Development of time-lapse photography for the population monitoring of a colonial seabird

Alice Jane Edney

A thesis submitted to the University of Gloucestershire in accordance with the requirements of the degree of MSc by Research in the School of Natural and Social Sciences

In collaboration with the
University of Oxford and University College Cork

September 2020
Word count: 25, 694

ABSTRACT

Seabirds are one of the most threatened groups of birds and large-scale monitoring is needed to link changing population trends to causative factors, in order to address population declines. Rapid advances in technology are offering new and exciting possibilities to expand monitoring over larger spatial and temporal scales, however, they also raise new challenges, such as dealing with increased amounts of data and ensuring the data obtained are comparable to that from 'traditional' monitoring methods. Specifically, this research focused on the use of time-lapse cameras to monitor the Black-legged Kittiwake *Rissa tridactyla,* a species listed as Vulnerable on the International Union for Conservation of Nature (IUCN) Red List. Chapters one and two used a case study on Skomer Island, Wales, to compare measurements of productivity and phenology obtained from fieldwork with expert analysis of time-lapse images. Chapter two then went on to explore the effects of weather on Kittiwake nest survival on Skomer. Chapter three used data from across a much wider area, to compare expert analysis of time-lapse images with citizen science analysis. This study showed that both field and image-derived data have inherent biases, but together can inform meaningful investigation into the factors contributing to Kittiwake decline. I found that strong westerly winds may be reducing egg and chick survival at the Wick colony, Skomer, and high daily maximum temperatures could also be lowering egg survival. If these results represent a longer-term pattern, then it could have important implications for Kittiwake population dynamics with climate change, which is predicted to increase the frequency and intensity of weather extremes. Expanding the scale of monitoring via the citizen science project, *Seabird Watch,* was found to have promising potential; although further work is needed to ensure volunteer data are as good as expert classification. Many factors affected the accuracy of citizen science results and these must be carefully considered before using the data to answer bigger scientific questions. Overall, this study has shown the potential of using time-lapse imagery to monitor a cliff-nesting seabird and will likely become an increasingly cost-effective monitoring solution in the coming years.

DECLARATION

I declare that the work in this thesis was carried out in accordance with the regulations of the University of Gloucestershire and is original except where indicated by specific reference in the text. No part of the thesis has been submitted as part of any other academic award. The thesis has not been presented to any other education institution in the United Kingdom or overseas.

Any views expressed in the thesis are those of the author and in no way represent those of the University.

Signed [redacted] Date: 28th September 2020

doi: 10.46289/CI36CO13

ACKNOWLEDGEMENTS

Firstly, I would to thank the University of Gloucestershire's Environmental Dynamics and Governance RPA for awarding me a bursary to undertake the MSc by Research, without which it would not have been possible. Equally, the research could not have taken place without field data collected by Alexa [redacted] on behalf of the Wildlife Trust of South and West Wales, photographs from Skomer Island provided by Matt Wood, photographs from the *Seabird Watch* citizen science project provided by Tom Hart and Mark Jessopp, and weather data from the Met Office.

I am extremely grateful for the help and support provided by my main supervisor, Matt Wood, and co-supervisors, Tom Hart and Mark Jessopp. Thank you to Matt for writing (presumably!) good references that helped me receive a PhD offer at Oxford with Tom, and thank you to Tom for agreeing to supervise me for another 4 years…

Thank you to my all-time favourite office buddy, Bryony [redacted], for offering endless advice, making me laugh and just generally bringing some life and fun to the FCH post-grad office. Now we just need to get back out to an island together! Also, thank you to my Cheltenham house mates for putting up with my strange hobbies and crazy ways. It is a shame we had to leave our Foley family early. Thank you to my parents and sister for welcoming me and my vast quantities of 'stuff' back home when lockdown began. Despite intending to remain in Cheltenham for a year, I enjoyed the 6 months I had there, and equally enjoyed the 4 months at home not having to cook. And of course, thank you to everyone who supported my decision to clear off to the Azores August – September 2020 with an uncompleted Masters dissertation weighing on my shoulders. With a few necessary nudges from Hannah [redacted] and Max [redacted], I was thankfully able to finish this thesis while still having plenty of time to monitor seabirds and enjoy island life. 2020 has thrown some curveballs, but I'm grateful to have been able to complete my MSc and am incredibly thankful to everyone who supported me along the way.

TABLE OF CONTENTS

INTRODUCTION

Seabirds are one of the most threatened groups of birds (Croxall *et al.* 2012). According to the International Union for Conservation of Nature (IUCN) Red List, 31% of seabird species are globally Threatened and 11% are Near-Threatened (BirdLife International 2018). Almost half of seabird species are undergoing population decline, primarily due to invasive species, climate change and fisheries bycatch (BirdLife International 2012, Dias *et al.* 2019). Changes in seabird population dynamics can have widespread impacts on marine ecosystem function via top-down effects, as seabirds are apex predators (Horswill *et al.* 2016, Lescroël *et al.* 2016, Suryan *et al.* 2006). Additionally, changes in abundance at lower trophic levels will influence seabirds from the bottom-up (Suryan *et al.* 2006, Horswill *et al.* 2016, Sydeman *et al.* 2017a, 2017b, Pacyna *et al.* 2019, Reynolds *et al.* 2019). As a result, seabirds have often been used as indicators of ecosystem health, and understanding the causes of their decline is necessary to implement successful conservation action at both the level of the species and ecosystem (Lescroël *et al.* 2016, Sydeman *et al.* 2017a, 2017b, Bland *et al.* 2018, Pacyna *et al.* 2019, Reynolds *et al.* 2019). In order to achieve this, effective monitoring of seabird populations is required.

UK seabird monitoring

Historically in the UK, the relationship between humans and seabirds was largely one of exploitation, with humans viewing seabirds as a source of food and feathers (Tasker 2000). However, diminishing harvests in the early 20th century began to indicate possible declines in UK seabird populations, and led to the foundation of the Royal Society for the Protection of Birds (RSPB) and some of the first bird conservation legislation (Tasker 2000). Seabird population counts began to be formally conducted at this time, including counts of Northern Gannet *Morus bassanus*, Northern Fulmar *Fulmarus glacialis*, and Black-legged Kittiwake *Rissa tridactyla* colonies in Britain and Ireland (Gurney 1913, Fisher 1959, Coulson 1963, Tasker 2000). These species-specific population counts were expanded to provide a census of all regularly breeding UK seabirds 1969-70, which was termed Operation Seafarer (Cramp *et al.* 1974, Tasker 2000). From the 1950s, biologists also began to study parameters other than population size (Tasker 2000). In 1991, Dunnet reported on 41 years of Northern Fulmar population ecology on Eynshallow, Orkney (Dunnet 1991). The study measured frequency of breeding, breeding success, age of first breeding and longevity, as well as population size (Dunnet 1991). These initial seabird surveys in the early to mid-20th century helped provide the impetus for focused national monitoring. Operation Seafarer showed that the UK held internationally important seabird populations, and that more comprehensive and coordinated monitoring would be needed to ensure these populations remained healthy (Tasker 2000). In 1986, the UK Seabird Monitoring Programme (SMP) was established.

The Seabird Monitoring Programme is an annual monitoring programme involving 19 partner organisations and is coordinated by the Joint Nature Conservation Committee (JNCC). It aims to collect sample data on the breeding abundance, breeding success, and, where possible, survival, phenology and diet of 25 seabird species that regularly breed in Britain and Ireland: Northern Fulmar *Fulmarus glacialis*, Manx Shearwater *Puffinus puffinus*, European Storm-petrel *Hydrobates pelagicus*, Leach's Storm-petrel *Oceanodroma leucorhoa*, Northern Gannet *Morus bassanus*, Great Cormorant *Phalacrocorax carbo*, European Shag *Phalacrocorax aristotelis*, Arctic Skua *Stercorarius parasiticus*, Great skua *Stercorarius skua*, Black-legged Kittiwake *Rissa tridactyla*, Black-headed Gull *Chroicocephalus ridibundus*, Mediterranean Gull *Larus melanocephalus*, Common Gull *Larus canus*, Lesser Black-backed Gull *Larus fuscus*, Herring Gull *Larus argentatus*, Great Black-backed Gull *Larus marinus*, Little Tern *Sternula albifrons*, Sandwich Tern *Sterna sandvicensis*, Common Tern *Sterna hirundo*, Roseate Tern *Sterna dougallii*, Arctic Tern *Sterna paradisaea*, Common Guillemot *Uria aalge*, Razorbill *Alca torda*, Black Guillemot *Cepphus grylle* and Atlantic Puffin *Fratercula arctica* (Joint Nature Conservation Committee 2020). This annual data is collected from keys sites spread geographically around the UK: Skomer Island (west Wales); Isle of May (east Scotland); Fair Isle (Northern Isles); Canna/Rum/St Kilda (west Scotland) (Tasker 2000). Annual monitoring is then complemented by complete censuses conducted every 10 to 15 years, to provide counts for all UK breeding seabird colonies. So far, national censuses have been undertaken 1969-70 (Operation Seafarer), 1985-88 (Seabird Colony

Register), 1998-2002 (Seabird 2000) and 2015-2021 (Seabirds Count) (Joint Nature Conservation Committee 2020).

Having a well-organised long-term monitoring programme is important so that trends in population numbers, productivity, survival and diet can be determined and species conservation status can be assessed. This information is also critical for assessing the potential impact of development and infrastructure on breeding seabirds, for example when licensing near-shore windfarms. Methods for the current Seabird Monitoring Programme are provided in the 'Seabird Monitoring Handbook for Britain and Ireland' to ensure the methodology used is consistent across colonies and the results are comparable (Walsh *et al.* 1995). However, these methods are not without their limitations and indeed monitoring seabird populations is often challenging.

Pelagic seabird species only visit land to breed and spend the rest of the year at sea. Many species nest on exposed cliffs which are difficult to access, particularly during inclement weather. Some species nest underground in burrows and those nesting on the ground surface may be concealed by vegetation or camouflage. (Mitchell & Parsons 2007, Robinson & Ratcliffe 2010). As well, frequently visiting breeding colonies can cause disturbance, is logistically challenging in remote locations, and expensive in terms of time and money (Anker-Nilssen *et al.* 1996, Field *et al.* 2005, Huffeldt & Merkel 2013, Southwell & Emmerson 2015). Consequently, the temporal and spatial scale of seabird monitoring is often small (Evans 1986, Lynch *et al.* 2012a, Paleczny *et al.* 2015).

Recent advances in digital photography and videography offer considerable potential for overcoming some of the difficulties associated with monitoring seabird populations. While time-lapse photography has been used for behavioural studies for decades, digital photography for wildlife monitoring has tended to be limited to small studies observing animals opportunistically, using animal-triggered or handheld cameras (Penney 1968, Harris 1982, Black 2018). Now, improved optics, increased battery life and data storage have transformed the potential of remote imaging and made it possible to monitor hard to access populations (Bolton *et al.* 2007, Kucera & Barrett 2011, Anderson & Gaston 2013, Black 2018). Nonetheless, choosing suitable equipment for a specific monitoring purpose can be difficult with such a wide array of technology now available. Furthermore, there are challenges associated with how to handle and analyse large volumes of digital data. For the first part of my MSc by Research, I conducted a literature review which discussed the potential uses of different types of digital imaging technology in seabird monitoring and research, as well as methods for processing the data. This review has since been published in *Ibis* and is included in Appendix 1 (Edney & Wood 2020). One type of digital imaging technology discussed in the review are time-lapse cameras, which shall be the focus of this thesis. Parts of the following text on 'Time-lapse cameras' are taken directly from Edney & Wood (2020) with permission.

Time-lapse cameras

Time-lapse cameras record photographs at predetermined time-intervals regardless of subject presence (Cutler & Swann 1999). They have been utilised for avian studies since the technology first became commercially available, however their potential uses in ornithology are quickly increasing with advances in digital technology (Dodge & Snyder 1960, Green & Anderson 1961, Cowardin & Ashe 1965, Temple 1972, Weller & Derksen 1972, Harris 1982, Harris 1987, Huffeldt & Merkel 2013). The increased availability of affordable cameras, requiring less frequent maintenance, reduced power consumption and larger storage capacity has seen the field of time-lapse photography rapidly expand in recent years (Bolton *et al.* 2007).

Time-lapse photography is most appropriate for studying animals frequently present at a location, where a single vantage point gives a representative view of individuals, and the measurement of interest will not activate a motion-triggered camera (Cutler & Swann 1999, Black 2018). Species that aggregate at high densities at some point in their life-history, such as breeding seabirds, therefore represent ideal candidates for use (Black 2018). Time-lapse cameras are suited for collecting data as part of long-term studies, principally time-series data such as annual breeding success and phenology, and can achieve this at scales otherwise unfeasible in terms of both time and money (Southwell & Emmerson 2015, Merkel *et al.* 2016, Black *et al.* 2018a, Hinke *et al.* 2018). They can capture images in locations and at times when human observation would be almost impossible, including harsh conditions in remote places and at night (Black 2018, Black *et al.* 2017, Black *et al.* 2018a, Southwell & Emmerson 2015). Table 1 offers a more detailed summary of the advantages and disadvantages of time-lapse photography as a tool for monitoring seabirds (Edney & Wood 2020).

Table 1 Advantages and disadvantages of using time-lapse photography to monitor breeding seabirds (Edney & Wood 2020).

Advantages	**Disadvantages**
Cost-effective: saves time and money during fieldwork. For example, difficult for a single researcher to record detailed nest activity across multiple nests at a colony. (Weller & Derksen 1972, Black 2018, De Pascalis *et al.* 2018)	Mechanical failures. (Cutler & Swann 1999, Merkel *et al.* 2016, Black 2018)
Increased spatial and temporal scale of monitoring. (Southwell & Emmerson 2015, Merkel *et al.* 2016)	Programming errors. (Cutler & Swann 1999, Black 2018)
Operates at locations and times when field observation would be near-impossible. For example, remote locations, harsh weather conditions, at night. (Cutler & Swann 1999, Southwell & Emmerson 2015, Black *et al.* 2017, Sinclair et al. 2017, Black 2018, Black *et al.* 2018)	Maintenance required. For example, images are vulnerable to camera movements caused by harsh weather conditions. (Merkel *et al.* 2016, Black 2018)
Removes observer bias from variation in surveyor experience and alertness over a long period. (Cowardin & Ashe 1965, Weller & Derksen 1972, Cutler & Swann 1999, Black 2018)	More affordable cameras take lower resolution images. (Black 2018)
More frequent observations than field workers allows observation of elusive species, obscure behaviours and phenology. (Cutler & Swann 1999, Black 2018)	Large camera networks needed to monitor an entire colony, which are expensive to install and maintain. (Black 2018)
Permanent record viewable any number of times and available for independent verification. (Cutler & Swann 1999, Merkel *et al.* 2016; Sinclair et al. 2017, Black, 2018)	Large amount of data to handle and analyse. (Merkel *et al.* 2016, Black 2018)
Easier to maintain comparable study effort between years. (Merkel *et al.* 2016)	Cameras rarely possess thermal imaging or infra-red sensors, making night monitoring difficult. (Black *et al.* 2018)

Infrequent visitation lowers nest and site disturbance. (Cutler & Swann 1999)	

The wide range of studies already using time-lapse cameras to monitor seabird populations demonstrates their considerable potential for improving the scale of seabird monitoring. Photographs have been used to obtain baseline ecological data on seabirds, such as breeding success and population counts, as well as finer scale ecological and behavioural information (Cowardin & Ashe 1965, Southwell & Emmerson 2015, Black *et al.* 2018a, Hinke *et al.* 2018, De Pascalis *et al.* 2018). This includes monitoring nest activity (such as nest attendance and division of labour between parents), re-sighting marked birds to determine adult survival and foraging behaviour (including foraging trip duration and frequency of foraging trips), predation and the timing and duration of phenological events (Weller & Derksen 1972, Mudge *et al.* 1987, Collins *et al.* 2014, Southwell & Emmerson 2015, Black *et al.* 2017, Black *et al.* 2018b, Hinke *et al.* 2018, De Pascalis *et al.* 2018).

Most studies so far have deployed time-lapse cameras at a small-scale, for example to monitor a single colony of interest, and so manual image analysis has been feasible (Southwell & Emmerson 2015, Black *et al.* 2017). However, as the scale of study increases, the number of raw images collected can quickly exceed researchers' processing capabilities (De Pascalis *et al.* 2018). This has likely prevented time-lapse cameras from monitoring seabird colonies across large spatial scales.

One method of processing huge volumes of data is to use volunteer citizen scientists to analyse photographs. Image annotation from citizen scientists can also be used to train computer algorithms to learn to recognise birds in photographs (Jones *et al.* 2020). Before machine learning can be deployed, researchers must be confident that observations from time-lapse photographs are comparable to field observations, and image analysis by citizen scientists is equivalent to examination by skilled image analysts who are 'experts' in the field. My thesis aims to compare expert observations from time-lapse images with field and citizen science observations, to assess the suitability of time-lapse imagery for monitoring a threatened seabird species: the Black-legged Kittiwake *Rissa tridactyla*.

Study species

Population status

Black-legged Kittiwakes (hereafter Kittiwake) are a small, pelagic gull found throughout the Northern Hemisphere. The global population is estimated at nine million adults, making them the world's most numerous gull (Coulson 2011). However, the species has declined rapidly over the past 40-50 years with no evidence of slowing, leading to their classification as 'Vulnerable' on the IUCN Red List (BirdLife International 2019). In the British Isles, the breeding Kittiwake population has declined by 23% since the mid-1980s (Mitchell *et al.* 2004). Nationally, Scotland has suffered the greatest loss, with the 2014 index being 72% below the 1986 baseline (Wilkie *et al.* 2019). Severe declines have also been observed at higher latitudes, namely Greenland, Iceland, mainland Norway and the Faroe Islands (Hentati-Sundberg, 2011; Sandvik *et al.,* 2014).

Productivity and phenology

Typically, changes in seabird populations are attributed to changes in post-fledging and adult survival (Weimerskirch *et al.* 1997, Mitchell *et al.* 2004). This is because seabird life-history is characterised by long-lifespan and low fecundity, as they mature late and produce few chicks per annum (Bennett & Owens 2002). This means their populations may be buffered by short-term variation in productivity due to a long life-span providing many opportunities to reproduce (Sæther & Bakke 2000, Jenouvrier *et al.* 2005). Seabirds are therefore 'k-selected' species. Conversely, the population growth rate of birds with low adult survival and high fecundity (r-selected species) is most sensitive to breeding success (Sæther & Bakke 2000). Nevertheless, recent evidence suggests poor breeding success can also be an important driver of seabird population decline alongside survival (Reiertsen 2013). This is particularly evident if productivity is reduced in multiple consecutive years, as has occurred at many Kittiwake colonies (Mitchell *et al.* 2004). Time-lapse cameras are well suited to monitoring productivity, as Kittiwakes usually return to the same locations to breed each year (Boulinier *et al.* 2008, Ponchon *et al.* 2015). This means a fixed-position camera can record the same nests every breeding season, providing comparable data over time. Furthermore, cameras can measure breeding phenology by recording the date of parent arrival, egg laying, egg hatching, chick fledging and parent departure. Associations between the precise timing of events and other variables can be explored to provide an insight into the factors contributing to low productivity and ultimately Kittiwake decline.

In the North Sea there is evidence that increasing sea surface temperature (SST) linked to climate change is altering the availability of Kittiwake's primary prey, Lesser Sandeel *Ammodytes marinus* (Eerkes-Medrando *et al.* 2017). This is likely due to a decrease in sandeel abundance (Frederiksen *et al.* 2004, 2005, 2007, 2013), as well as a shift in phenology causing a decline in prey energy-value at peak demand (Burthe *et al.* 2012). This decrease in prey availability has been compounded by sandeel fishing off south-east Scotland in the 1990s, which lowered adult survival and breeding success relative to years before the fishery opened and after it closed (Frederiksen *et al.* 2004).

Despite SST playing an important role in Kittiwake decline in some areas of Scotland, the same has not been shown across the UK. Winter SST was not a reliable indicator of Kittiwake breeding success at Fowlsheugh on the east coast of Scotland (Eerkes-Medrano *et al.* 2017) and there was no relationship between Celtic Sea STT and breeding success or population growth of Welsh breeding Kittiwakes (Lauria *et al.* 2012, Cook *et al.* 2014). Instead, the North Atlantic Oscillation (NAO) appeared to have a weak positive effect on Kittiwake population growth rate (Lauria *et al.* 2012). This may have been related to the direct effects of wind speed or storm frequency which are positively correlated with NAO (Lauria *et al.* 2012), although NAO can also be positively associated with SST making the results difficult to interpret (Qu *et al.* 2012). It is therefore evident that different mechanisms are likely contributing to lowered Kittiwake productivity in different locations (Carroll *et al.* 2015). SST is having a greater effect on colonies in Scotland, while there is less certainty regarding factors contributing to the decline in Wales. As well as assessing the suitability of time-lapse cameras for measuring Kittiwake phenology and productivity, this thesis will also explore the effect of weather on Kittiwake nest survival at a colony in Wales, to investigate the importance of prevailing weather conditions on Kittiwake breeding success.

Thesis aims and structure

Overall, there are three main aims to this thesis, which have been separated into three chapters.

1. **Chapter one**
   Aim: To determine the effectiveness of time-lapse photography for measurements of Black-legged Kittiwake nest success and phenology, by comparing field observations with expert analysis of photographs.

This chapter uses a case study of part of the Kittiwake colony on Skomer Island, Wales, to compare measurements of productivity and phenology from field observations and expert annotation of time-lapse images by researchers.

2. **Chapter two**
   Aim: To explore associations between Black-legged Kittiwake nest survival and weather events.

This chapter uses data from field observations and time-lapse images on Skomer in 2019 (i.e. from chapter one) to model the survival rate of Kittiwake eggs and chicks with different temporal, intrinsic, extrinsic variables. Based on the hypothesis that prevailing weather conditions could be having a greater impact on Kittiwake breeding success in Wales than SST, the extrinsic variables used were weather parameters (Lauria *et al.* 2012). This chapter aims to act as a preliminary study to see if there is any support for weather events contributing to the gradual reduction in Kittiwake breeding success on Skomer Island.

3. **Chapter three**
   Aim: To determine the effectiveness of time-lapse photography for measurements of Black-legged Kittiwake abundance, nest success and phenology at wider spatio-temporal scales than can be achieved by individual researchers, by comparing citizen science and expert analysis of photographs.

This chapter uses data from the citizen science project *Seabird Watch* (www.seabirdwatch.org) to compare counts of adult and juvenile Kittiwakes in time-lapse images by citizen scientists and 'gold standard' annotation by researchers.

Following the final chapter there is a general discussion of the use of time-lapse cameras for monitoring Kittiwakes, which brings together the results and conclusions from all three chapters.

CHAPTER ONE: Comparing field observations and expert analysis of time-lapse photographs

ABSTRACT

Monitoring seabird phenology and productivity is essential to understand long-term population trends. However, it can be logistically difficult, expensive in terms of time and money, and often precludes detailed phenology data depending on the frequency with which observations are made. Using time-lapse cameras to take photographs of cliff-nesting seabirds can potentially overcome some of these challenges. I compared Black-legged Kittiwake *Rissa tridactyla* phenology and productivity measured from time-lapse images and field observations, on Skomer Island, Wales, in 2019 to assess the effectiveness of digital monitoring relative to 'traditional' methods. Analysis of photographs recorded nests as 'apparently occupied' on average 3.7 days earlier than field observations. Photographs also observed chicks for the first and last time on average 13.9 and 9.6 days later respectively. Productivity measured from photographs ranged from 0.46 to 0.83 depending on the method used to determine the number of fledged chicks per nest. Field observations calculated productivity as 0.83. This study has shown that time-lapse photography has potential to improve seabird monitoring provided the camera can be positioned sufficiently close to the colony, or an appropriate magnification achieved with an optical lens. While the time-lapse camera on Skomer Island has potential to assist Kittiwake monitoring, improvements in camera lens magnification are needed to realise the full potential of this technology.

INTRODUCTION

Long-term monitoring programmes are essential for determining trends in seabird population ecology, which are necessary to assess species conservation status and understand potential causes of population change. The UK Seabird Monitoring Programme (SMP) has coordinated annual seabird monitoring in Britain and Ireland since 1986, with detailed data being collected at several key sites around the UK (Joint Nature Conservation Committee 2020). Annual estimates of breeding population size allow population trends to be determined, while measurements of phenology, breeding success, survival and diet can provide insights into the reasons for population change. The Seabird Monitoring Programme follows methodology provided in the 'Seabird Monitoring Handbook for Britain and Ireland', with consistent use of methods helping to ensure results are comparable across colonies (Walsh *et al.* 1995). However, most of the methods used are time-intensive, as fieldworkers must regularly visit sites to obtain the measurements required. For example, regular repeat visits are needed to resight large numbers of colour rings for survival studies and obtain time-series data to allow measurements of phenology and breeding success. Furthermore, many seabird species breed in remote locations, such as islands, which can be difficult and expensive to access. As a result, annual seabird monitoring is only conducted at a few key sites around the UK, limiting the spatial scale of our understanding of the UK breeding seabird population.

One solution to try and overcome some of the challenges of current field monitoring methods is to use remote sensing technologies (Edney & Wood 2020). In particular, time-lapse cameras could offer a means of collecting detailed time-series data without the need for regular field visits. Cameras set to record images once per hour may only need their SD cards and batteries changing every one to two years, allowing a single field visit per annum to collect data (Merkel *et al.* 2016). Installing a network of time-lapse cameras around the UK may therefore be a comparatively cost-effective solution to expand both the spatial and temporal scale of UK seabird monitoring. However, before such a task is undertaken, it is important to ensure that measurements from time-lapse cameras are comparable to traditional field observations. This requires a comparative study at a site where both field and time-lapse data have been collected.

Skomer Island is owned by Natural Resources Wales (NRW) and managed by the Wildlife Trust of South and West Wales (WTSWW). It is part of a Marine Conservation Zone and is a nationally important site for breeding seabirds (Stubbings *et al.* 2017). Long-term monitoring of seabirds has been conducted here since the 1960s and focuses on a number of species, including the Black-legged Kittiwake *Rissa Tridactyla* (hereafter Kittiwake) (Brown *et al.* 2004, Wilkie *et al.* 2019). Monitoring of Kittiwakes aims to gather four main pieces of information: the number of breeding birds, productivity (i.e. breeding success), timing of breeding and breeding adult survival. This data has shown that while the UK decline in breeding Kittiwakes has been most rapid in Scotland, a gradual decrease is occurring in Wales (Wilkie *et al.* 2019). On Skomer, Kittiwake numbers have fallen by 42% since 1986, likely a result of low productivity coupled with low survival (Wilkie *et al.* 2019). Preliminary Population Viability Analysis has shown that if current rates of survival and productivity continue, then the breeding population will decline to a few hundred by the end of the 21st century, and if parameters worsen the species may be locally extinct within 70 years (Horswill, Perrins & Wood, unpublished).

In 2019, a time-lapse camera was installed at one of the Skomer Kittiwake colonies to provide consistent year-round monitoring and hopefully give further insight into the factors affecting Kittiwake productivity and survival, which is the first step in developing measures to reverse their decline. This chapter aims to investigate how measurements of Kittiwake productivity and phenology from time-lapse photographs compare with field observations on Skomer Island. This will help assess whether time-lapse photography can improve the temporal and spatial scale of seabird monitoring.

METHODS

Data sources and collection

Field observations

Kittiwake productivity and phenology have been monitored at the same three sub-colonies on Skomer Island since 1989: South Stream (plots SS1 and SS2), High Cliff (plot HC2) and the Wick (plots W3, W4, W5 and W6) (Fig. 1). Observations are made in accordance with the Seabird Monitoring Handbook for Britain and Ireland (hereafter Seabird Monitoring Handbook) (Walsh *et al.* 1995). In 2019 the fieldworker was employed by the WTSWW and the data were gathered by Alex Piggott.

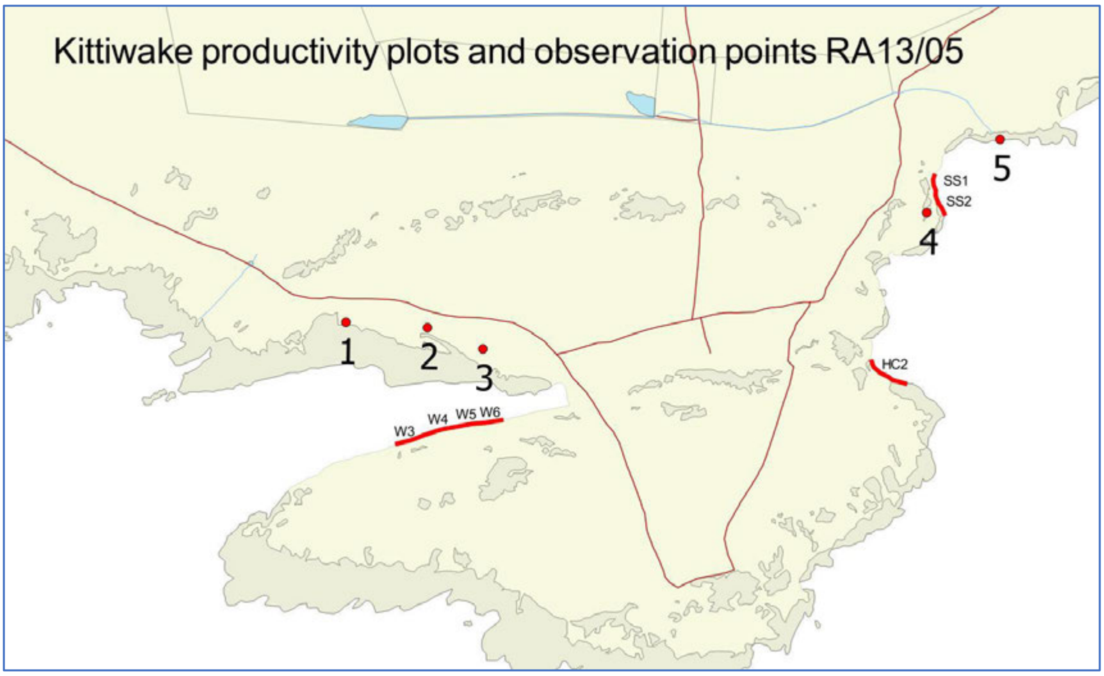


Figure 1 Kittiwake productivity monitoring on Skomer Island takes place at three sub-colonies which are divided into seven plots: the Wick (plots W3, W4, W5 and W6), High Cliff (plot HC2) and South Stream (plots SS1 and SS2). These plots are viewed by the fieldworker from observation points 1, 2, 3, 4 and 5. Image: Skomer Management Plan, WTSWW, 2015 (Mike Alexander).

To measure productivity, each plot is visited in late May and early June, and the following features are marked on a study map photo:

a) nests with birds apparently incubating;
b) other complete attended nests;
c) other site-holding birds with even a trace of a nest;
d) any unattended well-built nests (empty or otherwise).

Complete and trace nests are numbered sequentially, and from early June each plot is visited every 14 days until the first chicks reach medium size, when visits increase to a minimum of every five days (Walsh *et al.* 1995, Wilkie *et al.* 2019). In 2019, the Wick 4 and Wick 5 study plots were visited 13 times between the 22nd May and 11th August (Table 1). Nest state is recorded as one of the follow categories:

a) t/N = Trace nest, where N is the number of attending adults
b) 0 = No adults and well-built empty nest
c) 1, 2 = number of adults (1 or 2) attending a well-built but empty nest.
d) i = Adults apparently incubating
e) c/x = Adult at well-built nest, contents unknown (adult not incubating)
f) c/0 = No eggs seen but adult acting as incubating on and off
g) c/? = Egg/s seen, number in clutch unknown
h) c/N = Clutch size, where N is the number of eggs
i) b/x = Chick/s seen, number of chicks unknown (adult incubating)
j) b/N = Brood size, where N is the number of chicks
k) Chick size, according to Table 2.

Categories c-k are defined as apparently occupied nests (AONs).

Table 1 Number of days between nest checks at Wick 4 and Wick 5 study plots in 2019.

Visit dates	Interval between visits / days
22/05/2019	-
17/06/2019	26
25/06/2019	8
02/07/2019	9
08/07/2019	6
13/07/2019	5
17/07/2019	4
21/07/2019	4
25/07/2019	4
29/07/2019	4
02/08/2019	4
07/08/2019	6
10/08/2019 (Wick 4), 11/08/2019 (Wick 5)	3 (Wick 4), 4 (Wick 5)

Table 2 Guide for assessing size- or age-categories of Kittiwake chicks (Walsh *et al.* 1995).

Description	Size-category
Chick completely downy	Small (S)
Downy chick, but black tips to upper wing-coverts just visible	S
Clear grey/black pattern visible on upper-side of wing, but still some down on upper-wing, and mainly downy elsewhere	Medium (M) or M/S
No down on upper-side of wings, some down elsewhere	Large (L) or M/L

No down visible, wing tips at least equal to length of tail	L
Wing tips 1-2 cm longer than tail	'Fledgable' (F)
Wing tips 3-4 cm longer than tail	Fully fledged (FF)

The Seabird Monitoring Handbook suggests chicks with 'wing tips 1-2cm longer than the tail' are 'fledgable' and 'wing tips 3-4cm longer than the tail' are 'fully fledged'. However, this level of detail is not recorded on Skomer, and instead these chicks are categorised as 'large', and all large chicks are assumed to fledge. Productivity at each plot is calculated as: total number of large chicks divided by the total number of AONs (categories c-k).

This method of monitoring is sufficient to determine productivity but precludes detailed phenological data as nests are rarely checked more frequently than once every four days. If the exact date of key events, such as nest failure, are unknown then it can be difficult to correlate failure with external variables, like weather. On the other hand, time-lapse images have the potential to make observations more frequently and therefore provide precise timings of key biological events.

Observations from time-lapse photographs

Camera set-up

In 2019 a Reconyx time-lapse camera (HF2XODG HyperFire 2 Covert IR Camera OD) was installed opposite the Wick Kittiwake colony (Fig. 2). The distance between camera and colony is in the order of 90 m, and so a zoom lens (HF2 Telephoto Lens, Nar Illum & Matched PIR) was fitted to provide 2x magnification compared to the standard lens. The cost of the camera and lens was US$609.98 (US$459.99 camera, US$149.99 lens). The camera captured images of the Wick 4 and Wick 5 study plot every hour throughout the year.

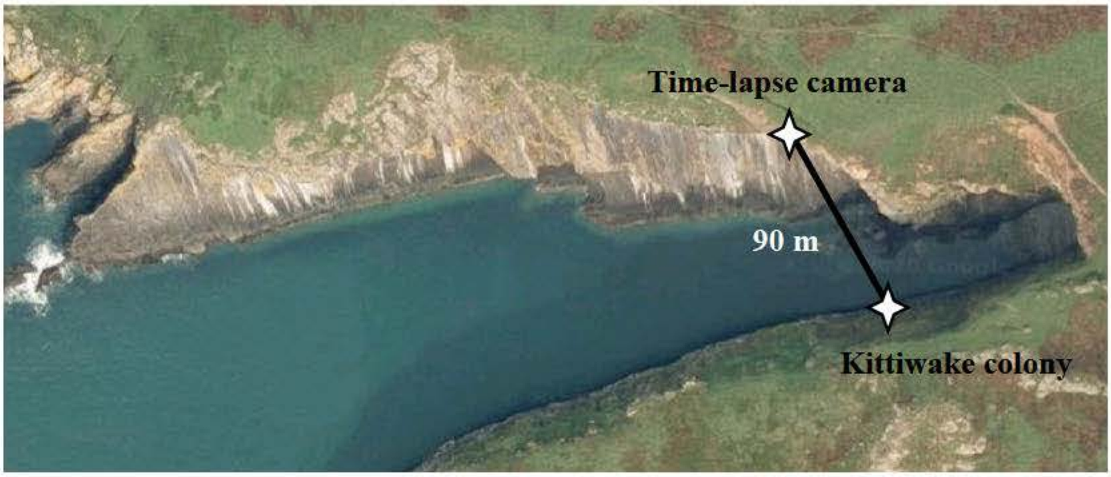


Figure 2 The time-lapse camera was positioned 90 m away from the Wick Kittiwake colony.

Monitoring

Nest monitoring from time-lapse photographs was conducted as similarly as possible to field observations on Skomer Island. Potential nest sites were marked and numbered on a photograph and each nest was followed from the 11th April, when the first photograph was taken, to the 15th August, when the breeding season was complete and Kittiwakes were not present in further images. Only the midday image was observed each day, as it was not achievable to thoroughly examine 24 photos per day in the time available. On some days the midday image was unsuitable due to poor weather conditions reducing visibility, or a bird in flight in the foreground obscuring part of the colony. In such cases the next suitable photograph either side of 12.00 was chosen. Ideally this was either the 11.00 or 13.00 image, although for a few days it was as late as 18.00. Nest state was recorded using the codes in Table 3. These were designed to match those used by the fieldworker as closely as possible, but image resolution meant eggs could not be seen and chick size could not be reliably determined. Only complete nests with an adult and/or chick present were 'apparently occupied' and included in the number of apparently occupied nests (AONs) used in the productivity calculation.

Table 3 Description of codes used to record nest state in time-lapse images.

Description	Code
One adult standing/sitting on the cliff but not occupying a trace nest or complete nest	1a
Two adults standing/sitting next to each other but not occupying a trace nest or complete nest	2a
Trace nest with no adults present	t0
Trace nest with one adult present	t1
Trace nest with two adults present	t2
Empty complete nest	n0
Complete nest with one adult present (AON)	n1
Complete nest with two adults present (AON)	n2
One adult and one chick at a complete nest (AON)	1a1c

One adult and two chicks at a complete nest (AON)	1a2c
Two adults and one chick at a complete nest (AON)	2a1c
Two adults and two chicks at a complete nest (AON)	2a2c

All image analysis was conducted before examining the field data. This meant that knowledge of nest location and outcome determined by the fieldworker did not bias results, and so the results from each method were completely independent. Likewise, all photographs were analysed prior to exploring the weather data used for survival analysis in chapter two. This helped ensure that the image analyser recorded exactly what they observed in the photographs, and not what they expected to be observed. For example, if the analyser knew a storm occurred on a particular day, then they might expect there to be fewer chicks in the photograph the following day, due to some nests being blown or washed off the cliffs. This could have biased the number of chicks 'seen' and recorded in post-storm images. Image analysis was conducted by one person (myself) to ensure consistency in observations.

Data analysis

Comparing observations in the field and from time-lapse photographs

The first step in the comparison was to match each nest monitored by the fieldworker with the corresponding nest identified in time-lapse images and assign each corresponding nest the same number. Nests only identified in the field or only identified in photos were assigned a unique number. The new numbers assigned to each nest are given in the Supplementary Material (Table S1).

Next, each code used to describe nest contents in the field had to be matched to the equivalent code used when analysing time-lapse photographs. Due to the differences between field and photo monitoring methods, all data were reassigned to a new category system, which is shown in Table 4. Inconsistencies in the field data made this a time-consuming task. For example, using a mix of capital and non-capital letters and including and not including forward dashes (e.g. B1, b1, B/1, b/1) vastly increased the number of individual codes used. Checking all of the field data for inaccuracies also identified two nests (86 and 89 at Wick 5, which were reassigned numbers 158 and 166 respectively) that had been recorded with an apparently incubating adult present (code i) yet were given an AON score of zero. The fieldworker did not count these nests as AONs because one was recorded as a trace nest, and the other was a well-built empty nest or contained one non-incubating adult on all other days. This delayed re-coding the data, as incubating should have implied that the nest went on to become an AON. Once data processing was complete, productivity and phenology determined from field and photo data were compared to assess the effectiveness of time-lapse photography as a monitoring tool at the Wick.

Table 4 Codes used when making field (field code) and photo (photo code) observations were reassigned to the same values for analysis (reassigned code). Codes that classify as apparently occupied nests (AONs) are indicated by a 'y' (yes), and those that do not are given an 'n' (no).

Field code	Field code description	Photo code	Photo code description	Reassigned code	AON (y/n)
0	No adults and well-built nest empty	n0	Empty complete nest	n0	n
1	One adult attending empty nest	n1	Complete nest with one adult present	N	y
2	Two adults attending empty nest	n2	Complete nest with two adults present	N	y
t	Trace nest	t0, t1, t2	Trace nest with zero, one or two adults present	T	n
t1	Trace nest with one adult	t1	Trace nest with no adults present	T	n
t2	Trace nest with two adults	t2	Trace nest with two adults present	T	n
i	Adults apparently incubating	n1, n2	Complete nest with one or two adults present	N	y
c/x	Adult at well-built nest, contents	n1, n2	Complete nest with one or two adults present	N	y

	unknown (adult not incubating)				
c/?	Egg/s seen, number in clutch unknown	n1, n2	Complete nest with one or two adults present	N	y
c/0	No eggs seen but adult acting as incubating on and off	n1, n2	Complete nest with one or two adults present	N	y
c/1+	Clutch of one egg minimum	n1, n2	Complete nest with one or two adults present	N	y
c/2+	Clutch of two eggs minimum	n1, n2	Complete nest with one or two adults present	N	y
c/3+	Clutch of three eggs minimum	n1, n2	Complete nest with one or two adults present	N	y
b/x	Chick/s seen, number of chicks unknown (adult incubating)	1c, 1a1c, 2a1c, 2c, 1a2c, 2a2c	One chick, one adult one chick, two adults one chick, two chicks, one adult two chicks, two adults two chicks at a complete nest	C	y
b/1+	Brood of one chick minimum	1c, 1a1c, 2a1c	One chick, one adult one chick, two adults one chick at a complete nest	C	y
b/2+	Brood of two chicks minimum	2c, 1a2c, 2a2c	Two chicks, one adult two chicks, two adults two chicks at a complete nest	C	y
(s)	Small chick	1c, 2c	One chick, two chicks	C	y
(m)	Medium chick	1c, 2c	One chick, two chicks	C	y
(l)	Large chick	1c, 2c	One chick, two chicks	C	y

Phenology

Three dates were determined for each nest site from field data and photographs: the date a nest was first recorded as an AON and the dates a chick was first and last seen. This allowed the length of the incubation and chick-rearing periods to be determined. The incubation period was defined as the number of days difference between the first time a nest was recorded as an AON and the first time a chick was seen, and the chick-rearing period was the difference between the first and last time a chick was seen at the nest. The date of first adult arrival could not be determined as Kittiwakes had already returned to the site before the first photograph was taken and first field observation was made. The average number of days difference between dates estimated from field observations and image analysis were calculated, along with the proportion of nest sites where the photo derived date was equal to, less than and greater than the field date.

Shapiro-Wilk normality tests assessed whether the differences in first AON, first chick and last chick date recorded in the field compared to photographs were normally distributed. Paired t-tests determined whether these differences were significant if the data were normally distributed, otherwise Wilcoxon signed rank tests were used.

The accuracy of methods cannot be directly assessed, as the ‘true’ state of each nest is unknown. However, predictions regarding the difference in dates between methods are possible. It was predicted that image analysis would record earlier first AON and later last chick dates, because the initiation and completion dates of time-lapse observations were earlier and later respectively than for field observations. In contrast, it was unclear how first chick dates would differ between field and photo observations. On the one hand, the increased frequency of observations provided by images (daily rather than every 3-26 days) might allow small chicks to be first seen closer to their true hatch date. Conversely, fieldworkers can wait for incubating birds to move and reveal nest contents, and behavioural information, such as incubating postures, may be easier to discern, allowing fieldworkers to spot small chicks for the first time before image analysis methods.

Productivity

Productivity was calculated as the total number of fledged chicks divided by the total number of AONs. For field observations, a chick was presumed to have fledged if it had reached ‘large’ size. Small and medium sized chicks that disappeared between visits were assumed to have died. The size of chicks was not recorded from time-lapse photographs as the images could not be viewed in sufficient detail to accurately determine chick size according to fieldworker methods. This made it difficult to determine from photographs which chicks fledged. The image data were therefore examined in greater detail and five methods were used to calculate the number of fledged chicks from photographs and thus productivity. The productivity calculated from each of these five methods was then compared with the productivity calculated from field observations to assess which method gave the most similar value to in the field. Methods deemed ‘similar enough’ to field observations, could then be used in future years to determine the number of chicks fledged from photographs and subsequently productivity if field data were not available. The methods used were as follows:

1. All chicks fledged

This assumed that all chicks seen in photographs fledged.

2. Chicks seen for a minimum of 21 days fledged.

Previous studies using time-lapse cameras have assumed that Kittiwake chicks present in the nest for a minimum of 21 days could fledge and those observed for less than 21 days do not survive (Ryan 2019, Barry 2020). This assumption was based on Coulson and White (1958) who reported a minimum fledge period of 36 days in Great Britain from field observations. Since small chicks can be very difficult to observe in the first one to two weeks after hatching due to parental brooding, it was assumed that chicks observed for more than 20 days survived, but less than 21 days was unrealistic despite allowances for variable incubation periods (Coulson & White 1958, Gabrielsen *et al.* 1992, Coulson 2011).

3. Chicks only seen on/after the 14th July fledged.

Table 5 shows the dates of key phenological events across all Kittiwake study plots on Skomer Island. In the past 14 years, the earliest date a chick was first seen was on the 8th June in 2010. Using Coulson and White’s (1958) minimum fledge period of 36 days, it would be expected that this chick fledged no earlier than the 14th July.

Therefore, based on this data, it might be assumed that only chicks still present on/after the 14th July could go on to fledge.

4. Chicks only seen on/after 21st July fledged.

The average date a chick was first observed in the past 14 years was the 15th June (Table 5). Adding 36 days onto this suggests the earliest hatched chicks were, on average, likely to fledge on the 21st July. From this, it could be assumed that only chicks seen on/after 21st July were large enough to fledge.

5. Chicks first seen on/after 1st August were not included in the number of fledged chicks.

Examination of the data showed that a chick was first seen on/after 1st August for nests 23, 26 and 102 in photographs and that these chicks were only seen once. It is most likely that the chicks were not from these nests but had instead fledged from other nests within the colony. Kittiwake chicks may leave the nest sometime after 36 days, but rarely head straight out to sea. Instead, they typically make short excursions around the colony and will attempt to land on suitable cliff sites. This means they may be found in nests other than their own (Tanedo & Hollémen 2020).

Table 5 Kittiwake phenology records from all study plots 2005 – 2018 (Taylor *et al.* 2010, Wilkie *et al.* 2019). Incubation period is: first chick – first egg. Fledge period is: first fledge – first chick.

Year	Nest building start / date	First egg / date	First chick / date	First fledge / date	Incubation period / days	Chick-rearing period /days
2005	29-Apr	18-May	11-Jun	-	24	-
2006	07-May	23-May	19-Jun	-	27	-
2007	07-May	19-May	16-Jun	-	28	-
2008	08-May	24-May	20-Jun	-	27	-
2009	30-Apr	11-May	11-Jun	-	31	-
2010	30-Apr	21-May	08-Jun	-	18	-
2011	07-May	13-May	10-Jun	-	28	-
2012	07-May	20-May	11-Jun	-	22	-
2013	10-May	28-May	23-Jun	-	26	-
2014	23-Apr	22-May	24-Jun	27-Jul	33	33

2015	04-May	19-May	20-Jun	24-Jul	32	34
2016	02-May	20-May	14-Jun	25-Jul	25	41
2017	06-May	23-May	14-Jun	25-Jul	22	41
2018	11-May	25-May	20-Jun	28-Jul	26	38
Mean	**04-May**	**20-May**	**15-Jun**	**25-Jul**	**26.4**	**37.4**

The McNemar test was used to test the null hypothesis that the relative probability of nest success did not differ when measured by a fieldworker or time-lapse camera. This test can be used when the dependent variable is dichotomous, such as success or fail. AONs that fledged one or more chicks were successful (1) while AONs that did not fledge a chick failed (0). Since the McNemar test did not take into account the number of chicks fledged per nest, paired t-tests or the non-parametric equivalent (Wilcoxon signed rank tests) were used to test the null hypothesis that the mean number of chicks fledged per AON did not differ when measured from field data compared to time-lapse photographs. Prior to this, the Shapiro-Wilk normality test determined whether the number of chicks fledged per nest was normally distributed. Paired t-tests were used if the data were normally distributed, whereas Wilcoxon signed rank tests were used if the data were not. All statistical analyses were conducted in the R environment (R Core Team 2019).

RESULTS

Individual nests

The time-lapse camera captured photographs of most of the Wick 4 and Wick 5 study plots. A total of 144 potential nests (those that at least reached trace stage) were recorded both in the field and in photographs (Fig. 3). An additional 18 potential nests were only identified in photographs: of these, 12 terminated as trace nests and six became apparently occupied nests (AONs) but did not produce a chick. Conversely, five potential nests were only recorded by the fieldworker. One of these terminated as a trace nest, two became AONs but without a chick, one successfully fledged one chick and the other successfully fledged two chicks. Out of the 144 potential nests, 129 were recorded as an AON in both field and photographs. An additional 14 nests were identified as an AON in photographs only compared to three AONs recorded only in the field (Table 6).

Table 6 The number of potential nests, trace nests and complete nests recorded in the field, in photographs and in both field and photographs. The number recorded in 'field and photo' refer to corresponding nests e.g. the same 129 AONs were recorded in both the field and photos.

Method	**Number of…**	
	Potential nests (trace & complete)	**Complete nests (AON)**
Photo	162	143
Field	149	132
Field and Photo	144	129

Phenology

The date trace nests were first recorded could not be compared between field and photographs, as some nests were already complete when the fieldworker made the first visit on the 22nd May. Figures 4-6 show the date a nest was first recorded as complete and a chick was first and last seen, for each complete nest observed in the field and/or in photographs. The first date a nest was complete was on average 3.7 ($\sigma = 8.2$) days earlier in photographs than in the field. The date a chick was first and last recorded was 13.9 ($\sigma = 7.6$) and 9.6 ($\sigma = 5.7$) days later respectively in photographs compared to field observations (Table 7, Fig. 7). In all cases, these differences in date were significant when comparing the 129 complete nests observed in both field and photos (first nest: Wilcoxon signed rank test, $V = 4210$, $P < 0.001$; first chick: Paired t-test, $t = -18.1$, $P < 0.001$; last chick: Wilcoxon signed rank test, $V = 28.5$, $P < 0.001$).

The duration of the egg stage, defined as the period between the date on which a nest was completed and a chick was first seen, was on average 16.8 ($\sigma = 10.5$) days longer when calculated from photographs compared to field data. Conversely, the duration of the chick stage, defined as the period between the date a chick was first and last seen, was 4.3 ($\sigma = 9.4$) days shorter when measured from photographs (Table 7, Fig. 7). Again, these differences were statistically significant (length of egg stage: Paired t-test, $t = -15.8$, $P < 0.001$; length of chick stage: Paired t-test, $t = 4.51$, $P < 0.001$).

Table 7 Comparison of key dates obtained from 129 complete nests observed in both the field and photographs at the Wick, Skomer, 2019. ‘Average difference’ is the mean of all differences between field and photograph dates for each nest. Each difference was calculated as field date – photograph date, meaning negative values show that dates from photographs were on average later than in the field, while positive values show that dates from photographs were earlier. σ is the standard deviation of the average difference. ‘Proportion of sites where photo = field’ is the proportion of nests where the date was the same in the field and image. ‘Proportion of sites where photo > field’ and ‘photo < field’ is the proportion of nests where the date was greater (i.e. later) in images than the field and smaller (i.e. earlier) in images than in the field, respectively.

Date of	**Average difference**	**σ**	**Proportion of nests where photo = field**	**Proportion of nests where photo > field**	**Proportion of nests where photo < field**
Complete nest first recorded	3.71	8.16	0.20	0.20	0.60
Chick first recorded	-13.90	7.61	0.01	0.93	0.06
Chick last recorded	-9.63	5.70	0.071	0.90	0.03
Length of egg stage	-16.82	10.52	0.03	0.94	0.03
Length of chick stage	4.27	9.36	0.03	0.36	0.36

Figure 3 Time-lapse photograph with all potential (trace & complete) nest sites labelled. Purple circles are nests recorded in the field and in time-lapse images (n = 144). Cyan triangles are nests recorded in time-lapse images only (n = 18). Orange diamonds are nests recorded in the field only (n = 5).

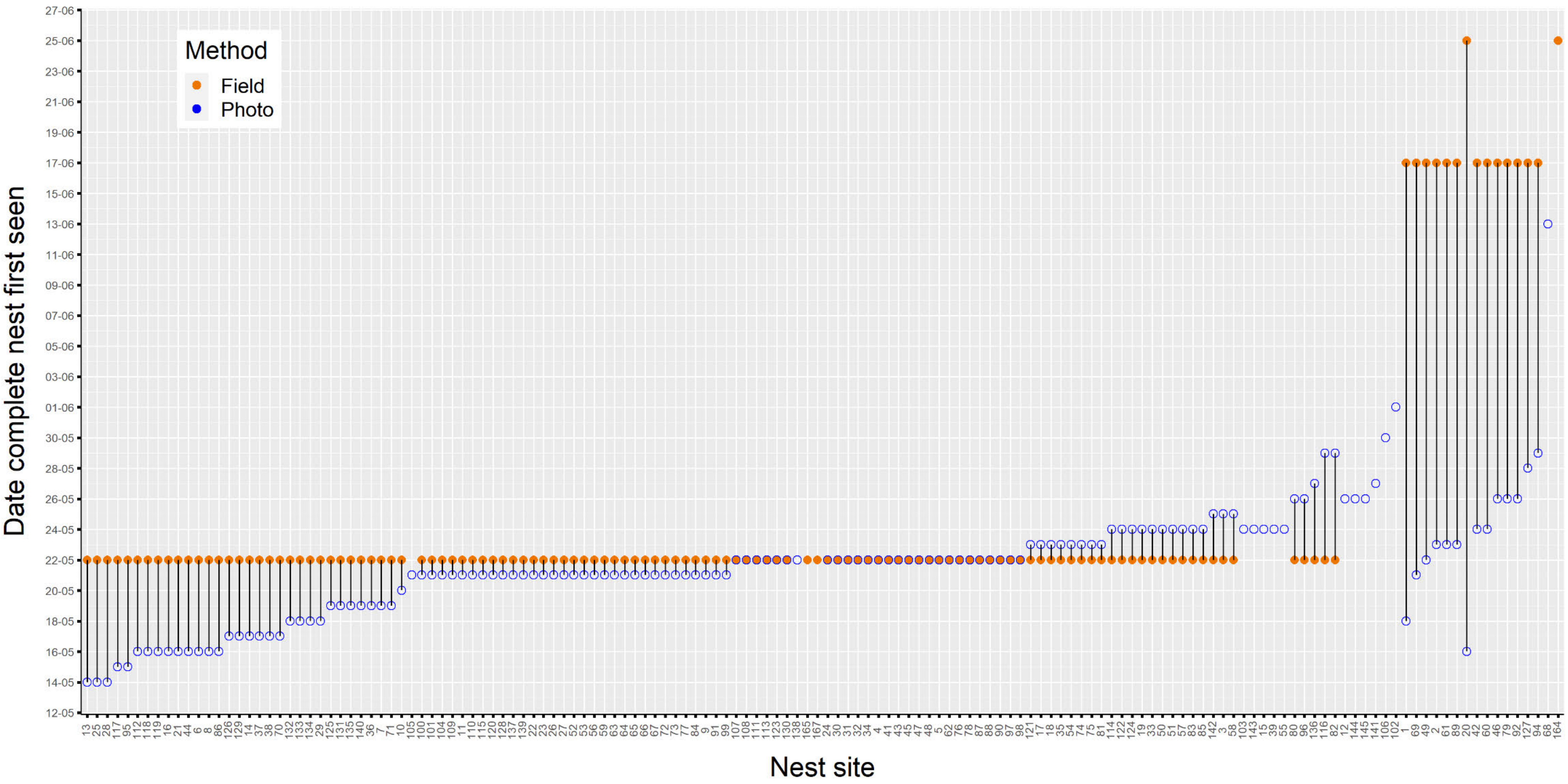


Figure 4 Difference in date a complete nest was first seen between field and photograph observations for each complete nest recorded in the field and/or photographs. The infrequency of site visits early in the season meant the fieldworker recorded most nests as complete on the 22nd May.

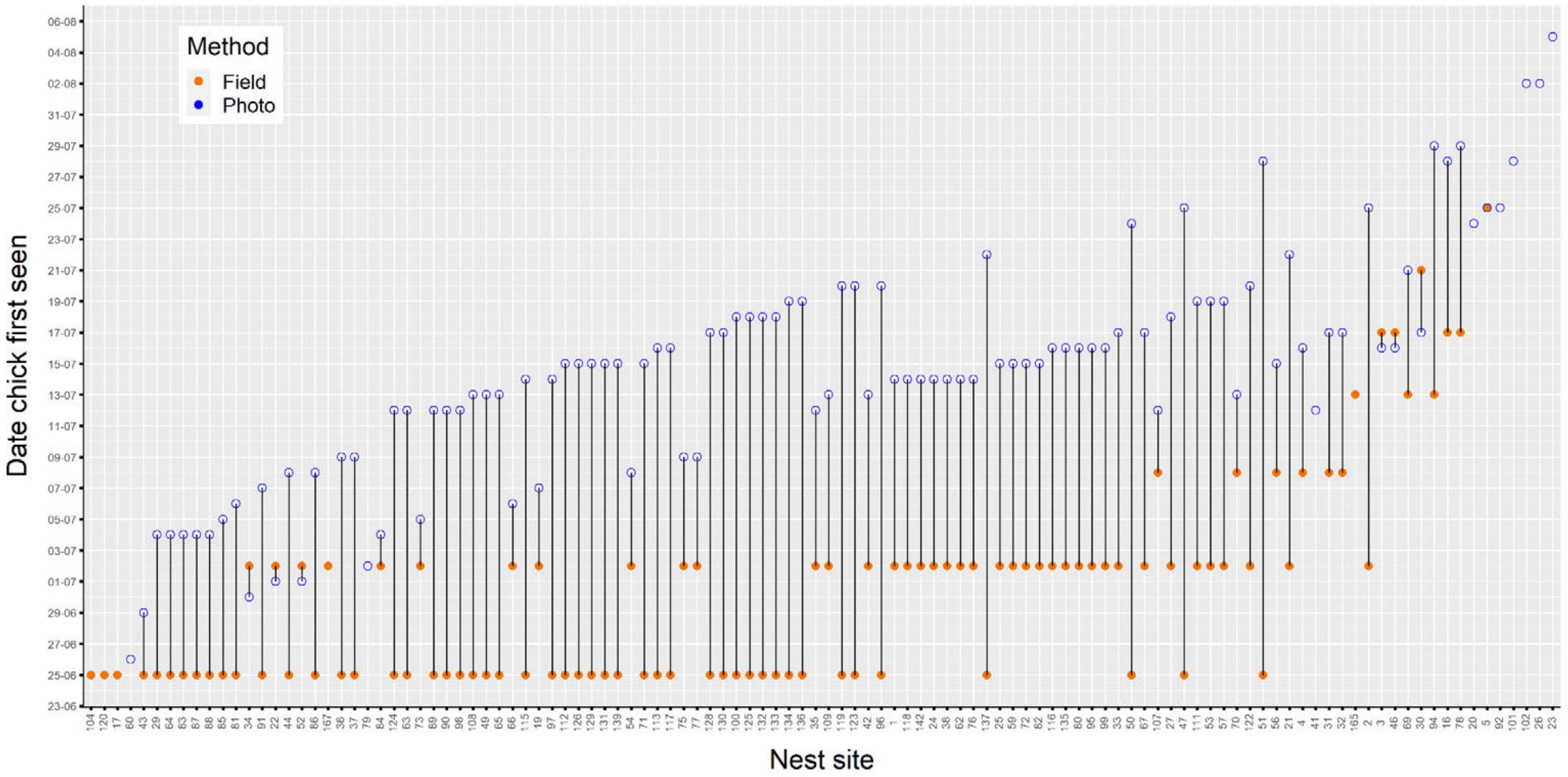


Figure 5 Difference in date a chick was first seen between field and photograph observations for each complete nest recorded in the field and/or photographs.

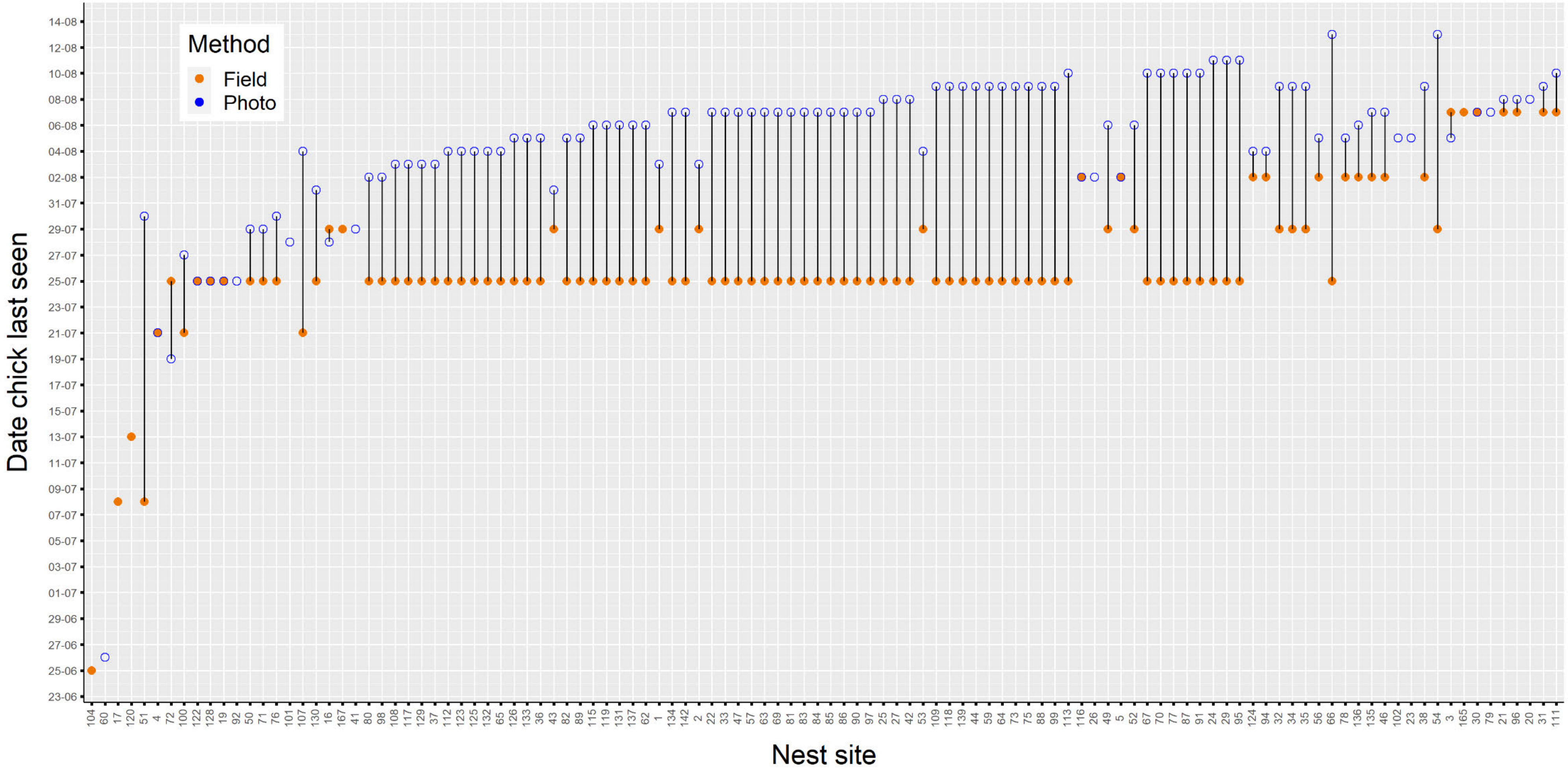


Figure 6 Difference in date a chick was last seen between field and photograph observations for each complete nest recorded in the field and/or photographs.

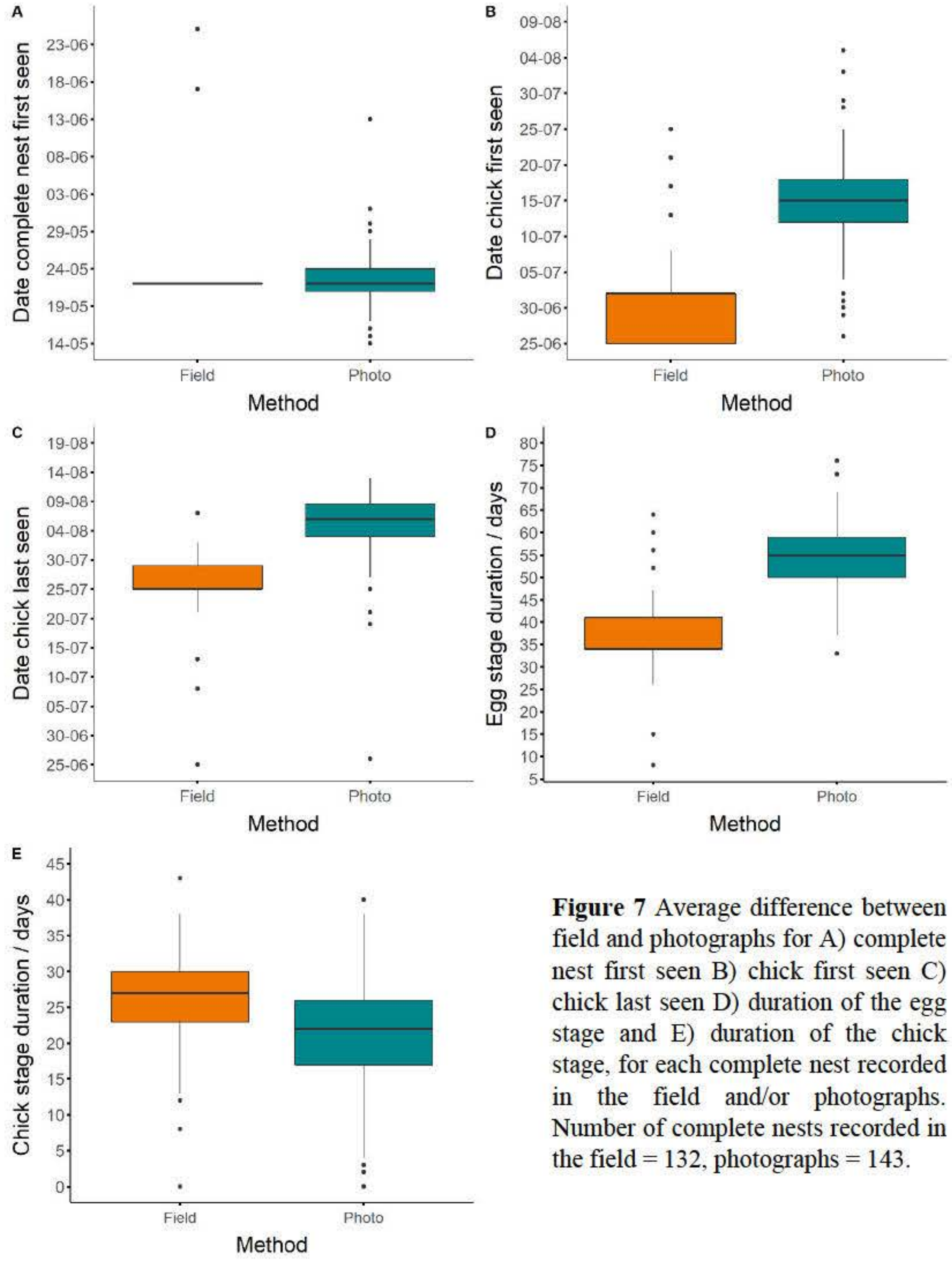


Figure 7 Average difference between field and photographs for A) complete nest first seen B) chick first seen C) chick last seen D) duration of the egg stage and E) duration of the chick stage, for each complete nest recorded in the field and/or photographs. Number of complete nests recorded in the field = 132, photographs = 143.

Productivity

Productivity was calculated as the total number of chicks that fledged divided by the total number of AONs. 143 AONs were recorded in time-lapse photographs and 132 were recorded in the field. The same 129 AONs were seen in both photographs and in the field. At least one chick was recorded in 107 out of 143 AONs (74.8%) seen in photographs, whereas 103 out of 132 AONs (78%) recorded a chick(s) in the field.

Productivity was calculated from photographs using five different methods to determine which gave the closest estimate to field observations (Table 8). Productivity was generally very similar (0.80 - 0.83) to that calculated by the fieldworker, who assumed that only large chicks fledged (0.83). When the 'minimum fledge period' of 21 days was used to determine the number of chicks that fledged, productivity dropped by 0.16 (19%) in the field and 0.4 (48%) in photos, compared to when only large chicks fledged. However, comparing the number of chicks seen per nest in the field and in photographs revealed numerous differences, despite apparently similar productivity values. These nest by nest differences were explored further to assess the reliability of each productivity estimate.

Table 8 Productivity estimates from field and photo data using five different methods to determine the number of fledged chicks.

Field		**Photo**	
Fledged chicks	**Productivity**	**Fledged chicks**	**Productivity**
Only large chicks	109 / 132 = 0.83	All chicks	119 / 143 = 0.83
Chicks seen for minimum of 21 days	89 / 132 = 0.67	Chicks seen for minimum of 21 days	62 / 143 = 0.43
Medium, medium/large and large chicks (i.e. chicks last seen on/after 14th July)	126 / 132 = 0.95	Only chicks last seen on/after 14th July (i.e. excludes nest 60)	118 / 143 = 0.83
		Only chicks last seen on/after 21st July (i.e. excludes nest 60 and 72)	117 / 143 = 0.82

		Only chicks last seen on/after 14th July (i.e. excludes nest 60) **AND** first seen on/before 31st July (i.e. excludes nest 23, 26 and 102)	115 / 143 = 0.81
		Only chicks last seen on/after 21st July (i.e. excludes nest 60 and 72) **AND** first seen on/before 31st July (i.e. excludes nest 23, 26 and 102)	114 / 143 = 0.80

Number of fledged chicks

Out of the 144 potential nests observed in both the field and in photographs, the same number of chicks (including zero) were recorded in the field and in images for 89 nests (62%). Of the 55 nests (38%) where there was a difference in the number of chicks seen, photographs were more likely to observe fewer chicks per nest compared to field observations (31%) (Fig. 8A). When comparing the number of chicks per nest in photographs, with the number of *large* chicks per nest (i.e. those that fledged) in the field, only 39 nests (27%) had a difference in the number of chicks (Fig. 8B).

Compared to Fig. 8A, Fig. 8B shows more nests where chick number was equal in photos and in field (photo = field) and greater in photos (photo > field), but fewer nests where there were less chicks observed in photos (photo < field). For 42 nests (29%), the fieldworker recorded two chicks when only one was observed from photographs; however, for 22 of these, one chick did not reach large size, and for 6 nests both chicks did not reach large size. Failure of one or both chicks in two chick broods, would have increased the proportion of nests where the number of large chicks in the field equalled (photo = field) and was less than (photo > field) the number of chicks in photos (Fig. 8B), relative to when comparing the number of chicks seen in the field and in photos irrespective of size (Fig. 8A). Table S3 in the Supplementary Material explains in further detail why there may have been discrepancies in the number of chicks fledged per nest from field and photo data. The implications of this for productivity estimates will be explored further in the discussion.

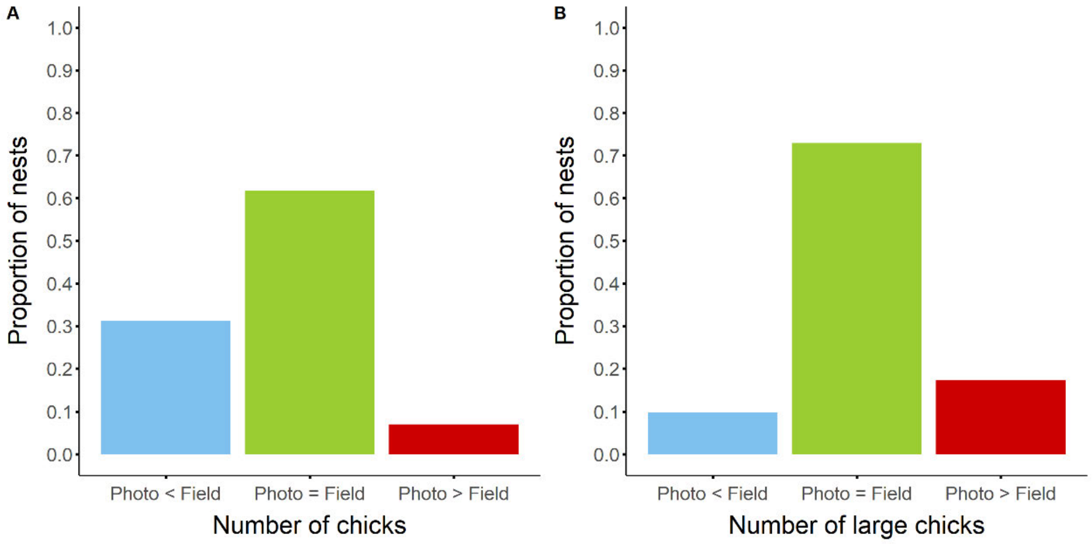


Figure 8 Proportion nests where the number of chicks seen in photographs was less than, equal to and greater than the number of A) chicks and B) large chicks seen in the field. n = 144 potential nests seen in both the field and photographs.

The mean number of chicks fledged per AON did not significantly differ between field and photo data when only large chicks fledged in the field and 1) all chicks, 2) chicks last seen after 13th July, or 3) chicks last seen after 20th July, fledged from photographs (Table 9). Conversely, there was a significant difference in the number of chicks fledged per AON when large chicks fledged from the field, but only 1) chicks last seen after 13th July and first seen before 1st August, or 2) last seen after 21st July and first seen before 1st August, fledged from photographs (Table 9). Equally, there was a significant difference in the number of chicks seen per nest, which is equivalent to assuming all chicks fledged from every nest seen in the field and in photographs. These results support Fig. 8 and Table S3 in the Supplementary Material, which suggest there were numerous discrepancies in the number of chicks seen per nest in field and photos. However, depending on the criteria used to determine fledging, the number of chicks fledged per nest and thus productivity could be very similar.

Table 9 Results of statistical analyses used to test the null hypothesis: the mean number of chicks fledged per AON does not differ when measured by a fieldworker compared to time-lapse photographs. The number of chicks fledged per AON for field and photo data are given in the comparison column. All Shapiro-Wilk normality tests were significant meaning the data were normally distributed and Wilcoxon signed rank tests were used to test the null hypothesis. n = 129 AONs seen in both the field and photographs.

Comparison		**Shapiro-Wilk normality test**	**Wilcoxon signed rank test**
Field	**Photo**		
Only large chicks fledge	All chicks fledge	W = 0.735, P < 0.001	V = 273, P = 0.0652
Only large chicks fledge	Only chicks last seen on/after 14th July fledge (i.e. excludes nest 60)	W = 0.729, P < 0.001	V = 266, P = 0.0873
Only large chicks fledge	Only chicks last seen on/after 21st July fledge (i.e. excludes nest 60 and 72)	W = 0.723, P < 0.001	V = 259, P = 0.116
Only large chicks fledge	Only chicks last seen on/after 14th July fledge (i.e. excludes nest 60) **AND** first seen on/before 31st July (i.e. excludes nest 23, 26 and 102)	W = 0.716, P < 0.001	V = 252, P = 0.152
Only large chicks fledge	Only chicks last seen on/after 21st July fledge (i.e. excludes nest 60 and 72) **AND** first seen on/before 31st July (i.e. excludes nest 23, 26 and 102)	W = 0.709, P < 0.001	V = 245, P = 0.198

All chicks	All chicks	W = 0.775, P < 0.001	V = 1265, P < 0.001

Time investment

The time required to analyse each image was between 20 and 30 minutes. This included identifying the nest contents of each pre-labelled nest and entering it into a spreadsheet. In total, 127 images were examined, meaning image analysis took approximately 42-64 hours. Additional time was needed pre-image analysis to extract the midday image for each day, and mark and number potential nest sites on a photograph. Identifying corresponding nests seen in both the field and in photographs and extracting information on phenology and productivity also required further time post-image analysis.

In comparison, the time taken by each field visit varied across the breeding season. On average, each visit took around 30 to 40 minutes per plot mid-season, when the chicks were mostly small/medium size, although later visits became much shorter in length as the fate of more and more nests were determined (i.e. either failure of fledging) (A. Piggott, pers. comm) . Assuming visits took 30 minutes and thirteen visits were made to each of the Wick 4 and Wick 5 study plots in 2019, then fieldwork took around 13 hours total. Extra time was spent inputting data into a spreadsheet and extracting information on phenology and productivity.

DISCUSSION

This chapter aimed to investigate how measurements of Kittiwake phenology and productivity from time-lapse photographs compared with field observations in 2019 on Skomer Island. Overall, measurements from photographs first recorded nests as ‘apparently occupied’ earlier than field observations, but chicks were observed for the first and last time later in photographs compared to in the field. Productivity measured from photographs was variable, ranging from 0.46 to 0.83, depending on the method used to determine the number of chicks that fledged. The highest productivity calculated from image analysis was the same as that calculated from field observations (0.83).

Individual nests

The same 144 potential nests (trace and complete) and 129 apparently occupied nests (AONs) were recorded by both field and photo analysis. Field and photographic methods each recorded additional nests not seen using the other method; however, in total, photographs recorded 13 more potential nests and 11 more AONs compared to field observations. Having ~150 nests per image meant most nest sites appeared to be small in photographs. Provisional examination of the time-lapse images showed that adult and juvenile Kittiwakes could be identified, although with reduced clarity than seen by the fieldworker, who was able to use a scope to zoom in further on the nests being monitored. It was therefore surprising that more potential nests and AONs were identified in photographs given the lower image resolution. One possible explanation is that some or all of the 18 trace nests seen only in images were abandoned before the fieldworker made their first visit at the end of May, and did not remain sufficiently intact to be recorded as a trace nest upon the fieldworker's arrival. Alternatively, these trace nests could have been misidentified in photographs, and been marks on the rock rather nesting material.

Phenology

The dates of key phenological events during the Kittiwake breeding season were significantly different when measured from time-lapse photographs compared to field observations. Complete nests were first seen earlier in images, while chicks were first and last seen later. Earlier first complete nest dates and later first chick dates resulted in the length of the egg stage being significantly longer when measured in photographs, while the length of the chick stage was significantly shorter.

Earlier first complete nests were observed in photographs because time-lapse images started being recorded before the fieldworker made their first observation. The first day a complete set of images were collected (i.e. one every hour) took place on the 11th April, whereas the fieldworker made their first visit on the 22nd May. This meant some nests were already completed when the fieldworker arrived. Equally, the fieldworker finished their observations on the 11th August, while the time-lapse camera records year-round. Kittiwakes were last observed in images on the 15th August, meaning some chicks from late nests may have still been present after the fieldworkers' last visit, resulting in the date a chick was last seen being later in images.

These differences highlight how time-lapse cameras can expand the temporal scale of monitoring, as they can record images throughout the year when fieldworkers may not be present at sites. This includes early and late in the breeding season, but also in winter, which can give an insight into colony attendance outside of breeding (Mudge *et al.* 1987, Black *et al.* 2017, Black *et al.* 2018b). However, it is important to consider whether the last chicks seen in time-lapse images were from the nest they were observed at. Juvenile Kittiwakes often make short flights around the colony sometime after 36 days post-hatching, but do not typically leave the colony until ~10 days after their first flight (Coulson 2011). While their parents will only feed them at their own nest, the juveniles may land at other empty nest sites, meaning offspring are increasingly likely to be seen at nests other than their own towards the end of the breeding season (Tanedo & Hollémen 2020). Consequently, without marking chicks it is difficult to be sure whether large chicks capable of fledging are from the nest they are observed at.

The later first chick dates recorded in photographs compared to in the field suggest that small chicks were harder to spot in images. Chicks are typically brooded continuously by a parent for the first 15-16 days after hatching, which can make them hard to see in a single static image (Gabrielsen *et al.* 1992). Analysis of photographs likely struggled to observe these young chicks for several reasons. Firstly, the time-lapse camera was positioned 90 m from the colony, which meant image resolution was fairly low. In the order of 150 nests were observable per image, which is greater than many other studies using time-lapse cameras (Southwell & Emmerson 2015, Hinke *et al.* 2018, Ryan 2019, Barry 2020). Increasing the number of nests per image (i.e. sample size) may give a more representative measure of colony productivity but will reduce the detail seen at each nest and therefore accuracy of observations. Conversely, while the fieldworker viewed the Wick Kittiwake colony from the same distance as the time-lapse camera, the use of a scope with a zoom lens allowing up to 50x magnification meant each nest could be viewed at much higher resolution compared to the 2x magnification afforded by the camera. This allowed the size of chicks to be established and small chicks were more likely to be spotted. Furthermore, field observations may be better able to detect chicks as the observer can spend time watching each nest, waiting for the adult and/or potential chick to move and allow confirmation of chick presence or absence (Tanedo & Hollémen 2020). The precise timing of imaging means such behaviour may be missed and so chicks are only likely to be seen once they are large enough to no longer need brooding.

The differences in dates between fieldworker and photo estimation, led to the length of the egg stage and chick stage being significantly longer and shorter respectively when measured in photographs compared to in the field. While this is fairly intuitive given previous explanations for the difference in dates, Fig. 4-6 suggest there may have also been a temporal aspect to the disparity between field and photo observations. For example, the disparity between first AON date appears to be greater for AONs completed earlier in the season compared to AONs completed later in the season. This needs further investigation, such as by doing a regression of disparity against time (e.g. days from first arrival).

Having accurate measurements of phenology is important for understanding the effect of environmental change on breeding. Successful reproduction depends on the seasonal peak in food availability matching the food needed to raise offspring (Moe *et al.* 2009). Changes in the timing of peak food availability associated with global climate change are therefore predicted to alter the timing of breeding in order to avoid mismatch (Walther *et al.* 2002, Frederiksen *et al.* 2004, Moe *et al.* 2009). Several studies have reported Kittiwakes breeding later, as measured by later first egg (i.e. egg-laying) dates (Frederiksen *et al.* 2004, Wanless *et al.* 2009) and hatch dates (Moe *et al.* 2009) across the time periods measured. Frederiksen *et al.* (2004) reported Kittiwake first egg dates becoming 5.1 days later per decade on the Isle of May, Scotland. These changes in Kittiwake phenology have been negatively associated with sea surface temperature (SST) (Frederiksen *et al.* 2004, Moe *et al.* 2009) and the North Atlantic Oscillation (NAO) index (Frederikesen *et al.* 2004, Wanless *et al.* 2009), meaning breeding occurred earlier in years with mild, wet winters (high NAO) and high winter (February-March) and spring (April-May) temperatures. Years with late breeding have been associated with low clutch size and mean annual breeding success at some locations, suggesting poor investment and food availability (Moe *et al.* 2009).

It therefore seems apparent that knowledge of only one phenological parameter, namely first egg or hatch dates, is required to understand the relationship between environmental change, phenology and breeding success. Time-lapse cameras may struggle to precisely determine chick hatch dates but are useful for determining when nests are finished being built. This is presumed to be a reasonable proxy for first egg dates, as Kittiwakes typically lay their first egg within one day of nest completion and the gap between laying subsequent eggs is normally no more than two days (Coulson 2011). Obtaining accurate chick fledge dates could also be useful to assess whether changing environmental variables are affecting breeding period duration, defined as the time between egg laying and chick fledging. Time-lapse cameras may provide more accurate fledge dates than fieldworkers when site visits finish before the final chicks have fledged. Accurate phenology measurements are further important for obtaining productivity estimates.

Productivity

Productivity calculated from photographs varied from 0.43 to 0.83 depending on the conditions used to determine the number of chicks that fledged. The productivity calculated by the fieldworker, assuming only large chicks fledged, was 0.83.

The greatest difference in productivity occurred when chicks seen for a minimum of 21 days were deemed to have fledged in photographs (0.43), compared to large chicks fledging in the field (0.83). Field productivity was also reduced (0.67) when re-calculated using this minimum fledge period. This suggests that overall, the 21 day minimum fledge time underestimates the number of chicks that fledged and lowers productivity estimates. Using a minimum fledge period to determine which chicks fledge relies on accurate hatch dates being obtained, as chicks are considered fledged after some number of days post-hatching (Tanedo & Hollémen 2020). Based on the *a priori* assumption that small chicks brooded by parents would be difficult to observe in photographs, leading to recorded hatch dates being later than the true hatch dates, the minimum fledge period chosen was 15 days less than the 36 days observed by Coulson and White (1958) in Great Britain (Ryan 2019, Barry 2020). Indeed, this assumption was supported, as phenology comparison showed that on average, the date a chick was first seen was 13.6 days later from photos than in the field. It is therefore surprising that productivity estimates from images were considerably reduced when using a 21 day fledge period, relative to field observations, as this means a large number of chicks were not even observed for 21 days. In future, it would be advised to try an alternative approach of adding 26 days incubation period and 36 days fledge period to the date of nest completion, which approximates egg-laying, to give a likely fledge date (Coulson 2011). Chicks last seen before their likely fledge date could be assumed to have failed. This would hopefully avoid the problem of not observing small chicks due to obscuration, for example by parents, rocks, shadows, or another chick in the nest.

All other methods of determining the number of chicks fledged from photographs gave similar productivity estimates to field data. The 'best' method was when all chicks fledged from photographs, as this gave the same productivity estimate as in the field: 0.83. However, when comparing the number of chicks per nest in photographs, with the number of chicks (irrespective of size) and number of large chicks per nest in the field, it became evident that fewer chicks were observed per nest in photographs. In 50% of cases where the fieldworker observed two chicks but only one was recorded from photos, the second chick did not reach large size. Second chicks may have died before they were large enough to be seen in images, meaning the number of chicks seen (and therefore fledged) in photos still equalled the number of large chicks in the field for these nests, despite the number of eggs hatched per nest being different. Differences like this would not have altered productivity values, explaining why productivity was so similar, regardless of the differences in the number of chicks seen per nest between field and photo data.

Nonetheless, knowing the number of chicks that fledged relative to the number of eggs that hatched is important for understanding other aspects of Kittiwake population dynamics, including factors influencing daily survival rate, such as food availability and weather. For example, loss of one chick in two chick broods could suggest insufficient food was available to raise both chicks (Benowitz-Fredericks *et al.* 2013). Of course, this insight would only be possible if monitoring recorded both eggs hatching but just one chick fledging. In this study, there were several nest sites where chicks were only recorded by the fieldworker and not in images. At nest 17, the fieldworker observed a small chick on the 8th July, but it was not seen in subsequent field visits, and image analysis recorded zero chicks. This supports the conclusion that small chicks were less likely to be spotted in photographs and demonstrates how only using time-lapse data from the Wick for survival analysis could incorrectly lower the number of failed chicks, despite colony level productivity being unaffected. This will be discussed further in chapter two. On the other hand, with sufficient image resolution, time-lapse imagery could increase detection of failed nests and improve the accuracy with which nest fate is determined. For example, fieldworkers observed an apparently incubating adult at nest 60 on the 17th June but did not record evidence of breeding on the next two checks (25th June and 2nd July) or any subsequent visits.

Conversely, one chick was seen at nest 60 in the 12.00 image on the 26th June yet was not seen again. This suggests the chick hatched, but also died, in between field observations on the 25th June and 2nd July. More frequent observations provided by time-lapse cameras could therefore increase the probability of recording chicks that failed shortly after hatching, provided the images have high enough resolution to observe small chicks.

Assuming chicks only fledged if they were seen for the last time on/after the 14th or 21st July and/or first seen on/before the 31st July, lowered productivity slightly (by up to 0.03) compared to field calculations. Using these thresholds increased the plausibility of productivity estimates, as chicks seen for the last time before the 14th or 21st July were more likely to have died than fledged, and those seen for the first time after the 31st July may have been chicks fledged from other nests. The number of chicks fledged per nest in photographs did not significantly differ from field observations when fledged chicks were last seen on/after the 14th or 21st July. However, the difference was significant when fledged chicks must have been first seen on/before 31st July, as well as last seen on/after the 14th or 21st July in images. This highlights the importance of accurately determining which chicks fledge, as the addition of one extra rule (first seen on/before 31st July) changed the difference between field and photo data from non-significant to significant. Selecting nests that are most visible, rather than trying to obtain data for every nest, could increase the probability of observing chicks and therefore obtaining accurate dates for key events. It would also reduce the amount of time spent analysing images if fewer nests were studied.

Time investment

Manual image analysis of time-lapse photographs took longer than field observations in 2019. In total, 42-64 hours were spent examining 127 images, compared to approximately 13 hours in the field. The large amount of time required to manually examine digital imagery has often limited the use of time-lapse cameras for seabird monitoring (De Pascalis *et al.* 2018).

In this study, one image was examined per day, despite the camera being programmed to take one image every hour. More frequent sampling might have allowed the metrics quantified from image observations to be closer to their true values. For example, it would be expected that as the number of images examined increased (i.e. the time between sampled images decreased), chicks would first be seen closer to their true hatch date. It is recommended that future studies using time-lapse cameras conduct a stand-alone assessment of the optimum number of images needed to quantify metrics of interest, in this case phenology and productivity. This should create an asymptotic view, showing the measured variable plateauing to a particular value as the number of images sampled increases, thus allowing the optimal number of images to be determined (Tanedo & Hollémen 2020).

Of course, sampling more images takes more time, and there is an evident trade-off between accuracy and effort. Outsourcing image analysis to paid employees or volunteer citizen scientists represent possible solutions to this problem (Arteta *et al.* 2016). Increasingly, many researchers are developing machine learning algorithms to identify birds in images and allow the scale of time-lapse monitoring to be expanded (Edney & Wood 2020).

Summary and future considerations

Comparison of phenology and productivity measurements from field and time-lapse camera data at the Wick, Skomer, has shown that photographs provide similar results to field observations and in some cases may be more accurate, but are currently unlikely to improve present monitoring methods. Analysis of photographs likely recorded more accurate dates for first date of nest completion and possibly date of last chick seen, but recorded later hatch dates and often only recorded one chick in two chick broods. These last two differences led to variable productivity estimates, especially when fledging was determined by chick age.

One of the main factors affecting the accuracy of identifying target reproductive stages and behaviours from photos and videos, is camera setup (Lorentzen *et al.* 2010). Increasing the horizontal distance between camera and colony increases the number of birds viewed per image, which should give measurements that are more representative of the entire colony but will lower image resolution. The ‘optimal’ distance is also dependent on a number of other factors, including the purpose of the study, topography, species and density. Typically, around 20 – 40 nests may be reliably monitored to measure breeding success (Southwell & Emmerson 2015, Hinke *et al.* 2018, Ryan 2019, Barry 2020), although for time-intensive studies requiring high temporal resolution, such as investigating foraging or incubation behaviour, fewer than five nests may be required (De Pascalis *et al.* 2018, Collins *et al.* 2014). Furthermore, vertical distance and camera angle is important for identifying nest contents, with cameras positioned above the colony and facing downwards being better able to see into nests and observe small chicks (Lorentzen *et al.* 2010, Merkel *et al.* 2016).

Despite field observations of the Wick colony taking place from the same distance as the time-lapse camera, photographs were unable to view nests at the same resolution as the fieldworker, who used a scope with a 25-50x magnification eyepiece lens to get a better view of nest contents. In comparison, the time-lapse camera was fitted with a 2x magnification zoom lens. The time-lapse camera cannot be positioned any closer to the Kittiwake colony, as the camera is on one side of a gully, and thus separated from the colony by an inlet of ocean. In order to improve image resolution, a different type of camera and/or lens with greater magnification would be needed to allow accurate identification of reproductive state. Merkel *et al.* (2016) used a DSLR Canon camera fitted with a 15-85mm or 18-105 mm Canon zoom lens, set at 60 mm and 105 mm respectively, to estimate breeding success of Brünnich's Guillemot *Uria lomvia.* This setup allowed identification of Guillemot chicks, which can be hard to spot as they are usually sheltered between the adults and the cliff-face. DSLR cameras should provide higher quality images than Reconyx time-lapse cameras but are more expensive and require heftier batteries. Using a better zoom lens could also improve phenology and productivity measurements, although optical zoom is preferred over digital zoom, as the latter reduces image quality (Lorentzen *et al.* 2010). Nonetheless, increased magnification reduces the number of nests monitored by the camera, meaning measurements may not be representative of the entire colony (Merkel *et al.* 2016, De Pascalis *et al.* 2018). This trade-off between sample size and resolution needs to be considered before using time-lapse cameras to monitor seabird colonies.

Time-lapse cameras have the advantage of recording observations more frequently than fieldworkers, and so it is expected that with improved image resolution, the Wick camera would record the timing of key events more precisely than field observations. Less frequent photo sampling significantly delayed detection of nest, incubation and hatch initiation in Kittiwakes in Resurrection Bay, Gulf of Alaska, compared to daily observations (Tanedo & Hollémen 2020). This discrepancy increased with greater intervals between sampling, as many of these events occurred between sampling points (Tanedo & Hollémen 2020). Accurate determination of breeding phenology is important for understanding how seabirds are responding to changes in their environment, which is in turn important for understanding long-term population trends (Tanedo & Hollémen 2020). Kittiwakes may alter the timing of breeding by as little as 0.58 - 0.88 days per year relative to sea surface temperatures, and so these small changes in breeding phenology could be lost in field observations spaced five to 14 days apart (Frederiksen *et al.* 2004, Byrd *et al.* 2008, Tanedo & Hollémen 2020). In addition, accurately recording breeding phenology is important for estimating productivity, especially when chick age is used to determine fledging (Tanedo & Hollémen 2020).

Having appropriate image resolution to accurately record nest state and the timing of key events is also essential if researchers intend to make the most of the high temporal resolution afforded by time-lapse cameras. For example, with increased image frequency, photographs can be used to re-sight marked birds to determine adult survival, nest activity (such as nest attendance and division of labour between parents) and foraging behaviour (Weller & Derksen 1972, Black *et al.* 2017, Black *et al.* 2018b, De Pascalis *et al.* 2018). Recording foraging duration from the timing of parental switch-overs could provide a non-invasive and time efficient method of indirectly assessing food availability, as adults are predicted to travel further and therefore spend longer foraging when food is scarce (Chivers *et al.* 2012, De Pascalis *et al.* 2018). Furthermore, the ability to observe each nest more often might increase the likelihood of capturing infrequent events not purposefully monitored (Harris 1982, Black *et al.* 2017, Black *et al.* 2018b). For example, Kittiwake chick predation by a Peregrine Falcon *Falco peregrinus* was recorded from time-lapse images on Puffin Island, Wales (Collins *et al.* 2014). Predation may be under-recorded by fieldworkers as human presence could deter predatory activity. Consequently, the high temporal resolution possible with time-lapse cameras may offer further insight into the causes of population decline, provided the images are of sufficient quality to record the measurements of interest. Increasing image resolution at the Wick is therefore a priority if we intend to use time-lapse cameras to help properly understand the causes of Kittiwake decline on Skomer.

Improving image quality would have the further benefit of making the photographs more suitable for citizen science and machine learning. Adult and juvenile Kittiwakes from time-lapse images at Protheroe's Dock, Skomer, are identified by citizens participating in the *Seabird Watch* project. This reduces the workload on researchers and is essential if photographs are recorded more frequently (e.g. every five minutes) to measure variables requiring high temporal resolution, like foraging duration. Moreover, with improved image resolution machine learning algorithms will be more likely to accurately identify Kittiwakes in photographs. The *Pengbot* algorithm, developed by the *Penguin Watch* team, automatically identifies and counts penguins in time-lapse images, and a similar tool is currently being developed for *Seabird Watch* (Jones *et al.* 2020, T. Hart pers. comm.).

CONCLUSION

This study has shown that cameras can provide important phenological and productivity information, with the potential to give more accurate data in some cases than can be delivered by infrequent repeat visits throughout the season. Such information will be crucial in understanding seabird responses to increasing anthropogenic pressures, like climate change and fisheries interactions. Effective camera placement and appropriate image resolution is, however, required to fully reach this potential. Understanding the different biases and detection abilities between field and time-lapse methods during the timeline of breeding is key to scaling up the use of cameras to non-fieldworker sites. It is hoped that such networks of time-lapse cameras will provide cost-effective means of large-scale seabird monitoring; a feat especially important when money is limited for conservation endeavours (Waldron *et al.* 2013).

SUPPLEMENTARY MATERIAL

Table S1 Nests were assigned different numbers in the field (field number) and from time-lapse photographs (photograph number). To allow comparison between the same nests observed in the field and from images, field numbers were re-assigned to match photograph numbers (field assigned number). Nests only identified in the field or only identified in photos were assigned a unique number. 'na' means the nest was not recorded in either the field or from photographs. An explanation of the reassigned number has been provided where necessary.

Field site	Field number	Photograph number	Field reassigned number

wick 5	11	1	1
wick 5	12	4	4
wick 5	14	12	12
wick 5	15	2	2
wick 5	16	3	3
wick 5	17	13	13
wick 5	18	14	14
wick 5	19	15	15
wick 5	20	16	16
wick 5	21	17	17
wick 5	22	18	18
wick 5	23	9	9
wick 5	24	10	10
wick 5	25	5	5
wick 5	26	11	11
wick 5	27	6	6
wick 5	28	7	7
wick 5	29	8	8
wick 5	30	na	165
wick 5	31	21	21
wick 5	32	22	22
wick 5	33	23	23
wick 5	34	19	19
wick 5	35	24	24
wick 5	36	28	28
wick 5	37	25	25
wick 5	38	26	26
wick 5	39	27	27
wick 5	40	29	29
wick 5	41	30	30
wick 5	42	31	31
wick 5	43	107	107
wick 5	44	32	32
wick 5	45	33	33
wick 5	46	36	36
wick 5	47	37	37
wick 5	48	34	34
wick 5	51	38	38
wick 5	52	39	39

wick 5	53	42	42
wick 5	54	43	43
wick 5	55	44	44
wick 5	56	41	41
wick 5	57	45	45
wick 5	58	47	47
wick 5	59	48	48
wick 5	60	49	49
wick 5	61	50	50
wick 5	62	51	51
wick 5	63	52	52
wick 5	64	53	53
wick 5	65	54	54
wick 5	66	55	55
wick 5	67	56	56
wick 5	68	57	57
wick 5	69	58	58
wick 5	70	144	144
wick 5	71	61	61
wick 5	72	62	62
wick 5	73	63	63
wick 5	74	68	na See nest 69 (photo) for explanation.
wick 5	74	69	69 Nest 68 and 69 (photo) are in the same approximate location as nest 74 (field). I decided that nest 74 (field) corresponds to nest 69 (photo) because both raised one chick. Nest 68 (photo) did not produce any chicks and was presumed to have not been seen in the field.
wick 5	75	70	70
wick 5	76	65	65
wick 5	77	71	71
wick 5	78	72	72

wick 5	79	73	73
wick 5	80	64	64
wick 5	81	59	59
wick 5	No nest was assigned number 82 in the field.	na	na
wick 5	83	60	60
wick 5	84	66	66
wick 5	85	67	67
wick 5	86	158	158
wick 5	87	153	153
wick 5	88	78	78
wick 5	89	na	166
wick 5	90	75	75
wick 5	91	74	74
wick 5	92	152	152
wick 5	93	147	147
wick 5	94	76	76
wick 5	95	77	77
wick 5	96	79	79
wick 5	97	80	80
wick 5	98	81	81
wick 5	99	150	150
wick 5	100	82	82
wick 5	101	92	92
wick 5	102	83	83
wick 5	103	84	84
wick 5	104	99	99
wick 5	107	46	46
wick 5	108	20	20
wick 5	49/50	35	49 field → 35 50 field → 167 Nest 49 and 50 (field) are in the same approximate location as nest 35 (photo). I decided that nest 49 (field) corresponds to nest 35 (photo) because both raised

			one chick. Nest 50 (field) had two successful chicks and so was presumed to have not been seen in photos. It was assigned a new unique number: 167.
wick 5	na	40	na
wick 5	na	145	na
wick 5	na	146	na
wick 5	na	148	na
wick 5	na	149	na
wick 5	na	155	na
wick 5	na	159	na
wick 5	na	160	na
wick 5	na	161	na
wick 5	na	162	na
wick 5	na	93	na
wick 5	na	102	na
wick 5	na	106	na
wick 4	10	139	139
wick 4	11	138	138
wick 4	12	140	140
wick 4	13	135	135
wick 4	14	134	134
wick 4	15	133	133
wick 4	16	137	137
wick 4	17	132	132
wick 4	18	130	130
wick 4	19	131	131
wick 4	20	136	136
wick 4	21	129	129
wick 4	22	128	128
wick 4	23	126	126
wick 4	24	125	125
wick 4	26	123	123
wick 4	27	122	122
wick 4	28	154	154
wick 4	29	114	114
wick 4	30	119	119

wick 4	31	118	118
wick 4	32	117	117
wick 4	33	116	116
wick 4	34	121	121
wick 4	35	120	120
wick 4	36	113	113
wick 4	37	112	112
wick 4	38	115	115
wick 4	39	110	110
wick 4	40	111	111
wick 4	41	109	109
wick 4	42	108	108
wick 4	43	142	142
wick 4	44	98	98
wick 4	45	97	97
wick 4	46	156	156
wick 4	47	104	104
wick 4	48	105	105
wick 4	49	103	103
wick 4	50	101	101
wick 4	51	95	95
wick 4	52	91	91
wick 4	53	90	90
wick 4	54	89	89
wick 4	55	88	88
wick 4	56	87	87
wick 4	57	86	86
wick 4	58	94	94
wick 4	59	85	85
wick 4	60	100	100
wick 4	61	96	96
wick 4	62	na	163
wick 4	63	127	127
wick 4	64	124	25 field →124 64 field → 164 Nest 25 and 64 (field) are in the same approximate location as nest 124 (photo). I decided that nest 24 (field)

			corresponds to nest 124 (photo) because both had one chick. Nest 64 (field) had zero chicks and so was presumed to have not been seen in photos. It was assigned a new unique number: 164.
wick 4	na	141	na
wick 4	na	143	na
wick 4	na	151	na
wick 4	na	157	na

Table S2 The difference in the number of chicks assumed to have fledged between field and photo data, when the fieldworker assumed only large chicks fledged but time-lapse photographs assumed all chicks fledge. ‘No. chicks field (A)’ is the number of chicks seen at each nest site in the field, regardless of size**.** **‘**No. large chicks field (B)’ is the number of large chicks seen at each nest site in the field**.** **‘**No. chicks photo (C)’ is number of chicks seen at each nest in photos regardless of size. ‘Difference (B – C)’ is the difference between the number of chicks assumed to have fledged based on field versus image observation. ‘Photo (C) = Field B’ is assigned a 1 if the ‘Difference’ is 0, and a 0 otherwise. ‘Photo (C) > Field (B)’ is assigned a 1 if the ‘Difference’ is -1 and a 0 otherwise, and ‘Photo (C) < Field (B) is 1 if the ‘Difference’ is 1 and a 0 otherwise. The 144 potential nests recorded in both in the field and in photographs are shown in this table. The 23 sites monitored only in the field or only in photographs are not included.

Nest site	**No. chicks field (A)**	**No. large chicks field (B)**	**No. chicks photo (C)**	**Difference (B – C)**	**Photo (C) = Field (B)**	**Photo (C) > Field (B)**	**Photo (C) < Field (B)**
1	2	2	1	1	0	0	1
2	1	0	1	-1	0	1	0
3	2	0	1	-1	0	1	0
4	1	0	1	-1	0	1	0
5	1	0	1	-1	0	1	0
6	0	0	0	0	1	0	0
7	0	0	0	0	1	0	0
8	0	0	0	0	1	0	0

9	0	0	0	0	1	0	0
10	0	0	0	0	1	0	0
11	0	0	0	0	1	0	0
12	0	0	0	0	1	0	0
13	0	0	0	0	1	0	0
14	0	0	0	0	1	0	0
15	0	0	0	0	1	0	0
16	2	0	1	-1	0	1	0
17	1	0	0	0	1	0	0
18	0	0	0	0	1	0	0
19	2	1	1	0	1	0	0
20	0	0	1	-1	0	1	0
21	2	1	1	0	1	0	0
22	2	2	1	1	0	0	1
23	0	0	1	-1	0	1	0
24	2	2	2	0	1	0	0
25	2	1	1	0	1	0	0
26	0	0	1	-1	0	1	0
27	2	1	1	0	1	0	0
28	0	0	0	0	1	0	0
29	1	1	1	0	1	0	0
30	1	1	1	0	1	0	0
31	2	2	1	1	0	0	1
32	2	1	1	0	1	0	0
33	1	1	1	0	1	0	0
34	2	1	1	0	1	0	0
35	1	1	1	0	1	0	0
36	1	1	1	0	1	0	0
37	2	2	1	1	0	0	1
38	2	2	2	0	1	0	0
39	0	0	0	0	1	0	0
41	0	0	1	-1	0	1	0
42	1	1	1	0	1	0	0
43	1	1	1	0	1	0	0
44	2	1	1	0	1	0	0
45	0	0	0	0	1	0	0
46	1	0	1	-1	0	1	0
47	1	1	1	0	1	0	0
48	0	0	0	0	1	0	0

49	1	1	1	0	1	0	0
50	1	1	1	0	1	0	0
51	2	0	1	-1	0	1	0
52	2	2	1	1	0	0	1
53	2	2	1	1	0	0	1
54	2	2	1	1	0	0	1
55	0	0	0	0	1	0	0
56	1	1	1	0	1	0	0
57	1	1	1	0	1	0	0
58	0	0	0	0	1	0	0
59	2	2	2	0	1	0	0
60	0	0	1	-1	0	1	0
61	0	0	0	0	1	0	0
62	1	1	1	0	1	0	0
63	1	1	2	-1	0	1	0
64	2	1	1	0	1	0	0
65	1	1	1	0	1	0	0
66	1	1	1	0	1	0	0
67	2	1	1	0	1	0	0
69	1	1	1	0	1	0	0
70	1	1	1	0	1	0	0
71	1	1	1	0	1	0	0
72	1	0	1	-1	0	1	0
73	2	2	2	0	1	0	0
74	0	0	0	0	1	0	0
75	2	2	2	0	1	0	0
76	2	0	2	-2	0	1	0
77	2	1	1	0	1	0	0
78	1	0	1	-1	0	1	0
79	0	0	1	-1	0	1	0
80	1	1	1	0	1	0	0
81	2	1	1	0	1	0	0
82	2	1	1	0	1	0	0
83	1	1	1	0	1	0	0
84	2	1	1	0	1	0	0
85	1	1	1	0	1	0	0
86	2	2	2	0	1	0	0
87	1	1	2	-1	0	1	0
88	2	2	2	0	1	0	0

89	1	1	1	0	1	0	0
90	2	2	1	1	0	0	1
91	1	1	1	0	1	0	0
92	0	0	1	-1	0	1	0
94	1	0	1	-1	0	1	0
95	1	1	1	0	1	0	0
96	1	1	1	0	1	0	0
97	2	2	1	1	0	0	1
98	2	2	2	0	1	0	0
99	2	1	1	0	1	0	0
100	1	0	1	-1	0	1	0
101	0	0	1	-1	0	1	0
103	0	0	0	0	1	0	0
104	1	0	0	0	1	0	0
105	0	0	0	0	1	0	0
107	2	0	1	-1	0	1	0
108	2	2	1	1	0	0	1
109	2	1	1	0	1	0	0
110	0	0	0	0	1	0	0
111	2	0	1	-1	0	1	0
112	1	1	1	0	1	0	0
113	1	1	1	0	1	0	0
114	0	0	0	0	1	0	0
115	1	1	1	0	1	0	0
116	2	1	1	0	1	0	0
117	1	1	1	0	1	0	0
118	1	1	1	0	1	0	0
119	2	1	1	0	1	0	0
120	2	0	0	0	1	0	0
121	0	0	0	0	1	0	0
122	2	0	1	-1	0	1	0
123	2	2	1	1	0	0	1
124	2	1	1	0	1	0	0
125	1	1	1	0	1	0	0
126	1	1	1	0	1	0	0
127	0	0	0	0	1	0	0
128	1	1	1	0	1	0	0
129	1	1	1	0	1	0	0
130	2	1	1	0	1	0	0

131	2	2	1	1	0	0	1
132	2	1	1	0	1	0	0
133	2	2	1	1	0	0	1
134	2	1	1	0	1	0	0
135	1	1	1	0	1	0	0
136	2	1	1	0	1	0	0
137	1	1	1	0	1	0	0
138	0	0	0	0	1	0	0
139	2	2	2	0	1	0	0
140	0	0	0	0	1	0	0
142	2	2	1	1	0	0	1
144	0	0	0	0	1	0	0
147	0	0	0	0	1	0	0
150	0	0	0	0	1	0	0
152	0	0	0	0	1	0	0
153	0	0	0	0	1	0	0
154	0	0	0	0	1	0	0
156	0	0	0	0	1	0	0
158	0	0	0	0	1	0	0
TOTAL (144)					**105**	**25**	**14**

Table S3 More detailed analysis of the 64 nest sites where the number of chicks assumed to have fledged differed between field and photograph, when the fieldworker assumed only large chicks fledged but all chicks were presumed to have fledged from photographs. Blue shading is where image analysis recorded fewer fledged chicks in the nest than field observations. Red shading is where image analysis recorded more fledged chicks than in the field. Green shading is where the number of fledged chicks was equal for field and photo data; however, the fieldworker recorded more non-large chicks which were presumed to have died and not fledged. The 'Explanation' column summarises why there were differences in the number of fledged chicks between field and photo where appropriate.

Nest site	**No. chicks field (A)**	**No. large chicks field (B)**	**No. chicks photo (C)**	**Difference (B – C)**	**Photo (C) = Field (B)**	**Photo (C) > Field (B)**	**Photo (C) < Field (B)**	**Explanation**	
								Field	**Photo**
1	2	2	1	1	0	0	1		
2	1	0	1	-1	0	1	0	One small/medium chick last seen 2019-07-29.	One chick seen 2019-07-25 to 2019-08-03.
3	2	0	1	-1	0	1	0	Two small/medium chicks last seen together 2019-07-25. One medium chick last seen 2019-08-07.	One chick seen 2019-07-16 to 2019-08-05.
4	1	0	1	-1	0	1	0	Two small chicks last seen together 2019-07-08. One medium chick last seen 2019-07-21.	One chick seen 2019-07-16 to 2019-07-21.

5	1	0	1	-1	0	1	0	One small/medium chick last seen 2019-08-02.	One chick seen 2019-07-25 to 2019-08-02.
16	2	0	1	-1	0	1	0	Two small chicks last seen 2019-07-29.	One chick seen once 2019-07-28.
17	1	0	0	0	1	0	0	One small chick last seen 2019-07-08.	No chicks observed.
19	2	1	1	0	1	0	0	Two small chicks last seen together 2019-07-08. One large chick last seen 2019-07-25.	
20	0	0	1	-1	0	1	0	Adult apparently incubating 2019-06-25, 2019-07-02 and 2019-07-08. No adult on the nest 2019-07-13 to 2019-07-25. No chicks seen.	One chick seen 2019-07-24 to 2019-08-08.
21	2	1	1	0	1	0	0	Two medium/large chicks last seen together 2019-08-02. One large chick last seen 2019-08-07.	
22	2	2	1	1	0	0	1		
23	0	0	1	-1	0	1	0	Adult apparently incubating 2019-05-22 to 2019-07-13. No incubation seen after	One chick seen once 2019-08-05. Could be a fledged chick from another nest.

								2019-07-13. No chicks seen.	
25	2	1	1	0	1	0	0	Two small chicks last seen together 2019-07-08. One large chick last seen 2019-07-25.	
26	0	0	1	-1	0	1	0	One egg 2019-07-02, apparently adult incubating 2019-7-08 and 2019-07-13. No incubation seen after 2019-07-13. No chicks seen.	One chick seen once 2019-08-02. Could be a fledged chick from another nest.
27	2	1	1	0	1	0	0	Two medium chicks last seen together 2019-07-13. One large chick last seen 2019-07-25.	
31	2	2	1	1	0	0	1		
32	2	1	1	0	1	0	0	Two small chicks last seen together 2019-7-13. One large chick last seen 2019-07-29.	

34	2	1	1	0	1	0	0	Two small chicks last seen together 2019-07-02, although two chicks may have been present until 2019-07-21. One large chick last seen 2019-07-29.	
37	2	2	1	1	0	0	1		
41	0	0	1	-1	0	1	0	Adult apparently incubating 2019-05-22 to 2019-07-08. No incubation seen after 2019-07-08. No chicks seen.	Chick first seen 2019-07-12 and recorded 11 times 2019-07-12 to 2019-07-29. Chick not seen after 2019-07-29.
44	2	1	1	0	1	0	0	Two small chicks last seen together 2019-07-08. One large chick last seen 2019-07-25.	
46	1	0	1	-1	0	1	0	One medium/large chick last seen 2019-08-02.	One chick seen 2019-07-12 to 2019-07-29.
51	2	0	1	-1	0	1	0	Two small chicks last seen 2019-07-08.	One chick seen 2019-07-28 and 2019-07-30 only.
52	2	2	1	1	0	0	1		
53	2	2	1	1	0	0	1		
54	2	2	1	1	0	0	1		

60	0	0	1	-1	0	1	0	Adult apparently incubating 2019-06-17. No incubation seen after 2019-06-17. No chicks seen.	One chick seen once 2019-06-26.
63	1	1	2	-1	0	1	0	One large chick last seen 2019-07-25.	One chick seen 2019-07-12 to 2019-08-07. Two chicks seen only once on 2019-07-22.
64	2	1	1	0	1	0	0	One or more medium chicks last seen together 2019-07-13. One large chick last seen 2019-07-25.	
67	2	1	1	0	1	0	0	Two small chicks last seen together 2019-07-13. One large chick last seen 2019-07-25.	
72	1	0	1	-1	0	1	0	One small chick last seen for definite 2019-07-21 and possibly seen 2019-07-25.	One chick seen 2019-07-15 and 2019-07-19 only.
76	2	0	2	-2	0	1	0	Two medium/large chicks last seen 2019-07-25.	One chick seen 2019-07-14 to 2019-07-30.

77	2	1	1	0	1	0	0	Two medium/large chicks last seen together 2019-07-21. One large chick last seen 2019-07-25.	
78	1	0	1	-1	0	1	0	One small/medium chick last seen 2019-08-02.	One chick seen 2019-07-29 to 2019-08-05.
79	0	0	1	-1	0	1	0	One egg 2019-06-25, apparently incubating adult 2019-07-02 and 2019-07-08. No incubation seen after 2019-07-08. No chicks seen.	One chick seen 2019-07-02 and 2019-08-07 only. Chick may have died and the one seen in August was a fledged chick from another nest.
81	2	1	1	0	1	0	0	Two medium/large chicks last seen together 2019-07-21. One large chick last seen 2019-07-25.	
82	2	1	1	0	1	0	0	Two medium chicks last seen together 2019-07-13. One large chick last seen 2019-07-25.	
84	2	1	1	0	1	0	0	Two small chicks last seen together 2019-07-	

								02. One large chick last seen 2019-07-25.	
87	1	1	2	-1	0	1	0	Only one chick seen.	Two chicks seen once 2019-07-26. Otherwise, only one chick recorded.
90	2	2	1	1	0	0	1		
92	0	0	1	-1	0	1	0	One egg 2019-07-21. Two adults not incubating 2019-07-25. No further observations recorded after 2019-07-25.	One chick seen once 2019-07-25.
94	1	0	1	-1	0	1	0	One or more small/medium chicks last seen 2019-07-21. One medium /large chick last seen 2019-08-02.	One chick first seen on 2019-07-29 and last seen on 2019-08-04.
97	2	2	1	1	0	0	1		
99	2	1	1	0	1	0	0	Two medium/large chicks last seen together 2019-07-21. One large chick last seen 2019-07-25.	
100	1	0	1	-1	0	1	0	One medium/large chick last seen 2019-07-21. One adult not incubating 2019-07-25.	One chick last seen 2019-07-28. Could be a fledged chick from another nest.

101	0	0	1	-1	0	1	0	One adult apparently incubating 2019-07-02 and 2019-07-08. One adult not incubating until 2019-07-25. No chicks seen.	One chick seen once 2019-07-28. Could be a fledged chick from another nest.
104	1	0	0	0	1	0	0	One small chick seen once 2019-06-25.	
107	2	0	1	-1	0	1	0	Two medium chicks last seen together 2019-07-21. One adult not incubating 2019-07-25.	One chick seen 2019-07-12 to 2019-08-04.
108	2	2	1	1	0	0	1		
109	2	1	1	0	1	0	0	Two medium/large chicks last seen together 2019-07-17. One large chick last seen 2019-07-25.	
111	2	0	1	-1	0	1	0	One medium chick and one large chick last seen together 2019-08-02. One medium/large chick last seen 2019-08-07. One adult not incubating 2019-08-10.	One chick last seen 2019-08-10.

116	2	1	1	0	1	0	0	Two small chicks last seen together 2019-07-08. One large chick last seen 2019-08-02.	
119	2	1	1	0	1	0	0	Two small chicks last seen together 2019-07-08. One large chick last seen 2019-07-25.	
120	2	0	0	0	1	0	0	Two medium chicks last seen together 2019-07-13.	
122	2	0	1	-1	0	1	0	Two medium/large chicks last seen together 2019-07-25.	One chick seen 2019-07-20 to 2019-07-25.
123	2	2	1	1	0	0	1		
124	2	1	1	0	1	0	0	One or more chicks last seen together 2019-07-08. One large chick last seen 2019-08-02.	
130	2	1	1	0	1	0	0	Two small chicks last seen together 2019-07-08. One large chick last seen 2019-07-25.	
131	2	2	1	1	0	0	1		
132	2	1	1	0	1	0	0	Two small chicks last seen together 2019-07-08.	

								One large chick last seen 2019-07-25.	
133	2	2	1	1	0	0	1		
134	2	1	1	0	1	0	0	Two small chicks last seen together 2019-07-08. One large chick last seen 2019-07-25.	
136	2	1	1	0	1	0	0	One or more small chicks last seen together 2019-07-08. One large chick last seen 2019-08-02.	
142	2	2	1	1	0	0	1		
TOTAL (64)					**105**	**25**	**14**		

CHAPTER TWO: Adverse weather conditions lower Black-legged Kittiwake *Rissa tridactyla* egg and chick survival

---

ABSTRACT

The Black-legged Kittiwake *Rissa tridactyla* is a species of conservation concern. In the UK, where 8% of the world population breeds, there has been a 23% decline since the mid-1980s. There is strong evidence for reduced breeding success in Scotland being associated with overfishing and increasing sea surface temperatures reducing prey availability; however, the cause of decline in Wales is far from certain. I investigated the effect of prevailing weather conditions on Kittiwake egg and chick survival during the 2019 breeding season on Skomer Island, Wales, to determine the relative importance of weather on breeding success. I found support for strong westerly winds reducing egg and chick daily survival rate, while high daily maximum temperature appeared to lower egg survival only. There was also evidence that egg and chick survival rate decreased towards the end of the incubation and chick-rearing period respectively. If the results of this study represent a longer-term pattern, then predicted increases in the frequency and intensity of extreme weather events with climate change could have severe adverse consequences on Kittiwake breeding success.

INTRODUCTION

The Black-legged Kittiwake *Rissa tridactyla* (hereafter Kittiwake) is the world's most numerous gull species and is found throughout the Northern Hemisphere (Coulson 2011). However, over their past three generations, Kittiwake populations have undergone rapid decline which shows no evidence of slowing (BirdLife International

2019). This has led to their classification as 'Vulnerable' on the IUCN Red List (BirdLife International 2019). In the British Isles, the breeding Kittiwake population has decreased by 23% since the mid-1980s, which is particularly concerning given that 8% of the world population breeds on UK cliffs (Mitchell *et al.* 2004). Nationally, the greatest loss has occurred in Scotland, where the 2014 index was 72% below the 1986 baseline. (Wilkie *et al.* 2019).

Changes in seabird populations are most often attributed to changes in post-fledging and adult survival (Weimerskirch *et al.* 1997, Mitchell *et al.* 2004). Seabird life-history is characterised by long-lifespan and low fecundity, meaning populations may be buffered by short-term variation in productivity due to a long life-span providing many opportunities to reproduce (Sæther & Bakke 2000). However, recent evidence suggests breeding success can also be an important driver of seabird population dynamics (Reiertsen 2013). In Scotland lowered Kittiwake breeding success has been associated with reduced prey availability as North Sea sea surface temperatures (SST) have increased. This shortage of food has been exacerbated by overfishing and increased predation (Regehr & Montevecchi 1997, Frederiksen *et al.* 2004, Cook *et al.* 2014). Conversely, no relationship was found between SST in the Celtic Sea and Kittiwake breeding success or population growth (Lauria *et al.* 2012, Cook *et al.* 2014). Instead, the North Atlantic Oscillation appeared to have a weak positive effect on Kittiwake population growth rate which may have been related to the direct effects of wind speed or storm frequency (Lauria *et al.* 2012). Although far less is known about the causes of Kittiwake decline in Wales compared to in Scotland, it has been suggested that increasingly extreme weather may be having a greater impact on Welsh Kittiwake productivity than SST.

Intensification of weather extremes has been described as 'one of the most important facets of climate change' but has received less attention compared to rising SSTs so far (Jentsch *et al.* 2007, Daunt & Mitchell 2013). Extreme weather events are typically defined as those having a low frequency of occurrence (< 5%) but large ecological impact and are increasing in frequency and intensity as the climate alters (Daunt *et al.* 2013, Descamps *et al.* 2015). In this study, I explored the relative importance of weather on Kittiwake nest survival on Skomer Island, Wales. Skomer Island is a nationally important site for breeding seabirds, but Kittiwake numbers have fallen by 42% since 1986 (Stubbings *et al.* 2017, Wilkie *et al.* 2019). Preliminary Population Viability Analysis has shown that the breeding population could decline to a few hundred by the end of the century if current rates of survival and productivity continue, or become locally extinct within 70 years if parameters worsen (Horswill, Perrins & Wood, unpublished). Understanding the factors affecting Kittiwake nest survival could provide crucial insight into falling productivity on Skomer and allow strategy development to help reverse the negative population trend. It may also direct research into the causes of Kittiwake decline at other colonies in Wales if primary factors are identified.

This chapter compares the effect of extrinsic weather variables, temporal and intrinsic factors on nest survival across the entire breeding season and during the egg and chick stages separately. I used data from field observations and time-lapse cameras in 2019 to test the three *a priori* hypotheses that nest survival is affected by: 1) time 2) weather conditions and 3) brood size. I predicted that survival would decrease across the breeding season as nests gradually failed (1), but the decline would be greatest at the end of the egg stage when parents deserted unhatched eggs. As well, I predicted that high temperatures and strong westerly winds seen in 2019 (2) could reduce survival of eggs and chicks through heat stress and knocking nest contents off the cliffs. Although I did not examine the effect of prey availability directly, I predicted that survival would be most affected by brood size (3) if prey was limited and parents could not supply sufficient food to support multiple offspring.

METHODS

Data collection

Study sites and nest monitoring

Monitoring was conducted at the Wick, Skomer Island, Wales (51°40N, 05°15W) during the 2019 breeding season (Fig. 1). A total of 132 nests were monitored at Wick 4 and 5 study plots both in the field and using a Reconyx time-lapse camera (HyperFire 2 Covert IR Camera OD) fitted with a 2x magnification lens (HF2 Telephoto Lens, Nar Illum & Matched PIR). Study plots were visited a total of 13 times between the 22nd May and 11th August, with visits occurring a minimum of every five days once the first chicks reached medium size (Walsh *et al.* 1995, Wilkie *et al.* 2019). One photograph was examined every day from the 11th April to 15th August, which was the last date Kittiwakes were seen in images.

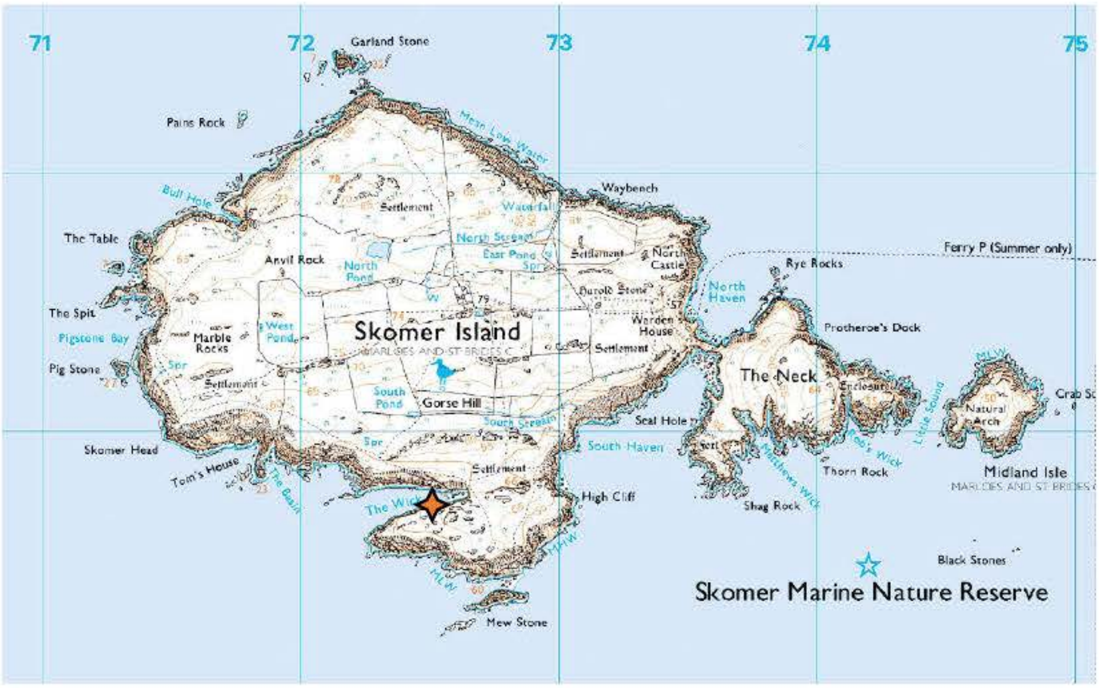


Figure 1 Map of Skomer Island, Wales. The Kittiwake colony studied is marked with an orange star.

Comparison of phenology data obtained from field and photo observations showed that on average, potential nests were first recorded as apparently occupied nests (AONs) 3.7 ($\sigma = 8.2$) days earlier in photos than in the field, while chicks were first and last seen 13.9 ($\sigma = 7.6$) and 9.6 ($\sigma = 5.7$) days later respectively in photographs (refer to chapter one). Time-lapse cameras likely recorded earlier first AONs and later last chick dates because they captured images before and after the fieldworkers' first and last visit. Earlier first chick dates may have been recorded in the field because small, recently hatched chicks were harder to spot in images. The fieldworker could spend time watching the nest, waiting for the adult and/or potential chick to move and confirm chick presence or absence, while the precise timing of imaging meant such behaviour could be missed (Tanedo & Hollémen 2020). To reduce these known biases in field and photo derived phenology, dates for survival analysis were chosen by comparing field and photo data. First AON and first chick dates were selected by examining both field and photo phenology and choosing the earliest dates per nest, while last chick dates were obtained by choosing the latest dates per nest.

Weather data

Daily total rainfall (mm), mean wind speed (kmh^{-1}) and mean wind direction were obtained from the Milford Haven Conservancy Board weather station (51°43N, 05°03W) ~17 km south-east of Skomer Island (Met Office 2012). Wind direction was converted from degrees to the cosines and sines of the radians. This removed the circularity of wind direction and generated a measure of northern wind direction for cosine values, from north = 1 to south = -1, and eastern wind direction for sine values, from east = 1 to west = -1 (Michielsen *et al.* 2019). Daily minimum, mean and maximum temperature (°C) were obtained from the time-lapse camera which had an inbuilt temperature sensor. During the study period, the daily minimum, mean and maximum temperature ranged from 7 to 16°C, 10.1 to 21.6°C and 11 to 34°C respectively. Daily total rainfall ranged from 0.2 to 18.6mm on days when it rained but was 0 mm on 58% of days. Daily mean wind speed was between 7.6 and 36.7 kmh^{-1}.

Data analysis

Collinearity of explanatory variables

Collinearity between explanatory variables can increase the probability of type I errors and was assessed using the 'usdm' R package (Lauria *et al.* 2012, Naimi *et al.* 2014, Mwangi *et al.* 2018). The 'vifstep' function calculated the variance inflation factor (VIF) for each variable and excluded those with VIF greater than three through a stepwise procedure. VIF greater than ten indicates high correlation with one or more predictor variables and a collinearity problem, although VIF greater than three may be used as a more stringent approach (Naimi *et al.* 2014, Zuur *et al.* 2010, Mwangi *et al.* 2018). This procedure was carried out for all explanatory variables across the entire nesting period (egg&chick stage), incubation period (egg stage) and chick-rearing period (chick stage). Variables with VIF < 3 for all stages were used in the survival analysis: daily maximum temperature, mean wind speed, mean eastern wind direction, mean northern wind direction and linear or quadratic time (Table 1). Linear and quadratic time could not be used in the same survival model as they were highly correlated. Total rainfall was excluded from the survival analysis as it showed a strong correlation (≥ 0.5) with maximum temperature ($r = -0.51$) across the entire nesting period; maximum temperature ($r = -0.65$) and mean temperature ($r = -0.56$) during the egg stage; and mean wind speed ($r = 0.51$) during the chick stage. As well, total rainfall was zero for 58% of the study period which could have led to zero-inflated data in the survival analysis (Zuur *et al.* 2010).

Table 1 Explanatory variables and their notation used in survival analysis

Explanatory variable (unit)	Notation
Time (days)	T
Time^2 (days^2)	TT
Daily maximum temperature (°C)	max_temp
Daily mean wind speed (km/h)	wind_speed
Daily mean eastern wind direction i.e. sine of the mean wind direction	sin_dir
Daily mean northern wind direction i.e. cosine of the mean wind direction	cos_dir
Number of chicks	brood_size

Survival of Kittiwake nests

Nest daily survival rate (S) was defined as the probability that a nest survives a single day and was calculated for the entire nesting period (egg&chick stage), incubation (egg stage) and chick-rearing period (chick stage) using Dinsmore *et al*'s (2002) nest survival model in the R package 'RMark' (Rotella *et al.* 2004, Laake 2013). This uses the R interface to run models in the program MARK (White & Burnham 1999, White 2011).

The nest survival model uses a generalised linear modelling approach based on binomial likelihood, where nest daily survival rates are modelled as a function of group-specific, nest-specific and/or time-specific covariates (Rotella *et al.* 2004). This means daily survival rates can vary among groups of nests, among individual nests and/or among days, respectively. The logit (logistic regression) link function is used to characterise the relationship between daily survival rate and covariates and is given by:

$$\mathrm{logit}(S_i)\text{: } \ln\left(S_i / (1 - S_i)\right) = \beta_0 + \Sigma_j\, \beta_j\, \chi_{ji}$$

where S_i is the daily survival rate (i.e. the probability that a nest survives from day i to day i +1); x_{ji} (j=1,2, …J) are values for j covariates on day i; and the $\{B_j\}$ are coefficients to be estimated from the data (Dinsmore *et al.* 2002, Rotella *et al.* 2004). These coefficients are referred to as 'beta' estimates since they are estimated 'slopes'/ beta terms from the linear model, and they can take any value (Rotella *et al.* 2004). The estimate of daily survival rate is obtained by back-transformation (i.e. calculating the inverse logit function) using the plogis function in R, to give:

$$S_i = \frac{\exp(\beta_0 + \Sigma_j \beta_j \chi_{ji})}{1 + \exp(\beta_0 + \Sigma_j \beta_j \chi_{ji})}$$

The coefficients here are referred to as 'real' estimates, as they have been back-transformed from the 'transformed' scale to the 'real' probability scale, and are bounded between 0 and 1 (Dinsmore *et al.* 2002).

In this study, the egg and chick stage were modelled separately, as well as together, to allow easy dissemination of the different factors affecting daily survival rate during the incubation and chick-rearing period. Daily survival rates of individual eggs and chicks were not calculated, due to the lack of independence between eggs/chicks from the same nest. Instead, when modelling the egg stage, nests were successful if *at least* one egg hatched, and nests were successful across the entire period and in the chick stage if *at least* one chick reached fledging (large) size. This definition of nest success follows Rotella *et al.* (2004). The number of chicks in each nest was included as a factor in chick stage models, to determine its effect on at least one chick surviving to fledge.

Construction of nest survival models required the date an egg/chick was 'first found' and 'last present' at each nest, the date the nest was 'last checked' and its fate (assigned zero if successful and one if unsuccessful) (Rotella *et al.* 2004, Dinsmore & Dinsmore 2007). Using a model that included both 'last present' and 'last checked' meant the exact date of nest failure did not need to be known or assumed; however, it was assumed that 'last checked' equalled 'last present' for successful nests (Rotella *et al.* 2004). Since eggs were difficult to observe both in the field and in time-lapse photographs, I used the first time a site was recorded as an 'apparently occupied nest' as a proxy for egg-laying date. Table 2 lists the dates used for failed and successful nests when modelling survival across the egg&chick, egg and chick stages.

Table 2 Dates used for ‘First found’, ‘Last present’ and ‘Last checked’ values in nest survival models for successful and failed nests during different breeding stages: entire nesting period (egg&chick stage), incubation period (egg stage) and chick-rearing period (chick stage). First AON and first chick dates were selected by examining both field and photo phenology and choosing the earliest dates per nest, while last chick dates were obtained by choosing the latest dates per nest.

Model parameter	**Egg&chick stage**	**Egg stage**	**Chick stage**
First found	First AON	First AON	First chick
Failed: Last present	Egg stage: Last time AON recorded Chick stage: Last time small/medium chick recorded	Last time AON recorded	Last time small/medium chick recorded
Failed: Last checked	Egg stage: First time empty nest recorded or adult not adopting incubation position Chick stage: First time an empty nest recorded	First time empty nest recorded or adult not adopting incubation position	First time an empty nest recorded
Successful: Last present	Last time large chick recorded	First chick	Last time large chick recorded
Successful: Last checked	Last time large chick recorded	First chick	Last time large chick recorded

Model development and selection

Subsets of models were developed according to three *a priori* biological hypotheses on factors affecting egg and chick survival: 1) time 2) weather conditions and 3) brood size. I included a maximum of two explanatory variables and one interaction per model to reduce the number of parameters to be estimated and thus model complexity (Christensen-Dalsgaard *et al.* 2018). Based on methods used by Christensen-Dalsgaard *et al.* (2018), each hypothesis group was considered separately and then all hypothesis groups were considered together to assess the relative importance of temporal, extrinsic and intrinsic sources of variation on daily survival rate (Fig. 2).

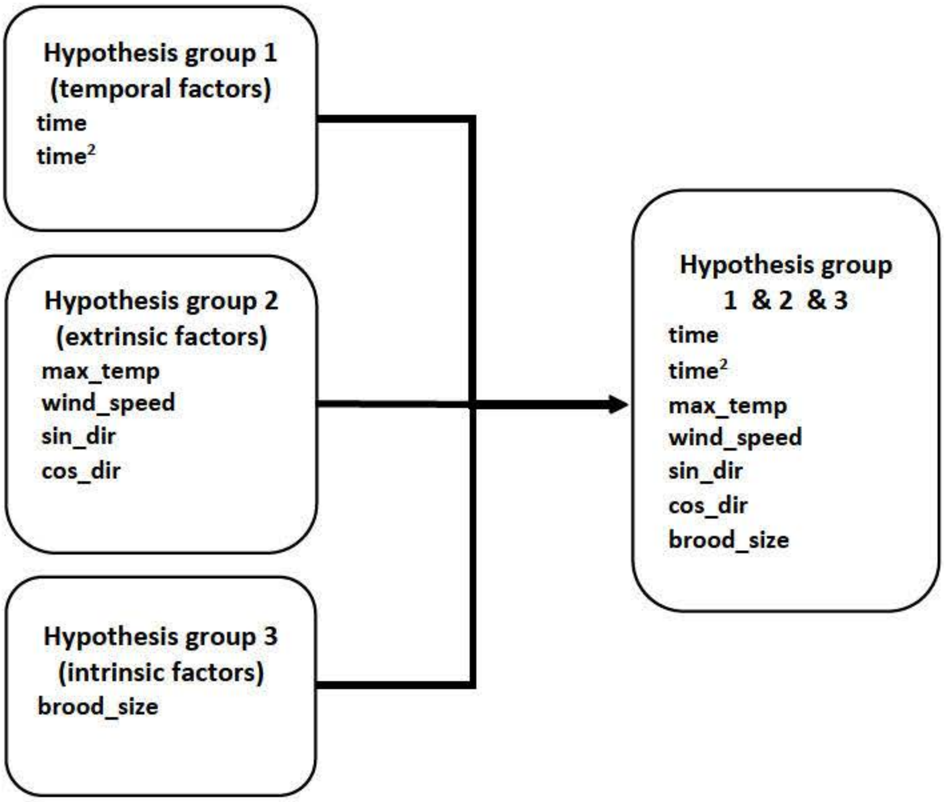


Figure 2 Flowchart showing the modelling process.

Model selection was performed using Akaike's information criterion adjusted for small sample size (AICc) (Burnham et al. 2011). The best model had the lowest AICc value and highest weight (w_i). Models within two AICc of the best model (Δ AIC < 2) are sometimes considered equally parsimonious, however for the purposes of this study, only the model with the lowest AIC was selected (Burnham & Anderson 2002, Burnham *et al.* 2011). Parameters within each model were statistically significant if the 95% confidence interval did not overlap zero. For the best model in each hypothesis group, daily survival rates from real parameter estimates were multiplied together to give the total probability of survival across the period of interest (i.e. the probability that a nest/egg/chick was successful) (Rotella *et al.* 2004, Conkling *et al.* 2015). To help evaluate the effect of covariates on the daily survival of Kittiwake nests/eggs/chicks, selected values of each covariate were substituted into the best model's logistic regression equation for each hypothesis group, and then back-transformed to give daily survival rates between 0 and 1 for each value. For models containing two covariates, one covariate was held at its minimum, mean and maximum value while the other changed on a continuous scale, to give three daily survival rate curves plotted on the same axes (Dinsmore *et al.* 2002, Pierce 2017).

RESULTS

Between the 11th April and 15th August 2019, 132 nests were monitored. The length of the entire nesting period from the first AON to the last chick observed was 92 days. The length of the entire incubation period (egg stage) from the first AON to the last chick hatched was 73 days, and the entire chick-rearing period (chick stage) from the first chick hatched to the last chick fledged was 50 days. For nests where an egg successfully hatched, the mean length of the incubation period was 39.5 days per nest, and for nests that fledged at least one chick, the mean length of the chick-rearing period was 38.2 days per nest. 22% of nests failed during the egg stage and 14% failed during the chick stage. 64% of nests successfully fledged at least one chick.

Hypothesis group one: temporal factors

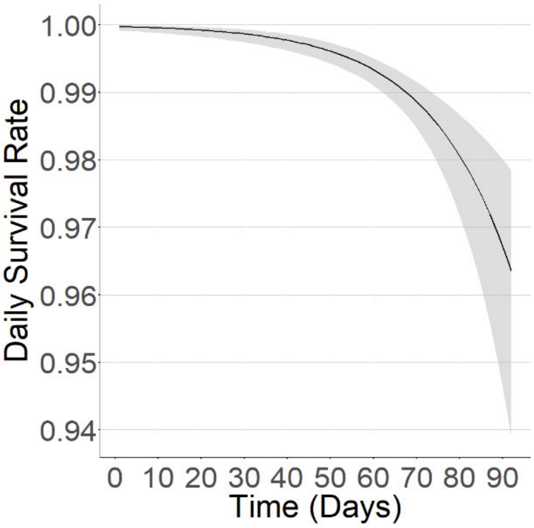


Figure 3 Estimated relationship between daily survival rate of Kittiwake nests and linear time across the entire nesting period. Estimates are from the best model in hypothesis group one, where logit(S) = 8.25 – 0.0541(T). (Table 3).

When considering the effect of temporal factors on survival across the entire breeding period, one model had Δ $AIC_c < 2$ ($w_i = 0.98$) which described nest daily survival rate as a function of linear time (Table 3). The model showed that daily survival rate significantly decreased as the breeding season progressed (β = -0.0541, SE = 0.00869, 95% CI -0.0711 to -0.0370) (Table 4, Fig. 3). The logistic regression equation for this model was: logit(S) = 8.25 – 0.0541(T).

When modelling the effect of temporal factors on egg survival, the best fit model (w_i = 0.92) showed daily survival rate decreasing with linear time (β = -0.222, SE = 0.0269 95% CI -0.274 to -0.169) (Table 3-4) and its logistic regression equation was: logit(S) = 15.6 – 0.222(T). During the chick-rearing period, the best fit temporal model (w_i = 0.81) described daily survival rate as a function of quadratic time (Table 3). This negative relationship was significant (β = -0.00201, SE = 0.000448, 95% CI -0.00280 to -0.00113) (Table 4) and the logistic regression equation for the model was given by: logit(S) = 7.28 – 0.00201(TT).

Hypothesis group two: extrinsic factors

For the effect of extrinsic variables on nest daily survival rate throughout the breeding period, two models had Δ $\text{AIC}_c < 2$ ($w_i = 0.68, 0.32$), both of which included maximum temperature and mean eastern wind direction (Table 3). The model with the lowest AIC_c showed that higher daily maximum temperatures significantly lowered the daily probability of survival (β = -0.585, SE = 0.0817, 95% CI -0.745 to -0.425) (Fig. 4A). Furthermore, daily survival rate was positively associated with mean eastern wind direction (β = 3.08, SE = 0.522 , 95% CI 2.06 to 4.10), meaning that increasingly westerly winds decreased the daily survival probability (Table 4). The logistic regression equation for this model was: logit(S) = 22.2 – 0.585(max_temp) + 3.08(sin_dir). Fig. 4A shows an additive effect of temperature and wind direction, suggesting that daily survival rate decreases at lower temperatures when the wind is more westerly. The best model for hypothesis group two was also the second most parsimonious model when considering temporal and extrinsic factors together across the entire breeding period ($w_i = 0.19$) (Table 3).

During the egg stage, one model had Δ $\text{AIC}_c < 2$ ($w_i = 0.84$) and showed a significant negative effect of maximum temperature (β = -1.07, SE = 0.150, 95% CI -1.37 to -0.778), mean eastern wind direction (β = -23.4, SE = 7.17, 95% CI -37.5 to -9.39) and a positive effect of the interaction between them, on egg daily survival rate (β = 0.883, SE = 0.234, 95% CI 0.425 to 1.34) (Table 3-4, Fig. 4B). The logistic regression equation for this model was: logit(S) = 34.5 – 1.07(max_temp) – 23.4(sin_dir) + 0.883(max_temp:sin_dir). This is very similar to the best supported weather model across the entire breeding period, except it also included the interaction term. For the chick stage, two models had Δ $\text{AIC}_c < 2$ ($w_i = 0.38, 0.17$) and the model with the lowest AIC_c suggested that mean wind speed significantly lowered chick daily survival rate (β = -0.134, SE = 0.0430, 95% CI -0.219 to -0.05) (Table 3-4). The effect of lowered daily survival rate with increasingly westerly winds was not statistically significant (β = 1.91, SE = 1.13, 95% CI -0.301 to 4.12) (Fig. 4C, Table 4). This model had a logistic regression equation of: logit(S) = 8.44 – 0.134(wind_speed) – 1.91(sin_dir).

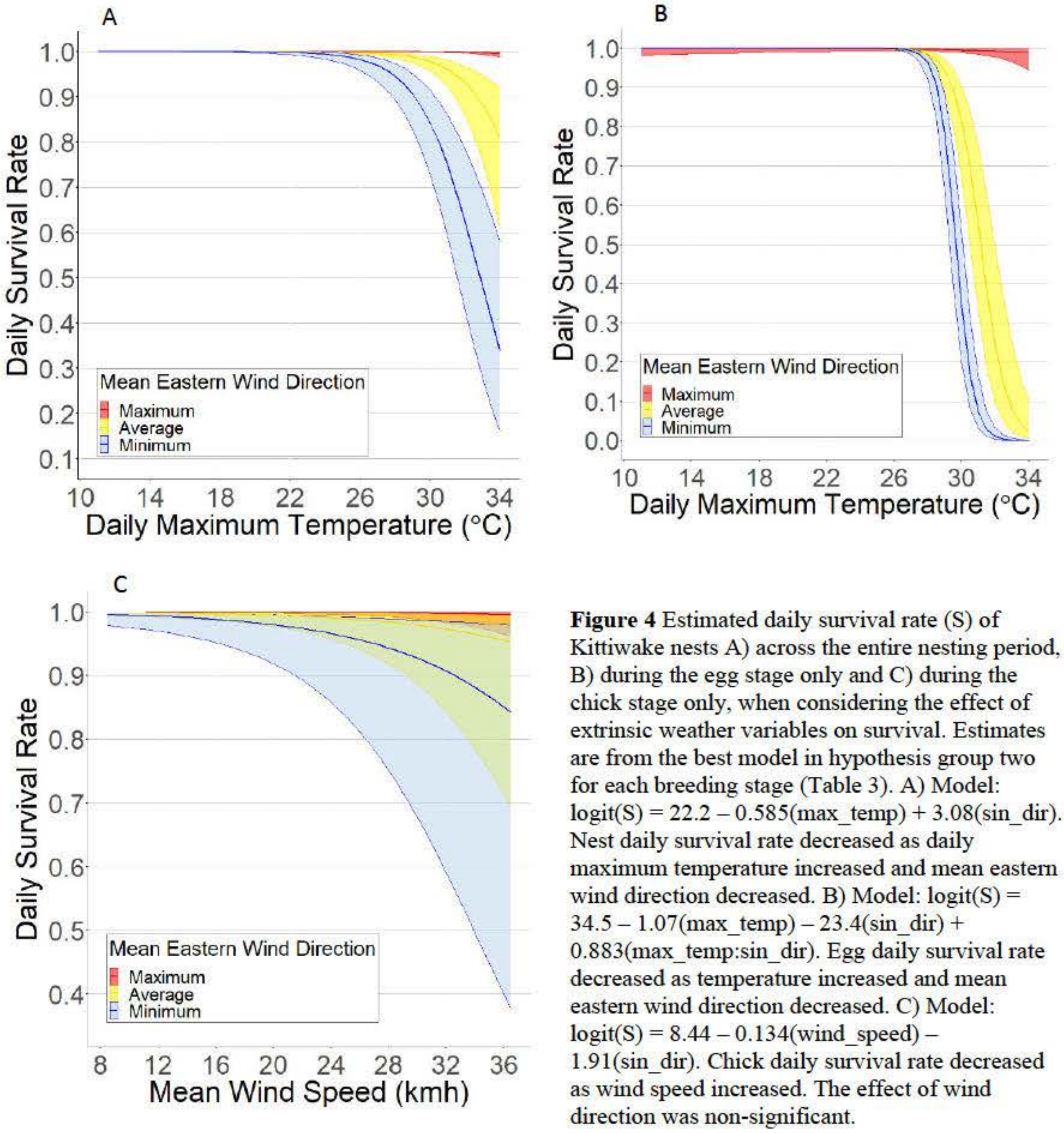


Figure 4 Estimated daily survival rate (S) of Kittiwake nests A) across the entire nesting period, B) during the egg stage only and C) during the chick stage only, when considering the effect of extrinsic weather variables on survival. Estimates are from the best model in hypothesis group two for each breeding stage (Table 3). A) Model: logit(S) = 22.2 – 0.585(max_temp) + 3.08(sin_dir). Nest daily survival rate decreased as daily maximum temperature increased and mean eastern wind direction decreased. B) Model: logit(S) = 34.5 – 1.07(max_temp) – 23.4(sin_dir) + 0.883(max_temp:sin_dir). Egg daily survival rate decreased as temperature increased and mean eastern wind direction decreased. C) Model: logit(S) = 8.44 – 0.134(wind_speed) – 1.91(sin_dir). Chick daily survival rate decreased as wind speed increased. The effect of wind direction was non-significant.

Hypothesis group three: intrinsic factors

For intrinsic variables, brood size had very little effect on the probability of at least one chick surviving on a given day ($w_i = 0.16$). Constant chick daily survival rate ($\beta = 5.32$, SE = 0.236, 95% CI 4.85 to 5.78) explained four times more variation in chick survival ($w_i = 0.84$) (Table 3-4).

All hypothesis groups: temporal, extrinsic and intrinsic factors

When temporal and extrinsic variables were considered together, two models had Δ $AIC_c < 2$ for the egg&chick stage (Table 3). The model with the lowest AIC_c value ($w_i = 0.49$) showed that nest daily survival rate was affected by time, mean wind speed and the interaction between time and wind speed; although only mean wind speed (β = 0.757, SE = 0.252 , 95% CI 0.263 to 1.25) and the interaction term were significant (β = -0.00879, SE = 0.00303 , 95% CI -0.0147 to -0.00285) (Table 3-4, Fig. 5A). The logistic regression equation for this model was given by: logit(S) = -0.452 + 0.0444(T) + 0.757(wind_speed) – 0.00879(T:wind_speed).

Four models had Δ AIC < 2 when temporal and extrinsic factors were considered during the egg stage (w_i = 0.25, 0.25, 0.17, 0.08) (Table 3). The top two models only differed by 0.02 AIC_c and therefore had almost equal support. The first model showed a significant decrease in egg daily survival rate with linear time (β = -0.227, SE = 0.0292 , 95% CI -0.284 to -0.169) and as mean wind direction became more westerly (β = 2.14, SE = 0.867, 95% CI 0.440 to 3.84) (Table 3-4, Fig. 5B). The logistic regression equation was: logit(S) = 16.9 + -0.227(T) + 2.14(sin_dir). The second model showed a decrease in egg daily survival rate with linear time and mean wind speed, as well as the interaction between them, although none of these parameters were significant (Table 3-4, Fig. 5C). When all hypotheses were addressed together for the chick stage, four models had Δ $AIC_c < 2$. The most parsimonious model was 1.25 AIC_c from the next best model and showed chick daily survival rate as a function of quadratic time, mean eastern wind direction and the interaction term, however only the interaction had a significant effect (β = 0.00245, SE = 0.00104, 95% 0.000412 to 0.00448) (Table 3-4). Fig. 5D shows that the probability of a chick surviving from one day to the next is much lower towards the end of the chick-rearing period, although the presence of strong westerly winds could lower daily survival rate earlier in the season. The logistic regression equation for this model was: logit(S) = 6.458 + -0.00046(TT) – 1.34(sin_dir) + 0.00245(TT:sin_dir).

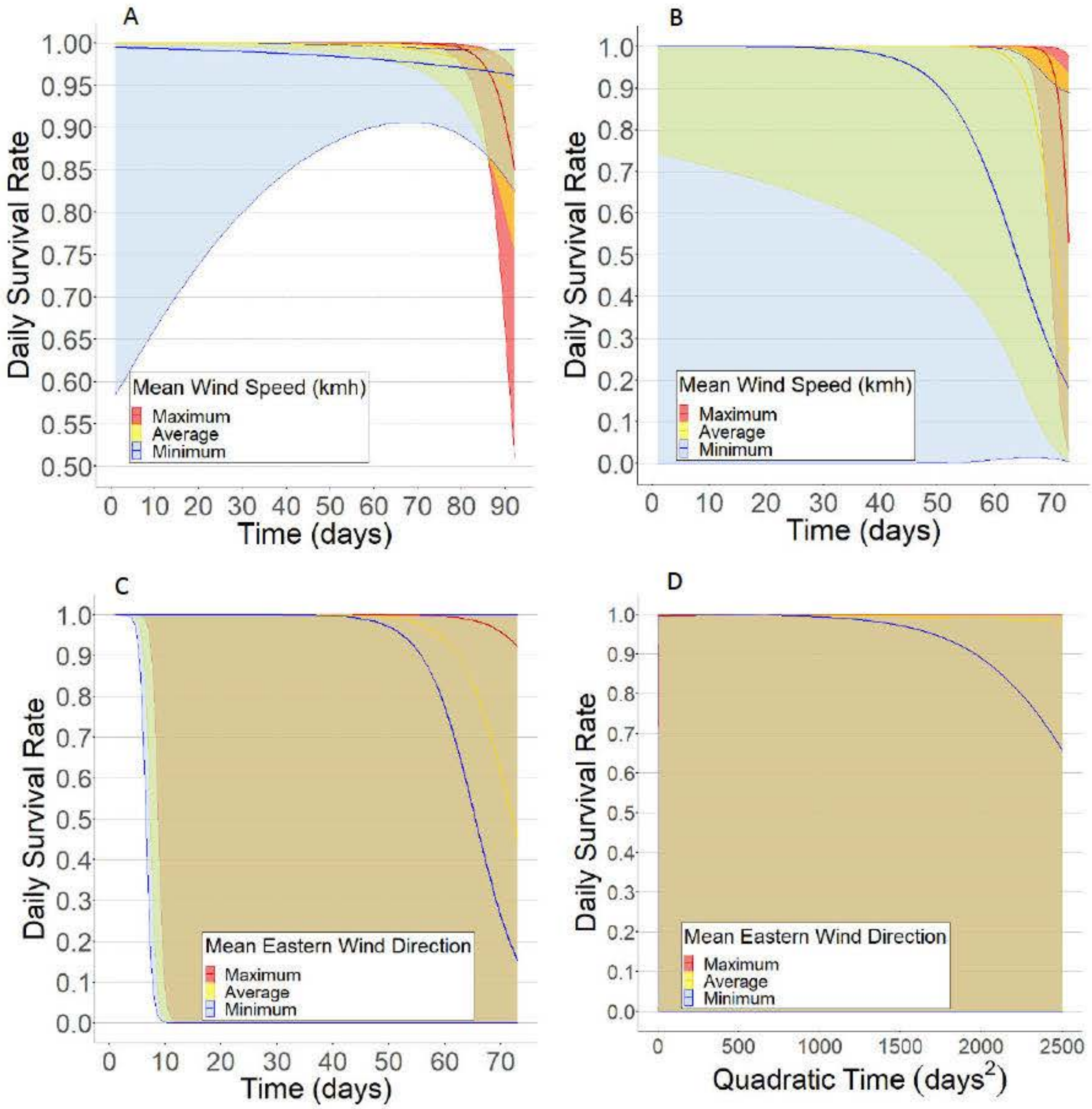


Figure 5 Estimated daily survival rate (S) of Kittiwake nests A) across the entire nesting period, B) during the egg stage only and C) during the chick stage only, when considering the effect of temporal and extrinsic weather variables on survival. Estimates are from the best model in hypothesis groups one and two combined for each breeding stage (Table 3). A) Model: $\text{logit}(S) = -0.452 + 0.0444(T) + 0.757(\text{wind\_speed}) - 0.00879(T{:}\text{wind\_speed})$. B) Model: $\text{logit}(S) = -9.09 + 0.0963(T) + 2.60(\text{wind\_speed}) - 0.0346(T{:}\text{wind\_speed})$. Egg daily survival rate decreased with time and at higher wind speed. High wind speed may contribute to a sharp decline in daily survival rate near the end of the egg stage. C) Model: $\text{logit}(S) = 16.9 + -0.227(T) + 2.14(\text{sin\_dir})$. Egg daily survival rated decreased with time and as the wind direction became more westerly. Strong westerly winds may lead to egg daily survival rate declining earlier in the season compared to when the wind is coming from the east. D) Model: $\text{logit}(S) = 6.458 + -0.00046(TT) - 1.34(\text{sin\_dir}) + 0.00245(TT{:}\text{sin\_dir})$. Chick daily survival rate decreased with quadratic time, meaning the probability of chick loss between two consecutive days increased towards the end of the chick-rearing period, and as the wind became more westerly. Strong westerly winds may lead to chick daily survival rate declining earlier in the chick-rearing period compared to when the wind is coming from the east.

Survival probability

Overall, assuming constant daily survival rate, mean daily survival probability was the same at 0.995 for the egg&chick stage (95% CI 0.993 to 0.996), egg stage only (95% CI 0.992 to 0.996) and chick stage only (95% CI 0.992, 0.997). Likewise, the total probability of survival across the period examined, assuming constant daily survival rate, was similar for all stages: 0.628 (95% CI 0.542 to 0.707) for the egg&chick stage, 0.686 (95% CI, 0.585 to 0.771) for the egg stage only, and 0.787 (95% CI 0.687 to 0.861) for the chick stage only (Table S5).

In comparison, the best model from hypothesis groups one and two combined for the egg&chick stage (logit(S) = -0.452 + 0.0444(T) + 0.757(wind_speed) – 0.00879(T:wind_speed)) predicted a mean daily survival rate of 0.992 and a total probability of survival across the entire nesting period of 0.493 (95% CI 0.347 to 0.640) when model parameters were kept at their average values (Table S5). For the egg stage, the best model for hypothesis groups one and two (logit(S) = 16.9 + -0.227(T) + 2.14(sin_dir)) gave a mean daily survival rate of 0.950 and a total probability of survival across the egg stage of 0.0120 (95% CI 0.00134 to 0.0988). For the chick stage, the most parsimonious model out of all hypotheses combined (logit(S) = 6.458 + -0.00046(TT) – 1.34(sin_dir) + 0.00245(TT:sin_dir)) gave a mean chick daily survival rate of 0.988 and a total probability of survival of 0.539 (95% CI 0.330, 0.735) (Table S5). However, it should be noted, that individual eggs/chicks are extremely unlikely to have survived the entire length of the egg stage (73 days)/ chick stage (50 days) without having hatched/fledged. On average, successful eggs hatched after 39.5 days and successful chicks fledged after 38.2 days. The mean daily survival rate and total probability of survival across the average length of the egg stage, from day nine to 48, were 0.999 and 0.971 respectively; while the corresponding values across the average length of the chick stage, from day six to 43, were 0.996 and 0.841. Day nine and day six are the mean day of egg-laying (first AON) and egg-hatching (first chick) respectively.

Table 3 Model selection results for models where $\Delta\,\mathrm{AIC}_c < 2$. Explanatory variable codes: *1* constant; *T* time; *TT* time^2; *max_temp* daily maximum temperature; *wind_speed* daily mean wind speed; *sin_dir* daily mean eastern wind direction; *cos_dir* daily mean northern wind direction. npar is the number of parameters in each model. See Supplementary Material for models where $\Delta\,\mathrm{AIC}_c > 2$.

Stage	**Hypothesis group**	**Model**	**npar**	**AIC$_c$**	**Δ AIC$_c$**	**w_i**	**Deviance**
Egg + Chick	1	~T	2	408.19	0	0.979	404.19
Egg	1	~T	2	128.96	0	0.920	124.96
Chick	1	~TT	2	164.72	0	0.809	160.72
Egg + Chick	2	~max_temp + sin_dir	3	397.27	0	0.683	391.27
Egg + Chick	2	~max_temp + sin_dir + max_temp:sin_dir	4	398.81	1.53	0.317	390.80
Egg	2	~max_temp + sin_dir + max_temp:sin_dir	4	131.98	0	0.842	123.97
Chick	2	~wind_speed + sin_dir	3	178.10	0	0.380	172.09
Chick	2	~wind_speed	2	179.73	1.63	0.168	175.72
Chick	3	~1	1	185.72	0	0.842	183.72
Egg + Chick	1 & 2	~T + wind_speed + T:wind_speed	4	395.38	0	0.486	387.38
Egg + Chick	1 & 2	~max_temp + sin_dir	3	397.27	1.89	0.189	391.27
Egg	1 & 2	~sin_dir + T	3	122.50	0	0.247	116.50
Egg	1 & 2	~T + wind_speed + T:wind_speed	4	122.52	0.016	0.245	114.51
Egg	1 & 2	~wind_speed + T	3	123.23	0.73	0.171	117.23
Egg	1 & 2	~sin_dir + T + sin_dir:T	4	124.43	1.93	0.0939	116.42
Chick	1, 2 & 3	~sin_dir + TT + sin_dir:TT	4	163.08	0	0.267	155.07
Chick	1, 2 & 3	~sin_dir + T + sin_dir:T	4	164.34	1.25	0.143	156.33

Chick	1, 2 & 3	~sin_dir + TT	3	164.65	1.56	0.122	158.64
Chick	1, 2 & 3	~TT	2	164.72	1.64	0.118	160.72

Table 4 Parameter estimates, standard error (SE), 95% lower (lcl) and upper (ucl) intervals for models within 2 Δ AIC$_c$ of the most parsimonious model. Parameter estimates are given on the logit scale (beta β estimates), rather than real estimates. Parameters that had a significant effect on daily survival rate are in bold. Explanatory variable codes: *1* constant; *T* time; *TT* time^2; *max_temp* daily maximum temperature; *wind_speed* daily mean wind speed; *sin_dir* daily mean eastern wind direction; *cos_dir* daily mean northern wind direction.

Stage	Hypothesis group	Model	Parameters	β Estimate	SE	lcl	ucl
Egg + Chick	1	~ T	**Intercept**	8.250	0.584	7.105	9.395
			T	-0.0541	0.00869	-0.0711	-0.0370
Egg	1	~ T	**Intercept**	15.560	1.553	12.515	18.604
			T	-0.222	0.0269	-0.274	-0.169
Chick	1	~ TT	Intercept	7.276	0.628	6.0450	8.508
			TT	-0.00201	0.000448	-0.00289	-0.00113
Egg + Chick	2	~max_temp + sin_dir	**Intercept**	22.241	2.438	17.463	27.0197
			max_temp	-0.585	0.0817	-0.745	-0.425
			sin_dir	3.082	0.522	2.0587	4.105
Egg + Chick	2	~max_temp + sin_dir + max_temp:sin_dir	**Intercept**	19.0255	5.323	8.593	29.458
			max_temp	-0.482	0.176	-0.826	-0.137
			sin_dir	-1.431	6.898	-14.952	12.090
			max_temp:sin_dir	0.146	0.225	-0.295	0.586
Egg	2	~max_temp + sin_dir + max_temp:sin_dir	**Intercept**	34.531	4.405	25.897	43.165
			max_temp	-1.073	0.150	-1.368	-0.778
			sin_dir	-23.443	7.172	-37.499	-9.386
			max_temp:sin_dir	**0.883**	**0.234**	**0.425**	**1.340**
Chick	2	~wind_speed + sin_dir	Intercept	8.435	1.042	6.394	10.477
			wind_speed	-0.134	0.0430	-0.219	-0.0501
			sin_dir	1.911	1.128	-0.301	4.122
Chick	2	~wind_speed	Intercept	7.460	0.777	5.936	8.984

			wind_speed	-0.128	0.0383	-0.203	-0.0528
Chick	3	~1	Intercept	5.316	0.236	4.853	5.780
Egg + Chick	1 & 2	~T + wind_speed + T:wind_speed	Intercept	-0.452	2.536	-5.423	4.519
			T	0.0444	0.0317	-0.0178	0.107
			wind_speed	0.757	0.252	0.263	1.250
			T:wind_speed	-0.00879	0.00303	-0.0147	-0.00285
Egg + Chick	1 & 2	~max_temp + sin_dir	**Intercept**	22.241	2.438	17.463	27.0197
			max_temp	-0.585	0.0817	-0.745	-0.425
			sin_dir	3.082	0.522	2.0587	4.105
Egg	1 & 2	~sin_dir + T	**Intercept**	16.918	1.884	13.225	20.612
			sin_dir	2.140	0.867	0.440	3.840
			T	-0.227	0.0292	-0.284	-0.169
Egg	1 & 2	~T + wind_speed + T:wind_speed	Intercept	-9.086	14.854	-38.200	20.0288
			T	0.0963	0.2109	-0.317	0.510
			wind_speed	2.595	1.690	-0.718	5.908
			T:wind_speed	-0.0346	0.0241	-0.0818	0.0126
Egg	1 & 2	~wind_speed + T	**Intercept**	13.834	2.0432	9.829	17.839
			wind_speed	0.243	0.118	0.0120	0.475
			T	-0.241	0.0351	-0.310	-0.172
Egg	1 & 2	~sin_dir + T + sin_dir:T	**Intercept**	17.947	4.471	9.184	26.711
			sin_dir	3.747	6.189	-8.379	15.873
			T	-0.244	0.0729	-0.387	-0.101
			sin_dir:T	-0.0271	0.103	-0.228	0.174
Chick	1, 2 & 3	~sin_dir + TT + sin_dir:TT	Intercept	6.458	0.587	5.307	7.610
			sin_dir	-1.344	1.112	-3.522	0.835
			TT	-0.00046	0.0009	-0.00223	0.00130
			sin_dir:TT	0.00245	0.00104	0.000412	0.00448
Chick	1, 2 & 3	~sin_dir + T + sin_dir:T	Intercept	7.394	0.869	5.691	9.0974
			sin_dir	-2.825	1.32073	-5.413	-0.236
			T	-0.0483	0.0376	-0.122	0.0254
			sin_dir:T	0.130	0.0396	0.0519	0.207
Chick	1, 2 & 3	~ sin_dir + TT	Intercept	7.628	0.701	6.255	9.00224
			sin_dir	1.190	0.907	-0.588	2.967
			TT	-0.00188	0.000442	-0.00275	-0.00102
Chick	1, 2 & 3	~TT	Intercept	7.276	0.628	6.0450	8.508
			TT	-0.00201	0.000448	-0.00289	-0.00113

DISCUSSION

Kittiwake nests were monitored during the 2019 breeding season to determine the relative importance of time, weather parameters, brood size and stage of the breeding season on nest survival. Overall, I found a strong time effect on nest survival when considering the egg and chick stages separately and egg and chick stages combined. I also found that mean eastern wind direction and wind speed were important during the egg and chick stage, while maximum temperature appeared to have an effect mainly during the egg stage. Brood size did not significantly affect chick survival.

Time

Time showed a negative relationship with overall nest, egg and chick survival (Fig. 3). The increased probability of an egg not surviving on a given day towards the end of the egg stage (as shown by the drop in egg daily survival rate) could suggest that eggs were more likely to fail nearer their hatching than laying date. However, it could also be explained by parental perception of egg failure. Unless eggs suffer an obvious misfortune (i.e. disappearance or damage) parents are unlikely to ‘know’ that their egg(s) have failed, so would continue to incubate eggs until they were ‘expected’ to hatch. Waiting a sufficient period of time before nest desertion could therefore lead to more eggs appearing to fail towards the end of the incubation period.

Equally, the increased probability of chick failure towards the end of the breeding season could suggest that chicks were more likely to fail just prior to fledging than just after hatching. The best supported model ($w_i = 0.49$) across the entire nesting period when considering all variables (hypothesis groups one & two) predicted that daily survival rate was a function of time, mean wind speed and the interaction term, with a large predicted increase in failures near the end of the breeding season at high wind speed. Nearly fledged Glaucous Gull *Larus hyperboreus* chicks have been blown off cliffs by strong downdrafts when testing their wings, suggesting that strong winds could hinder fledging success (Mallory *et al.* 2009). However, the drop off in chick daily survival rate at the end of the chick-rearing period, could partly be due to the nature of field observations. Following the UK Seabird Monitoring Programme guidelines, fieldworkers assume that only ‘large’ chicks are capable of fledging, and so a few medium and medium/large chicks remaining in late August were presumed to have failed (Walsh *et al.* 1995). This meant several chicks appeared to fail around the same time, despite their fate being unknown. The ability of cameras to record data outside of the field season could reduce this uncertainty.

Wind

Alongside temporal factors, eastern wind direction also consistently featured in the top models across the entire nesting period, egg and chick stage, while northern wind direction was not present. Daily survival rate typically lowered with increasingly westerly winds. When considering extrinsic factors (hypothesis group two) during the entire nesting period and egg stage only, eastern wind direction interacted with maximum temperature, such that daily survival rate was predicted to decrease at lower temperatures when the wind was coming from the west rather than the east. The orientation of Skomer Island means the Kittiwake colony is buffered from north, south and easterly winds by land (Fig. 1), however it is exposed to westerly winds which blow straight into the Wick. Previous studies have shown that Kittiwake chicks are vulnerable to chilling at low temperatures, the effects of which are exacerbated by wind and rain (Kennedy 1970, Kildaw 1999). However, temperature did not appear to have an important role in chick stage models, and 2019 was a particularly dry and warm year, with temperatures above average in May, June and July (Wilkie *et al.* 2019). It is perhaps more likely that westerly winds lowered survival rates by directly knocking eggs and/or chicks from their nests or generating large waves that washed them away. Across the entire breeding season and particularly during the chick stage, there was also support for increasing wind speed lowering daily survival rate, supporting the idea that strong westerly winds could have knocked or washed nests from the cliffs. It is also possible that chicks may drown or succumb to hypothermia after being drenched by waves, an effect made worse at lower temperatures, even if they were not washed from their nests into the sea (Seddon & van Heezik 1991, Sherley *et al.* 2012).

In 2020, many Guillemot *Uria aalge*, Razorbill *Alca torda* and Kittiwake eggs on Skomer and its neighbouring Island, Skokholm, were lost due to large waves generated by gales from the 21st - 23rd May (Brown & Eagle pers. obs., Wilkie & Zbijewska pers. obs.). In one plot on Skokholm 60% of Razorbill pairs lost their eggs, although 2/3rds re-laid from the 6th - 8th June (Brown & Eagle pers. obs.). Loss of nests due to storms is not an uncommon phenomenon and has been shown to reduce seabird nest survival (Mallory *et al.* 2009, Sherley *et al.* 2012, Newell *et al.* 2015, Christensen-Dalsgaard *et al.* 2018). Orientation of nests relative to wind direction is critical, as is the timing of inclement weather relative to the breeding season. Newell *et al.* (2015) recorded failure in 15.6% of Kittiwake nests in exposed plots compared to only 1.9% in sheltered plots following a summer storm. While some pairs re-laid, the success of re-lays was lower than pairs that survived, meaning re-laying only provided partial compensation (Newell *et al.* 2015). Storms occurring later in the breeding season are expected to result in even greater loss, as fewer pairs are likely to re-lay (Golet *et al.* 2004). Predicted increases in storm frequency with climate change could therefore have severe implications for Kittiwake breeding success (Seneviratne *et al.* 2012).

Temperature

This study also suggests that predicted increases in global surface temperature over the coming years may lower Kittiwake productivity (Seneviratne *et al.* 2012). During the egg stage there was support for high daily maximum temperatures having a negative effect on egg daily survival rate, although the magnitude of this effect was dependent on wind direction. At high temperatures eggs may fail due to heat stress (Ayo *et al.* 2011) or from indirect effects, like predation and disease. For example, heat-stressed Great Skua *Catharacta skua* on Foula, Shetland, regularly left chicks unattended at air temperatures >14°C in order to bathe (Oswald *et al.* 2008). Nest attendance was lowered further on hot days in years of poor food availability, when more time was also spent foraging. This reduced nest attendance increased the risk of chick predation by neighbouring conspecifics and likely contributed to low chick survival in warm, poor food years (Oswald *et al.* 2008). Furthermore, egg loss among Brünnich's Guillemot *Uria lomvia* at Coats Island, northern Hudson Bay, was associated with high daily maximum temperatures and mosquito abundance, but not high temperatures only (Gaston *et al.* 2002). During periods with high numbers of mosquitos, many nests were deserted, and in three instances egg predation was observed at deserted nests by Glaucous Gulls (Gaston *et al.* 2002).

On Skomer, it is therefore possible that Kittiwake nest attendance may be lower at higher temperatures, which could indirectly reduce egg daily survival rate. The fieldworker on Skomer witnessed nest predation at the Wick by Great Black-backed Gulls *Larus marinus* and a Peregrine Falcon *Falco peregrinus*, although it is unclear whether this was facilitated by reduced parental attendance at high temperatures (Wilkie *et al.* 2019). Predation was not seen when analysing one time-lapse photograph per day, although it may be observed by increasing the frequency of photo sampling. This could allow further exploration into the relationship between abiotic conditions and Kittiwake nest predation. Moreover, it is possible that high temperatures combined with strong westerly winds could leave eggs vulnerable to being blown or washed off cliffs, if parents abandon incubation duties at high temperatures. This relationship between temperature and wind direction may be less pertinent during the chick stage, because larger chicks are regularly left alone in the nest as they are better able to cope with external conditions. Conversely, eggs and small chicks are very vulnerable if left unattended. Eggs would normally be incubated continuously until hatching and then small chicks brooded until ~15-16 days post-hatching when they become homeothermic (Gabrielsen *et al.* 1992). Further research into the exact mechanism by which daily maximum temperature and wind direction may be affecting Kittiwake egg survival is needed.

Alternative factors

Daily maximum temperature did not appear to negatively affect chick survival, and in general there was less support for models explaining survival as a function of extrinsic variables during the chick stage, compared to the egg stage and egg&chick stage. The best models for hypothesis group two (extrinsic factors) in the chick stage explained 38% of the variation in daily survival rate, compared to 84% for the egg stage and 68% for the egg&chick stage. This could suggest that weather variables had a greater effect on egg than chick survival. During storms, parents may leave nests to avoid being buffeted by waves, leaving eggs and small chicks vulnerable to being blown or washed from nests, while medium/large chicks should be better able to withstand the assault (Sherley *et al.* 2011). Equally, parents may be more likely to brood chicks and provide protection during inclement weather compared to eggs, as they have invested considerable energy in reproduction by the time chicks have hatched, but much less when their reproductive attempt is only at the egg stage. Furthermore, other factors not investigated in this study, such as prey availability, may have been more important than the direct effects of weather in influencing the probability of a chick surviving to fledge.

Food availability has a direct effect on chick survival but only an indirect effect on egg survival. Prey availability can, in turn, be affected by weather, as wind and rain can inhibit foraging and lower food provisioning rates to chicks. Increased wind speed can reduce prey visibility to surface feeders, like Kittiwakes, by increasing turbidity of surface waters (Taylor 1983, Mitchell *et al.* 2018). High wind speed will also increase energy expenditure during flight and mean adults have to work harder to catch prey (Christensen-Dalsgaard *et al.* 2018).

I predicted that higher brood size would lower survival if food was limited. The probability of at least one chick fledging was not affected by the number of chicks per nest, suggesting food provisioning may have been sufficient. However, this analysis did not assess how brood size affects the likelihood of different numbers of chicks fledging (e.g. zero, one or two) and it did not consider how many chicks fledged relative to the number of eggs laid. Moreover, the largest brood size on Skomer was two, despite three-chick broods being observed at other Kittiwake colonies (Coulson 2011). Prey availability could therefore already be limiting the number of eggs laid, although this in itself is dependent on a variety of factors, including parental fitness, age and conditions during the non-breeding season (Coulson & White 1961, Coulson & Porter 1985, Chastel *et al.* 1995, Coulson 2011).

CONCLUSION

The relationship between temporal, extrinsic and intrinsic factors on Kittiwake nest survival is complicated. This study provides initial support for the hypothesis that weather variables may be contributing to nest failure on Skomer Island, Wales. Digitising nest monitoring data going back to the early 1990s is of priority to allow this preliminary survival analysis to be extended and evaluate the impact of weather on Kittiwake productivity over the past few decades. With more years of data, the effect of ‘extreme’ versus ‘average’ weather could also be investigated, to assess the impact of climate change on breeding success. This could provide insight into the gradual decline in Kittiwake productivity on Skomer Island, and with expansion to landscape level monitoring enabled by novel imaging technologies, like time-lapse cameras, it could help identify factors contributing to population decline in Wales.

SUPPLEMENTARY MATERIAL

Table S1 VIF for explanatory variables with VIF < 3 when assessing collinearity of all candidate variables. NA indicates that this variable was excluded through the stepwise procedure implemented using the ‘vifstep’ function in R. Explanatory variable codes: *T* time; *TT* time^2; *max_temp* daily maximum temperature; *min_temp* daily minimum temperature*;* *mean_temp* daily mean temperature*; total rain* daily total rainfall; *wind_speed* daily mean wind speed; *sin_dir* daily mean eastern wind direction; *cos_dir* daily mean northern wind direction.

Explanatory variables	**Egg + chick stage**	**Egg stage**	**Chick stage**
max_temp	2.070827	2.442336	NA
min_temp	NA	NA	2.581015
mean_temp	NA	NA	1.267133
total_rain	1.695375	1.940486	1.635243
wind_speed	1.632525	1.487364	1.779399
cos _dir	1.207371	1.244018	2.305048
sin _dir	1.080126	1.082939	1.186121
T	NA	1.289431	NA
TT	1.524245	NA	1.422533

Table S2 VIF for explanatory variables with VIF < 3 when assessing collinearity of time or quadratic time, maximum temperature and mean wind speed and direction. ‘Not included’ means the variable was not used in the collinearity assessment. Explanatory variable codes: *T* time; *TT* time^2; *max_temp* daily maximum temperature; *wind_speed* daily mean wind speed; *sin_dir* daily mean eastern wind direction; *cos_dir* daily mean northern wind direction.

Explanatory variables	**Egg + chick stage**	**Egg stage**	**Chick stage**
max_temp	1.560995	1.518730	1.366613
wind_speed	1.520868	1.481102	1.551015
cos_ dir	1.204095	1.234330	1.393939
sin_ dir	1.019785	1.033824	1.184155
T	1.382341	Not included	1.278736
TT	Not included	1.269624	Not included

Table S3 Model selection results for the top three models in each stage and hypothesis group. Explanatory variable codes: 1 constant; *T* time; *TT* time^2; *max_temp* daily maximum temperature; *wind_speed* daily mean wind speed; *sin_dir* daily mean eastern wind direction; *cos_dir* daily mean northern wind direction. npar is the number of parameters in each model.

Stage	**Hypothesis group**	**Model**	**npar**	**AIC$_c$**	**Δ AIC$_c$**	**w$_i$**	**Deviance**
Egg + Chick	1	~T	2	408.1934	0	0.979075	404.1921
Egg + Chick	1	~TT	2	415.885	7.69131	0.020925	411.8834
Egg + Chick	1	~1	1	457.253	49.05958	2.18E-11	455.2526
Egg	1	~T	2	128.961	0	0.91972	124.9591
Egg	1	~TT	2	133.838	4.8771	0.08028	129.8362
Egg	1	~1	1	273.481	144.5198	0	271.4804
Chick	1	~TT	2	164.721	0	0.809136	160.7178
Chick	1	~T	2	167.6101	2.88904	0.190842	163.6068
Chick	1	~1	1	185.7219	21.00086	2.23E-05	183.7208
Egg + Chick	2	~max_temp + sin_dir	3	397.2741	0	0.682776	391.2715
Egg + Chick	2	~max_temp + sin_dir + max_temp:sin_dir	4	398.8072	1.533151	0.317218	390.8029
Egg + Chick	2	~max_temp + wind_speed + max_temp:wind_speed	4	420.5746	23.30051	5.95E-06	412.5702
Egg	2	~max_temp + sin_dir + max_temp:sin_dir	4	131.9793	0	0.842331	123.972
Egg	2	~max_temp + wind_speed + max_temp:wind_speed	4	135.3773	3.39806	0.154029	127.37
Egg	2	~max_temp + sin_dir	3	142.8692	10.88998	0.003637	136.8649
Chick	2	~wind_speed + sin_dir	3	178.0967	0	0.379768	172.0901

Chick	2	~wind_speed	2	179.7262	1.629511	0.168141	175.7229
Chick	2	~wind_speed + sin_dir + sin_dir:wind_speed	4	180.098	2.001308	0.139617	172.087
Chick	3	~1	1	185.7219	0	0.84225	183.7208
Chick	3	~brood_size	3	189.072	3.350132	0.15775	183.0655
Egg + Chick	1 & 2	~T + wind_speed + T:wind_speed	4	395.3827	0	0.486063	387.3783
Egg + Chick	1 & 2	~max_temp + sin_dir	3	397.2741	1.891379	0.188793	391.2715
Egg + Chick	1 & 2	~sin_dir + T	3	397.9549	2.572199	0.134322	391.9523
Egg	1 & 2	~sin_dir + T	3	122.4998	0	0.24658	116.4955
Egg	1 & 2	~T + wind_speed + T:wind_speed	4	122.516	0.016137	0.244598	114.5087
Egg	1 & 2	~wind_speed + T	3	123.2311	0.7313	0.171064	117.2268
Chick	1, 2 & 3	~sin_dir + TT + sin_dir:TT	4	163.0838	0	0.267391	155.0728
Chick	1, 2 & 3	~sin_dir + T + sin_dir:T	4	164.3387	1.2549	0.142774	156.3277
Chick	1, 2 & 3	~sin_dir + TT	3	164.6483	1.564502	0.122298	158.6417

Table S4 Real parameter estimates, standard error (SE) and 95% lower (lcl) and upper (ucl) levels for significant parameters in the most parsimonious model for each stage and hypothesis group. Explanatory variable codes: 1 constant; *T* time; *TT* time^2; *max_temp* daily maximum temperature; *wind_speed* daily mean wind speed; *sin_dir* daily mean eastern wind direction; *cos_dir* daily mean northern wind direction.

Stage	**Hypothesis group**	**Model**	**npar**	**Real Estimate**	**SE**	**lcl**	**ucl**
Egg + Chick	1	~T	T	0.486	0.502	0.482	0.491
Egg	1	~T	T	0.445	0.507	0.432	0.458
Chick	1	~TT	TT	0.499	0.500	0.499	0.500
Egg + Chick	2	~max_temp + sin_dir	max_temp	0.358	0.520	0.322	0.395
			sin_dir	0.956	0.628	0.887	0.984
Egg	2	~ max_temp + sin_dir + max_temp:sin_dir	max_temp	0.255	0.538	0.203	0.315
			sin_dir	6.59E-11	0.999	5.18E-17	8.39E-05
			max_temp:sin_dir	0.707	0.558	0.605	0.793
Chick	2	~ wind_speed + sin_dir	wind_speed	0.466469	0.511	0.446	0.487
Egg + Chick	1 & 2	~T + wind_speed + T:wind_speed	wind_speed	0.680663	0.563	0.565	0.777
			T:wind_speed	0.498	0.501	0.496	0.499
Egg	1 & 2	~ sin_dir + T	T	0.444	0.507	0.429	0.458
			sin_dir	0.895	0.704	0.608	0.979
Chick	1 & 2	~ sin_dir + TT + sin_dir:TT	sin_dir:TT	0.501	0.500	0.500	0.501

Table S5 Mean daily survival probability and total probability of survival for the null hypothesis (Survival~1) and the most parsimonious model for each stage and hypothesis group. Model parameters were kept at their average values and real parameter estimates were used, rather than beta (β) parameter estimates. Explanatory variable codes: 1 constant; *T* time; *TT* time^2; *max_temp* daily maximum temperature; *wind_speed* daily mean wind speed; *sin_dir* daily mean eastern wind direction; *cos_dir* daily mean northern wind direction.

Stage	**Hypothesis group**	**Model**	**Mean daily survival probability**	**Probability of surviving the duration of the study**
Egg + Chick	1	~1	0.9949 (95% CI 0.993-0.996)	0.6284 (95% CI 0.542-0.707)

Egg + Chick	1	~T	0.9931	0.530 (95% CI 0.425-0.632)
Egg	1	~1	0.9948 (95% CI 0.992-0.996)	0.686 (95% CI 0.585-0.771)
Egg	1	~T	0.947	0.00897 (95% CI 0.000856- 0.0873)
Chick	1	~1	0.995 (95% CI 0.992, 0.997)	0.787 (95% CI 0.687, 0.861)
Chick	1	~TT	0.991	0.631 (95% CI 0.447, 0.783)
Egg + Chick	2	~max_temp + sin_dir	0.9951	0.633 (95% CI 0.547-0.712)
Egg	2	~ max_temp + sin_dir + max_temp:sin_dir	0.986	0.266 (95% CI 0.146-0.435)
Chick	2	~ wind_speed + sin_dir	0.992	0.673 (95% CI 0.475, 0.825)
Egg + Chick	1 & 2	~T + wind_speed + T:wind_speed	0.9923	0.493 (95% CI 0.347-0.640)
Egg	1 & 2	~sin_dir + T	0.950	0.0120 (95% CI 0.00134-0.0988)
Chick	1 & 2	~ sin_dir + TT + sin_dir:TT	0.988	0.539 (95% CI 0.330, 0.735)

CHAPTER THREE: Comparing citizen science and expert annotated imagery from time-lapse photographs

ABSTRACT

Citizen science has the potential to expand the magnitude and scope of scientific investigations and enable research that would not otherwise be possible or practical, in terms of time and resources. *Seabird Watch* is one such citizen science project, which asks volunteers to count seabirds in time-lapse images. However, before data from this approach can be used to answer larger scientific questions related to seabird productivity and phenology, it is essential to validate volunteer data against 'gold standard' classification provided by experts. This study compared aggregated citizen science counts and gold standard counts of adult and juvenile Black-legged Kittiwake *Rissa tridactyla* in images from five sites and spanning three years. Using the optimal aggregation threshold of > 2 (meaning at least three users must have clicked on a bird for it to be counted as a Kittiwake), the lowest average percentage differences between gold standard and citizen science counts ranged from 3.35% to 21.80% Kittiwakes, but could exceed 70.00% when comparing counts of chicks only. Accuracy of volunteer counts varied depending on a number of factors, including aggregation threshold, whether adults and chicks were considered together or separately, and site/camera, which was related to camera setup, the number of birds in the field of view and species present. I conclude that validation of *Seabird Watch* data must occur independently for each site, and that whilst aggregated volunteer results are currently not 'as good as' a single expert, this is likely due to issues with the clustering algorithm rather than the performance of *Seabird Watch* volunteers,

INTRODUCTION

Citizen science, the engagement of the general public in the process of science, has the potential to greatly expand the magnitude and scope of research (Swanson *et al.* 2016). In the last decade there have been significant advances in online citizen science, with the creation of the *Zooniverse* platform being one such example (Cox *et al.* 2015). The *Zooniverse* is 'the world's largest and most popular platform for people-powered research' (Zooniverse 2020a) and engages > 1.1. million registered users worldwide (Cox *et al.* 2015). The website (www.zooniverse.org) comprises numerous projects that span a range of disciplines from astronomy to ecology, but all of which use volunteers to analyse large datasets and therefore enable research that would not otherwise be possible or practical (Cox *et al.* 2015, Zooniverse 2020a). Motivation to participate in projects largely stems from a desire to contribute to science and the ability to engage in scientific discussion via the *Zooniverse* discussion forums (Raddick *et al.* 2010, 2013, Mugar *et al.* 2014, Swanson *et al.* 2016). However, despite the enormous potential of citizen science, many researchers remain apprehensive of non-expert derived data (Foster-Smith & Evans 2003, Dickinson *et al.* 2010, Bonter & Cooper 2012, Swanson *et al.* 2016). Citizen science projects have used a range of methods to improve quality of volunteer data, including training volunteers prior to participation, requiring volunteers to pass a competency test, or, in the case of many *Zooniverse* projects, aggregating the results of multiple users to form a consensus (Dickinson *et al.* 2010, Swanson *et al.* 2016).

Seabird Watch is a *Zooniverse* project established in 2017, which uses volunteers to mark seabirds in images, from which a count can then be obtained. A clustering algorithm aggregates the raw clicks of multiple users to form 'consensus' clicks which are summed to give the number of birds in the image. Specifically, adult and juvenile Kittiwakes and Guillemots are counted. So far, *Seabird Watch* has 10,782 registered users who have made 533,604 classifications across 54,116 images. These images are from a network of cameras located at 13 sites and span more than five years, meaning there is now sufficient data across a large enough temporal and spatial scale to inform meaningful analyses, such as investigating the phenology and nest success of breeding seabirds (Zooniverse 2020b). Nonetheless, before such analyses are undertaken, it is essential to validate the data against gold standard classifications provided by experts, and therefore ensure the citizen science output is reliable (Swanson *et al.* 2016, Bruggemann *et al.* 2018).

This chapter aims to assess the accuracy of the *Seabird Watch* dataset and determine whether aggregated volunteer data are of equal or greater accuracy than the results from a single expert. The methods used here are similar to those developed to validate *Penguin Watch* (Jones *et al.* 2018), so all conclusions made are comparable for these two sister projects. Accuracy was predicted to vary by aggregation threshold level, site/camera (related to camera setup, number of birds per image and image resolution) and whether adults and chicks were considered together or separately.

METHODS

Data collection

Study sites and camera setup

Photographs were obtained from time-lapse cameras (Reconyx Hyperfire or Reconyx Ultrafire) setup as part of the citizen science project *Seabird Watch.* Each camera is programmed to take one photo every hour throughout the year, although gaps in the dataset exist during some winters and when annual visits to change the SD cards and batteries were not possible. Images from five cameras deployed at five colonies (3 in Svalbard, 1 in Iceland, and 1 in Ireland) were chosen to represent a range of 'views', as the cameras are positioned at different distances and angles to the colonies of interest (Table 1). Based on differing 'views' each camera was ranked according to the perceived likelihood of accurately recording nest contents (Table 1). In a few instances camera position changed between years (MITTa2015a to MITTa2017a, OSSIa2016a to OSSIa2017a), however these cameras were still treated as the same site to prevent small sample sizes (< 30 images) limiting the reliability of analysis.

Citizen science image annotation

Images were annotated by members of the public via the citizen science project, *Seabird Watch* (www.seabirdwatch.org), on the *Zooniverse* platform (www.zooniverse.org). Users are given a brief tutorial, which provides examples of how birds appear in images and explains how to select the correct identification tool, before they are shown their first image and follow the workflow outlined in Fig. 1 to identify adult and chick Kittiwakes and Guillemots. A field guide of animals likely to be seen in images is also available to aid identification, as well as a Frequently Asked Questions page and online discussion forum which is moderated by scientists.

Each image is first shown to four participants and if no birds are identified or the image is too dark/blurry to classify, the image is retired. This means it is removed from the active dataset and not seen by further volunteers. However, if any of the four participants identifies an animal, then the image is shown to an additional 11 people before being retired, meaning 15 users view each image containing birds. Previously, 20 users viewed each image containing birds (i.e. the subject retirement was 20), although this was lowered to 15 to increase the rate of image retirement. Participants do not have to classify every bird in each image to help prevent loss of interest when the photograph contains many individuals. It is assumed that every bird will be classified between the 15 independent viewers, meaning completed images could, in theory, have up to 15 marks for each animal. These classifications are aggregated using a clustering algorithm to generate a single ‘consensus click’ for each marked object. This method uses agglomerative hierarchical clustering to group coordinates of clicks by their spatial position and produce an average x,y coordinate for each tightly clustered group of markings, thus averaging volunteer input (Jones *et al.* 2018). Clicks made by the same user are placed in separate clusters.

Output data were processed to produce comma-separated values (csv) files containing: image name (image_id), consensus click coordinates, number of markings contributing to each consensus click (num_markings), number of people that marked all birds in the image (marked_all), number of people that said the image was too dark/blurry to classify (dark_num_votes), and date, time and temperature when the image was taken. Images were not carried forward for analysis if zero volunteers marked all birds, as these do not reflect the true number of birds per image. Equally, images were excluded if greater than three people said the image was too dark/blurry to classify. Consensus clicks had to be formed from greater than one mark (num_markings > 1) to be included in the dataset. This helped filter out accidental markings.

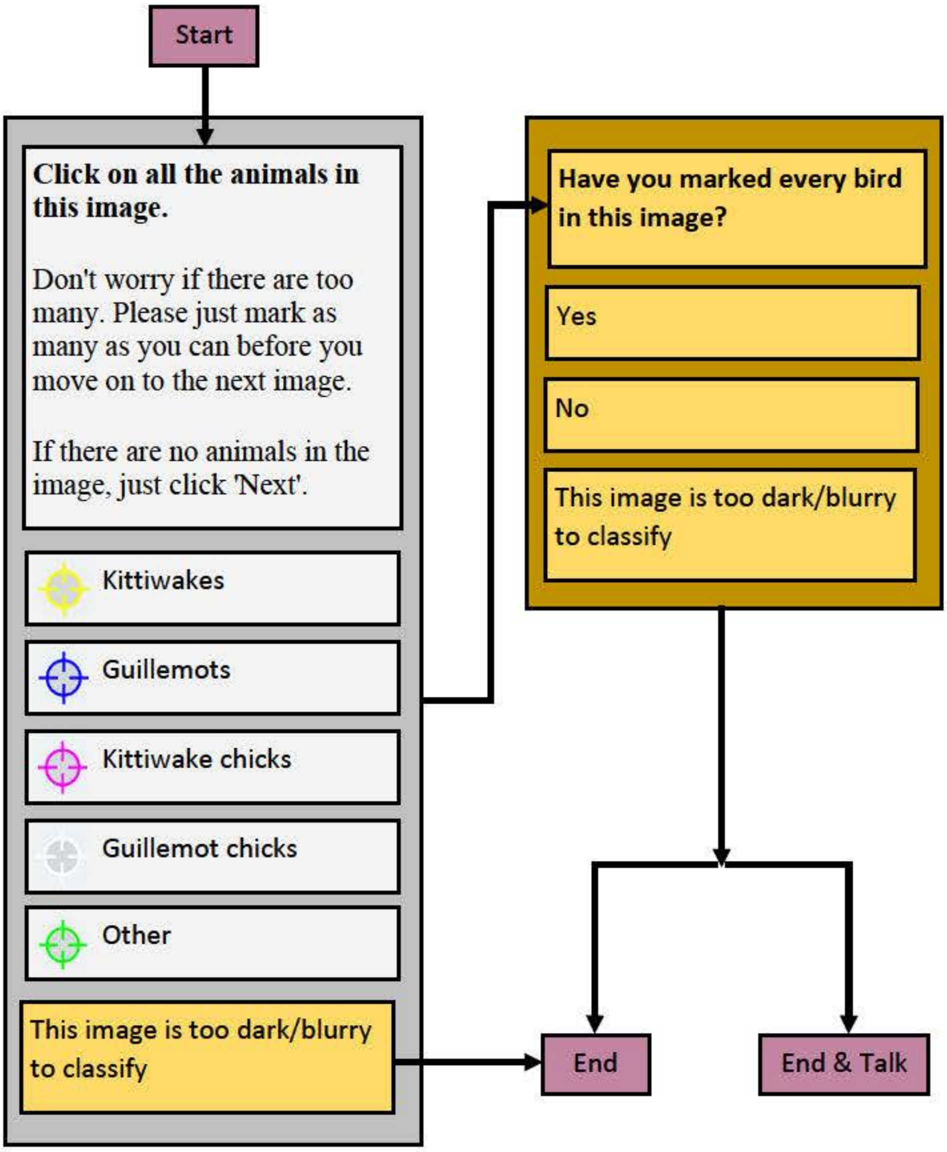


Figure 1 The *Seabird Watch* workflow. If the image contains an animal, then participants are asked to mark individuals by clicking on them, and classify individuals as 'Kittiwakes', 'Guillemots', 'Kittiwake chick', 'Guillemot chick' and 'Other'. After classification, participants can choose whether they would like to 'talk' about it on a *Seabird Watch* forum. Brown boxes mean participants must give an answer, grey boxes mean a process must be carried out, such as clicking on animals. This workflow is based on *Penguin Watch* (Jones *et al*. 2018).

Gold standard image annotation

Gold standard annotation was performed using an inactive version of the *Seabird Watch* project (*kittiwake project*), which is only available to researchers. This allowed specific sets of images to be uploaded and viewed. Conversely, *Seabird Watch* presents the user with random images from a very large dataset.

Images from MITTa, OSSIa, HVITa and SKELa were annotated by myself, while images from ALKEa were classified by three other researchers (Rachel Ryan, Mark Jessopp and Tom Hart) as part of a previous project. For HVITa the 13.00 images during the breeding period were annotated; whereas the 10.00, 12.00 and 13.00 images were randomly sampled across the dataset at OSSIa, and the 12.00 images were randomly sampled at SKELa and ALKEa. At MITTa some 11.00-13.00 images were randomly selected for annotation, but a continuous series of 12.00 images from 2014-2015 were also analysed. By analysing a continuous time-series of images (i.e. one image per day for a set time period) at MITTa and HVITa it allows future work to compare phenology measured from gold standard and citizen science annotation, rather than just comparing counts. Photographs from MITTb were also classified but too few images had citizen science annotations available so they could not be used in this study.

Table 1 Location, sampling method and rank (R) of each camera providing images for gold standard and citizen science comparison. Cameras were ranked according to perceived likelihood of accurately recording nest contents from the images, where 1 gave the best view of nest contents (i.e. highest quality image) and 5 gave the worst view. Random sampling means images were chosen at random from the dates and times provided. Time-series sampling means one image per day at the given time, was selected between a specific date range.

Camera name	Location	Country	GPS	Sampling method	R
MITTa (MITTa2015a, MITTa2017a)	Midterhukfjellet	Svalbard	77.663908°N, 14.87916°E	Random Times: 11.00, 12.00, 13.00 Dates: 10/06/2014 - 20/06/2015 & 02/07/2016 - 24/03/2017 Time-series Time: 12.00 Dates: 10/06/2014 - 20/06/2015	4
OSSIa (OSSIa2016a, OSSIa2017a)	Ossian Sarsfjellet	Svalbard	78.941015°N, 12.491033°E	Random Times: 10.00, 12.00, 13.00	2

				Dates: 12/07/2015 – 7/08/2015 & 6/07/2016 - 5/06/2017	
ALKEa2016a	Alkhornet	Svalbard	78.211362°N, 13.783972°E	Random Time: 12.00 Dates: 25/06/2015 – 19/06/2016	3
HVITa (HVITa2016a, HVITa2016b)	Hvitabjarnarey	Iceland	65.08°N, 22.68°W	Time-series Time: 13.00 Dates: 22/06/2016 – 20/08/2016	1
SKELa2015a	Skellig Michael	Ireland	51.771486°N, 10.539689°W	Random Time: 12.00 Dates: 08/09/2014 – 29/08/2015	5

Data analysis

Comparing citizen science and gold standard analysis

Images without Kittiwakes

For images where gold standard annotation stated that no Kittiwakes were present, I compared the total number of Kittiwakes (adults and chicks combined) per image between gold standard (GS) and citizen science (CS) classification. I calculated the average difference (GS – CS) in the number of Kittiwakes per image, and the proportion of images where the CS count was equal to (GS = CS) and greater than (GS < CS) the GS count (i.e. matches and overestimates). Accuracy was dependent on the average count difference, where higher accuracy was defined as a smaller average difference between GS and CS counts.

Images with Kittiwakes

For images where GS annotation stated that at least one Kittiwake (adult or chick) was present, I calculated the number of adults, number of chicks and number of adults and chicks combined for GS and CS classification (consensus clicks) at different threshold levels of num_markings and marked_all. The threshold levels used ranged from two to 10 for num_markings and one to 10 for marked_all. The lowest threshold level of num_markings was > 2 because the minimum number of raw clicks that formed a consensus click was three, meaning num_markings > 1 and num_markings > 2 would give the same results. I calculated the average difference (GS – CS), average percentage difference ((GS – CS / GS) * 100) and proportion of images where the CS count was equal to (GS = CS), greater than (GS < CS) and less than (GS > CS) the GS count, for adults only, chicks only and adults and chicks combined at each threshold level. For comparison of chicks only, I used images where either GS or CS annotation recorded at least one chick. This prevented images with adult birds but no chicks being included in the chick only comparison. For calculations of average percentage difference, images containing zero birds in GS classification were removed, as it is not possible to divide by zero.

Comparisons were conducted separately for each site/camera to assess how different image 'views' might affect the reliability of citizen science classification relative to gold standard. For images with large differences between gold standard and citizen science counts, the x,y coordinates of each click (consensus click for citizen science classifications) were plotted onto the image to investigate the cause of the discrepancy.

RESULTS

Comparing citizen science and gold standard analysis

Images without Kittiwakes

The percentage agreement between GS and CS classifications in images where GS classification recorded zero Kittiwakes, was highest at HVITa and OSSIa (100%). All images were marked as containing no Kittiwakes. This was followed by MITTa (97%), ALKEa (96%), and SKELa (88%) (Table 2). The single SKELa image where GS and CS classifications differed had been erroneously marked by *Seabird Watch* volunteers, leading to a count difference of one. For MITTa, nine images recorded Kittiwakes in CS classifications but not in GS classifications. In one image, two Kittiwakes had been erroneously marked by *Seabird Watch* volunteers, and for two images GS classification missed one Kittiwake. Three images were partially obscured by a bird in the foreground and were marked as 'too dark/blurry to classify' by the gold standard analyser and one to three *Seabird Watch* volunteers, but still had some Kittiwakes marked from volunteer consensus clicks. In the remaining three images, *Seabird Watch* volunteers recorded greater than 30 Kittiwakes (49, 71 and 117 adults) compared to zero in gold standard classification, despite these images being of good quality. Likewise, three images at ALKEa had between 26 and 69 adult Kittiwakes recorded from CS classification but zero from GS classification. It should be noted that image sample size was < 30 for HVITa (n = 2), OSSIa (n = 26) and SKELa (n = 8) meaning the results may not be representative of the true agreement between GS and CS classifications in images with zero Kittiwakes.

Table 2. Comparison of counts between gold standard (GS) and citizen science (CS) classification in images where GS classification recorded zero Kittiwakes. 'Average difference' is the mean GS minus CS count and σ is the standard deviation. 'Proportion of differences 0 or 1' is the proportion of images for which the CS count was equal to the GS or different by one individual. The proportion of GS classifications that are equal to (GS = CS) and less than (GS < CS) CS classifications is also given. 'n' is the number of images in the sample. The results presented are for images where GS classification did not record any Kittiwakes.

Camera	**Average difference**	σ	**Proportion of differences 0 or 1**	**GS < CS**	**GS = CS**	**n**
HVITa	0	0	1	0	1	2
OSSIa	0	0	1	0	1	26
ALKEa	-1.58228	8.871747	0.962025	0.037975	0.962025	79
MITTa	-0.95	8.761127	0.982143	0.032143	0.967857	280
SKELa	-0.125	0.353553	1	0.125	0.875	8

Images with Kittiwakes

Accuracy was predicted to vary by threshold level, site/camera and whether adults and chicks were considered together or separately.

Threshold values

When comparing Kittiwake counts between GS and CS classifications at different threshold levels of num_markings (the number of volunteer clicks aggregated to produce a consensus click), the most accurate results (i.e. lowest average differences) occurred when num_markings was > 2 for all cameras (Table S1-3). At this threshold level, the average differences ranged from 0.12 (σ = 11.64) to 17.09 (σ = 43.49) Kittiwakes, and the average percentage differences ranged from -0.26% (σ = 70.05) to 79.50% (σ = 20.68). Increasing num_markings typically increased both the average difference and average percentage difference and therefore decreased accuracy.

Using num_markings > 2, the percentage agreement between GS and CS classifications (GS = CS) was low for all cameras. It ranged from 0% at OSSIa and SKELa when comparing counts of adults and chicks combined and chicks only respectively, to 15% at SKELa when comparing counts of adults and chicks combined (Fig. 2).

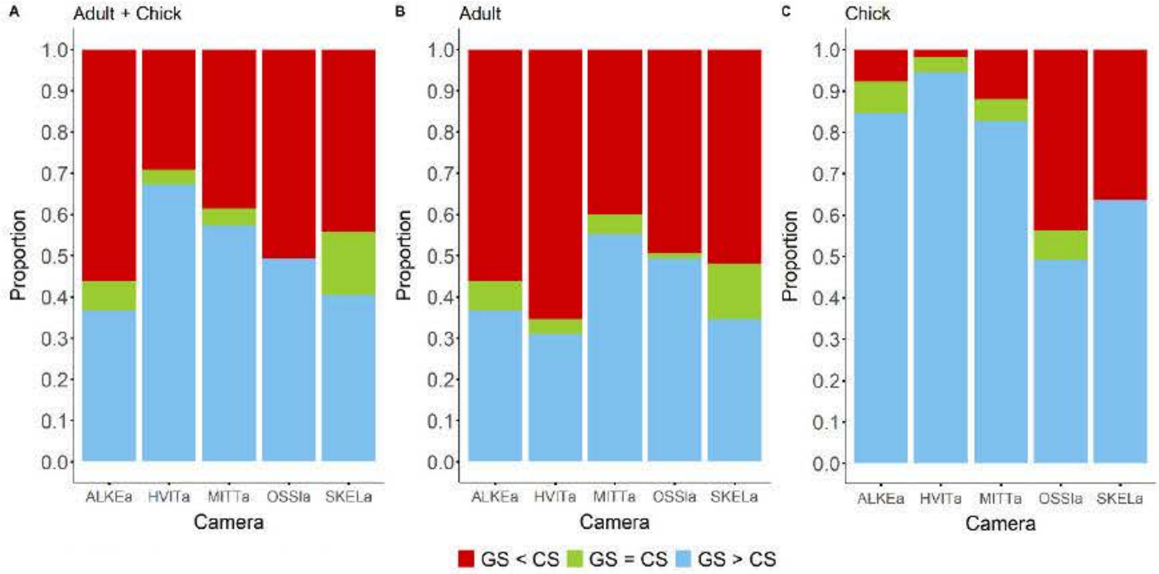


Figure 2 Proportion of gold standard (GS) classifications that are greater than (GS > CS), equal to (GS = CS) and less than (GS < CS) citizen science (CS) classifications for num_markings > 2 when comparing counts of A) adults and chicks, B) adults and C) chicks at each camera.

When comparing Kittiwake counts at different threshold levels of marked_all (the number of volunteers that marked all birds in the image), the threshold that gave the most accurate results varied by camera. Generally, the most accurate results were obtained at higher threshold levels of marked_all compared to num_markings (> 2) and varying marked_all gave more accurate results than varying num_markings. The lowest average differences obtained when adjusting the marked_all threshold ranged from 0 ($\sigma = 13.63$) to 5.19 ($\sigma = 4.79$) Kittiwakes and 0.01% ($\sigma = 27.85$) to 17.49% ($\sigma = 45.22$) average percentage difference (Table S1-3). Figures 3-4 show how average difference for adults and chicks changed with increased num_markings and marked_all thresholds.

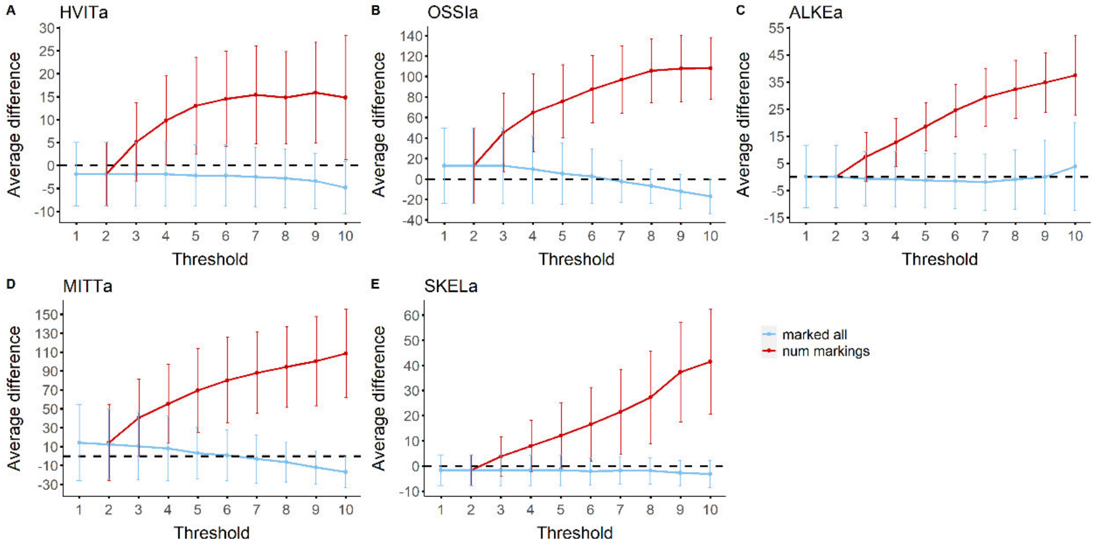


Figure 3 Average differences (+/- σ) between gold standard (GS) and citizen science (CS) **adult** Kittiwake counts for different threshold levels of num_markings and marked_all at cameras A) HVITa, B) OSSIa, C) ALKEa, D) MITTa and E) SKELa. Average difference is the mean GS minus CS count. A 'num_markings' threshold of '> 2' means at least three *Seabird Watch* volunteers must have clicked an area for it be counted as a Kittiwake. A 'marked_all' threshold of '> 1' means at least two *Seabird Watch* volunteers marked all birds in the image.

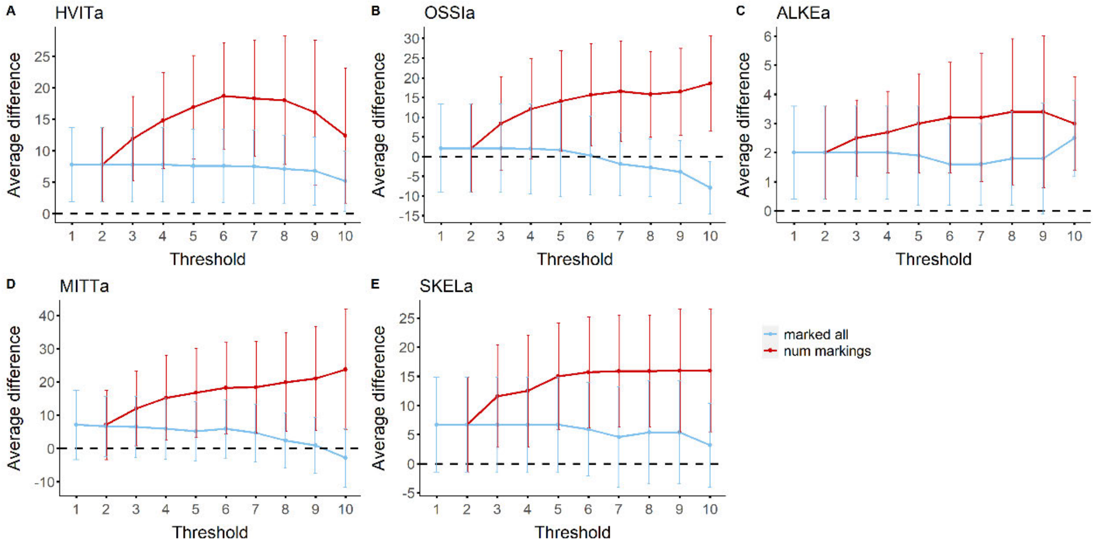


Figure 4 Average differences (+/- σ) between gold standard (GS) and citizen science (CS) Kittiwake **chick** counts for different threshold levels of num_markings and marked_all at cameras A) HVITa, B) OSSIa, C) ALKEa, D) MITTa and E) SKELa. Average difference is the mean GS minus CS count. A 'num_markings' threshold of '> 2' means at least three *Seabird Watch* volunteers must have clicked an area for it be counted as a Kittiwake. A 'marked_all' threshold of '> 1' means at least two *Seabird Watch* volunteers marked all birds in the image.

Camera

Accuracy also varied markedly by site/camera. For num_markings > 2, the average difference between GS and CS classifications for adults and chicks combined was greatest at MITTa (17.09, σ = 43.49), followed by OSSIa (14.47, σ = 43.67), HVITa (5.47, σ = 10.05) and ALKEa (0.76 σ = 11.61), with the lowest average difference at SKELa (-0.21, σ = 6.01). This order changed when adults and chicks were considered separately. The average difference in the number of adults was highest at MITTa (14.3, σ = 40.32), followed by OSSIa (12.97, σ = 36.55), HVITa (-1.94, σ = 6.91), SKELa (-1.63, σ = 6.11) and then ALKEa (0.12 +/- 11.64); while the average difference in the number of chicks was highest at HVITa (7.82, σ = 5.87) followed by MITTa (7.15, σ = 10.50), SKELa (6.73, σ = 8.08), OSSIa (2.07, σ = 11.24), and ALKEa (2.00, σ = 1.63) (Table 3, Table S1-3). Having said this, average difference does not consider the number of birds per image. Larger average differences would be expected for cameras viewing a greater number of Kittiwakes. The highest number of Kittiwakes recorded in a single image from GS classification was at MITTa (229), followed by OSSIa (198), SKELa (88), HVITa (62) and finally ALKEa (60) (Table 3). Table 3 compares the average absolute and percentage differences between GS and CS counts for each site.

As well, accuracy could have been affected by the proportion of images where a high number of *Seabird Watch* volunteers (≥ 10) marked all birds. HVITa and SKELa had the highest proportion of images where ≥ 10 *Seabird Watch* users marked all birds (71% and 65% respectively), whereas the highest proportion of images (~50%) were marked by 7, 8 or 9 users at OSSIa, ALKEa and MITTa (Table S4, Fig. 3).

Table 3 Average absolute and percentage differences between gold standard (GS) and citizen science (CS) counts of adults and chicks combined, adults only and chicks only for num_markings > 2. Average difference is the mean GS minus CS count and σ_a is the standard deviation. Average percentage difference is the mean GS minus CS count divided by the GS count and multiplied by 100, and σ_p is the standard deviation. Maximum number is the maximum number of adults, chicks, and adults and chicks combined recorded from GS classification for each camera.

Camera	**Average difference**	$\boldsymbol{\sigma_a}$	**Average % difference**	$\boldsymbol{\sigma_p}$	**Maximum number**
HVITa					
Adults + Chicks	5.47	10.05	21.80	36.52	62
Adults	-1.95	6.91	-25.38	86.06	36
Chicks	7.82	5.87	50.80	34.24	31
OSSIa					
Adults + Chicks	14.48	43.67	10.48	30.64	198
Adults	12.87	36.55	10.29	29.81	153
Chicks	2.07	11.24	-0.26	70.05	55
ALKEa					
Adults + Chicks	0.76	11.61	3.35	35.56	60
Adults	0.12	11.64	1.74	35.78	57
Chicks	2.00	1.63	70.77	42.71	10
MITTa					
Adults + Chicks	17.09	43.49	14.43	34.77	229
Adults	14.30	40.32	13.23	35.06	209
Chicks	7.15	10.50	46.92	44.89	57
SKELa					
Adults + Chicks	-0.21	6.01	6.06	33.96	88
Adults	-1.63	6.11	3.76	34.79	70
Chicks	6.73	8.08	79.50	20.68	26

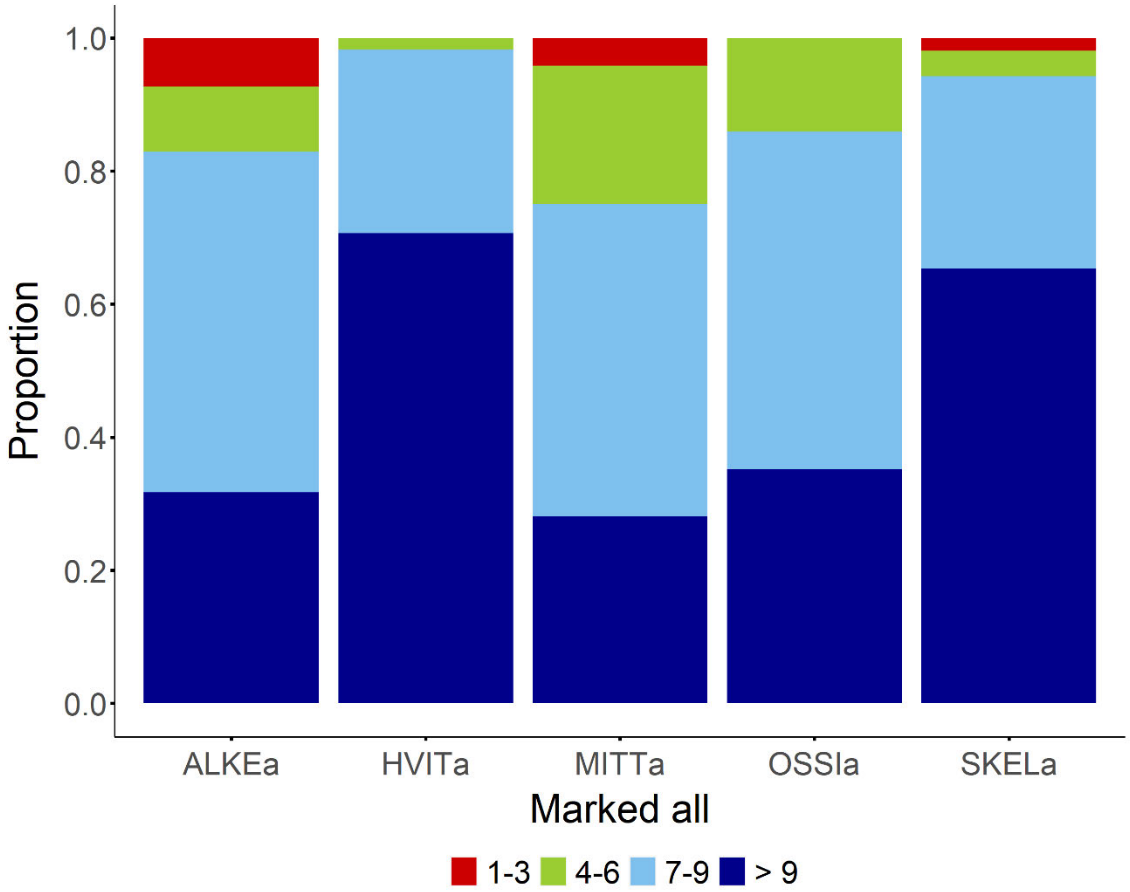


Figure 5 Proportion of images for each camera where 1-3, 4-6, 7-9 and ≥10 users marked all birds, in images where at least one Kittiwake was seen in gold standard (GS) classifications.

Adults and chicks combined or separate

For each site, accuracy differed depending on whether adults and chicks were considered together or separately. Average difference between GS and CS counts was lowest at MITTa (7.15, $\sigma = 10.50$) and OSSIa (2.07, $\sigma = 11.24$) when considering chicks only, at ALKEa (0.12, $\sigma = 11.64$) when considering adults only, and at HVITa (5.57, $\sigma = 10.05$) and SKELa (-0.21, $\sigma = 6.01$) for adults and chicks combined. This changed at MITTa and SKELa when using average percentage difference, which was lowest for counts of adults (Table 3). The proportion of GS counts that equalled CS counts for adults only was typically either the same or higher than for chicks (Fig. 2).

DISCUSSION

Counts of adult and juvenile Kittiwakes in time-lapse images made by citizen science (CS) and gold standard (GS) analysers were compared across five sites. Percentage difference between GS and CS classification was very low (< 5%) in images where GS classification recorded zero Kittiwakes. Percentage difference was generally higher for images containing Kittiwakes, especially when comparing counts of chicks only. Using num_markings > 2 gave the lowest percentage difference for all sites, while a single optimal marked_all threshold was not found. Overall, accuracy of volunteer counts was dependent on several factors, including threshold values, whether adults and chicks were considered together or separately, and site/camera, which was related to camera setup, the number of birds in the field of view and species present.

Images without Kittiwakes

For images where GS classification recorded zero Kittiwakes, percentage agreement between GS and CS classifications was generally high (> 95%) and was only lower at SKELa (88 %) for which sample size was small (n = 8 images). This suggests that *Seabird Watch* volunteers and GS classification perform almost equally well when the image contains zero Kittiwakes. In a few cases volunteer consensus clicks erroneously marked one or more Kittiwakes, and GS classification occasionally missed a bird. At MITTa and ALKEa, GS classification appeared to miss greater than 25 Kittiwakes in several images. While any user may accidentally miss some birds, it seems improbable that gold standard analysis would fail to notice so many individuals. It is more likely that the image was skipped by the gold standard analyser who intended to classify the image later, however the image was incorrectly retired and recorded with zero Kittiwakes instead.

Images with Kittiwakes

For images containing Kittiwakes, accuracy varied depending on threshold level, site/camera and whether adults and chicks were considered together or separately.

Threshold values

Exploring the effect of num_markings and marked_all thresholds on the difference between GS and CS classification was important to try and find an optimal threshold that minimises those differences. For all cameras, the average difference was lowest with num_markings > 2, meaning consensus clicks drawn from a cluster with fewer than three classifications were discarded. At this threshold the number of false positives (where a CS consensus click incorrectly marked an area as a Kittiwake) and false negatives (where an area containing a Kittiwake was not marked by a consensus click) was at a minimum. As the threshold level was increased further, the proportion of overestimates by CS classification (GS < CS) decreased, as more marks were required per consensus click for it to be counted (Table S1-3). This helped filter out erroneous marks by volunteers, which were likely made by accident. Increasing the num_markings threshold also increased the proportion of underestimates (GS > CS) as more consensus clicks were excluded from the Kittiwake count, despite their x,y coordinates pertaining to the location of a Kittiwake (i.e. a false negative) (Table S1-3). Overall, the average difference between GS and CS counts tended to increase and become more positive as num_markings increased, due to GS classification recording increasingly more Kittiwakes than CS classification (Fig 3-4, Table S1-3).

The increasing average count difference with num_markings could suggest that consensus clicks comprising few raw clicks are not necessarily erroneous (i.e. false positives), and in fact excluding them from the results means numerous Kittiwakes are incorrectly ‘lost’ from the CS count. Alternatively, the average difference might generally increase as num_markings is increased, because volunteers are not required to mark all birds in the image. As num_markings is raised, the number of volunteers that must have marked the image increases, and so the probability that any one observer will miss any one bird also increases. Nevertheless, Fig 3-4 show that the average difference between GS and CS counts does not consistently increase with num_markings for every camera. At HVITa, the average difference reaches a maximum at num_markings > 9 and num_markings > 6 when comparing adult only and chick only counts respectively, and then decreases beyond these values. It is therefore not necessarily a mathematical certainty that accuracy decreases as num_markings is increased. Importantly, this analysis has established that regardless of the trend with increasing num_markings, num_markings > 2 is optimal for all cameras, and reflects the best balance between over- and underestimation.

Using marked_all to filter out incorrect classifications appeared to increase accuracy more than num_markings. The optimal marked_all threshold gave either equal or higher accuracy than the optimal num_markings threshold (Fig. 3-4). For example, discarding images where fewer than eight people marked all birds (marked_all > 7) gave an average difference of -1.13 ($\sigma = 27.50$) Kittiwakes at MITTa, compared to 17.09 ($\sigma = 43.49$) Kittiwakes when at least three clicks contributed to each consensus click (num_markings > 2). As the marked_all threshold increased, the proportion of overestimates (GS < CS) increased and underestimates (GS > CS) decreased, whereas when increasing num_markings the opposite was true (Table S1-3). This suggests that CS classification is more likely to miss birds and therefore underestimate the number of Kittiwakes when fewer users mark all birds in the image. Altering the *Seabird Watch* tutorial to encourage users to mark all birds could reduce underestimation, however some users may just skip images containing too many birds to avoid having to annotate them all.

Of course, it should be noted that marked_all and num_markings thresholds have different effects on the data. The num_markings threshold prevents consensus clicks being included if they contain less than a certain number of raw clicks. This means photographs will only be removed from the dataset if all consensus clicks in the image are discarded, and so num_markings primarily affects the image bird count. Conversely, the marked_all threshold discards entire images if less than a certain number of people marked all birds in the photograph. This means the image bird count does not change, but the number of images in the sample is reduced. Determining an optimal num_markings and marked_all threshold that will maximise accuracy while minimising the number of images discarded, thus maximising sample size, could be an important step for using *Seabird Watch* data in future analysis. While num_markings > 2 was consistently best across all cameras used in this study, the optimal marked_all threshold varied by camera and was perhaps more dependent on the proportion of images that fell into different marked_all categories (e.g. 1-3, 4-6, 7-9 and ≥ 10 users marked all birds). The optimal num_markings and marked_all threshold combination may therefore need to be established on a camera by camera basis and should be the initial focus of future work using the *Seabird Watch* dataset. For the purpose of this study, investigation into the effect of other factors on CS count accuracy was conducted using num_markings > 2 and all marked_all values (i.e. marked_all > 1).

Clustering algorithm

At num_markings > 2 the average difference between GS and CS counts was still high for some cameras and was often associated with large standard deviations. When comparing the number of adults and chicks combined between GS and CS classification at MITTa, the lowest average difference was 17.09 ($\sigma = 43.49$) Kittiwakes and the lowest average percentage difference was 14.43% ($\sigma = 34.77\%$). Such a large standard deviation shows how variable the GS and CS counts were, and that CS classification both under- and overestimated the number of Kittiwakes. Plotting the GS and CS consensus clicks onto images with a large difference between GS and CS classification, showed that there were frequently two CS consensus clicks per Kittiwake (Fig. 6). Both CS and GS classification missed some birds, but the greatest difference in counts appeared to be due to the clustering algorithm generating multiple consensus clicks per bird from CS data. While *Seabird Watch* users are asked to click on the centre of birds, it seems that this is not always followed, with some users consistently clicking on the head and others on the tail. If the head and tail of the same Kittiwake are far enough apart, then the clustering algorithm appears to group the head and tail marks into separate clusters, therefore generating two consensus clicks for one bird. This issue might be exacerbated at cameras positioned close to the colony and for birds in the foreground, as these birds will appear larger in the frame and so have a greater difference between their head and tail coordinates. Small birds in the background will have similar coordinates for marks on the head and tail, so are more likely to be aggregated into the same consensus click.

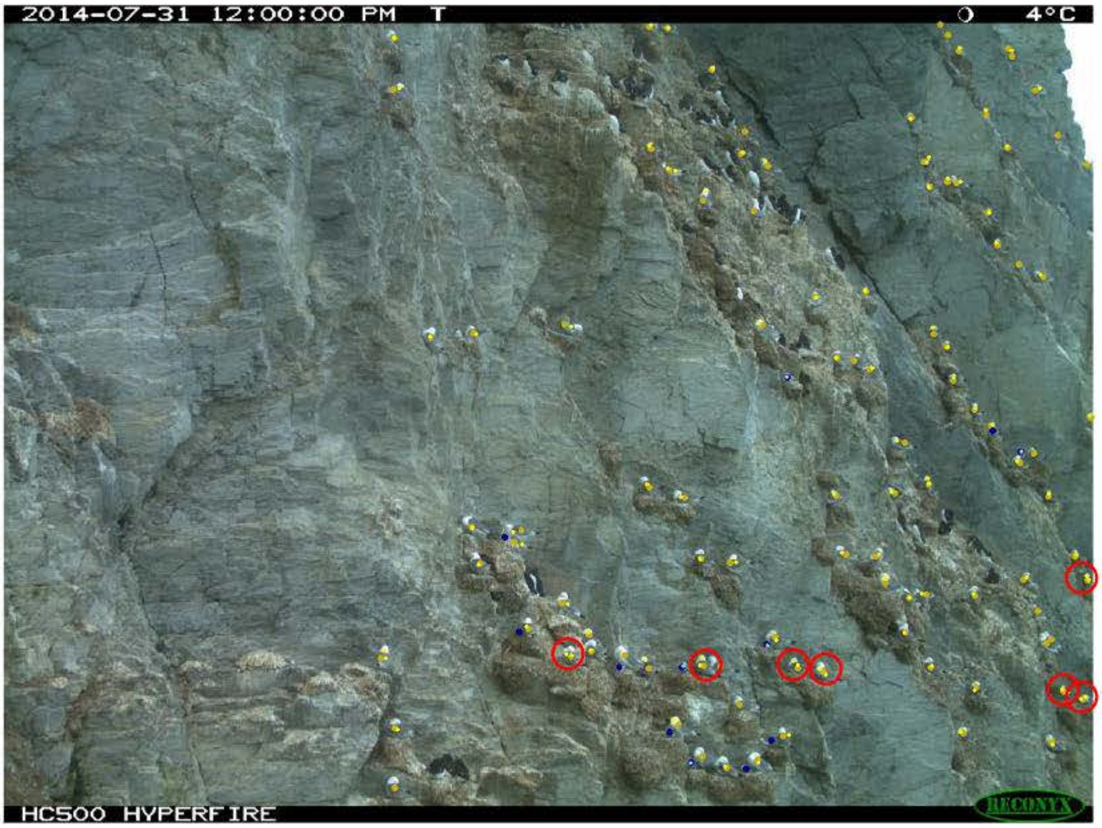


Figure 6 Gold standard and citizen science classification of the 12.00 image on 31/07/2014 at MITTa. The following colour codes were used: gold standard adult = gold, citizen science adult = yellow, gold standard chick = dark blue, citizen science chick = light blue. Citizen science marks are consensus clicks rather than raw clicks. Red circles show examples of where two consensus clicks were marked per bird.

Adjusting the clustering algorithm to generate one consensus click per bird from CS raw clicks is of priority for the development of *Seabird Watch*, however it could prove challenging. Like *Seabird Watch,* validation of *Penguin Watch* data showed a decrease in overestimates and increase in underestimates as num_markings increased, although it did not highlight an issue with the clustering algorithm. Accuracy was generally much higher for *Penguin Watch,* and the proportion of differences that were zero or one ranged from 0.15 to 0.94 for images containing penguins, compared to 0.00 to 0.40 for *Seabird Watch* images containing Kittiwakes (Jones *et al.* 2018).

One reason for such differing results between *Seabird Watch* and *Penguin Watch* data is likely the nature of the photographs. Most *Penguin Watch* images contain fewer nests per photo and so the penguins are larger in the frame and further apart. Penguins also typically nest further apart than Kittiwakes. As a result, neighbouring Kittiwakes appear to be touching in some images, especially those from MITTa and OSSIa. Raw clicks intended for differing individuals may therefore merge, leading to extra consensus clicks being generated at the boundaries between birds, as well as near the centre of each individual. It will be interesting to see how the algorithm fares when comparing GS and CS classification of Guillemots, which nest even closer (pers. obs.). Such little spatial separation could increase the likelihood of underestimation (GS > CS), if users struggle to distinguish between individuals and/or raw clicks intended for different individuals are grouped in the same consensus click. Discussing the operation of the clustering algorithm with the *Zooniverse* team is recommended to ascertain how it could be improved for use with *Seabird Watch* data.

Camera

The accuracy of results varied greatly depending on the site/camera being examined. When using the optimal num_markings threshold (> 2), the largest average difference between GS and CS counts was associated with MITTa (17.09, $\sigma = 43.49$ Kittiwakes), followed by OSSIa, HVITa, ALKEa and SKELa. However, average count difference does not account for the number of birds present in images, which was expected to influence accuracy. In fact, the greatest average percentage difference between GS and CS counts was associated with HVITa (21.80%, $\sigma = 36.52\%$) followed by MITTa (14.43%, $\sigma = 34.77\%$), OSSIa (10.48, $\sigma = 30.64\%$), SKELa (6.06%, $\sigma = 33.96\%$) and ALKEa (3.35%, $\sigma = 35.56\%$). This broadly supports the assumption that accuracy decreases with colony size, as the maximum number of Kittiwakes recorded from GS classification at MITTa was 229, compared to 198, 88, 62 and 60 at OSSIa, SKELa, HVITa and ALKEa respectively.

The high percentage difference at HVITa is unusual given the colony size is small, and the camera is positioned close to the colony providing high image resolution. Further examination of HVITa 13.00 images shows that the sun casts dark shadows over the site on some days, which may make it harder to spot birds. Equally, bright sunlight sometimes reflects off the white guano on the cliffs. This could make it difficult to see the Kittiwakes, which have a white body and pale grey back in adult plumage. It is therefore evident that both the number of Kittiwakes viewed, and other site-specific factors affect volunteer counts. Additionally, some cameras captured

Guillemots as well as Kittiwakes (MITTa, OSSIa, SKELa), and so the maximum count of all species, not just Kittiwakes, could influence accuracy.

Examination of *Penguin Watch* data suggested that citizen scientists find it harder to distinguish individuals when presented with larger groups (Jones *et al.* 2018). Furthermore, users might begin to lose patience and become more careless as they mark an increasing number of birds in an image, which could lead to lower accuracy. This could also explain why the average difference between GS and CS counts is larger for *Seabird Watch* than *Penguin Watch,* as *Seabird Watch* images typically contain more birds than images used on *Penguin Watch*. Cropping images with many birds to focus on those in the foreground could help increase *Seabird Watch* accuracy, both in terms of number of Kittiwakes counted (adults and chicks combined) and bird identity (whether the Kittiwake is an adult or chick). Users may struggle to spot chicks and distinguish between adult and juvenile Kittiwakes at the back of images, especially at sites like MITTa where a large number of nests are present towards the rear. Segmentation of images does not, however, prevent partial or complete obscuration of individuals by conspecifics, which increases the probability of some birds being missed (Jones *et al.* 2018). Again, this problem may be greater for *Seabird Watch* than *Penguin Watch,* as Kittiwakes and Guillemots nest closer together than penguins, meaning individuals are more likely to be obscured by their neighbours.

As well, the number of volunteers that marked all birds in the image might contribute to variable accuracy between sites. Fewer birds were expected to be missed when more users marked all birds in the image. Cameras where a high proportion of images had all birds marked by many users (i.e. 7 - 9 or ≥ 10) should have fewer underestimates (GS > CS) than if few users marked all birds (i.e. 1 - 3 or 4 - 6). This appeared to hold true when comparing adults only (num_markings > 2), as HVITa had the greatest proportion of images where 7 - 9 and ≥ 10 users marked all birds (99 %), followed by SKELa (94 %), OSSIa (86 %), ALKEa (83 %), then MITTa (75 %); and the proportion of underestimates (GS > CS) was lowest at HVITa (31 %), followed by SKELa (35 %), ALKEa (37 %), OSSIa (49 %), then MITTa (55 %). However, this pattern did not apply when comparing chicks only and adults and chicks combined. This could be because at each camera ≥ 75% of images had at least seven users mark all birds, and so there were not many images where few users marked all birds. Any expected trend would therefore be less apparent.

Adults and chicks separate or combined

Using num_markings > 2, chick counts appeared to be less accurate (i.e. higher average difference) than adult counts at HVITa (average difference 7.82, σ = 5.87 chicks; -1.95, σ = 6.91 adults), ALKEa (2.00, σ = 1.63 chicks; 0.12, σ = 11.64 adults) and SKELa (6.73, σ = 8.08 chicks; -1.63, σ = 6.11 adults), but more accurate at MITTa (7.15, σ = 10.50 chicks; 14.30, σ = 40.32 adults) and OSSIa (2.07, σ = 11.24 chicks; 12.87, σ = 36.55 adults) (Table S2-3). Average percentage difference was lower for chick counts than adult counts at OSSIa only. Despite this, the proportion of images where GS and CS counts were equal (GS = CS) for chicks only was equal to or higher than adults only at ALKEa (8% chicks, 7% adults), HVITa (4% chicks, 3% adults), MITTa (5% chicks, 5% adults) and OSSIa (7% chicks, 1% adults), but was much lower at SKELa (0% chicks, 13% adults).

Lower accuracy of chick estimation at HVITa, ALKEa and SKELa could be due to incorrect identification. Chicks may be labelled as adults or vice versa, especially when juvenile Kittiwakes are near fledging and look similar to adults. This is perhaps more likely to have occurred at SKELa, where the camera is positioned considerably further away from the colony compared to the other sites, and so the birds appear smaller and are harder to identify. However, if incorrect identification was the main issue, then accuracy should have been highest when adult and chick counts were combined (Jones *et al.* 2018), which was not the case for HVITa and ALKEa. Accuracy was highest when comparing adult counts only at these two sites. The proportion of overestimates (GS < CS) was greater than the proportion of underestimates (GS > CS) for adults at num_markings > 2, but the proportion of underestimates was greater than overestimates at all threshold levels when comparing chick counts (Table S1-3). This suggests that chicks are often missed by volunteers at HVITa and ALKEa (Jones *et al.* 2018). Chicks may be partially concealed by a parent, particularly during the first two weeks post-hatching when they are continuously brooded (Gabrielsen *et al.* 1992) and might be more likely to be missed at HVITa and ALKEa than other sites due to camera setup. For example, at ALKEa the camera is angled upwards making it harder to observe nest contents. Cameras are best positioned slightly above the colony and facing down (Lorentzen *et al* .2010). Including more images of young chicks and information on expected lay dates on the *Seabird Watch* interface could improve volunteers' abilities to notice chicks (Jones *et al.* 2018).

Validation of *Penguin Watch* data also found that chicks were frequently missed during classification, which could in part stem from the project workflow. Users are not required to mark all birds in each image in order to maintain interest, which does not detrimentally affect the accuracy of adult counts (Jones *et al.* 2018). However, if most volunteers choose to mark adults first, perhaps because they are easier to identify, then they might consistently move onto the next image before marking any chicks. As suggested for *Penguin Watch*, adjusting the *Seabird Watch* tutorial to request that volunteers classify a range of seabird 'types' before moving onto the next image, could help minimise chick underestimation (Jones *et al.* 2018).

Adult underestimation was also a problem at some *Seabird Watch* sites, namely MITTa and OSSIa. The proportion of underestimates (GS > CS) was greater than or equal to the proportion of overestimates (GS < CS) for counts of adults only and chicks only at all threshold levels of num_markings, suggesting adult and chick counts were underestimated when considered separately. Counts of adults and chicks combined were likewise underestimated for all threshold levels at MITTa, but were overestimated with num_markings > 2 at OSSIa. This suggests that all birds, regardless of 'type' were underestimated by volunteers at MITTa, which recorded the highest maximum number of birds in a single image (209). Users could become fatigued or start to lose interest after marking a large number of birds in one image, thus increasing the likelihood of missing individuals, even if they thought they had marked all. Equally, some of the marks themselves could obscure other birds, especially in colonies with many individuals breeding close together. As previously suggested, asking users to only mark birds in the foreground or cropping/sub-sampling images with many birds could provide one option for minimising underestimation. Identifying these potential sources of error and developing solutions to overcome them, is an essential process in all citizen science projects to ensure integrity in non-expert derived data.

CONCLUSION

Comparison of the *Seabird Watch* dataset with gold standard classification has highlighted sources of inaccuracy that will need to be resolved going forwards. The lowest average percentage differences between GS and CS Kittiwake counts (achieved using num_markings > 2) ranged from 3.35% to 21.80% across the five sites examined, but could exceed 70% when comparing counts of chicks only. Count accuracy is lower for *Seabird Watch* than its sister project, *Penguin Watch,* most likely due to the nature of the images and study species. Understanding how these differences could be affecting the aggregation of raw volunteer marks into consensus clicks may be important for improving count accuracy, without needing to change the *Seabird Watch* interface. Nonetheless, changes to the *Seabird Watch* tutorial and field guide might still be necessary to resolve other problems, like chick underestimation. Comparing citizen science and gold standard datasets from time-lapse cameras is challenging, given the range of factors that can affect results. This study alone has highlighted several variables affecting count accuracy, including: camera setup and study site, species, 'type' of animal (adult or chick), number of animals per image, proportion of users marking all animals in an image and aggregation threshold. It is therefore essential that validation of citizen science datasets occurs on a case-by-case, or for this study, camera-by-camera basis.

SUPPLEMENTARY MATERIAL

Table S1 Comparison between gold standard (GS) and citizen science (CS) counts of **adult and chicks (combined)** for five cameras at different num_markings and marked_all thresholds. num_markings threshold of '> 5' means at least six people must have clicked an area for it be counted as a Kittiwake, while marked_all threshold of ' > 5' means at least six people marked all birds in the image. 'D_a' refers to 'average difference' and is the mean GS minus CS count, and σ_a is the standard deviation. 'D_p' refers to 'average percentage difference' and is the mean GS minus CS count divided by the GS count and multiplied by 100, and σ_p is the standard deviation. 'PD 0/1' is the 'Proportion of differences 0 or 1'; the proportion of images for which the CS count was equal to the GS or different by one individual. The proportion of gold standard (GS) classifications that are greater than (GS > CS), equal to (GS = CS) and less than (GS < CS) citizen science (CS) classifications for the threshold value with the lowest average difference is also given. 'n_a' is the number of images in the sample for average difference and proportion calculations and 'n_p' is the number of images used for percentage difference calculations. The results presented are for images where Kittiwakes (adults and/or chicks) were present according to gold standard and/or citizen science classifications.

Threshold	**D_a**	**σ_a**	**PD 0/1**	**GS < CS**	**GS = CS**	**GS > CS**	**n_a**	**D_p**	**σ_p**	**n_p**
HVITa										
num_markings										
>2	5.47	10.05	0.16	0.29	0.03	0.67	58	21.80	36.52	58
>3	16.42	12.65	0.12	0.00	0.04	0.96	57	47.13	28.86	57
>4	23.86	14.63	0.05	0.00	0.02	0.98	57	65.25	24.85	57
>5	29.00	15.39	0.04	0.00	0.00	1.00	56	77.62	19.23	56
>6	32.12	15.48	0.04	0.00	0.00	1.00	52	85.07	14.01	52
>7	32.36	16.00	0.05	0.00	0.00	1.00	42	90.06	9.77	42
>8	30.90	17.21	0.07	0.00	0.00	1.00	29	92.96	7.07	29
>9	29.71	19.38	0.10	0.00	0.00	1.00	21	95.25	4.85	21
>10	24.50	21.65	0.14	0.00	0.00	1.00	14	97.30	3.04	14
marked_all										
>1	5.47	10.05	0.16	0.29	0.03	0.67	58	21.80	36.52	58
>2	5.47	10.05	0.16	0.29	0.03	0.67	58	21.80	36.52	58
>3	5.47	10.05	0.16	0.29	0.03	0.67	58	21.80	36.52	58
>4	5.47	10.05	0.16	0.29	0.03	0.67	58	21.80	36.52	58
>5	5.07	9.67	0.16	0.30	0.04	0.67	57	21.20	36.55	57
>6	5.07	9.67	0.16	0.30	0.04	0.67	57	21.20	36.55	57
>7	4.53	9.24	0.16	0.31	0.04	0.65	55	19.42	35.73	55
>8	3.94	8.46	0.16	0.31	0.02	0.67	51	18.79	35.84	51
>9	2.91	7.50	0.15	0.35	0.02	0.63	46	17.84	37.01	46
>10	-0.17	5.50	0.24	0.48	0.03	0.48	29	17.49	45.22	29
OSSIa										
num_markings										
>2	14.48	43.67	0.03	0.51	0.00	0.49	71	10.48	30.64	71
>3	51.82	44.99	0.06	0.10	0.00	0.90	71	38.75	30.25	71
>4	74.39	45.61	0.01	0.00	0.01	0.99	70	55.77	29.07	70
>5	86.79	43.31	0.00	0.00	0.00	1.00	62	65.16	25.58	62
>6	99.91	40.51	0.00	0.00	0.00	1.00	56	74.88	22.00	56
>7	110.55	40.61	0.00	0.00	0.00	1.00	49	82.88	18.49	49
>8	118.32	37.55	0.00	0.00	0.00	1.00	38	88.84	14.73	38
>9	120.04	38.56	0.00	0.00	0.00	1.00	27	92.38	13.25	27
>10	120.92	36.88	0.00	0.00	0.00	1.00	13	93.28	10.61	13
marked_all										
>1	14.48	43.67	0.03	0.51	0.00	0.49	71	10.48	30.64	71
>2	14.48	43.67	0.03	0.51	0.00	0.49	71	10.48	30.64	71
>3	14.48	43.67	0.03	0.51	0.00	0.49	71	10.48	30.64	71
>4	10.94	40.78	0.03	0.53	0.00	0.47	68	7.87	28.31	68
>5	6.48	38.61	0.03	0.57	0.00	0.43	63	4.22	25.97	63
>6	2.97	33.53	0.03	0.59	0.00	0.41	61	2.08	23.46	61
>7	-4.00	25.06	0.04	0.67	0.00	0.33	54	-2.14	20.16	54
>8	-9.05	21.12	0.05	0.73	0.00	0.27	44	-5.91	17.32	44
>9	-15.04	21.64	0.00	0.80	0.00	0.20	25	-10.17	16.46	25
>10	-22.38	20.24	0.00	0.92	0.00	0.08	13	-17.01	14.85	13

ALKEa										
num_markings										
>2	0.76	11.61	0.24	0.56	0.07	0.37	41	3.35	35.56	41
>3	8.10	9.17	0.10	0.10	0.05	0.85	39	22.55	28.68	39
>4	13.59	8.88	0.03	0.03	0.00	0.97	39	37.30	26.23	39
>5	19.42	8.75	0.00	0.00	0.00	1.00	38	52.86	22.93	38
>6	25.56	9.64	0.00	0.00	0.00	1.00	36	66.06	19.79	36
>7	30.48	10.84	0.00	0.00	0.00	1.00	33	77.48	16.34	33
>8	33.76	11.11	0.00	0.00	0.00	1.00	29	86.22	11.54	29
>9	36.50	11.94	0.00	0.00	0.00	1.00	20	92.48	7.67	20
>10	38.60	14.82	0.00	0.00	0.00	1.00	10	95.37	4.38	10
marked_all										
>1	0.76	11.61	0.24	0.56	0.07	0.37	41	3.35	35.56	41
>2	0.76	11.61	0.24	0.56	0.07	0.37	41	3.35	35.56	41
>3	-0.18	10.09	0.25	0.58	0.08	0.35	40	1.22	33.27	40
>4	-0.28	10.20	0.26	0.59	0.08	0.33	39	0.92	33.65	39
>5	-0.76	9.88	0.26	0.61	0.08	0.32	38	-0.24	33.31	38
>6	-1.11	9.92	0.25	0.61	0.08	0.31	36	-0.99	33.87	36
>7	-1.38	10.16	0.24	0.69	0.07	0.24	29	-2.51	34.52	29
>8	-0.38	10.70	0.25	0.63	0.08	0.29	24	0.85	36.00	24
>9	0.60	13.36	0.20	0.60	0.07	0.33	15	3.81	45.50	15
>10	5.00	15.42	0.25	0.63	0.00	0.38	8	18.19	51.47	8
MITTa										
num_markings										
>2	17.09	43.49	0.07	0.39	0.04	0.57	192	14.43	34.77	192
>3	45.01	43.46	0.07	0.06	0.03	0.91	190	39.92	32.73	190
>4	61.23	45.86	0.04	0.01	0.03	0.97	174	53.53	29.38	174
>5	75.81	48.12	0.02	0.01	0.01	0.98	161	64.84	27.06	161
>6	87.20	48.55	0.02	0.01	0.01	0.99	146	74.90	23.31	146
>7	95.10	46.22	0.02	0.01	0.00	0.99	118	81.57	19.05	118
>8	101.60	46.46	0.02	0.00	0.00	1.00	95	87.20	15.33	95
>9	108.67	51.56	0.01	0.00	0.00	1.00	67	91.34	11.55	67
>10	118.50	53.24	0.03	0.00	0.00	1.00	36	94.10	7.45%	36
marked_all										
>1	17.09	43.49	0.07	0.39	0.04	0.57	192	14.43	34.77	192
>2	15.01	39.93	0.07	0.39	0.04	0.57	189	13.16	33.53	189
>3	12.86	37.95	0.08	0.40	0.04	0.55	184	11.63	32.45	184
>4	10.42	36.32	0.08	0.42	0.05	0.53	176	9.74	31.57	176
>5	5.01	29.57	0.09	0.46	0.05	0.49	161	5.91	28.98	161
>6	2.94	29.48	0.09	0.49	0.05	0.46	145	3.27	28.27	145
>7	-1.13	27.50	0.10	0.54	0.05	0.41	124	0.01	27.85	124
>8	-5.70	22.62	0.11	0.62	0.04	0.34	91	-4.45	26.06	91
>9	-11.71	18.70	0.05	0.76	0.02	0.22	55	-9.42	28.33	55
>10	-17.81	19.01	0.06	0.87	0.03	0.10	31	-15.45	32.97	31
SKELa										
num_markings										
>2	-0.21	6.01	0.40	0.44	0.15	0.40	52	6.06	33.96	52
>3	5.61	8.96	0.33	0.12	0.16	0.73	51	24.68	28.93	51
>4	10.02	12.40	0.16	0.02	0.08	0.90	50	34.39	28.87	50
>5	14.27	15.42	0.12	0.02	0.06	0.92	49	45.22	30.26	49
>6	18.91	17.90	0.09	0.00	0.06	0.94	47	56.30	30.10	47
>7	24.38	20.55	0.08	0.00	0.05	0.95	40	63.38	29.44	40

>8	30.43	22.90	0.06	0.00	0.00	1.00	35	76.51	21.84	35
>9	40.09	23.41	0.04	0.00	0.00	1.00	23	85.03	17.35	23
>10	45.06	25.61	0.06	0.00	0.00	1.00	18	91.08	12.94	18
marked_all										
>1	-0.21	6.01	0.40	0.44	0.15	0.40	52	6.06	33.96	52
>2	-0.21	6.01	0.40	0.44	0.15	0.40	52	6.06	33.96	52
>3	-0.21	6.01	0.40	0.44	0.15	0.40	52	6.06	33.96	52
>4	-0.21	6.01	0.40	0.44	0.15	0.40	52	6.06	33.96	52
>5	-0.21	6.01	0.40	0.44	0.15	0.40	52	6.06	33.96	52
>6	-0.92	4.93	0.42	0.46	0.16	0.38	50	4.91	34.06	50
>7	-1.02	4.87	0.44	0.46	0.17	0.38	48	5.02	34.74	48
>8	-1.00	4.79	0.44	0.46	0.15	0.39	41	6.44	37.21	41
>9	-1.63	4.55	0.43	0.51	0.11	0.37	35	6.08	40.04	35
>10	-2.19	5.08	0.38	0.52	0.14	0.33	21	12.15	44.01	21

Table S2 Comparison between gold standard (GS) and citizen science (CS) counts of **adults (only)** for five cameras at different num_markings and marked_all thresholds. num_markings threshold of '> 5' means at least six people must have clicked an area for it be counted as a Kittiwake, while marked_all threshold of '> 5' means at least six people marked all birds in the image. 'D_a' refers to 'average difference' and is the mean GS minus CS count, and σ_a is the standard deviation. 'D_p' refers to 'average percentage difference' and is the mean GS minus CS count divided by the GS count and multiplied by 100, and σ_p is the standard deviation. 'PD 0/1' is the 'Proportion of differences 0 or 1'; the proportion of images for which the CS count was equal to the GS or different by one individual. The proportion of gold standard (GS) classifications that are greater than (GS > CS), equal to (GS = CS) and less than (GS < CS) citizen science (CS) classifications for the threshold value with the lowest average difference is also given. 'n_a' is the number of images in the sample for average difference and proportion calculations and 'n_p' is the number of images used for percentage difference calculations. The results presented are for images where adults were present according to gold standard and/or citizen science classifications.

Threshold	D_a	σ_a	PD 0/1	GS < CS	GS = CS	GS > CS	n_a	D_p	σ_p	n_p
HVITa										
num_markings										
>2	-1.95	6.91	0.22	0.66	0.03	0.31	58	25.38	86.06	55
>3	5.12	8.54	0.25	0.19	0.11	0.70	57	17.13	60.67	54
>4	9.84	9.74	0.21	0.07	0.11	0.82	57	46.14	38.46	54
>5	12.96	10.49	0.20	0.00	0.05	0.95	56	63.39	28.63	54
>6	14.52	10.44	0.15	0.00	0.04	0.96	52	71.63	24.41	50
>7	15.36	10.70	0.12	0.00	0.05	0.95	42	79.52	18.85	40
>8	14.76	9.97	0.14	0.00	0.07	0.93	29	86.49	11.24	27
>9	15.90	10.98	0.19	0.00	0.10	0.90	21	92.18	07.08	19
>10	14.79	13.50	0.29	0.00	0.14	0.86	14	95.66	04.51	12
marked_all										
>1	-1.95	6.91	0.22	0.66	0.03	0.31	58	-25.38	86.06	55
>2	-1.95	6.91	0.22	0.66	0.03	0.31	58	-25.38	86.06	55
>3	-1.95	6.91	0.22	0.66	0.03	0.31	58	-25.38	86.06	55
>4	-1.95	6.91	0.22	0.66	0.03	0.31	58	-25.38	86.06	55
>5	-2.18	6.75	0.23	0.67	0.04	0.30	57	-26.47	86.48	54
>6	-2.18	6.75	0.23	0.67	0.04	0.30	57	-26.47	86.48	54
>7	-2.53	6.55	0.24	0.69	0.04	0.27	55	-29.70	86.45	52
>8	-2.78	6.39	0.24	0.71	0.04	0.25	51	-30.20	87.18	49
>9	-3.43	6.03	0.24	0.74	0.04	0.22	46	-33.32	90.31	44
>10	-4.83	5.71	0.28	0.79	0.07	0.14	29	-35.98	103.60	27
OSSIa										
num_markings										
>2	12.87	36.55	0.07	0.49	0.01	0.49	71	10.29	29.81	71
>3	45.34	38.39	0.04	0.07	0.00	0.93	71	38.25	29.60	71
>4	64.84	37.96	0.01	0.01	0.00	0.99	70	55.26	28.62	70
>5	75.87	35.77	0.00	0.00	0.00	1.00	62	64.68	25.53	62
>6	87.55	32.85	0.00	0.00	0.00	1.00	56	74.46	22.02	56
>7	96.96	33.00	0.00	0.00	0.00	1.00	49	82.40	18.66	49
>8	105.82	31.02	0.00	0.00	0.00	1.00	38	88.40	15.21	38
>9	107.81	32.16	0.00	0.00	0.00	1.00	27	92.03	13.92	27
>10	108.08	29.87	0.00	0.00	0.00	1.00	13	92.97	11.02	13
marked_all										
>1	12.87	36.55	0.07	0.49	0.01	0.49	71	10.29	29.81	71
>2	12.87	36.55	0.07	0.49	0.01	0.49	71	10.29	29.81	71
>3	12.87	36.55	0.07	0.49	0.01	0.49	71	10.29	29.81	71
>4	9.37	32.75	0.07	0.51	0.01	0.47	68	7.69	27.38	68
>5	5.19	29.90	0.08	0.56	0.02	0.43	63	4.10	25.02	63
>6	2.74	26.65	0.08	0.57	0.02	0.41	61	2.13	22.84	61
>7	-2.59	20.29	0.07	0.65	0.02	0.33	54	-1.71	19.95	54
>8	-6.93	16.76	0.07	0.73	0.00	0.27	44	-5.43	17.53	44

>9	-11.96	16.71	0.04	0.76	0.00	0.24	25	-9.80	17.73	25
>10	-16.92	16.94	0.00	0.92	0.00	0.08	13	-15.89	19.22	13
ALKEa										
num_markings										
>2	0.12	11.64	0.22	0.56	0.07	0.37	41	1.74	35.78	41
>3	7.28	9.09	0.13	0.13	0.08	0.79	39	21.12	28.82	39
>4	12.69	8.84	0.05	0.03	0.03	0.95	39	36.06	26.55	39
>5	18.47	8.92	0.00	0.00	0.00	1.00	38	51.77	23.49	38
>6	24.50	9.75	0.00	0.00	0.00	1.00	36	65.19	20.29	36
>7	29.30	10.80	0.00	0.00	0.00	1.00	33	76.86	16.80	33
>8	32.34	10.69	0.00	0.00	0.00	1.00	29	85.75	11.70	29
>9	34.80	11.02	0.00	0.00	0.00	1.00	20	92.26	07.73	20
>10	37.40	14.67	0.00	0.00	0.00	1.00	10	95.22	04.41	10
marked_all										
>1	0.12	11.64	0.22	0.56	0.07	0.37	41	1.74	35.78	41
>2	0.12	11.64	0.22	0.56	0.07	0.37	41	1.74	35.78	41
>3	-0.83	10.06	0.23	0.58	0.08	0.35	40	-0.43	33.40	40
>4	-0.95	10.16	0.23	0.59	0.08	0.33	39	-0.77	33.77	39
>5	-1.37	9.95	0.24	0.61	0.08	0.32	38	-1.86	33.52	38
>6	-1.61	10.13	0.22	0.61	0.08	0.31	36	-2.43	34.28	36
>7	-2.00	10.39	0.21	0.69	0.07	0.24	29	-4.30	34.92	29
>8	-1.04	10.98	0.21	0.63	0.08	0.29	24	-1.09	36.60	24
>9	0.00	13.63	0.13	0.60	0.07	0.33	15	1.84	46.20	15
>10	3.75	16.32	0.13	0.63	0.00	0.38	8	14.63	53.66	8
MITTa										
num_markings										
>2	14.30	40.32	0.07	0.40	0.05	0.55	192	13.23	35.06	192
>3	40.47	40.78	0.07	0.09	0.03	0.88	190	38.6	33.06	190
>4	55.39	41.88	0.04	0.01	0.03	0.97	174	52.40	29.74	174
>5	69.47	44.58	0.02	0.01	0.01	0.98	161	63.82	27.43	161
>6	80.35	45.42	0.02	0.01	0.01	0.99	146	74.04	23.77	146
>7	88.07	42.99	0.02	0.01	0.00	0.99	118	80.83	19.58	118
>8	94.48	42.95	0.02	0.00	0.00	1.00	95	86.62	15.73	95
>9	100.51	47.05	0.01	0.00	0.00	1.00	67	90.91	11.80	67
>10	108.61	46.69	0.03	0.00	0.00	1.00	36	93.82	7.48	36
marked_all										
>1	14.30	40.32	0.07	0.40	0.05	0.55	192	13.23	35.06	192
>2	12.48	37.73	0.07	0.41	0.05	0.54	189	11.95	33.80	189
>3	10.43	35.66	0.07	0.42	0.05	0.53	184	10.40	32.68	184
>4	8.24	34.10	0.07	0.44	0.05	0.51	176	8.49	31.75	176
>5	3.12	27.41	0.08	0.48	0.06	0.47	161	4.62	29.12	161
>6	0.79	26.98	0.08	0.52	0.06	0.43	145	1.76	28.35	145
>7	-2.79	25.50	0.09	0.56	0.06	0.38	124	-1.36	28.06	124
>8	-6.48	21.04	0.10	0.64	0.05	0.31	91	-5.57	26.37	91
>9	-12.02	17.26	0.05	0.80	0.04	0.16	55	-10.66	28.52	55
>10	-16.90	16.21	0.06	0.87	0.03	0.10	31	-16.12	32.99	31
SKELa										
num_markings										
>2	-1.63	6.11	0.38	0.52	0.13	0.35	52	3.76	34.79	52
>3	3.78	7.73	0.35	0.18	0.18	0.65	51	22.36	29.82	51

>4	8.02	10.15	0.16	0.02	0.08	0.90	50	32.38	29.24	50
>5	12.12	12.89	0.12	0.02	0.06	0.92	49	43.46	30.69	49
>6	16.57	14.75	0.09	0.00	0.06	0.94	47	54.82	30.37	47
>7	21.60	16.78	0.08	0.00	0.05	0.95	40	62.17	29.43	40
>8	27.26	18.39	0.06	0.00	0.00	1.00	35	75.75	21.89	35
>9	37.30	19.77	0.04	0.00	0.00	1.00	23	84.59	17.35	23
>10	41.50	20.84	0.06	0.00	0.00	1.00	18	90.96	12.88	18
marked_all										
>1	-1.63	6.11	0.38	0.52	0.13	0.35	52	3.76	34.79	52
>2	-1.63	6.11	0.38	0.52	0.13	0.35	52	3.76	34.79	52
>3	-1.63	6.11	0.38	0.52	0.13	0.35	52	3.76	34.79	52
>4	-1.63	6.11	0.38	0.52	0.13	0.35	52	3.76	34.79	52
>5	-1.63	6.11	0.38	0.52	0.13	0.35	52	3.76	34.79	52
>6	-2.10	5.52	0.40	0.54	0.14	0.32	50	2.91	34.97	50
>7	-1.79	5.30	0.42	0.52	0.15	0.33	48	3.67	35.45	48
>8	-1.93	5.26	0.44	0.54	0.12	0.34	41	4.78	38.05	41
>9	-2.71	4.97	0.43	0.60	0.09	0.31	35	4.13	40.93	35
>10	-3.10	5.40	0.38	0.62	0.10	0.29	21	10.41	45.16	21

Table S3 Comparison between gold standard (GS) and citizen science (CS) counts of **chicks (only)** for five cameras at different num_markings and marked_all thresholds. num_markings threshold of '> 5' means at least six people must have clicked an area for it be counted as a Kittiwake, while marked_all threshold of '> 5' means at least six people marked all birds in the image. 'D_a' refers to 'average difference' and is the mean GS minus CS count, and σ_a is the standard deviation. 'D_p' refers to 'average percentage difference' and is the mean GS minus CS count divided by the GS count and multiplied by 100, and σ_p is the standard deviation. 'PD 0/1' is the 'Proportion of differences 0 or 1'; the proportion of images for which the CS count was equal to the GS or different by one individual. The proportion of gold standard (GS) classifications that are greater than (GS > CS), equal to (GS = CS) and less than (GS < CS) citizen science (CS) classifications for the threshold value with the lowest average difference is also given. 'n_a' is the number of images in the sample for average difference and proportion calculations and 'n_p' is the number of images used for percentage difference calculations. The results presented are for images where chicks were present according to gold standard and/or citizen science classifications.

Threshold	**D_a**	**σ_a**	**PD 0/1**	**GS < CS**	**GS = CS**	**GS > CS**	**n_a**	**D_p**	**σ_p**	**n_p**
HVITa										
num_markings										
>2	7.82	5.87	0.11	0.02	0.04	0.95	55	50.80	34.24	55
>3	11.93	6.66	0.09	0.00	0.00	1.00	54	67.97	26.20	54
>4	14.80	7.60	0.06	0.00	0.00	1.00	54	80.41	20.31	54
>5	16.94	8.22	0.04	0.00	0.00	1.00	53	87.68	16.10	53
>6	18.67	8.36	0.04	0.00	0.00	1.00	49	93.99	9.37	49
>7	18.31	9.19	0.05	0.00	0.00	1.00	39	96.90	6.00	39
>8	18.00	10.22	0.08	0.00	0.00	1.00	26	97.87	4.24	26
>9	16.11	11.54	0.11	0.00	0.00	1.00	18	98.68	3.07	18
>10	12.36	10.65	0.18	0.00	0.00	1.00	11	98.86	2.65	11
marked_all										
>1	7.82	5.87	0.11	0.02	0.04	0.95	55	50.80	34.24	55
>2	7.82	5.87	0.11	0.02	0.04	0.95	55	50.80	34.24	55
>3	7.82	5.87	0.11	0.02	0.04	0.95	55	50.80	34.24	55
>4	7.82	5.87	0.11	0.02	0.04	0.95	55	50.80	34.24	55
>5	7.65	5.79	0.11	0.02	0.04	0.94	54	49.89	33.89	54
>6	7.65	5.79	0.11	0.02	0.04	0.94	54	49.89	33.89	54
>7	7.46	5.75	0.12	0.02	0.04	0.94	52	48.80	33.79	52
>8	7.15	5.44	0.10	0.02	0.04	0.94	48	46.39	33.28	48
>9	6.79	5.37	0.12	0.02	0.05	0.93	43	45.43	34.01	43
>10	5.19	4.79	0.19	0.04	0.08	0.88	26	44.16	36.09	26
OSSIa										
num_markings										
>2	2.07	11.24	0.13	0.44	0.07	0.49	55	-0.26	70.05	55
>3	8.36	11.86	0.20	0.20	0.09	0.71	55	37.07	52.45	55
>4	12.15	12.73	0.16	0.05	0.09	0.85	55	59.91	39.66	55
>5	14.10	12.76	0.15	0.02	0.02	0.96	48	70.98	29.84	48
>6	15.73	13.03	0.09	0.00	0.00	1.00	44	81.51	24.20	44
>7	16.65	12.77	0.08	0.00	0.00	1.00	40	89.83	17.35	40
>8	15.83	10.82	0.07	0.00	0.00	1.00	30	96.02	10.18	30
>9	16.50	11.01	0.05	0.00	0.00	1.00	20	97.32	8.16	20
>10	18.56	11.96	0.00	0.00	0.00	1.00	9	97.49	7.53	9
marked_all										
>1	2.07	11.24	0.13	0.44	0.07	0.49	55	-0.26	70.05	55
>2	2.07	11.24	0.13	0.44	0.07	0.49	55	-0.26	70.05	55
>3	2.02	11.44	0.11	0.45	0.08	0.47	53	-4.04	68.51	53
>4	1.65	11.82	0.12	0.49	0.08	0.43	49	-10.53	66.97	49
>5	0.30	9.99	0.13	0.51	0.09	0.40	47	-14.54	65.39	47
>6	-1.85	8.00	0.15	0.59	0.10	0.32	41	-25.69	62.19	41
>7	-2.82	7.34	0.15	0.64	0.09	0.27	33	-34.72	61.03	33
>8	-3.85	7.98	0.20	0.65	0.15	0.20	20	-28.00	44.42	20
>9	-7.89	6.60	0.22	0.89	0.11	0.00	9	-54.70	33.84	9
>10	2.07	11.24	0.13	0.44	0.07	0.49	55	-0.26	70.05	55
ALKEa										
num_markings										
>2	2.00	1.63	0.38	0.08	0.08	0.85	13	70.77	42.71	13

>3	2.46	1.33	0.31	0.00	0.00	1.00	13	84.10	30.49	13
>4	2.69	1.38	0.23	0.00	0.00	1.00	13	90.00	25.17	13
>5	3.00	1.71	0.17	0.00	0.00	1.00	12	94.17	13.79	12
>6	3.17	1.95	0.17	0.00	0.00	1.00	12	96.67	7.78	12
>7	3.25	2.18	0.17	0.00	0.00	1.00	12	97.50	6.22	12
>8	3.42	2.47	0.17	0.00	0.00	1.00	12	100.00	0.00	12
>9	3.40	2.63	0.20	0.00	0.00	1.00	10	100.00	0.00	10
>10	3.00	1.63	0.25	0.00	0.00	1.00	4	100.00	0.00	4
marked_all										
>1	2.00	1.63	0.38	0.08	0.08	0.85	13	70.77	42.71	13
>2	2.00	1.63	0.38	0.08	0.08	0.85	13	70.77	42.71	13
>3	2.00	1.63	0.38	0.08	0.08	0.85	13	70.77	42.71	13
>4	2.00	1.63	0.38	0.08	0.08	0.85	13	70.77	42.71	13
>5	1.92	1.68	0.42	0.08	0.08	0.83	12	68.33	43.66	12
>6	1.64	1.43	0.45	0.09	0.09	0.82	11	65.45	44.58	11
>7	1.64	1.43	0.45	0.09	0.09	0.82	11	65.45	44.58	11
>8	1.78	1.56	0.33	0.11	0.11	0.78	9	57.78	46.04	9
>9	1.80	1.92	0.40	0.20	0.00	0.80	5	44.00	42.78	5
>10	2.50	1.29	0.25	0.00	0.00	1.00	4	60.00	27.08	4
MITTa										
num_markings										
>2	7.15	10.50	0.16	0.12	0.05	0.83	75	46.92	44.89	74
>3	11.99	11.25	0.14	0.03	0.01	0.96	72	70.17	33.13	72
>4	15.16	12.68	0.10	0.01	0.00	0.99	67	80.60	24.48	67
>5	16.74	13.44	0.10	0.02	0.00	0.98	61	88.11	19.90	61
>6	18.18	13.81	0.11	0.00	0.00	1.00	55	92.90	14.79	55
>7	18.44	13.81	0.09	0.00	0.00	1.00	45	95.89	9.94	45
>8	19.88	14.80	0.09	0.00	0.00	1.00	34	97.75	5.93	34
>9	21.04	15.75	0.08	0.00	0.00	1.00	26	99.14	2.75	26
>10	23.73	18.06	0.00	0.00	0.00	1.00	15	100.00	0.00	15
marked_all										
>1	7.15	10.50	0.16	0.12	0.05	0.83	75	46.92	44.89	74
>2	6.64	9.11	0.17	0.13	0.06	0.82	72	44.67	44.45	71
>3	6.49	9.25	0.17	0.13	0.06	0.81	69	43.87	45.02	68
>4	5.91	9.14	0.18	0.14	0.06	0.80	65	43.62	46.24	64
>5	5.24	8.88	0.17	0.16	0.07	0.78	58	38.80	46.35	57
>6	5.87	8.80	0.17	0.13	0.08	0.79	53	40.85	44.11	52
>7	4.68	8.73	0.20	0.16	0.09	0.75	44	35.62	46.12	43
>8	2.37	8.25	0.27	0.23	0.13	0.63	30	25.65	48.10	29
>9	0.89	8.43	0.26	0.21	0.16	0.63	19	19.06	45.58	19
>10	-2.80	8.82	0.40	0.30	0.30	0.40	10	1.65	49.46	10
SKELa										
num_markings										
>2	6.73	8.08	0.36	0.36	0.00	0.64	11	79.50	20.68	7
>3	11.63	8.80	0.25	0.13	0.00	0.88	8	89.60	14.17	7
>4	12.50	9.59	0.25	0.13	0.00	0.88	8	93.55	7.67	7
>5	15.00	9.11	0.14	0.00	0.00	1.00	7	95.82	5.33	7
>6	15.71	9.52	0.14	0.00	0.00	1.00	7	99.32	1.80	7
>7	15.86	9.60	0.14	0.00	0.00	1.00	7	100.00	0.00	7
>8	15.86	9.60	0.14	0.00	0.00	1.00	7	100.00	0.00	7
>9	16.00	10.52	0.25	0.00	0.00	1.00	4	100.00	0.00	4
>10	16.00	10.52	0.25	0.00	0.00	1.00	4	100.00	0.00	4

marked_all										
>1	6.73	8.08	0.36	0.36	0.00	0.64	11	79.50	20.68	7
>2	6.73	8.08	0.36	0.36	0.00	0.64	11	79.50	20.68	7
>3	6.73	8.08	0.36	0.36	0.00	0.64	11	79.50	20.68	7
>4	6.73	8.08	0.36	0.36	0.00	0.64	11	79.50	20.68	7
>5	6.73	8.08	0.36	0.36	0.00	0.64	11	79.50	20.68	7
>6	5.90	8.01	0.40	0.40	0.00	0.60	10	77.12	21.58	6
>7	4.63	8.55	0.50	0.50	0.00	0.50	8	88.51	13.65	4
>8	5.43	8.90	0.43	0.43	0.00	0.57	7	88.51	13.65	4
>9	5.43	8.90	0.43	0.43	0.00	0.57	7	88.51	13.65	4
>10	3.17	7.22	0.50	0.50	0.00	0.50	6	93.65	11.00	3

Table S4 Average number of *Seabird Watch* users and proportion of images where 1-3, 4-6, 7-9 and ≥10 users marked all birds, in images where at least one Kittiwake was seen in gold standard (GS) classifications. ‘n’ is the number of images in the sample for each camera.

Camera	**Average**	σ	**1-3**	**4-6**	**7-9**	**≥10**	**n**
HVITa	10.06897	1.745793	0	0.017241	0.275862	**0.706897**	58
OSSIa	8.760563	2.154496	0	0.140845	**0.507042**	0.352113	71
ALKEa	8.243902	2.332172	0.073171	0.097561	**0.512195**	0.317073	41
MITTa	8.078125	2.427836	0.041667	0.208333	**0.46875**	0.28125	192
SKELa	10	2.249183	0.019231	0.038462	0.288462	**0.653846**	52

DISCUSSION

The premise behind this thesis was to assess the use of time-lapse imagery for monitoring Black-legged Kittiwake productivity and phenology and to provide a provisional insight into some of the potential factors contributing to population change. Kittiwakes are listed as Vulnerable on the IUCN Red List and so effective monitoring is increasingly important to understand the causes of decline (BirdLife International 2019). Time-lapse cameras could allow cost-effective monitoring over much larger spatial and temporal scales than current fieldwork enables, provided that methods are in place to process and analyse the large volume of images collected. In chapters one and two, I investigated the use of Kittiwake nest monitoring data from expert analysis of time-lapse images compared to field observations, using a case study on Skomer Island, Wales. Chapter three then went on to explore the reliability of citizen science derived data compared to expert analysis of time-lapse images.

The case study on Skomer Island showed that time-lapse imagery has the potential to provide reliable measurements of Kittiwake nest success and phenology, but achieving results 'as good as' current field monitoring methods is very much dependent on camera type and positioning (Lorentzen *et al.* 2010, Merkel *et al.* 2016). The distance between camera and colony at The Wick, Skomer, is in the order of 90 m, which is much higher than previous studies capturing photographs of Kittiwakes (Collins *et al.* 2014, De Pascalis *et al.* 2018). Despite having a 2x magnification lens, which increased image resolution compared to the 'standard' camera, the distance was still too large to reliably record nest contents in each of over 100 nests. In particular, annotation of images was less likely to identify both chicks in two chick broods, and chick hatch dates were significantly later than for field observations, as small chicks being brooded by parents were hard to spot in images. While the results of this investigation are specific to Skomer Island, the general principles can be applied more broadly and highlight the importance of considering site suitability and setup when using any form of digital camera for monitoring wildlife populations (Edney & Wood 2020). This concept was further reiterated in chapter three when comparing Kittiwake counts between images annotated by *Seabird Watch* volunteers and gold standard analysers.

The agreement between gold standard and citizen science counts varied according to a number of factors, including aggregation threshold, whether adults and chick counts were considered together or separately, and site/camera, which was linked to the number of birds per image and species present. Camera setup followed the same overall protocol at each site, but the nature of each location meant some cameras were a greater distance from the colony of interest than desired, leading to a large number of nests being present in each image. Images with a high number of birds in both the foreground and background can result in some users only counting birds in the foreground, while others count everything in the photograph, leading to a wide range of counts (Swanson *et al.* 2016). Objects in the background of images are also more likely to be missed completely or misidentified, such as juvenile Kittiwakes being mistaken for adults. Equally, some cameras were positioned slightly below the colony of interest, which made it harder to accurately determine nest contents, compared to if the camera had been positioned above and pointing downwards (Lorentzen *et al.* 2010).

Aggregating the classifications of multiple users per image can help overcome some of the challenges posed by variable image quality. Having multiple users classify images is typically more reliable than a single person, even when that single person is an expert (Swanson *et al.* 2016, Jones *et al.* 2018). While volunteer aggregations are not yet more reliable than gold standard classifications for *Seabird Watch*, it is true that accuracy was greater when at least three volunteers must have marked an area for it to be counted as a bird. Further investigation of the clustering algorithm used to aggregate volunteer classifications should improve reliability of *Seabird Watch* data, such that it becomes of equal or greater accuracy than classification by a single expert. When this is achieved, we may expect accuracy to increase asymptotically with the number of classifiers (Swanson *et al.* 2016). Optimising the trade-off between effort (number of classifiers) and accuracy will then be critical to ensure a steady stream of image processing without compromising data quality and thus ensuring time-lapse imagery remains a credible monitoring tool (Swanson *et al.* 2016).

In the future, with the expansion of *Seabird Watch* to other species, further considerations will need to be made when processing images. For images with mixed species groups, accuracy is likely to vary by species and may relate to rarity (Swanson *et al.* 2016). Another Zooniverse project, *Snapshot Serengeti*, found that visually striking species which are clearly identifiable, like giraffes, were almost always correctly identified by volunteers. Conversely, rare species, like aardwolves, presented fewer opportunities for learning and people were more eager to report them, leading to much lower accuracy. Based on this study and the findings of my own research, I would recommend that a sample of aggregated citizen science data are always validated against images verified by experts when deploying new cameras. It is not enough to make a comparison at just a few sites within a dataset and assume the trend applies to all. Instead, data from every camera should be validated to account for site-specific variation. Further comparison is therefore needed for the other cameras in the *Seabird Watch* network which were not considered here.

Validation studies may seem less 'exciting' than actually using citizen science to answer bigger scientific questions, but it is essential work to ensure the reliability and credibility of citizen science. Many researchers remain sceptical of data from non-experts (Foster-Smith & Evans 2003, Dickinson *et al.* 2010, Bonter & Cooper 2012, Swanson *et al.* 2016), making it even more important to be able to confirm its reliability. Citizen science has the potential to massively increase the rate of image annotation and expand the scope and scale of monitoring (Swanson *et al.* 2015, 2016). Within three days of launching *Snapshot Serengeti*, an 18-month backlog of images had been processed (Swanson *et al.* 2016) and on average, Zooniverse projects save 34 full-time working years due to volunteer involvement (Cox *et al.* 2015). Using volunteer classifications to train computer algorithms to classify images is the next step for optimising image processing. This has already been achieved for several Zooniverse projects, including *Penguin Watch,* and is in the process of being developed for *Seabird Watch*. Images from sites where algorithms consistently perform poorly in species identification, can be purposefully assigned to volunteers, to create an optimal balance between citizen science and machine learning for image classification (Jones *et al.* 2020). This will allow larger scientific questions to be answered. For *Seabird Watch* this could help understand why threatened seabirds are declining.

One of the main causes of Kittiwake decline is thought to be changes in prey availability associated with climate change, and linked to this, weather extremes. Chapter two provided a preliminary investigation into the effect of weather on Kittiwake nest survival and highlighted the potential adverse impacts of high temperatures during the egg stage, and strong winds buffeting colonies during the chick stage. However, further investigation is required across multiple years and multiples sites. It is hoped that these data could in part be provided by *Seabird Watch.*

Annotating images using a citizen science project, like *Seabird Watch*, will become increasingly important if the deployment of time-lapse cameras is to be expanded to increase the spatial and temporal scale of seabird monitoring and better understand the reasons for population change. Manually identifying birds and nest contents in images is time-consuming, and as chapter one revealed, can sometimes take longer than fieldwork. While citizen science can significantly reduce the workload of researchers and increase the rate of image processing, it is limited by volunteer interest. Both the *Seabird Watch* and *Penguin Watch* programmes invest substantial time promoting their projects and developing effective public engagement strategies to maintain interest and a high rate of image annotation (T. Hart, pers. comm.). As the technology develops, it is hoped that an increasing number of projects using digital imagery will develop machine learning algorithms to automatically detect objects in images. While many annotated images may initially be needed to train the algorithm, in the long-term this should minimise the need for manual image analysis and rapidly increase image processing time.

Having a network of cameras monitoring Kittiwake colonies over the coming years and a streamlined image processing system to supplement sites already monitored by fieldworkers, would greatly benefit this species and indeed other cliff- and ground-nesting seabirds of conservation concern. Combining camera data from the breeding grounds with other sources of information, such as GPS tracking, could provide novel insights into the ecology and behaviour of seabirds, and hopefully a better understanding of how we can assist threatened species.

REFERENCES

Anderson, K. & Gaston, K.J. 2013. Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment* **11**: 138–146.

Anker-Nilssen, T., Erikstad, K.E. & Lorentsen, S.-H. 1996. Aims and effort in seabird monitoring: an assessment based on Norwegian data. *wbio* **2**: 17–26.

Arteta, C., Lempitsky, V. & Zisserman, A. 2016. Counting in the wild. *Eur. Conf. Comput. Vis. LNCS.*, 9911: 483–498.

Ayo, J.O., Obidi, J.A. & Rekwot, P.I. 2011. Effects of Heat Stress on the Well-Being, Fertility, and Hatchability of Chickens in the Northern Guinea Savannah Zone of Nigeria: A Review. *ISRN Veterinary Science* **2011**: 1–10.

Barry, C. 2020. Latitudinal variation in the reproductive success of black-legged kittiwakes *Rissa tridactyla*. Master's thesis, University College Cork, Cork, Ireland.

Benowitz-Fredericks, Z.M., Kitaysky, A.S., Welcker, J. & Hatch, S.A. 2013. Effects of Food Availability on Yolk Androgen Deposition in the Black-Legged Kittiwake (Rissa tridactyla), a Seabird with Facultative Brood Reduction. *PLOS ONE* **8**: e62949.

Bennett, P. M. and Owens, I. P. F. 2002. *Evolutionary ecology of birds: life histories, mating systems, and extinction*. New York, USA: Oxford University Press.

Boulinier, T., McCoy, K.D., Yoccoz, N.G., Gasparini, J. & Tveraa, T. 2008. Public information affects breeding dispersal in a colonial bird: kittiwakes cue on neighbours. *Biology Letters* **4**: 538–540.

BirdLife International. 2012. Spotlight on seabirds. Presented as part of the BirdLife State of the world's birds website. Available at: http://www.birdlife.org/datazone (accessed 24 October 2019).

BirdLife International. 2018. State of the World's Birds: Taking the Pulse of the Planet. BirdLife International, Cambridge, UK.

BirdLife International. 2019. *Rissa tridactyla* (amended version of 2018 assessment). *The IUCN Red List of Threatened Species* 2019: e.T22694497A155617539. Available at: https://dx.doi.org/10.2305/IUCN.UK.2018-2.RLTS.T22694497A155617539.en (accessed 18 July 2020).

Black, C. 2018. Spying on seabirds: a review of time-lapse photography capabilities and limitations. *Seabird* **31**:1-14.

Black, C., Collen, B., Lunn, D., Filby, D., Winnard, S. & Hart, T. 2018a. Time-lapse cameras reveal latitude and season influence breeding phenology durations in penguins. *Ecology and Evolution* **8**: 8286–8296.

Black, C., Rey, A.R. & Hart, T. 2017. Peeking into the bleak midwinter: Investigating nonbreeding strategies of Gentoo Penguins using a camera network. *The Auk* **134**: 520–529.

Black, C., Southwell, C., Emmerson, L., Lunn, D. & Hart, T. 2018b. Time-lapse imagery of Adélie penguins reveals differential winter strategies and breeding site occupation. *PLOS ONE* **13**: e0193532.

Bland, L.M., Watermeyer, K.E., Keith, D.A., Nicholson, E., Regan, T.J. & Shannon, L.J. 2018. Assessing risks to marine ecosystems with indicators, ecosystem models and experts. *Biological Conservation* **227**: 19–28.

Bolton, M., Butcher, N., Sharpe, F., Stevens, D. & Fisher, G. 2007. Remote monitoring of nests using digital camera technology. *Journal of Field Ornithology* **78**: 213–220.

Bonter, D.N. & Cooper, C.B. 2012. Data validation in citizen science: a case study from Project FeederWatch. *Frontiers in Ecology and the Environment* **10**: 305–307.

Brown, JG, Perrins, CM, Boyle, D, Duffield, S, Dustan, A, Easton, J, Smith, S and Parsons, M. 2004. Seabird monitoring on Skomer Island in 1999-2002. *JNCC Report No.339.*

Bruggemann, J., Lander, G.C. & Su, A.I. 2018. Exploring applications of crowdsourcing to cryo-EM. *Journal of Structural Biology* **203**: 37–45.

Burnham, K.P. & Anderson, D.R. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. New York, USA: Springer.

Burnham, K.P., Anderson, D.R. & Huyvaert, K.P. 2011. AIC model selection and multimodel inference in behavioral ecology: some background, observations, and comparisons. *Behav Ecol Sociobiol* **65**: 23–35.

Burthe, S., Daunt, F., Butler, A., Elston, D.A., Frederiksen, M., Johns, D., Newell, M., Thackeray, S.J. & Wanless, S. 2012. Phenological trends and trophic mismatch across multiple levels of a North Sea pelagic food web. *Marine Ecology Progress Series* **454**: 119–133.

Byrd, G.V., Sydeman, W.J., Renner, H.M. & Minobe, S. 2008. Responses of piscivorous seabirds at the Pribilof Islands to ocean climate. *Deep Sea Research Part II: Topical Studies in Oceanography* **55**: 1856–1867.

Carroll, M.J., Butler, A., Owen, E., Ewing, S.R., Cole, T., Green, J.A., Soanes, L.M., Arnould, J.P.Y., Newton, S.F., Baer, J., Daunt, F., Wanless, S., Newell, M.A., Robertson, G.S., Mavor, R.A. & Bolton, M. 2015. Effects of sea temperature and stratification changes on seabird breeding success. *Climate Research* **66**: 75–89.

Chastel, O., Weimerskirch, H. & Jouventin, P. 1995. Body Condition and Seabird Reproductive Performance: A Study of Three Petrel Species. *Ecology* **76**: 2240–2246.

Chivers, L.S., Lundy, M.G., Colhoun, K., Newton, S.F., Houghton, J.D.R. & Reid, N. 2012. Foraging trip time-activity budgets and reproductive success in the black-legged kittiwake. *Marine Ecology Progress Series* **456**: 269–277.

Christensen-Dalsgaard, S., May, R., Barrett, R., Langset, M., Sandercock, B. & Lorentsen, S. 2018. Prevailing weather conditions and diet composition affect chick growth and survival in the black-legged kittiwake. *Mar. Ecol. Prog. Ser.* **604**: 237–249.

Collins, P.M., Green, J.A., Dodd, S., Shaw, P.J.A. & Halsey, L.G. 2014. Predation of Black-legged Kittiwake Chicks *Rissa tridactyla* by a Peregrine Falcon *Falco peregrinus* : Insights from Time-lapse Cameras. *The Wilson Journal of Ornithology* **126**: 158–161.

Conkling, T.J., Belant, J.L., DeVault, T.L., Wang, G. & Martin, J.A. 2015. Assessment of variation of nest survival for grassland birds due to method of nest discovery. *Bird Study* **62**: 223–231.

Cook, A.S.C.P., Dadam, D., Mitchell, I., Ross-Smith, V.H. & Robinson, R.A. 2014. Indicators of seabird reproductive performance demonstrate the impact of commercial fisheries on seabird populations in the North Sea. *Ecological Indicators* **38**: 1–11.

Coulson, J. 2011. *The Kittiwake*. London, UK: T. and A.D. Polser.

Coulson, J.C. & Porter, J.M. 1985. Reproductive success of the Kittiwake Rissa tridactyla: the roles of clutch size, chick growth rates and parental quality. *Ibis* **127**: 450–466.

Coulson, J.C. & White, E. 1958. Observations on the Breeding Kittiwake. *Bird Study* **5**: 74–83.

Coulson, J.C. & White, E. 1961. An Analysis of the Factors Influencing the Clutch Size of the Kittiwake. *Proceedings of the Zoological Society of London* **136**: 207–217.

Cowardin, L.M. & Ashe, J.E. 1965. An Automatic Camera Device for Measuring Waterfowl Use. *The Journal of Wildlife Management* **29**: 636–640.

Cox, J., Oh, E.Y., Simmons, B., Lintott, C., Masters, K., Greenhill, A., Graham, G. & Holmes, K. 2015. Defining and Measuring Success in Online Citizen Science: A Case Study of Zooniverse Projects. *Computing in Science Engineering* **17**: 28–41.

Cramp, S., Bourne, W.R.P. & Saunders, D. 1974. The seabirds of Britain and Ireland. London, UK: Collins.

Croxall, J.P., Butchart, S.H.M., Lascelles, B., Stattersfield, A.J., Sullivan, B., Symes, A. & Taylor, P. 2012. Seabird conservation status, threats and priority actions: a global assessment. *Bird Conservation International* **22**: 1–34.

Cutler, T.L. & Swann, D.E. 1999. Using Remote Photography in Wildlife Ecology: A Review. *Wildlife Society Bulletin (1973-2006)* **27**: 571–581.

Daunt, F. & Mitchell, I. 2013. Impacts of climate change on seabirds. MCCIP Science Review 2013, 125-133.

De Pascalis, F., Collins, P. & Green, J. 2018. Utility of time-lapse photography in studies of seabird ecology. *PLOS ONE* **13**: e0208995.

Descamps, S., Tarroux, A., Varpe, Ø., Yoccoz, N.G., Tveraa, T. & Lorentsen, S.-H. 2015. Demographic effects of extreme weather events: snow storms, breeding success, and population growth rate in a long-lived Antarctic seabird. *Ecology and Evolution* **5**: 314–325.

Dias, M.P., Martin, R., Pearmain, E.J., Burfield, I.J., Small, C., Phillips, R.A., Yates, O., Lascelles, B., Borboroglu, P.G. & Croxall, J.P. 2019. Threats to seabirds: A global assessment. *Biological Conservation* **237**: 525–537.

Dickinson, J.L., Zuckerberg, B. & Bonter, D.N. 2010. Citizen Science as an Ecological Research Tool: Challenges and Benefits. *Annual Review of Ecology, Evolution, and Systematics* **41**: 149–172.

Dinsmore, S. & Dinsmore, J. 2007. Modeling Avian Nest Survival in Program MARK. *Studies in Avian Biology* 73–83.

Dodge, W.E. & Snyder, D.P. 1960. An Automatic Camera Device for Recording Wildlife Activity. *The Journal of Wildlife Management* **24**: 340–342.

Dunnet, G.M. 1991. Population studies of the Fulmar on Eynhallow, Orkney Islands. *Ibis* **133**: 24–27.

Eerkes-Medrano, D., Fryer, R.J., Cook, K.B. & Wright, P.J. 2017. Are simple environmental indicators of food web dynamics reliable: Exploring the kittiwake–temperature relationship. *Ecological Indicators* **75**: 36–47.

Evans, P.G.H. 1986. Monitoring Seabirds in the North Atlantic. In: *Mediterranean Marine Avifauna* (X. Monbailliu, ed), pp. 179–206. Springer, Berlin, Heidelberg.

Field, S.A., Tyre, A.J. & Possingham, H.P. 2005. Optimizing Allocation of Monitoring Effort Under Economic and Observational Constraints. *The Journal of Wildlife Management* **69**: 473–482.

Foster-Smith, J. & Evans, S.M. 2003. The value of marine ecological data collected by volunteers. *Biological Conservation* **113**: 199–213.

Frederiksen, M., Anker-Nilssen, T., Beaugrand, G. & Wanless, S. 2013. Climate, copepods and seabirds in the boreal Northeast Atlantic – current state and future outlook. *Global Change Biology* **19**: 364–372.

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

Frederiksen, M., Mavor, R.A. & Wanless, S. 2007. Seabirds as environmental indicators: the advantages of combining data sets. *Marine Ecology Progress Series* **352**: 205–211.

Frederiksen, M., Wright, P.J., Harris, M.P., Mavor, R.A., Heubeck, M. & Wanless, S. 2005. Regional patterns of kittiwake Rissa tridactyla breeding success are related to variability in sandeel recruitment. *Marine Ecology Progress Series* **300**: 201–211.

Gabrielsen, G.W., Klaassen, M. & Mehlum, F. 1992. Energetics of Black-legged Kittiwake Rissa-Tridactyla chicks. *Ardea* **80**: 29–40.

Gaston, A.J., Hipfner, J.M. & Campbell, D. 2002. Heat and mosquitoes cause breeding failures and adult mortality in an Arctic-nesting seabird. *Ibis* **144**: 185–191.

Golet, G.H., Schmutz, J.A., Irons, D.B. & Estes, J.A. 2004. Determinants of Reproductive Costs in the Long-Lived Black-Legged Kittiwake: A Multiyear Experiment. *Ecological Monographs* **74**: 353–372.

Green, G.W. & Anderson, D.C. 1961. A Simple and Inexpensive Apparatus for Photographing Events at Pre-set Intervals1. *The Canadian Entomologist* **93**: 741–745.

Gurney, J.H. 1913. *The gannet : a bird with a history*. Witherby, London.

Harris, M.P. 1982. Promiscuity in the Shag as Shown by Time-Lapse Photography. Bird Study **29**: 149-154.

Harris, M.P. 1987. A low-input method of monitoring kittiwake Rissa tridactyla breeding success. *Biological Conservation* **41**: 1–10.

Hentati-Sundberg, J. 2011. Arctic Seabirds Breeding in the African-Eurasian Waterbird Agreement (AEWA) Area. CAFF's Circumpolar Seabird expert group (CBird), CAFF Assessment Series Report No. 1, CAFF International Secretariat, Akureyri, Iceland. ISBN 978-9935-431-03-5.

Hinke, J.T., Barbosa, A., Emmerson, L.M., Hart, T., Juáres, M.A., Korczak-Abshire, M., Milinevsky, G., Santos, M., Trathan, P.N., Watters, G.M. & Southwell, C. 2018. Estimating nest-level phenology and reproductive success of colonial seabirds using time-lapse cameras. *Methods in Ecology and Evolution* **9**: 1853–1863.

Horswill, C., Ratcliffe, N., Green, J.A., Phillips, R.A., Trathan, P.N. & Matthiopoulos, J. 2016. Unravelling the relative roles of top-down and bottom-up forces driving population change in an oceanic predator. *Ecology* **97**: 1919–1928.

Huffeldt, N.P. & Merkel, F.R. 2013. Remote Time-lapse Photography as a Monitoring Tool for Colonial Breeding Seabirds: A Case Study Using Thick-billed Murres (Uria lomvia). *Waterbirds: The International Journal of Waterbird Biology* **36**: 330–341.

Jenouvrier, S., Barbraud, C., Cazelles, B. & Weimerskirch, H. 2005. Modelling population dynamics of seabirds: importance of the effects of climate fluctuations on breeding proportions. *Oikos* **108**: 511–522.

Jentsch, A., Kreyling, J. & Beierkuhnlein, C. 2007. A new generation of climate-change experiments: events, not trends. *Frontiers in Ecology and the Environment* **5**: 365–374.

Joint Nature Conservation Committee. 2020. *Seabird Monitoring.* Available at: https://jncc.gov.uk/our-work/seabird-monitoring-programme/ (accessed 15 March 2021)

Jones, F.M., Allen, C., Arteta, C., Arthur, J., Black, C., Emmerson, L.M., Freeman, R., Hines, G., Lintott, C.J., Macháčková, Z., Miller, G., Simpson, R., Southwell, C., Torsey, H.R., Zisserman, A. & Hart, T. 2018. Time-lapse imagery and volunteer classifications from the Zooniverse Penguin Watch project. *Sci Data* **5**: 180124.

Jones, F.M., Arteta, C., Zisserman, A., Lempitsky, V., Lintott, C.J. & Hart, T. 2020. Processing citizen science- and machine-annotated time-lapse imagery for biologically meaningful metrics. *Scientific Data* **7**: 1–15.

Kennedy, R.J. 1970. Direct effects of rain on birds: a review. *Br Birds* **63**: 401-414.

Kildaw, S.D. 1999. Effect of wind on the growth rate of kittiwake chicks. Pacific Seabirds, Washington, 26 pp.

Kucera, T.E. & Barrett, R.H. 2011. A History of Camera Trapping. In: *Camera Traps in Animal Ecology: Methods and Analyses* (A. F. O'Connell, J. D. Nichols, & K. U. Karanth, eds), pp. 9–26. Springer Japan, Tokyo.

Laake, J. 2013. "RMark: An R Interface for Analysis of Capture-Recapture Data with MARK." AFSC Processed Rep. 2013-01, Alaska Fish. Sci. Cent., NOAA, Natl. Mar. Fish. Serv., Seattle, WA. http://www.afsc.noaa.gov/Publications/ProcRpt/PR2013-01.pdf.

Lauria, V., Attrill, M.J., Pinnegar, J.K., Brown, A., Edwards, M. & Votier, S.C. 2012. Influence of Climate Change and Trophic Coupling across Four Trophic Levels in the Celtic Sea. *PLOS ONE* **7**: e47408.

Lescroël, A., Mathevet, R., Péron, C., Authier, M., Provost, P., Takahashi, A. & Grémillet, D. 2016. Seeing the ocean through the eyes of seabirds: A new path for marine conservation? *Marine Policy* **68**: 212–220.

Lorentzen, E., Steen, H. & Strøm, H. 2010. Estimating chick survival in cliff-nesting seabirds – a hazard made easy with monitoring cameras. *SEAPOP Short Report* **8**. 1-10.

Lynch, H.J., White, R., Black, A.D. & Naveen, R. 2012. Detection, differentiation, and abundance estimation of penguin species by high-resolution satellite imagery. *Polar Biol* **35**: 963–968.

Mallory, M.L., Gaston, A.J. & Gilchrist, H.G. 2009. Sources of Breeding Season Mortality in Canadian Arctic Seabirds. *Arctic* **62**: 333–341.

Merkel, F.R., Johansen, K.L. & Kristensen, A.J. 2016. Use of time-lapse photography and digital image analysis to estimate breeding success of a cliff-nesting seabird. *Journal of Field Ornithology* **87**: 84–95.

Met Office. 2012. Met Office Integrated Data Archive System (MIDAS) Land and Marine Surface Stations Data (1853-current). NCAS British Atmospheric Data Centre (accessed 18th April 2020).

Michielsen, R.J., Ausems, A.N.M.A., Jakubas, D., Pętlicki, M., Plenzler, J., Shamoun-Baranes, J. & Wojczulanis-Jakubas, K. 2019. Nest characteristics determine nest microclimate and affect breeding output in an Antarctic seabird, the Wilson's storm-petrel. *PLOS ONE* **14**: e0217708.

Mitchell, I., Cook, A., Douse, A., Foster, S., Kershaw, M., McCulloch, N., Murphy, M., Hawkridge.,, J. 2018. Kittiwake breeding success. UK Marine Online Assessment Tool, available at: https://moat.cefas.co.uk/biodiversity-food-webs-and-marine-protected-areas/birds/kittiwake-breeding-success/.

Mitchell, I., Newton, S., Ratcliffe, N., Dunn, T. 2004. Seabird Populations of Britain and Ireland: results of the Seabird 2000 census (1998-2002). London, UK: T. and A.D. Poyser.

Mitchell, P.I. & Parsons, M. 2007. Strategic Review of the UK Seabird Monitoring Programme. *JNCC Unpublished Report.*

Moe, B., Stempniewicz, L., Jakubas, D., Angelier, F., Chastel, O., Dinessen, F., Gabrielsen, G., Hanssen, F., Karnovsky, N., Rønning, B., Welcker, J., Wojczulanis-Jakubas, K. & Bech, C. 2009. Climate change and phenological responses of two seabird species breeding in the high-Arctic. *Mar. Ecol. Prog. Ser.* **393**: 235–246.

Mudge, G.P., Aspinall, S.J. & Crooke, C.H. 1987. A photographic study of seabird attendance at Moray Firth colonies outside the breeding season. *Bird Study* **34**: 28–36.

Mugar, G., Østerlund, C., Hassman, K.D., Crowston, K. & Jackson, C.B. 2014. Planet hunters and seafloor explorers: legitimate peripheral participation through practice proxies in online citizen science. In: *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing - CSCW '14*, pp. 109–119. ACM Press, Baltimore, Maryland, USA.

Mwangi, J., Ndithia, H.K., Kentie, R., Muchai, M. & Tieleman, B.I. 2018. Nest survival in year-round breeding tropical red-capped larks *Calandrella cinerea* increases with higher nest abundance but decreases with higher invertebrate availability and rainfall. *J Avian Biol* **49**: e01645.

Naimi, B., Hamm, N.A.S., Groen, T.A., Skidmore, A.K., & Toxopeus, A.G. 2014. Where is positional uncertainty a problem for species distribution modelling?, Ecography **37**: 191-203.

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

Oswald, S.A., Bearhop, S., Furness, R.W., Huntley, B. & Hamer, K.C. 2008. Heat stress in a high-latitude seabird: effects of temperature and food supply on bathing and nest attendance of great skuas Catharacta skua. *Journal of Avian Biology* **39**: 163–169.

Pacyna, A.D., Jakubas, D., Ausems, A.N.M.A., Frankowski, M., Polkowska, Ż. & Wojczulanis-Jakubas, K. 2019. Storm petrels as indicators of pelagic seabird exposure to chemical elements in the Antarctic marine ecosystem. *Science of The Total Environment* **692**: 382–392.

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

Penney, R.L. 1968. Territorial and Social Behavior in the Adélie Penguin. In: *Antarctic Bird Studies*, pp. 83–131. American Geophysical Union (AGU).

Pierce, A. K. 2017. Population Status of Mountain Plover in South Park, Colorado. Master's thesis, University of Colorao, Denver, US.

Ponchon, A., Chambert, T., Lobato, E., Tveraa, T., Grémillet, D. & Boulinier, T. 2015. Breeding failure induces large scale prospecting movements in the black-legged kittiwake. *Journal of Experimental Marine Biology and Ecology* **473**: 138–145.

Qu, B., Gabric, A.J., Zhu, J., Lin, D., Qian, F. & Zhao, M. 2012. Correlation between sea surface temperature and wind speed in Greenland Sea and their relationships with NAO variability. *Water Science and Engineering* **5**: 304–315.

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raddick, M.J., Bracey, G., Gay, P.L., Lintott, C.J., Cardamone, C., Murray, P., Schawinski, K., Szalay, A.S. & Vandenberg, J. 2013. Galaxy Zoo: Motivations of Citizen Scientists. *arXiv:1303.6886 [astro-ph, physics:physics]*Available at: http://arxiv.org/abs/1303.6886.

Raddick, M.J., Bracey, G., Gay, P.L., Lintott, C.J., Murray, P., Schawinski, K., Szalay, A.S. & Vandenberg, J. 2010. Galaxy Zoo: Exploring the Motivations of Citizen Science Volunteers. *Astronomy Education Review* **9**.

Regehr, H. & Montevecchi, W. 1997. Interactive effects of food shortage and predation on breeding failure of black-legged kittiwakes:indirect effects of fisheries activities and implications for indicator species. *Mar. Ecol. Prog. Ser.* **155**: 249–260.

Reiertsen, T.K. 2013. Seabirds, climate and prey. A population study of two seabird species. PhD dissertation, University of Tromsø.

Reynolds, S.J., Hughes, B.J., Wearn, C.P., Dickey, R.C., Brown, J., Weber, N.L., Weber, S.B., Paiva, V.H. & Ramos, J.A. 2019. Long-term dietary shift and population decline of a pelagic seabird—A health check on the tropical Atlantic? *Global Change Biology* **25**: 1383–1394.

Robinson, R.A. & Ratcliffe, N. 2010. The Feasibility of Integrated Population Monitoring of Britain's Seabirds. 73.

Rotella, J.J., Dinsmore, S.J. & Shaffer, T.L. 2004. Modeling nest–survival data: a comparison of recently developed methods that can be implemented in MARK and SAS. *Animal Biodiversity and Conservation* 19.

Ryan, R. 2019. Variations in black-legged kittiwakes (*Rissa tridactyla*) breeding success in in their arctic distribution and environmental influences on the timing of phenological events. Master's thesis, University College Cork, Cork, Ireland.

Sæther, B.-E. & Bakke, Ø. 2000. Avian Life History Variation and Contribution of Demographic Traits to the Population Growth Rate. *Ecology* **81**: 642–653.

Sandvik, H., Reiertsen, T.K., Erikstad, K.E., Anker-Nilssen, T., Barrett, R.T., Lorentsen, S.-H., Systad, G.H. & Myksvoll, M.S. 2014. The decline of Norwegian kittiwake populations: modelling the role of ocean warming. *Climate Research* **60**: 91–102.

Seddon, P.J. & van Heezik, Y. 1991. Effects of Hatching Order, Sibling Asymmetries, and Nest Site on Survival Analysis of Jackass Penguin Chicks. *The Auk* **108**: 548–555.

Seneviratne, S.I., Nicholls, N., Easterling, D., Goodess, C.M., Kanae, S., Kossin, J., Luo, Y., Marengo, J., McInnes, K., Rahimi, M., Reichstein, M., Sorteberg, A., Vera, C., Zhang, X., Rusticucci, M., Semenov, V., Alexander, L.V., Allen, S., Benito, G., Cavazos, T., Clague, J., Conway, D., Della-Marta, P.M., Gerber, M., Gong, S., Goswami, B.N., Hemer, M., Huggel, C., van den Hurk, B., Kharin, V.V., Kitoh, A., Tank, A.M.G.K., Li, G., Mason, S., McGuire, W., van Oldenborgh, G.J., Orlowsky, B., Smith, S., Thiaw, W., Velegrakis, A., Yiou, P., Zhang, T., Zhou, T. & Zwiers, F.W. 2012. Changes in Climate Extremes and their Impacts on the Natural Physical Environment. In: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation* (C. B. Field, V. Barros, T. F. Stocker, & Q. Dahe, eds), pp. 109–230. Cambridge University Press, Cambridge.

Sherley, R.B., Ludynia, K., Underhill, L.G., Jones, R. & Kemper, J. 2012. Storms and heat limit the nest success of Bank Cormorants: implications of future climate change for a surface-nesting seabird in southern Africa. *J Ornithol* **153**: 441–455.

Sinclair, N.C., Harris, M.P., Nager, R.G., Leakey, C.D.B. & Robbins, A.M. 2017. Nocturnal colony attendance by common guillemots Uria aalge at colony in Shetland during the pre-breeding season. *Seabird* **30**: 51–62.

Southwell, C. & Emmerson, L. 2015. Remotely-operating camera network expands Antarctic seabird observations of key breeding parameters for ecosystem monitoring and management. *Journal for Nature Conservation* **23**: 1–8.

Stubbings, E.M., Büche, B.I., Riordan, J.A., Moss, J. & Wood, M.J. (2017). Seabird monitoring on Skomer Island in 2017. *JNCC Report*.

Suryan, R.M., Irons, D.B., Brown, E.D., Jodice, P.G.R. & Roby, D.D. 2006. Site-specific effects on productivity of an upper trophic-level marine predator: Bottom-up, top-down, and mismatch effects on reproduction in a colonial seabird. *Progress in Oceanography* **68**: 303–328.

Swanson, A., Kosmala, M., Lintott, C. & Packer, C. 2016. A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. *Conservation Biology* **30**: 520–531.

Swanson, A., Kosmala, M., Lintott, C., Simpson, R., Smith, A. & Packer, C. 2015. Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. *Scientific Data* **2**: 150026.

Sydeman, W.J., Piatt, J.F., Thompson, S.A., García-Reyes, M., Hatch, S.A., Arimitsu, M.L., Slater, L., Williams, J.C., Rojek, N.A., Zador, S.G., Renner, H.M. 2017a. Puffins reveal contrasting relationships between forage fish and ocean climate in the North Pacific. *Fisheries Oceanography* **26**: 379–395.

Sydeman, W.J., Thompson, S.A., Piatt, J.F., García-Reyes, M., Zador, S., Williams, J.C., Romano, M., Renner, H.M. 2017b. Regionalizing indicators for marine ecosystems: Bering Sea–Aleutian Island seabirds, climate, and competitors. *Ecological Indicators* **78**: 458–469.

Tanedo, S.A. & Hollmén, T.E. 2020. Refining remote observation techniques to estimate productivity of Black-legged Kittiwakes Rissa tridactyla in Resurrection Bay, Gulf of Alaska. *Marine Ornithology* **48**: 61–69.

Tasker, M.L. 2000. The UK and Ireland seabird monitoring programme – a history and introduction. Atlantic seabirds **2**: 97–102.

Taylor, C.J., Baer J., Boyle, D., and Perrins, C.M. 2010. Seabird monitoring on Skomer Island in 2010. *JNCC Report.*

Taylor, I.R. 1983. Effect of Wind on the Foraging Behaviour of Common and Sandwich Terns. *Ornis Scandinavica (Scandinavian Journal of Ornithology)* **14**: 90–96.

Temple, S.A. 1972. A Portable Time-Lapse Camera for Recording Wildlife Activity. *The Journal of Wildlife Management* **36**: 944–947.

Waldron, A., Mooers, A.O., Miller, D.C., Nibbelink, N., Redding, D., Kuhn, T.S., Roberts, J.T. & Gittleman, J.L. 2013. Targeting global conservation funding to limit immediate biodiversity declines. *PNAS* **110**: 12144–12148.

Walsh, P.M., Halley, D.J., Harris, M.P., Del Nevo, A., Sim, I.M.W. & Tasker, M.L. 1995. *Seabird monitoring handbook for Britain and Ireland*. Peterborough, UK: JNCC / RSPB /ITE / Seabird Group.

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T.J.C., Fromentin, J.-M., Hoegh-Guldberg, O. & Bairlein, F. 2002. Ecological responses to recent climate change. *Nature* **416**: 389–395.

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

Weimerskirch, H., Brothers, N. & Jouventin, P. 1997. Population dynamics of wandering albatross Diomedea exulans and Amsterdam albatross D. amsterdamensis in the Indian Ocean and their relationships with long-line fisheries: Conservation implications. *Biological Conservation* **79**: 257–270.

Weller, M.W. & Derksen, D.V. 1972. Use of Time-Lapse photography to Study Nesting Activities of Birds. *The Auk* **89**: 196–200.

White, G.C. & Burnham, K.P. 1999. Program MARK: survival estimation from populations of marked animals. Bird Study **46**: 120–139.

White, G.C. 2011. Program MARK. Mark and recapture survival rate estimation (Version 6.1). http://warnercnr.colostate.edu/~gwhite/ mark/mark.htm

Wilkie, N.G., Zbijewska, S.M., Piggott, A.R., Hastie, V. & Wood, M.J. 2019. Seabird monitoring on Skomer Island in 2019. *JNCC Report.*

Zooniverse. 2020a. *About*. Available at: https://www.zooniverse.org/about (accessed XX September 2020).

Zooniverse. 2020b. *Seabirdwatch*. Available at: https://www.zooniverse.org/projects/penguintom79/seabirdwatch (accessed XX September 2020).

Zuur, A.F., Ieno, E.N. & Elphick, C.S. 2010. A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution* **1**: 3–14.

APPENDIX

APPENDIX 1

The following manuscript is the final draft submitted to *Ibis* prior to acceptance and publication. The published article can be accessed here:

Edney, A.J. and Wood, M.J. (2020), Applications of digital imaging and analysis in seabird monitoring and research. Ibis. doi:10.1111/ibi.12871

Running head: ***Digital Imaging Technology in Seabird Monitoring***

Applications of Digital Imaging and Analysis in Seabird Monitoring and Research

ALICE J. EDNEY[1*] & MATT J. WOOD[1]

[1]*School of Natural & Social Sciences, University of Gloucestershire, Francis Close Hall, Cheltenham GL50 4AZ*

[*]Corresponding author.

Email: [redacted]

Twitter: [redacted]

Rapid advances in digital imaging technology offer efficient and cost-effective methods for measuring seabird abundance, breeding success, phenology, survival and diet. These methods can facilitate understanding of long-term population trends, and design and implementation of successful conservation strategies. This paper reviews the suitability of satellites, manned aircraft, Unmanned Aerial Vehicles (UAVs), fixed position, handheld and animal-borne cameras for recording digital photographs and videos used to measure seabird demographic and behavioural parameters. It considers the disturbance impacts, accuracy of results obtained, cost-effectiveness, and scale of monitoring possible compared to 'traditional' fieldworker methods. Given the ease of collecting large amounts of imagery, image processing is an important step in realising the potential of this technology. The effectiveness of manual, semi-automated and automated image processing are also reviewed. Satellites, manned aircraft and UAVs have most commonly been used for population counts. Spatial resolution is lowest in satellites, limiting monitoring to large species and those with obvious signs of presence, such as penguins. Conversely, UAVs have the highest spatial resolution, which has allowed fine-scale measurements of foraging behaviour. Time-lapse cameras are more cost-effective for collecting time-series data such as breeding success and phenology, as human visits are only required infrequently for maintenance. However, the colony of interest must be observable from a single vantage point. Handheld, animal-borne and motion-triggered cameras have fewer cost-effective uses, but have provided information on seabird diet, foraging behaviour and nest predation. The latter has been important for understanding the impact of invasive mammals on seabird breeding success. Advances in automated image analysis are increasing the suitability of digital photography and videography to facilitate and/or replace traditional seabird monitoring methods. Machine learning algorithms, such as *Pengbot,* have allowed rapid identification of birds, although training requires thousands of pre-annotated photographs. Digital imaging has considerable potential in seabird monitoring, provided that appropriate choices are available for both image capture technology and image processing. These technologies offer opportunities to collect data in remote locations and increase the number of sites monitored. The potential to include such solutions in seabird monitoring and research will develop as the technology evolves, which will be of benefit given funding challenges in monitoring and conservation.

Keywords: remote sensing, photography, videography, population ecology, conservation, seabirds

Seabirds are one of the most threatened groups of birds, with almost half of seabird species experiencing population declines (Croxall *et al.* 2012). Effective monitoring is essential to understand long-term population trends, so that conservation action can be implemented (Walsh *et al.* 1995, Anker-Nilssen *et al.* 1996, Petersen *et al.* 2008). However, monitoring seabird populations can be challenging. Pelagic species spend most of the year at sea, only returning to land to breed. Many nest on exposed cliffs with difficult access, especially during periods of inclement weather, while ground nesting birds may be concealed by camouflage or vegetation and some species nest underground (Mitchell & Parsons 2007, Robinson & Ratcliffe 2010). Furthermore, regularly visiting breeding colonies is logistically difficult in remote locations, can cause disturbance, and is often expensive in terms of time and money (Anker-Nilssen *et al.* 1996, Field *et al.* 2005, Huffeldt & Merkel 2013, Southwell & Emmerson 2015). As a result, monitoring efforts are often restricted to small temporal and spatial scales (Evans 1986, Lynch *et al.* 2012a, Paleczny *et al.* 2015).

Recent advances in digital imaging technology offer considerable potential for overcoming some of the challenges associated with monitoring seabird populations. Digital photography has a long-history in wildlife monitoring but has previously been limited to small studies that observe animals opportunistically, using handheld or animal-triggered cameras (Black 2018). Now, increased battery life, data storage and better optics, have transformed the potential of remote photography and videography and made it possible to monitor populations that are hard to access (Bolton *et al.* 2007, Kucera & Barrett 2011, Anderson & Gaston 2013, Black 2018). Nevertheless, the wide range of technology available can make it challenging to decide which type of equipment is most suitable for a specific monitoring purpose, and how to handle and analyse large amounts of digital data.

Here we summarise the main technologies available for collecting digital data on seabird populations, and offer a critical assessment of each data collection methods. The suitability of each technology for measuring demographic and behavioural parameters is assessed in relation to the disturbance caused, accuracy of results obtained, cost-effectiveness, and scale of monitoring possible, compared to non-digital (termed 'traditional') methods. In particular, we focus on the ability of satellites, manned aircraft, Unmanned Aerial Vehicles (UAVs), and handheld, animal-borne and fixed position (including time-lapse, video and motion-triggered) cameras to monitor the abundance, breeding success, phenology, survival, and diet of seabird populations at sea and on land. This includes surface-nesting and cliff-nesting seabirds, seaducks on inland bodies of water, and seabirds at sea. We assess the accuracy and cost of manual, semi-automated, and automated image analysis methods, as well as considering future developments needed in the field. Our hope is that by drawing information together from many individual studies, this review can help researchers decide where digital photography and videography could facilitate seabird monitoring, in a world that can be short of time and money for conservation endeavours (Waldron *et al.* 2013).

COLLECTION OF DIGITAL IMAGERY

Satellites

One of the first developments in remote sensing technology was the use of satellites for aerial surveys. Although more commonly used to survey vegetation, satellite imagery was used for seabird monitoring as early as the 1980s (Schwaller *et al.* 1989, Nowak *et al.* 2019). Images have been used to locate and count seabird populations, including penguins, Masked Booby *Sula dactylatra*, and Wandering Albatross

Diomedea exulans (Fig. 1) (Schwaller *et al.* 1989, Guinet *et al.* 1995, Fretwell & Trathan 2009, Hughes *et al.* 2011, Fretwell *et al.* 2012, Lynch *et al.* 2012b, Fretwell *et al.* 2014, Waluda *et al.* 2014, Fretwell *et al.* 2017, Borowicz *et al.* 2018, Dolliver *et al.* 2019). The downward-facing perspective of satellites means images are unlikely to provide a representative view of cliff-nesting species, but they are suitable for observing surface-nesting seabirds, seabirds at sea and seaducks inland.

The primary advantage of satellite imagery is its global coverage. This has allowed the discovery of previously unknown populations, often in remote, inaccessible areas (Fretwell & Trathan 2009, Fretwell *et al.* 2012, 2014, Ancel *et al.* 2017, Borowicz *et al.* 2018). Moreover, satellite data collection occurs at such high altitude that it does not disturb birds or habitats, unlike ground, boat or other aerial surveys, making satellites ideal for monitoring sensitive species and locations.

This high-altitude view and the lack of control over the spectral, spatial and temporal resolution of images means that many populations are not visible in sufficient detail to be accurately counted from satellites (Rush *et al.* 2018, Nowak *et al.* 2019). The trade-off between spatial and temporal resolution also limits their ability to collect the frequent, high resolution images needed to measure breeding success. Terra and Aqua satellites with MODIS sensors have high temporal resolution (four images every 24 hours), but very low spatial resolution; whereas Landsat or Sentinel-2 satellites have high spatial resolution (10-30 m) but low temporal resolution (one image every 16 days) (Nowak *et al.* 2019). A fixed re-visit time means image frequency may be further reduced if poor weather conditions such as low cloud obscure the area of interest when in the satellite's view (Müllerova *et al.* 2017, Nowak *et al.* 2019). Ground cover will also affect bird visibility, making satellite imagery unsuitable for monitoring burrow-nesting species and those nesting in dense habitat, like long grass. Furthermore, none of the freely available satellite images have < 1 m spatial resolution and acquiring images from commercial suppliers is expensive (Nowak *et al.* 2019). Consequently, satellites can offer a cost-effective method of counting some seabird populations, but only if they can be viewed at the necessary spatial and temporal resolution from freely available images. This means satellites are most likely to be cost-effective in remote locations that are not readily accessible, and are more suitable for monitoring bigger species, like penguins, and those that leave obvious signs of presence, such as substantial areas of faecal staining (Fretwell & Trathan 2009).

Satellites are unlikely to facilitate monitoring of large numbers of small breeding seabird colonies.

Manned aircraft

Aerial seabird surveys are more commonly conducted with manned aircraft (Fig. 2) or UAVs, rather than satellites (Loarie *et al.* 2007, Rush *et al.* 2018). Compared to boats, manned aircraft afford a more cost-effective technique for surveying large areas of sea and inland bodies of water (Camphuysen *et al.* 2004). The shorter survey time of manned flights at high speed reduces the risk of double counting, which increases count-accuracy. However, this benefit may be negated by reduced time to detect and identify smaller or less abundant species, meaning that surveys of inshore seaducks rarely detect grebes, Common Goldeneye *Bucephala clangula*, or Black-throated Diver *Gavia arctica* (Joint Nature Conservation Committee 2010). As a result, land- and boat-based counts are often used alongside aerial surveys to ensure that birds are not missed (Joint Nature Conservation Committee 2010). This increases survey effort and thus time and money required for effective monitoring using manned aircraft.

Installing manned aircraft with cameras might reduce the need for accompanying land- and boat-based surveys. Photographs and videos provide a permanent record that can be used to identify additional individuals that surveyors might have missed (Hutchinson 1980). This is supported by a study in Carmarthen Bay (Wales, UK) which found that visual aerial surveys gave lower estimates of Common Scoter *Melanitta nigra* abundance compared to digital images and videos taken from an aeroplane (Buckland *et al.* 2012). Digital aerial surveys could also be used to count surface-nesting seabirds, including Arctic Skua Stercorarius parasiticus, terns, and Lesser Black-backed *Larus fuscus* and Great Black-backed Gull *Larus marinus*. Aerial surveys reduce habitat disturbance compared to traditional colony walk-through methods and also reduce disruption to nesting birds (Brisson-Curadeau *et al.* 2017, Rush *et al.* 2018). Taking digital photographs and videos from manned aircraft can further lower behavioural stress responses, as it allows the vehicle to be flown at higher altitude. This is because aerial surveyors must be close enough to the birds to allow accurate identification and counts, whereas images can be magnified during analysis (Thaxter & Burton 2009, Kemper *et al.* 2016).

The benefits of reduced disturbance and increased accuracy must be balanced against the high purchase and operation costs of manned aircraft (Hutchinson 1980). This includes the price of fuel, hiring a pilot with a professional aviation licence, and, if photographs or videos are desired, camera installation and hiring a camera operator (Wilhelm *et al.* 2015, Nowak *et al.* 2019). In the past, photographs were taken through windows using handheld cameras, while most studies today install fixed cameras to improve image quality and consistency (Hutchinson 1980, Wilhelm *et al.* 2015). Additionally, manned aircraft are restricted in where they can operate, as they require a nearby airport, fulfilment of aviation procedures and are not manoeuvrable over small areas (Nowak *et al.* 2019). Moreover, at sea aerial surveys are not advised in winds greater than Beaufort 4 to reduce the likelihood of inaccurate counts as can happen, for example, if white wave caps are confused with gulls (Thaxter & Burton 2009).

Monitoring seabirds using imagery from manned aircraft has some disadvantages. The high cost means that temporal resolution is typically low, so manned flight surveys are best deployed to obtain infrequent population counts, rather than time-series data such as breeding success (Anderson & Gaston 2013, Lyons *et al.* 2019). Manned aircraft are unlikely to reduce disturbance to surface-nesting seabirds that can be monitored from a single vantage point, although they could be a useful alternative to walk-through surveys. The benefits of using manned aircraft are therefore context and species dependent, but could appreciably benefit sensitive species and sites. Increasingly, many studies are now turning to UAVs for aerial monitoring, to overcome some of the challenges faced by manned aircraft (Anderson & Gaston 2013).

Unmanned Aerial Vehicles (UAVs)

The number of environmental biology papers using UAVs has increased markedly in the past 20 years, particularly since 2011 (Nowak *et al.* 2019). UAVs are known under a variety of terms, including: Unmanned Aerial Systems, Remotely Piloted Aircraft, and colloquially as ‘drones’. They are small, powered aerial vehicles that can be flown remotely or autonomously and can carry a payload, such as a camera (Fig. 3).

To date, UAV imagery has mainly been used for counting nests or individuals and has even identified 'new' populations (Nowak *et al.* 2019, Pfeifer *et al.* 2019). A wide range of seabirds have been monitored using UAVs, including penguins (Spheniscidae) (Hodgson *et al.* 2016, Borowicz *et al.* 2018, Korczak-Abshire *et al.* 2019, Pfeifer *et al.* 2019), albatrosses (Diomedeidae) (McClelland *et al.* 2016), terns and gulls (Laridae) (Sardà-Palomera *et al.* 2012, Grenzdörffer 2013, Chabot *et al.* 2015, Hodgson *et al.* 2016, Brisson-Curadeau *et al.* 2017, Rush *et al.* 2018), shags and cormorants (Phalacrocoracidae) (Irigoin-Lovera *et al.* 2019, Korczak-Abshire *et al.* 2019, Oosthuizen *et al.* 2020), auks (Alcidae) (Brisson-Curadeau *et al.* 2017), frigatebirds (Fregatidae) (Hodgson *et al.* 2016, Villegas *et al.* 2018), boobies (Sulidae) (Irigoin-Lovera *et al.* 2019), pelicans (Pelecanidae) (Irigoin-Lovera *et al.* 2019), and giant petrel species (*Macronectes spp.*) (Korczak-Abshire *et al.* 2019). Since most UAVs allow camera rotation, cliff-nesting seabirds can be readily surveyed – a feat more difficult to achieve with satellites (Brisson-Curadeau *et al.* 2017). However, unlike satellites and manned aircraft, distant sea surveys are limited as UAVs must typically remain in the line of the sight of the controller to satisfy flight regulations (Nowak *et al.* 2019).

Increasingly, UAVs are being used for monitoring purposes other than population or nest counts. For example, UAV surveys have recorded fine-scale foraging behaviour of terns in relation to wakes created by strong currents interacting with man-made structures (Lieber *et al.* 2019). In addition, UAVs might also collect time-series data, for example to measure nesting success. They create less disturbance than manned aircraft due to being smaller and less noisy and are cheaper to purchase and operate (Goebel *et al.* 2015). This means multiple flights throughout the breeding season are more feasible in terms of animal welfare and cost, and flight height can be lower, which increases spatial resolution and accuracy. For example, minimum flight height for at-sea surveys using manned aircraft is 450 m, whereas UAVs are regularly flown at < 100 m (Thaxter & Burton 2009). Nonetheless, UAVs can still disturb breeding seabirds. The behavioural response to UAV flight should be measured before studies to ensure that it does not exceed that of traditional field monitoring methods such as walk-through surveys.

The magnitude of behavioural response depends on the type of UAV; flight parameters, including altitude and speed; take-off location relative to the colony; and the species being monitored (Rümmler *et al.* 2016, Brisson-Curadeau *et al.* 2017, Mulero-Pázmány *et al.* 2017, Rush *et al.* 2018, Weimerskirch *et al.* 2018, Irigoin-Lovera *et al.* 2019). Rümmler *et al.* (2018) found that Adélie Penguins Pygoscelis adeliae reacted to a small octocopter UAV at the highest test altitude of 50 m, whereas Gentoo Penguins *Pygoscelis papua* only reacted below 30 m. In another study, Adélie Penguins did not respond to fixed-wing electric UAVs at 350 m altitude but did show vigilance and increased activity levels in response to UAVs flown at the same height but powered by a piston engine (Korczak-Abshire *et al.* 2016). Deciding on a suitable flight protocol to minimise disturbance is therefore difficult, as it will vary between and within species depending on a variety of factors. For example, there may be intra-species variation in response at different locations due to variable aerial predation levels or variation in response by the same colony at different times of the year. Consequently, it seems wise that test flights should always be conducted before using UAVs for seabird monitoring.

National flight regulations mean both a pilot and ground level observer are often required for UAV flights (Nowak *et al.* 2019). While this increases the cost of UAV studies, especially as pilots require training, it allows a dedicated ground level observer to focus on monitoring disturbance levels to ensure that flights are conducted safely. Legal restrictions also limit UAV flight parameters, including maximum altitude, speed, and use over reserves, which can affect the possibility of data acquisition (Nowak *et al.* 2019). This may be further limited by adverse weather conditions, as UAVs are more vulnerable to damage during aerial surveys than manned aircraft and satellites. For example, many small, lightweight UAVs, such as the Al-Multi (by Aerial Insight, Brandon, MB, Canada), cannot operate in precipitation and wind often reduces image quality due to camera movements during flight (Chabot *et al.* 2015, Goebel *et al.* 2015).

Overall, UAV-based monitoring is likely to be effective for measuring breeding success or counting nesting seabirds, provided disturbance is not greater than traditional monitoring methods. UAVs are particularly cost-effective if the window for fieldwork is short, and they can survey areas inaccessible by foot or vehicle, such as sea-stacks (Lyons *et al.* 2019, Oosthuizen *et al.* 2020). On the other hand, aerial surveys are not necessary for seabirds that can be viewed from a single vantage point (Table 1). Instead, time-lapse photography may be a better alternative to traditional point surveys than UAVs.

Fixed position cameras

Time-lapse cameras

Time-lapse photography records images at predetermined time-intervals regardless of subject presence (Cutler & Swann 1999). It has been used for avian studies since the technology first became commercially available, although its potential uses in ornithology are quickly increasing with advances in digital technology (Dodge & Snyder 1960, Green & Anderson 1961, Cowardin & Ashe 1965, Temple 1972, Weller & Derksen 1972, Harris 1982, Huffeldt & Merkel 2013). The increased availability of affordable cameras, requiring less frequent maintenance, reduced power consumption, and larger data storage capacity has seen the field of time-lapse photography rapidly expand in recent years (Bolton *et al.* 2007).

Time-lapse cameras are most appropriate for studying animals frequently present at a location, where a single vantage point gives a representative view of individuals, and the measurement of interest will not activate a motion-triggered camera (Cutler & Swann 1999, Black 2018). Species that aggregate at high densities at some point in their life-history, such as breeding seabirds, therefore represent ideal candidates for use (Fig. 4) (Black 2018). Time-lapse cameras are suited for collecting data as part of long-term studies, principally time-series data such as annual breeding success and phenology, and have a number of advantages over traditional field observations (Southwell & Emmerson 2015, Merkel *et al.* 2016, Black *et al.* 2018a, Hinke *et al.* 2018).

Firstly, time-constraints placed on fieldworkers and external conditions such as weather mean direct observations of nesting success are typically recorded less frequently than time-lapse photographs (Walsh *et al.* 1995). Most studies set cameras to record one image per hour and are only returned to once per year to change SD cards and batteries (Southwell & Emmerson 2015, Black *et al.* 2018a). This means that time-lapse photography can improve temporal resolution and data accuracy with reduced time investment.

High temporal resolution also makes time-lapse photography suitable for measuring numerous other parameters. This includes nest activity (such as nest attendance and division of labour between parents), re-sighting marked birds to determine adult survival and foraging behaviour, and population counts of breeding birds year-round, allowing insights into over-winter site attendance (Weller &

Derksen 1972, Mudge *et al.* 1987, Black *et al.* 2017, Black *et al.* 2018b, Pascalis *et al.* 2018). Additionally, time-lapse cameras can provide evidence of infrequent events not purposefully monitored (Harris 1982, Black *et al.* 2017, Black *et al.* 2018b). For example, time-lapse photographs have recorded Black-legged Kittiwake *Rissa tridactyla* chick predation by a Peregrine Falcon *Falco peregrinus* (Collins *et al.* 2014). Predation may be under-recorded by fieldworkers and aerial surveys, as both human and aircraft presence could deter predatory activity. Similarly, cameras might capture adult seabirds carrying prey, which could give information on chick diet.

As well as their diverse range of uses, time-lapse cameras are unlikely to adversely affect the wildlife they monitor, provided that they are installed and maintained outside the breeding season and are located at a safe distance from breeding birds (Merkel *et al.* 2016). Determining a ‘safe’ distance is difficult, but the distance kept by fieldworkers could be a provisional minimum (Joint Nature Conservation Committee 2016). Limited disturbance also means cameras can collect data regardless of abiotic conditions. For example, the UK ‘Seabird Count’ instructs surveyors to avoid visiting colonies in winds stronger than Beaufort 4 or during heavy and continuous rain, as disturbance during wet weather can leave eggs and chicks vulnerable to chilling, and weather conditions can affect colony attendance so that this strict protocol helps to ensure count comparability across years and colonies (Joint Nature Conservation Committee 2016).

The infrequency of human visits (i.e., yearly maintenance) allows time-lapse cameras to capture images in locations and at scales otherwise unfeasible in terms of time, money, and human capabilities, such as in harsh conditions and remote places (Weller & Derksen 1972, Black *et al.* 2017, Black 2018, Black *et al.* 2018a, Black *et al.* 2018b , Pascalis *et al.* 2018). Already, extensive camera networks in the Antarctic have provided data on previously unmonitored penguin colonies (Southwell & Emmerson 2015). Nonetheless, maintaining camera networks is expensive, and if only visited once annually, a large amount of data could be lost from mechanical failure between visits. Camera set-up is also a crucial consideration to ensure useful and reliable data is obtained (Lorentzen *et al.* 2010). Increasing the distance between camera and colony will increase the number of birds viewed per frame but will lower image resolution. A study on pygoscelid penguins suggested approximately 20 nests could be reliably monitored for the duration of the breeding season, but this depended on nest density and topography (Hinke *et al.* 2018). The optimal camera angle and horizontal and vertical distance from the colony will therefore be specific to location, study species and study purpose (Lorentzen *et al.* 2010). A summary of the advantages and disadvantages of time-lapse photography as a tool for monitoring seabirds is given below (Table 2).

Video cameras

Videography is similar to time-lapse photography, except that observations are recorded continuously. It may be preferable when constant field measurements are required, as time-lapse cameras might miss an event that occurred between photographs and results would not be comparable with field observations, introducing bias into long-term studies. Examples include recording incubation behaviour, thermoregulatory responses and rate of adult provisioning (Frederiksen *et al.* 2019, Cook *et al*. 2020, Williams & DeLeon 2020).

Frederiksen *et al.* (2019) used video surveillance to measure chick feeding rates of Little Auk *Alle alle* in north-east Greenland. Traditional methods required 12- or 24-hour surveillance in the field, which is time-consuming, physically demanding, and renders results liable to error from observer fatigue, even when monitoring is conducted in shifts (Harding *et al.* 2007, Mosbech *et al.* 2017). Although videos take a long time to analyse manually, the ability to increase playback speed means that periods of inactivity can be watched quickly, while important events can be slowed down, re-wound, and re-watched innumerable times, to ensure that accurate records are made. Moreover, processing can take place independent of external abiotic conditions that inhibit direct observations in the field. Having said this, poor weather can reduce image quality, meaning that neither video and nor time-lapse cameras deliver useable data in all conditions.

Unfortunately, the large amount of data recorded by video cameras per unit time means that SD cards and batteries must be replaced regularly, often daily (Mosbech *et al.* 2017, Frederiksen *et al.* 2019). This makes continuous videography only suitable in locations readily accessible by humans. For most studies requiring data with high temporal resolution, time-lapse cameras are a more cost-effective option.

Motion-triggered cameras

For studies where measurements do not need to be made at regular intervals or continuously, motion-triggered cameras are an alternative to static time-lapse or video cameras. Movement in front of the sensor triggers photographs or a short video sequence to be recorded, allowing capture of individual, instantaneous events (Black 2018). So far, motion-triggered cameras have been most frequently deployed in seabird research to examine the impact of nest predation on breeding success (Hervías *et al.* 2013, Thiebot *et al.* 2014, Davies *et al.* 2015, Ekanayake *et al.* 2015, Luna *et al.* 2018, Whelan *et al.* 2018, Stolpmann *et al.* 2019). In some cases, this has provided support for removal of introduced predators at seabird colonies (Davies *et al.* 2015). Motion-triggered cameras are likely to be more effective at monitoring predation than time-lapse cameras, as the camera should be triggered whenever a predator enters the field of view, rather than at specific time points. Another use of motion-triggered cameras has been to understand nesting seabird behaviours, such as incubation and foraging patterns. They can record the time at which parents exchange incubation duty or when one parent returns from a foraging trip to feed the young (Hart *et al.* 2016, Mendez *et al.* 2017). This could allow assessment of seabird diet for species that load prey in their bills.

One of the difficulties of deploying motion-triggered cameras is to prevent irrelevant motion in the surrounding environment causing false triggers. This is often due to vegetation moving in the wind, and while some vegetation could be removed from the camera's zone of detection, the environment should ideally be altered as little as possible (Van Berkel 2014). Alternatively, positioning cameras closer to the object of interest, such as a seabird nest, can reduce false triggers but severely limits spatial coverage (Van Berkel 2014). Each camera might therefore only view one or two nests. This greatly increases the number of cameras required, and thus cost, if many nests need to be monitored.

Handheld cameras

Another form of digital photography that has been used to investigate seabird diet is the handheld camera (Table 3). Although time-lapse and motion-triggered cameras may capture seabirds with prey, purposefully taking photographs of prey-carrying seabirds can record the diet of a greater number of individuals, given that handheld cameras do not have a fixed field of view. Traditional techniques to investigate seabird diet predominantly focus on morphological analysis and include visual identification of prey species and size in the field, as well as mist-netting adults to obtain whole prey or regurgitates or collecting regurgitates from chicks, either from the ground or using ligatures (Votier *et al.* 2003, Barrett *et al.* 2007, Forys & Havesh 2017, Gaglio *et al.* 2017). More recently, the use of molecular and biochemical techniques in diet studies has dramatically increased, particularly DNA and stable isotope analysis of blood and faecal samples (Horswill *et al.* 2018). Each method of diet analysis has its own advantages and limitations. In general, sample collection has the obvious disadvantage of disturbing birds, while direct observation is more likely to result in incorrect identification, especially when trying to estimate prey size in the field. Conversely, taking photographs of adult seabirds carrying prey is non-invasive provided that a safe distance is kept between bird and photographer. It also produces a permanent record for checking identification of species and size and is more likely to capture the entire prey item. For example, terns often only regurgitate the posterior body and caudal fin, making identification of similar species challenging (McLeay *et al.* 2009).

Gaglio *et al.* (2017) showed that photo-sampling produced similar estimates of Greater Crested Tern *Thalasseus bergii* prey composition and size compared to regurgitations, and at a faster species accumulation rate. Over three breeding seasons they were able to double the known diversity of prey taken by two Great Crested Tern colonies. Likewise, photo-sampling increased the known number of fish species fed to Black Skimmer Rynchops niger chicks by 29 % (Forys & Havesh 2017).

Handheld cameras could allow seabird diet to be monitored at greater scales than before, as photographs can be accumulated faster than prey samples. There is already a vast wildlife photography community capturing seabirds with prey, offering a rich source of diet data. This use of citizen science was recognised by Forys & Havesh (2017), who used Facebook and Flickr to ask for photographs of Black Skimmer adults carrying prey. From 211 photographs, they conducted a small study of chick diet during the 2015-2016 breeding season. At a much larger scale, the RSPB's Project Puffin UK is currently requesting photographs of Atlantic Puffins *Fratercula arctica* carrying prey from any year, to better understand spatial and temporal variation in diet (Fig. 5) (RSPB 2020).

Of course, using photography to investigate diet is only feasible for seabirds that carry prey in their bills. Moreover, photographs take time to process and strict protocols are required to minimise bias. For example, birds should be photographed at random, rather than focussing on individuals carrying large, interesting or multiple prey items. For studies using citizen science, it can be hard to ensure that protocols have been followed, especially when mining existing databases, and so quantification of suspected biases in method is essential. Project Puffin UK suspected that puffins carrying large prey were more likely to be spotted and photographed than those carrying small prey. To quantify this potential size-bias, a researcher took photographs of any puffin approaching a colony on the Farne Islands (England, UK) regardless of whether it appeared to have prey. Photographs containing prey will be compared with images taken by members of the public at the same location and in the same year to quantify any size-bias (E. Owen pers. comm.).

Animal-borne cameras

The final type of digital imaging device to consider are animal-borne cameras. Originally, the large size of these devices limited their deployment to mammals and captive and/or flightless birds (Watanuki *et al.* 2008). However, gradual miniaturisation of the technology has since allowed use on unhabituated, free-ranging seabirds (Moll *et al.* 2007, Watanuki *et al.* 2008). Bird-borne cameras can record still images or videos and are unique in that they provide observations from the perspective of the animal (Moll *et al.* 2007, Tremblay *et al.* 2014). This makes them particularly well suited for understanding fine-scale interactions between seabirds and their environment (Moll *et al.* 2007). Cameras on seabirds have been particularly useful for providing insight into foraging behaviours. This includes foraging habitat selection (Watanuki *et al.* 2008); movement patterns (Ponganis *et al.* 2000, Tremblay *et al.* 2014); and interactions with prey (Grémillet *et al.* 2006, Handley *et al.* 2016, Handley *et al.* 2018), fisheries (Votier *et al.* 2013), conspecifics (Takahashi *et al.* 2004, Yoda *et al.* 2011) and other predator species during foraging (Sakamoto *et al.* 2009, Yoda *et al.* 2011, Thiebault *et al.* 2014).

One of the main limitations of animal-borne cameras is system lifespan (Moll *et al.* 2007). The ethical requirement for minimised camera size limits battery capacity and means that recording duration is often under two hours, especially for continuous video recordings (Grémillet *et al.* 2006, Moll *et al.* 2007, Hooker *et al.* 2008, Yoda *et al.* 2011, Thiebault *et al.* 2014, Tremblay *et al.* 2014, Handley *et al.* 2018). Battery power is a greater limitation than data storage capacity for cameras connected to a transmitter, because the data can be relayed to a remote downloading station (Moll *et al.* 2007, Hays 2015). Downloading data remotely is advantageous, because it means data is not lost if the device cannot be retrieved. However, constraints on bandwidth available through data relay platforms, such as the Argos service, can again limit the duration of camera deployment (Hays 2015).

Recapturing birds to recover cameras can be challenging and frequently restricts studies to breeding adults that can be re-caught on the nest (Watanuki *et al.* 2008, Sakamoto *et al.* 2009, Votier *et al.* 2013, Tremblay *et al.* 2014). Equally, the difficulty of recapture, ethical implications of handling and attaching devices to birds, and high cost of each device, means most studies only deploy cameras on a small number of individuals, commonly < 10 (Ponganis *et al.* 2000, Takahashi *et al.* 2004, Grémillet *et al.* 2006, Moll *et al.* 2007, Bluff & Rutz 2008, Watanuki *et al.* 2008, Sakamoto *et al.* 2009, Yoda *et al.* 2011, Bicknell *et al.* 2016). Small sample size can sacrifice robust population-level inferences, although the ability to collect novel data from the field of view of the seabird should not be overlooked (Hebblewhite & Haydon 2010).

Night-vision

One advantage of using any form of digital imaging technology for seabird monitoring is the improved ability to make observations at night using infra-red illumination and/or thermography. Infra-red illumination allows cameras to take photographs and videos in the dark, by shining infra-red light on the area of interest. This reduces disturbance to burrow nesting seabirds and seabirds being monitored at night compared to visible light flash photography, as infra-red wavelengths are invisible to birds and mammals (Perkins *et al*. 2018). Collins *et al.* (2014) were able to observe night-time predation of Black-legged Kittiwake nests on Puffin Island (Wales, UK) from infra-red images captured by a Ltl-Acorn 5210MC time-lapse camera. Conversely, infra-red thermography (thermal imaging) does not itself emit light, but instead detects infra-red radiation (heat) emitted by animals (McCafferty 2013). It is often used to detect and count nesting sites, with Israel & Reinhard (2017) using a UAV-borne thermal camera to detect camouflaged Northern Lapwing *Vanellus*

vanellus nests. This has potential for locating inconspicuous nests of surface-nesting seabirds, like gulls and terns.

Limitations of digital image collection

In summary, digital imaging technology has the potential to increase accuracy, cost-effectiveness and scale of seabird data collection, while reducing disturbance to breeding birds. Nevertheless, it is not a 'silver bullet' solution. Different technologies have different uses and some species cannot easily be monitored using digital imagery, like burrow-nesting seabirds. Perkins *et al.* (2018) concluded that infra-red filming was a costly and inefficient method for counting European Storm Petrels *Hydrobates pelagicus* relative to tape playback, due to the large amount of expensive equipment and reviewing time needed. It would only be beneficial at sites that cannot otherwise be surveyed safely or where disturbance is a concern.

One of the main trade-offs for most digital imaging technologies is between cost and image resolution, which affects how well the object(s) of interest can be identified in photographs and videos. Image resolution is clearly affected by the choice of camera, including the number of pixels and optical quality of the lens. However, it is also influenced by factors specific to the image capture method. For example, reducing the flight speed of manned aircraft will improve video quality but increase flight time, and the latter increases costs of fuel and pilot hire (Mellor *et al.* 2007). Conversely, faster speeds can be achieved with less reduction in quality if higher frame rates are used or the number of pixels is increased, both of which increase camera cost (Mellor *et al.* 2007). Increasing the depth of frame from 1000 pixels to 2500 pixels means a bird would stay in frame for the same amount of time at double the flight speed, or alternatively stay in frame for over twice as long at a given speed (Mellor *et al.* 2007). Dealing with this trade-off between cost and image resolution is difficult when funding is limited for wildlife monitoring (Waldron *et al.* 2013). Users must remember that image quality should be 'good enough' to provide data of equal or better accuracy than traditional non-digital methods, but it does not need to be 'exceptional'. Selecting an affordable method that will provide imagery of sufficient quality for the monitoring purpose is therefore all that can be recommended.

DATA PROCESSING AND ANALYSIS

For seabird species where digital photography and videography could aid data collection, it is important to consider current data processing and analysis methods. These must be feasible, in terms of time and money, and provide accurate data for the technology to be of value to seabird monitoring. So far, manual methods have been deployed most commonly, although rapid advances in semi-automated and automated information extraction are revealing that digital imagery can be a powerful and cost-effective monitoring technique.

Manual image analysis

Manual image analysis requires researchers to examine photographs individually and make the appropriate measurement, such as count the number of each species present or record re-sighted birds. If multiple images have been collected over time, then parameters such as breeding success and phenology can be calculated. For photographs taken by UAVs and manned aircraft, the images must be orthorectified prior to analysis to produce an orthomosaic (mosaic image with positional accuracy) using software like Agisoft Photoscan (Rush *et al.* 2018).

Accuracy

One of the most important considerations to make when deciding whether to manually analyse images is accuracy. An ‘accurate’ estimate can be defined as one that is close to the true value, for example the true population count (Gregory *et al.* 2004, Hodgson *et al.* 2016). The accuracy of manual counts firstly depends on the researchers’ intrinsic ability to correctly identify and count individuals in an image. This can be termed ‘count-accuracy’. Secondly, it depends on the image itself, and whether it has captured all the individuals of interest, for example all active nests on the section of cliff being examined. This is ‘image-accuracy’.

To increase count-accuracy by reducing misidentification of birds and counting errors most studies have used counting tools. Users click on a bird to mark it, and the computer programme automatically sums the number of marks to give a total count per image. Software commonly used includes ImageJ (Merkel *et al.* 2016, Hurford 2017, Hodgson *et al.* 2018), Adobe Photoshop’s count tool (Chabot *et al.*

2015, Goebel *et al.* 2015, Hodgson *et al.* 2016, Sinclair *et al.* 2017) and GIS environments (Sardà-Palomera *et al.* 2012, Lyons *et al.* 2019). ImageJ and QGIS are free, whereas users must pay for ArcGIS and Adobe Photoshop. Some researchers have built their own purpose designed annotation software, such as 'Penguin Nest Picture Analyser' in Java (Southwell & Emmerson 2015) and the *Penguin Watch* interface on Zooniverse (Black *et al.* 2017, Jones *et al.* 2018, Jones *et al.* 2020). Overall, the availability of free, easy-to-use counting tools means that researchers should not be limited by software in their ability to manually analyse digital images. To further assist with manual counting, several studies have overlaid grid cells on photographs and then made systematic, cell-specific counts (Hodgson *et al.* 2016, Korczak-Abshire *et al.* 2019). Count-accuracy can also be increased by brightening dull photographs (Sinclair *et al.* 2017).

It is not possible to directly assess count-accuracy, unless the true image count is known. Instead, precision within counts of the same and different observers should be calculated (Sinclair *et al.* 2017). This means calculating the variance and/or standard deviation between replicated counts by the same and different counters attempting to count the same sample (Gregory *et al.* 2004, Hodgson *et al.* 2016, Sinclair *et al.* 2017, Korczak-Abshire *et al.* 2019). Unfortunately, this increases the time required for an already laborious task, so it has not become common practice.

As with count-accuracy, it is not possible to assess image-accuracy, unless the true count in the wild is known. Nevertheless, comparison between traditional and digital photography methods can be informative. If results from traditional monitoring and digital image analysis do not significantly differ, then digital photography is at least 'as accurate as' traditional techniques. For example, no significant difference was found between ground and UAV-derived counts of penguins in Antarctica and terns in Australia, suggesting UAVs were suitable for these population counts (Goebel *et al.* 2015, Hodgson *et al.* 2016). Equally, a significant correlation between direct and time-lapse photography measurements of penguin breeding success in Antarctica, supports the use of time-lapse cameras for measuring nesting success (Southwell & Emmerson 2015, Hinke *et al.* 2018). It is important that different researchers conduct ground surveys and image analysis to allow valid comparison of methods (Goebel *et al.* 2015).

Alternatively, if there is a statistically significant difference between traditional and digital image derived results, then interpreting the accuracy of digital photography is more complicated. Further analysis of the data is required to assess whether traditional or digital methods are more accurate. For example, counts of Common Tern Sterna hirundo from UAV-derived images were 93-94 % of traditional ground counts in North America (Chabot *et al.* 2015). UAV-derived counts were presumed to be less accurate, due to variable visibility of birds with ground cover, weather conditions and image quality. Conversely, UAV-derived counts of penguins and frigatebirds in Australia were significantly larger than ground counts. The authors suggested the downward-facing perspective of UAV images reduced the number of birds missed by topography and other birds obscuring the counters' line of sight in ground surveys (Hodgson *et al.* 2016). Such problems are likely to be species and habitat specific, thus reinforcing that assessment of accuracy should be made on a case-by-case basis.

Depending on the parameters being measured, it may not be possible to perform statistical analyses with small sample size. Southwell and Emmerson (2015) found that the first date of Adélie Penguin arrival was zero to two days later in time-lapse images compared to direct observation, over eight years; and the first egg was seen two to six days later in camera images, over two years. Later detection of first arrival was expected given the cameras' restricted spatial coverage compared to direct observers, and first egg detection was limited by temporal resolution. Incubating parents huddle tightly on the egg and reliable detection requires near-continuous observation (Southwell & Emmerson 2015). But, should these small differences in dates prevent time-lapse cameras being used to measure penguin phenology? The answer will largely depend on the individual situation. Do the other advantages of time-lapse cameras compared to direct observation, outweigh the costs of marginally different phenology measurements?

Moreover, in some locations, monitoring has only occurred with digital photography, making comparison to traditional methods impossible. This is typical of remote locations at high latitude with harsh environmental conditions, and highlights how digital imaging technology can greatly increase the scale of monitoring (Black *et al.* 2017, 2018a, b, Korczak-Abshire *et al.* 2019). For these studies, it is particularly important to calculate the variance of intra- and inter-observer counts of the same image to ensure high count-accuracy.

Cost

One of the main disadvantages of manual image analysis is the time required. This has likely prevented the wide-scale use of digital imaging methods such as time-lapse photography to date, as the volume of raw imagery collected can quickly exceed researchers' processing capabilities (Pascalis *et al.* 2018). To date, most studies have monitored only a single colony of interest (Southwell & Emmerson 2015, Black *et al.* 2017). The time required per image depends on the number of birds per photograph, image quality and experience of the analyser, although this can be decreased using a variety of methods. Sinclair *et al.* (2017) assessed how manual counts of Common Guillemot *Uria aalge* were affected if only one-quarter of the original image was counted. They found that counts from all quarters of an image were significantly correlated, meaning only the top-right hand corner needed to be sampled. This reduced post-processing from seven to three minutes per photo. However, this method is only possible when seabirds are evenly distributed across the image.

Another, and increasingly common method to reduce the time researchers spend processing images, at little extra cost, is to engage volunteer citizen scientists. Two projects currently advocating citizen science for seabird monitoring are *Penguin Watch* and S*eabird Watch* on the Zooniverse platform (https://www.zooniverse.org). Time-lapse photographs are uploaded onto the platform, and volunteers click on birds to classify them as either adult or juvenile penguins (*Penguin Watch*), Black-legged Kittiwakes or guillemots (*Seabird Watch*) (Fig. 6). Each image is shown to four participants and if no animals are identified or the image is too dark/blurry to classify, the image is retired from the active dataset and not seen by further volunteers. If any of the four participants identifies an animal, then the image is shown to an additional six people before being retired (Jones *et al.* 2018). Having multiple people view each image increases data reliability and a field guide is available to aid bird identification and increase accuracy. For *Penguin Watch*, comparison between annotations made by citizen scientists and 'gold standard' researchers, has validated the use of citizen science for identifying penguins in time-lapse photographs (Jones *et al.* 2018). This process is currently being undertaken for *Seabird Watch*, as well as a comparison between results from field observations and 'gold standard' researcher analysed images (A. Edney unpubl. data).

Although citizen scientists reduce researcher post-processing time, the total amount of time for images to be analysed is often much longer. Volunteers cannot be given strict deadlines like paid researchers, meaning a large number of volunteers are needed for images to be analysed quickly. There are also concerns that an increasing number of citizen science projects will effectively ‘flood the market’, resulting in fewer participants per project. The most effective way to increase cost-efficiency of digital image analysis is to develop semi-automated and automated techniques.

Semi-automated image analysis

Semi-automated classification is a form of supervised classification. It is user-driven and cannot identify and count birds without human guidance (Fretwell *et al.* 2012, Rush *et al.* 2018). Most semi-automated classification involves finding a unique spectral signature for the object in question (e.g., the head of a gull) that can be used to identify all occurrences of this object in the image (Schwaller *et al.* 1989, Fretwell *et al.* 2012, Grenzdörffer 2013, Waluda *et al.,* 2014, Hodgson *et al.* 2018).

Rush *et al.* (2018) offer a comprehensive description of one approach to semi-automated classification of nesting Lesser Black-backed Gulls counts from UAV images. In brief, the training sample manager tool in ArcGIS identified different spectral signatures of three species of gull and surrounding habitat features. The maximum likelihood tool performed supervised classification and identified the gull species in each image. A shapefile, with the outlines of objects identified as birds, was overlaid on every original image for manual editing. This process was fast to complete and involved systematically scanning the image and confirming if objects in the shapefile were indeed birds. Non-bird objects were deleted. The number of Lesser Black-backed Gulls from semiautomated classification had a mean agreement of 104 % with manual counts, due to some non-bird objects being incorrectly identified as gulls. Agreement was reduced to 98 % via manual editing. This demonstrates that semi-automated classification of UAV images can provide accurate counts of a surface-nesting seabird with minimal disturbance.

Semi-automated classification would be especially useful for classifying birds in time-lapse photographs, as the sheer number of raw images collected can make manual classification unfeasible. While it may be difficult for species that do not have good contrast with their surroundings, such as shags and cormorants on dark rocks, initial spectral analysis can quickly determine this (Grenzdörffer 2013, Lyons *et al.* 2019). It is also important to remember that human vision is limited to visible light, so different objects that appear the same colour to the human eye, might still have a unique spectral signature that allows them to be separated.

Automated image analysis

Automated image analysis is a rapidly developing field that has the potential to vastly increase the scale of seabird monitoring. Automatic cell counting is frequently performed by cell biologists in ImageJ, using the 'Automatic cell counter' tool, but its transferability to seabird monitoring is limited due to the complexity of seabird colonies (Grishagin 2015).

ImageJ's automated counter cannot differentiate between species and is most accurate when birds occur against a plain background (Hurford 2017). It is liable to underestimate the true count, due to overlapping birds being counted as one object, while birds with strongly contrasting plumage patterns may be overestimated (Hurford 2017). The high nest density of cliff-breeding species and the complex background created by the natural environment, mean automated counts in ImageJ are unlikely to be accurate for most seabird colonies. Nest density, terrain, and vegetation should be carefully considered when deciding on seabird colonies suitable for automated image analysis (Hinke *et al.* 2018).

Recently, more studies are developing machine learning algorithms to identify birds in images, including those obtained from videos (Williams & DeLeon 2020). One example is the *Pengbot* algorithm, developed by the *Penguin Watch* team, to automatically identify and count penguins in time-lapse photographs (Jones *et al.* 2020). A similar tool is in the process of being developed for *Seabird Watch* (T. Hart, pers. comm.). *Pengbot* uses a Convolutional Neural Network (CNN) to estimate an object (penguin) density map, from which the number of objects (penguins) can be obtained. Training the algorithm to recognise penguins and then testing it required in the order of 82 000 pre-annotated images, which were provided by citizen scientists via *Penguin Watch* (Arteta *et al.* 2016). Without citizen science, labelling photographs is expensive in terms of time and money, especially if professionals are paid to do so via micropayment sites like Amazon Mechanical Turk (Arteta *et al.* 2016, Wang *et al.* 2019).

Consequently, whilst automated image analysis can be cost-effective once machine learning algorithms are up and running, it is important to remember the effort that goes into their development. For small-scale studies on a single species, manual or semi-automated analysis may be more achievable. Nevertheless, automated analysis of time-lapse photographs could monitor species at very large scales, by installing time-lapse cameras across their range.

CONCLUSIONS

We have critically assessed the use of a wide range of digital imaging methods for seabird monitoring, both from a data collection and data analysis perspective. All types of digital photography and videography create a permanent record of observations that can be validated and re-analysed. Many offer a cost-effective means of overcoming challenges associated with ‘traditional’ methods for measuring specific demographic and behavioural parameters. The extent to which digital imaging methods are incorporated into seabird monitoring over the coming years, will largely depend on advances in automated image analysis.

This leaves researchers to consider whether digital imaging technology could facilitate and/or replace their traditional monitoring techniques. There is a trade-off between potentially increased accuracy, cost-effectiveness, and reduced disturbance, with reduced consistency in long-term studies. Long-term research conducted in the same way for many years needs to consider the risk of new methods biasing results. Where digital imaging could replace traditional methods, researchers must also consider the transition time required: how long should traditional and new methods be run in parallel before traditional methods are phased out? Decisions like this must be made on a case-by-case basis. Despite these unanswered questions, digital imaging technology has the potential to greatly assist seabird monitoring in a research environment with increasingly limited time and funding for conservation (Waldron *et al.* 2013).

We would like to thank Sarah Money and Peter Fretwell for providing photographs to include in the paper. We would also like to thank the University of Gloucestershire’s Environmental Dynamics and Governance RPA for awarding AJE a bursary to undertake an MSc by Research. Two anonymous reviewers provided helpful feedback on the manuscript.

DATA AVAILABILITY STATEMENT

This work has no associated data.

REFERENCES

Ancel, A., Cristofari, R., Trathan, P.N., Gilbert, C., Fretwell, P.T. & Beaulieu, M. 2017. Looking for new emperor penguin colonies? Filling the gaps. *Global Ecology and Conservation* **9**: 171–179.

Anderson, K. & Gaston, K.J. 2013. Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Frontiers in Ecology and the Environment* **11**: 138–146.

Anker-Nilssen, T., Erikstad, K.E. & Lorentsen, S.-H. 1996. Aims and effort in seabird monitoring: an assessment based on Norwegian data. *wbio* **2**: 17–26.

Arteta, C., Lempitsky, V. & Zisserman, A. 2016. Counting in the wild. *Eur. Conf. Comput. Vis. LNCS.*, 9911: 483–498.

Barrett, R.T., Camphuysen, K. (C J.), Anker-Nilssen, T., Chardine, J.W., Furness, R.W., Garthe, S., Hüppop, O., Leopold, M.F., Montevecchi, W.A. & Veit, R.R. 2007. Diet studies of seabirds: a review and recommendations. *ICES J Mar Sci* **64**: 1675–1691.

Bibby, C.J., Burgess, N.D., Hill, D.A. & Mustoe, S. 2000. *Bird Census Techniques*. London: Elsevier.

Bicknell, A.W., Godley, B.J., Sheehan, E.V., Votier, S.C. & Witt, M.J. 2016. Camera technology for monitoring marine biodiversity and human impact. *Frontiers in Ecology and the Environment* **14**: 424–432.

Black, C. 2018. Spying on seabirds: a review of time-lapse photography capabilities and limitations. *Seabird* **31**:1-14.

Black, C., Collen, B., Lunn, D., Filby, D., Winnard, S. & Hart, T. 2018. Time-lapse cameras reveal latitude and season influence breeding phenology durations in penguins. *Ecology and Evolution* **8**: 8286–8296.

Black, C., Rey, A.R. & Hart, T. 2017. Peeking into the bleak midwinter: Investigating nonbreeding strategies of Gentoo Penguins using a camera network. *The Auk* **134**: 520–529.

Black, C., Southwell, C., Emmerson, L., Lunn, D. & Hart, T. 2018. Time-lapse imagery of Adélie penguins reveals differential winter strategies and breeding site occupation. *PLOS ONE* **13**: e0193532.

Bluff, L.A. & Rutz, C. 2008. A quick guide to video-tracking birds. *Biol Lett* **4**: 319–322.

Bolton, M., Butcher, N., Sharpe, F., Stevens, D. & Fisher, G. 2007. Remote monitoring of nests using digital camera technology. *Journal of Field Ornithology* **78**: 213–220.

Borowicz, A., McDowall, P., Youngflesh, C., Sayre-McCord, T., Clucas, G., Herman, R., Forrest, S., Rider, M., Schwaller, M., Hart, T., Jenouvrier, S., Polito, M.J., Singh, H. & Lynch, H.J. 2018. Multi-modal survey of Adélie penguin mega-colonies reveals the Danger Islands as a seabird hotspot. *Scientific Reports* **8**: 1–9.

Brisson-Curadeau, É., Bird, D., Burke, C., Fifield, D.A., Pace, P., Sherley, R.B. & Elliott, K.H. 2017. Seabird species vary in behavioural response to drone census. *Scientific Reports* **7**: 1–9.

Buckland, S.T., Burt, M.L., Rexstad, E.A., Mellor, M., Williams, A.E. & Woodward, R. 2012. Aerial surveys of seabirds: the advent of digital methods. *Journal of Applied Ecology* **49**: 960–967.

Camphuysen, K.J., Fox, A.D., Leopold, M.F. & Petersen, I.K. 2004. *Towards standardised seabirds at sea census techniques in connection with environmental impact assessments for offshore wind farms in the UK: a comparison of ship and aerial sampling methods for marine birds, and their applicability to offshore wind farm assessments.* NIOZ Report Commissioned by COWRIE Ltd. Texel, Netherlands: NIOZ.

Chabot, D., Craik, S.R. & Bird, D.M. 2015. Population Census of a Large Common Tern Colony with a Small Unmanned Aircraft. *PLOS ONE* **10**: e0122588.

Collins, P.M., Green, J.A., Dodd, S., Shaw, P.J.A. & Halsey, L.G. 2014. Predation of Black-legged Kittiwake Chicks *Rissa tridactyla* by a Peregrine Falcon *Falco peregrinus*: Insights from Time-lapse Cameras. *The Wilson Journal of Ornithology* **126**: 158–161.

Cook, T.R., Martin, R., Roberts, J., Häkkinen, H., Botha, P., Meyer, C., Sparks, E., Underhill, L.G., Ryan, P.G. & Sherley, R.B. 2020. Parenting in a warming world: thermoregulatory responses to heat stress in an endangered seabird. *Conservation Physiology* **8**: coz109.

Cowardin, L.M. & Ashe, J.E. 1965. An Automatic Camera Device for Measuring Waterfowl Use. *The Journal of Wildlife Management* **29**: 636–640.

Croxall, J.P., Butchart, S.H.M., Lascelles, B., Stattersfield, A.J., Sullivan, B., Symes, A. & Taylor, P. 2012. Seabird conservation status, threats and priority actions: a global assessment. *Bird Conservation International* **22**: 1–34.

Cutler, T.L. & Swann, D.E. 1999. Using Remote Photography in Wildlife Ecology: A Review. *Wildlife Society Bulletin (1973-2006)* **27**: 571–581.

Davies, D., Dilley, B., Bond, A., Cuthbert, R. & Ryan, P. 2015. Trends and tactics of mouse predation on Tristan Albatross *Diomedea dabbenena* chicks at Gough Island, South Atlantic Ocean. *Avian Conservation and Ecology* **10**.

Dodge, W.E. & Snyder, D.P. 1960. An Automatic Camera Device for Recording Wildlife Activity. *The Journal of Wildlife Management* **24**: 340–342.

Dolliver, J.E. 2019. Using Satellite Imagery to Count Nesting Albatross from Space. Master's thesis, Oregon State University, Corvallis, OR.

Ekanayake, K.B., Sutherland, D.R., Dann, P. & Weston, M.A. 2015. Out of sight but not out of mind: corvids prey extensively on eggs of burrow-nesting penguins. *Wildl. Res.* **42**: 509–517.

Evans, P.G.H. 1986. Monitoring Seabirds in the North Atlantic. In: *Mediterranean Marine Avifauna* (X. Monbailliu, ed), pp. 179–206. Springer, Berlin, Heidelberg.

Field, S.A., Tyre, A.J. & Possingham, H.P. 2005. Optimizing Allocation of Monitoring Effort Under Economic and Observational Constraints. *The Journal of Wildlife Management* **69**: 473–482.

Forys, E.A. & Hevesh, A.R. 2017. Investigating Black Skimmer Chick Diets Using Citizen Science and Digital Photography. *sena* **16**: 317–325.

Frederiksen, M., Mosbech, A., Andersson, A.W., Castro, A.C., Egevang, C., Fort, J., Grémillet, D., Linnebjerg, J., Lyngs, P., Haaning Nielsen, H. & Rømer, J.K. 2019. *POPULATION SIZE AND HABITAT USE OF BREEDING SEABIRDS IN NORTHEAST GREENLAND. Field studies 2017-2018. Danish Centre for Environment and Energy Scientific Report 337.* Roskilde, Denmark: Aarhus University.

Fretwell, P.T. & Trathan, P.N. 2009. Penguins from space: faecal stains reveal the location of emperor penguin colonies. *Global Ecology and Biogeography* **18**: 543–552.

Fretwell, P.T., LaRue, M.A., Morin, P., Kooyman, G.L., Wienecke, B., Ratcliffe, N., Fox, A.J., Fleming, A.H., Porter, C. & Trathan, P.N. 2012. An Emperor Penguin Population Estimate: The First Global, Synoptic Survey of a Species from Space. *PLoS ONE* **7**: e33751.

Fretwell, P.T., Scofield, P. & Phillips, R.A. 2017. Using super-high resolution satellite imagery to census threatened albatrosses. *Ibis* **159**: 481–490.

Fretwell, P.T., Trathan, P.N., Wienecke, B. & Kooyman, G.L. 2014. Emperor Penguins Breeding on Iceshelves. *PLoS ONE* **9**: e85285.

Gaglio, D., Cook, T.R., Connan, M., Ryan, P.G. & Sherley, R.B. 2017. Dietary studies in birds: testing a non-invasive method using digital photography in seabirds. *Methods in Ecology and Evolution* **8**: 214–222.

Goebel, M.E., Perryman, W.L., Hinke, J.T., Krause, D.J., Hann, N.A., Gardner, S. & LeRoi, D.J. 2015. A small unmanned aerial system for estimating abundance and size of Antarctic predators. *Polar Biol* **38**: 619–630.

Green, G.W. & Anderson, D.C. 1961. A Simple and Inexpensive Apparatus for Photographing Events at Pre-set Intervals1. *The Canadian Entomologist* **93**: 741–745.

Gregory, R.D., Gibbons, D.W. & Donald, P.F. 2004. *Bird census and survey techniques.* In: **Sutherland, W.J., Newton, I. & Rhys, G.** 2004. *Bird Ecology and Conservation: A Handbook of Techniques.* New York, USA: Oxford University Press. 17-52.

Grémillet, D., Enstipp, M.R., Boudiffa, M. & Liu, H. 2006. Do cormorants injure fish without eating them? An underwater video study. *Marine Biology* **148**: 1081–1087.

Grenzdörffer, G.J. 2013. UAS-based automatic bird count of a common gull colony. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **XL-1/W2**: 169–174.

Grishagin, I.V. 2015. Automatic cell counting with ImageJ. *Analytical Biochemistry* **473**: 63–65.

Guinet, C., Jouventin, P. & Malacamp, J. 1995. Satellite remote sensing in monitoring change of seabirds: use of Spot Image in king penguin population increase at Ile aux Cochons, Crozet Archipelago. *Polar Biol* **15**: 511–515.

Handley, J.M. & Pistorius, P. 2016. Kleptoparasitism in foraging gentoo penguins Pygoscelis papua. *Polar Biol* **39**: 391–395.

Handley, J.M., Thiebault, A., Stanworth, A., Schutt, D. & Pistorius, P. 2018. Behaviourally mediated predation avoidance in penguin prey: in situ evidence from animal-borne camera loggers. *Royal Society Open Science* **5**: 171449.

Harding, A.M.A., Piatt, J.F., Schmutz, J.A., Shultz, M.T., Pelt, T.I.V., Kettle, A.B. & Speckman, S.G. 2007. Prey Density and the Behavioral Flexibility of a Marine Predator: The Common Murre (Uria aalge). *Ecology* **88**: 2024–2033.

Harris, M.P. 1982. Promiscuity in the Shag as shown by timelapse photography. *Bird Study* **29**: 149–154.

Hart, L.A., Downs, C.T. & Brown, M. 2016. Sitting in the sun: Nest microhabitat affects incubation temperatures in seabirds. *Journal of Thermal Biology* **60**: 149–154.

Hays, G.C. 2015. New insights: animal-borne cameras and accelerometers reveal the secret lives of cryptic species. *Journal of Animal Ecology* **84**: 587–589.

Hebblewhite, M. & Haydon, D.T. 2010. Distinguishing technology from biology: a critical review of the use of GPS telemetry data in ecology. *Philos Trans R Soc Lond B Biol Sci* **365**: 2303–2312.

Hervías, S., Henriques, A., Oliveira, N., Pipa, T., Cowen, H., Ramos, J.A., Nogales, M., Geraldes, P., Silva, C., de Ybáñez, R.R. & Oppel, S. 2013. Studying the effects of multiple invasive mammals on Cory's shearwater nest survival. *Biol Invasions* **15**: 143–155.

Hinke, J.T., Barbosa, A., Emmerson, L.M., Hart, T., Juáres, M.A., Korczak-Abshire, M., Milinevsky, G., Santos, M., Trathan, P.N., Watters, G.M. & Southwell, C. 2018. Estimating nest-level phenology and reproductive success of colonial seabirds using time-lapse cameras. *Methods in Ecology and Evolution* **9**: 1853–1863.

Hodgson, J.C., Baylis, S.M., Mott, R., Herrod, A. & Clarke, R.H. 2016. Precision wildlife monitoring using unmanned aerial vehicles. *Scientific Reports* **6**: 1–7.

Hodgson, J.C., Mott, R., Baylis, S.M., Pham, T.T., Wotherspoon, S., Kilpatrick, A.D., Segaran, R.R., Reid, I., Terauds, A. & Koh, L.P. 2018. Drones count wildlife more accurately and precisely than humans. *Methods in Ecology and Evolution* **9**: 1160–1167.

Horswill, C., Jackson, J.A., Medeiros, R., Nowell, R.W., Trathan, P.N. & O'Connell, T.C. 2018. Minimising the limitations of using dietary analysis to assess foodweb changes by combining multiple techniques. *Ecological Indicators* **94**: 218–225.

Huffeldt, N.P. & Merkel, F.R. 2013. Remote Time-lapse Photography as a Monitoring Tool for Colonial Breeding Seabirds: A Case Study Using Thick-billed Murres (Uria lomvia). *Waterbirds: The International Journal of Waterbird Biology* **36**: 330–341.

Hughes, B.J., Martin, G.R. & Reynolds, S.J. 2011. The use of Google EarthTM satellite imagery to detect the nests of masked boobies Sula dactylatra. *wbio* **17**: 210–216.

Hurford, C. 2017. Improving the Accuracy of Bird Counts Using Manual and Automated Counts in ImageJ: An Open-Source Image Processing Program. In: *The Roles of Remote Sensing in Nature Conservation: A Practical Guide and Case Studies* (R. Díaz-Delgado, R. Lucas, & C. Hurford, eds), pp. 249–276. Springer International Publishing, Cham.

Hutchinson, A.E. 1980. Estimating Numbers of Colonial Nesting Seabirds: A Comparison of Techniques. *Proceedings of the Colonial Waterbird Group* **3**: 235–244.

Irigoin-Lovera, C., Luna, D.M., Acosta, D.A. & Zavalaga, C.B. 2019. Response of colonial Peruvian guano birds to flying UAVs: effects and feasibility for implementing new population monitoring methods. *PeerJ* **7**: e8129.

Israel, M. & Reinhard, A. 2017. Detecting nests of lapwing birds with the aid of a small unmanned aerial vehicle with thermal camera. In: *2017 International Conference on Unmanned Aircraft Systems (ICUAS)*, pp. 1199–1207. IEEE, Miami, FL, USA.

Joint Nature Conservation Committee. 2010. *Seaduck survey data.* Available at: http://archive.jncc.gov.uk/page-4570 (accessed 21 January 2020)

Joint Nature Conservation Committee. 2016. *Seabirds Count: A census of breeding seabirds of Britain and Ireland. Census Instructions and Recording Form.* Available at: http://archive.jncc.gov.uk/page-7485 (accessed 21 January 2020)

Jones, F.M., Allen, C., Arteta, C., Arthur, J., Black, C., Emmerson, L.M., Freeman, R., Hines, G., Lintott, C.J., Macháčková, Z., Miller, G., Simpson, R., Southwell, C., Torsey, H.R., Zisserman, A. & Hart, T. 2018. Time-lapse imagery and volunteer classifications from the Zooniverse Penguin Watch project. *Sci Data* **5**: 180124.

Jones, F.M., Arteta, C., Zisserman, A., Lempitsky, V., Lintott, C.J. & Hart, T. 2020. Processing citizen science- and machine-annotated time-lapse imagery for biologically meaningful metrics. *Scientific Data* **7**: 1–15.

Kemper, G., Weidauer, A. & Coppack, T. 2016. MONITORING SEABIRDS AND MARINE MAMMALS BY GEOREFERENCED AERIAL PHOTOGRAPHY. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **XLI-B8**: 689–694.

Korczak-Abshire, M., Kidawa, A. & Zmarz, A. 2016. Preliminary study on nesting Adélie penguins disturbance by unmanned aerial vehicles. *CCAMLR Science* **23**: 1-16.

Korczak-Abshire, M., Zmarz, A., Rodzewicz, M., Kycko, M., Karsznia, I. & Chwedorzewska, K.J. 2019. Study of fauna population changes on Penguin Island and Turret Point Oasis (King George Island, Antarctica) using an unmanned aerial vehicle. *Polar Biol* **42**: 217–224.

Kucera, T.E. & Barrett, R.H. 2011. A History of Camera Trapping. In: *Camera Traps in Animal Ecology: Methods and Analyses* (A. F. O'Connell, J. D. Nichols, & K. U. Karanth, eds), pp. 9–26. Springer Japan, Tokyo.

Lieber, L., Nimmo-Smith, W.A.M., Waggitt, J.J. & Kregting, L. 2019. Localised anthropogenic wake generates a predictable foraging hotspot for top predators. *Commun Biol* **2**: 123.

Loarie, S.R., Joppa, L.N. & Pimm, S.L. 2007. Satellites miss environmental priorities. *Trends in Ecology & Evolution* **22**: 630–632.

Lorentzen, E., Steen, H. & Strøm, H. 2010. Estimating chick survival in cliff-nesting seabirds – a hazard made easy with monitoring cameras. *SEAPOP Short Report* **8**. 1-10.

Luna, N., Varela, A.I., Brokordt, K. & Luna-Jorquera, G. 2018. Assessing Potential Predation Risk by Introduced Predators on Unattended Eggs in the Red-Tailed Tropicbird, Phaethon rubricauda, on Rapa Nui (Easter Island). *Tropical Conservation Science* **11**: 1940082918785079.

Lynch, H.J., Naveen, R., Trathan, P.N. & Fagan, W.F. 2012a. Spatially integrated assessment reveals widespread changes in penguin populations on the Antarctic Peninsula. *Ecology* **93**: 1367–1377.

Lynch, H.J., White, R., Black, A.D. & Naveen, R. 2012b. Detection, differentiation, and abundance estimation of penguin species by high-resolution satellite imagery. *Polar Biol* **35**: 963–968.

Lyons, M., Brandis, K., Wilshire, J., Murray, N., McCann, J., Kingsford, R. & Callaghan, C. 2019. *A protocol for using drones to assist monitoring of large breeding bird colonies. EcoEvoRxiv.*

McCafferty, D.J. 2013. Applications of thermal imaging in avian science. *Ibis* **155**: 4–15.

McClelland, G.T., Bond, A.L., Sardana, A. & Glass, T. 2016. Rapid population estimate of a surface-nesting seabird on a remote island using a low-cost unmanned aerial vehicle. *Marine Ornithology* **44**: 215–220.

McLeay, L.J., Page, B., Goldsworthy, S.D., Ward, T.M. & Paton, D.C. 2009. Size matters: variation in the diet of chick and adult crested terns. *Mar Biol* **156**: 1765–1780.

Mellor, M., Craig, T., Baillie, D. & Woolaghan, P. 2007. *Trial High Definition Video Survey of Seabirds.* COWRIE Ltd.

Mendez, L., Prudor, A. & Weimerskirch, H. 2017. Ontogeny of foraging behaviour in juvenile red-footed boobies (Sula sula). *Scientific Reports* **7**: 13886.

Merkel, F.R., Johansen, K.L. & Kristensen, A.J. 2016. Use of time-lapse photography and digital image analysis to estimate breeding success of a cliff-nesting seabird. *Journal of Field Ornithology* **87**: 84–95.

Mitchell, P.I. & Parsons, M. 2007. Strategic Review of the UK Seabird Monitoring Programme. *JNCC Unpublished Report.*

Moll, R.J., Millspaugh, J.J., Beringer, J., Sartwell, J. & He, Z. 2007. A new 'view' of ecology and conservation through animal-borne video systems. *Trends in Ecology & Evolution* **22**: 660–668.

Mosbech, A., Lyngs, P. & Johansen, K.L. 2017. Estimating little auk (Alle alle) breeding density and chick-feeding rate using video surveillance. *Polar Research* **36**: 1374122.

Mudge, G.P., Aspinall, S.J. & Crooke, C.H. 1987. A photographic study of seabird attendance at Moray Firth colonies outside the breeding season. *Bird Study* **34**: 28–36.

Mulero-Pázmány, M., Jenni-Eiermann, S., Strebel, N., Sattler, T., Negro, J.J. & Tablado, Z. 2017. Unmanned aircraft systems as a new source of disturbance for wildlife: A systematic review. *PLOS ONE* **12**: e0178448.

Müllerová, J., Brůna, J., Bartaloš, T., Dvořák, P., Vítková, M. & Pyšek, P. 2017. Timing Is Important: Unmanned Aircraft vs. Satellite Imagery in Plant Invasion Monitoring. *Front. Plant Sci.* **8**.

Nowak, M.M., Dzióbr, K. & Bogawski, P. 2019. Unmanned Aerial Vehicles (UAVs) in environmental biology: a review. *European Journal of Ecology* **4**: 56–74.

Oosthuizen, W.C., Krüger, L., Jouanneau, W. & Lowther, A.D. 2020. Unmanned aerial vehicle (UAV) survey of the Antarctic shag (Leucocarbo bransfieldensis) breeding colony at Harmony Point, Nelson Island, South Shetland Islands. *Polar Biol* **43**: 187–191.

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

Pascalis, F., Collins, P. & Green, J. 2018. Utility of time-lapse photography in studies of seabird ecology. *PLOS ONE* **13**: e0208995.

Perkins, A.J., Bingham, C.J. & Bolton, M. 2018. Testing the use of infra-red video cameras to census a nocturnal burrow-nesting seabird, the European Storm Petrel Hydrobates pelagicus. *Ibis* **160**: 365–378.

Petersen, A., Irons, D., Anker-Nilssen, T., Artukhin, Y., Barrett, R., Boertmann, D., Egevang, C., Gavrilo, M.V., Gilchrist, G., Hario, M., Mallory, M., Mosbech, A., Olsen, B., Osterblom, H., Robertson, G. & Strøm, H.. 2008. *Framework for a Circumpolar Arctic Seabird Monitoring Network. CAFF CBMP Report No.15.* Akureyri, Iceland: CAFF International Secretariat.

Pfeifer, C., Barbosa, A., Mustafa, O., Peter, H.-U., Rümmler, M.-C. & Brenning, A. 2019. Using Fixed-Wing UAV for Detecting and Mapping the Distribution and Abundance of Penguins on the South Shetlands Islands, Antarctica. *Drones* **3**: 39.

Ponganis, P.J., Dam, R.P.V., Marshall, G., Knower, T. & Levenson, D.H. 2000. Sub-ice foraging behavior of emperor penguins. *Journal of Experimental Biology* **203**: 3275–3278.

Radjawali, I., Pye, O. & Flitner, M. 2017. Recognition through reconnaissance? Using drones for counter-mapping in Indonesia. *The Journal of Peasant Studies* **44**: 817–833.

Robinson, R.A. & Ratcliffe, N. 2010. *The Feasibility of Integrated Population Monitoring of Britain's Seabirds. BTO Research Report No. 526.* Thetford, UK: British Trust for Ornithology.

RSPB. 2020. *Puffarazzi.* Available at: https://www.rspb.org.uk/reserves-and-events/events-dates-and-inspiration/puffarazzi/ (accessed 27 January 2020).

Rümmler, M.-C., Mustafa, O., Maercker, J., Peter, H.-U. & Esefeld, J. 2016. Measuring the influence of unmanned aerial vehicles on Adélie penguins. *Polar Biol* **39**: 1329–1334.

Rümmler, M.-C., Mustafa, O., Maercker, J., Peter, H.-U. & Esefeld, J. 2018. Sensitivity of Adélie and Gentoo penguins to various flight activities of a micro UAV. *Polar Biol* **41**: 2481–2493.

Rush, G.P., Clarke, L.E., Stone, M. & Wood, M.J. 2018. Can drones count gulls? Minimal disturbance and semiautomated image processing with an unmanned aerial vehicle for colony-nesting seabirds. *Ecology and Evolution* **8**: 12322–12334.

Sakamoto, K.Q., Takahashi, A., Iwata, T. & Trathan, P.N. 2009. From the Eye of the Albatrosses: A Bird-Borne Camera Shows an Association between Albatrosses and a Killer Whale in the Southern Ocean. *PLOS ONE* **4**: e7322.

Sardà-Palomera, F., Bota, G., Viñolo, C., Pallarés, O., Sazatornil, V., Brotons, L., Gomáriz, S. & Sardà, F. 2012. Fine-scale bird monitoring from light unmanned aircraft systems. *Ibis* **154**: 177–183.

Schwaller, M.R., Olson, C.E., Ma, Z., Zhu, Z. & Dahmer, P. 1989. A remote sensing analysis of Adélie penguin rookeries. *Remote Sensing of Environment* **28**: 199–206.

Sinclair, N.C., Harris, M.P., Nager, R.G., Leakey, C.D.B. & Robbins, A.M. 2017. Nocturnal colony attendance by common guillemots Uria aalge at colony in Shetland during the pre-breeding season. *Seabird* **30**: 51–62.

Southwell, C. & Emmerson, L. 2015. Remotely-operating camera network expands Antarctic seabird observations of key breeding parameters for ecosystem monitoring and management. *Journal for Nature Conservation* **23**: 1–8.

Stolpmann, L.M., Landers, T.J. & Russell, J.C. 2019. Camera trapping of Grey-faced Petrel (*Pterodroma gouldi*) breeding burrows reveals interactions with introduced mammals throughout the breeding season. *Emu - Austral Ornithology* **119**: 391–396.

Takahashi, A., Sato, K., Naito, Y., Dunn, M.J., Trathan, P.N. & Croxall, J.P. 2004. Penguin–mounted cameras glimpse underwater group behaviour. *Proc. R. Soc. Lond. B* **271**.

Temple, S.A. 1972. A Portable Time-Lapse Camera for Recording Wildlife Activity. *The Journal of Wildlife Management* **36**: 944–947.

Thaxter, C.B. & Burton, N.H.K. 2009. *High Definition Imagery for Surveying Seabirds and Marine Mammals: A Review of Recent Trials and Development of Protocols. British Trust for Ornithology Report Commissioned by COWRIE Ltd.* Thetford, UK: British Trust for Ornithology.

Thiebault, A., Mullers, R.H.E., Pistorius, P.A. & Tremblay, Y. 2014. Local enhancement in a seabird: reaction distances and foraging consequence of predator aggregations. *Behav Ecol* **25**: 1302–1310.

Thiebot, J.-B., Barbraud, C., Delord, K., Marteau, C. & Weimerskirch, H. 2014. Do Introduced Mammals Chronically Impact the Breeding Success of the World's Rarest Albatross? *jorn* **13**: 41–46.

Tremblay, Y., Thiebault, A., Mullers, R. & Pistorius, P. 2014. Bird-Borne Video-Cameras Show That Seabird Movement Patterns Relate to Previously Unrevealed Proximate Environment, Not Prey. *PLOS ONE* **9**: e88424.

Van Berkel, T. 2014. *Camera trapping for wildlife conservation: expedition field techniques*. Geography Outdoors, London.

Villegas, P., Mena, L., Constantine, A., Villalba, R. & Ochoa, D. 2018. Data Imaging Acquisition and Processing as a Methodology for Estimating the Population of Frigates Using UAVs. In: *2018 IEEE ANDESCON*, pp. 1–4.

Votier, S.C., Bearhop, S., MacCormick, A., Ratcliffe, N. & Furness, R.W. 2003. Assessing the diet of great skuas, Catharacta skua, using five different techniques. *Polar Biol* **26**: 20–26.

Votier, S.C., Bicknell, A., Cox, S.L., Scales, K.L. & Patrick, S.C. 2013. A Bird's Eye View of Discard Reforms: Bird-Borne Cameras Reveal Seabird/Fishery Interactions. *PLoS One* **8**.

Waldron, A., Mooers, A.O., Miller, D.C., Nibbelink, N., Redding, D., Kuhn, T.S., Roberts, J.T. & Gittleman, J.L. 2013. Targeting global conservation funding to limit immediate biodiversity declines. *PNAS* **110**: 12144–12148.

Walsh, P.M., Halley, D.J., Harris, M.P., Del Nevo, A., Sim, I.M.W. & Tasker, M.L. 1995. *Seabird monitoring handbook for Britain and Ireland*. Peterborough, UK: JNCC / RSPB /ITE / Seabird Group.

Waluda, C.M., Dunn, M.J., Curtis, M.L. & Fretwell, P.T. 2014. Assessing penguin colony size and distribution using digital mapping and satellite remote sensing. *Polar Biol* **37**: 1849–1855.

Wang, D., Shao, Q. & Yue, H. 2019. Surveying Wild Animals from Satellites, Manned Aircraft and Unmanned Aerial Systems (UASs): A Review. *Remote Sensing* **11**: 1308.

Watanuki, Y., Daunt, F., Takahashi, A., Newell, M., Wanless, S., Sato, K. & Miyazaki, N. 2008. Microhabitat use and prey capture of a bottom-feeding top predator, the European shag, shown by camera loggers. *Marine Ecology Progress Series* **356**: 283–293.

Weimerskirch, H., Prudor, A. & Schull, Q. 2018. Flights of drones over sub-Antarctic seabirds show species- and status-specific behavioural and physiological responses. *Polar Biol* **41**: 259–266.

Weller, M.W. & Derksen, D.V. 1972. Use of Time-Lapse photography to Study Nesting Activities of Birds. *The Auk* **89**: 196–200.

Whelan, R., Clarke, C., Almansoori, N., Jaradat, A., Qadi, N.S.A. & Muzaffar, S.B. 2018. Demographic consequences of native fox predation on Socotra cormorants on Siniya Island, United Arab Emirates. *wbio* **2018**. 1-13.

Wilhelm, S.I., Mailhiot, J., Arany, J., Chardine, J.W., Robertson, G.J. & Ryan, P.C. 2015. Update and trends of three important seabird populations in the western North Atlantic using a geographic information system approach. *Marine Ornithology* **43**: 211–222.

Williams, H.M. & DeLeon, R.L. 2020. Deep learning analysis of nest camera video recordings reveals temperature-sensitive incubation behavior in the purple martin (Progne subis). *Behav Ecol Sociobiol* **74**: 7.

Yoda, K., Murakoshi, M., Tsutsui, K. & Kohno, H. 2011. Social Interactions of Juvenile Brown Boobies at Sea as Observed with Animal-Borne Video Cameras. *PLoS One* **6**.

Table 1 Advantages and disadvantages of using Unmanned Aerial Vehicles (UAVs) to monitor seabirds.

Advantages	**Disadvantages**
Cost-effective: short survey time, low purchase and operation costs. (Bibby *et al.* 2000, Buckland *et al.* 2012, Rush *et al.* 2018, Villegas *et al.* 2018, Nowak *et al.* 2019)	More affordable UAVs take lower resolution images. (Nowak *et al.* 2019)
Portability and limited launch requirements allow operation in most locations and terrains. (Goebel *et al.* 2015)	National and regional administrative regulations can affect possibility of data acquisition. (Nowak *et al.* 2019)
Manoeuvrable, so can operate over small areas and monitor small objects. (Nowak *et al.* 2019)	Reduced use in areas with limited electricity. (Radjawali *et al.* 2017, Nowak *et al.* 2019)
Operate at locations and times when ground-based field observations would be near-impossible. For example, remote locations, onshore and offshore, difficult terrain, at night. (Rush *et al.* 2018)	Vulnerable to damage in adverse weather conditions. (McClelland *et al.* 2016)
Greater control over the scale, quality, and temporal and spatial resolution of images. (Thaxter & Burton 2009, Rush *et al.* 2018, Korczak-Abshire *et al.* 2019, Nowak *et al.* 2019)	Large amount of data to handle and analyse. (Rush *et al.* 2018)
Downward-facing view can observe birds in a range of habitats and help reduce missed counts. (Rush *et al.* 2018, Villegas *et al.* 2018)	Data quality depends on operator skill, environment and meteorological conditions during flight. (Nowak *et al.* 2019)
Combine habitat mapping and seabird occupancy from images, to investigate how habitat features affect populations. (Oosthuizen *et al.* 2020)	Animals may modify their behaviour in response to a flying object, increasing intra-specific aggression, predation of eggs/chicks and nest abandonment. (Rush *et al.* 2018, Nowak *et al.* 2019)

Permanent record viewable any number of times and available for independent verification. (Thaxter & Burton 2009, Buckland *et al.* 2012, Rush *et al.* 2018)	
Reduced nest and site disturbance compared to walk-through surveys. (Rush *et al.* 2018)	
Reduced disturbance when flown at the same height as manned aircraft. (Goebel *et al.* 2015, Korczak-Abshire *et al.* 2019)	
Removes observer bias from variation in surveyor experience and alertness over a long period. This is useful when observers are swamped with a large number of birds to count. (Bibby *et al.* 2000, Rush *et al.* 2018, Thaxter & Burton 2009)	

Table 2 Advantages and disadvantages of using time-lapse photography to monitor seabirds.

Advantages	**Disadvantages**
Cost-effective: saves time and money during fieldwork. For example, difficult for a single researcher to record detailed nest activity across multiple nests at a colony. (Weller & Derksen 1972, Black 2018, Pascalis *et al.* 2018)	Mechanical failures. (Cutler & Swann 1999, Merkel *et al.* 2016, Black 2018)
Increased spatial and temporal scale of monitoring. (Southwell & Emmerson 2015, Merkel *et al.* 2016)	Programming errors. (Cutler & Swann 1999, Black 2018)
Operates at locations and times when field observation would be near-impossible. For example, remote locations, harsh weather conditions, at night. (Cutler & Swann 1999, Southwell & Emmerson 2015, Black *et al.* 2017, Sinclair et al. 2017, Black 2018, Black *et al.* 2018)	Maintenance required. For example, images are vulnerable to camera movements caused by harsh weather conditions. (Merkel *et al.* 2016, Black 2018)
Removes observer bias from variation in surveyor experience and alertness over a long period. (Cowardin & Ashe 1965, Weller & Derksen 1972, Cutler & Swann 1999, Black 2018)	More affordable cameras take lower resolution images. (Black 2018)
More frequent observations than field workers allows observation of elusive species, obscure behaviours and phenology. (Cutler & Swann 1999, Black 2018)	Large camera networks needed to monitor an entire colony, which are expensive to install and maintain. (Black 2018)
Permanent record viewable any number of times and available for independent verification. (Cutler & Swann 1999, Merkel *et al.* 2016; Sinclair et al. 2017, Black, 2018)	Large amount of data to handle and analyse. (Merkel *et al.* 2016, Black 2018)

Easier to maintain comparable study effort between years. (Merkel *et al.* 2016)	Cameras rarely possess thermal imaging or infra-red sensors, making night monitoring difficult. (Black *et al.* 2018)
Infrequent visitation lowers nest and site disturbance. (Cutler & Swann 1999)	

Table 3 Advantages and disadvantages of photo-sampling for obtaining information on seabird chick diet.

Advantages	**Disadvantages**
Non-invasive, assuming photographers remain a safe distance from birds. (Gaglio *et al.* 2017)	Only suitable for species that carry prey in their bills. (Gaglio *et al.* 2017)
Possible in a range of locations, including from land and boat. (Gaglio *et al.* 2017)	Large amount of data to handle and process. (Gaglio *et al.* 2017)
Large amounts of data can be collected in a short time-period.	Repeated photography of individuals carrying the same prey load. (Gaglio *et al.* 2017)
Minimal training to use cameras. (Gaglio *et al.* 2017)	Observer bias. (Gaglio *et al.* 2017)
Equipment relatively affordable and durable. (Gaglio *et al.* 2017)	Chick diet is not always representative of adult diet, or diet outside the breeding season. (McLeay *et al.* 2009, Gaglio *et al.* 2017)
Permanent record of observations available for independent verification and re-analysis without loss of quality. For example, prey samples degrade over time. (Gaglio *et al.* 2017)	Challenging in poor weather conditions. (Gaglio *et al.* 2017)
Only requires one individual to collect photographs. (Gaglio *et al.* 2017)	Large-scale studies across multiple locations/species are time-consuming, unless multiple people are deployed.
More likely to record the entire prey item than regurgitations, aiding accurate identification. (Gaglio *et al.* 2017, McLeay *et al.* 2009)	

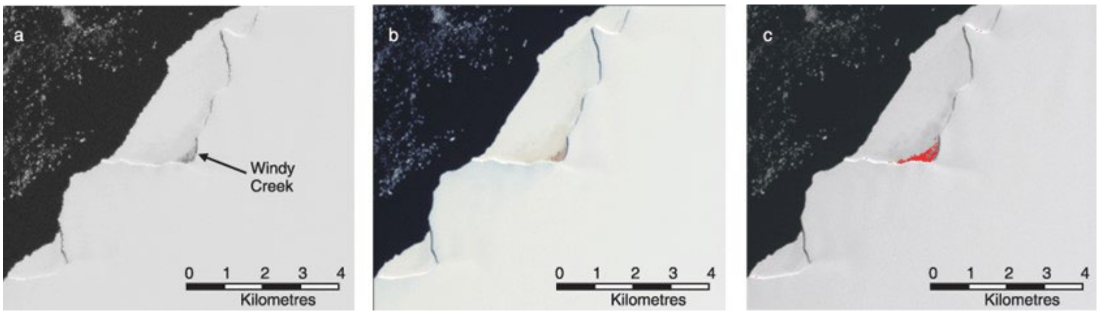


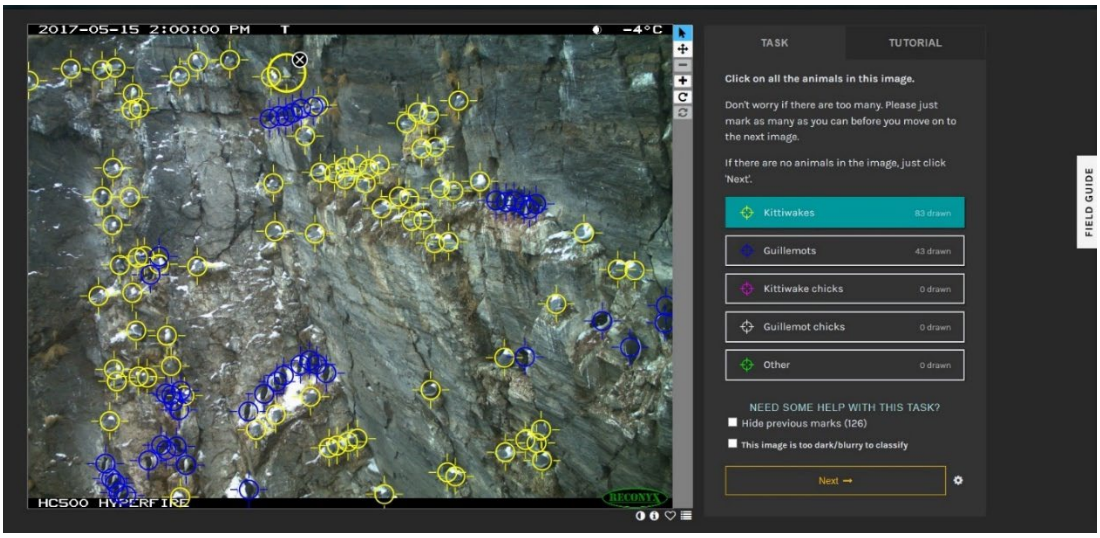
Figure 1 Landsat ETM imagery used to identify the Windy Creek Emperor Penguin *Aptenodytes forsteri* colony from faecal stains. a) Data viewed online from the Landsat Image Mosaic of Antarctica (LIMA) website showed a potential penguin colony. b) Data downloaded from the LIMA website and viewed in GIS clearly showed the brown faecal staining of the colony. c) Spectral analysis identified areas where the red band had a higher value than the blue band. The resulting positive area, shown in red, located the exact area of the colony (Fretwell & Trathan, 2009). Images: MAXAR.

Figure 2 Aerial photograph of the gannetry on Grassholm Island, UK, 2015. Image: Sarah Money.

Figure 3 DJI Inspire 1 quadcopter Unmanned Aerial Vehicle (UAV) fitted with a DJI FC350 camera being used to survey Lesser Black-backed Gull nests on Skokholm Island, UK, 2016 (Rush *et al.* 2018). Image: Matt Wood.

Figure 4 Time-lapse photograph of nesting Black-legged Kittiwakes at Protheroe's Dock, Skomer Island, UK, 2018. Image: *Seabird Watch*.

Figure 5 Photograph of an Atlantic Puffin carrying prey, submitted to RSPB Project Puffin UK. Image: Alice Edney.



Click on all the animals in this image.

Don't worry if there are too many. Please just mark as many as you can before you move on to the next image.

If there are no animals in the image, just click 'Next'.

Kittiwakes 83 drawn
Guillemots 43 drawn
Kittiwake chicks 0 drawn
Guillemot chicks 0 drawn
Other 0 drawn

NEED SOME HELP WITH THIS TASK?

☐ Hide previous marks (126)
☐ This image is too dark/blurry to classify

Figure 6 Annotated Seabird Watch image on the Zooniverse platform. Yellow circles mark adult Black-legged Kittiwakes and blue circles mark adult guillemots. Image annotated by Alice Edney.