

Universidade de Lisboa  
Faculdade de Ciências  
Departamento de Estatística e Investigação Operacional



# Estimating the spatio-temporal variation of bird phenology using citizen science data

Pedro Guilherme Antunes Lopes da Silva Nicolau

Mestrado em Bioestatística

Dissertação orientada por:  
Doutor Tiago André Marques  
Doutora Alison Johnston

2017



# Acknowledgements

First and foremost, I want to express endless gratitude to my two incredible supervisors. To Ali, for giving me a chance, for initiating this project and for being my personal mentor. For teaching me a number of statistical methodologies, to code more efficiently, to write science, and numerous life hacks which have deeply improved my work productivity. To Tiago, from bringing calm amidst storm, for the knowledge and for opening my horizons, luring me into the stats world.

A special thanks goes to the amazing organisation that is the British Trust for Ornithology. My home for 6 months, where the full project was conceived and developed, and whose data I used in this work. In addition to thanking all the staff for their good will, friendliness, baking abilities (cakes!) and treating me as one of their own, I want to specifically show gratitude to a number of people.

To the team members who directly worked on this project, providing various input in a number of project meetings and exchanges of emails: Dave Leech, Nick Moran and Stephen Baillie. Particularly to Dave, for all the vast ecological knowledge and brainstorming sessions.

To the Pops office team, for all the friendly chats and help provided: Dario Massimino, Jenny Border, Sam Franks, Stuart Newson, Simon Gillings and my friend Fabrizio.

For diverse reasons, to Andy Clements, Andy Musgrove, Dawn Balmer, Carole Showell, Carl Barimore, Sam Massimino, Paul Stancliffe, Gavin Siriwardena and Bonita Johnson.

To Malcolm Burgess for providing additional data to empirically validate our models.

To Stephen, my great companion and birding mate, for lending me a room while in the UK, for making me company and helping me with any problems I had along the way. For becoming my friend.

To all my friends, for the endless fun and lengthy discussions on controversial topics, providing a stable and stimulating environment for me to thrive.

To Filipa. For making me a better person, for support and warmth when times get hard, always encouraging me to grow, staying true to myself and never settle for anything less than great.

To my sister, Joana, for taking good care of me and helping me prepare for this big journey.

To my parents, Jorge and Luísa, for making me a critical person, rigorous with myself, inciting me to always fight for what I think is right and for dreams.

To the Biostatistics Masters' professors, which taught me immensely in these last two years and gave me the basis to grow upon.

Finally, to the ERASMUS+ program, which gave me financial support to go abroad, meet new cultures, different people, and experience life in the UK.

# Contents

<b>Resumo</b>	<b>ix</b>
<b>Abstract</b>	<b>xii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Biological Problem . . . . .	2
1.2 Citizen Science . . . . .	3
1.3 Study Species . . . . .	5
1.3.1 Barn Swallow . . . . .	5
1.3.2 Pied Flycatcher . . . . .	6
1.4 Approach . . . . .	7
1.4.1 Content of the thesis . . . . .	7
<b>2 Statistical Background - Generalized Additive Models</b>	<b>8</b>
2.1 Generalized Additive Model . . . . .	9
2.1.1 Definition . . . . .	9
2.1.2 Smooth Functions . . . . .	10
2.1.3 Types of Smoothers . . . . .	11
2.2 Package <i>mgecv</i> . . . . .	13
2.2.1 <i>gam</i> function . . . . .	13
2.2.2 <i>bam</i> function . . . . .	13
2.2.3 Shrinkage Selection with <i>mgecv</i> . . . . .	13
2.3 Bootstrapping . . . . .	14
<b>3 Arrival on Breeding Grounds</b>	<b>16</b>
3.1 BirdTrack Data . . . . .	17
3.2 Data Processing . . . . .	17

3.2.1	Variables Description . . . . .	19
3.3	Model Specifications . . . . .	21
3.3.1	Barn Swallow Model . . . . .	21
3.3.2	Pied Flycatcher . . . . .	23
3.4	Model Validation . . . . .	25
3.5	Extracting the Phenological Signal . . . . .	25
3.6	Bootstrapping . . . . .	28
3.6.1	Results . . . . .	28
3.7	Discussion . . . . .	29
<b>4</b>	<b>Breeding Onset</b>	<b>31</b>
4.1	Defining Breeding Onset . . . . .	32
4.2	The Nest Record Scheme Data . . . . .	32
4.3	Data Processing . . . . .	32
4.4	Model Specifications . . . . .	34
4.4.1	Barn Swallow . . . . .	34
4.4.2	Pied Flycatcher . . . . .	35
4.5	Model Validation . . . . .	35
4.6	Extracting the Phenological Signal . . . . .	36
4.7	Bootstrapping . . . . .	38
4.7.1	Results . . . . .	38
4.7.2	Pied Flycatcher . . . . .	39
4.8	Discussion . . . . .	40
<b>5</b>	<b>Arrival &amp; Breeding Interval</b>	<b>41</b>
5.1	Introduction . . . . .	42
5.2	Obtaining the Gap . . . . .	42
5.3	Results . . . . .	43
5.4	Discussion . . . . .	44
<b>6</b>	<b>Multi-year Analysis for Pied Flycatcher</b>	<b>46</b>
<b>7</b>	<b>Conclusions</b>	<b>50</b>
7.1	Project Overview . . . . .	51

7.2	Acquired Competencies & Outcomes . . . . .	51
7.3	Final Remark . . . . .	52
	<b>Appendices</b>	<b>57</b>

# List of Figures

- 1.1 Barn Swallow Image . . . . . 5
- 1.2 Pied Flycatcher Image . . . . . 6
  
- 2.1 Basis functions using polynomial basis . . . . . 11
- 2.2 Effect of k on smoothness . . . . . 12
- 2.3 Bootstrapping example . . . . . 15
  
- 3.1 Pied Flycatcher filtering scheme . . . . . 18
- 3.2 Arrival across all years . . . . . 19
- 3.3 Pied Flycatcher: observer score distribution . . . . . 24
- 3.4 Cross-validation sampling scheme . . . . . 25
- 3.5 Timing of Arrival vs Northing . . . . . 26
- 3.6 Arrival curves for different latitudes . . . . . 27
- 3.7 Arrival Bootstraps . . . . . 28
  
- 4.1 NRS data distribution . . . . . 33
- 4.2 Breeding curves for different latitudes in Swallow . . . . . 37
- 4.3 NRS 2D . . . . . 37
- 4.4 Breeding Onset Bootstraps . . . . . 38
  
- 5.1 Arrival vs Breeding Onset . . . . . 42
- 5.2 Arrival-Breeding Interval . . . . . 43
  
- 6.1 Arrival & Breeding: multi-year analysis . . . . . 47
- 6.2 Gap: multi-year analysis . . . . . 49
  
- 1 Birdtrack Model Swallow: Covariates 1 . . . . . 58
- 2 Birdtrack Model Swallow: Covariates 2 . . . . . 59
- 3 Birdtrack Model Pied Flycatcher: Covariates . . . . . 60

4	NRS Model Swallow: Covariates . . . . .	61
5	NRS Model Pied Flycatcher: Covariates . . . . .	62
6	Breeding curves for different latitudes in Pied Flycatcher . . . . .	63



# List of Tables

- 3.1 Sample Sizes . . . . . 18
- 3.2 Habitat variables description . . . . . 20
- 3.3 Swallow arrival model: coefficients, deviance & BIC . . . . . 22
- 3.4 Flycatcher arrival model: coefficients, deviance & BIC . . . . . 24
- 3.5 Predictive abilities of the arrival models . . . . . 26
- 3.6 Barn Swallow arrival day bootstrap estimates . . . . . 29
- 3.7 Pied Flycatcher arrival day bootstrap estimates . . . . . 30
  
- 4.1 Sample Sizes: NRS data . . . . . 34
- 4.2 Swallow’s clutch initiation model: coefficients, deviance & BIC . . . . . 35
- 4.3 Flycatcher’s clutch initiation model: coefficients, deviance & BIC . . . . . 36
- 4.4 Predictive abilities of the arrival models . . . . . 36
- 4.5 Barn Swallow breeding onset bootstrap estimates . . . . . 39
- 4.6 Pied Flycatcher breeding onset bootstrap estimates . . . . . 39

# Resumo

O impacto do aquecimento global em processos biológicos por todo o mundo tem sido amplamente documentado: desde alterações na distribuição de diferentes espécies à mudança temporal de eventos ecológicos.

As aves têm sido usadas como bioindicadores há muito tempo por diversas razões, entre as quais por serem fáceis de estudar e por ligarem níveis diferentes das cadeias biológicas, pelo que a sua relação com o clima tem sido amplamente analisada. Em particular, tem-se verificado que aves migradoras de longa distância têm sido particularmente afectadas pelo clima, por não se conseguirem ajustar às suas recentes e drásticas mudanças. Entre estas, incluem-se as migradoras afro-paleárticas, que invernam na África sub-Sahariana e migram para a Europa e áreas adjacentes para se reproduzirem até ao Outono, onde voltarão a migrar de volta.

A fenologia engloba o estudo de ciclos periódicos biológicos e a forma como estes se relacionam com factores externos, que no caso das aves migradoras são os tempos de migração e reprodução, dois aspectos cruciais no seu ciclo de vida. Migrar demasiado cedo pode significar encontrar tempo desfavorável e levar à morte, reproduzir-se demasiado tarde pode levar à falta de recursos aquando da alimentação da ninhada. Todos os tempos são, nesta medida, críticos para o sucesso de um indivíduo.

O aumento das temperaturas e o avanço do começo da Primavera têm levado a um desfasamento da fenologia de vários migradores com a abundância de recursos locais, necessária à reprodução e bom desenvolvimento da prole. Algumas populações não têm conseguido adaptar-se, o que tem levado a declínio.

Para melhor compreender a severidade dos impactos do clima na fenologia das aves, é necessário estudar os mecanismos de adaptação e um conjunto de vários indicadores fenológicos em larga escala. Contudo, a dificuldade em obter bases de dados de grande dimensão e extensão espaço-temporal tem restringido a grande maioria dos estudos a populações locais, cujas conclusões podem não reflectir necessariamente padrões globais.

Em décadas recentes, a expansão da ciência cidadã (*citizen science*) tem permitido solucionar o problema da escassez dos dados e aberto portas para o estudo de padrões em larga escala. Na base deste fenómeno, está o acesso a milhares de cidadãos com interesse em ciência que, no seu dia-a-dia, podem recolher dados com grande valor científico. Em particular, o *birdwatching* é um hobby praticado por muitas de pessoas em todo o mundo, com particular expressão no Reino Unido onde conta com milhões de praticantes, que podem ajudar no estudo da ecologia de várias espécies. Este estudo faz uso de dados provenientes de ciência cidadã para perceber a variação, em larga escala, da fenologia de duas espécies de passeriformes, o papa-moscas-preto (*Ficedula hypoleuca*) e a andorinha-das-chaminés (*Hirundo rustica*), na Grã-Bretanha, usando o ano de 2014 como ponto de partida.

Os objectivos concretos são: estimar a data de chegada da população de cada uma das espécies aos territórios de reprodução, ao longo do gradiente latitudinal; estimar a data de

início de reprodução e a respectiva variação com a latitude; combinar as estimativas obtidas, anteriormente, para estimar o intervalo desde a chegada de migração até ao início da reprodução. Com este processo, pretendemos averiguar qual é a variação dos processos de chegada e início de reprodução, e o modo como se relacionam entre si, nomeadamente, perceber se a data de chegada pode restringir a data de início da reprodução, como verificado em estudos individuais de algumas populações.

O projecto divide-se em três fases distintas, sendo que nas duas primeiras são utilizados modelos lineares generalizados (GAM) para obter estimativas dos parâmetros de interesse. Os GAM são modelos que incorporam funções *smooth* ao preditor linear do modelo, permitindo formas flexíveis de descrever complexas relações não lineares, sem necessidade de transformação de variáveis. Validação formal dos modelos foi feita através de *ten-folded cross validation* e, para obter a variância associada às estimativas, recorreu-se a um método de *bootstrap* não paramétrico.

A primeira fase do projecto consistiu em estimar a data de chegada da população, que aqui é definida como a data em que metade dos indivíduos reprodutores chegou aos territórios de nidificação. Para tal, usou-se uma base de dados *online* de ciência cidadã, *BirdTrack*, em que observadores submetem listas completas de espécies registadas num determinado local e data, fornecendo informação sobre todas as espécies que estão presentes e ausentes. Isto permite modelar a probabilidade de detecção da espécie numa lista em função do dia e latitude, assim como outras covariáveis. A probabilidade de detecção em cada lista é equiparada à proporção de indivíduos que chegam aos territórios de reprodução, num processo cumulativo. Assim, obtendo-se a estimativa da data em que 50% dos indivíduos podem ser detectados em cada latitude, obtemos uma aproximação da data em que metade da população chega aos territórios de reprodução. Um método de *bootstrap*, com reamostragem aleatória simples, foi aplicado de modo a obter estimativas de incerteza associadas à mediana das datas de chegada para cada latitude.

A segunda fase do projecto consistiu em estimar a data de início de reprodução, definida como a data em que metade dos indivíduos coloca o primeiro ovo (início da postura). Para tal, usaram-se dados de outra base de ciência cidadã (Nest Record Scheme), em que voluntários descobrem ninhos e os acompanham ao longo da temporada, permitindo perceber em que data as aves iniciam a postura. Usando um processo semelhante ao anterior, modelou-se a probabilidade de detecção de um evento de início de postura, ao longo do tempo e do gradiente latitudinal. A data em que a probabilidade de detecção de um evento de postura é máxima corresponde à data em que metade dos casais iniciou a postura. O mesmo método de *bootstrap* usado na fase anterior foi aqui aplicado para obter a incerteza associada às estimativas da data em que metade da população iniciou a reprodução.

A última fase do projecto combinou as estimativas das datas de chegada e início de reprodução para obter uma estimativa do intervalo entre estes processos, para cada latitude. O facto das metodologias das duas secções antecedentes serem muito semelhantes permitiu que a simples diferença analítica entre os conjuntos de valores obtidos (via *bootstrap*) obtivesse as estimativas para este intervalo, com a respectiva variabilidade associada.

Os resultados deste projecto variaram em ambas as espécies, na medida em que as estimativas associadas à andorinha-das-chaminés não permitiram que conclusões ecológicas sólidas pudessem ser tiradas, maioritariamente por não se conseguir garantir que os indivíduos amostrados eram indivíduos reprodutores e não apenas migradores de passagem. Já no caso do papamoscas-preto, foi possível separar claramente territórios de reprodução dos territórios de migração, e as estimativas foram boas o suficiente para tirar relações. Assim, verificou-se que, no ano de 2014, houve uma redução da duração do intervalo de reprodução com um aumento na

latitude, sendo que os indivíduos no extremo norte da distribuição demoraram, em média, menos 15 dias a iniciar a reprodução comparativamente aos indivíduos do extremo sul. Isto verificou-se uma vez que o início da reprodução de metade da população desta espécie se deu num intervalo de 3 dias apenas, em média, nos 550km de extensão aqui estudados.

Por forma a complementar este estudo, num projecto paralelo, estendeu-se a análise ao período de 2013-2016 por forma a perceber o quão variável o padrão observado em 2014 poderia ser. Os resultados indicam que a redução do intervalo de reprodução com o aumento na latitude nem sempre se verifica, no entanto, o processo de chegada é sempre muito mais flexível do que o início de reprodução, que varia muito pouco entre anos. Isto é, a data de chegada aparenta não influenciar a data de início de postura, que tende a acontecer num curto espaço de dias.

**Palavras-chave:** fenologia, modelos generalizados aditivos, bootstrap, papa-moscas-preto, andorinhadas-chaminés.

# Abstract

Migratory bird populations all over Europe are deeply changing. In particular, long-distance migrants are being increasingly affected by climate change and have shown difficulties in adapting to the recent and fast changes. Studying the spatial variation in ecological processes is of fundamental importance when unfolding the mechanisms underlying bird population change. The purpose of this study was understanding the variation in arrival and breeding onset with latitude, estimating the length of time between these two processes and describing their relationship. Using data from two citizen science databases, we estimated the latitudinal variation of arrival date and breeding onset for two migratory bird species, Pied Flycatcher (*Ficedula hypoleuca*) and Barn Swallow (*Hirundo rustica*), in Great Britain in the year of 2014. Unlike many of the past site-specific localized studies, here we focused on broad population patterns. Generalized additive models were used to produce estimates of both phenological processes, using a spatio-temporal interaction smooth. Arrival was estimated using data from citizen science database, BirdTrack, by modelling the probability of occurrence of each species on complete birding checklists. This allowed the estimation of the date representing arrival of 50% of the population at different locations. For the breeding onset, we used data from citizen science based Nest Record Scheme to model clutch initiation events, as a measure of breeding onset. From this we extracted the date where the estimated peak of breeding occurred at different locations. A ten-folded cross-validation procedure was used to provide metrics of model validation. Both models were bootstrapped to estimate uncertainty, and bootstrap estimates were used to calculate the "gap" between arrival and breeding. Our results proved inconclusive for Barn Swallow, yet the Pied Flycatcher showed a marked reduction in gap length with an increase in latitude. This is the first time that these ecological processes have been examined and compared across these broad spatial scales.

**Keywords:** phenology, generalized additive models, bootstrap, pied flycatcher, barn swallow

# Chapter 1

## Introduction

## 1.1 Biological Problem

Evidence supporting the impact of climate change on biological events all around the world has piled up: from changes in the distribution of different species to the alteration of the timings of important life events (Root *et al.*, 2003; Walther *et al.*, 2002). There has been a growing need for the development of bioindicators to assess these impacts and understand their magnitude, raise awareness of their consequences and act to minimize them (European Environment Agency, 2007). Birds have been used as important environmental indicators for a long time for different reasons, namely their well studied ecology, their link to different levels of the ecological chain and their conspicuity, allowing them to be detected and recorded efficiently (Bunce *et al.*, 1981; Padoa-Schioppa *et al.*, 2006). Thus, birds' relationship with the climate has been extensively studied in an effort to understand how they are being impacted, what sources of pressure they face and how well they can adapt.

Afro-Palaearctic migrants are birds that overwinter in Africa (mostly below the Sahara desert) which, in spring, migrate to Europe and adjacent areas to breed until they migrate back to Africa in Autumn (Hahn *et al.*, 2009). This process has been repeating itself for thousands of years and birds have developed flexible strategies to be able to keep up with all the external factors that influence migration (Pulido, 2007). However, the rapid changes in climate in recent decades, have been affecting many of these migrants which may not be able to keep up with the fast pace of change (Visser, 2008). In particular, longer distance migrants are subject to a greater impact because they depend on a greater number of factors and thus are less flexible to adapt to changes (Hewson and Noble, 2009).

Phenology comprehends the study of periodic plant and animal life cycle events and how these are driven by external factors, including habitat, seasonal or climatic events. The phenology of migration is a crucial aspect of a migrant's life cycle as it affects both its productivity and survival. Migrating too early might mean encountering unfavourable weather en route (Brown and Brown, 2000; Newton, 2013), whereas arriving too late to the breeding grounds may limit nesting opportunities or narrow the window to find a mate (e.g. Cooper *et al.* 2010, Canal *et al.* 2012). Therefore, the timing of migration is a key determinant of an individual's annual survival and breeding success (Cotton, 2003).

Timing is critical when successfully raising the young. For this reason, many species tend to match their breeding timing with peaks in food abundance to optimize the brood's growth. This means that a bird needs to arrive at the right time to set a territory, build its nest and lay its eggs, so that when they hatch, food is plentiful and juveniles can thrive (Verhulst and Nilsson, 2008; Both *et al.*, 2005). Increased temperatures and the advancement of spring events have led to a disconnection between birds' phenology and the abundance of lower trophic-level resources. Some populations have not been able to advance their migration, so they are now arriving 'late' relative to local resources. This causes a mismatch between their timing of breeding and the timing of optimal resources for raising young. (Both *et al.*, 2005; Saino *et al.*, 2010). Nonetheless, contradictory trends have been verified between and within species. While some populations have shown to be tracking climatic changes and have successfully advanced both arrival and breeding dates (Valtonen *et al.*, 2017), matching the resources, others have not adapted their phenology (Both *et al.*, 2005; Miller-Rushing *et al.*, 2010; Saino *et al.*, 2010).

To understand the severity of the impacts of climate on bird phenology, past studies have looked at a multitude of phenology-related aspects. These include the study of breeding and arrival timings through time, the dynamics of migration speed, the impact of weather on individual fitness, among others. However, most of these studies have been conducted at relatively small scales, for example: small populations followed through long periods of time; a few tagged individuals from a specific population; or bird counts at single migration observatories.

Particularly in Europe, few studies have looked at large population "chunks", with all their associated variability. Individual variation does not always reflect general patterns, and conclusions drawn from small sets of birds might differ from conclusions based on full communities (Orme *et al.*, 2006; Bohning-Gaese and Bauer, 1996). This happens because different sub-populations are subject to different factors, and while some may have set of conditions favourable to a certain outcome, the majority might actually be responding differently. This might explain some of the contradicting trends found by different studies, and exposes the need to look at relationships at broader scales. Large studies in ecology, however, often require many hours of specialized field work, consuming many resources which are usually unattainable. A major and important solution to this problem will be discussed in the following section.

## 1.2 Citizen Science

As the names suggests, citizen science consists of research conducted using data collected by volunteers who are not necessarily linked to scientific institutions in a direct way (Cohn, 2008).

Despite existing for centuries with amateur natural historians, only recently citizen science has gained serious expression and an increasing number of papers have been published using data collected by everyday people, in several different fields, from ornithology to ecotoxicology (Silvertown, 2009). Easily usable on-line tools have made it possible that anyone, with the required specific knowledge, can contribute fast, at any time, anywhere (Cohn, 2008).

When it comes to understanding bird migration, it is essential that coverage is extensive both in space and time, so citizen science starts taking a central role in many studies all over the world. Platforms, such as BirdTrack and eBird, have now thousands of avid birdwatchers contributing with daily sightings to help fill in the knowledge gaps (Sullivan *et al.*, 2009; BTO *et al.*, 2017). These voluntary citizen-based schemes not only allow the collection of massive amounts of data, as they also motivate the public to contribute to something that serves as a basis for future population monitoring and conservation (Sullivan *et al.*, 2009).

Using voluntary citizen-based data does require a great deal of understanding of the data's limitations and sources of potential biases, so that they can be acknowledged and, if possible, addressed (Snäll *et al.*, 2011). This is a common aspect involving the use of citizen science data across all fields. Due to the fact that it is based on voluntary contributions, it can be plagued by sources of biases and confounding. Therefore, a key aspect when using citizen science data is not only knowing what it can be used for, but most especially what it cannot be used for. Certain inferences are valid, others are not. It is a currently flourishing research field the development of more or less sophisticated statistical approaches to account for the potential biases in citizen science data, allowing to address questions that might otherwise not be possible to answer based on such data (e.g. (Gimenez *et al.*, 2014)) .

Volunteer-based schemes in ornithology are based, mostly, on an on-line checklist system. Users must have their own account to where they upload their periodic bird observations resulting from their field trips. Each visit is restricted to a certain location and, in addition to the recorded species, the observer may be required to provide a number of details which may include survey duration, amount of area covered, type of observation protocol (stationary, travelling, occasional), breeding behaviour, among others.

On-line checklist databases serve two main functions: gathering quality data from a scientific point of view and allowing the user to store and explore the data in interactive and dynamic ways. It is fundamental that this balance is well kept: while focussing too much on the scientific side might diminish the reach of the database and exclude potential users



from contributing, attracting and enticing users to contribute too much (regardless of their experience and motivations) may mean risking the data's quality and hamper its potential. Thus, it is critical that good practices are promoted among users and they understand the protocols, purpose of their contributions and what they can do to improve their input (Tiago, 2017).

Despite the great amount of work done in bird migration using citizen science databases, many of the more ambitious, large-scale studies, have been America-centred, with fewer studies coming from the "Old World". This may happen because in America only one database is mainly used (eBird), and manages to gather data from Canada to Peru (Sullivan *et al.*, 2014), while in Europe there is a wide variety of recording schemes (such as Ornitho, Observado, BirdTrack, iGoTerra, GBIF, eBird and others). This limits the access to the data which, in turn, is not standardized nor collected in the same way. This is one of the reasons that our study is focused in the UK, rather than the full latitudinal breeding range of the considered species. In contrast to most other countries in Europe, however, the UK possesses a vast culture in birdwatching, which extends over centuries, resulting in millions of voluntary bird recorders, many of which contribute to recording schemes such as BirdTrack (recording bird's distributions) and Nest Record Scheme (recording breeding parameters) (CBI Ministry of Foreign Affairs, 2015).

## 1.3 Study Species

Our selected species are two Afro-Palearctic migrants that winter mostly below the Sahara desert and breed throughout Europe. They are both easy to identify within the UK context, reducing possible misidentification mistakes, and have enough data to support robust estimates.

### 1.3.1 Barn Swallow

The Barn Swallow (*Hirundo rustica*), commonly referred to as Swallow, is a widespread passerine all around the world, with different populations and subspecies in five continents. In Europe, it is a common migrant breeder that can be found almost anywhere during the hot season. In the UK, it has the most extensive distribution of any summer migrant, being only absent in the most densely populated urban areas (Balmer *et al.*, 2013).



Figure 1.1: Barn Swallow (*Hirundo rustica*). *author: Pedro Nicolau*

Barn swallow often use man-made structures build their nests on, particularly in areas where insect abundance is high and a source of water is available, being mostly a generalist and adaptable species. Its proximity to humans makes it a well liked and recognizable species, often symbolizing the onset of spring (Turner, 2006).

Despite being a common bird, population declines in Barn Swallow have been reported in a number of countries for a few decades, with BirdLife International (2016b) reporting a negative trend on its world population. Additionally, some populations of this species have shown to be directly affected by weather patterns, and to be changing their breeding phenology, making it a species vulnerable to climate change (García-Pérez *et al.*, 2014; Burman, 2016; BirdLife International, 2016b).

Barn Swallow's widespread distribution and conspicuous behaviour make it a good species for migration studies, with plenty of sightings available on most citizen science databases.

### 1.3.2 Pied Flycatcher

The Pied Flycatcher (*Ficedula hypoleuca*) is a small passerine that breeds throughout most of Europe and parts of Asia, migrating to tropical Africa to winter. It has a large population throughout its range, but in the UK it is a localized breeder confined to western and northern Great Britain. Unlike Swallow, it is highly localized in terms of breeding habitat, favouring mature upland woodlands (preferentially deciduous forests) (Ouwehand and Both, 2016; BirdLife International, 2016a). This species has been a scientist-favourite model for many decades due to its local abundance, tameness and loyalty to nesting sites, favouring nest boxes to breed in (Lundberg and Alatalo, 1992).



Figure 1.2: Pied Flycatcher (*Ficedula hypoleuca*). author: Pedro Nicolau

Its population has steadily declined in the UK in the past decades, following a several decades long trend that has been matched in Europe (Robinson *et al.*, 2016; BirdLife International, 2016a).

The impact of climate change on the Pied Flycatcher has been a hot topic in research in recent decades, being one of the first species where phenological mismatch was shown to be happening (Both and Visser, 2001; Ahola *et al.*, 2004; Both *et al.*, 2005). This meant that in a few populations, at least, birds were being constrained by their arrival date to not adjust their breeding onset and keep matching the resource peak. Other populations, however, have revealed opposite trends and have been adjusting to the climate, successfully advancing both arrival and breeding (Burger *et al.*, 2012; Valtonen *et al.*, 2017). In the UK, it has been postulated that arrival date wouldn't be affecting the species as much due to the large arrival-breeding gap shown by a population in southwestern England (Goodenough *et al.*, 2010). The fact that most studies have been site-specific, with localized sub-populations and different methodologies, might explain the lack of a clear large-scale picture when it comes to understanding Pied Flycatcher's relationship with the climate.

## 1.4 Approach

This project will analyse how the phenology of two migratory bird species varied throughout Great Britain, in the year of 2014. More specifically, we will use citizen-science databases to estimate arrival date and breeding onset, here considered the date of first laid egg of the season in a nest, subsequently obtaining the temporal gap between the two processes and evaluating whether that gap changes with latitude.

Three key steps are necessary for our analysis:

1. Understand when populations arrive at the breeding grounds at different latitudes, using a checklist database from which we will estimate the date of arrival of 50% of the individuals.
2. Use a nest-related dataset to find out when birds start breeding, by estimating the peak in clutch initiation events (laying of the first egg).
3. Combining outputs from the previous steps, evaluate the gap between breeding and arrival, exploring potential patterns, in particular with latitude.

We will fit generalized additive models to both datasets, in stage one and two, and uncertainty estimates for derived parameters will be obtained via bootstrapping. Through this process, we will be looking at the following questions:

- How did arrival vary with latitude in 2014?
- Did breeding onset follow a similar pattern in the same year?
- Was breeding onset being constrained by arrival?
- Is there a fixed gap between arrival and breeding, or does it change with latitude?

### 1.4.1 Content of the thesis

This thesis, being on Biostatistics, uses statistical methods to answer practical ecological questions. The structure of the thesis is as follows. We begin with the above introduction reflecting the importance of this project and reasons for its relevance in today's scientific overview. Next, a theoretical chapter describes the statistical methods used. The juicy bits come with the next three analysis chapters, each of them containing one of the three distinct phases of this project as noted above, including their respective methods, results and individual implications. At the end of the project, it made sense to us to extend the analysis beyond what was initially defined and proposed in the thesis. This analysis was conducted in a side project, aiming at a scientific publication, and its results and main conclusions are summarized in chapter 6. We end up with a conclusions chapter, where I take a step back and evaluate what I have learned in this process, in addition to laying the foundations for possible future work.

## Chapter 2

# Statistical Background - Generalized Additive Models

## 2.1 Generalized Additive Model

The key statistical methodology considered in this work is the Generalized Additive Model (GAM). GAMs are an extension to basic linear models, by allowing one to fit non-linear responses as a function of potential explanatory covariates.

To define a Generalized Additive Model (GAM) we begin by introducing the simpler linear models which GAMs build upon. The definition of GAM here presented follows Wood (2017). The methods were implemented through package *mgcv* in software R (R Development Core Team, 2017).

### 2.1.1 Definition

A Linear Model describes a response variable ( $Y_i$ ) with mean value  $\mu_i = E(Y_i)$  as the sum of two terms: a function of linear combinations of  $M$  independent variables or predictors  $X_m$ , ( $m=1,2,\dots,M$ ),  $\mu_i = E(Y_i) = \beta_0 + \beta_1 x_{i1} + \dots + \beta_M x_{iM}$  (called the linear predictor) and a random error  $\epsilon_i$ . The equation of the model is given by

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_M X_{iM} + \epsilon_i, i = 1, \dots, n \quad (2.1)$$

where  $\beta_0, \beta_1, \dots, \beta_M$  are the unknown parameters, with  $\beta_0$  representing an intercept and  $\beta_m$  representing the parameter associated with the  $m^{th}$  predictor ( $m = 1, 2, \dots, M$ ). A fundamental assumption about the linear model is that the random errors  $\epsilon_i$  are independent and identically distributed, following a Gaussian distribution with mean 0 and standard deviation  $\sigma$ .

Generalized Linear Models (GLM) extend this basic formulation by allowing a degree of non-linearity between the response and the linear predictor, provided that the response variable follows a distribution from the Exponential Family. This provides us with a rigorous framework to model counts, proportions, or strictly positive responses, using distributions like the Poisson, the Binomial or the Gamma, respectively. This was strictly not possible with a simple linear model given the Gaussian distributional assumption for the random errors. In the old days LMs would be used (and often abused!) by forcing residuals to look like Gaussian via transformations of the data.

In a GLM we relate the expected value of the response variable to the linear predictor of explanatory variables via a link function. This enforces linearity on the scale of the link function. A GLM is therefore given by

$$g(E(Y_i)) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_M X_{iM} \quad (2.2)$$

Assessing the residuals' patterns often gives a good idea if the relationships are indeed linear on the scale of the link function, but in certain situations this does not hold true - relationships are simply not linear. Imagine for example the abundance of a species as a function of a given variable, say latitude. There will be an optimum, but then there will be a decrease on abundance with deviations from that optimum. Using GLMs to model such data would not be reasonable. Again, the solution could come at the expense of data transformations, but these may obscure existing patterns of interest, and may not always solve the problem or be the most appropriate solution. Adding polynomial functions of covariates may be a solution in some cases, but these constrain the form of relationships. An elegant solution comes in the form of Generalized Additive Models (GAMs), which fit a smoothing function through the data,

allowing for a flexible way of describing complex non-linear relationships, as frequently happens with biological data.

A GAM can be defined as a generalized linear model in which the linear predictor contains a sum of (smooth) functions of covariates. The response variable,  $Y$ , analogously to GLM, must belong to the Exponential Family. As before, a link function is specified,  $g()$ , relating the expected value of  $Y$  to smoothing functions of the covariates,  $x_k$ . The general model has the following structure:

$$g(E(Y_i)) = \beta_0 + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}) + f_4(x_{4i}) + \dots \quad (2.3)$$

Rather than specifying detailed parametric relationships with each variable, we specify the model in terms of flexible and dynamic smooth functions. This raises the need to define these smooth functions and to know how smooth/flexible they ought to be. One of the many advantages of working with smooths is the possibility to combine a number of covariates in a single smooth function. This would extend the above formulation to look like

$$g(E(Y_i)) = \beta_0 + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + \dots \quad (2.4)$$

The flexibility provided by the incorporation of such bivariate smooths in a GAM model will prove essential for this project.

### 2.1.2 Smooth Functions

Smooth terms are obtained as a linear sum of simple basis functions. The general equation for a smooth term is:

$$f(x) = \sum_{j=1}^k b_j(x)\beta_j \quad (2.5)$$

representing the sum of  $k$  basis functions,  $b_j$ , multiplied by their respective coefficients,  $\beta_j$ .

The amount of flexibility, or smoothness, of the function is a function of the number of basis functions underlying its construction, also referred to as the basis dimension,  $k$ . The higher the number of basis functions the more flexible the function will be (see fig. 2.2). In particular a number of basis functions larger than the sample size would allow to interpolate the data.

Finding the optimal  $k$  is not straightforward, and one could feel inclined to do it via hypothesis testing or backwards selection of the models. However, a more efficient way is through Penalized Least Squares, minimizing the following expression (for the univariate case):

$$\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|^2 + \lambda \boldsymbol{\beta}^T \mathbf{S} \boldsymbol{\beta} \quad (2.6)$$

where the bolded terms correspond to the matricial form, with  $\mathbf{T}$  symbolizing the inverse matrix.  $\lambda$  is a smoothing parameter and  $\mathbf{S}$  is a complex matrix involving a number of terms necessary for penalization.

The first term corresponds to the residual sum of squares of the regression, while the added term is a penalty term that measures the smoothness through a smoothing parameter,  $\lambda$ . The higher the value of  $\lambda$ , the higher the degree of smoothing. This is calculated using an iterative method, with the most commonly used being Generalized Cross-Validation (GCV),

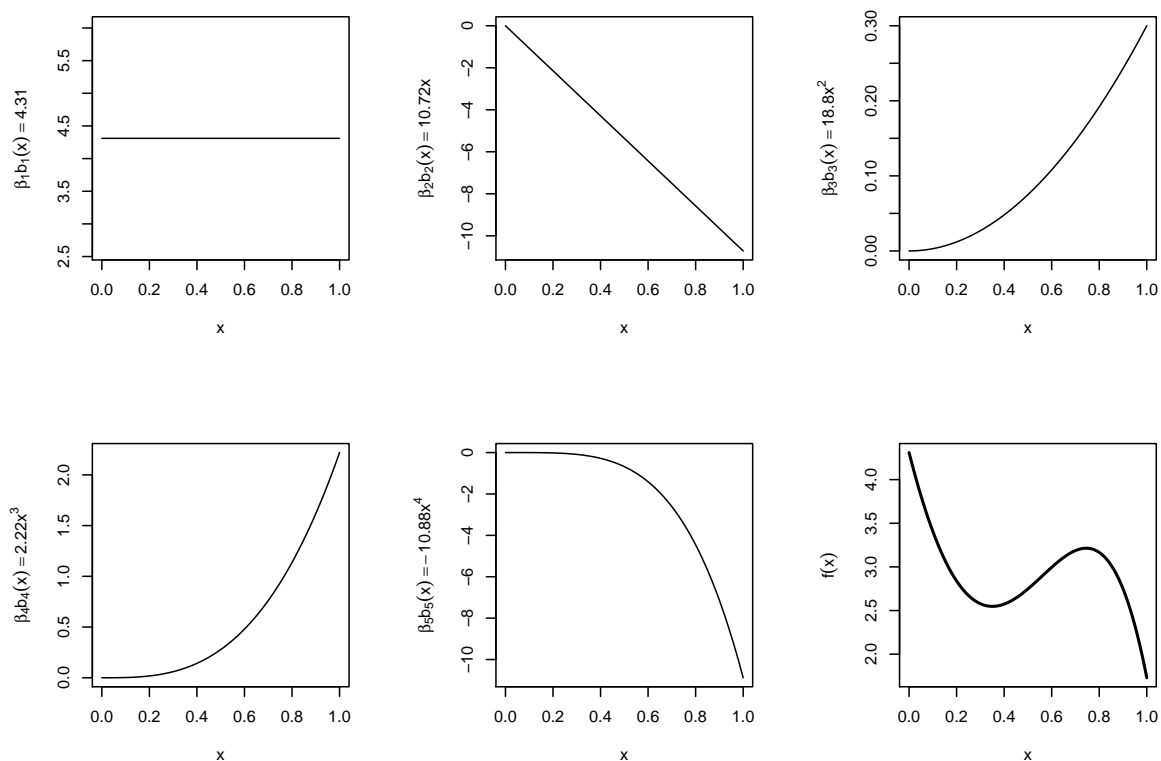


Figure 2.1: Basis functions using polynomial basis. The first five functions correspond to different basis functions which are summed to obtain the bolded function at the bottom right panel. source: Wood (2017)

whose score is automatically given in the output of the model in R. Adding more smooth terms to the model will result in added penalty terms, with each term having its associated  $\lambda$ .

### 2.1.3 Types of Smoothers

There are many different types of basis-penalty smoothers, also called splines, which may be applied to single variables or interaction terms of up to three variables. Each use different methods to estimate  $\lambda$  and have different penalty terms. Examples of smoothers are:

- Cubic regression splines
- P-splines
- Thin plate regression splines
- Tensor product bases (multivariate only)

Because of their unique properties, tensor product smooth terms are particularly relevant to this project, so a brief description will be provided.



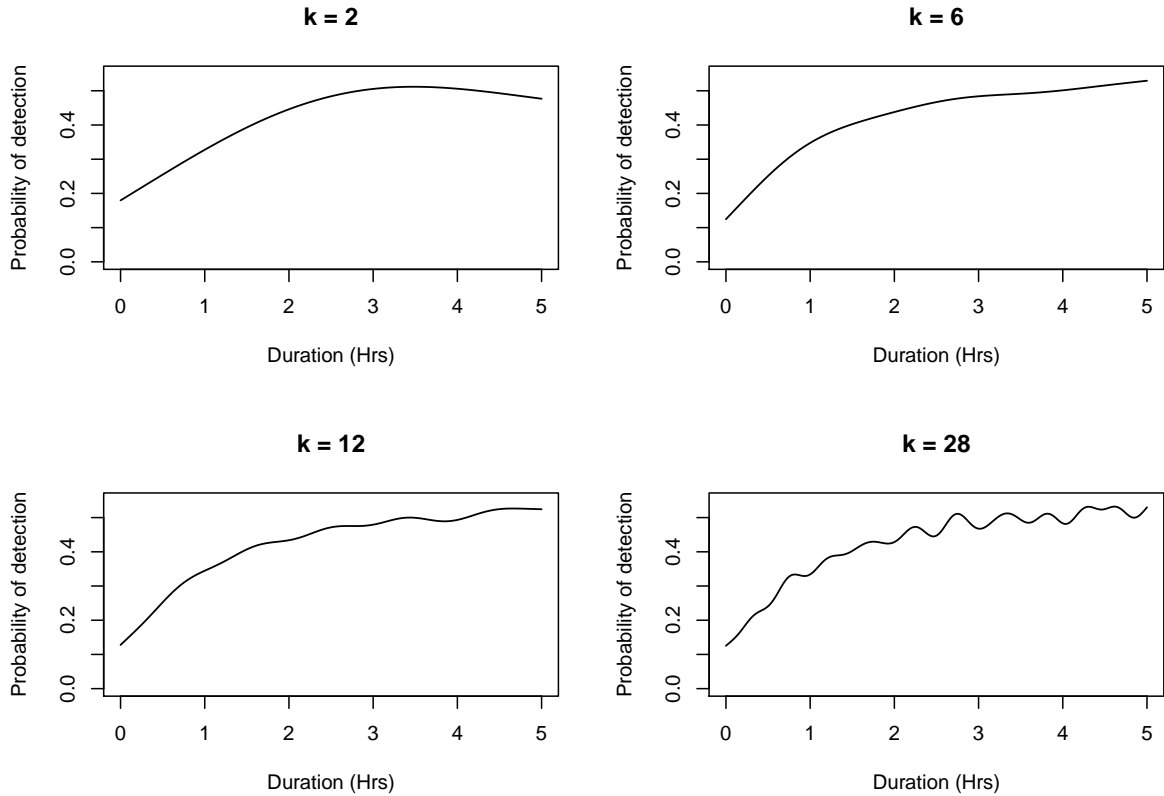


Figure 2.2: Effect of basis dimension ( $k$ ) on the "wiggleness" of the function. These plots show the effect of visit duration on the probability of detection of Barn Swallow. With ( $k = 2$ ), an approximation of a quadratic function is obtained, which fails to represent the most biologically reasonable relationship, but the optimal  $k$  selected automatically using GCV ( $k = 28$ ) shows a clear over-fitting. Somewhere between 6 and 12 would be ideal.

## Tensor product bases

The main idea behind these smooth terms is starting with smooths for single covariates, so-called marginal smooths, from which a "tensor product" construction is applied to build smooths of multiple variables. From the equations

$$f_x(x) = \sum_{i=1}^I a_i(x)\alpha_i \quad \text{and} \quad f_z(z) = \sum_{l=1}^L d_l(z)\delta_l \quad (2.7)$$

we obtain

$$f_{xz}(x, z) = \sum_{i=1}^I \sum_{l=1}^L a_i(x)d_l(z)\delta_{il} \quad (2.8)$$

where  $a(x)$  and  $d(z)$  are basis functions multiplied by their respective coefficients,  $\alpha$  and  $\delta$ , associated with each marginal function,  $f_x$  and  $f_z$ . Iterative methods are again used to estimate the optimal parameters with an associated penalty to each marginal smooth function. Although thin plate splines may also be constructed for more than one variable, the key difference is them being sensitive to variable scaling, whereas tensor product bases are not.

## 2.2 Package *mgcv*

As an updated, extended and more efficient version of the **GAM** package, developed by Hastie and Tibshirani (1990), the **mgcv** package was introduced by Wood (2006) and is the most popular package nowadays to fit GAMs in R.

Below we provide some of **mgcv**'s functionalities, focusing on those that were relevant to consider in this work, using R-specific terminology.

### 2.2.1 *gam* function

The main function used to fit GAMs, *gam*, is analogous in its use to those for fitting LMs or GLMs, with the main difference being that it can introduce a variety of smooth terms. `s()` is indicated for univariate smooths, isotropic smooths of multiple variables and random effects, while `te()` is a tensor product smooth term, indicated for joint modelling in a single smooth term of variables measured in different scales.

A number of different options allow to control the smoothing specifications. The most important one is perhaps the dimension of the basis functions dimension selected by the user (specified by the argument `k` in a given smooth). A smoothness selection criterion, may be used to estimate the optimal  $k$  given the maximum provided, usually via a generalized cross validation procedure. The optimum selected is often referred to as the effective degrees of freedom (edf) of the smooth, and reflects the degree of non-linearity of the relationship (an edf of 1 corresponds to a linear function) between the response and the predictor(s). The type of basis function used can also be selected by the user, via the `bs` argument. Options include cubic regression splines, the default, or thin regression spline with shrinkage (`bs="ts"`).

### 2.2.2 *bam* function

For very large datasets the **gam** function may not be computationally efficient especially when dealing with large basis dimensions and a big number of variables leading. To solve this problem, the **bam** function was introduced. It provides a good approximation of the results obtained when fitting a regular GAM with the **gam** function, an efficient algorithm. It makes use of iterative update schemes to obtain factors of the model matrix while requiring only subsets of the model matrix to be computed (Wood *et al.*, 2014).

### 2.2.3 Shrinkage Selection with *mgcv*

Variable selection is necessary to assess which covariates in the model have the strongest effect on the response variable. From a statistical point of view, it is desirable to achieve a balance between goodness of fit and parsimony, so that model intelligibility and prediction accuracy are enhanced (Marra and Wood, 2011).

Different methods for variable selection are available in literature, namely subset selection, stepwise selection and shrinkage selection (Guisan *et al.*, 2002). Subset selection aims at testing every possible combination of predictors to optimize certain criteria, but gets very computationally inefficient when working with very large datasets. Stepwise selection, on the contrary, leaves far too many options untested and even if it helps improve computational cost, its outputs are dependent of chosen path and there's no guarantee to find the best model, even for the model selection criteria used. Shrinkage methods provide an alternative and appear as

a valid option in terms of stability and prediction (Marra and Wood, 2011). Furthermore, they are particularly convenient when working with a high number of covariates and large datasets as estimation and model selection are carried out in a single step. (Hesterberg *et al.*, 2008; Marra and Wood, 2011) Additionally, they avoid the need to employ testing methods for which no general distributional theory exists (Marra and Wood, 2011).

Shrinkage selection acts upon spline construction by adding an extra quadratic penalty term to the spline (see equation 2.5) that allows the effective degrees of freedom of each covariate to be reduced down to 0, therefore effectively excluding it from the model (Guisan *et al.*, 2002; Wood, 2017). A detailed explanation on shrinkage selection is provided in Marra and Wood (2011). The package `mgcv` offers the option to incorporate shrinkage when specifying the type of smooth selected for each variable.

## 2.3 Bootstrapping

The bootstrap is a resampling method that allows one to obtain variance estimates under situations where these are not available analytically. Further, it can be a simple way of incorporating and propagating uncertainty in multi-stage inferential procedures.

The method that consists of re-sampling the original sample, obtaining so called pseudo-samples, and at each time computing the parameters of interest based on the new pseudo-samples. This is called a non-parametric bootstrap. Then, by repeating the process enough times, the empirical distribution of the computed parameters approximates the real parameter distribution (Efron, 1979). It is based on the idea that the distribution of values contained within a sample  $n$  from a population is the best indicator of that population's distribution. Like so, to approximate the outcome of resampling the population, it is sensible to resample the data (Manly, 2006). An alternative form of bootstrap, which we do not consider in this thesis, is a parametric bootstrap, in which we resample from an assumed distribution for the parameter of interest to get the distribution of derived quantities of those parameters.

Bootstrapping allows the construction of simple percentile confidence intervals for the estimated parameters. The  $100(1 - \alpha)\%$  confidence limits for a parameter are simply the corresponding percentiles that include the central  $100(1 - \alpha)\%$  estimates obtained from re-sampling the original sample (Manly, 2006). An illustrative example is presented in Figure 2.3, where we obtain a 95% confidence interval for the mean of a Poisson distribution.

Bootstrap has become an extremely useful tool in general and in particular in Ecology, where data are often non-parametric and follow unknown distributions. It presents a simple and direct manner of obtaining often complex parameters and confidence intervals for derived statistics, such as median or variance. In addition, it allows to check the stability of the results and robustness of the dataset (DiCiccio *et al.*, 1996; Manly, 2006).

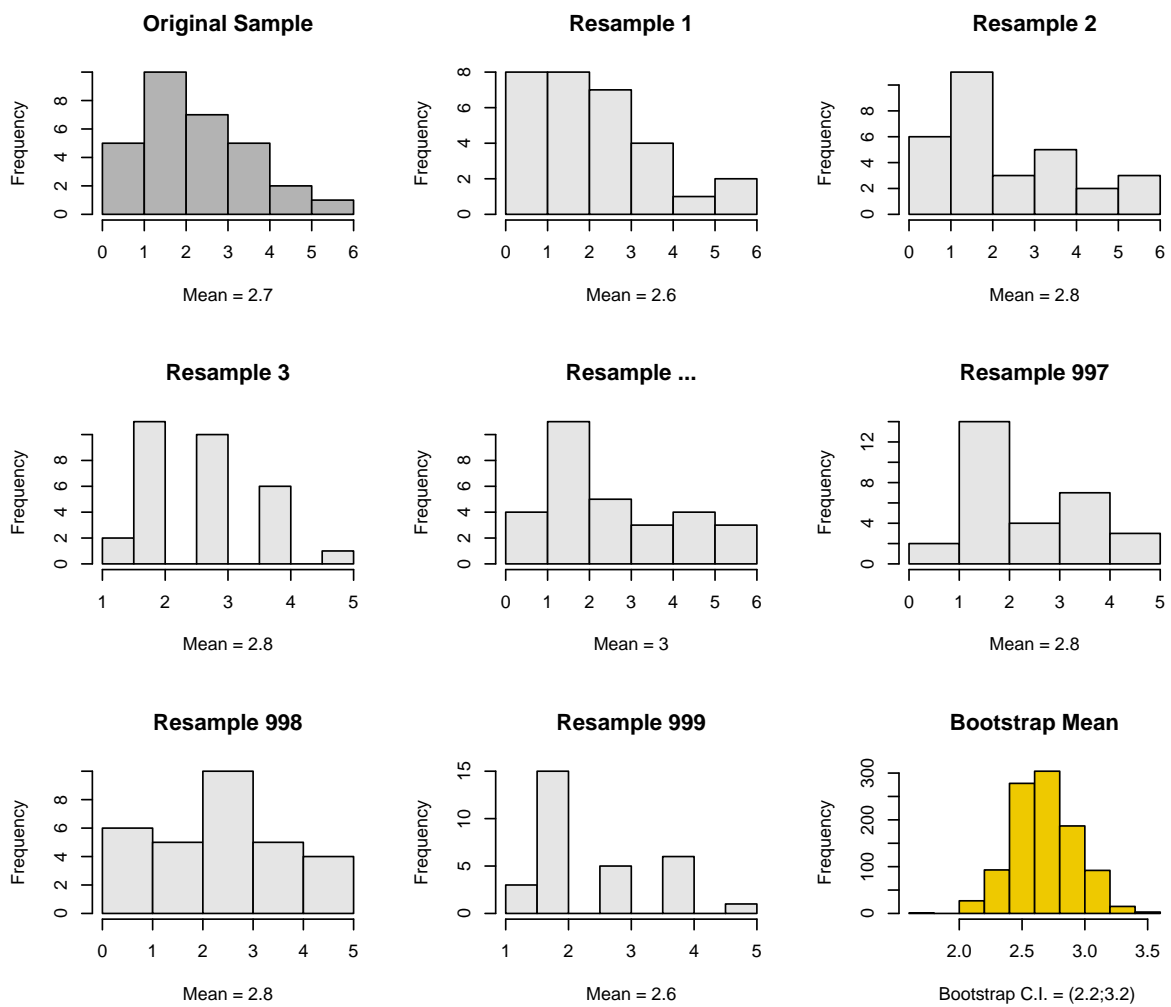


Figure 2.3: Bootstrapping example. A count sample was generated with an underlying Poisson distribution with mean 3. The bootstrap method was used and the generated sample was re-sampled 999 times in order to obtain the mean distribution and its subsequent bootstrap confidence interval (C.I.), shown in the bottom left panel.

## Chapter 3

# Arrival on Breeding Grounds

Through the following chapter we will focus on estimating arrival dates for different latitudes, at the population level, for our study species. First we describe the data and covariates considered for our analysis, proceeding to describe the specifications of the fitted models and metrics used in model validation, with the respective graphical analysis. At the end we present a short discussion.

### 3.1 BirdTrack Data

BirdTrack (<http://www.birdtrack.net>) is a citizen science database, where each participant inserts checklists for bird observations made at a certain location and time. The lists are subsequently reviewed by nationwide reviewers to ensure quality of the records. It has thousands of volunteers that produce hundreds of checklists per day, and overall has millions of bird observations for different species.

Each checklist can be classified as either complete or incomplete by the observer, in addition to other specifications. Those checklists marked as complete are visits where the observer states they have reported every single detected and positively identified species. This feature allows the extraction of presence and absence for every species, more accurately described as detection/non-detection.

BirdTrack comes from a partnership between the British Trust for Ornithology (BTO), the Royal Society for the Protection of Birds (RSPB), BirdWatch Ireland, the Scottish Ornithologists' Club (SOC) and the Welsh Ornithological Society (WOS).

### 3.2 Data Processing

To work with such a large database it is necessary to deal with a huge variability across all levels, ranging from observer skill to spatio-temporal distribution, and potential biases. This happens because each birdwatcher tends to go to the same places, pay more attention to certain species, go out at certain times of the day and favour certain habitats. To reduce some of these effects, a lot of pre-processing and filtering needs to be performed.

It is important to note that the fact that a species is not detected does not mean it is necessarily absent. More competent observers will have a higher likelihood of detecting certain species, but even so the conditions of the observation (such as the weather or time of the day) might reduce the detectability. For the Barn Swallow, we consider false absences to be minimal given this species often calls in flight, tends to choose open habitats where it is easily detected and can easily be recognised by any observer. The Pied Flycatcher is certainly more difficult to find, but as it is a habitat-specific species, loyal to nesting sites (often nesting boxes), this makes it easier to predict its occurrence. Furthermore, in the spring, all males sing to secure their territory which increases detectability upon arrival at the breeding grounds. The uniformity in their breeding habitat makes it reasonable to assume that its detectability remains constant and, thus, the biases are contained.

We used data from Great Britain, selecting the 2014 year. Only complete checklists under 5 hours were chosen, to make sure overly long and less accurate lists were discarded. The complete checklist option enabled us to obtain the detection/non-detection of the species for each reported location (checklist), which we used as our response variable.

To select the optimal time frame for each species' arrival period, we used *birdtrack.net*'s exploratory tools, making sure the window was wide enough to surround the full arrival period,

Table 3.1: Sample Sizes

	Detections	Non-detections	Proportion of Detections	Sampled 10-km Squares
Barn Swallow	8684	19400	30.9%	24192
Pied Flycatcher	290	7087	3.93%	4955

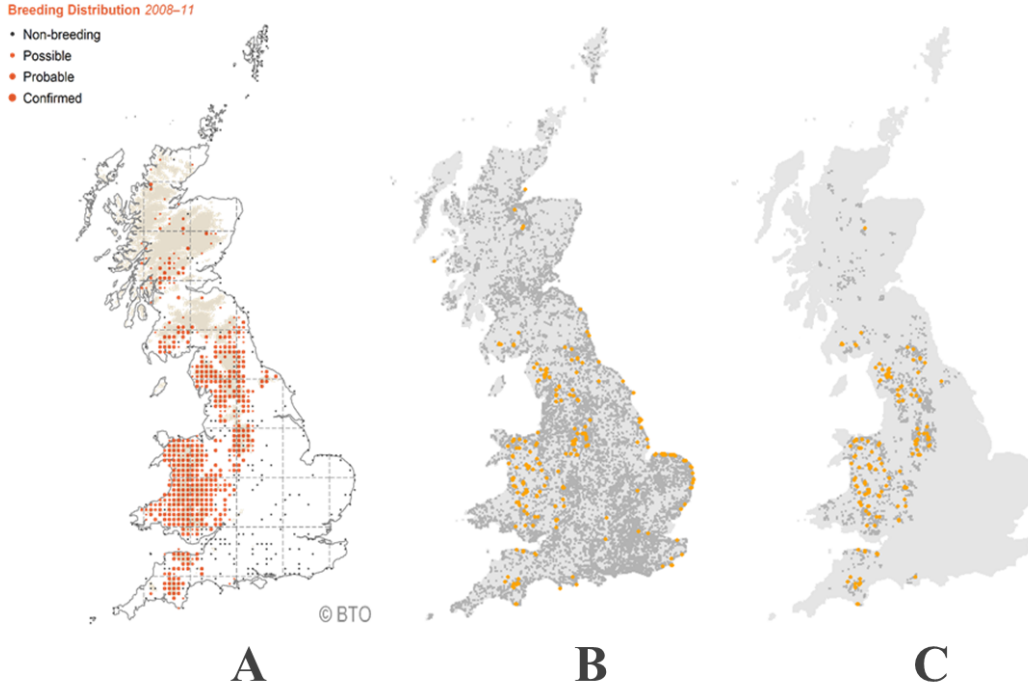


Figure 3.1: Pied Flycatcher filtering scheme: (A) corresponds to Pied Flycatcher’s breeding distribution map, obtained from the 2007-2011 Bird Atlas survey (*source: BTO*); (B) shows locations of BirdTrack checklists - the coloured dots represent checklists with Pied Flycatcher detections while the grey ones represent non-detections; (C) shows locations of BirdTrack checklists, filtered by the Atlas distribution and arrival time period.

but excluding the second migration peak (fig. 3.2). For the swallow the sampling period was 1st-Mar to 15th-Jun and for the Pied Flycatcher it was 1st-Mar to 15-Jul. To minimize the number of observations which did not correspond to arrivals at the breeding grounds but instead to passing birds (still en route), some measures needed to be implemented. The first measure consisted of only accepting records from the known breeding distribution of the study species. For this, we only kept records from the 10-kilometre squares breeding behaviour had been recorded during the Bird Atlas 2007-2011 survey (Balmer *et al.*, 2013), thus excluding several migration-only watch points. This proved to be particularly relevant for the Pied Flycatcher, which showed an 83% drop in number of checklists. The reason for this being that the Pied Flycatcher breeds almost exclusively in woodland habitat and has a restricted distribution in the UK, so all migration observations near the coast are immediately excluded (fig.3.1). The Barn Swallow, on the other hand, showed a simple 1% decrease due to being such a widespread breeder, in all kinds of habitats (this lack of differences made us not to include the swallow plot in Figure 3.1).

To further mitigate this problem, a second measure for Barn Swallow was adopted, in which we excluded all observations within 1-kilometre squares which contained coastal habitat due to swallows being regularly detected on coastal watch points during spring migration.

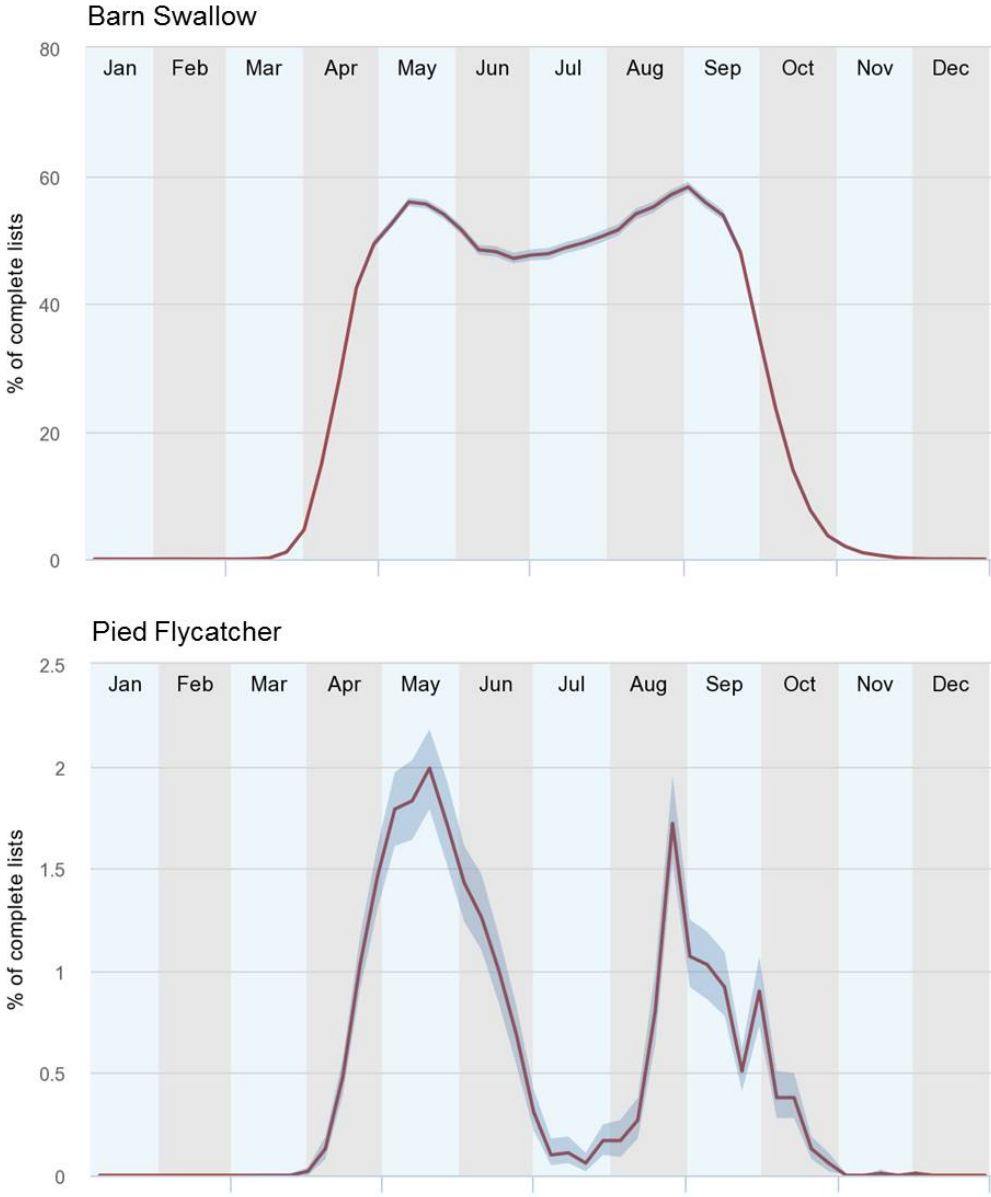


Figure 3.2: Reporting rates across all years contained in Birdtrack. The drop after the arrival is a possible combination of lower bird detectability and less directed effort. source: *birdtrack.net*

### 3.2.1 Variables Description

#### Dependent Variable

To model the presence of each species at the breeding grounds, we used the detection/non-detection of the species on a complete checklist as the response variable. This is a Bernoulli variable, which takes a value of 0 whenever the species was not detected on a visit, and 1 whenever the observer positively identified and reported the species.



## Covariates

Three different groups of covariates were used as explanatory variables: spatio-temporal, effort related and environmental.

1. The first group has two spatio-temporal variables whose interaction is modelled as a joint smooth: *day of year* and *northing* (a measure of latitude in kilometres). These are meant to describe the large-scale arrival pattern of the population. We decided not to include longitude, based on the assumption that both species do not show a relevant longitudinal gradient (they migrate northwards).
2. The second group incorporates observer (as a random effect) and visit duration to account for variations in detectability. We expect visit duration to be more relevant for the Pied Flycatcher as it is a more secretive species that takes more effort to detect, compared to Barn Swallow.
3. The third group includes the altitude and habitat variables (see Table 3.2) and is meant to describe the species' ecological associations. BirdTrack records are associated with 1km x 1km nationally referenced squares. The habitat covariates correspond to the total area covered by the respective habitat (in percentage), within each of these squares, while the altitude variable corresponds to the mean elevation of the land. No significant correlations between any of the environmental variables were detected.

Table 3.2: Habitat variables description

Variable	Habitat description
Broadleaf	Broadleaved, mixed and yew woodland areas
Coniferous	Coniferous woodland cover;
Arable	Arable and horticulture areas;
Grassland	Aggregated variable : rough, neutral, calcareous and acid grassland areas (semi-natural habitat);
Improved	Grazed grassland
Marsh	Fen, marsh and swamp areas;
Heather&Bog	Aggregated variable: bog and dwarf shrub heather;
Mountain	Aggregated variable: montane habitats and inland rock;
Saltwater	Saltwater areas;
Freshwater	Freshwater areas:
Urban	Densely populated, urbanised, built-up areas and gardens
Suburban	Built-up areas and gardens
Coastal	Aggregated variable: supra-littoral rock and sediments, littoral rock and sediments, saltmarsh (not included);

### 3.3 Model Specifications

A binomial generalized additive model was fitted separately for both species. The package `mgcv` was used to fit the models, using the functions `gam` or `bam` (Wood, 2017).

Because we expect arrival to vary differently with latitude, we designed a model structure that would incorporate this flexibility and account for the interaction between timing and location. Therefore, *day of year* and *northing* were introduced as a joint variable via a tensor product smooth term (*te*), which allows arrival curves to vary smoothly with latitude. The tensor product smooths are invariant to linear rescaling of covariates, being recommended when quantities are measured in very different units and when there is not a natural way of rescaling them (Wood, 2017).

Initially, we decided to include observer as a random effect in our model due to each person having similar habits and specific identification skills. On the `mgcv` package, random effects may be specified when defining the type of smooth term to use  $s(\dots, bs="re")$ , making it a more efficient way of running a model and avoiding the need for specific mixed modelling function *gamm*.

The *mgcv* package allows for model selection to be done in one single step, by choosing splines with shrinkage terms. The shrinkage selection procedure removes variables by lowering their effective degrees of freedom below 0.1. Two different types of spline are available: shrinkage cubic splines (`cs`) or thin plate regression splines (`ts`). We tested both splines to see which models would showcase the lowest BIC<sup>1</sup> and decided to choose `ts` splines. The fact that these `ts` splines use an extra penalty term (Marra and Wood, 2011) might explain the more conservative and parsimonious models in our case.

To produce biologically reasonable relationships between the covariates and the response variable, we restricted the maximum degrees of freedom (*k*) to 5 for all smooth terms, allowing the function to select the optimal edf given the maximum provided. After fitting the model, we proceeded to ascertain all the relationships between the covariates and the independent variable, and reduced the *k* in cases where the relationships appeared unreasonable from a biological point of view.

#### 3.3.1 Barn Swallow Model

The fitted model included 16 explanatory variables as smooth terms and a random effect, and an explained deviance of 45%. We set the maximum number of degrees of freedom for the joint smooth at 48, and for the remaining variables we started with a maximum of 5, restricting it to 4 for three variables (*broadleaf*, *improved* and *freshwater*), to 3 for *grassland* and to 2 for *suburban*.

$$\begin{aligned} \text{logit}(E(\text{Detection}_i)) = & \beta_0 + t_1(\text{dayofyear}_i, \text{northing}_i) + f_1(\text{duration}_i) \\ & + f_2(\text{altitude}_i) + \sum_{j=3}^{15} f_j(\text{habitat}_{ij}) + \text{observer}_k \end{aligned} \quad (3.1)$$

The *f* functions correspond to different thin regression splines applied to each covariate and the *t* function corresponds to a tensor product spline. Here *i* indexes the checklist, *j* indexes the

---

<sup>1</sup>Bayesian Information Criterion (BIC) has been shown to perform better, compared to other metrics (namely Akaike's Information Criterion), when using large datasets with great variability (Brewer *et al.*, 2016).

habitat, and  $k$  the observer. *habitat* represents the 12 habitat variables described in Table 3.2, with the exception of *coastal*.

The size of the Swallow dataset and complexity of the model increased the running duration to around 72 hours. This would hamper model validation, as well as our uncertainty estimates via bootstrapping. For this reason, we opted to use the function `bam`, which fits a GAM to large datasets using more efficient computation, obtaining good results (Wood *et al.*, 2014). In this case, using this function reduced the running time to slightly over two hours.

Table 3.3: Swallow arrival model: coefficients, deviance & BIC

Variable	EDF	Chi.sq	p-value
te(dayofyear,northing)	29	4953.747	< 0.001
s(broadleaf)	3.5	165.950	< 0.001
s(altitude)	2.1	111.288	< 0.001
s(urban)	2.9	69.753	< 0.001
s(suburban)	1.1	112.397	< 0.001
s(grassland)	2.2	150.100	< 0.001
s(heather bog)	< 0.1	0.000	0.856
s(arable)	1.4	264.814	< 0.001
s(freshwater)	3.8	294.235	< 0.001
s(mountain)	0.7	13.308	0.032
s(marsh)	0.8	8.448	0.019
s(improved grassland)	3.4	602.189	< 0.001
s(saltwater)	0.3	0.614	0.218
s(observer)	657	2119.533	< 0.001
s(duration)	4.6	3242.657	< 0.001

*Deviance explained: 45%*

*BIC: 26469.2*

*sample size: 28034*

Swallows build nests on a wide range of structures, including houses, bridges and ruins, using mainly mud to form the nest. Therefore, they do not have very specific breeding habitat requirements, breeding in a variety of habitats (BirdLife International, 2016b). Nonetheless, we do expect them to be at the same density in all habitats, due to the number of breeding structures or differences in food availability, so we include all habitat covariates in the model to account for variation in breeding density caused by these factors (excluding *coastal*, as previously explained).

The shrinkage selection procedure maintained most variables in the model, with the exception of *heatherbog* whose effective degrees of freedom were shrunk below 0.1. *Saltwater* shows up as non-significant but is, nonetheless, maintained.

The strongest environmental variables associated with predicting swallow presence on checklists were *improved grassland*, *freshwater* and *arable*, all of which represent key habitats for both foraging and nesting (Turner, 2006). At the same time, *duration* appears to be very relevant and the second most statistically significant covariate in our model. The explained deviance, close to 50%, is very much satisfactory given the big sample ( $n = 28034$ ) and large variability contained in the data (see Table 3.3).

### 3.3.2 Pied Flycatcher

The model was fitted using the `gam` function and included 8 explanatory variables, as smooth terms, with an explained deviance of 39%. The number of maximum degrees of freedom was set at 24 for the spatio-temporal covariate, and kept at 5 for each of the remaining covariates. The absence of the *observer* covariate will be discussed further down.

$$\begin{aligned} \text{logit}(E(\text{Detection}_i)) = & \beta_0 + t_1(\text{dayofyear}_i, \text{northing}_i) + f_1(\text{duration}_i) \\ & + f_2(\text{altitude}_i) + \sum_{j=3}^7 f_j(\text{habitat}_{ij}) \end{aligned} \quad (3.2)$$

The  $f$  functions correspond to different thin regression splines applied to each covariate and the  $t$  function corresponds to a tensor product spline.  $i$  indexes the checklist and  $j$  indexes the habitats broadleaf, coniferous, grassland, urban and suburban.

Pied Flycatchers have a well defined habitat to breed, contrary to the Barn Swallow, and for this reason we opted for a different route, by selecting covariates which have been linked to its breeding ecology:

1. *Broadleaf* and *coniferous* were selected given they correspond to prime breeding habitat of pied flycatchers which breed exclusively on woodland (see section 1.3)
2. *Urban* and *suburban* were selected due this species being very sensitive to human presence, being important in explaining absences.
3. *Grassland*, an aggregated variable which includes open habitats that often surround woodland, was included to encompass two types of “edge effect”: detectability of birds is improved in more accessible, open areas, surrounding woodland; and the edges of the forest are associated with food abundance and higher habitat productivity (Brotons and Herando, 2003).

Albeit being initially included, the random effect covariate, *observer*, was not included in the final model. The Pied Flycatcher, being a relatively inconspicuous species that breeds in dense woodland habitat, often needs directed effort to be seen during spring. This means that observers, which do not search specifically for it in favourable habitat, will most likely not see one. This generates two main groups of observers: the ones that never record pied flycatchers; and the ones actively search for them or regularly go out in favourable habitat, subsequently showing a high report rate (see fig 3.3). Therefore, the observer scores do not follow a normal distribution and we decided to excluded this covariate from the model.

Shrinkage selection removed no variable from the model. The model’s most important covariate was *broadleaf*, which comes as no surprise given the extensive bibliography linking this particular habitat to the species’ nesting in Britain (Balmer *et al.*, 2013). *Duration* of

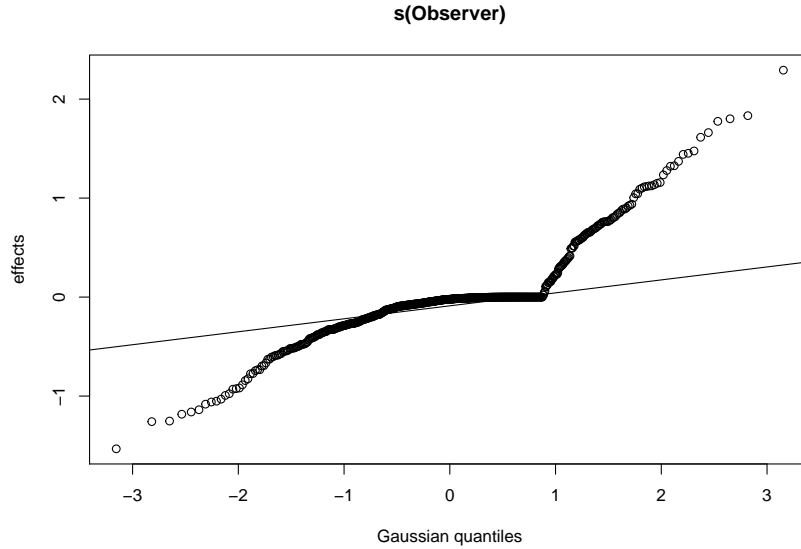


Figure 3.3: Distribution of the observer effects in the Pied Flycatcher model. Variability at both ends enhances the fact that there are many observers that show a bias towards either frequently detect or rarely detect this species. Normality is not verified.

the checklist, as expected, proved to be very important given the fact that this species usually requires some effort to be seen. *Urban* and *suburban* showed a negative linear relationship with this species' absence (see appendices), as we would expect. The lower deviance was inferior to Swallow's model, possibly as a result of the poorer dataset with lower proportion of detections (3.9%), as well as the absence of a variable accounting for observer skill.

Table 3.4: Flycatcher arrival model: coefficients, deviance & BIC

Variable	EDF	Chi.sq	p-value
te(dayofyear,northing)	12.4718	113.951	<0.001
s(broadleaf)	4.5	171.516	<0.001
s(altitude)	2.5	26.496	<0.001
s(urban)	0.9	5.322	0.012
s(suburban)	0.9	9.713	0.001
s(grassland)	3.3	50.145	<0.001
s(coniferous)	2.4	15.171	<0.001
s(duration)	2.2	79.674	<0.001

*Deviance explained: 39.2%*

*BIC: 1804.5*

*sample size: 7373*

### 3.4 Model Validation

To test the predictive abilities of the arrival models, we performed a ten-fold cross-validation on our models. For this procedure, we first divided our datasets into 10 different subsets (groups), each containing 10% of the observers. That is, each of the ten sets contained all the observations from 10% of the birdwatchers. We then fitted models for each of the 90%-sets and used them to predict over each one of the smaller 10% test sets. By doing this, we intended to see if our predictions were robust, and so prediction is made for a set with completely different observers than the ones included in the model. As each observer reported a different number of lists, each group ends up with a different length (see fig. 3.4).

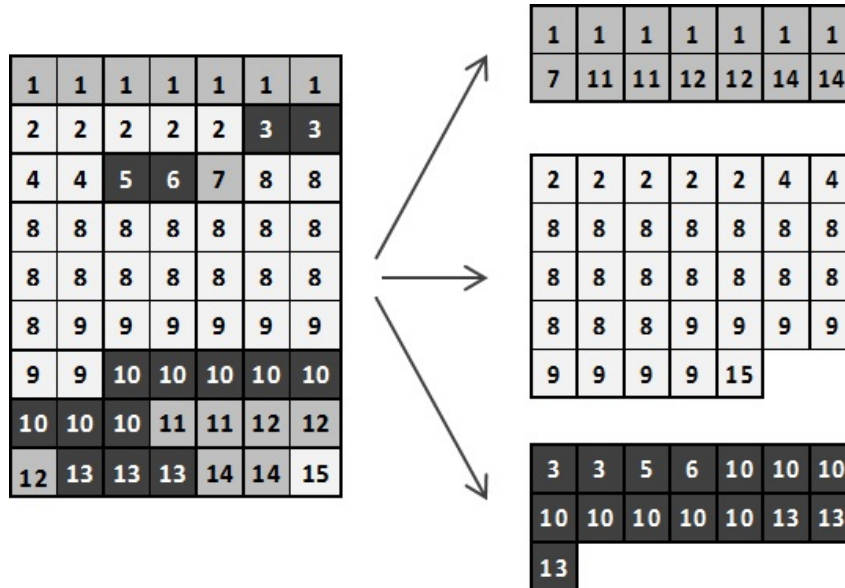


Figure 3.4: Cross-validation sampling scheme. The dataset (left) is randomly split into smaller groups, subsetting by observer. Each number represents one observer, each square represents a checklist and each colour corresponds to a different subset group. By dividing the dataset into subsets composed by the same number of observers, it means that each of the subsets has a different number of observations because each observer has a different number of checklists. In this example, the dataset is split into only 3 groups (opposed to 10) for purposes of image simplicity.

The metric selected for formal validation was AUC (area under the curve) (Sing *et al.*, 2004), obtained via the `PresenceAbsence` package (Freeman and Moisen, 2008). In addition to it, we looked into the sensitivity and specificity using the species observed prevalence as the threshold. Both models performed very well, with AUC equalling to 87.7% and 92.1%, for the swallow and flycatcher models, respectively. The high number of absences may inflate the AUC values, however, both sensitivity and specificity were very satisfactory (see Table 3.5). Considering the enormous variability contained in our data, this adds to the conclusion that our models are fairly robust at predicting detections on checklists.

### 3.5 Extracting the Phenological Signal

In the previous subsections, we modelled each species' probability of detection through space and time. However, we are interested in obtaining arrival dates per latitude, which we will later use to estimate the gap (chapter 5).

Table 3.5: Predictive abilities of the arrival models

Model	AUC	Sensitivity	Specificity
Barn Swallow	87.7%	86.5%	74.7%
Pied Flycatcher	92.1%	81.9%	84.5%

Defining arrival is not easy, because there is not a single day when a full population arrives – this occurs through the course of several weeks in a cumulative process. This results in a lot of individual variation in terms of arrival day at the breeding grounds, so it is important that we define an objective metric representative of arrival for the population. It has been shown that different percentages of population arrival, in spring, give similar results for different migratory species (Miles *et al.*, 2016). We selected the timing when 50% of the population had been detected as our metric to approximate the 50<sup>th</sup> quantile of an underlying cumulative distribution of arrival times.

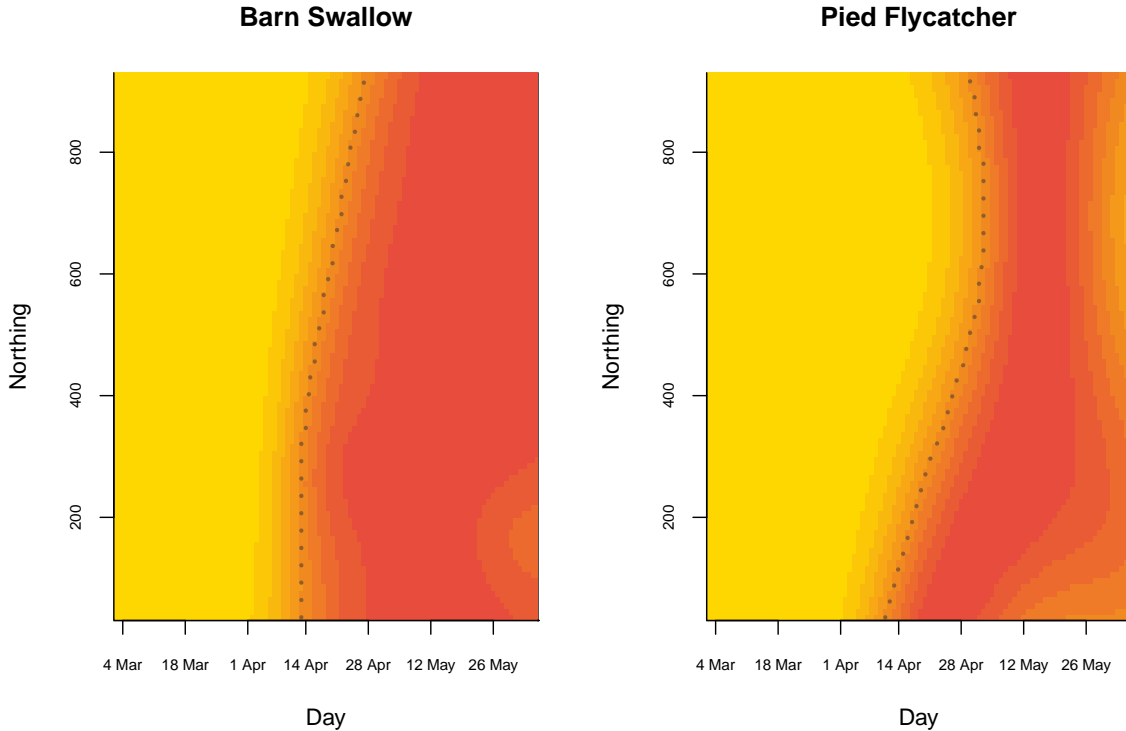


Figure 3.5: Timing of Arrival vs Northing. These plots show the species’ probability of occurrence of each species on a checklist with day and northing. The dotted lined corresponds to the approximation of the 50<sup>th</sup> quantile in arrival given by the date when 50% of the population has been detected.

To estimate arrival date, we obtained the probability of detection on a checklist from the GAM model for each day and across a sequence of northings. All other covariates included in the models were fixed at their median values in locations where the species had been detected, hence effectively predicting the occurrence for an average habitat composition. In the case of the Swallow, observer’s random effect was excluded from the prediction, so that we predicted for an average skilled birdwatcher instead of a specific person from our dataset. This process

resulted in a prediction matrix containing the probability of occurrence of each species on a checklist at each day and northing. To estimate the proportion of the population that had arrived by each date, we standardized the prediction matrix using the maximum probability of detection for each northing. This is represented in Figure 3.5, and is justified by assuming that the maximum corresponds to the date where all the individuals that could potentially arrive have already arrived. Any drops in the estimated probabilities after that date are assumed to be the consequence of drops in detectability, which is expected once breeding starts. Given this point as a reference, we can now estimate the date that corresponds to half of the population having arrived. This allows the extraction of standardized arrival curves for different northings, showcasing the arrival progression throughout the season for different latitudes. These standardized arrival curves at each latitude are essentially slices of the prediction matrix mentioned above, and some examples are presented in Figure 3.6.

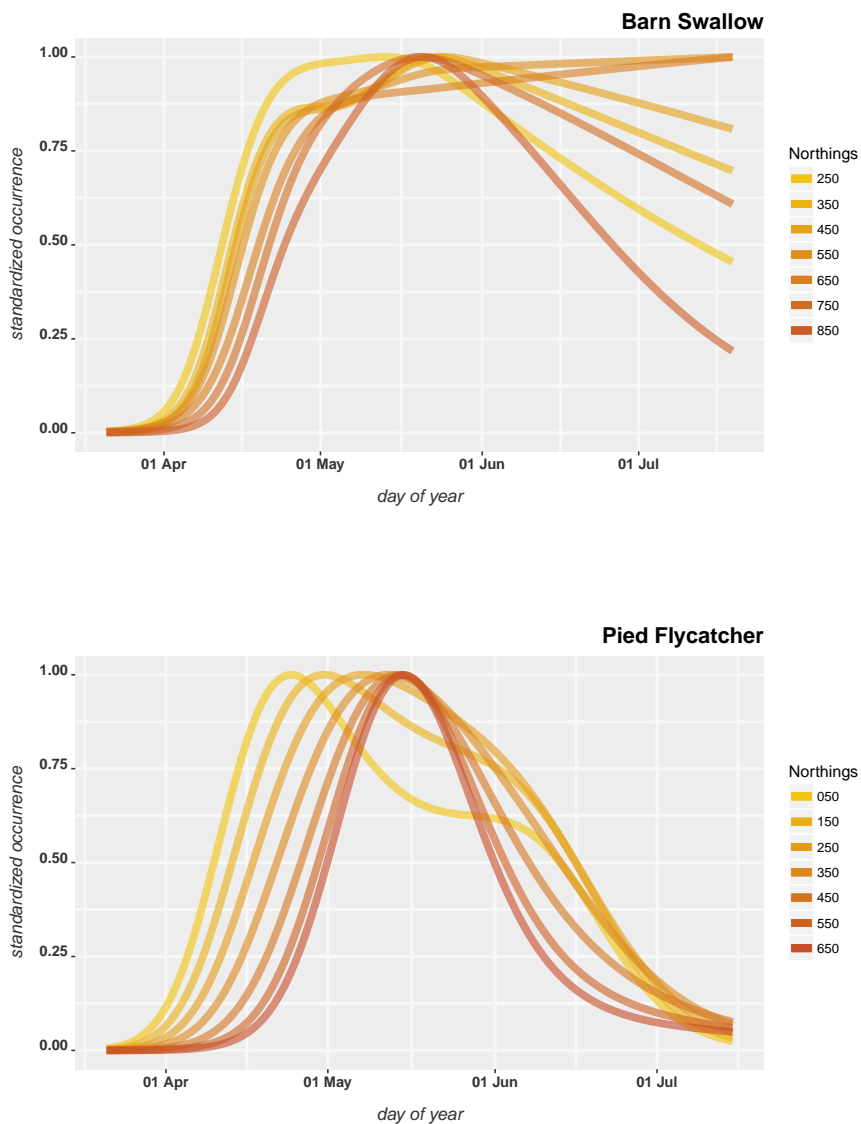


Figure 3.6: Each curve shows the progression of the probability of detection on a checklist through the arrival period for each of the selected northings. The peak in detection is assumed to correspond to the date of when the population has arrived.



## 3.6 Bootstrapping

The scope of this project is describing the arrival pattern and spatio-temporal distribution of our study species, rather than trying to explain where each species occurs. For this reason, and given that we are working with binomial GAMs whose residuals are notoriously difficult to interpret (Wood, 2017), we decided to bootstrap our data, both to test our models' robustness, and provide a direct measure of the uncertainty associated with our estimates. Two-hundred bootstraps were performed for each of the two models, for each bootstrap randomly selecting the same number of checklists with replacement. We fitted our models to each of the bootstrapped datasets and extracted the median date of population arrival at each northing. For each bootstrap, a random seed<sup>2</sup> was set and stored to ensure the random sample and the whole procedure were reproducible.

### 3.6.1 Results

Our bootstrapped estimates for arrival day of 50% of the population can be seen on Figure 3.7. Barn Swallow showed a larger latitudinal range than Pied Flycatcher due to having a more extensive distribution northwards. Roughly, both species presented the same pattern: the further north, the later the arrival day. Detailed results are presented for each species and, for easier reading, a Table is shown, with median values for arrival dates at different latitudes, in Julian days.

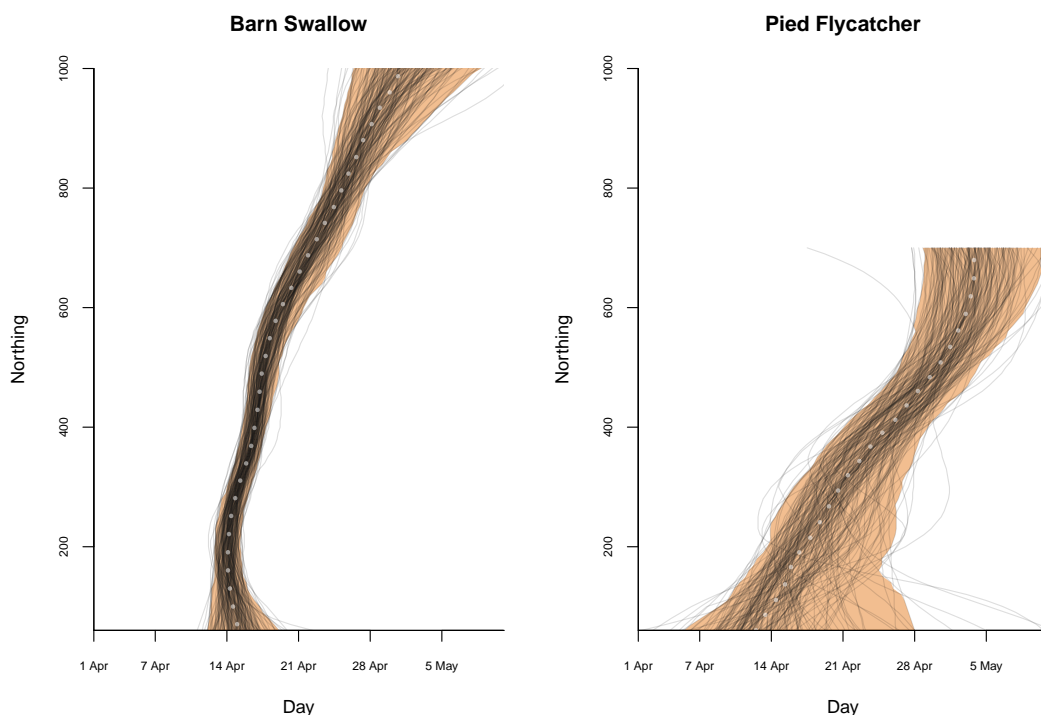


Figure 3.7: Arrival day variation with northing. Each dark line represents the set of estimated dates of when 50% of birds have been detected for each northing, in the 200 bootstraps; the coloured band shows a 95% confidence interval for each latitude; and the dotted line gives the median value across all northings.

<sup>2</sup>”A random seed is an integer vector, containing the random number generator (RNG) state for random number generation in R.” (R Development Core Team, 2017)

## Barn Swallow

The set of bootstrapped values for the Barn Swallow showed a noticeable coherence through its range, with less than a week of uncertainty in most of its core distribution. A bigger variability occurred at the edges of the distribution, most likely as a result of smaller sample sizes and modelling edge effects (Wood, 2017). The earliest median estimate for population arrival date was April 13<sup>th</sup> (Julian day 103), while the latest occurred at the far north of the distribution, 19 days later.

Table 3.6: Barn Swallow arrival day bootstrap estimates

Northing	Median Estimate	Bootstrap Interval
50	104	(101 , 109)
150	103	(102 , 105)
250	103	(102 , 105)
350	105	(104 , 106)
450	106	(105 , 108)
550	107	(106 , 109)
650	110	(108 , 113)
750	113	(111 , 116)
850	116	(113 , 119)
950	119	(115 , 124)
1050	122	(116 , 132)

## Pied Flycatcher

The Pied Flycatcher showed a similar trend to Swallow, with birds arriving later at the northern grounds compared to the south. However, these estimates are less precise, showing larger variability, a likely reflection of patchier data with fewer presences compared to the other species.

The earliest arrival coincided with the earliest in Swallow, April 13<sup>th</sup>, as well as the latest arrival, May 2<sup>nd</sup>. Comparing to Swallow, however, Pied Flycatcher populations showed a larger difference between populations arriving in the south and at northing 600. This can be seen from the shallower slope of the arrival curve.

## 3.7 Discussion

Along the chapter, we used BirdTrack's complete checklists, submitted by citizen scientists, to obtain arrival dates of two different migrant species, at the population level, for different latitudes (measured in northing). This was done by modelling the probability of detection of each species through time and northing, and subsequently obtaining detection curves for each

Table 3.7: Pied Flycatcher arrival day bootstrap estimates

Northing	Median Estimate	Bootstrap Interval
100	103	(98 , 116)
200	106	(103 , 115)
300	110	(105 , 116)
400	114	(110 , 118)
500	119	(116 , 123)
600	122	(117 , 128)

northing, from which we estimated an approximation of the 50<sup>th</sup> arrival quantile. Using a bootstrap methodology, we obtained uncertainty intervals for our estimates.

Both Swallow and Pied Flycatcher’s estimates revealed that birds in the northern part of the UK arrive later. This may be because: a) they travel further, b) they leave later, or c) they migrate slower; or a combination of these factors. We know that spring migration can be really fast with some birds covering up to 370km/day in some parts of the migration route, especially when crossing unfavourable habitat areas (Ouwehand and Both, 2016). This suggests that the pace of arrival in the UK is not restricted by migration speed, but rather is limited by something else. It may be that birds can alter their speed of migration in response to weather, for example, migrating more quickly during warm years when they may be at risk of being late for the peak of resources for breeding.

Swallow showed precise estimates in terms of arrival variability, however, it had a less effective data filtering than the Flycatcher when restricting to breeding birds in arrival grounds, rather than passage ones. This happens because swallows are very generalist species, and breeding birds often co-occur with passage ones in breeding areas. And because they don’t have well-defined habitats, we would need some sort of breeding behaviour-related variable, associated with each observation, if we wanted to be sure we were dealing with individuals at the breeding grounds. As a consequence of this, we can’t be sure if we are dealing with local breeding birds alone, or if migratory ones are involved, which may mean our arrival dates may be blurred by migrants. This effect would have a larger impact in the south where there are more migrants relative to breeders, and would be expected to make the estimate of 50% population arrival earlier in the south. Therefore the estimated arrival curve up the country would be shallower than the true population arrival.

In contrast to Swallow, Pied Flycatchers are much more scarcely seen on migration and this usually happens by the coast. Moreover, they show a clear separation between breeding locations and migratory ones, which means that filtering by breeding areas will be very effective in excluding the migratory individuals.

## Chapter 4

# Breeding Onset

In this chapter, we will proceed to estimate the clutch initiation date, a measure of the breeding onset, for the barn swallow and the pied flycatcher. We will use an analogous methodology to the previous chapter in order to estimate the spatio-temporal variation of this parameter.

## 4.1 Defining Breeding Onset

After arriving at the breeding grounds, birds initiate their breeding season. This involves a lot of preparation work: get in physical condition to sustain the reproductive period, if they are not already fit; display to find a mate to copulate with; find and defend a suitable territory; build a proper nest; lay the eggs, one by one, and incubate; raise the brood after eggs hatch; possibly, repeat the full cycle if something fails or if the species is multi-brooded. Therefore, it appears difficult to select one key moment when breeding has begun. Although knowing when a bird starts displaying or finishes building a nest may be extremely difficult, once the nest is found and tracked, one can find when bird laid its first egg (clutch initiation). This provides a concrete metric of breeding onset which is often used in similar studies (Valtonen *et al.*, 2017).

## 4.2 The Nest Record Scheme Data

Probably the largest and longest running breeding scheme in the world, BTO's Nest Record Scheme (NRS) has been gathering high-quality data on the breeding birds of the UK, since 1939. It relies on a network of over 500 knowledgeable volunteers who locate and monitor the progress of single nests for a variety of species (Crick *et al.*, 2003). Nest recorders register the species, location (10-km square) and date, as well as numbers of eggs and young, bird activity and outcome of the nest. Each person is encouraged to make multiple visits to each nest from the first visit down to the completion of the breeding attempt, so that a number of different variables can be obtained, such as clutch initiation day, brood size or breeding success. After online submission, the data is processed by a curating team and a number of derived variables are subsequently obtained.

## 4.3 Data Processing

Considering very few nests are found in their early stages, due to secretive bird behaviour, clutch initiation day, our selected metric for breeding onset, requires back-calculation based on the species' documented ecology, a process that provides a minimum and a maximum date for when the first egg of the breeding season was laid (Crick *et al.*, 2003). All nests with a day range bigger than 10 ( $\pm 5$  days uncertainty) were excluded from our analysis, and the midpoint in this uncertainty interval was used as our phenology estimate. The sample size for Swallow was 628 nests, with a median uncertainty of  $\pm 2$  days, and 803 nests for the Pied Flycatcher, with a  $\pm 1.5$  days median uncertainty.

A more intuitive way of estimating clutch initiation date would be modelling laying day as a function of a number of spatial and environmental covariates. However, neither species is exclusively single-brooded, meaning that clutch initiation does not follow a unimodal distribution, but rather a multivariate normal with two peaks. This double-brood phenomenon is extremely common in swallows where most individuals breed twice within the same season, whereas most flycatchers will not, but can re-lay in case there is an early failure in the breeding

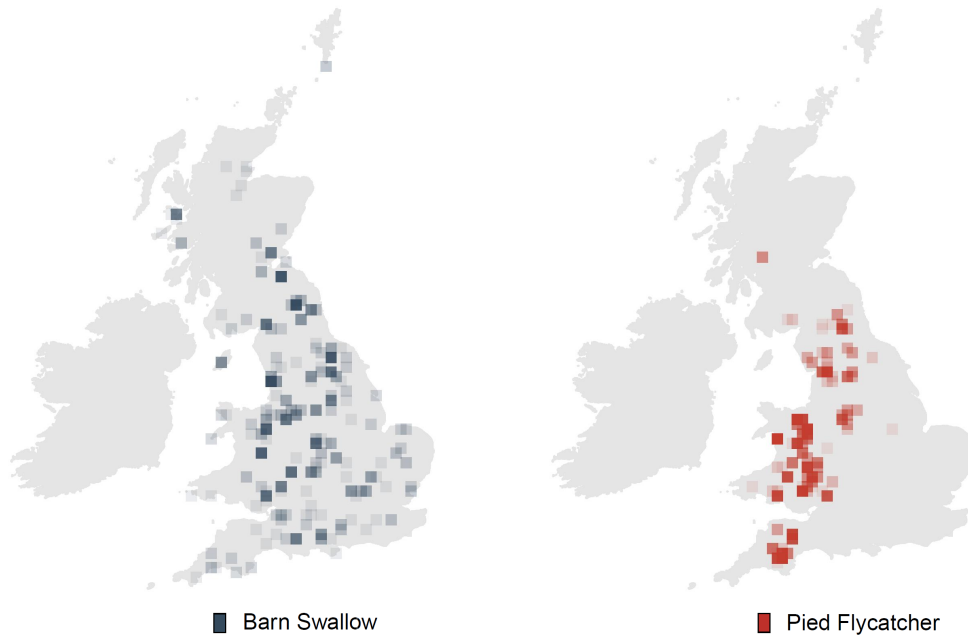


Figure 4.1: NRS Data Distribution: more intense colours represent higher density of recorded nests.

attempt (Lundberg and Alatalo, 1992). This means that the distribution of laying dates is not distributed according to standard forms, such as a Gaussian distribution and can't naturally be modelled in this way. For this reason, and to be able to combine these results with our Birdtrack analysis, we felt it was important that our framework remained the same, and so we analysed the nest data using a similar model structure.

We created an detection/non-detection model from this dataset in particular, with the same environmental variables, smooth terms and selection method. We expected to obtain the spatio-temporal distribution of clutch initiation and, subsequently, extract the bootstrapped median laying dates.

Contrary to the previous dataset, where we had a detection and non-detection of each species for every sampled point, the NRS data we does not provide explicit non-detections: we merely possess location and date of occupied nests. Nonetheless, by knowing when the laying event occurred (clutch initiation), we can infer all the days when it didn't, for that particular location. Because of the patchiness of our data and lack of continuity between points, clutch initiation events were aggregated within 10km x 10km squares and two day periods. In doing this, we assume that nests' detectability remains the same in every sample square. Our response variable was the detection/non-detection of a recorded clutch initiation event within a square during a two day period.

Table 4.1: Sample Sizes: NRS data

	Detections	Non-detections	Sampled 10-km Squares
Barn Swallow	513	9279	153
Pied Flycatcher	416	2219	162

## 4.4 Model Specifications

As mentioned in the previous section, we fitted a binomial GAM for the two species to estimate the presence of clutch initiation events through time and space. We wanted to keep the model structure and specifications as similar as possible, so we used the same type of smooth terms, both for the smoothing of variables and for the selection procedure (see section 2.3).

The smaller sample sizes, comparing to the previous section models, and respective shorter running periods, allowed both models to be fitted using the *gam* function directly.

Northing, at a 10km scale (*northing10*), and time, at a two day scale (*twoday*), composed our two-dimensional space-time joint variable. The same environmental variables used before were included as a proxy for available breeding habitat in each 10-km square. Given that most individuals breed in similar habitat, the importance of these variables was expected to be relatively low in the model, particularly for Pied Flycatchers which are restricted to woodland to breed. Ideally, we would have some kind of effort variable which, to our view, would be important when locating nest sites, such as the total amount of nests of any species recorded in each square or the number of nest recorders per square. However, these were not available and, unfortunately, we could not include them.

The general equation for the breeding onset models is:

$$\text{logit}(E(\text{Presence}_i)) = \beta_0 + t_1(\text{twoday}_i, \text{northing10}_i) + f_1(\text{altitude}_i) + \sum_{j=2}^9 f_j(\text{habitat}_{ij}) \quad (4.1)$$

The  $t$  function corresponds to a tensor product spline used for the spatio-temporal interaction variable, and  $f$  functions are different thin regression splines applied to *altitude* and 8 environmental covariates. Here  $i$  refers to the 2-day period and 10km square, and  $j$  refers to the environmental covariates.

### 4.4.1 Barn Swallow

This model includes 8 environmental covariates in addition to the spatio-temporal joint smooth. This number is lower than before, because a few previously included variables were now excluded due to not being present in most of the sampled locations: *marsh*, *mountain*, *saltwater*, *freshwater* and *urban*. Some of these areas represent feeding areas for the species, but do not necessarily support breeding structures or are under-surveyed by the nest recorder. The maximum  $k$  for the tensor product smooth was set at 24, and for the remaining covariates at 4.

The total explained deviance for this model was 11.6% (see Table 4.2). The shrinkage procedure removed the *improved grassland* covariate from the model, with all other covariates

showing up as significant. The explained deviance was low and is a probable consequence of the absence of an effort variable, massive variability in swallow nesting and behaviour, and small number of detections in of our dataset. Despite the sample size of 9792 units, these correspond to only 513 clutch initiation events, from 153 different squares, scattered through both clutch periods (most swallows are double-brooded). This means that only about half of this number (about 51%) correspond to nests recorded during the first brood period, our interest period. This makes our effective sample size decidedly small. Model validation and the bootstrapping procedure will, nonetheless, provide a better way of understanding this model’s uncertainty and quality of adjustment.

Table 4.2: Swallow’s clutch initiation model: coefficients, deviance & BIC

Variable	EDF	Chi.sq	p-value
te(twoday,northing10)	20.41	172.5	< 0.001
s(altitude)	3.86	24.30	< 0.001
s(broadleaf)	2.33	9.60	0.007
s(coniferous)	3.77	24.90	< 0.001
s(arable)	3.04	9.76	0.015
s(grassland)	3.36	18.92	< 0.001
s(improved)	0.00	0.00	0.637
s(heatherbog)	3.40	16.93	< 0.001
s(suburban)	3.79	15.28	0.003

*Deviance explained: 11.6%*  
*BIC: 3970.6*  
*sample size: 9792*

#### 4.4.2 Pied Flycatcher

This model contained 5 environmental variables in addition to the spatio-temporal joint smooth. The environmental variables were the ones used in the corresponding arrival model (3.3.2), with the exception of *urban* which was absent from this dataset. The maximum k for the tensor product smooth was set at 24, and for the remaining covariates at 4.

The explained deviance was 27.3%, notably higher than swallow’s analogous model. No environmental covariates were excluded on the selection procedure, although a few appeared not to be statistically significant, surprisingly *broadleaf* and *grassland*. Both these variables showed great importance in predicting the species’ presence upon arrival, but since they are highly linked to the flycatcher’s distribution, most recorded nests are within optimal habitat and, therefore, they are worse at explaining variability.

### 4.5 Model Validation

Repeating a similar procedure described in section 3.4, we performed ten-fold cross-validation to our models, subsetting the data by nest location instead of observer. The same metrics to assess quality of prediction were used (AUC, Sensitivity and Specificity) and can be seen on



Table 4.3: Flycatcher’s clutch initiation model: coefficients, deviance & BIC

Variable	EDF	Chi.sq	p-value
te(twoday,northing10)	20.98	304.69	< 0.001
s(altitude)	2.57	7.96	0.048
s(suburban)	0.90	7.26	0.004
s(grassland)	1.84	2.25	0.271
s(broadleaf)	2.62	3.54	0.238
s(conifer)	3.86	16.31	0.002

*Deviance explained: 27.3%*  
*BIC: 1929.7*  
*sample size: 2635*

Table 4.4. The swallow model performed poorly when predicting, with a disappointing AUC value, below 70%, as well as low sensitivity and specificity. On the contrary, the pied flycatcher model performed very well, with an AUC above 80%, and satisfactory individual values from sensitivity and specificity.

Table 4.4: Predictive abilities of the arrival models

Model	AUC	Sensitivity	Specificity
Barn Swallow	66.7%	65.9%	57.6%
Pied Flycatcher	80.6%	77.7%	71.5%

## 4.6 Extracting the Phenological Signal

Arrival was a cumulative process, because a bird that arrives will not leave until the end of the summer, and can be detected and observed multiple times throughout the full breeding period. However, clutch initiation events are single occurrences. In order to obtain the date when 50% of the population have initiated clutches, we assume clutch initiation events of first nesting attempts follow a normal distribution (Lundberg and Alatalo, 1992). The peak of clutch initiation events will give us the median date: when 50% of birds have laid their eggs. To estimate the peak laying date, we recorded the date when the first derivative of the distribution changed from positive to negative.

Figure 4.2 shows the double-brood phenomenon very well in Swallow, with two peaks in egg laying (one in spring, and another in summer). The change in the sign of the derivative will thus record the first peak, representing 50% of clutch initiation events. The corresponding Figure for Pied Flycatcher is in the appendices, showing mostly a single peak for the different northings. The same effect can be seen in Figure 4.3 for Swallow, which shows the variation of the probability of detection of a clutch initiation event with day and northing.

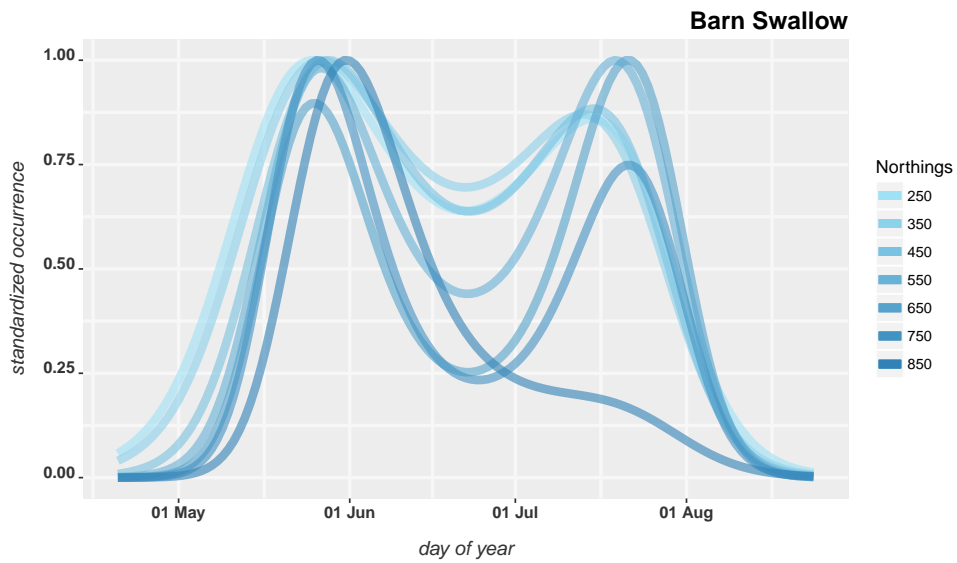


Figure 4.2: Breeding curves for different latitudes in Swallow

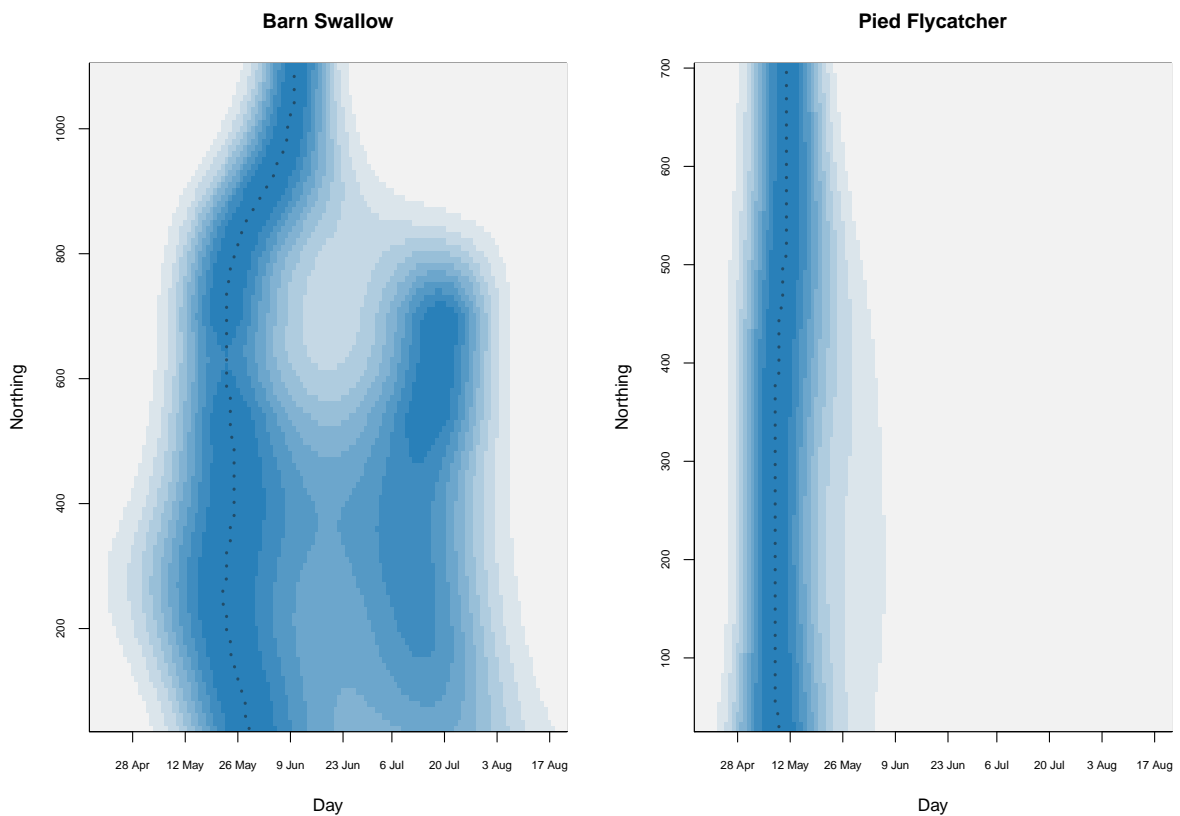


Figure 4.3: Breeding onset with latitude. This image shows our models' predictions for the probability of detection of a clutch initiation for different days and different northings. The intensity of blue varies from light blue, lowest probability, to dark blue, corresponding to the highest probability of a clutch initiation occurrence. The dotted line corresponds to the first peak in clutch initiation events.

## 4.7 Bootstrapping

Repeating the same procedure explained in section 3.6, we performed 200 bootstraps for each model, retrieving the derived date associated with the peak in clutch initiation events, our measure for breeding onset.

### 4.7.1 Results

The bootstrapped estimates for the breeding onset of 50% of the population are shown in Figure 4.4. Each species shows a different pattern due to very different associated variabilities. It is evident, however, that swallows breed later than flycatchers. Like in section 3.6, we describe the bootstrapped results individually and provide a table with sectioned results for easier reading.

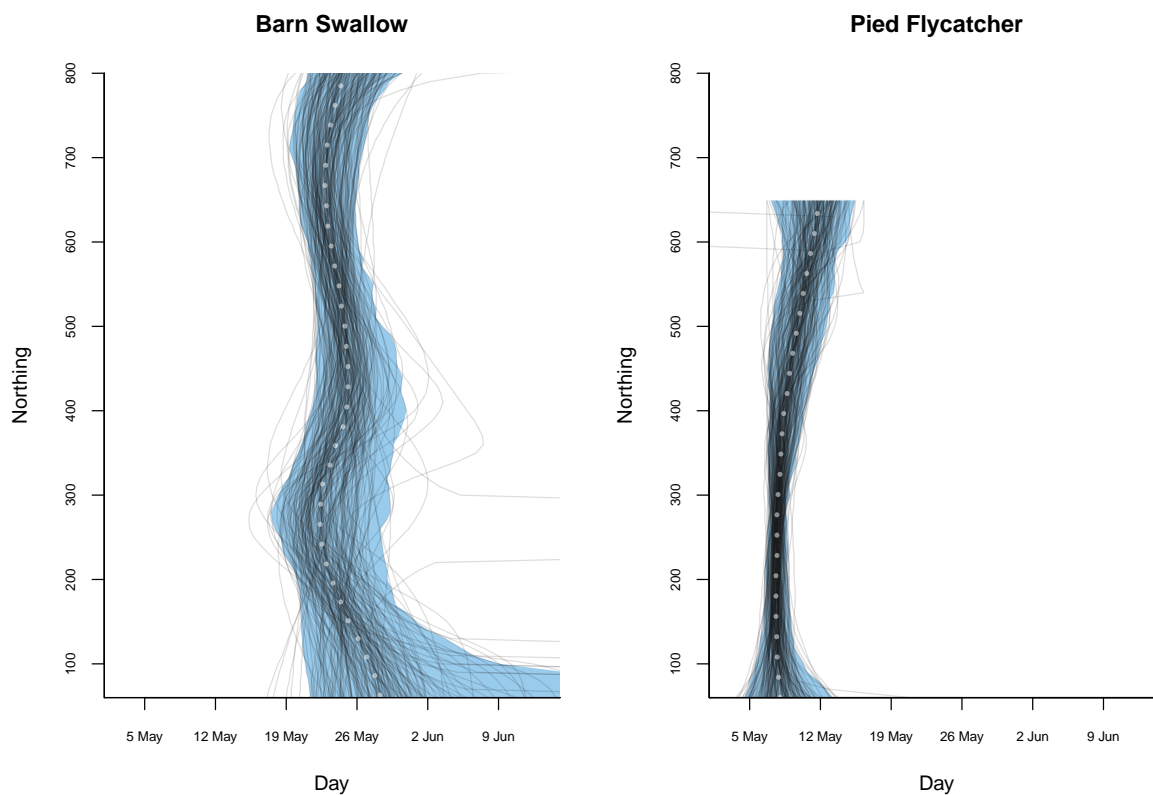


Figure 4.4: Breeding onset variation with northing. Each dark line represents the date when the peak in clutch initiation occurred for each latitude in each of the 200 bootstraps; the coloured band shows a 95% confidence interval for each latitude; and the dotted line gives the median value of the peak initiation date across all latitudes.

### Barn Swallow

Contrary to the arrival date estimates, which had low variability and were very precise for most of the range (section 4.7), the breeding onset estimates have broad confidence intervals, from the minimum of 7 days in the middle of the breeding range, increase to over 20 days at the edges. All estimates above northing 800km were disregarded given the enormous variability and are not plotted on Figure 4.4. All median dates for the 800km breeding range fall within a period

of 8 days, and dates appear to be rather close from 250km to 750km, with the medians for peak in clutch initiation occurring between the 23<sup>rd</sup> and 26<sup>th</sup> of May.

Table 4.5: Barn Swallow breeding onset bootstrap estimates

Northing	Median Estimate	Bootstrap Interval
50	148	(140 , 182)
150	144	(140 , 151)
250	141	(137 , 147)
350	143	(139 , 149)
450	144	(141 , 149)
550	143	(141 , 147)
650	142	(139 , 145)
750	143	(139 , 146)
850	147	(141 , 158)

#### 4.7.2 Pied Flycatcher

The estimates for breeding onset on the Pied Flycatcher show a remarkable similarity and high precision, with uncertainties varying from a minimum of 3 days (around 400km) to a maximum of 8 days at the northern distribution limit (600km). This contrasts with the arrival estimates which had larger variability, and with swallow's breeding onset estimates which show much larger variability. The median estimates increase little from to the south (7<sup>th</sup> of May) to the north (10<sup>th</sup> of May), varying only by 3 day. Therefore, our estimates show that, in 2014, median breeding onset of the population was close to being synchronous for most of the population of pied flycatchers.

Table 4.6: Pied Flycatcher breeding onset bootstrap estimates

Northing	Median Estimate	Bootstrap Interval
100	127	(125 ; 129)
200	127	(126 ; 128)
300	127	(126 ; 128)
400	128	(126 ; 129)
500	129	(126 ; 132)
600	130	(127 ; 134)

## 4.8 Discussion

In this chapter, we used data from Nest Record Scheme, a nest-related citizen science database, to estimate the date when our two study species initiated breeding in 2014, in Great Britain. To measure breeding onset at the population level, we estimated the peak in clutch initiation events as a derived estimate of modelling the probability of detecting a clutch event. Once again, uncertainty was provided via bootstrapping of our dataset and fitting the same model for each of the bootstrapped sets.

The quality of the data was very different in our two species: Pied Flycatcher had a high quality sample which provided good estimates for a 550km breeding range, whereas Barn Swallow had a lower sample dispersed between two broods and over an 800km breeding range. This means that our estimates, in the case of the Flycatcher, have narrow confidence intervals and allow to interpret broad patterns, opposed to swallow whose uncertainties prevent us from taking such solid conclusions.

Looking at the median dates, both species showed a narrow variation in breeding onset in 2014. The Pied Flycatcher showed a remarkably short delay in breeding onset from the south compared the north, with birds laying, on average, only 3 days later despite being 500km apart. In the case of swallow this happened with about a week of difference for a larger range, although we must be careful when interpreting these results.

This analysis has shown that laying date shows little variability with latitude, especially when compared to arrival. It suggests that laying date is a constrained phenological trait, with a broad scale set of factors largely impacting the full population at similar times. This could be a mixture of day-length, a temperature threshold, and possibly others (Valtonen *et al.*, 2017).

Additionally, while modelling the Swallow data, we noted that double brooding didn't occur in the north, being more frequent at the centre of the Swallow's distribution. More research, across more years and better data, could prove fruitful.

It should be noted that the much more common species (Swallow), with a much broader distribution, had a poorer quality dataset then the more localized, less abundant, species (Pied Flycatcher). Although the raw sample for the former included many thousands of nests, most had to be discarded because they had too few visits to obtain precise estimates of clutch initiation day, meaning that most volunteers, after finding a Swallow nest, don't follow it up. The exact opposite happened with the Pied Flycatcher, where the majority of nests had been properly followed and were integrated into our analysis. Volunteers should be encouraged to change practices and try to follow up Swallow nests to improve the quality and potential of NRS data regarding this species, and possibly others.

## Chapter 5

# Arrival & Breeding Interval

## 5.1 Introduction

In the previous two chapters, we focused on modelling two large-scale, population-level, processes in order to obtain two derived parameters: timing of arrival at the breeding grounds and timing of clutch initiation. These are two key processes of a bird's life cycle which relate to one another, as breeding can only occur after arrival at breeding habitat. In this chapter we intend to look further into this relationship, and understand how arrival and breeding onset varied through the territory in 2014. How fast can birds begin to breed? Do they need a fixed amount of time to do so? Is arrival constraining breeding onset? Here we will integrate the analysis conducted in the previous chapters to help solve some of these questions, which have also been the focus of past studies.

## 5.2 Obtaining the Gap

Analytically, the gap corresponds to the difference between the breeding onset day and the arrival day. To obtain this interval in a simple manner, we kept the model structure as similar as we could along the project and bootstrapped the results to obtain a direct and intuitive measure of uncertainty. Because of this, the extraction of the gap itself was straightforward - we subtracted each of the 200 sets of median clutch initiation dates by a randomly paired set of median arrival dates. Figure 5.1 allows the an intuitive visualization of the arrival-breeding gap, where arrival and breeding bootstrap estimates are plotted next to one another. It is visible how Pied Flycatchers on average breed sooner after arrival compared to Swallows, with the confidence intervals between arrival and breeding overlapping with one another.

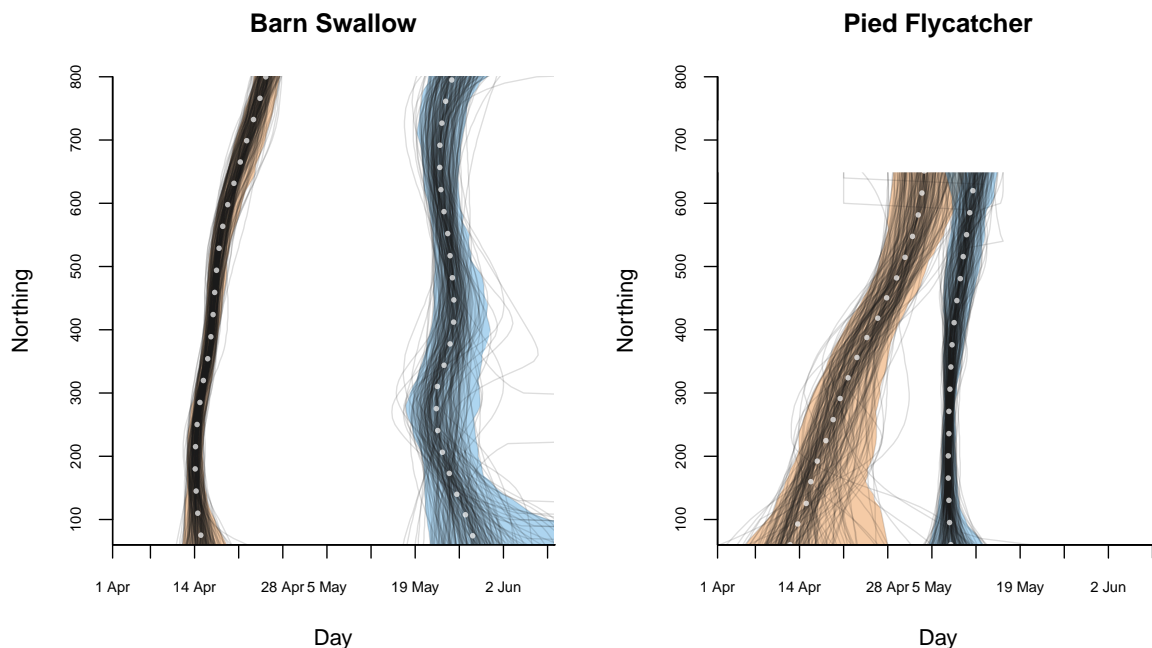


Figure 5.1: Arrival vs Breeding Onset. For each species arrival date bootstraps are plotted on the left with an orange background providing the bootstrap confidence intervals, while breeding onset is plotted on the right with the blue background. The dotted line provides the median value for each of the parameters.

## 5.3 Results

The main results of this section are provided with Figure 5.2, where the difference in days between breeding and arrival is plotted against northing. In both species, the median gap reduces with an increase in northing. Pied Flycatcher shows a shorter median gap than swallow in all its range.

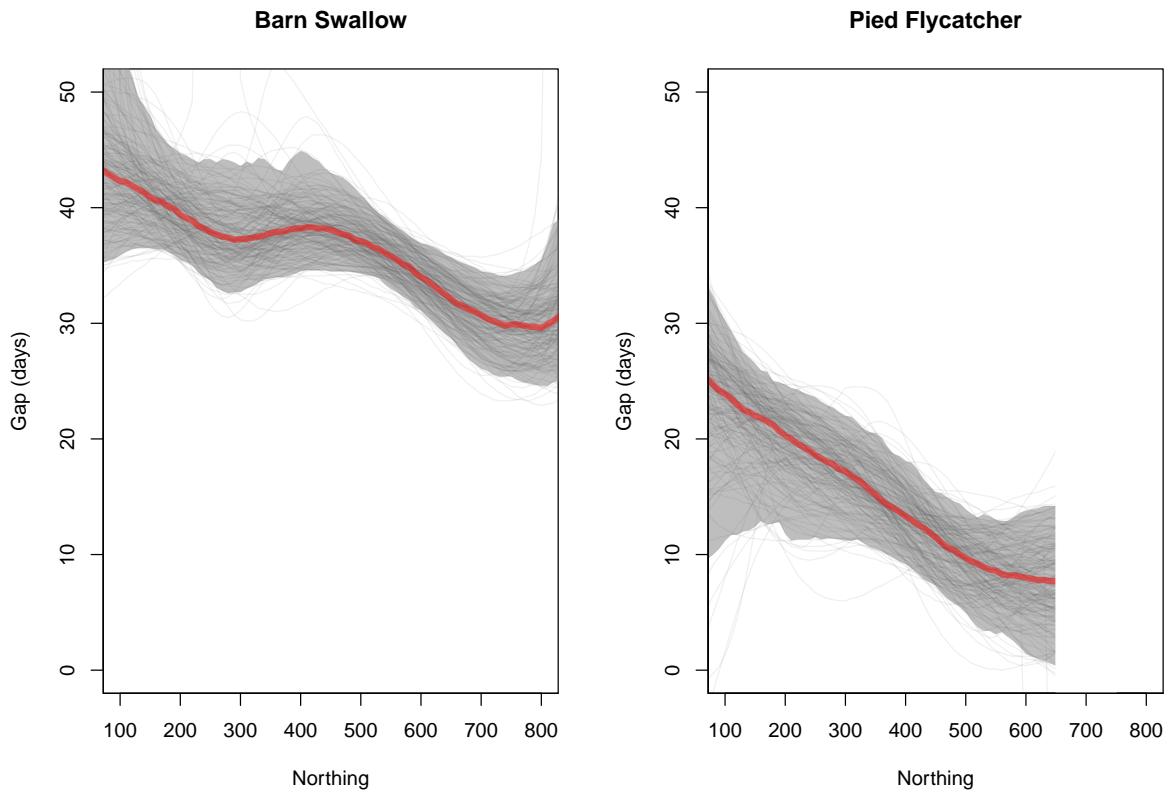


Figure 5.2: Arrival-Breeding Gap with latitude. Each line corresponds to the difference between randomly paired sets of dates from both models analysed in chapters 3 and 4. The grey band corresponds to 95% confidence intervals and the red line represents the median gap.

### Barn Swallow

Barn Swallow shows a "wiggly" reduction in the median values of the gap, but with broad and uneven variable confidence intervals which are a product of large associated uncertainties, resulting mostly from the breeding onset estimates. The median values at the southern end of the range were above 40 days and around 30 days in the north.

### Pied Flycatcher

Contrary to Swallow, Pied Flycatcher shows a straight reduction in gap duration in the south compared to the north. Values at the south start at around 25 days, lowering to below 10 in the northernmost territories, in a 15 days median gap reduction. Considering these measures correspond to a range of about 550km (smaller than Swallow's), this decrease can be considered rather notable, even if the associated confidence intervals are quite broad.



## 5.4 Discussion

Before delving into our results, it is important to situate 2014 in terms of weather to understand how regularly these patterns can be verified. Given the long route that long-distance migrants take, weather becomes a critical factor which deeply affects arrival (Both *et al.*, 2005). We cannot assess all the weather variations birds faced en route, however, breeding onset should be affected mostly by the local weather and MetOffice (2017) records provide detailed local and general weather reports which may be used. Spring of 2014 season was within the normal range of temperatures albeit being relatively warm, being slightly above the 30-years average for the UK, with a very average amount of rainfall (MetOffice, 2017). Locally, there were also not any drastic weather changes which, to our view, could justify a large impact to our conclusions. In conclusion, spring 2014 was a stable season in terms of weather in the UK.

Throughout this analysis, it has become clear that arrival is the more flexible process and has a larger variation than breeding onset. In 2014, when looking specifically at the interval between our two key phenological processes, our estimates indicate a reduction in gap length with an increase in latitude. However, each species should be interpreted very differently due to their own particularities.

For the Pied Flycatcher, we managed to successfully restrain our arrival data to individuals at the breeding grounds, due to this species being more scarcely seen on migration and normally at the coast, where breeding habitat is absent. Therefore, even if the sample size is not very large, and thus lead to bigger uncertainties in the estimates, we can be confident our estimates correspond to the real arrival dates of the population. In addition, our breeding onset model performed well on validation, and our breeding onset estimates were very precise with low variability. This sustains the observed pattern, in which Pied Flycatchers on average breed sooner after arrival in the north when compared to the south, with peak clutch initiation happening almost simultaneously (3 days difference) in the north and the south, 550km apart.

Barn Swallow, despite showing a very large sample size for arrival, does not allow the separation of migratory and breeding grounds, as they occur all around the territory and have unspecific habitat requirements. In an attempt to reduce this, all observations from main migratory watch points, which occur at the coast, were discarded. Still, the error is present and we must be careful when interpreting the results. This directly affects the interpretation of the absolute gap values, especially if we consider the southern areas might have more migrants passing north than the areas at the northern edge of their physical distribution. This might mean that our estimates for the southern dates are advanced in time and the median gap reduction we see might be caused by a confounding variable we cannot account for: detection of migratory birds. We would expect the false advance in arrival dates in the south to lead to a gap estimate that decreases with northing at a greater rate than reality. Therefore, the true decrease in gap with northing will be shallower than our estimate, if the detection of migrant birds in the south is affecting our results. In addition to this, our breeding onset estimations resulted from a relatively poor model, with a low explained deviance and disappointing validation values. Interpreting the relationship between arrival and breeding in this species, and the respective gap values, are made difficult by the handicaps in each of the stages, and more data and a different approach are required if variation in Swallow's phenology is to be more fully understood.

Birds that have a smaller gap between arrival and breeding may have less time to prepare for breeding. For example, less time to improve condition, or less time to find and defend a high quality territory. These factors could affect their breeding and this may be most evident in the number of eggs they lay (clutch size). Trying to understand if the reduction in gap length could have had consequences in terms of breeding success, we assessed the correlation between the gap and clutch size (directly linked to productivity in the Pied Flycatcher (Lundberg

and Alatalo, 1992)), which turned out to be weak and not statistically significant in both cases (Swallow:  $r = 0.008$ ,  $p\text{-value} = 0.85$ ; Pied Flycatcher:  $r = -0.005$ ,  $p\text{-value} = 0.93$ ).

Expanding this analysis to include more years, focusing on the Pied Flycatcher, will give us a better idea of how consistent the phenological pattern is and how it could fluctuate between years.

## Chapter 6

# Multi-year Analysis for Pied Flycatcher

At the start of this project, we chose a standard year weather-wise by talking to a number of specialists in the matter, deciding to study the relationship between arrival and breeding in a number of different species to understand common patterns. Further on, during the study species selection process, we realised that only two species presented enough data to work with. By the end of this project, we conclude that only the Pied Flycatcher analysis provided solid conclusions.

Although our findings did not provide multi-species conclusions as initially intended, they highlighted a single-species approach as a follow-up to this study. Is the same pattern observed in all years? Is there always a gap reduction associated with northern birds? To test this hypothesis, we extended the analysis we developed in this thesis, applying most of the same principles, to model the arrival and breeding onset of the Pied Flycatcher across several years. Additionally, we compared our model estimates, derived from our broad-scale analysis, to precise data collected on individuals at one location in southern England using GPS transmitters and field observations. Although all the results are not presented here (after all a thesis must have a start and an end!), the main results and conclusions of that effort are summarized next.

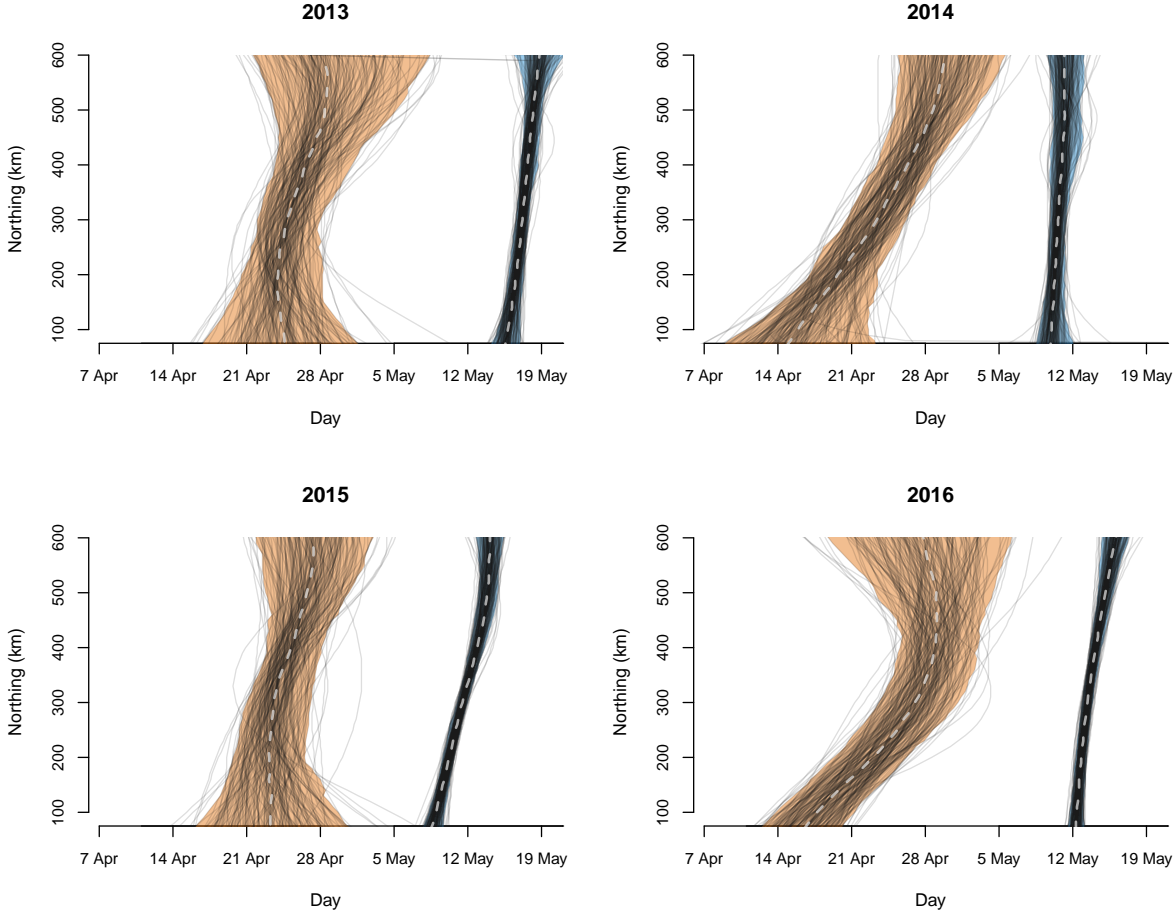


Figure 6.1: Arrival and Breeding initiation bootstrap plots.

The period for this multi-year analysis was 2013-2016. The same methodology described in the past chapters was applied to all years, although some parameters in the models were changed to be suitable for all years. Figure 6.1 shows arrival and breeding onset estimates for the Pied Flycatcher's breeding range in Great Britain.

In all years, arrival in the north was later than in the south. However, there were between year differences in arrival times, mostly in the south of the range. In 2014 and 2016, the years with the earliest arrival at the southern grounds, birds still arrived at the north around the same week compared to 2013 and 2015. Regarding breeding onset, a different pattern was observed. Even though birds laid their first egg later in the north compared to the south, this difference was much less pronounced: it was as short as 2 days and as large as 6 days. In terms of gap, we saw a marked reduction from north to south in two years, the ones with the earliest arrival, while the other two years showed an approximately constant gap with no clear trend.

In 2015 and 2016, Pied Flycatchers in southern England (northing: 78 km) were GPS-tagged and provided estimates of individual arrivals for a total of 86 birds (Malcolm Burgess, unpublished data). Intense fieldwork following these individuals also provided estimates of the first egg dates for the nests of these birds. For each individual, we calculated the gap between arrival and the onset of breeding. We used the median gap across all individuals to compare with the population estimates of the gap from our models for the same northing. These served as practical validation of our models and also enabled us to compare individual estimates with the population. In both years, our estimated population medians were within the confidence intervals for the medians using the individual data (cf. point estimate and CI in Figure 6.2). This two-year empirical validation exercise helped validate the quality of adjustment in our models, and gives us confidence that our study provided solid conclusions and that our methodology can be used in further studies with different species or similar ends.

This analysis highlighted that arrival timing is a very flexible process, with a great variation between and within years, especially in the southern territories. On the contrary, breeding appears to be a much less variable process, and birds tend to lay within few days apart through the range, varying less than a week between the studied years. In years where arrival is earlier, birds tend to wait longer to breed than years where birds arrive later, which suggests that arriving earlier at the breeding grounds does not appear to bring any advantage in terms of clutch initiation.

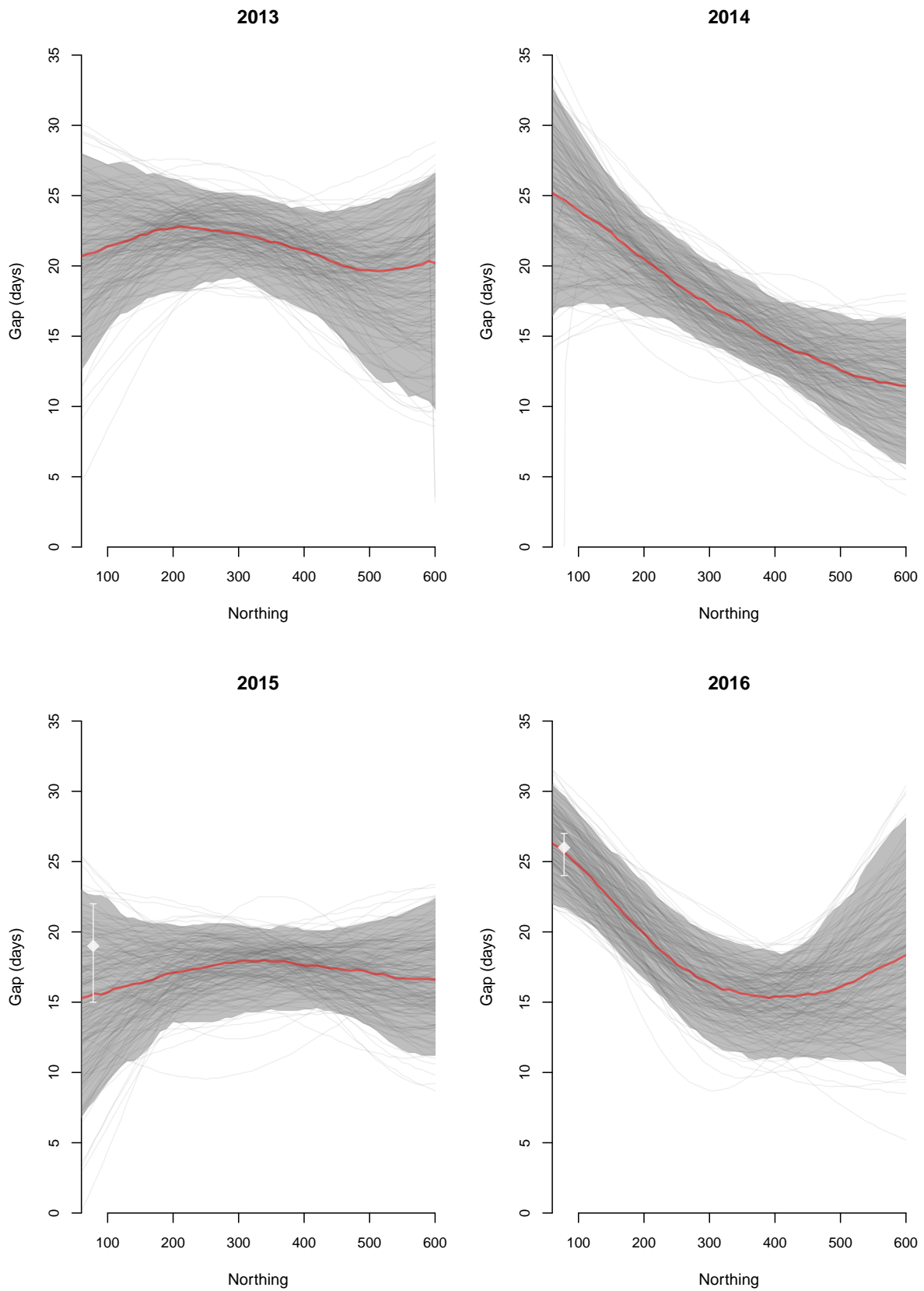


Figure 6.2: Arrival and Breeding initiation bootstrap plots. Northing is measured in kilometres.

## Chapter 7

# Conclusions

## 7.1 Project Overview

This thesis aimed to obtain derived estimates of GAMs to approximate phenological parameters, based on a number of theoretical assumptions. This was a rather complex process with many small intermediate steps. Nonetheless, we managed to produce biologically reasonable and partially empirically verifiable results. All this while using data collected by ordinary people, potentially subject to a number of biases. Even if in the end only the results for one of the species are ecologically interesting, instead of two as initially intended, this outcome is nothing short of extraordinary.

Initially we wanted to work with a number of different species, showing how phenology of reproduction varied with latitude in a single year. Soon we realized we would not have enough data in both datasets to do it: some species had really good arrival data but lacked good breeding datasets, while others showed the opposite. We had to settle for two species which we thought would give us decent estimates, but through the project, we realised one of them would probably not meet the goal. We ended up sticking to one species which in 2014 showed a striking pattern. This hinted that a strong relationship between breeding and arrival was happening: the later the arrival, the faster the timing to breed. However, when expanding this analysis to more years, we actually saw that perhaps no relationship was happening whatsoever. It seems that it really does not matter when birds arrive, they will still breed around the same time. Obviously, we do not know if there are consequences of needing to breed sooner after arrival. Is the breeding productivity affected? We do not know if birds are in condition to breed as they arrive at the breeding grounds or if they make use of longer days in the north to get in condition faster. How much is productivity affected by these nuances? There is so much more to unfold about this topic, especially if we consider different species with different ecological strategies. This is now a possible task for future work, both ours and from others.

## 7.2 Acquired Competencies & Outcomes

At a personal level, this project has taught me what it is to be a scientist. What it is to be in a place bubbling with ideas to solve real life problems (BTO). It allowed me to share and absorb ideas from specialists who know so much and from whom I have learnt tremendously. Many of the skills I acquired along this journey are worth mentioning here.

Broadly, through this project I have:

1. learnt about the scientific thought process and work flow;
2. improved my scientific writing skills and the English language;
3. improved my presentation skills while conducting a seminar and project meetings;
4. improved my programming skills and increased my confidence in programming;
5. improved my overall problem-solving skills.

Specifically, I've learnt to:

1. deal with large matrices and process them, using a number of R-packages such as `dplyr`;
2. model spatio-temporal data;



3. apply skills acquired in the masters program to learning completely new statistical methodologies;
4. apply and (partially) comprehend GAMs, using R-package `mgcv` and many of its functions to fit models, obtain predictions and work with derived estimates;
5. work with different kinds of splines and apply shrinkage selection;
6. test models using ten-folded cross-validation;
7. obtain uncertainty estimates and propagating uncertainty across multi-stage inferential procedures using non-parametric bootstraps;
8. develop a series of functions to improve analysis efficiency;
9. use an external server to run many different models simultaneously;
10. create elaborate plots using R's default function and package `ggplot2`;
11. write a thesis in LaTeX;

The main outcomes from this project are:

1. a Masters dissertation;
2. a seminar presented at the BTO;
3. an oral communication presented in the Portuguese Statistical Society congress;
4. an article for BTO News;
5. a scientific paper to be published in a peer-reviewed journal.

### 7.3 Final Remark

Through this thesis, statistical methodologies were applied to answer practical ecological questions. Specifically, a critical period in migratory birds' life cycle was studied, leading to a better understanding of the variation of key ecological processes at a broad scale.

This thesis shows how initial ideas do not always lead to what would be expected. But, along the way, some unexpected answers arise and even more questions are raised. This provides a good example of what truly is the scientific process.

# Bibliography

- Ahola, M., T. Laaksonen, K. Sippola, T. Eeva, K. Rainio, and E. Lehikoinen (2004). Variation in climate warming along the migration route uncouples arrival and breeding dates. *Global Change Biology* 10, 1610–1617.
- Balmer, D., S. Gillings, B. Caffrey, B. Swann, and I. Downie (2013). *Bird Atlas 2007-11*. British Trust for Ornithology.
- BirdLife International (2016a). *Ficedula hypoleuca*. The IUCN red list of threatened species 2016: e.t22709308a87938122.
- BirdLife International (2016b). *Hirundo rustica*. The IUCN red list of threatened species 2016: e.t22712252a87461332.
- Bohning-Gaese, K. and H.-G. Bauer (1996). Changes in species abundance, distribution, and diversity in a Central European bird community. *Conservation Biology* 10, 175–187.
- Both, C., R. G. Bijlsma, and M. E. Visser (2005). Climatic effects on timing of spring migration and breeding in a long-distance migrant, the pied flycatcher *Ficedula hypoleuca*. *Journal of Avian Biology* 36, 368–373.
- Both, C. and M. E. Visser (2001). Adjustment to climate change is constrained by arrival date in a long-distance migrant bird. *Nature* 411, 296–298.
- Brewer, M. J., A. Butler, and S. L. Cooksley (2016). The relative performance of AIC, AICC and BIC in the presence of unobserved heterogeneity. *Methods in Ecology and Evolution* 7, 679–692.
- Brotons, L. and S. Herrando (2003). Effect of increased food abundance near forest edges on flocking patterns of coal tit *Parus ater* winter groups in mountain coniferous forests. *Bird Study* 50, 106–111.
- Brown, C. R. and M. B. Brown (2000). Heritable basis for choice of group size in a colonial bird. *Proceedings of the National Academy of Sciences* 97, 14825–14830.
- BTO, RSPB, B. Ireland, SOC, and WOS (2017). Birdtrack. [www.birdtrack.net](http://www.birdtrack.net). Accessed: August 2017.
- Bunce, R. G., C. J. Barr, and H. A. Whittaker (1981). An integrated system of land classification. *Annual Report of the Research Institute of Terrestrial Ecology*, 28–33.
- Burger, C., E. Belskii, T. Eeva, T. Laaksonen, M. Mägi, R. Mänd, A. Qvarnström, T. Slagsvold, T. Veen, M. E. Visser, K. L. Wiebe, C. Wiley, J. Wright, and C. Both (2012). Climate change, breeding date and nestling diet: how temperature differentially affects seasonal changes in pied flycatcher diet depending on habitat variation. *Journal of Animal Ecology* 81, 926–936.

- Burman, M. S. (2016). *Citizen science reveals complex changes in barn swallow phenology in South Africa over three decades*. Ph. D. thesis, University of Cape Town.
- CBI Ministry of Foreign Affairs (2015). CBI product fact sheet: birdwatching tourism from Europe. Technical report.
- Cohn, J. P. (2008). Citizen science: Can volunteers do real research? *BioScience* 58, 192.
- Cotton, P. A. (2003). Avian migration phenology and global climate change. *Proceedings of the National Academy of Sciences* 100, 12219–12222.
- Crick, H. Q., S. R. Baillie, and D. I. Leech (2003). The UK nest record scheme: its value for science and conservation. *Bird Study* 50, 254–270.
- DiCiccio, T. J., B. Efron, P. Hall, M. A. Martin, A. J. Canty, A. C. Davison, D. V. Hinkley, L. J. Gleser, S. M. S. Lee, G. A. Young, T. J. DiCiccio, and B. Efron (1996). Bootstrap confidence. *Statistical Science* 11, 189–228.
- Efron, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann. Statist.* 7, 1–26.
- European Environment Agency (2007). Halting the loss of biodiversity by 2010: proposal for a first set of indicators to monitor progress in Europe. Technical Report 11.
- Freeman, E. A. and G. Moisen (2008). PresenceAbsence: An R Package for presence absence analysis. *Journal of Statistical Software* 23.
- García-Pérez, B., K. A. Hobson, G. Albrecht, M. D. Cadman, and A. Salvadori (2014). Influence of climate on annual survival of Barn Swallows (*Hirundo rustica*) breeding in North America. *The Auk* 131, 351–362.
- Gimenez, O., S. T. Buckland, B. J. T. Morgan, N. Bez, S. Bertrand, R. Choquet, S. Dray, M.-P. Etienne, R. Fewster, F. Gosselin, B. Mérigot, P. Monestiez, J. M. Morales, F. Mortier, F. Munoz, O. Ovaskainen, S. Pavoine, R. Pradel, F. M. Schurr, L. Thomas, W. Thuiller, V. Trenkel, P. de Valpine, and E. Rexstad (2014). Statistical ecology comes of age. *Biology Letters* 10, 1–4.
- Goodenough, A. E., A. G. Hart, and S. L. Elliot (2010). What prevents phenological adjustment to climate change in migrant bird species? Evidence against the “arrival constraint” hypothesis. *International Journal of Biometeorology* 55, 97–102.
- Guisan, A., T. C. Edwards, and T. Hastie (2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling* 157, 89–100.
- Hahn, S., S. Bauer, and F. Liechti (2009). The natural link between Europe and Africa - 2.1 billion birds on migration. *Oikos* 118, 624–626.
- Hesterberg, T., N. H. Choi, L. Meier, and C. Fraley (2008). Least angle and  $\ell_1$  penalized regression: A review. *Statistics Surveys* 2, 61–93.
- Hewson, C. M. and D. G. Noble (2009). Population trends of breeding birds in British woodlands over a 32-year period: relationships with food, habitat use and migratory behaviour. *Ibis* 151, 464–486.
- Lundberg, A. and R. V. Alatalo (1992). *The Pied Flycatcher (T & AD Poyser)*. Academic Press.
- Manly, B. (2006). *Randomization, Bootstrap and Monte Carlo Methods in Biology* (3 ed.). Taylor & Francis Inc.

- Marra, G. and S. N. Wood (2011). Practical variable selection for generalized additive models. *Computational Statistics & Data Analysis* 55, 2372–2387.
- MetOffice (2017). Weather and climate change. [www.metoffice.gov.uk](http://www.metoffice.gov.uk). Accessed: 2017-06-15.
- Miles, W. T. S., M. Bolton, P. Davis, R. Dennis, R. Broad, I. Robertson, N. J. Riddiford, P. V. Harvey, R. Riddington, D. N. Shaw, D. Parnaby, and J. M. Reid (2016). Quantifying full phenological event distributions reveals simultaneous advances, temporal stability and delays in spring and autumn migration timing in long-distance migratory birds. *Global Change Biology* 23, 1400–1414.
- Miller-Rushing, A. J., T. T. Hoye, D. W. Inouye, and E. Post (2010). The effects of phenological mismatches on demography. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365, 3177–3186.
- Newton, I. (2013). *The Migration Ecology of Birds* (1 ed.). Elsevier.
- Orme, C. D. L., R. G. Davies, V. A. Olson, G. H. Thomas, T.-S. Ding, P. C. Rasmussen, R. S. Ridgely, A. J. Stattersfield, P. M. Bennett, I. P. F. Owens, T. M. Blackburn, and K. J. Gaston (2006). Global patterns of geographic range size in birds. *PLoS Biology* 4, e208.
- Ouwehand, J. and C. Both (2016). African departure rather than migration speed determines variation in spring arrival in pied flycatchers. *Journal of Animal Ecology* 86, 88–97.
- Padoa-Schioppa, E., M. Baietto, R. Massa, and L. Bottoni (2006). Bird communities as bioindicators: The focal species concept in agricultural landscapes. *Ecological Indicators* 6, 83–93.
- Pulido, F. (2007). The genetics and evolution of avian migration. *BioScience* 57, 165–174.
- R Development Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Robinson, R., J. Marchant, D. Leech, D. Massimino, M. Sullivan, S. Eglinton, C. Barimore, D. Dadam, I. Woodward, M. Hammond, S. Harris, D. Noble, R. Walker, and S. Baillie (2016). Birdtrends 2016: trends in numbers, breeding success and survival for uk breeding birds. resreport 691, BTO, Thetford.
- Root, T. L., J. T. Price, K. R. Hall, S. H. Schneider, C. Rosenzweig, and J. A. Pounds (2003). Fingerprints of global warming on wild animals and plants. *Nature* 421, 57–60.
- Saino, N., R. Ambrosini, D. Rubolini, J. von Hardenberg, A. Provenzale, K. Huppopp, O. Huppopp, A. Lehikoinen, E. Lehikoinen, K. Rainio, M. Romano, and L. Sokolov (2010). Climate warming, ecological mismatch at arrival and population decline in migratory birds. *Proceedings of the Royal Society B: Biological Sciences* 278, 835–842.
- Silvertown, J. (2009). A new dawn for citizen science. *Trends in Ecology & Evolution* 24, 467–471.
- Sing, T., N. Beerenwinkel, and T. Lengauer (2004). Learning mixtures of localized rules by maximizing the area under the ROC curve. In J. Hernández-Orallo, C. Ferri, N. Lachiche, and P. A. Falch (Eds.), *ROCAI-2004: 1st International Workshop on ROC Analysis in Artificial Intelligence*, Valencia, Spain, pp. 89–96. University of Valencia.
- Snäll, T., O. Kindvall, J. Nilsson, and T. Pärt (2011). Evaluating citizen-based presence data for bird monitoring. *Biological Conservation* 144, 804–810.

- Sullivan, B. L., J. L. Aycrigg, J. H. Barry, R. E. Bonney, N. Bruns, C. B. Cooper, T. Damoulas, A. A. Dhondt, T. Dietterich, A. Farnsworth, D. Fink, J. W. Fitzpatrick, T. Fredericks, J. Gerbracht, C. Gomes, W. M. Hochachka, M. J. Iliff, C. Lagoze, F. A. L. Sorte, M. Merrifield, W. Morris, T. B. Phillips, M. Reynolds, A. D. Rodewald, K. V. Rosenberg, N. M. Trautmann, A. Wiggins, D. W. Winkler, W.-K. Wong, C. L. Wood, J. Yu, and S. Kelling (2014). The eBird enterprise: An integrated approach to development and application of citizen science. *Biological Conservation* 169, 31–40.
- Sullivan, B. L., C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling (2009). eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation* 142, 2282–2292.
- Tiago, P. M. N. (2017). *Improving citizen science as a tool for biodiversity monitoring*. Ph. D. thesis, Doutoramento em Biologia e Biodiversidade, Faculdade De Ciências, Universidade De Lisboa.
- Turner, A. (2006). *The Pied Flycatcher (T & AD Poyser)*. Academic Press.
- Valtonen, A., R. Latja, R. Leinonen, and H. Pöysä (2017). Arrival and onset of breeding of three passerine birds in eastern finland tracks climatic variation and phenology of insects. *Journal of Avian Biology* 48, 785–795.
- Verhulst, S. and J.-A. Nilsson (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.
- Visser, M. E. (2008). Keeping up with a warming world; assessing the rate of adaptation to climate change. *Proceedings of the Royal Society B: Biological Sciences* 275, 649–659.
- Walther, G.-R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J.-M. Fromentin, O. Hoegh-Guldberg, and F. Bairlein (2002). Ecological responses to recent climate change. *Nature* 416, 389–395.
- Wood, S. (2017). *An Introduction to Generalized Additive Models with R*. Taylor & Francis Inc.
- Wood, S. N., Y. Goude, and S. Shaw (2014). Generalized additive models for large data sets. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 64, 139–155.

# Appendices

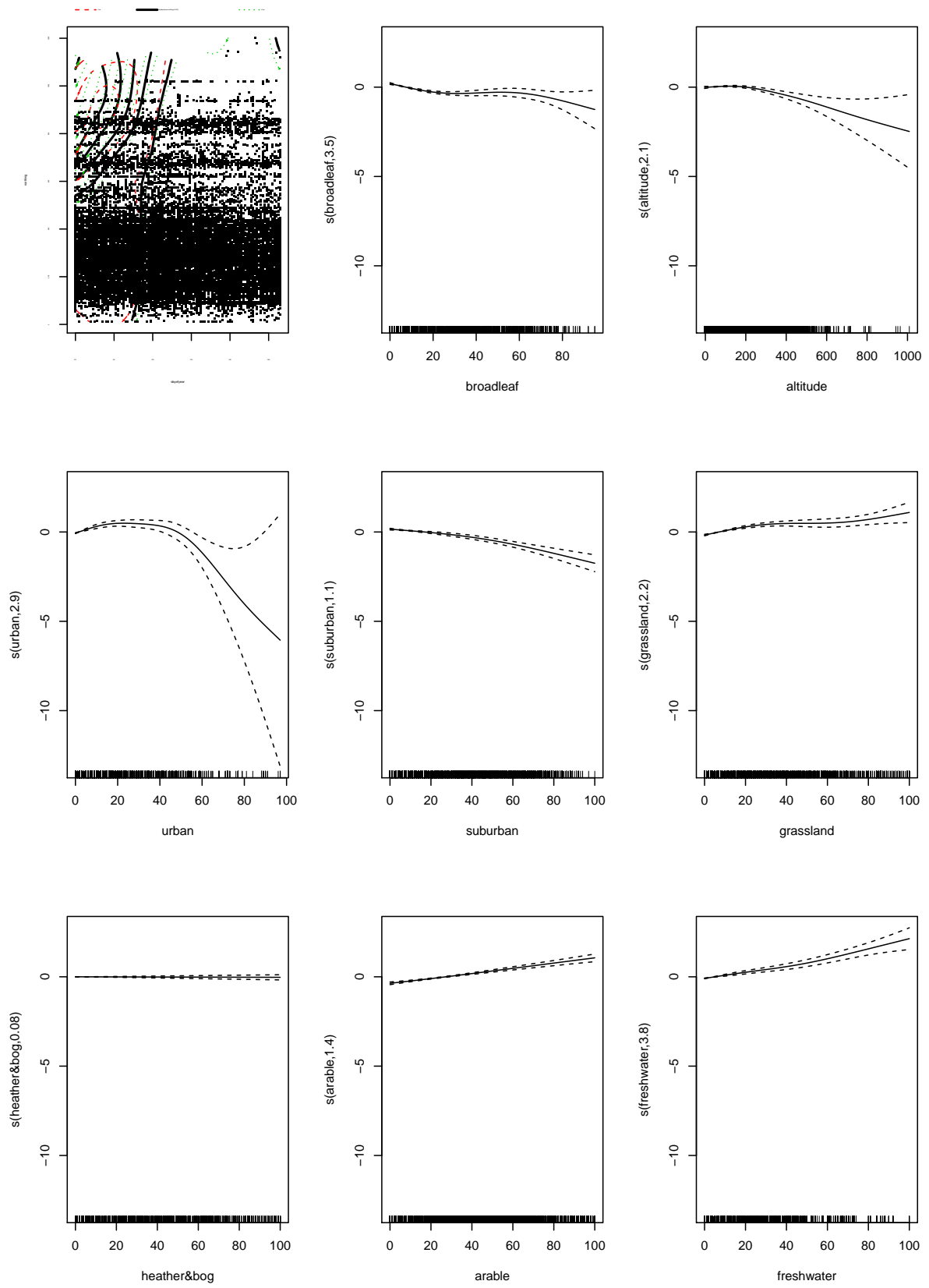


Figure 1: Birdtrack Model Swallow: Covariates

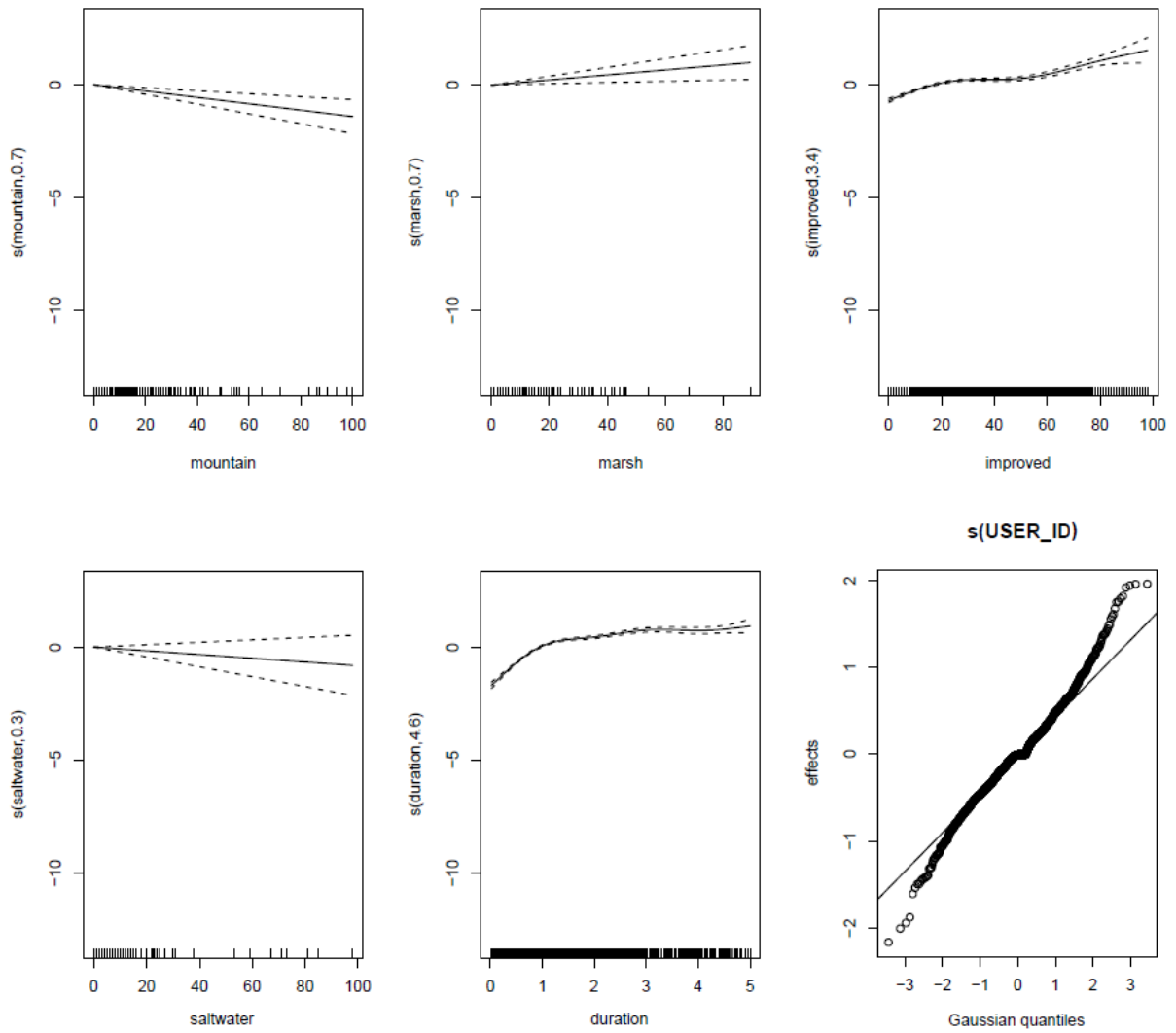


Figure 2: Birdtrack Model Swallow: Covariates



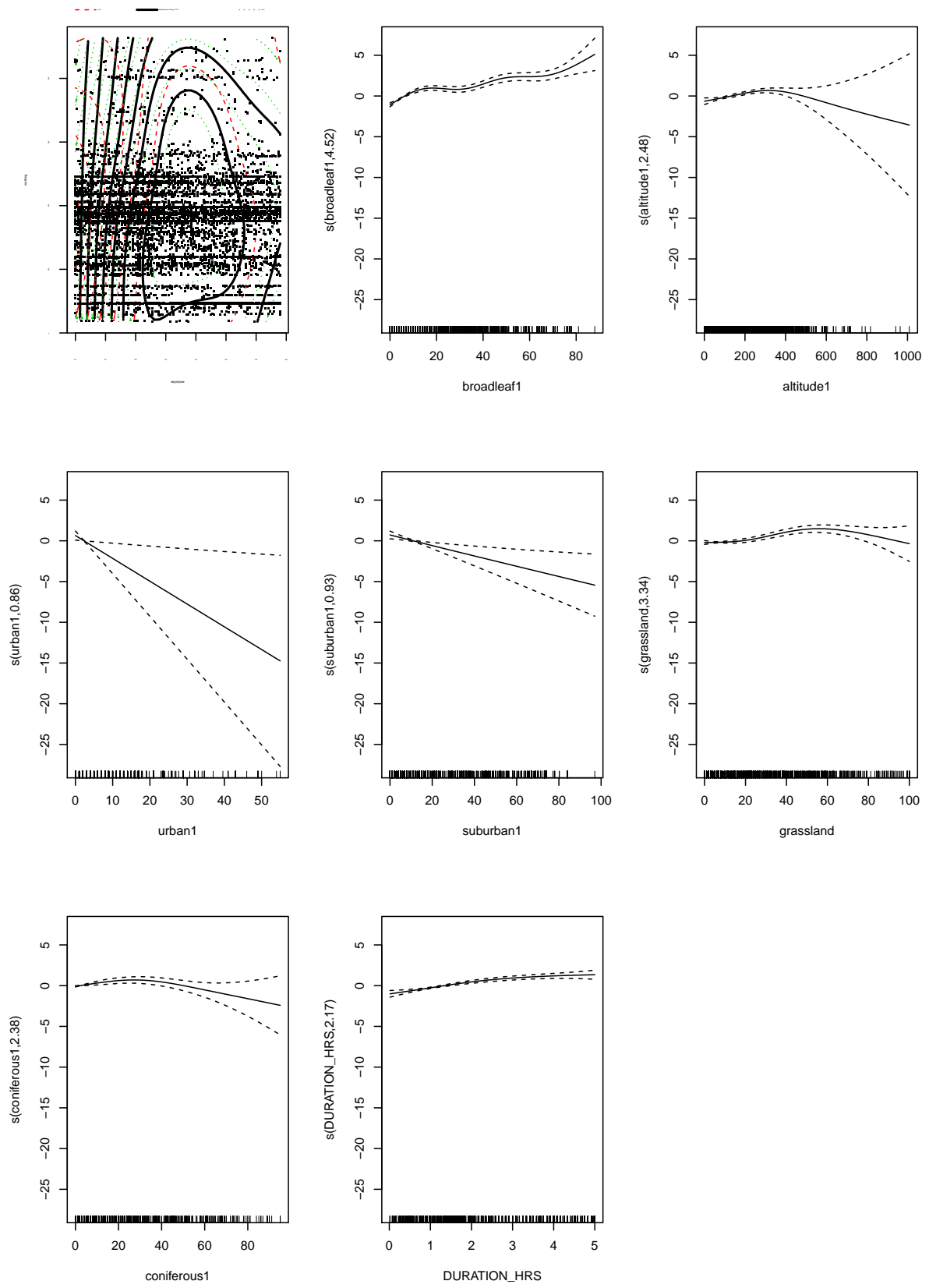


Figure 3: Birdtrack Model Pied Flycatcher: Covariates

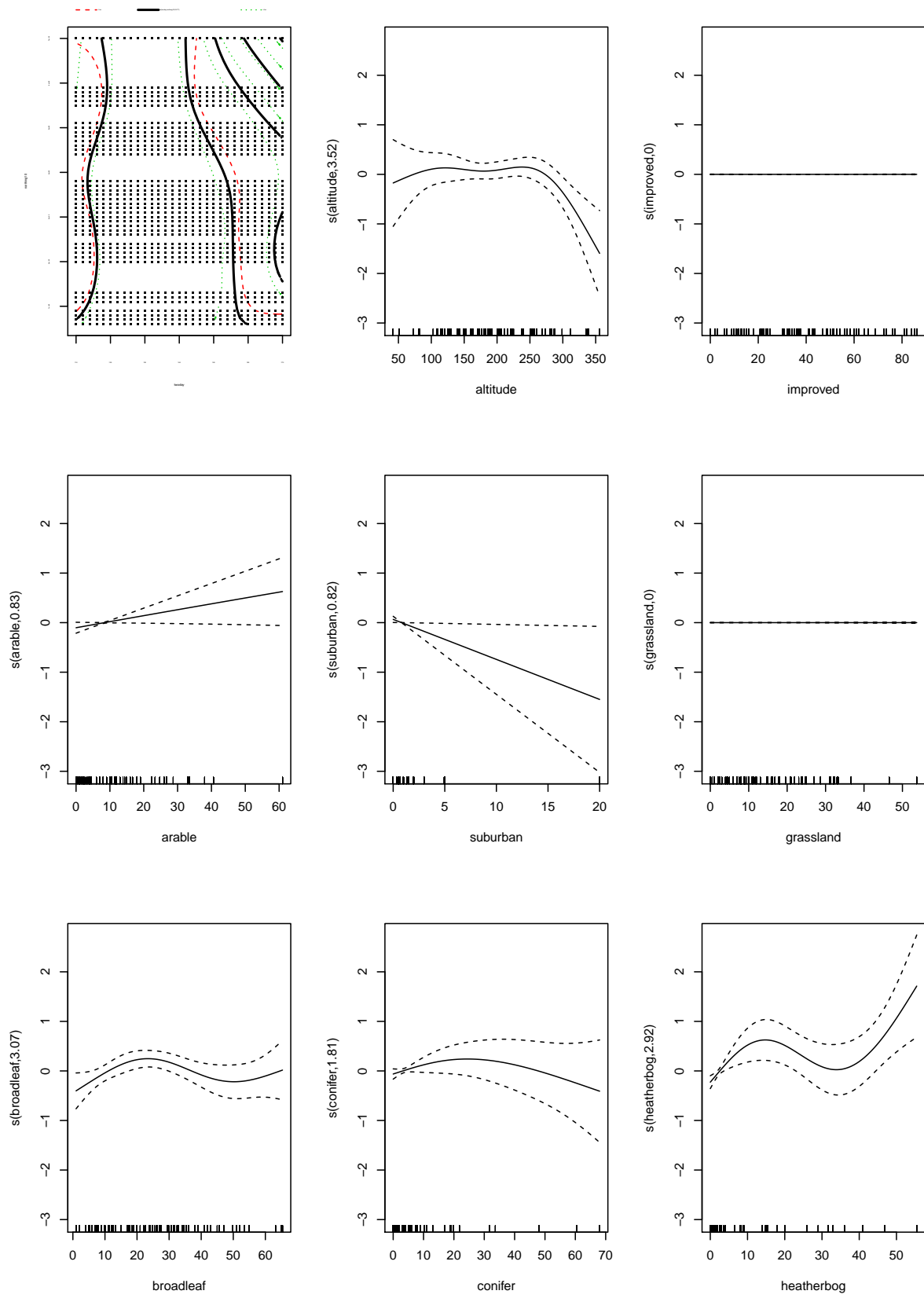


Figure 4: NRS Model Swallow: Covariates

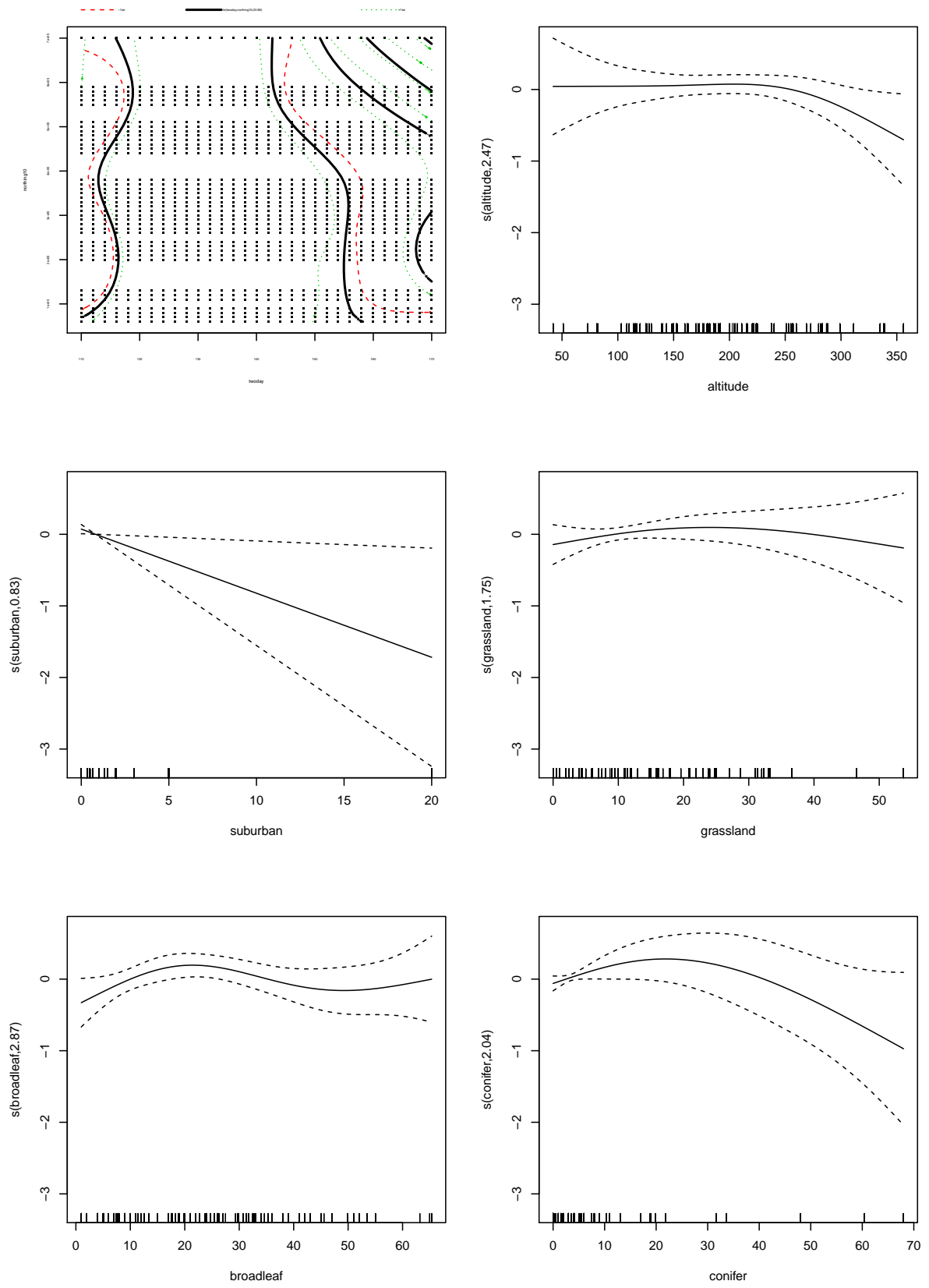


Figure 5: NRS Model Pied Flycatcher: Covariates

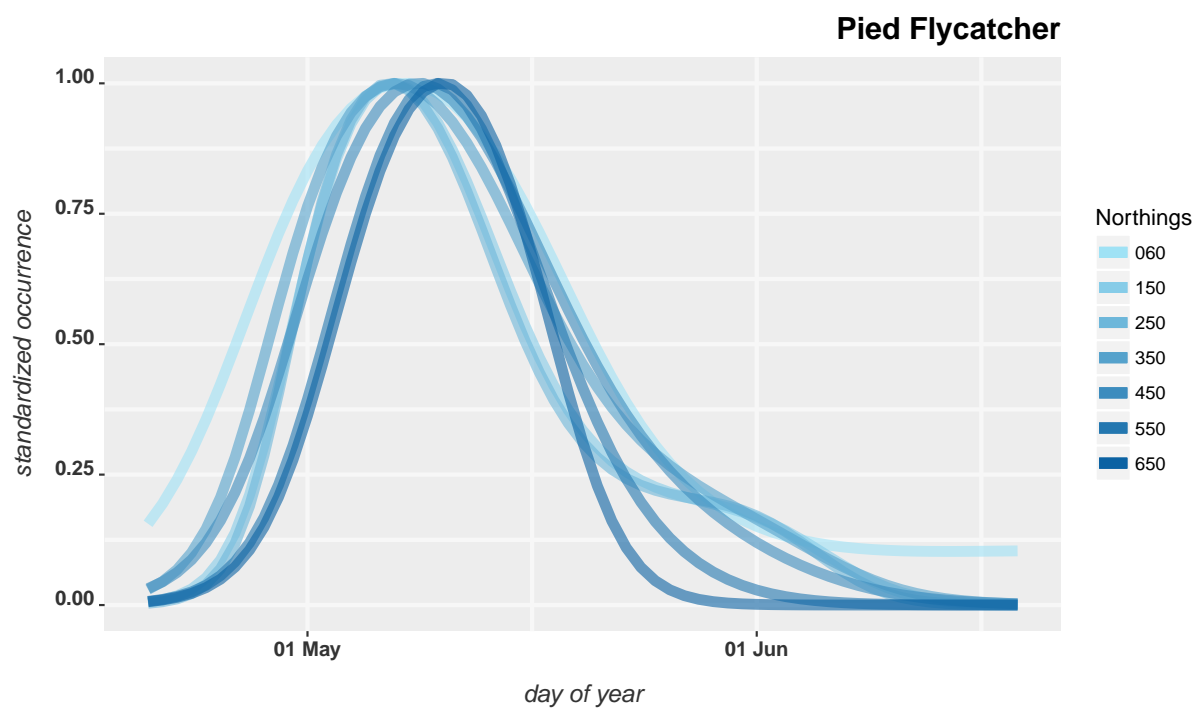


Figure 6: Detection of clutch initiation events for different latitudes in Pied Flycatcher, with a single brood peak.