
Year (centred):(Intercept).population	0.00011 (-0.00016, 0.00038)	9374
Temperature (centred):(Intercept).population	-0.00315 (-0.00608, -0.00048)	3632
Mean lay date (centred):(Intercept).population	-9e-05 (-0.00041, 2e-04)	15652
(Intercept):Year (centred).population	0.00011 (-0.00016, 0.00038)	9374
Year (centred):Year (centred).population	1e-05 (0, 3e-05)	14385
Temperature (centred):Year (centred).population	-3e-05 (-9e-05, 3e-05)	13832
Mean lay date (centred):Year (centred).population	0 (-1e-05, 1e-05)	18510
(Intercept):Temperature (centred).population	-0.00315 (-0.00608, -0.00048)	3632
Year (centred):Temperature (centred).population	-3e-05 (-9e-05, 3e-05)	13832
Temperature (centred):Temperature (centred).population	0.00083 (7e-05, 0.00178)	20671
Mean lay date (centred):Temperature (centred).population	2e-05 (-4e-05, 1e-04)	16647
(Intercept):Mean lay date (centred).population	-9e-05 (-0.00041, 2e-04)	15652
Year (centred):Mean lay date (centred).population	0 (-1e-05, 1e-05)	18510
Temperature (centred):Mean lay date (centred).population	2e-05 (-4e-05, 1e-04)	16647

Mean lay date (centred):Mean lay date (centred).population	2e-05 (0, 3e-05)	13798
units	0.01303 (0.01154, 0.01455)	20997

Table C.13. Coefficients (and Credible Intervals) from the post hoc model which tests the effect of time on average annual population breeding success. Year is centred on 2003.

Response = Breeding success	Coefficient/Variance (Mean and CI)	Effective sample size
Fixed Terms		
Intercept	0.45 (0.387, 0.514)	20414
Year (centred)	-0.001 (-0.003, 0.001)	24700
Random Terms		
Year (factor)	0.00081 (0, 0.00175)	14827
Site	0.00923 (0, 0.02569)	1432
Year (factor):Site	0.00816 (0.00563, 0.01082)	21446
(Intercept):(Intercept).population	0.02606 (0.01407, 0.03895)	3174
Year (centred):(Intercept).population	1e-04 (-0.00023, 0.00041)	14881
(Intercept):Year (centred).population	1e-04 (-0.00023, 0.00041)	14881
Year (centred):Year (centred).population	3e-05 (1e-05, 4e-05)	18341
Units	0.01476 (0.01319, 0.01645)	23240

Table C.14. Meta data used in this thesis are available at

https://github.com/katkeogan/PhDthesis
Chapter 5 - Meta data

Table C.15. Population level results are available at

https://github.com/katkeogan/PhDthesis
Chapter 5 - Results