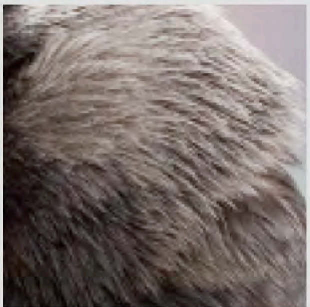
Heathland Birds Bioacoustic Monitoring

June 2019





baker consultants

Who we are:
Baker Consultants is an ecology and sustainability consultancy. We work in terrestrial, freshwater and marine environments, providing a range of services to industry, government, developers, public services and utilities.

Baker Consultants comprises a highly experienced team of professional ecologists. We do wildlife surveys - but they are only the first steps in the process for most projects. We are also involved in ecological assessment, environmental law, biodiversity management and design planning.

We don't just work with wildlife, because we know that communication with clients, design teams and conservation bodies is the key to project success. Explaining the implications of survey data, and interpreting legislation, policy and best practice is one of our strengths. We help decisions to be made and actions taken, allowing constraints to be kept to a minimum and project risks to be managed.

Our approach is scientific, pragmatic and creative. Alongside tried and tested methods, we seek to innovate, introduce clients to new ways of thinking and always deliver sound commercial awareness. You will find us honest and approachable, but we're not afraid to be robust and challenging - or to ask difficult questions.

We do believe in nature conservation. But we also believe in good development, well delivered. We know that, with our input, projects and plans can provide benefits for both nature and people.

That's not the whole story.
For more information, look at our web site www.bakerconsultants.co.uk, subscribe to our blog, or call us on 01629 593958.

Client	Natural England
Project	Heathland Birds Bioacoustic Monitoring
Report title	Final Report
File reference	972_FINAL_CA _MW_V2.docx
Team leader	Carlos Abrahams
Contact details	c.abrahams@bakerconsultants.co.uk

	Name	Position	Date
Author	Carlos Abrahams	Technical Director	2 May 2019
Reviewed	Mark Woods	Principal Ecologist	2 May 2019
Revised	Carlos Abrahams	Technical Director	7 June 2019

Baker Consultants Ltd
Cromford Station
Cromford Bridge
Matlock
Derbyshire DE4 5JJ
info@bakerconsultants.co.uk
www.bakerconsultants.co.uk
01629 593958

Company No. 6702156

Disclosure and Limitation:

Baker Consultants has prepared this document for the sole use of the commissioning client in accordance with the agreed scope of works and Terms and Conditions under which our services were performed. The evidence and opinion provided is true and has been prepared in accordance with the guidance of our professional institution's Code of Professional Conduct. No other warranty is made as to the professional advice included in this document or any other services provided by us. This document may not be relied upon by any third party without the prior and express written agreement of Baker Consultants.

Unless otherwise stated in this document, the assessments made assume that the study site referred to will continue to be used for its current purpose without significant change. The assessment, recommendations and conclusions contained in this document may be based upon information provided by third parties and upon the assumption that the information is relevant, correct and complete. There has been no independent verification of information obtained from third parties, unless otherwise stated in the report.

Where field investigations have been carried out, these have been restricted to the agreed scope of works and carried out to a level of detail required to achieve the stated objectives of the services. Natural habitats and species distributions may change over time and further data should be sought following any significant delay from the publication of this document.

Report Contents

Executive Summary

1. Effective long-term monitoring of rare and declining species is critical to enable their conservation, particularly when environmental processes such as climate change may cause significant threats. However, gathering consistent and reliable data on target species can often be difficult due to issues including detectability during surveys, or the seasonal and diurnal timing requirements for survey. Due to such reasons, many existing monitoring practices and protocols are often constrained in terms of their resource efficiency and the value of their outcomes. A review of bird monitoring programmes from around Europe concluded that the design of most monitoring schemes could be significantly improved, particularly in terms of the number of repeat sampling events, the use of unbiased random sampling locations, the optimization of sampling effort, and more efficient statistical use of the data gathered.

2. The Thames Basin and Wealden Heaths Special Protection Areas (SPAs) are internationally designated under the EU Birds Directive 2009 due to their important breeding populations of the Annex 1 bird species, European nightjar *Caprimulgus europaeus*, woodlark *Lullula arborea* and Dartford warbler *Sylvia undata*. Despite significant legal and policy protection, these sites are threatened by air pollution, urban development, inappropriate management and recreational disturbance and, as a result, the heathland specialist birds are under considerable risk and require effective monitoring.

3. Monitoring for the three target bird species is currently undertaken by a range of surveyors, who undertake site transect surveys during the breeding season to map bird registrations, with the aim of determining population levels. While this approach is an established standard survey technique, it is open to recognised errors and bias, such as observer bias, the availability of sufficiently skilled/experienced surveyors, and the infrequent and short-term nature of the site visits. These issues are likely to produce variations in data, particularly with nocturnal nightjar surveys, where it is hard to differentiate individuals and accurately map territories.

4. As an alternative survey approach, the use of acoustic recorders, used by themselves or in conjunction with existing methods, has great potential to reduce bias and variability in survey results and account for the effects of detectability between sites and surveys, producing more reliable and consistent population estimates. This method uses automated recording units, which can be deployed in the field for days or weeks at a time to capture wildlife sounds. It is increasingly being used as an alternative technique to survey for birds, and a variety of other taxa. Offering a number of advantages, acoustic monitoring can produce a standardised, long-duration, permanent dataset and record of species identification, which can be repeatedly analysed and subject to validation by independent reviewers. Automated recorders also potentially offer lower costs and greater data volume compared to traditional survey methods.

5. This study deployed automated recording units on sites within the Thames Basin Heaths and Wealden Heaths complexes, to record bird song/calls at dawn and dusk for a period of six days. Forty-four sampling locations were covered in May-June 2018. The audio data

recorded on the site was analysed to identify the numbers and locations of vocalizations of European nightjar, woodlark and Dartford warbler. The presence/absence of each species on each day was determined for each sampling location, and this information was then assessed to determine the quality of the survey (in terms of detection rate) and the distributions and numbers of the target bird species.

6. The data was assessed using an 'occupancy modelling' framework, to determine both occupancy and detectability estimates for each species. Occupancy is the proportion of an area, or number of sample sites, occupied by a species – and is an effective measure of population size. Detectability is the frequency with which a species is repeatedly recorded at each sampling site. It provides a measure to assess how often the species in question is missed during survey events, and false negative results are created. It can therefore be used to assess for, and correct, imperfect detection during surveys.

7. Occupancy/detectability values are given on a 0-1 index scale. The relevant scores gained in this study for the three species were: nightjar 0.682/0.733, Dartford warbler 0.382/0.258 and woodlark 0.162/0.491 – i.e. nightjar was recorded at 68% of all sampling locations, and during 73% of all daily survey events.

8. Including environmental variables within the occupancy models indicated that tree density, wetland area and heather cover were important habitat variables, particularly influencing detectability. Nightjar detectability (i.e. the quality of the survey data in correctly determining presence/absence) was much higher than that normally achieved through traditional walked transect survey methods, while the detection estimates for Dartford warbler and woodlark were similar to traditional surveys.

9. The occupancy figures showed the proportions of habitat occupied by the three species was highest for nightjar, moderate for Dartford warbler and low for woodlark. A provisional conversion of the occupancy estimates to number of pairs for each species provided the following figures (with 90% Confidence Intervals) for the areas studied:

Nightjar	51 pairs (42-59)
Dartford warbler	140 pairs (85-195)
Woodlark	8 pairs (3-13)

10. This study demonstrates the suitability of the bioacoustics approach to identify the distributions and assess the populations of target bird species on heathland study areas. The sampling protocol allowed robust and consistent data to be gathered over 10km^2 of heathland habitat, with limited fieldwork resourcing required. Occupancy and detectability estimates were produced to assess population sizes for the three target bird species, taking into account imperfect detection – a critical factor that is not currently assessed in site monitoring. If carried out on a regular basis, the bioacoustics approach could provide a valuable new method for monitoring of population levels and favourable conservation status of heathland birds on the SPAs.

1 Introduction

1.1 Thames Basin and Wealden Heaths

1.1.1 The Thames Basin Heaths Special Protection Area (SPA) and the Wealden Heaths SPA are large, internationally important, nature conservation sites in southern England. Together, they include 18 Sites of Special Scientific Interest located across the counties of Surrey, Hampshire, Berkshire and West Sussex (Figure 1), and cover a total of 12,199 ha, of which 5,702 ha is classified as lowland heath (Clark & Eyre 2012). The sites support important breeding populations of a number of birds of lowland heathland, especially the Annex 1 bird species of European importance, nightjar *Caprimulgus europaeus*, woodlark *Lullula arborea* and Dartford warbler *Sylvia undata* (Clark & Eyre 2012). The Thames Basin Heaths SPA is designated for supporting 445 pairs of Dartford warbler, 264 pairs of nightjar and 149 pairs of woodlark, while the Wealden Heaths SPA is designated for supporting 123 pairs of Dartford warbler, 103 pairs of nightjar and 105 pairs of woodlark.

1.1.2 In terms of their habitats, the SPAs are made up of open heathland on sandy or peaty soils. These habitats comprise a mix of dry heath vegetation on well-drained slopes, wet heath on low-lying shallow slopes, and waterbodies and bogs in valleys. The sites also include permanent grassland, scrub and blocks of woodland, within which are scattered areas of open heath and mire (JNCC 2001).

1.2 Heathland Bird Monitoring

1.2.1 Long-term biodiversity monitoring of designated nature conservation sites is needed to identify population trends and inform management planning efforts, especially in the context of environmental factors such as climate change and habitat loss/severance affecting wildlife populations (BirdLife International 2006; Butchart et al. 2012; Furnas & Callas, 2015). As part of the SPA designation and management process, monitoring of the relevant target species is required. For the Thames Basin and Wealden Heaths, this has been undertaken on a regular annual basis since 2003, with two or more surveys taking place between April and July to record bird contacts (mostly the presence of singing males), and to assess the number and locations of bird territories. This monitoring is critical in evaluating the status of the bird populations within the SPAs, determining the effects of management, and assessing potential impacts of development and visitor pressure on the site.

Figure 1. Location of heathland study sites

1.2.2 One example, illustrating the value of monitoring, is the population status of Dartford warbler. This species was at a population high in 2006, with more than 2,500 pairs across its range in the UK. However, severe winters in 2009-10 caused a huge crash in numbers. On the Thames Basin Heaths, nearly 1000 territories were recorded in 2004, but only 50 in 2010 (Holling 2012). Surveys for Dartford warbler in 2009, 2010 and 2011 recorded 7-14 territories at Chobham Common SSSI, none at Horsell Common SSSI, and 3-20 at Thursley, Hankley and Frensham Commons SSSI - after respective maxima in previous years of 119, 20 and 248 (Clark & Eyre 2012). Following this rapid decline, the UK population is now projected to continue to increase. This recovery is particularly important, as climate-change impacts may make more than 60% of the current European range no longer suitable, and hence the UK will become increasingly important for global conservation of this species (Hayhow et al. 2017).

1.2.3 A report by Liley & Fearnley (2014) compiled survey information for the three heathland target species relating to the nearby Dorset Heaths from 1991-2013. These indicated that nightjar and woodlark numbers had stayed stable over this period, but with some marked fluctuations between years, possibly due to weather and land management changes. As described above, the Dartford warbler numbers rose to a peak in 2000 in Dorset, but then declined sharply in 2010-11 due to harsh winter conditions.

1.2.4 The most recent data accessed for the Thames Basin Heaths as a whole is for the 2010-2016 years, and is shown in Table 1 below. The 2016 numbers for nightjar were in line with previous years, while woodlark was the lowest since the survey programme began in 2003, probably due to reducing habitat availability from vegetation growth. The trend in Dartford warbler followed the pattern discussed above, with a recovery since 2010/11, but with a possible local decline in 2016 at Horsell Common[1].

Table 1. Thames Basin Heaths SPA Annex 1 bird survey results

Survey year	2010	2011	2012	2013	2014	2015	2016
Nightjar	326	337	320	325	355	306	332
Woodlark	159	161	202	135	155	137	117
Dartford warbler	38	47	87	118	292	456	427

1.2.5 Survey data showing the number of territories for Chobham and Horsell Commons in 2018 was provided by 2Js Ecology (From : Annex 1 Birds on the Thames Basin Heaths SPA Results of the 2018 Monitoring programme for Natural England. J.Eyre & J.Clark, December 2019. 2Js Ecology) - and is summarised in Table 2 below. The area covered by these surveys was slightly wider than that included for the current study and so the numbers have been amended to correspond to the area investigated in this study .

1.2.6 Similar data for Thursley Common (D.A. Boyd, 2018) has also been provided, and is listed in Table 2.

1 https://surreyheath.moderngov.co.uk/documents/s8821/2016%20Thames%20Basin%20Heaths%20Special%20Protection%20Area%20Annex%201%20bird%20survey%20results.pdf, accessed 25-10-2018)

1.2.7 One noticeable feature of the data provided on nightjar from the standard transect surveys is the close proximity of the assumed 'territories' indicated on the survey maps. It is not clear from the data provided how the number and layout of these territories has been derived from the data gathered on site, and how the territories relate to the bird registrations gathered from visual and audible clues during the field visits. Many of the territories represented are less than 100 m distant from each other, and as a result indicate very high densities of breeding nightjars on the site. The national survey for nightjar, however, used a threshold of 350 m distance between nightjar registrations to differentiate between separate male territories (Conway *et al*. 2007). If this rule was to be applied to the 2018 data as set out in Table 2 below, then the numbers of pairs of nightjar indicated would be reduced by at least half.

Table 2. Monitoring Results 2018 (no. of territories at each site indicated)

Species	Horsell Common	Chobham Common	Thursley Common
Nightjar	7	51	39
Dartford Warbler	10	56	20
Woodlark	0	8	7

1.3 Automated Recording Units for Bird Monitoring

1.3.1 Monitoring of bird numbers and distribution can be carried out in a number of ways, although the most commonly used methods are transect or point count surveys by human observers, potentially with territory mapping. These are effective methods, but have recognised disadvantages, such as observer bias, the availability of sufficiently skilled/experienced surveyors, and the infrequent and short term nature of the site visits used during surveys. Many of these problems are recognised in the monitoring work done for the three Annex 1 heathland species.

1.3.2 In response to these recognised constraints, automated recorders are increasingly being used to survey a variety of taxa, including birds. As an alternative to traditional bird survey methods, the use of automated units, which can be positioned in the field and left for days/weeks at a time to record bird calls/song, is highly applicable (Rempel et al. 2005; Acevedo & Villanueva-Rivera 2006; Brandes, 2008; Celis-Murillo *et al*. 2009; Depraetere *et al.* 2012; Browning *et al*. 2017; Shonfield & Bayne 2017). The benefits of using automated recording, especially alongside traditional surveys, are well documented and proven (see Box 1). In particular, the ability to produce a standardised, long-duration, permanent dataset and record of species identification, which can be repeatedly analysed and subject to validation by independent reviewers, is a major advantage. In addition, the approach also potentially offers lower costs and increased standardisation compared to traditional survey methods. Due to potential benefits such as these, the use of automated recorders in scientific research has increased significantly over the last ten years (Figure 2), and some researchers have advocated the use of automated recorders instead of expert personnel conducting surveys (Haselmayer & Quinn 2000; Hobson 2002; Rempel *et al*. 2005, Brandes 2008; Zwart *et al*. 2014; Darras *et al*. 2018).

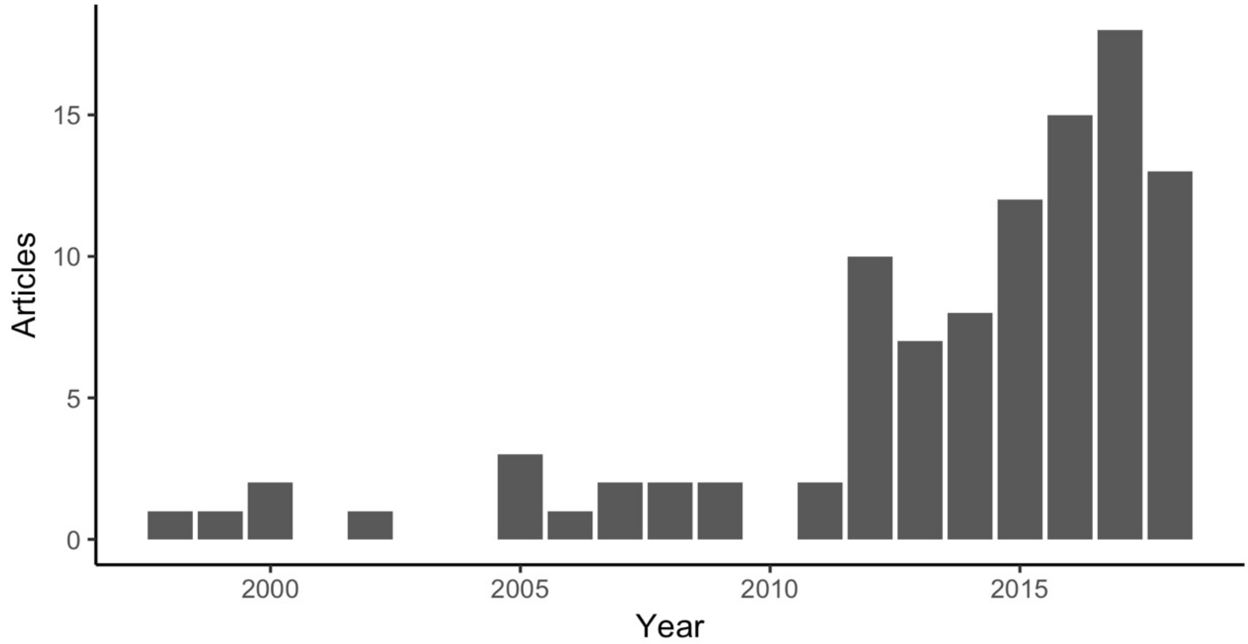


Figure 2. Numbers of original research articles published on bird bioacoustics. Search conducted on Web of Science database in September 2018 using the following search term: (bird* OR avian) (automated OR autonomous OR *acoustic) (recorder OR aru OR ard).

1.3.3 There are potential disadvantages to using a bioacoustic approach on its own (see Box 2) – principally the lack of visual cues that would be used by a human surveyor in the field, especially for birds that may not vocalize regularly. In addition, this approach does not lend itself to recording bird locations and preparing the territory maps often used in bird surveys in the UK.

1.3.4 The bioacoustics approach using automated recorders is a parallel to point counts, rather than transect surveys, and can be used to develop a bird species list for a discrete area of habitat. In addition, the method is excellent for single-species approaches, and has been utilized in research and conservation projects for species such as corncrake, bittern, capercaillie and nightjar. A study on the latter species (Zwart *et al*. 2014) found that bioacoustic recorders offered substantial improvements over human surveyors when deployed throughout the night and "detected nightjars during 19 of 22 survey periods, while surveyors detected nightjars on only six of these occasions". Similarly, Darras *et al*. (2017) concluded that sound recording provides a more powerful and promising tool to monitor birds in a standardized, verifiable, and exhaustive way than manual point counts. In particular, the higher detection probability of rare species, such as those being studied here at the heathland SPAs, is a notable advantage for the current study.

Box 1. Advantages of the bioacoustic approach

1. Ability to repeatedly listen to and re-analyse data
2. Greater standardisation in data collection, by unit programming
3. H&S – avoids night-time work, less frequent visits required to remote areas
4. Less disturbance to surveyed birds
5. Less reliance on availability of expert surveyors, deployments can be done by local staff
6. Long-duration data capture - ability to generate large consistent datasets
7. Permanent raw data record created of the survey event and opportunity to share this
8. Quality assurance, i.e. independent verification of ID by multiple reviewers is possible
9. Reduced subjectivity and observer bias compared to human surveyors varying skill levels
10. Reliable, indisputable biological survey data in the form of recordings can also avoid legal challenges and disputes that can delay projects.
11. Constant all day/every day recording possible
12. Simultaneous recording across sites, or between different locations, eliminating time-of-day and time-of-season temporal differences between samples
13. Less frequent visits required – e.g. to remote or inaccessible areas
14. Better for recording birds that only vocalise infrequently –can be used to cover longer survey periods
15. Bioacoustics approaches are cheaper as the data/effort ratio is very high

Box 2. Disadvantages of the bioacoustic approach

1. Greater post-processing time required to analyse recordings.
2. No visual recording (problems with vocally cryptic species)
3. Loss of data if units fail (and no warning of this until data downloaded)
4. Less observer bias – but equipment bias still possible
5. Automated classification systems require more development
6. Not possible to cover a wide spatial area so easily
7. Capital and maintenance costs of hardware and software
8. ARUs are mostly comparable to human observers in terms of species richness, but in some cases, they detect fewer species and at shorter distances.
9. Data storage requirements of recordings
10. Potential sampling trade-offs in spatial vs. temporal coverage.

1.4 Occupancy Modelling

1.4.1 The recent advances in bioacoustics and other methods, have allowed improved sampling of animal populations in the wild. Alongside these technological gains, there has been a dramatic increase in the development and application of occupancy models that take into account species detectability (MacKenzie *et al*. 2002; MacKenzie & Nichols 2004; MacKenzie *et al*. 2006; Welsh *et al*. 2013; Furnas & McGrann 2018). These models can be used with bioacoustic methods to assess the population status of threatened species, because they can estimate the probability that a species occurs in a sample location even though it was not detected (Shonfield & Bayne 2017). The ability to factor this in to assessments allows improved estimates of occupancy/populations and greater understanding of ecological patterns such as species/habitat relationships (MacKenzie *et al*., 2006) (see Appendix).

1.4.2 Occupancy is defined as the proportion of an area, patches or number of sample sites occupied by a species (MacKenzie *et al*. 2006) – or the probability that a species is present at a particular site. For example, if surveys took place at ten locations within a woodland and a species was recorded at three of these, then occupancy would be 0.3. This simple measure of presence/absence is, however, termed 'naive occupancy', as it does not take account of detectability, i.e. the chance of missing species that are actually present, but not recorded during the survey. For the majority of survey methods, there is always the potential that a species will be missed when it was actually present within the studied location. In such cases, the resulting occupancy (presence/absence) of a species at a site may not be correct, due to detectability issues. The error that arises from this eventuality needs to be accounted for in data analysis, as detection issues will cause an under-estimate of occupancy, and hence population size, as individuals are missed.

1.4.3 Unlike the naive occupancy estimates described above, true occupancy estimation takes the detectability issue into account, by comparing the results of repeated presence/absence surveys to allow for the estimation of, and correction for, imperfect detection. For example, if the same survey is undertaken at a site on three separate occasions, and a species is recorded during all visits, then detectability is high. But if the species is only recorded on one of the three surveys (despite actually being present throughout), then detectability is low.

1.4.4 Alongside the ability to take detectability into account, a significant benefit of an occupancy modelling approach is that it relies only on presence/absence data, rather than metrics of abundance such as counts of individuals. This normally requires substantially less effort to determine, but still provides a surrogate for abundance that can be used for investigating population dynamics or spatial variation in a population. As presence/absence at a site is more easily definable than numbers, it also requires less interpretation in the field/lab and counteracts the potential for inter-observer or inter-survey error (MacKenzie *et al.* 2006; Rovero & Zimmermann 2016). In addition, when using acoustic recording units, the number of samples and sites can be increased without additional field time, improving the accuracy of occupancy estimates and/or allowing greater spatial/temporal coverage (MacKenzie *et al.* 2002; MacKenzie & Nichols 2004). Data collection with automated recorders can also be synchronized across sites to occur at matching times, thereby reducing temporal variability in detections (Brandes 2008).

1.4.5 For rare species in particular, it can be very difficult to estimate abundance during surveys, whereas estimation of occupancy may still be possible. Occupancy and abundance will be linked in most populations, and at small spatial scales and with territorial species, occupancy and population size are, in effect, the same. As a result, the combination of acoustic monitoring and occupancy modelling, as used in this study, provides an effective and efficient way to generate meaningful ecological data that can be applied to understand and map the distribution of rare, elusive and threatened species (Campos-Cerqueira & Aide 2016).

1.4.6 Fine-grained, highly-detailed data can be gained from acoustic recorders, but occupancy modelling, using simple assessments of presence/absence across a range of sampling locations, is resource efficient and well suited for monitoring programs especially across large spatial extents (MacKenzie & Nichols 2004). Occupancy models have been successfully used to assess bird assemblages across large geographical areas (Furnas & Callas 2015), to improve distribution data of threatened bird species (Campos-Cerqueira & Aide 2016; Crates *et al.* 2017; Stiffler *et al.* 2018), to monitor population dynamics of owls (Olson *et al*. 2005) and to detect dates of peak vocal activity in a range of birds (Furnas & McGrann 2018).

1.4.7 Despite the utility of combining bioacoustic techniques and occupancy models, both are relatively recent methodological developments and only a few studies have combined these techniques (Yates & Muzika 2006; Furnas & Callas 2015; Kalan *et al*. 2015; Campos-Cerqueira & Aide 2016; Stiffler *et al*. 2018). This study, therefore, provides a very useful additional case-study to further develop the evidence base for this approach.

1.5 Objectives

1.5.1 This document reports on a trial project undertaken in spring/summer 2018, which aimed to investigate the potential of automated bioacoustic recording for monitoring populations of the three target bird species of interest on the Thames Basin Heaths and Wealden Heaths SPAs. The project entailed placing a number of recorders in the field during the bird breeding season to record the presence/absence of birds within habitat patches.

1.5.2 Establishing presence/absence of the target species with the acoustic data would allow occupancy, and hence a population index, to be calculated for the surveyed sites within the SPAs, providing numerical data that could be compared between years, and potentially correlated against other population survey methods.

1.5.3 Based upon the results of the trial project, we considered the overall utility of automated recorders, used by themselves or in conjunction with other survey methods, for application in species monitoring on internationally important designated nature conservation sites.

2 Methods

2.1 Sample Sites

2.1.1 This study has been carried out on three component SSSI sites of the SPAs, for which access could be readily arranged through Natural England (Figure 1). These are:

- Chobham Common - Thames Basin Heaths SPA (518.5 ha)
- Horsell Common - Thames Basin Heaths SPA (151.6 ha)
- Thursley Common - Wealden Heaths SPA (321.8 ha)

2.2 Field Recording - Temporal Coverage

2.2.1 Previous bird bioacoustic studies have used a wide range of recording protocols, and no standardised survey recommendation yet exists in the UK (Abrahams 2018).. However, many studies have utilised an interval time-sampling method, where short sections of recording are undertaken, with longer breaks in between (Depraetere *et al*. 2012; Oppel *et al*. 2014; Furnas & Callas 2015; La & Nudds 2016). This approach helps to sample across a broad survey time 'window' without generating excessive amounts of data, which is then difficult to store and analyse.

2.2.2 For this project, the following protocol was employed: a 1 minute sample was taken every ten minutes during the recording period (i.e. one minute on, nine minutes off), with units programmed to record from two hours before sunrise, until three hours after and then from one hour before sunset until two hours after. Each day therefore comprised a 5 hour dawn survey, and 3 hours at dusk, with recording covering 10% of the time within these periods.

2.2.3 Each deployment at a single sample site covered a period of six days of recording, to provide multiple day samples that would allow detection probability and true occupancy to be calculated. With this program, each site therefore had a minimum of 288 minutes of recording (10% coverage x 8 hours x 6 days). In some cases, more recording was undertaken, but this has been excluded from analysis.

2.2.4 Three sessions of field recording took place in May-June 2018, each with 16 recorders set out at different locations during each session - and hence covering 48 sample sites. Despite differences in date, these are all treated as individual site samples, within a single season. The assumption has therefore been made that bird distribution, population size and density did not change over the course of the three survey events.

2.3 Field recording - Spatial Coverage

2.3.1 For spatial coverage of a survey area, the aim should be to establish a number of sampling locations that are spread widely across the area to encompass the range of the habitats and species of interest. Recorders can potentially be located on site using a regular or stratified grid-based sampling system, with recorder locations being randomly assigned to a selection of nodes within the grid.

2.3.2 In this study, a 250m grid spacing between recorder locations was used to provide 16 potential sampling points/km^2. This is a sufficient distance for most recordings to be independent of other each other, dense enough to provide a good level of survey data, and also likely to be relevant to the territory sizes of the nightjar and Dartford warbler, and a high sampling density for woodlark (see Appendix). To allow the selection of actual sampling locations, GIS was used to place a regular 250m continuous grid over the SPA study area, and then cut to the study site boundaries. This provided a grid comprising 166 potential sample nodes within the three designated sites. From this list of nodes, a random sample of 48 nodes was generated, stratified to the relative area of each site, to provide 9 nodes at Horsell Common, 15 at Thursley Common, and 24 at Chobham Common (Figure 3). These were randomly assigned to one of three survey events, so that 3 nodes at Horsell Common, 5 at Thursley Common, and 8 at Chobham Common would be sampled at each survey. Each sampling site was allocated a unique numerical identification code, followed with a suffix of C for Chobham, H for Horsell and T for Thursley, e.g. 80C, 430H and 401T (see Figure 3).

2.3.3 All deployment and collection of recording units was undertaken by Baker Consultants staff, as was the analysis and reporting.

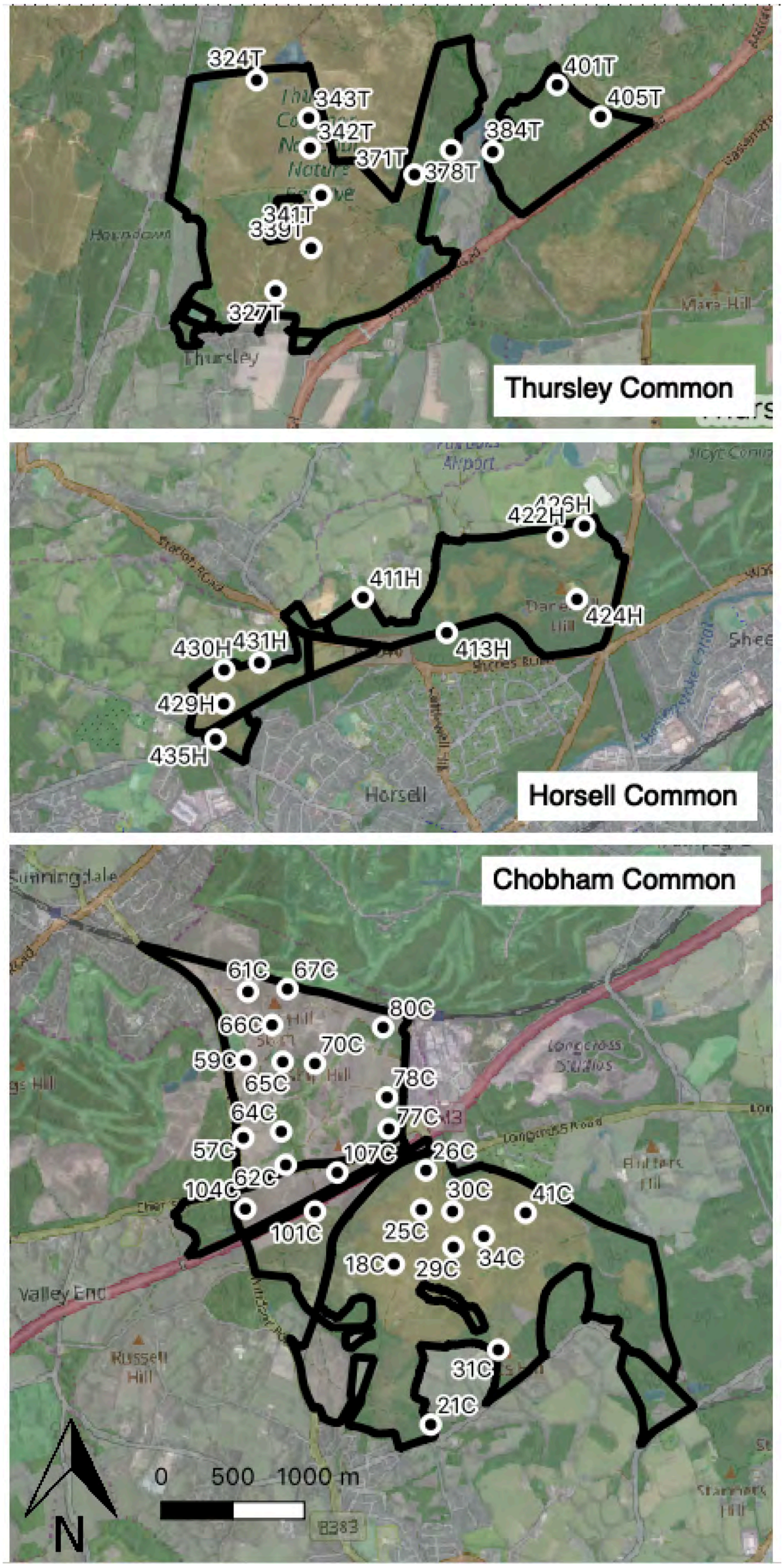


Figure 3. Sample site locations

2.4 Call Analysis

Data sets

2.4.1 The field recording covered three sessions in May-June 2018, spanning 6 days at each sample site. In each session, 16 recorders were deployed across the three SPA areas and recorded for the following periods:

- Session 1 - 26 May to 31 May
- Session 2 - 5 June to 10 June
- Session 3 - 16 June to 21 June

2.4.2 These deployments produced acoustic data for each sampling site consisting of 288 one-minute .wav audio files (=1.67GB per site). These sound files can be played back on computer and displayed as a 'spectrogram' graph using appropriate software (Figure 4).

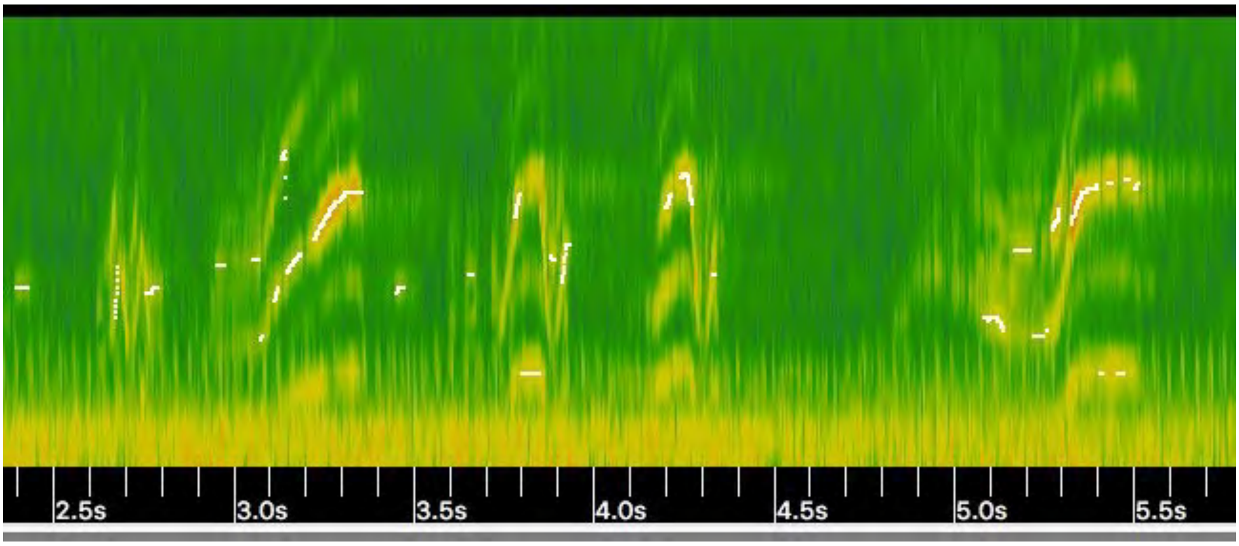


Figure 4. Spectrogram of nightjar and lapwing calls

2.4.3 The audio recordings taken from the field were analysed using a semi-automated system to identify target species vocalisations in the recordings. Kaleidoscope Pro 4.3.2 software (Wildlife Acoustics 2017) was employed, using its 'cluster analysis' method. This process analyses the time and frequency characteristics of the audio files using Hidden Markov Models to search for repeated phrases in the recordings (e.g. the song of a particular bird species). It then groups these phrases into a number of clusters based on their similarity. The Kaleidoscope software settings employed were as listed below:

- Split to 60sec max duration
- Signal of Interest 1500-7000Hz, 2-20s, 1s max inter-syllable
- Scan and cluster:
  - 1.0 max distance
  - FFT Window 5.33ms
  - Max states 12
  - 0.5 max distance to centre
  - 500 max clusters

2.4.4 This analysis scanned the data recordings to identify bird vocalisations (or other sounds), grouped them into clusters of similar sound characteristics and saved the results. The output from the software was a spreadsheet of call data, with one row per vocalisation (hyperlinked to the associated .wav file), each row providing information on time and frequency parameters, and including metadata information on the recorder, date and time of day.

2.4.5 The analysis process allowed the vocalisations to be rapidly sorted into different types, and the songs/calls of the three target species to be rapidly identified. This was done by visual identification of spectrograms and by listening to playback of the recordings. The spectrograms of each species are distinctive, based upon their frequency range (nightjar 1-3 kHz, woodlark 2.5-6 kHz, Dartford warbler 2-7 kHz) and the shape/pattern of their vocalisations.

2.4.6 All vocalisations detected by the software were reviewed and manually identified, and this information was then compiled into daily presence/absence data for each of the three species at each sampling site. This summarised data was then used to generate detectability and occupancy estimates for each species.

2.5 Environmental data sources and methods

2.5.1 To accompany the acoustic data, habitat information for the SPA sites was obtained from a combination of satellite and terrestrial mapping sources, and analysed using the QGIS Geographical Information System. These comprised Copernicus satellite data, Land Cover Map 2015 and Ordnance Survey mapping, as detailed below. Data relating to each acoustic sampling site was generated from these data sources and incorporated as habitat covariates into occupancy models.

2.5.2 Ordnance Survey base mapping in raster format was accessed from EDINA under an Educational Use Licence, with the following citations:

- 1:25 000 Scale Colour Raster [TIFF geospatial data], Scale 1:25000, Tiles: su96, Updated: 17 May 2018, Ordnance Survey (GB), Using: EDINA Digimap Ordnance Survey Service, https://digimap.edina.ac.uk, Downloaded: 2018-09-14 22:54:06.95
- 1:25 000 Scale Colour Raster [TIFF geospatial data], Scale 1:25000, Tiles: su95,su96,tq05,tq06, Updated: 17 May 2018, Ordnance Survey (GB), Using: EDINA Digimap Ordnance Survey Service, https://digimap.edina.ac.uk, Downloaded: 2018-09-14 22:54:06.95
- 1:25 000 Scale Colour Raster [TIFF geospatial data], Scale 1:25000, Tiles: su83,su84,su93,su94, Updated: 17 May 2018, Ordnance Survey (GB), Using: EDINA Digimap Ordnance Survey Service, https://digimap.edina.ac.uk, Downloaded: 2018-09-14 22:54:06.95

2.5.3 Ordnance Survey OpenMap-Local vector data was also downloaded and used to provide road network mapping (https://www.ordnancesurvey.co.uk/business-and-government/products/os-open-map-local.html). This was used to calculate the distance from each sampling location to the nearest road using the QGIS 'Distance to Nearest Hub' function (QGIS Development Team 2018).

2.5.4 To provide broad habitat information, Land Cover Map 2015 (LCM2015) vector data was accessed from the Centre for Ecology and Hydrology (Rowland et al. 2017). LCM2015 is derived from satellite images and digital cartography and provides land cover information for the entire UK. The vector mapping is made up of polygons, each representing a parcel of land and with attributes describing land cover. The land cover classification used is based on UK Biodiversity Action Plan Broad Habitats and includes 21 different broad habitat types (see Figure 5). The area of each of these habitat types within a 100m radius of each sampling site was measured using the 'Intersect' and 'Field Calculator' functions in QGIS.

2.5.5 Processed satellite imagery was accessed from Copernicus Pan-European High Resolution Layers (https://land.copernicus.eu/pan-european/high-resolution-layers). These High Resolution Layers (HRL) provide information on specific land cover characteristics, produced from Sentinel satellite imagery. The data used was the 20m resolution raster map, with imagery dating from 2015 (Figure 6). There are separate layers of information for tree cover density (TCD), water and wetness (WAW), and imperviousness degree (IMD). This data was analysed using QGIS and the properties of each layer were measured within a 100m radius of each sampling site, using the 'Zonal Statistics' plugin.

- The imperviousness degree HRL captures the spatial distribution of artificially sealed (i.e urbanized/road) areas, including the level of sealing of the soil per area unit. The level of sealed soil is given as an imperviousness degree score of 1-100% for each pixel.
- The tree cover density (forest) HRL provides the level of tree cover in a range from 0-100% for each pixel, defined by the vertical projection of tree crowns to the earth's surface.
- The water and wetness HRL shows the occurrence of water and wet surfaces over the period from 2009 to 2015. This layer is based on intra-annual dynamics, allowing the definition of the following hydrological classes for each pixel: (1) permanent water, (2) temporary water, (3) permanent wetness and (4) temporary wetness. This shows the occurrence of water/wetness, independently of the actual vegetation cover and is, therefore, not limited to a specific land cover class.

2.5.6 A search was made for weather data for the area, and MIDAS surface weather data was accessed from ceda.ac.uk for the weather station at Wisley, Surrey (Ref. src_id 719, WGS 84 51.3108, -0.47634, Grid ref: TQ 062579, Code DCNN 5237). However, records for the required time period were sparse and the only data available was for 'derived 24hr sun duration'.

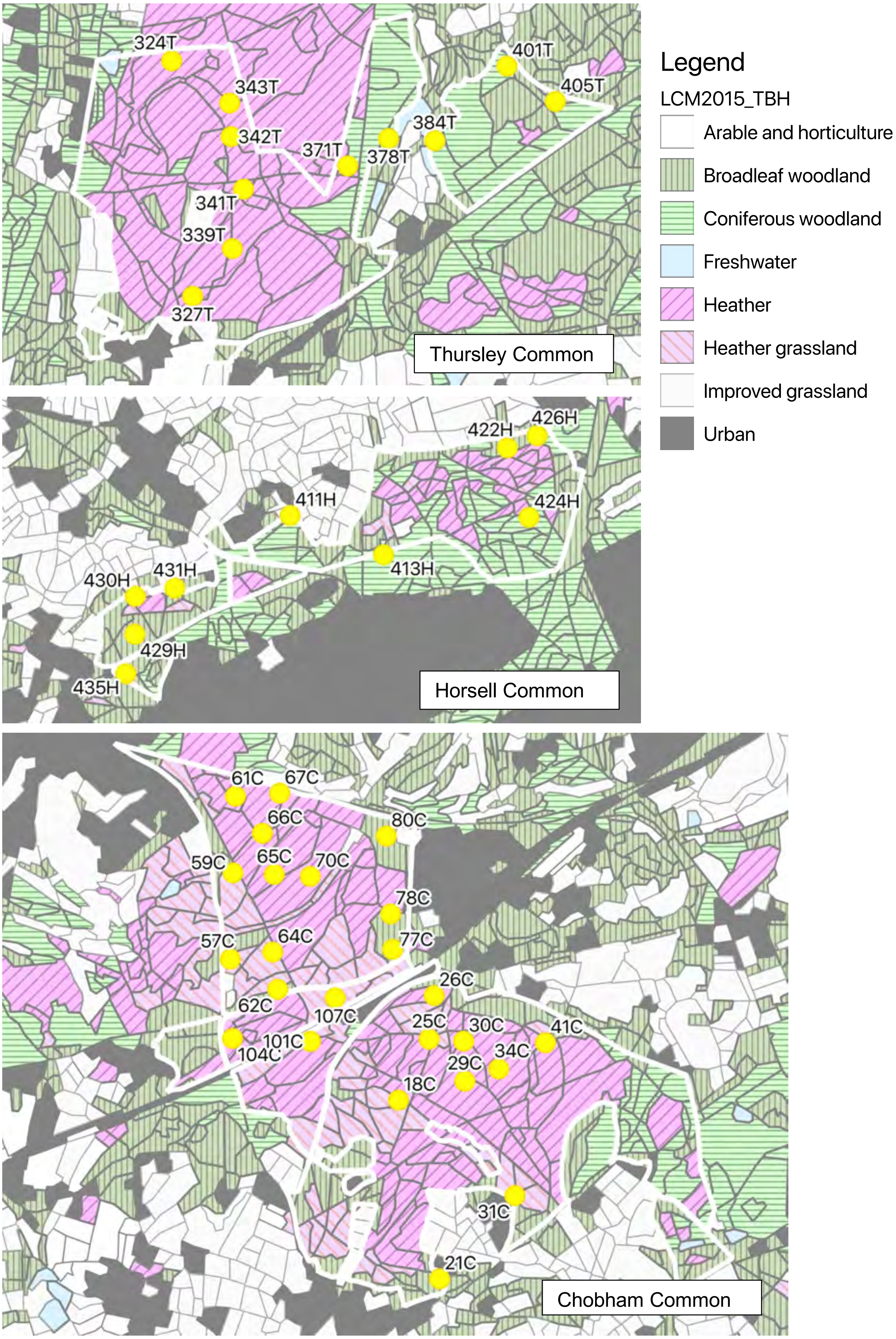


Figure 5. Land Cover Map 2015 data

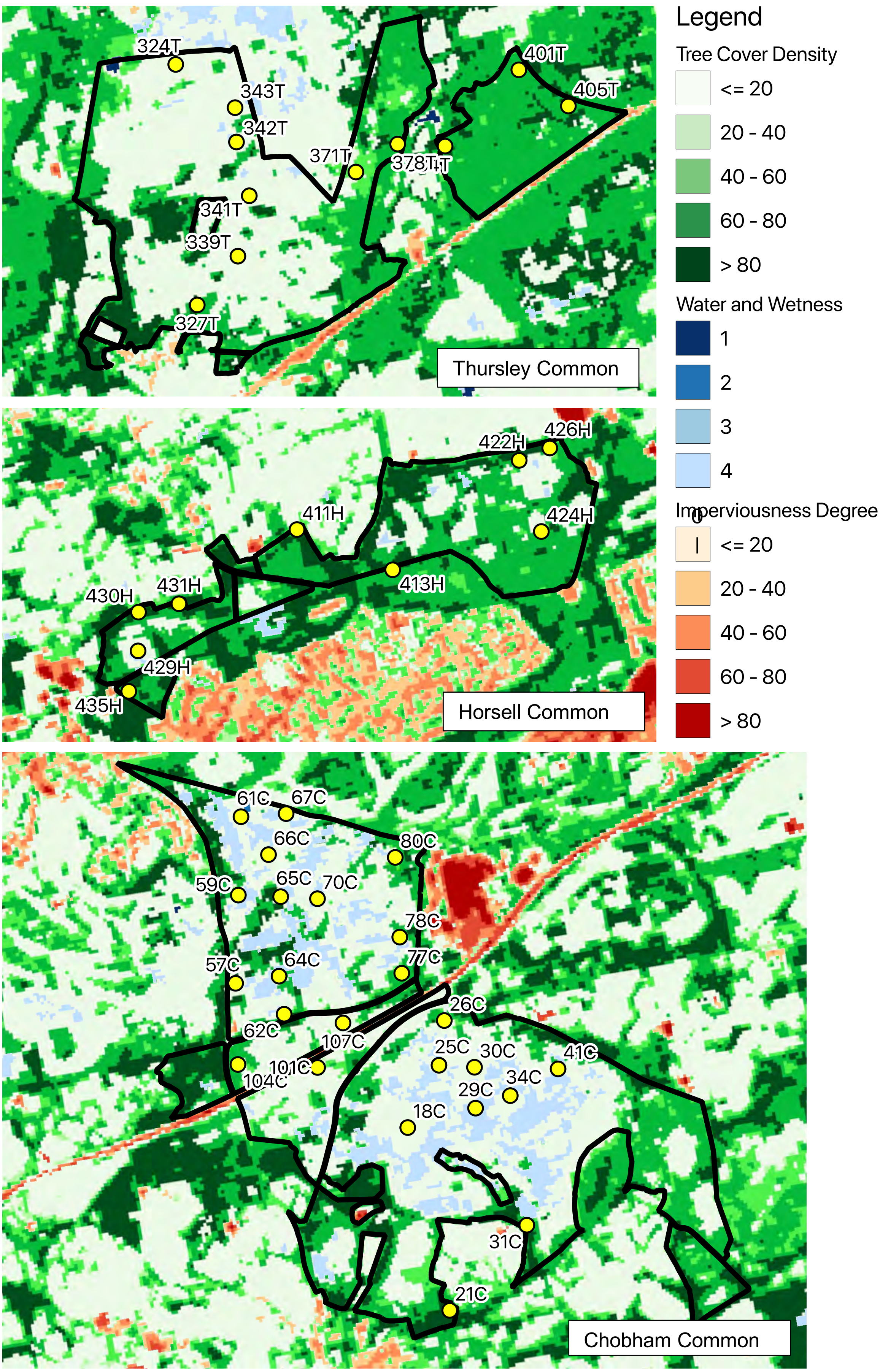


Figure 6. Copernicus Satellite Data

2.6 Occupancy modelling

2.6.1 The daily presence/absence data gained from the acoustic surveys was used in combination with the environmental variables to determine detection and occupancy probabilities for the three target species and each sampling location. The occupancy of each of the three target species was modelled separately using a single-species, single-season modelling approach with observation and habitat covariates (MacKenzie *et al*. 2002; MacKenzie *et al*. 2006; Furnas & Callas 2015; Stiffler *et al*. 2018), using established protocols with the 'Unmarked' package in R (Fiske & Chandler 2011; R Core Team 2013; RStudio Team 2015).

2.6.2 The acoustic data was summarised to day-level resolution of presence/absence, to produce a detection history at each sampling site comprising six replicate surveys. The naive occupancy for each species was checked and confirmed to be >0.1, so that detection histories were not too sparse to fit single-species models.

2.6.3 A null model, with no detection or occupancy covariates, was first created for each species, followed by a range of other models with combinations of covariates to investigate their effects on detectability and occupancy. Each model was inspected to check estimates, standard errors and convergence. The null model was created separately for each species to estimate its average occurrence frequency, providing information about the relative population levels of the three species.

2.6.4 After null models had been created, additional models were evaluated using a range of spatial and temporal covariates. The models included data on the areas of different habitat types within 100m of the sampling location (from LCM2015 and Copernicus data), and distance to the nearest road. It was anticipated that detection probability might change over the course of the survey period (Campos-Cerqueira & Aide 2016; Furnas & McGrann 2018) due to seasonal and weather reasons, and Julian day of survey and 24-hour sun duration were used to represent this information. A range of observation and site covariates were first fitted to the detection parameter to produce the best model for each species. Once these had been established, they were kept constant while site covariates were applied to the occupancy parameter. Using this process, all covariates were fitted, developing models based on the review of model outputs of estimates, standard errors and probabilities to select the most parsimonious complete models. Model fit was assessed using Akaike's Information Criterion (AIC), ranking and comparing models based on AIC relative differences between the top ranked model and each other model (delta AIC) and AIC weights. Models with delta AIC <2 were considered to be equally supported [Burnham & Anderson 2002), and were combined by applying model averaging (Barton 2018) to estimate occupancy and detection for each species.

2.6.5 The analysis procedure employed allowed the occupancy probability of each sampling site to be estimated, taking into account imperfect detection, following a standard maximum-likelihood hierarchical approach (MacKenzie *et al*. 2002).

3 Results

3.1 Environmental Parameters

3.1.1 The three study sites vary in relation to their habitat components, and the acoustic recorders were placed in habitats that varied from open heath to mature forest. Data from LCM2015 and Copernicus imagery (see Figures 5 and 6) indicates that Thursley Common comprises two halves - the western part of the site is dominated by heather, with the eastern part being coniferous and broadleaved woodland. Chobham Common is a mosaic of heather and heather grassland, with broadleaved and coniferous woodland around its fringes. This site has a much larger extent of 'Water and Wetness' (WAWsum) than the two other sites indicated by the Copernicus data. Horsell Common is the smallest of the sites and is mostly coniferous and broadleaved woodland, with patches of heather at its eastern end. Not shown on the habitat mapping for this site is an area of cleared ground at the western end of the Common (around location 429H), where trees have recently been removed to allow heather to regenerate.

3.1.2 The GIS analysis of the habitat at each sampling site allowed quantitative values to be generated from the LCM2015, Copernicus and Ordnance Survey data. The means and ranges of the measured environmental parameters are listed in Table 3 below.

Table 3. Measured habitat parameters (n=44 sampling sites)

Habitat variable	Mean value	Range	Units
TCDsum	2570	0-6209	Sum of % per pixel
WAWsum	36.8	0-252	Sum of 1-4 index per pixel
Distance to Road (HubDist)	351	29-961	Metres
Broadleaf	7260	0-31315	Sum of pixels
Coniferous	3537	0-24905	Sum of pixels
Heather	14459	0-31318	Sum of pixels
Heather grassland	4204	0-31060	Sum of pixels

3.1.3 Initial data exploration (Figures 7 and 8) identified the following habitat variables as suitable for use in analysis: TCDsum and WAWsum from Copernicus, Distance to Road (HubDist) from Ordnance Survey, heather and heather grassland from LCM2015. Broadleaf woodland and coniferous were excluded as these duplicated the TCDsum habitat type, and the LCM2015 data were more zero-inflated than the Copernicus data. The following variables were also excluded as these were generally not present within 100m of the recorder locations, and hence measured values were too sparse within the dataset: IMD from Copernicus, and freshwater, improved, surburban and urban from LCM2015.

3.1.4 The five habitat variables generated from the GIS datasets, all numerical, were prepared for analysis and standardised to the same scale for inclusion within the occupancy models.

3.1.5 The heatmap at Figure 9 provides a classification of the sampling sites based on their constituent LCM2015 habitats. This shows that sites are divided fairly clearly into the four main habitat types, with little overlap - i.e. the sampling sites are generally surrounded by a single habitat type, rather than occurring within a mosaic. This is especially the case for heather habitat, although there are some sites with heather and heather grassland, and others with broadleaf woodland and heather.

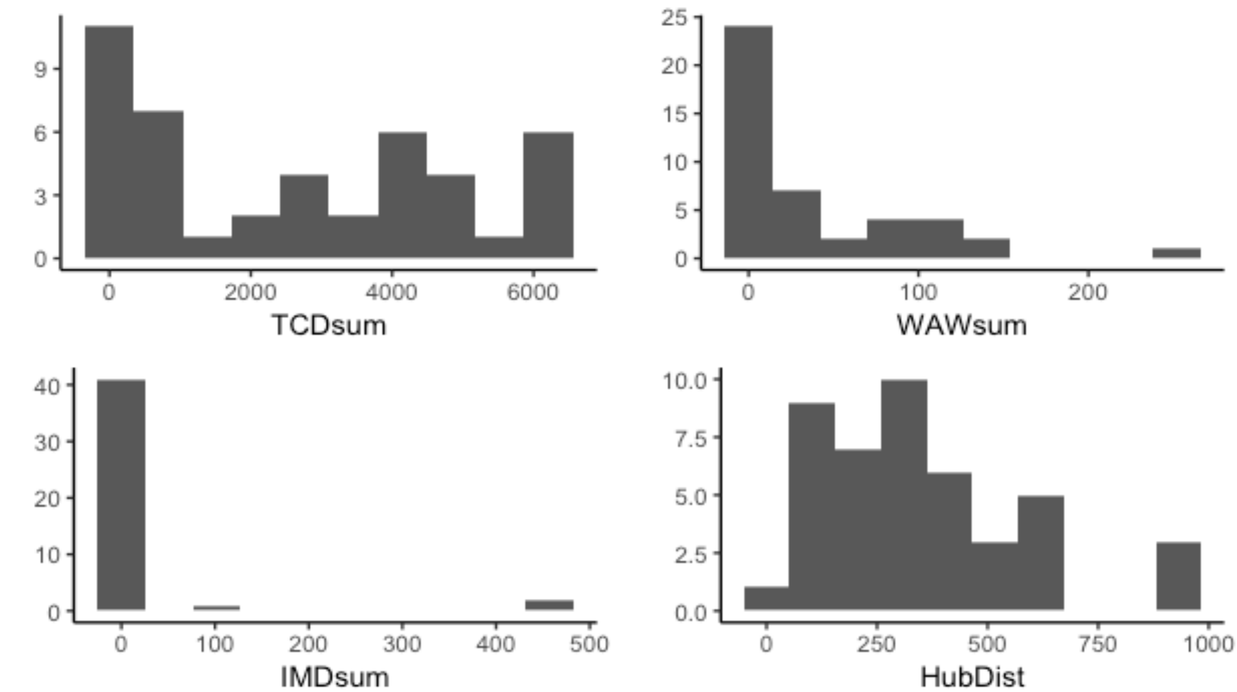


Figure 7. Histogram of Copernicus and Ordnance Survey data values

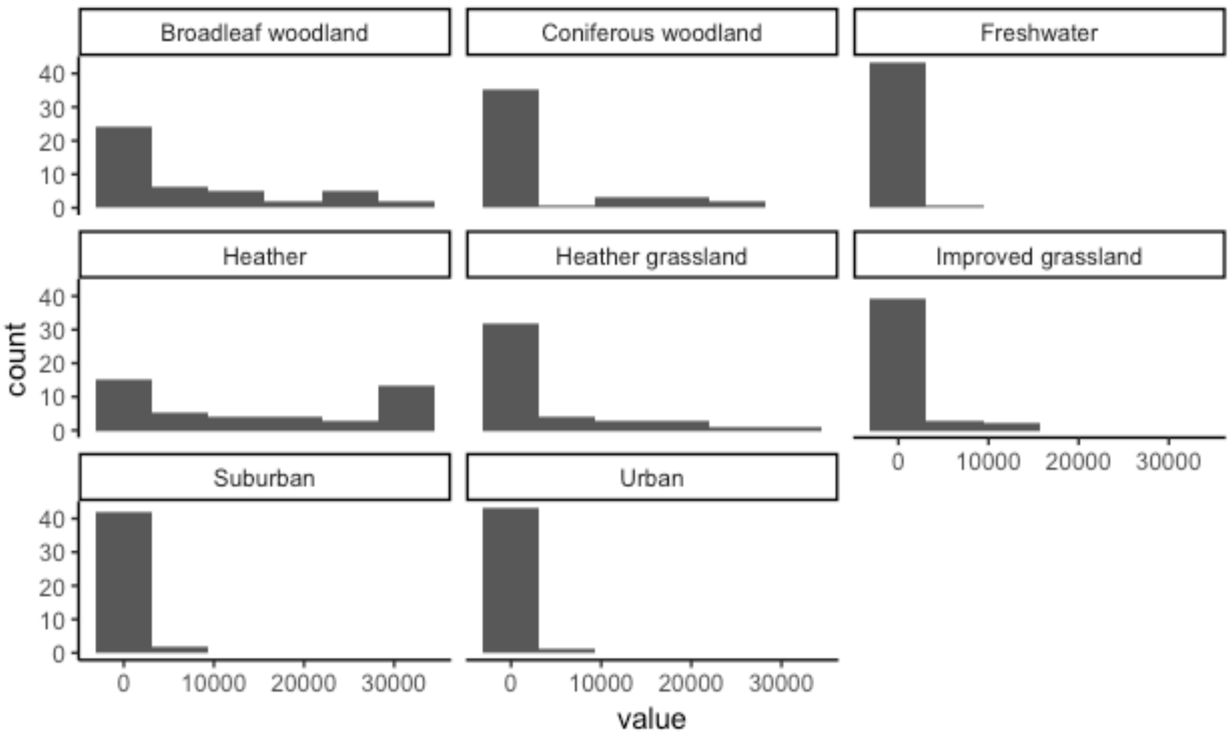


Figure 8. Histograms of LCM2015 data values

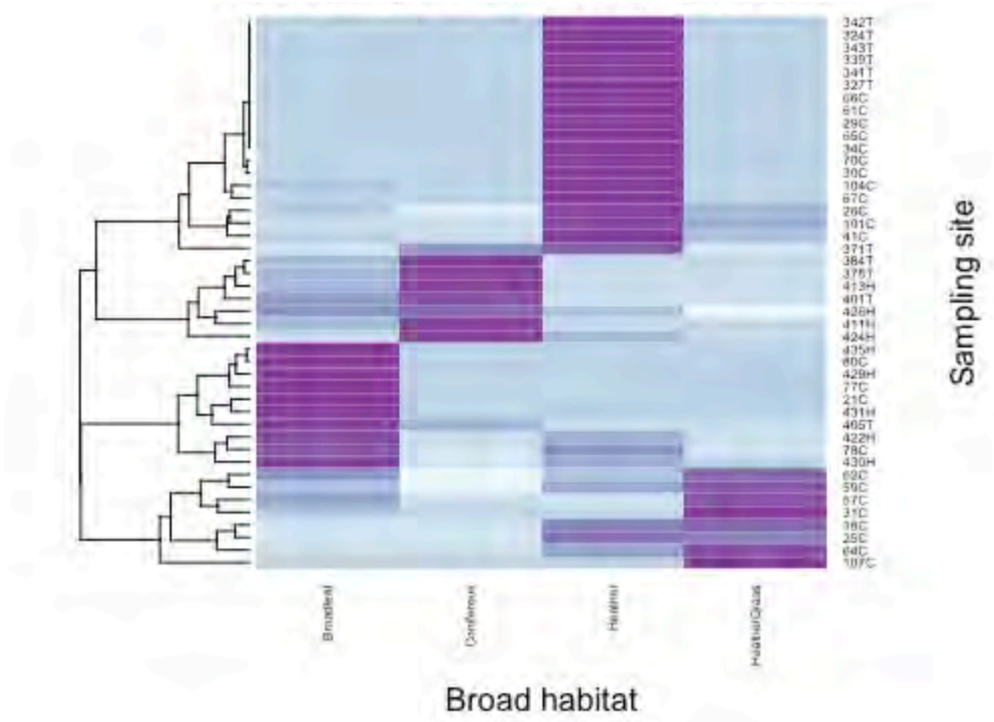


Figure 9. Heatmap of LCM2015 broad habitats at each site

3.2 Audio Data Collection

3.2.1 A total of 48 sampling sites were used, across the three SPA sites. However, one recorder failed to record evening sessions repeatedly (at three sampling sites), and another suffered battery failure on one occasion. These failures were all at Thursley Common (sites 315T, 319T, 332T, 391T) and the sites were removed from the dataset, leaving 44 sampling locations. Due to all of these being at Thursley, this site is under-represented in terms of its area, compared to Chobham and Horsell Commons. An additional result is that the second recording session of 5-10 June is missing a recorder in comparison to the first and third sessions, and so the total number of vocalisations in the dataset will be lower for this survey period. With the 44 sampling locations, and 288 minutes of recording at each, a total of 211.2 hours of audio was recorded and analysed during the study.

3.3 Clustered Audio Segments

3.3.1 Kaleidoscope clustering of the complete audio dataset produced 28,775 .wav files, an average of 109 files per site/day. Each processed .wav file represented an audio segment with bird vocalisations and other sounds, potentially up to 60 seconds in length. The mean duration of the segments was 6 seconds (range 2-21 sec), and together, the clustered segments comprised 48 hours of audio - 23% of the recorded dataset. The files were grouped into 55 clusters by the software.

3.4 Temporal Patterns in Activity

3.4.1 The total number of phrases recorded per day across all sampling sites varied from 1,974 on 30 May to 1,145 on 17 June. The daily number of phrases was relatively even between recording sessions 1 and 2, but declined for session 3 in mid-June. This pattern was matched somewhat by the daily numbers of target species vocalizations (Figure 10). Nightjar and Dartford warbler vocalizations were recorded throughout all three recording sessions, with Dartford warbler most common in mid-June. Woodlark was mostly confined to the early June session only - although this is likely to be related to presence at the sites being sampled at that time, rather than any reason to do with seasonal timing.

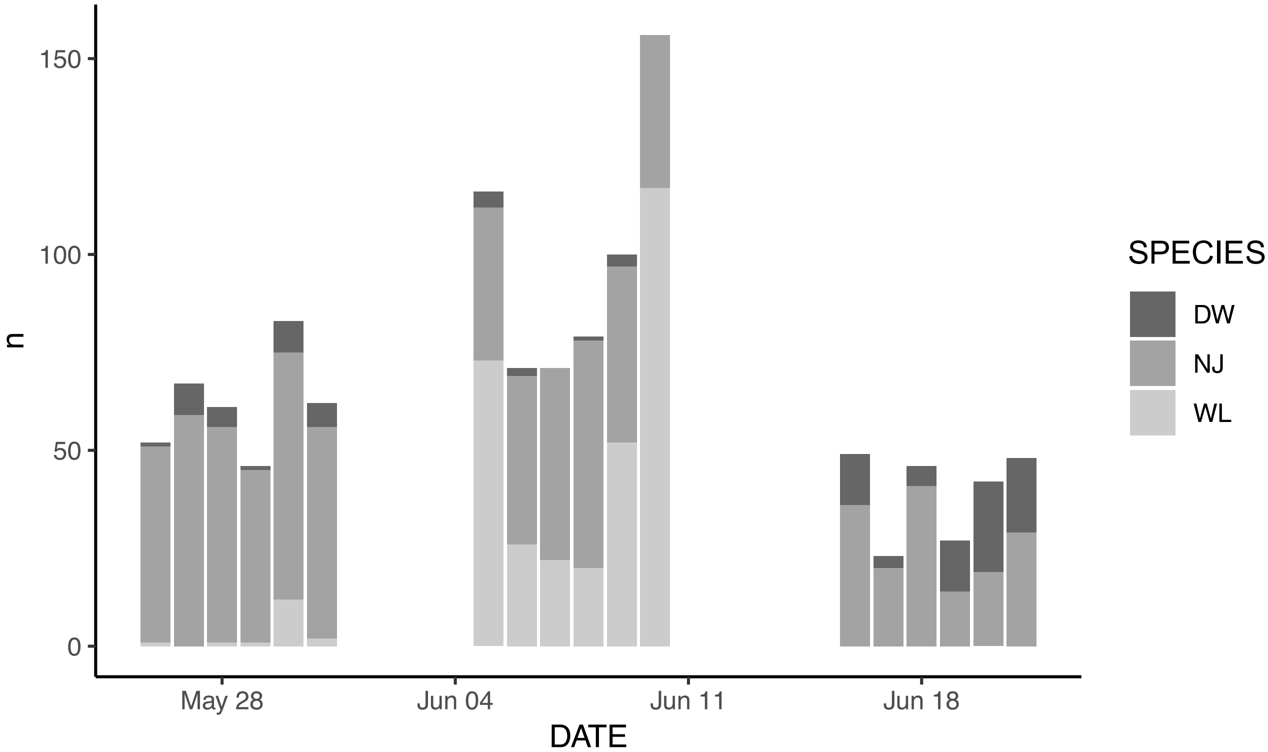


Figure 10. Number of target species recorded per day

3.4.2 The number of target species vocalisations recorded at each site is illustrated in Figure 11. This shows that the most active sites were 61C and 70C for nightjar, 29C and 25C for woodlark and 339T and 343T for Dartford warbler. Significant numbers of calls were not recorded for any species at the Horsell Common sites.

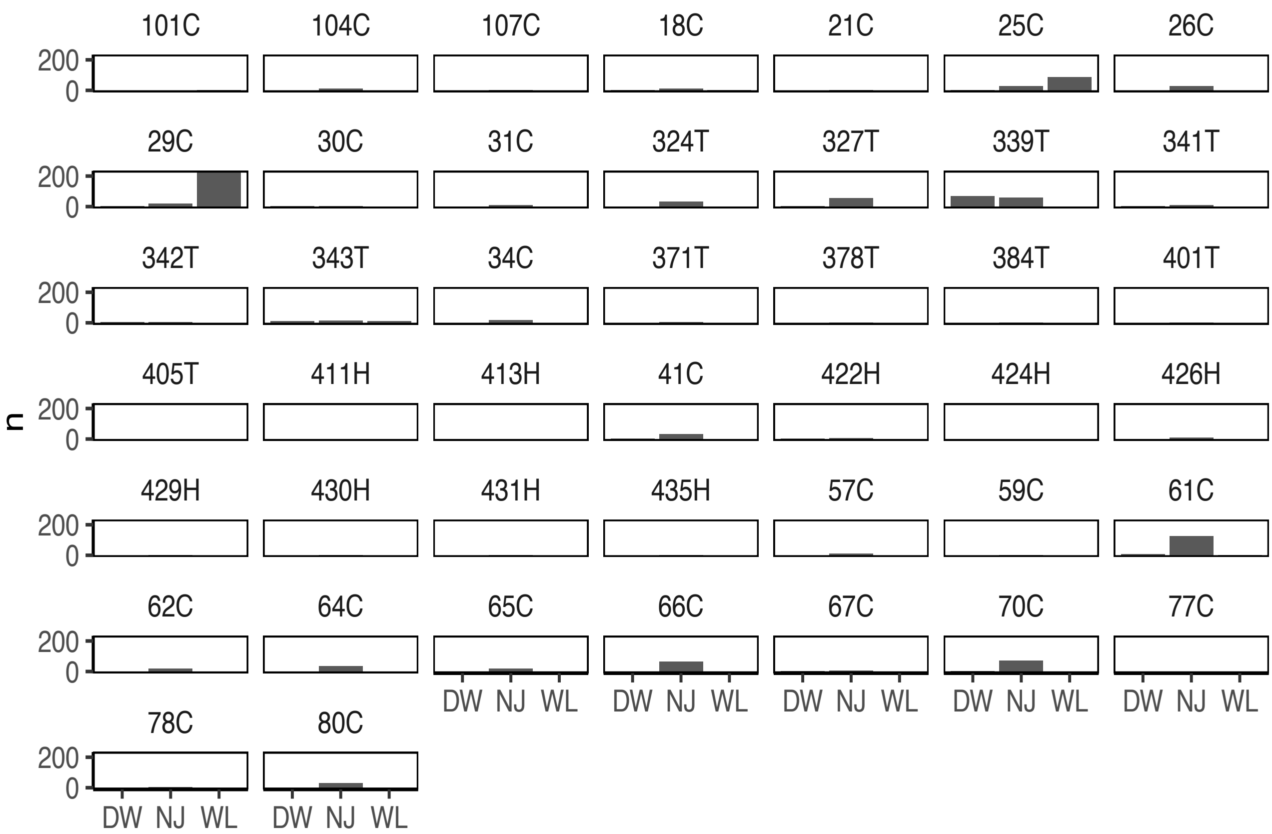


Figure 11. Number of target species vocalizations recorded per site

3.5 Naive Occupancy

3.5.1 Manual review of all the clustered recording segments identified the three target species in the dataset, with 757 segments having vocalisations of nightjar across 30 sites, 327 segments of woodlark at 7 sites, and 115 segments of Dartford warbler at 14 sites. This gave a total of 1,199 vocal segments recorded for the three target species. Nightjar and Dartford warbler were recorded at all three SPA sites, but woodlark was only recorded at Chobham and Thursley Commons, and not at Horsell Common. The low number of Dartford warbler vocalisations recorded is perhaps notable, as this is supposedly the most numerous species of the three, based on the SPA citations. However, it was recorded as present in one-third of the sites, so was reasonably widespread despite the low total number of calls identified.

3.5.2 The most vocally active sites were 61C and 70C (north Chobham) for nightjar, 29C and 25C (south Chobham) for woodlark and 339T and 343T (central Thursley) for Dartford warbler - see locations at Figures 6 & 7. Significant numbers of calls were not recorded for any species at the Horsell Common sites.

3.5.3 Although no exhaustive efforts were made to identify non-target species, the list of birds in Table 4 was recorded during the manual review of the acoustic data. Common species such as wren, blackbird, chiffchaff and willow warbler, in fact, made up the majority of the vocalizations recorded in the field - as might be expected given their dominance in many lowland bird assemblages.

3.5.4 All deployment and collection of recording units was undertaken by Baker Consultants staff, as was the analysis and reporting.

Table 4. List of non-target species recorded

Common name	Scientific Name	Common name	Scientific Name
Blackbird	*Turdus merula*	Little grebe	*Tachybaptus ruficollis*
Blue tit	*Cyanistes caeruleus*	Long-tailed tit	*Aegithalos caudatus*
Canada goose	*Branta canadensis*	Magpie	*Pica pica*
Carrion crow	*Corvus corone*	Mallard	*Anas platyrhynchos*
Chaffinch	*Fringilla coelebs*	Meadow pipit	*Anthus pratensis*
Chiffchaff	*Phylloscopus collybita*	Pheasant	*Phasianus colchicus*
Cuckoo	*Cuculus canorus*	Robin	*Erithacus rubecula*
Curlew	*Numenius arquata*	Skylark	*Alauda arvensis*
Great tit	*Parus major*	Song thrush	*Turdus philomelos*
Green woodpecker	*Picus viridis*	Treecreeper	*Certhia familiaris*
Jay	*Garrulus glandarius*	Willow warbler	*Phylloscopus trochilus*
Lapwing	*Vanellus vanellus*	Wren	*Troglodytes troglodytes*

3.6 Occupancy Modelling

3.6.1 Naive occupancy was calculated for each species, based on the presence of the species across all 44 sample sites in the study. The naive occupancy values, equal to the proportion of sites with positive detections, were 0.68 for nightjar, 0.32 for Dartford warbler and 0.16 for woodlark (Table 5).

3.6.2 The naive occupancy for each of the three SPA sites is also shown below in Table 5. This indicates that occupancy of the target species is relatively high at Chobham and Thursley Commons, but is low at Horsell. Nightjar occupancy is particularly high at Chobham (0.83), while Dartford warbler is highest at Thursley (0.45). Woodlark occupancy is practically the same at Chobham and Thursley (~0.2). These figures give a simple measure of the relative population size at each site.

Table 5. Naive occupancy rates at each SPA site (No. of sites and occupancy)

Species	All Sites (n=44)	Chobham (n=24)	Horsell (n=9)	Thursley (n=11)
Dartford warbler	14 (0.32)	8 (0.33)	1 (0.11)	5 (0.45)
Nightjar	30 (0.68)	20 (0.83)	3 (0.33)	7 (0.64)
Woodlark	7 (0.16)	5 (0.21)	0 (0.00)	2 (0.18)

3.6.3 Figure 12 shows the number of days on which each species was recorded as present at each site. Nightjars, if present, were commonly recorded on all or most of the survey days, indicating a high detectability. Dartford warbler was only recorded on 1-2 days out of the six that were recorded during each session. Hence detectability for this species was low.

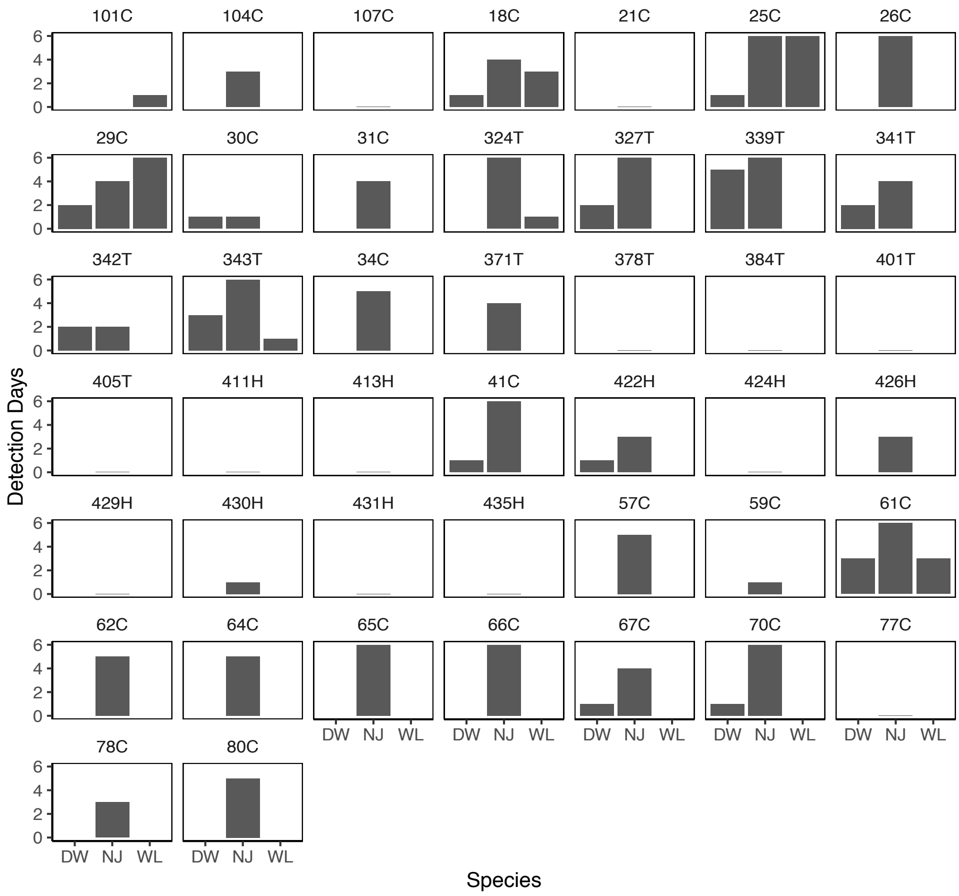


Figure 12. Number of detection days per species at each site

3.6.4 To refine the naïve occupancy values, intercept-only models were calculated, taking into account the detectability of each species. It is presumed that these are likely to be more accurate than the naive frequencies shown at Table 5, and the outputs for each species are shown in Table 6 below. Dartford warbler had a relatively low detectability value indicating that it was only recorded sporadically at each sampling location, and was likely to have been missed in some locations. Hence, for this species, the improved model provided an increased occupancy value against the naive estimate – giving an occupancy of 0.38 instead of 0.32.

Table 6. Intercept-only occupancy and detectability estimates

Species	Occupancy estimate (SE)	Detectability estimate (SE)
Dartford warbler	0.382 (0.0914)	0.258 (0.0571)
Nightjar	0.682 (0.0702)	0.733 (0.0331)
Woodlark	0.162 (0.0562)	0.491 (0.0807)

3.7 Spatial Patterns in Activity

3.7.1 A large variation was found in the total number of audio segments produced at each sampling location (including all sounds) - with site 80C having 1134 segments, and site 101C having only 114 segments. The site average was 654 segments across the six-day recording period. The data shows that all but one of the top ten sites in terms of vocal activity were in woodland placements, rather than more open heathland habitat - presumably due to the more numerous, diverse and vocal bird assemblage in these locations.

3.7.2 Maps showing the distribution of each of the three species at the sampling sites are provided below in Figures 13-18. The maps show the daily number of vocalisations recorded at each site, with data overlaid for each day of the survey. These show that the recorded nightjar vocal activity was greatest in the north of Chobham Common and through the central part of Thursley Common (Figure 13-14). Activity in Horsell Common was limited to small numbers of calls on the northern boundary of the site. Dartford warbler calls were generally in low numbers, scattered through Chobham and Horsell, although there were two sampling sites with more regular high numbers in the central part of Thursley. Woodlark calls were only recorded in significant numbers at two sampling sites, both in the centre of Chobham Common South.

Figure 13. Nightjar vocalizations per day – Chobham and Horsell

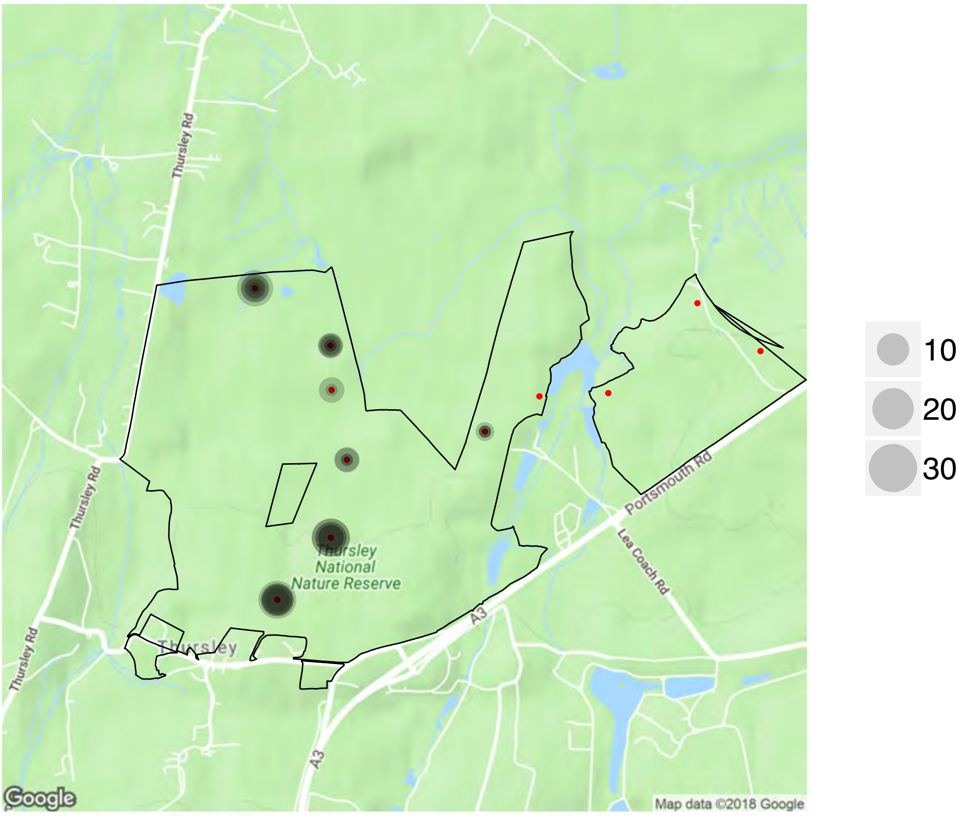


Figure 14. Nightjar vocalizations per day – Thursley

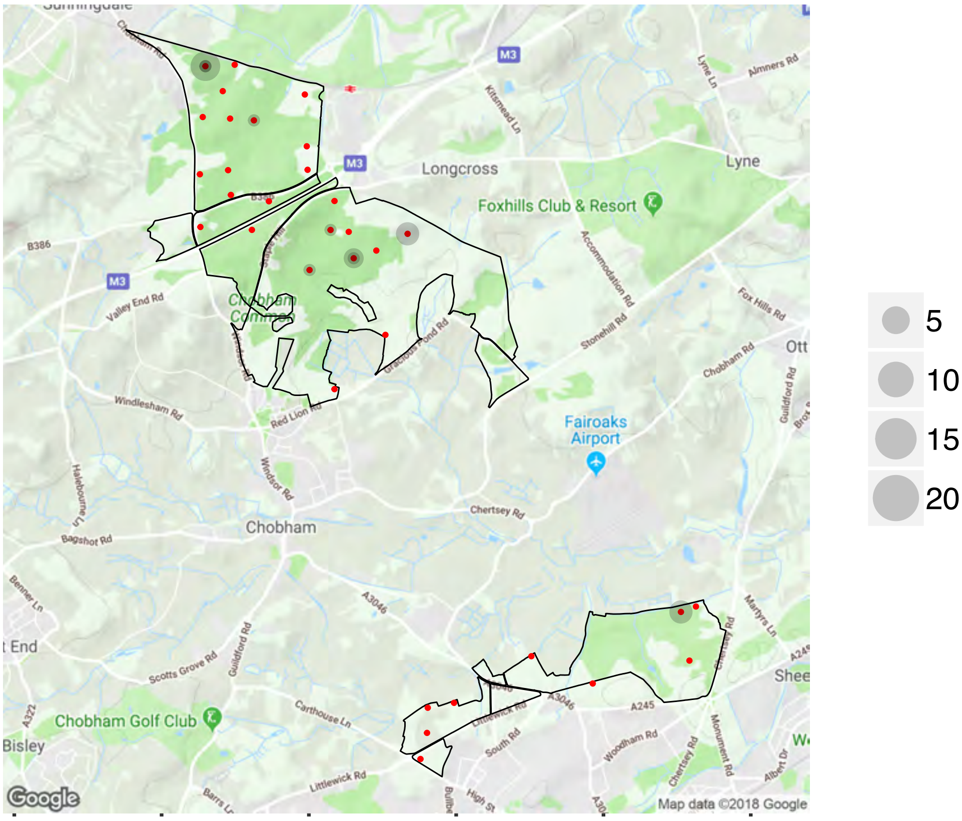


Figure 15. Dartford warbler vocalizations per day – Chobham and Horsell

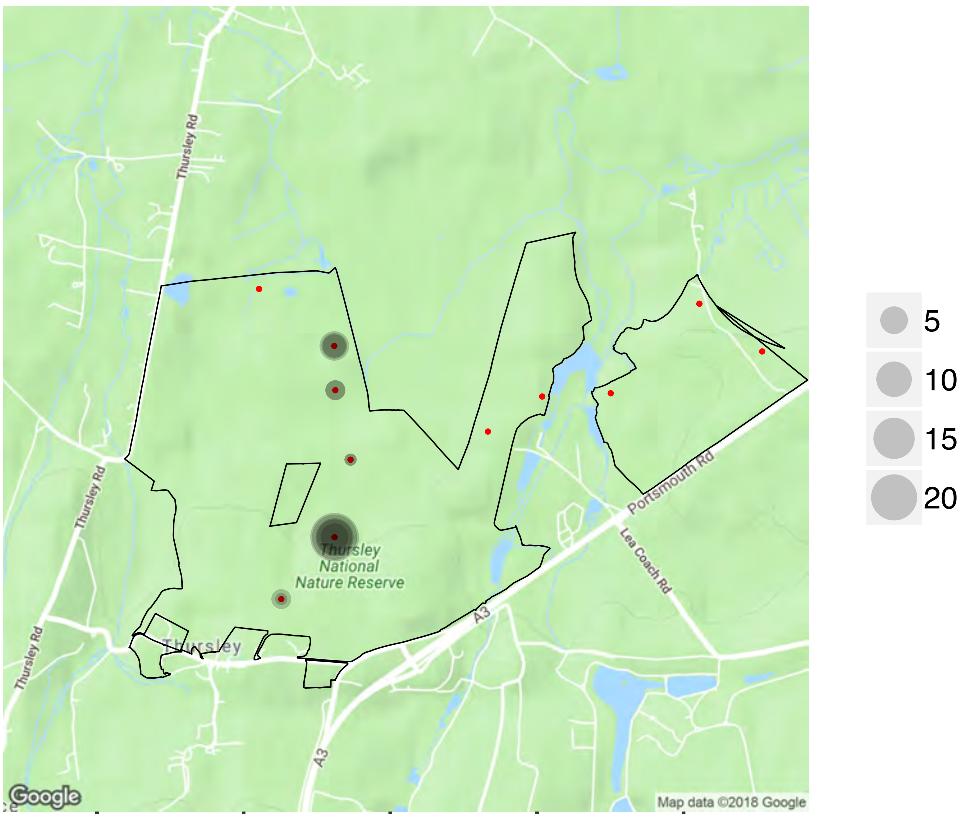


Figure 16. Dartford warbler vocalizations per day – Thursley

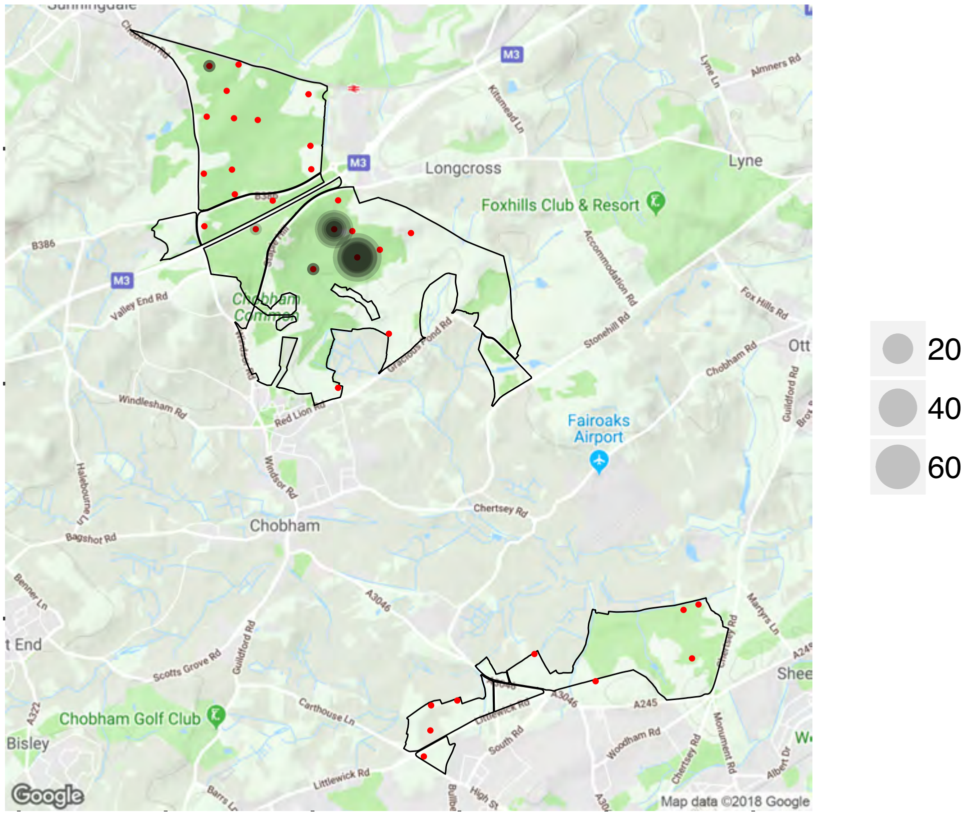


Figure 17. Woodlark vocalizations per day – Chobham and Horsell

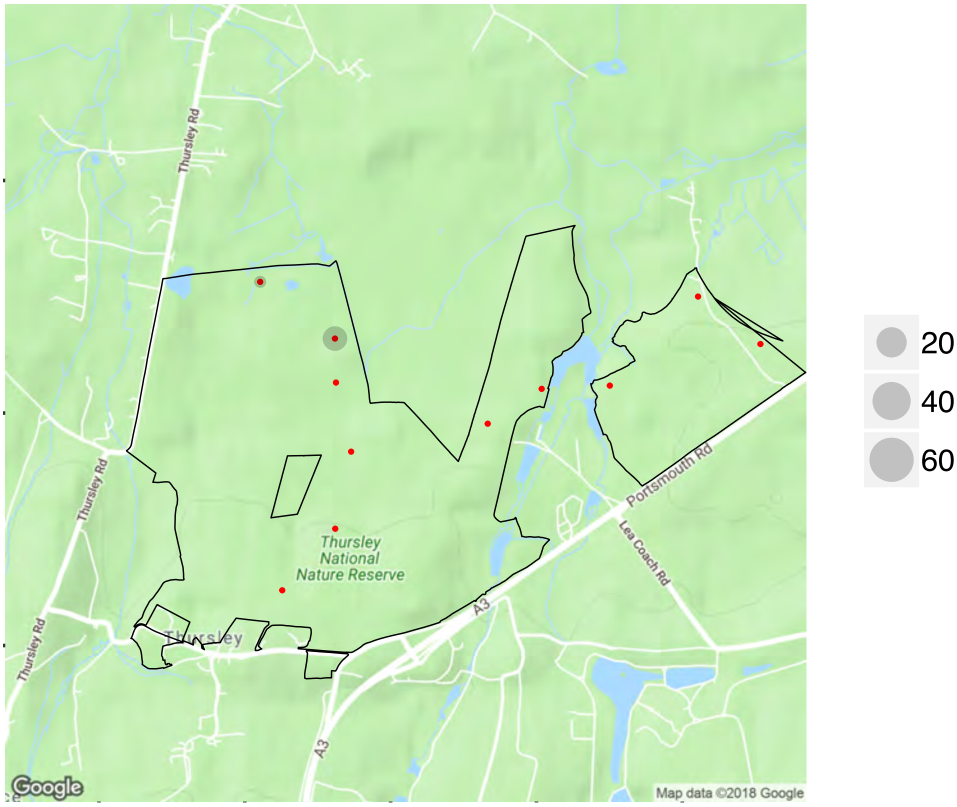


Figure 18. Woodlark vocalizations per day – Chobham and Horsell

3.8 True occupancy with covariates

3.8.1 Five habitat variables were used for occupancy modelling, together with Julian date and daily sun duration. A null model was created first, without any of these covariates. In accordance with standard practice, covariates were then applied to the detectability parameter, before being applied to the occupancy parameter. This was done for each of the three species separately.

3.8.2 The occupancy modelling, initially using detectability covariates only, produced varying models for each species. The best performing were NJmdet3, DWmdet5, and WLmdet3 (Tables 7-9). These models showed that Julian date, TCDsum and WAWsum were among the most valuable covariates for modelling detectability of the three species, with HubDist also important for Dartford warbler, and heather and heather grassland useful for woodlark. Sun duration did not feature in any of the best fit models.

3.8.3 Tables 7-9 below compare the null model (m0) to the other models created, based on minimizing the Akaike Information Criteria (AIC) score (Burnham and Anderson 2002). All models with AIC differences of less than 2 from the lowest score have a substantial level of support, and are the most useful in defining the important covariates.

Table 7. Model selection list for Nightjar detectability

No.	Model	Formula	AIC	delta	AICwt
4	NJmdet3	~JULIAN + TCDsum + WAWsum ~ 1	259.62	0.000	0.78
5	NJmdet4	~JULIAN + TCDsum ~ 1	263.02	3.408	0.14
2	NJmdet1	~JULIAN + SUNDUR + TCDsum + WAWsum + HubDist + Heather + HeatherGrass ~ 1	265.47	5.856	0.04
1	**NJm0**	~1 ~ 1	267.79	8.170	0.01
6	NJmdet5	~JULIAN + WAWsum ~ 1	268.10	8.488	0.01
3	NJmdet2	~JULIAN + SUNDUR ~ 1	268.50	8.883	0.01

Table 8. Model selection list for Dartford warbler detectability

No.	Model	Formula	AIC	delta	AICwt
6	DWmdet5	~TCDsum + HubDist ~ 1	157.10	0.000	0.436
5	DWmdet4	~TCDsum + WAWsum + HubDist ~ 1	158.39	1.289	0.228
1	**DWm0**	~1 ~ 1	158.99	1.887	0.169
4	DWmdet3	~JULIAN + TCDsum + HubDist + Heather ~ 1	160.21	3.104	0.092
2	DWmdet1	~JULIAN + SUNDUR + TCDsum + WAWsum + HubDist + Heather + HeatherGrass ~ 1	161.84	4.738	0.040
3	DWmdet2	~JULIAN + SUNDUR ~ 1	162.32	5.213	0.032

Table 9. Model selection list for Woodlark detectability

No.	Model	Formula	AIC	delta	AICwt
4	WLmdet3	~JULIAN + WAWsum + HubDist + Heather + HeatherGrass ~ 1	75.28	0.000	0.866
2	WLmdet1	~JULIAN + SUNDUR + TCDsum + WAWsum + HubDist + Heather + HeatherGrass ~ 1	79.03	3.754	0.132
6	WLmdet5	~JULIAN + HubDist + Heather ~ 1	89.72	14.441	0.001
3	WLmdet2	~JULIAN + SUNDUR ~ 1	94.74	19.460	0.001
5	WLmdet4	~TCDsum + HubDist ~ 1	98.97	23.694	0.001
1	**WLm0**	~1 ~ 1	100.5	25.265	0.001

3.8.4 The best-fit models for detectability were taken forwards for the application of occupancy covariates. The inclusion of occupancy covariates allowed further development of the models as set out in Table 10, incorporating data on habitat preferences, and all improving on the null models for the three species. Table 10 indicates that there were relatively high weights (AICwt) for WLmocc2 (0.59) and NJMdet3 (0.53), but low for DWmdet5 (0.36) indicating that there are a number of potentially useful models for Dartford warbler. To make use of the variety of models generated, a model-averaging approach was taken.

Table 10. Model selection list - with detectability and occupancy covariates

Model	Formula	AIC	delta	AICwt
NJmdet3	~JULIAN + TCDsum + WAWsum ~ 1	259.62052	0.000000	0.5276062
NJmocc3	~JULIAN + TCDsum + WAWsum ~ TCDsum	260.63638	1.015857	0.3174821
NJmocc2	~JULIAN + TCDsum + WAWsum ~ TCDsum + HubDist	262.32509	2.704568	0.1364647
NJmocc1	~JULIAN + TCDsum + WAWsum ~ TCDsum + WAWsum + HubDist + Heather + HeatherGrass	267.63928	8.018754	0.0095733
NJm0	~1 ~ 1	267.79106	8.170535	0.0088736
DWmdet5	~TCDsum + HubDist ~ 1	157.10760	0.000000	0.3642909
DWmocc3	~HubDist + TCDsum ~ HeatherGrass	158.19094	1.083339	0.2119358
DWmdet4	~TCDsum + WAWsum + HubDist ~ 1	158.39678	1.289175	0.1912087
DWm0	~1 ~ 1	158.99528	1.887677	0.1417570
DWmocc2	~HubDist + TCDsum ~ WAWsum + HeatherGrass	160.05667	2.949064	0.0833810
DWmocc1	~HubDist + TCDsum ~ TCDsum + WAWsum + HubDist + Heather + HeatherGrass	164.89340	7.785793	0.0074265
WLmocc2	~JULIAN + WAWsum + HubDist + Heather + HeatherGrass ~ WAWsum + Heather + HeatherGrass	69.30793	0.000000	0.5929403
WLmocc3	~JULIAN + WAWsum + HubDist + Heather + HeatherGrass ~ WAWsum + HeatherGrass	70.75162	1.443682	0.2880841
WLmocc1	~JULIAN + WAWsum + HubDist + Heather + HeatherGrass ~ TCDsum + WAWsum + HubDist + Heather + HeatherGrass	73.09827	3.790332	0.0891150
WLmdet3	~JULIAN + WAWsum + HubDist + Heather + HeatherGrass ~ 1	75.28505	5.977120	0.0298604
WLm0	~1 ~ 1	100.55037	31.242437	0.0000001

3.8.5 Two models for nightjar had equal support (ΔAIC <2) and so were averaged to produce covariate estimates. The averaged model included Julian date (JULIAN), Tree Cover Density (TCDsum) and Water and Wetness (WAWsum) as detectability covariates with no covariates acting on occupancy. The best fit model for nightjar (NJmdet3), with an AICwt of 53%, indicates an occupancy of 0.684 (SE 0.071) with a detectability of 0.740 (SE 0.035), varying only slightly from the null model (occupancy = 0.682, detectability = 0.733).

3.8.6 There were four favoured models for Dartford warbler, including the null model, with TCDsum, WAWsum, and distance to road (HubDist) featuring on the detectability parameter - Heather grassland being the only indicator for occupancy. The averaged model for Dartford warbler used only distance to road as a detectability covariate, with no covariates acting on occupancy. The best-fit model for Dartford warbler (DWmdet5), with an AICwt of 36%, indicates an occupancy of 0.449 (SE 0.107), with a detectability of 0.196

(SE 0.053), an increase from the null model occupancy of 0.382 (SE 0.091), but decrease in detectability from 0.258 (SE 0.057) (see Tables 11 and 12).

3.8.7 Woodlark had two favoured models, sharing Julian date, WAWsum, distance to road, Heather and Heather grassland as detectability covariates, and WAWsum, Heather and Heather grassland for occupancy covariates. The averaged model for woodlark had five significant covariates, and again, these were all on the detection parameter. Julian date, WAWsum and Heather were all positively related to detectability, while distance to road and Heather grassland were negative indicators. For woodlark, the best-fit model (WLmocc2), with an AICwt of 59%, indicated an occupancy of 0.13 (SE 0.117), lower than the null model figure of 0.162 (SE 0.056), and a detectability of 0.996 (SE 0.012), which varied substantially from the null model detectability of 0.491 (SE 0.081).

Table 11. Back-transformed occupancy estimates

Model	Occupancy Estimate	Standard Error	LinComb
NJmdet3	0.684	0.0705	0.773
NJmocc3	0.688	0.0712	0.791
DWdet5	0.449	0.107	-0.207
DWmocc3	0.47	0.116	-0.12
DWdet4	0.452	0.105	-0.193
DWm0	0.382	0.0914	-0.481
WLmocc2	0.132	0.117	-1.88
WLmocc3	0.239	0.102	-1.16

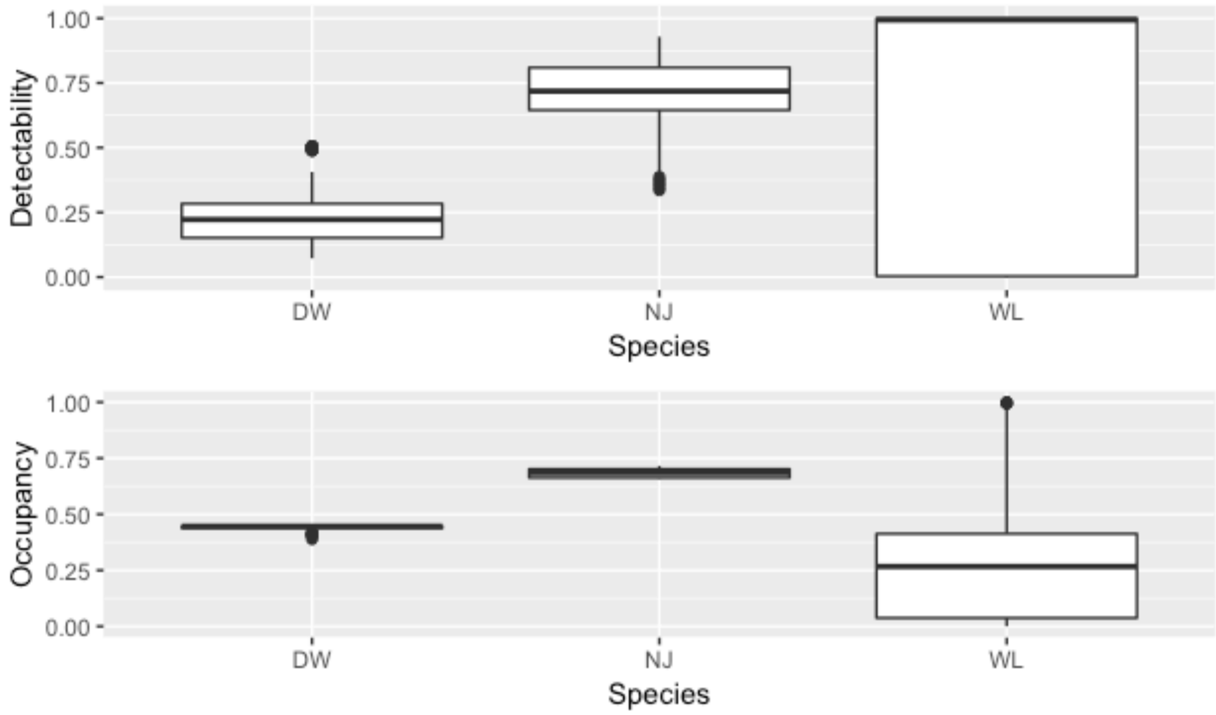
Table 12. Back-transformed detection probabilities

Model	Detection Probability	Standard Error	LinComb
NJmdet3	0.74	0.035	1.04
NJmocc3	0.74	0.0349	1.05
DWdet5	0.196	0.0528	-1.41
DWmocc3	0.188	0.0518	-1.46
DWdet4	0.19	0.0516	-1.45
DWm0	0.258	0.0571	-1.06
WLmocc2	0.996	0.0122	5.48
WLmocc3	0.996	0.0109	5.53

3.8.8 Predicted occupancy varied little between sampling sites for nightjar and Dartford warbler (Figure 19), as only single covariates were acting on these species - TCDsum and Heather grassland respectively. Woodlark occupancy predictions varied more widely due to the number of habitat covariates acting on the models for this species - including WAWsum, Heather and Heather grassland. Detectability predictions were sensible for nightjar and Dartford warbler, but highly polarised to 0-1 between sampling sites in the models for woodlark, and hence difficult to interpret.

3.8.9 Welsh *et al*. (2013) acknowledge that occupancy models may be difficult to fit in some cases, and the estimating equations can give boundary estimates which produce fitted probabilities of zero or one. For the woodlark data in this study it would appear that the small number of positive sampling sites for this species may have produced unstable estimates for occupancy and detectability, making them difficult to interpret.

Figure 19. Model-averaged predicted occupancy and detectability across all sampling sites



3.9 Population Size Assessment

3.9.1 The results presented here can be used to provide a baseline for assessing the population of the three heathland bird species studied. We assumed that occupancy is a good surrogate for abundance (MacKenzie & Nichols 2004) and that we could quantify the relative abundances of the bird species, based on the proportion of sampling sites in which they were recorded to be present. This has already been done above, giving intercept-only occupancy values of 0.68 for nightjar, 0.38 for Dartford warbler and 0.16 for woodlark, and an increased modelled occupancy values of 0.45 for Dartford warbler, when detectability is considered.

3.9.2 Given the separation distances between recorder locations in this study, it is considered reasonable to assume that each occupied sampling site represented a separate territory/pair. Using the occupancy estimates from the null models for the three species we can calculate that the areas of occupied habitat for each species, from a total 992 ha, are: nightjar 676 ha, Dartford warbler 379 ha, and woodlark 161 ha (Table 13). Combining these habitat areas with published breeding densities of 0.074-0.078 males/ha for nightjar (Berry 1979; Conway et al. 2007), 0.32-0.42 pairs/ha for Dartford warbler (Bibby & Tubbs 1975), and 0.05 pairs/ha for woodlark (Langston et al. 2007; Sitters et al. 1996), gives estimated population levels of: nightjar 51 males, Dartford warbler 140 pairs, and woodlark 8 pairs (Table 13).

Table 13. Calculated areas of occupied habitat, based on intercept-only occupancy estimates

Species	Occupancy (SE)	Occupied habitat (90% CI)	Density ha-1	Pairs (90% CI)
Nightjar	0.682 (0.0702)	676 ha (562-791)	0.075	51 (42-59)
Dartford warbler	0.382 (0.0914)	379 ha (230-528)	0.37	140 (85-195)
Woodlark	0.162 (0.0562)	161ha (69-252)	0.05	8 (3-13)

3.9.3 The same calculation can be done for each site, and these are presented below in Table 14 using the naive occupancy for each of the three sites.

Table 14. Naive occupancy rates and area of occupied habitat at each SPA site

Species	Chobham (321.8 ha)	Horsell (151.6 ha)	Thursley (321.8 ha)
Dartford warbler	0.33 (106 ha)	0.11 (17 ha)	0.45 (145 ha)
Nightjar	0.83 (267 ha)	0.33 (50 ha)	0.64 (206 ha)
Woodlark	0.21 (68 ha)	0.00 (0 ha)	0.18 (58 ha)

4 Discussion

4.1 Study Outcome

4.1.1 To our knowledge, this is the first study in Europe to combine bioacoustic survey with occupancy modelling. It is also the first in the UK to undertake a large scale survey for multiple bird species using automated recorders. We used species detection data from six repeated days of recording at 44 sampling sites to estimate the probability of occupancy for three heathland bird species. We combined this with environmental covariates to account for imperfect species detection. Our results showed that the bioacoustic approach can be used effectively for monitoring heathland bird populations through the application of occupancy models - providing quantitative data on population levels.

4.1.2 Although we included models where habitat covariates could influence occupancy, the 'best' models for each species suggested that the habitat variables were not important indicators of occupancy at the scale studied. This is possibly due to the fact that the study areas were all lowland heathland sites, generally suitable for the study species, and so the distribution of individuals was likely to relate to micro-habitat features that were not detectable at the scale of the satellite and map data applied. This corresponds to the finding of Niedballa *et al*. (2015), that both the spatial scale of habitat covariate data, and the radius sampled around survey sites, can affect the fit of occupancy models. Importantly, however, suitable models could be produced without the environmental covariates, indicating that a simple approach without this information, could be used effectively for future monitoring.

4.2 Monitoring Approaches

4.2.1 Biodiversity monitoring is central to nature conservation, allowing species status to be evaluated or assessments to be made of biological responses to environmental changes. However, monitoring practices and protocols are often not optimal in terms of resource efficiency and the value of their outcomes. Schmeller *et al*. (2012) reviewed the large number of bird monitoring practices from around Europe, covering programmes from local scale studies by experts, to national scale volunteer surveys. They concluded that the monitoring design in a majority of the programs could be improved, notably in terms of unbiased spatial coverage, sampling effort optimization, replicated sampling to account for variations in detection probability, and more efficient statistical use of the data. Only 28% of the surveys reviewed had randomized sampling, and stratified methods were applied in only 22-37% of cases. However, the largest deficit in methods was the lack of repeated sampling, making many studies prone to misinterpretation of trends due to variations in detection probability. This failing may be potentially important in a large number of studies, as even low differences in detection probability between site or years can cause significant biases and errors in conclusions.

4.2.2 Monitoring a variety of bird species, with differing behaviours, over extensive heathland sites presents significant challenges for conservation managers. In particular, a number of different surveyors are normally involved in the heathland bird surveys used for monitoring the target species – this is inevitable given the geographic and time coverage required. Inter-observer differences are therefore likely to produce variations in data, particularly with nocturnal nightjar surveys, which make it hard to differentiate individuals and accurately map territories (Liley & Fearnley 2014). Automated recorders, used by themselves or in conjunction with other methods, have great potential to reduce the logistical burden on site managers of finding and deploying skilled bird surveyors to undertake regular site monitoring, and in reducing variability in occupancy and detection probability.

4.2.3 Occupancy modelling requires that the detection process is independent at each site. On heathland, this is likely to be an issue mostly for nightjar, so sufficient distance needs to be maintained between sampling sites to avoid double-counting. Single nightjar males can hold large territories and may churr from different parts of the territory in a short period of time (Liley & Fearnley 2014). Static recording could therefore potentially over-estimate the numbers of this species, if recorder spacing was too close. Conway *et al*. (2007) used a threshold of 350 m distance between nightjar registrations to differentiate between male territories. Within the current study, the closest spacing between sampling sites was set by the 250m sampling grid - the mean nearest neighbour distances of the recorder sites were 316m (range 218-703) for Chobham, 346m (202-635) for Horsell, and 329m (210-450) for Thursley. Due to the sampling sites being spread across three survey sessions, the mean distances between actual recorders would be greater than these figures. Overall, the spacing of the sites and recorders related well to the guidance from Conway *et al*. (2007), and as a result, there can be a reasonable amount of confidence that there was no double-counting for the bird species being studied. Interestingly, it is not certain that such a precaution has been taken for the standard transect surveys undertaken on the site, with 'territories' (or registrations) less than 100m apart presented as being different pairs. This may cause an over-inflation of the derived population counts, and warrants further clarification of the current monitoring data.

4.3 Detectability

4.3.1 The probability of detecting a species during a bioacoustic survey is a function of both the probability of it vocalizing and the recorder detecting the call. The vocalisation rates of many birds vary due to age, sex, breeding status, time of day, and seasonal variation (Campos-Cerqueira & Aide 2016; La & Nudds 2016; Ehnes *et al.* 2018; Furnas & McGrann 2018). Age and sex-specific variation in vocalisation rates cannot be accounted for easily when using automated recorders, but samples were recorded over a relatively short period of time during the breeding season, minimising the potential for seasonal variation in call rates, and covering a wide timeframe every day.

4.3.2 We estimated that the probability the bioacoustic system would detect a species was 0.76 for nightjar, 0.34 for woodlark and 0.12 for Dartford warbler. Perhaps unsurprisingly, previous studies using traditional survey methods found different detectabilities for the same species, with 0.30 for nightjar, 0.47 for woodlark and 0.37 for Dartford warbler (Johnston *et al*. 2014). So, in this comparison, nightjar is much better detected by recorders, but woodlark and Dartford warbler less so.

4.3.3 Traditional bird surveying requires repeated visits to sites across the sampling season in order to assess detection probabilities – which may be low for some cryptic species.

However, automated recorders can potentially circumvent this low detectability by surveying consistently for longer periods. The normal timespan between site visits by human observers may result in violation of the assumption that detection probability remains constant across the surveys and sites. In contrast, the use of automated recorders makes meeting this assumption easier by allowing for repeated back-to-back surveys within short timespans (Stiffler *et al.* 2018). Furthermore, using automated recorders allows simultaneous surveys at multiple locations over extended periods of time. Thus, recorders can provide a broader spatial distribution and a greater number of survey replicates, both of which improve occupancy and detection estimates (MacKenzie *et al.* 2006; Stiffler *et al.* 2018).

4.3.4 The site-level detection probabilities in this study were high for nightjar, variable for Dartford warbler and polarized to 0-1 for woodlark (due to the small number of occupied sites). This potentially indicates that a higher sampling effort, in terms of sites and/or sampling days, may be required for the two latter species.

4.3.5 Our results, together with those of Johnston *et al*. (2014), show how detection probability varies significantly by species. This should be taken into account to aid decisions about study design when planning to survey birds using either automated recorders or traditional methods. Survey date, combined with habitat characteristics, explained detectability and improved model performance for the species studied here.

4.4 Occupancy

4.4.1 Our results provide a good baseline for assessing relative abundances and habitat associations of the three heathland bird species studies. To do so, we assumed that occupancy is a good surrogate for abundance (MacKenzie & Nichols 2004) and that we could quantify the relative abundances of the bird species, based on the proportion of sampling sites in which they were recorded to be present. Furthermore, we provided an example of how occupancy associations with habitat covariates differed for the three species.

4.4.2 Broken down by the areas of each site, the calculated numbers of pairs are as shown in Table 15 . The figures given in brackets are the numbers of pairs/males provided by the existing traditional survey methods on the sites (J. Eyre and D. Boyd pers.comm.). However, the numbers for nightjar provided here should be treated with caution. A figure of 51 males is provided for the Chobham Common survey area, equivalent to 0.16 males/ha, and therefore much higher than the breeding densities previously recorded for this species (quoted above). In addition, many of the ‘territories’ plotted for nightjar at Chobham are for registrations much closer to each other than the 350m minimum distance recommended by Conway *et al.*, (2007). If this threshold were applied to the 2018 survey data, then the actual number of territories for Chobham would be in the region of 15-30, not 51.

Table 15. Population estimates from occupancy models (and from 2018 transects)

Species	Horsell	Chobham	Thursley	Total, 90% CI
Nightjar	8 (7)	27 (51*)	16 (39)	51, 42-59 (97*)
Dartford warbler	21 (10)	73 (56)	45 (20)	140, 85-195 (86)
Woodlark	1 (0)	4 (8)	3 (7)	8, 3-13 (15)

4.5 Identification of calls

4.5.1 For occupancy modelling, there needs to be awareness that automated species identification from acoustic data can often include high levels of false-positive detections (Zwart *et a*l. 2014). This is important because misclassification error can lead to substantial errors in occupancy estimates (MacKenzie *et al*. 2006; Royle & Link 2006). The issue can be addressed by complete manual identification, but this will be highly time-consuming. Alternative approaches are to (i) implement more complex occupancy models that take into account both false-negative and false-positive errors; or (ii) use automated species identification models to reduce the size of the data set, and then validate all positive identifications to eliminate any false-positive detections; as used in this study (Campos-Cerqueira & Aide 2016).

4.5.2 Analysis time to determine species presence may be substantially reduced by using automated classification methods, especially for studies interested in a few target species (e.g., Bardeli *et al.* 2010; Venier *et al.* 2012; Potamitis *et al*. 2014). Automated classifiers have been developed mainly for songbirds (e.g., Anderson *et al.* 1996; Briggs *et al.* 2012; Stowell & Plumbley 2014). For non-songbirds such as waterfowl, automated classifiers may be more difficult to develop because their vocalisations are lower frequency (low-frequency noise filters would erase them from recordings) and are structurally short and simple, reducing the number of features available for classification. Nonetheless, automated classifiers for non-songbirds would be a worthwhile endeavour that can substantially increase analysis efficiency of acoustic recordings.

4.5.3 All calls in this study were manually identified by one person (CA), after initial sorting using the clustering process. This study focussed on only three species, with distinctive song patterns, which made classification easier. A number of previous studies have used prior testing of observers to ensure that identification standards are robust and comparable, and that error rates are minimised. This approach and/or the use of a second person to quality check identifications is to be recommended, particularly when the survey scope is the identification of complete assemblages.

4.5.4 In the near future, automated interpretation of recordings will become increasingly viable as an option, allowing effective differentiation of recordings of a range of bird species (Brandes 2008; Acevedo & Villanueva-Rivera 2009; Shonfield & Bayne 2017; Knight *et al.* 2017). The permanent nature of recordings from bioacoustic research will allow any such ongoing developments in call analysis and automated identification to be used to re-analyse previously collected data in the future (Shonfield & Bayne 2017; Stiffler *et al.* 2018). Thus, it will be possible to re-evaluate existing sets of recordings, perhaps alongside new data, improving understanding of species distribution and population trends for a range of species and study locations.

5 Conclusion

5.1.1 Conservation science and practice is becoming increasingly interdisciplinary, and is continuously adopting improved, cheaper and more easily available technologies. The new tools that have developed out of this partnership, including bioacoustics, alongside GPS, remote sensing and eDNA, are indispensable as a part of conservation work today. Such methods allow the collection of more and better data to improve the monitoring of wildlife, habitats and threats, thereby assisting management decisions (Berger-Tal & Lahoz-Monfort 2018).

5.1.2 The work reported here allowed the detection of the target bird species, and enabled daily presence/absence to be determined for each sampling location during the survey period. This illustrated the spatial distribution for each species and allowed the proportion of occupied habitat to be assessed, making use of habitat survey and satellite data. The study therefore demonstrates the suitability of the bioacoustics approach to estimate the populations, distribution and habitat associations of target bird species on the heathland SPA sites, while taking detectability into account.

5.1.3 The use of bioacoustics should be an indispensable part of the wider move towards digital technologies in nature conservation, providing new methods for conducting long-term and potentially continuous monitoring over large spatial scales. As such, the adoption of the bioacoustic approach may be an important component of understanding the ongoing effects of threats and management practices on heathland bird populations.

6 References

Abrahams, Carlos. 2018. “Bird Bioacoustic Surveys – developing a standard protocol.” *In Practice* 102: 20-23

Acevedo, Miguel A., and Luis J. Villanueva-Rivera. 2009. “Using automated digital recording systems as effective tools for the monitoring of birds and amphibians.” *Wildlife Soc B* 34.

Acevedo, Miguel A., And Luis J. Villanueva-Rivera. 2006. “Using Automated Digital Recording Systems as Effective Tools for the Monitoring of Birds and Amphibians.” *Wildlife Society Bulletin* 34: 211–14. doi:10.2193/0091-7648(2006)34[211:UADRSA]2.0.CO;2.

Barton, Kamil. 2018. “tMuMIn: Multi-Model Inference.” CRAN.R-project.org/package=MuMIn.

Bayne, Erin, Michelle Knaggs, and P Solymos. 2017. “How to Most Effectively Use Autonomous Recording Units When Data are Processed by Human Listeners.” The Bioacoustic Unit.

Berger-Tal, Oded, and José J. Lahoz-Monfort. 2018. “Conservation technology: The next generation.” *Conservation Letters*, April, e12458. doi:10.1111/conl.12458.

Berry, Rob. 1979. “Nightjar habitats and breeding in East Anglia.” *British Birds* 72 (5): 207–18.

Bibby, C J, and C R Tubbs. 1975. “Status, habitats and conservation of the Dartford Warbler in England.” *British Birds* 68 (5): 177–95.

BirdLife International. 2006. “Monitoring Important Bird Areas: a global framework. Version 1.2.” Cambridge, UK: BirdLife International.

Brandes, T.S. 2008. “Automated sound recording and analysis techniques for bird surveys and conservation.” *Bird Conservation International* 18 (2008): S163–S173. doi:10.1017/S0959270908000415.

Bright, J A, R H W Langston, and S Bierman. 2007. “Habitat associations of nightjar Caprimulgus europaeus breeding on heathland in England.” 25. Vol. 25.

Browning, Ella, Rory Gibb, Paul Glover-Kapfer, and Kate E. Jones. 2017. “Passive acoustic monitoring in ecology and conservation.” Vol. 2. https://www.wwf.org.uk/conservationtechnology/documents/Acousticmonitoring-WWF-guidelines.pdf.

Burnham, K.P., and D.R. Anderson. 2002. *Model Selection and Multimodel Inference. A Practical Information-Theoretic Approach*. Second Edi. Springer. doi:10.1002/1521-3773(20010316)40:6<9823::AID-ANIE9823>3.3.CO;2-C.

Butchart, Stuart H.M., Jörn P.W. Scharlemann, Mike I. Evans, Suhel Quader, Salvatore Aricò, Julius Arinaitwe, Mark Balman, et al. 2012. “Protecting important sites for biodiversity contributes to meeting global conservation targets.” Edited by Peter M. Bennett. *PLoS ONE* 7 (3). Public Library of Science: e32529. doi:10.1371/journal.pone.0032529.

Campos-Cerqueira, Marconi, and T. Mitchell Aide. 2016. “Improving distribution data of threatened species by combining acoustic monitoring and occupancy modelling.” Edited by Kate Jones. *Methods in Ecology and Evolution* 7 (11): 1340–8. doi:10.1111/2041-210X.12599.

Celis-Murillo, Antonio, Jill L. Deppe, and Michael F. Allen. 2009. “Using soundscape recordings to estimate bird species abundance, richness, and composition.” *Journal of Field Ornithology* 80 (1): 64–78. doi:10.1111/j.1557-9263.2009.00206.x.

Clark, John M., and John Eyre. 2012. “Dartford Warblers on the Thames Basin and Wealden Heaths.” *British Birds* 105 (6): 308–17.

Conway, Greg, Simon Wotton, Ian Henderson, Rowena Langston, Allan Drewitt, and Fred Currie. 2007. “Status and distribution of European Nightjars Caprimulgus europaeus in the UK in 2004.” *Bird Study* 54 (1): 98–111. doi:10.1080/00063650709461461.

Crates, Ross, Aleks Terauds, Laura Rayner, Dejan Stojanovic, Robert Heinsohn, Dean Ingwersen, and Matthew Webb. 2017. “An occupancy approach to monitoring regent honeyeaters.” *Journal of Wildlife Management* 81 (4): 669–77. doi:10.1002/jwmg.21222.

Darras, Kevin, Péter Batáry, Brett Furnas, Antonio Celis-Murillo, Steven L. Van Wilgenburg, Yeni A. Mulyani, and Teja Tscharntke. 2018. “Comparing the sampling performance of sound recorders versus point counts in bird surveys: A meta-analysis.” Edited by Steve Willis. *Journal of Applied Ecology*, July, 1–12. doi:10.1111/1365-2664.13229.

Darras, Kevin, Péter Batáry, Brett Furnas, Irfan Fitriawan, Yeni Mulyani, and Teja Tscharntke. 2017. “Autonomous bird sound recording outperforms direct human observation: Synthesis and new evidence,” 1–37. doi:10.1101/117119.

Depraetere, Marion, Sandrine Pavoine, Fréderic Jiguet, Amandine Gasc, Stéphanie Duvail, and Jérôme Sueur. 2012. “Monitoring animal diversity using acoustic indices: Implementation in a temperate woodland.” *Ecological Indicators* 13 (1): 46–54. doi:10.1016/j.ecolind.2011.05.006.

Ehnes, Mandy, Jeffrey P. Dech, and Jennifer R. Foote. 2018. “Seasonal changes in acoustic detection of forest birds.” *Journal of Ecoacoustics* 2 (April): QVDZO7. doi:10.22261/JEA.QVDZO7.

Fiske, Ian, and Richard Chandler. 2011. “unmarked: An R Package for Fitting Hierarchical Models of Wildlife Occurrence and Abundance.” *Journal of Statistical Software* 43 (10): 1–23. http://www.jstatsoft.org/v43/i10/.

Furnas, Brett J., and Richard L. Callas. 2015. “Using automated recorders and occupancy models to monitor common forest birds across a large geographic region.” *Journal of Wildlife Management* 79 (2). Wiley-Blackwell: 325–37. doi:10.1002/jwmg.821.

Furnas, Brett J., and Michael C. McGrann. 2018. “Using occupancy modeling to monitor dates of peak vocal activity for passerines in California.” *The Condor* 120 (1): 188–200. doi:10.1650/CONDOR-17-165.1.

Haselmayer, John, and James S Quinn. 2000. “A Comparison of Point Counts and Sound Recording as Bird Survey Methods in Amazonian Southeast Peru.” *The Condor* 102 (102): 887–93. doi:10.1650/0010-5422(2000)102[0887:ACOPCA]2.0.CO;2.

Hayhow DB, Ausden MA, Bradbury RB, Burnell D, Copeland AI, Crick HQP, Eaton MA, et al. 2017. “The state of the UK’s birds 2017.” https://www.bto.org/sites/default/files/publications/state-of-uk-birds-{_}2017.pdf.

Holling, Mark. 2012. “Rare breeding birds in the United Kingdom in 2010.” *British Birds* 105 (7): 352–416. doi:10.1016/j.jdent.2003.11.006.

Johnston, Alison, Stuart E. Newson, Kate Risely, Andy J. Musgrove, Dario Massimino, Stephen R. Baillie, and James W. Pearce-Higgins. 2014. “Species traits explain variation in detectability of UK birds.” *Bird Study* 61 (3). Taylor and Francis: 340–50. doi:10.1080/00063657.2014.941787.

Kalan, Ammie K., Roger Mundry, Oliver J.J. Wagner, Stefanie Heinicke, Christophe Boesch, and Hjalmar S. Kühl. 2015. “Towards the automated detection and occupancy estimation of primates using passive acoustic monitoring.” *Ecological Indicators*. doi:10.1016/j.ecolind.2015.02.023.

Knight, Elly C., Kevin C. Hannah, Gabriel J. Foley, Chris D. Scott, R. Mark Brigham, and Erin Bayne. 2017. “Recommendations for acoustic recognizer performance assessment with application to five common automated signal recognition programs.” *Avian Conservation and Ecology* 12 (2). The Resilience Alliance: art14. doi:10.5751/ACE-01114-120214.

La, Van T., and Thomas D. Nudds. 2016. "Estimation of avian species richness: biases in morning surveys and efficient sampling from acoustic recordings." Edited by W. A. Boyle. *Ecosphere* 7 (4): e01294. doi:10.1002/ecs2.1294.

Langston, R H W, S R Wotton, G J Conway, L J Wright, J W Mallord, F A Currie, A L Drewitt, P V Grice, D G Hoccom, and N. Symes. 2007. "Nightjar Caprimulgus europaeus and Woodlark Lullula arborea - Recovering species in Britain?" *Ibis* 149 (SUPPL. 2): 250–60. doi:10.1111/j.1474-919X.2007.00709.x.

Liley, D, and H Fearnley. 2014. "Trends in Nightjar, Woodlark and Dartford Warbler on the Dorset Heaths, 1991-2013." https://www.footprint-ecology.co.uk/reports/Liley and Fearnley - 2014 - Trends in Nightjar, Woodlark and Dartford Warbler .pdf.

MacKenzie, Darryl I., and James D Nichols. 2004. "Occupancy as a surrogate for abundance estimation." *Animal Biodiversity and Conservation* 27 (1): 461–67. http://abc.museucienciesjournals.cat/files/ABC-27-1-pp-461-467.pdf.

MacKenzie, Darryl I., James D. Nichols, Gideon B. Lachman, Sam Droege, Andrew A. Royle, and Catherine A. Langtimm. 2002. "Estimating site occupancy rates when detection probabilities are less than one." *Ecology* 83 (8): 2248–55. doi:10.1890/0012-9658(2002)083[2248:ESORWD]2.0.CO;2.

MacKenzie, DI, JD Nichols, JA Royle, KH Pollock, LL Bailey, and JE Hines. 2006. *Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence*. Elsevier/Academic Press.

Niedballa, Jürgen, Rahel Sollmann, Azlan bin Mohamed, Johannes Bender, and Andreas Wilting. 2015. "Defining habitat covariates in camera-trap based occupancy studies." *Scientific Reports* 5 (1). Nature Publishing Group: 17041. doi:10.1038/srep17041.

Olson, Gail S., Robert G. Anthony, Eric D. Forsman, Steven H. Ackers, Peter J. Loschl, Janice A. Reid, M. Katie, Elizabeth M. Glenn, and William J. Ripple. 2005. "Modeling of site occupancy dynamics for northern spotted owls , with emphasis on the effects of barred owls" *Journal of Wildlife Management* 69 (3). Wiley-Blackwell: 918–32. doi:10.2193/0022-541X(2005)069[0918:MOSODF]2.0.CO;2.

Oppel, Steffen, Sandra Hervias, Nuno Oliveira, Tania Pipa, Carlos Silva, Pedro Geraldes, Michelle Goh, Eva Immler, and Matthew McKown. 2014. "Estimating population size of a nocturnal burrow-nesting seabird using acoustic monitoring and habitat mapping." *Nature Conservation* 7: 1–13. doi:10.3897/natureconservation.7.6890.

Piña-Covarrubias, Evelyn, Andrew P. Hill, Peter Prince, Jake L. Snaddon, Alex Rogers, and C. Patrick Doncaster. 2018. "Optimization of sensor deployment for acoustic detection and localization in terrestrial environments." Edited by Nathalie Pettorelli and Jean Guillard. *Remote Sensing in Ecology and Conservation*, October. Wiley-Blackwell. doi:10.1002/rse2.97.

QGIS Development Team. 2018. "QGIS Geographic Information System." Open Source Geospatial Foundation Project. http://qgis.osgeo.org.

R Core Team. 2013. "R: A language and environment for statistical computing." R Foundation for Statistical Computing, Vienna, Austria. http://www.r-project.org/.

Rempel, Robert S, Keith A Hobson, George W. Holborn, Steve L Van Wilgenburg, and Julie Elliott. 2005. "Bioacoustic monitoring of forest songbirds: interpreter variability and effects of configuration and digital processing methods in the laboratory." *Journal of Field Ornithology* 76: 1–11. doi:10.1648/0273-8570(2005)076.

Rovero, Francesco, and Fridolin Zimmermann. 2016. *Camera Trapping for Wildlife Research*. Pelagic Publishing.

Rowland, C.S., R.D. Morton, L. Carrasco, G. McShane, A.W. O'Neil, and C.M. Wood. 2017. "Land Cover Map 2015 (vector, GB)." NERC Environmental Information Data Centre. doi:https://doi.org/10.5285/6c6c9203-7333-4d96-88ab-78925e7a4e73.

Royle, J. Andrew, and William A. Link. 2006. "Generalized site occupancy models allowing for false positive and false negative errors." *Ecology* 87 (4): 835–941. doi:10.1890/0012-9658(2006)87[835:GSOMAF]2.0.CO;2.

RStudio Team. 2015. "RSudio: Integrated Development for R." RStudio, Inc., Boston, MA. http://www.rstudio.com/ http://www.rstudio.com.

Schmeller, Dirk, Klaus Henle, Adeline Loyau, Aurelien Besnard, and Pierre-Yves Henry. 2012. "Bird-monitoring in Europe – a first overview of practices, motivations and aims." *Nature Conservation* 2: 41–57. doi:10.3897/natureconservation.2.3644.

Shannon, Graeme, Jesse S. Lewis, and Brian D. Gerber. 2014. "Recommended survey designs for occupancy modelling using motion-activated cameras: insights from empirical wildlife data." *PeerJ* 2 (August). PeerJ Inc.: e532. doi:10.7717/peerj.532.

Shonfield, Julia, and Erin M. Bayne. 2017. "Autonomous recording units in avian ecological research: current use and future applications." *Avian Conservation and Ecology* 12(1) (1): 14. doi:10.5751/ACE-00974-120114.

Sitters, H. P., R. J. Fuller, R. A. Hoblyn, M. T. Wright, N. Cowie, and C. G.R. Bowden. 1996. "The Woodlark Lullula arborea in Britain: Population trends, distribution and habitat occupancy." *Bird Study* 43 (2): 172–87. doi:10.1080/00063659609461010.

Stiffler, Lydia L., James T. Anderson, and Todd E. Katzner. 2018. "Occupancy Modeling of Autonomously Recorded Vocalisations to Predict Distribution of Rallids in Tidal Wetlands." *Wetlands*. Wetlands, 1–8. doi:10.1007/s13157-018-1003-z.

Welsh, Alan H., David B. Lindenmayer, and Christine F. Donnelly. 2013. "Fitting and Interpreting Occupancy Models." Edited by Ethan P. White. *PLoS ONE* 8 (4). Public Library of Science: e52015. doi:10.1371/journal.pone.0052015.

Wildlife Acoustics. 2017. "Kaleidoscope Pro 3." Wildlife Acoustics.

Yates, M. D., and R. M. Muzika. 2006. "Effect of Forest Structure and Fragmentation on Site Occupancy of Bat Species in Missouri Ozark Forests." *Journal of Wildlife Management*. doi:10.2193/0022-541X(2006)70[1238:EOFSAF]2.0.CO;2.

Yip, Daniel A., Lionel Leston, Erin M. Bayne, Péter Sólymos, and Alison Grover. 2017. "Experimentally derived detection distances from audio recordings and human observers enable integrated analysis of point count data." *Avian Conservation and Ecology* 12 (1): art11. doi:10.5751/ACE-00997-120111.

Zwart, Mieke C., Andrew Baker, Philip J K McGowan, and Mark J. Whittingham. 2014. "The use of automated bioacoustic recorders to replace human wildlife surveys: An example using nightjars." *PLoS ONE* 9 (7). doi:10.1371/journal.pone.0102770.

7 Appendix

7.1 Occupancy and Detectability Considerations

7.1.1 Proper occupancy estimation depends on certain sampling assumptions, which must be taken into account during survey and data analysis. These are: (i) that sites are closed to changes in occupancy state between sampling occasions, (ii) that the detection process is independent at each sampling site, (iii) that the probability of detection of a species is either constant across sites, or that it can be explained by habitat covariates (such as vegetation type or altitude). These assumptions can be addressed by conducting field recording during a single breeding season, when breeding territories are established and stable, ensuring that sampling areas have no spatial overlap, and by placing recorders in similar habitat types (or by recording standard habitat data at each sample site) (MacKenzie *et al*., 2006; Rovero & Zimmermann 2016).

7.1.2 Any bioacoustic survey design operating with limited resources has a trade-off to make between the number of sample sites than can be covered within a survey period and the length of time that a recorder can be left at each site, i.e. should the survey go for "more sites-shorter time" or "fewer sites-longer time". A correctly balanced approach in gathering data for occupancy modelling within this framework should reflect the detectability and habitat occupancy of the target species. Bird species will vary in both of these aspects based on their behaviour and population densities (Johnston *et al*., 2014), and so to avoid bias, optimal survey methods may need to be different for a range of taxa. Either that, or methods need to be robust enough to allow for variation in these factors.

7.1.3 Estimates of detectability for 195 UK bird species have been produced, based upon the national Breeding Bird Survey method (Johnston *et al.* 2014). For this survey, which uses both visual and audible cues to surveyors, detectability has been found to be significantly affected by bird size, diet and habitat specialization. Of the seven largest orders considered, passerines (Passeriformes) had the lowest median detectability of 0.37, while shorebirds (Charadriiformes) had the highest median detectability of 0.65. Species most associated with closed habitats such as woodland and urban areas had the lowest detectability. Smaller species had lower detectability than larger species. If not taken into account, these differences in species detectability can lead to biased conclusions, particularly when calculating multi-species indices such as species richness or diversity (Johnston *et al*. 2014).

7.1.4 Estimated detectability for the three heathland species is shown in Table A1, with a higher index value indicating species with higher detectability. The figures shown in Table A1 for Dartford warbler and woodlark are higher than the median detectability for passerines of 0.37, and so these two species are more detectable than average for their taxonomic order. Nightjar detectability from the BBS results is low, as might be expected for a crepuscular species.

7.1.5 Alongside detectability, occupancy for the three heathland species is also fairly well known from previous surveys which have assessed breeding densities (Liley & Fearnley, 2014), and from the population levels provided within the published SPA citation. In relative terms for the three target species, woodlark density can be characterised as low, nightjar moderate and Dartford warbler high (Table A1) - although the population crash for the latter species in 2010-2011 would obviously have altered this temporarily.

Table A1. Detectability estimates and relative breeding densities for the SPA bird species (detectability values from Johnston et al. 2014)

Species	Detectability	Breeding density
Woodlark	0.472	Low
Dartford Warbler	0.374	High
Nightjar	0.297	Moderate

7.1.6 In any occupancy study, the survey effort (number of recorders deployed and the length of sampling period) affects the accuracy and precision (i.e. error) of the occupancy estimate. It has been found (Shannon *et al.* 2014) that increasing total sampling effort generally decreases error associated with the occupancy estimate, but changing the number of sites or sampling duration can have very different results, depending on whether a species is spatially common or rare and easy or hard to detect when available. For rare species with a high probability of detection (i.e. woodlark for this study) the required survey effort should maximize the number of sites covered. For common species with low detection (i.e. nightjar) the most efficient sampling approach is to increase the number of occasions (survey days). However, for common species that are moderately detectable (i.e. Dartford warbler) occupancy should be reliably estimated with comparatively low numbers of recorders over a short sampling period. Clearly any survey design that aims to gather data on all three species will have to balance spatial and temporal coverage of the recording units.

7.2 Spatial Coverage

7.2.1 In occupancy modelling, the distance between sampling nodes (and hence recording units) should be relevant to the territory size of the taxa being recorded, ideally being slightly larger than the diameter of the average home range of the species of interest (Niedballa *et al*., 2015; Rovero & Zimmermann 2016). This ensures that spacing is sufficient to limit recording the same individuals, so that counts are independent and spatial autocorrelation is avoided. This can deal with the issue of home-range, but with audio recording, the spacing also needs to take into account the distance between a vocalising bird and the receiving microphone (which is not an issue when using similar methods with camera traps). For bird sound, the effective recording radius of most detectors is likely to be in the region of 50m, dependent on the volume and frequency of the bird species, the surrounding habitat, weather conditions, the detector type and microphone condition (Furnas & Callas, 2015; Yip *et al*., 2017), so a distance between recorders of 100m or more should normally be sufficient to prevent overlap between recorders, i.e. vocalisations from one bird being recorded at two (or more) sampling sites.

7.2.2 In heathland situations, Dartford warbler and nightjar have been found with nesting densities of 10-20 and 8-10 pairs per km^2 respectively, or average separation distances of c.170m (Bibby & Tubbs 1975; Berry 1979; Bright *et al*., 2007). The 2004 UK nightjar survey recorded a median distance between simultaneously calling nightjar males of 314 m (Conway *et al*., 2007). Based on these data, a sampling grid with nodes around 200-300m apart would be appropriate for the species being studied here.

7.2.3 For any future bioacoustic studies, additional refinement of detector placement may be warranted to maximise coverage of sites. This study utilised a stratified, randomised sampling approach within a regular 250m grid. However, recent research (Piña-Covarrubias *et al*. 2018) has indicated that, for a desired threshold of detection efficiency, careful selection of placements based on topography, vegetation and weather patterns may be more efficient. When this is done, near-optimal placements on hilly terrain can halve the number of devices otherwise needed for square or random grids, thereby more than halving monitoring costs. Developing such an optimised recorder layout may be more complex, but could be worth undertaking if the supply of recorders was limited or if a greater area coverage was required.



bakerconsultants