Seabird GPS tracking on the Isle of May, Fowlsheugh and St Abb's Head in 2021 in relation to offshore wind farms in the Forth/Tay region

Maria I. Bogdanova, Charlotte Regan, Carrie Gunn, Mark Newell, Ella Benninghaus, Samuel Langlois Lopez, Marine Quintin, Kate Searle & Francis Daunt (UKCEH)

Saskia Wischnewski, Ian Cleasby, Emma Witcutt, David Kinchin-Smith, Elizabeth Holmes, Derren Fox & Aly McCluskie (RSPB)

Katherine Whyte, Adam Butler & Esther L. Jones (BioSS)

Issue Number 2.1

Date 05/12/2022

Title	Seabird GPS tracking on the Isle of May, St Abb's Head and Fowlsheugh in 2021 in relation to offshore wind farms in the Forth/Tay region
Authors	Maria I. Bogdanova, Saskia Wischnewski, Ian Cleasby, Katherine Whyte, Charlotte Regan, Carrie Gunn, Mark Newell, Ella Benninghaus, Samuel Langlois Lopez, Marine Quintin, Emma Witcutt, David Kinchin-Smith, Elizabeth Holmes, Derren Fox, Kate Searle, Adam Butler, Esther Jones, Aly McCluskie & Francis Daunt
Client	Neart na Gaoithe Offshore Wind Limited and SSE Renewables
Client reference	NNG-CEH-ECF-REP-0007
UKCEH reference	07956
UKCEH contact details	Francis Daunt Bush Estate, Penicuik, Midlothian, EH26 0QB t: 0131 445 8526 f: 0131 445 3943 e: frada@ceh.ac.uk
Author	Bogdanova, Maria I.
Approved by	Daunt, Francis
Date	05/12/2022

Contents

Executive summary

- This report presents the results of GPS tracking of adult common guillemots, razorbills, and Atlantic puffins breeding on the Isle of May, and of black-legged kittiwakes breeding on the Isle of May, St Abb's Head and Fowlsheugh (east Scotland) in 2021, as well as an assessment of connectivity with consented and proposed offshore wind farms in the Forth/Tay region (Neart na Gaoithe, Seagreen and Berwick Bank).

- Locational data were obtained from 23 guillemots, 11 razorbills, 24 puffins and 50 kittiwakes on the Isle of May, 40 kittiwakes at Fowlsheugh and 37 kittiwakes at St Abb's Head, in June and July 2021. Kittiwake deployments involved two types of GPS loggers (Pathtrack and UvA-Bits), in order to collect data on distributions and fine-scale behaviour at sea, and to further investigate device effects. A resampling procedure suggested that the sample sizes of tracked birds were adequate to estimate the at-sea area used by the study populations during the deployment period.

- The data were partitioned into non-commuting behaviours (foraging and resting), relevant to displacement effects, and commuting behaviours, relevant to collision risk and barrier effects (guillemot and razorbill from TDR; kittiwake and puffin from HMM). For kittiwake, movement models differentiated between three states that most likely represent commuting, foraging, and resting. For puffins, two states comprising flight and non-flight were identified. Caution is advised interpreting these classifications because there was residual autocorrelation within the data.

- The at-sea distributions of the four study species breeding on the Isle of May included both inshore and offshore areas, as found in previous GPS tracking studies in 2010-20. Differences among the species were apparent, with guillemots and razorbills using coastal areas within the Firth of Forth extensively, and puffins and kittiwakes using mainly offshore waters. Core areas (50% utilisation distribution contours) of all species included waters around the colony. The mean maximum range (± SE) was 26.5 ± 2.2 km in guillemots, 36.0 ± 2.3 km in razorbills, 40.5 ± 1.8 km in puffins and

71.3 ± 3.6 km in kittiwakes. Kittiwakes from Fowlsheugh ranged from north-east to south-east of the colony, with core areas located mostly to the east. The mean maximum range was 81.9 ± 3.8 km. Kittiwakes from St Abb's Head used coastal waters and offshore areas spanning north to east of the colony, with core areas located mostly to the north-east. The mean maximum range at this colony was 63.7 ± 4.8 km.

The distribution of commuting flight trajectories matched closely the distributions based on non-commuting activities.

- Overlap between distributions associated with non-commuting behaviours and the Neart na Gaoithe footprint varied between species. In guillemots, core foraging and resting areas did not overlap with the wind farm footprint and the overall area used overlapped with it to a small extent (<2%). In razorbills, overlap of foraging areas (both core and overall) was low whereas overlap of resting areas was slightly higher (up to 8%). In puffins, overlap was substantially higher, in particular for core areas (nearly 20%). In kittiwakes from the Isle of May, there was some overlap between core areas and Neart na Gaoithe (up to 14%), whereas overall areas overlapped little (<4%). However, the entire footprint fell within the 50% and 90% UD contours of puffins and kittiwakes from the Isle of May, and large parts of the footprint within the 90% UD contours of razorbills. The areas used for foraging and resting by kittiwakes from Fowlsheugh did not overlap with this wind farm. Similarly, there was negligible (<0.05%) overlap between the areas used by birds from St Abb's Head and Neart na Gaoithe.

  Overlap between foraging and resting areas and the Seagreen wind farm was low overall. Areas used by birds from the Isle of May (all four species) did not overlap with the Seagreen footprint. In kittiwakes from the other two colonies the overlap was between 0 and 5%.

  There was negligible (<0.5%) overlap between core foraging and resting areas of birds from all species and colonies and the Berwick Bank footprint. Overlap with the overall areas used for non-commuting behaviours varied between species and colonies, from 0 in guillemots, ≤2% in razorbills and puffins to nearly 20% in kittiwakes from St Abb's Head. All overlap values are provided in tabulated form in the report.

- The overlap of commuting activities with the Neart na Gaoithe footprint was generally higher than the overlap of non-commuting activities. All species except guillemots travelled through the wind farm extensively (64% of razorbills, 96% of puffins and 100% of kittiwakes). Zero and one bird from Fowlsheugh and St Abb's Head, respectively, were recorded flying through the Neart na Gaoithe footprint.

  Commuting flight activities of guillemots and razorbills did not overlap with the Seagreen footprint, and overlap in flight activities of puffins and kittiwakes from the Isle of May was low. Kittiwakes from Fowlsheugh used Seagreen more extensively, with 58% of the study birds passing through it at least once. No birds from St Abb's Head were recorded flying through the Seagreen footprint.

  Commuting flights of guillemots and razorbills did not overlap with Berwick Bank. Puffin flight activities showed more extensive overlap with this wind farm than the other auk species (17% of birds). Kittiwakes from all three colonies used the wind farm area, with highest proportion of birds from the Isle of May (72%), followed by birds from Fowlsheugh (35%) and St Abb's Head (11%).

- Using data collected by the multi-sensor (GPS-altimeter-accelerometer) UvA-Bits loggers we estimated flight heights and associated error in kittiwakes with a variety of newly developed methods and compared the performance of each approach. Our results indicate that GPS altitude and barometric pressure can both be used to estimate flight altitude. GPS and barometric altitude were correlated, but there were instances when estimates varied quite significantly. Precision increased at higher sampling rates. In addition, we observed a consistent tendency for estimates of altitude from barometers to be slightly higher than GPS estimates (median ~3.5 - 4 metres) during periods when birds were classified as being in flight. Further work is needed to understand what drives this pattern.

- In contrast to previous years, there was reduced evidence for negative effects of handling or GPS logger deployment on chick feeding rates in puffins. There was some evidence for a reduction in feeding rates by both colour-ringed and GPS tagged individuals; however, no difference in feeding rates was observed between treatment and control burrows, indicating compensation by the unmanipulated partner was effective. As we carried out supplementary feeding of all chicks at treatment burrows (informed by findings in previous years), the chicks from treatment burrows fledged in better condition compared to controls. Survival of all study chicks was high, with all but two fledging.

- We recorded measurable negative effects of UvA logger deployment on kittiwake at-colony behaviour (parent changeover rates, indicative of foraging trip duration, and chick attendance). At all study colonies, there were fewer parent changeovers per day (indicative of longer foraging trips) at nests where one bird was tagged with a UvA logger than at control and Pathtrack nests. Chicks of UvA-tagged birds were left unattended more often than those at control and Pathtrack nests. Breeding success of UvA nests at Fowlsheugh was lower, whereas Pathtrack nests fledged more chicks than controls. Device effects on at-sea behaviour were found at Fowlsheugh only, where UvA-tagged birds made longer foraging trips compared to Pathtrack-tagged birds, and core areas were located further offshore.

Conclusions: This study demonstrates variation in at-sea distributions between seabird species breeding on the Isle of May and, when comparing with previous GPS tracking studies undertaken between 2010 and 2020, variation within species between years. Our results from 2021, with the addition of two colonies to the study, indicate that the extent to which local seabirds use the Forth/Tay wind farm areas, for both commuting and non-commuting activities, varies between species and colonies. We also found that both GPS altitude and barometric pressure recorded by the multi-sensor UvA-Bits loggers can be used to provide estimates of flight altitude. We found some evidence of device effects in kittiwakes carrying UvA tags, and propose to focus future work on the smaller Pathtrack tags, where no device effects were apparent. Although there was little evidence of device effects in puffins in 2021 compared to previous years, we will continue to monitor this closely and consider options for smaller loggers should they become available. The interannual variation in distribution indicates that the structured monitoring plan being undertaken in this study, including GPS tracking before, during and after wind farm construction, is valuable and will maximise opportunities for quantifying any impacts of the Forth/Tay wind farms on these protected seabird populations.

1 Introduction

1.1 Background

Understanding the potential ecological impact of offshore wind farms (OWFs) on seabird populations has become a crucial issue in seabird conservation (Furness et al. 2013, Dierschke et al. 2016, Peschko et al. 2020). In the UK, the number of offshore wind developments is anticipated to increase substantially over the next few decades in delivering to UK and devolved policy on energy security and net zero (British energy security strategy, Net Zero Strategy: Build Back Greener, Climate Change (Emissions Reduction Targets) (Scotland) Act 2019). OWFs have the potential to impact on protected seabird populations, mainly due to collisions with turbine blades, displacement from important habitat and barrier effects to movements (Drewitt & Langston 2006; Larsen & Guillemette 2007; Masden et al. 2010; Grecian et al. 2010, Langton et al. 2011; Searle et al. 2014, 2018). These effects may be particularly important for breeding seabirds that are constrained to forage within a certain distance from the colony because of the requirement to return regularly to the nest to relieve the attending mate and feed the young (Daunt et al. 2002; Enstipp et al 2006). Accordingly, for the purposes of Habitats Regulations Appraisal, there is a need to estimate the potential impact of OWFs on seabirds breeding at Special Protection Areas (SPAs). To achieve this requires two steps: (1) to determine the extent of interaction of birds from colony SPAs with OWFs, and (2) to estimate whether such interactions are having a detrimental effect at the population level. Empirical studies that quantify key behaviours (foraging, flight and resting at sea), energetics and demographic rates of individuals before, during and after construction, provide the evidence base to answer these questions, with a view to ultimately establishing whether any population-level consequences cause the SPA interest features to no longer meet the site's Conservation Objectives.

1.2 Sub-lethal and lethal effects of OWFs

A key challenge is quantifying sub-lethal effects of OWFs on seabirds, notably displacement and barrier effects. Sub-lethal effects are challenging to study because they do not cause immediate mortality, but instead cause changes in behaviour which may alter energy budgets and demography (breeding success and survival) of the affected individuals. Ultimately, if the breeding success and/or survival of enough individuals are being affected in a sufficiently strong way, there would be an impact on population size. The most powerful approach to quantifying these effects is to compare the fate of individual birds that vary in the extent to which their behaviour is altered by OWFs. The effects on all individuals are then combined to estimate a population-level impact. A combination of bio-logging and monitoring approaches are required to obtain the relevant individual-level data to achieve this. GPS tracking of breeding adults at SPAs enables their distribution and key areas at sea to be established, and the extent to which these overlap with the OWF(s) of interest to be quantified. The use of additional bio-logging technologies, such as accelerometry and time-depth recording, allows classification of the GPS data into key behaviours and calculation of time-activity and energy budgets more accurately than can be achieved from GPS data alone. Further targeted monitoring at the colony provides information on the condition of the study individuals, and the condition and survival of their chicks.

Estimates of collision risk with OWFs and hence mortality (lethal effects) are currently based on well-established collision risk models (e.g. Band 2012, Masden 2015, McGregor et al. 2018) which are parameterised using our understanding of seabird ecology and wind turbine design (reviewed in: Masden & Cook 2016). Focussing on the seabird ecology aspect of such models, practitioners need to supply estimates of seabird density, flight speed, nocturnal activity etc. One crucial aspect of such models is the provision of representative distributions of flight altitudes as this information can influence collision risk model outputs (Cook et al. 2014, Johnston et al. 2014, Largey et al. 2021, Masden et al. 2021). However, while flight altitude is an important component of collision risk models, obtaining accurate and precise measures of this parameter is challenging. There are multiple ways in which the flight altitude of birds can be estimated, each with their own strengths and weaknesses. For example, flight altitude can be assessed by: (1) boat-based observers assigning birds to height categories in order to generate flight height distributions (Johnston et al. 2014); (2) using radar, digital high-definition aerial surveys and rangefinders (e.g. Shamoun-Baranes et al. 2006; Mendel et al. 2014, Cook et al. 2018); (3) attaching data loggers such as GPS units or barometers to birds and recording altitude and/or air pressure (Garthe et al. 2014, Cleasby et al. 2015, Ross-Smith et al. 2016, Lane et al. 2019, Peschko et al. 2021). Given the ongoing vast expansion of tracking studies, the estimation of flight altitude from bird-borne loggers such as GPS and altimeters is of particular relevance. However, the use of bio-logging data is associated with a number of challenges, described in detail in Appendix 1. As part of this study, we developed methods for estimating flight altitude from bio-logging data, whilst taking into account such challenges, and assessed their performance. Understanding the performance of such approaches is crucial as both GPS and barometric sensors can be used to estimate flight altitude and there is currently no standardized way of using such data. Consequently, different approaches may yield different altitude estimates with down-stream effects on any impact assessments such as collision risk modelling.

Monitoring pre- and during construction is fundamental to interpreting potential effects of OWFs and a key strand of the structured before-during-after design. Accordingly, we were tasked by Neart na Gaoithe Offshore Wind Limited (NnGOWL), Seagreen Wind Energy Ltd and SSE Renewables, in the context of the consented Neart na Gaoithe and Seagreen OWFs currently under construction and the proposed Berwick Bank OWF currently under development in the Forth/Tay region, to undertake GPS tracking of seabirds breeding at three colonies along the east coast of Scotland during the 2021 breeding season. The work was tasked with addressing two questions: (1) What is the extent of interaction of breeding birds from colony SPAs with OWFs in the Forth/Tay region? (2) What are the population-level consequences of these interactions? The work on the Isle of May, part of the Forth Islands SPA, focussed on four species that have been central to HRA/EIA assessments of these developments: common guillemot *Uria aalge* (hereafter 'guillemot'), razorbill *Alca torda,* Atlantic puffin *Fratercula arctica* (hereafter 'puffin') and black-legged kittiwake *Rissa tridactlya* (hereafter 'kittiwake'). The work at Fowlsheugh and St Abb's Head involved tracking of kittiwakes only. This project constitutes part of monitoring during construction following GPS tracking undertaken on the same species on the Isle of May in 2018-2020 as part of pre-construction monitoring (Bogdanova et al. 2018, 2020, 2021). Note, however, that from a seabird disturbance perspective the 2021 breeding season may be considered more similar to pre-construction years since no structures were installed above the water at that time.

1.3 Device effects

An important consideration for studies using bio-logging technologies, including GPS tracking, is the potential for negative device effects on the study animals. Device effects on birds are being increasingly recognised as is the awareness of the importance of reducing these (Bodey et al. 2018). Key considerations in this respect are the choice of device (dimensions, weight, shape), optimal placement on the bird and attachment methods, and minimising handling-related disturbance (Vandenabeele et al. 2012, 2014; Thaxter et al. 2014, 2017). It is imperative that studies take these factors into account in order to minimise negative impacts on the study populations and increase the representativeness of resulting data.

Previous work has shown that puffins on the Isle of May and at other colonies are susceptible to disturbance, including handling and attachment of data loggers (Rodway et al 1996, Harris & Wanless 2011; Harris et al. 2012; Bogdanova et al. 2018, 2020, 2021; Daunt & Bogdanova 2022). Accordingly, as in previous years, we used the smallest suitable loggers available on the market and captured birds at burrow entrances, thereby ensuring only one adult per pair was tracked and that the chicks of all instrumented birds were identified and could be supplementary fed to secure their wellbeing should provisioning rates be affected. We also carried out deployments when the puffin chicks were older and more robust. We then, as in 2018-2020, undertook an assessment of the effects of GPS deployment on chick feeding rates and chick condition and survival.

Recent work also indicates that GPS logger deployment can negatively affect the behaviour and physiology of kittiwakes, reporting elevated levels of stress hormones, reduced nest attendance and longer foraging trips (Heggøy et al. 2015) or reduced time spent flying (Chivers et al. 2016) in tagged compared to control birds. Our previous work on the Isle of May (during the 2019 breeding season) showed no negative device effects on kittiwake foraging behaviour (parent changeover rates, indicative of foraging trip duration), chick condition or breeding success when using 4.0g loggers attached to tail feathers with tape (Bogdanova et al. 2020). However, we found small but measurable negative effects of GPS deployment on parent changeover rates, chick attendance and breeding success when using more sophisticated but larger (9.0g) loggers with a different attachment method (glued to back feathers) and longer deployment duration during the 2020 breeding season (Bogdanova et al. 2021). The negative effects on breeding success on the Isle of May were in contrast to findings at Flamborough and Filey in 2017 and 2018 (Wischnewski et al. 2018). In order to tease apart device effects from interannual and between-colony variation and to inform future tracking work on this species, in 2021 we deployed both logger types simultaneously at three colonies (Isle of May, Fowlsheugh and St Abb's Head) and tested for device effects on 1) at-colony behaviour (parent changeover rates at the nest as a proxy of foraging trip duration and chick attendance), 2) at-sea behaviour (foraging trip duration, total distance and range, and utilisation distributions), 3) adult and chick condition (Isle of May only) and 4) breeding success.

1.4 Objectives

The objectives of this project were to:

1) carry out GPS tracking of guillemots, razorbills, puffins and kittiwakes breeding on the Isle of May, and kittiwakes breeding at Fowlsheugh and St Abb's Head during the 2021 breeding season in order to establish their at-sea distribution and flight trajectories, and to estimate overlap with the Neart na Gaoithe, Seagreen and Berwick Bank wind farms. This work built on GPS tracking undertaken on these four species on the Isle of May in 2018, 2019 and 2020 (Bogdanova et al. 2018, 2020, 2021) and was expanded to include two other major colonies for kittiwakes along the east coast of Scotland;

2) develop methods for estimating flight altitude using data from multi-sensor loggers, attached to kittiwakes, which incorporated GPS units, altimeters and accelerometers (UvA-Bits) and assess their performance; and

3) assess potential effects of GPS deployment on puffins and kittiwakes.

The report contains a series of maps of distributions and flight trajectories, estimates of overlap with the planned wind farm and analyses of minimum adequate sample size, where we assess whether our data were sufficient to reliably estimate the population distribution over the sampling period. We also present different approaches to estimating flight altitude from bio-loggers using data from a variety of in-built sensors (GPS, altimeter and accelerometer). Finally, we report the results from the analysis of device effects in puffins and kittiwakes. The report represents joint work by UKCEH, RSPB and BioSS.

2 Methods

2.1 GPS tracking

2.1.1 Data collection

The data were collected on the Isle of May National Nature Reserve (56°11' N, 2°33' W), St Abb's Head (55°55' N, 2°08' W) and Fowlsheugh (56°55' N, 2°11' W) in June and July 2021. For the auk species, we used remote download Pathtrack nanoFix-GEO+RF GPS loggers (guillemot: 59x24x11mm, 16.5g, *ca.*2.1% of minimum body mass; razorbill: 54x24x10mm, 12.8g, *ca.*2.3% of minimum body mass, and puffin: 42x14x8mm, 3.4g, *ca.*1% of minimum body mass; all with ~50mm external whip antenna; Fig. 1a-c). For guillemot and razorbill, the GPS loggers had Time-Depth Recorder (TDR) added, allowing us to accurately identify foraging behaviour. For puffins, we selected the smallest and lightest GPS-only loggers available in order to minimise any negative device effects (Bogdanova et al. 2020, 2021). For kittiwakes, we used two types of loggers: Pathtrack nanoFix-GEO+RF (42x14x8mm, 4.0g, *ca.*1.3% of minimum body mass) and remote download GPS-accelerometer-altimeter loggers (approx. 62x30x12 mm, 9.0g, *ca.*2.9% of minimum body mass) developed at the University of Amsterdam (UvA, Fig. 1d, e). Deployment details are provided in Table 1.

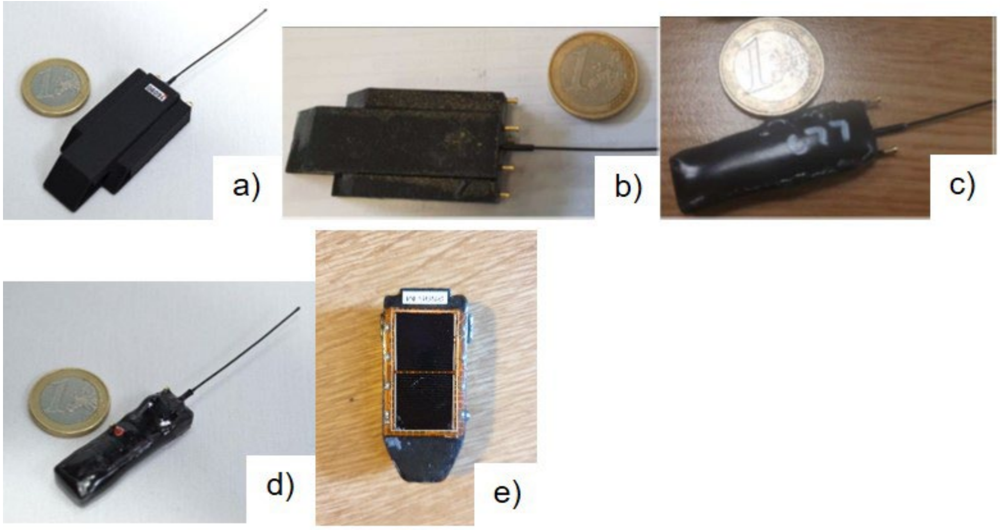


Figure 1: GPS devices used in the project for: a) guillemot; b) razorbill; c) puffin; d) kittiwake (Pathtrack) and e) kittiwake (UvA). €1 coin provided for scale. See text for full details.

Species	Colony	Deployment dates	Number deployed	Number with data
Guillemot	Isle of May	21 – 22 June	25	23
Razorbill	Isle of May	21 – 22 June	15	11
Puffin	Isle of May	2 – 6 July	25	24
Kittiwake	Isle of May	28 June – 1 July	50 (37 Pt+13 UvA)	50
Kittiwake	Fowlsheugh	1 – 10 July	40 (26 Pt+14 UvA)	40
Kittiwake	St Abb's	27 June – 7 July	39 (26 Pt+13 UvA)	37

Table 1: Details of logger deployment dates and number of birds tracked for the four study species. Pt = Pathtrack tag, UvA = University of Amsterdam tag.

All data were collected during chick rearing, except in three kittiwakes on the Isle of May that were tracked towards the end of incubation. In three of the species (guillemot, razorbill and kittiwake), breeding adults were captured at the nest site with a noose at the end of an extendable pole. Puffins were captured using purse nets at the entrance to their nesting burrows. In guillemot, razorbill and puffin, the loggers were attached to lower back feathers using waterproof Tesa tape (Fig. 2a). In kittiwakes, the Pathtrack loggers were attached to central tail feathers using waterproof Tesa tape (Fig. 2b). The UvA loggers were deployed on the upper back with superglue to ensure the solar panels were exposed even when the birds were at the nest with their wings tucked in and the loggers remained attached for a sufficient length of time (Fig. 2c). The attachment method for this species required for the feathers in the attachment site to be trimmed down to 5mm or less. Any rough edges were trimmed away with surgical scissors.

On the Isle of May, average handling time for auks and kittiwakes equipped with Pathtrack loggers was 5 minutes (range 1-11 mins). Handling time for kittiwakes equipped with UvA loggers averaged 8 minutes (range 5-12). At Fowlsheugh, handling time for kittiwakes deployed with Pathtrack loggers was 10 minutes (range 8-14), and with UvA loggers 14 minutes (range 10-18). At St Abb's Head average handling time was 12 minutes (range 8-20) and 17 minutes (range 13-25) for Pathtrack and UvA loggers, respectively.

Pathtrack deployments lasted up to *ca.* two weeks before the loggers fell off; UvA deployments lasted longer, with some birds carrying loggers for over four weeks.

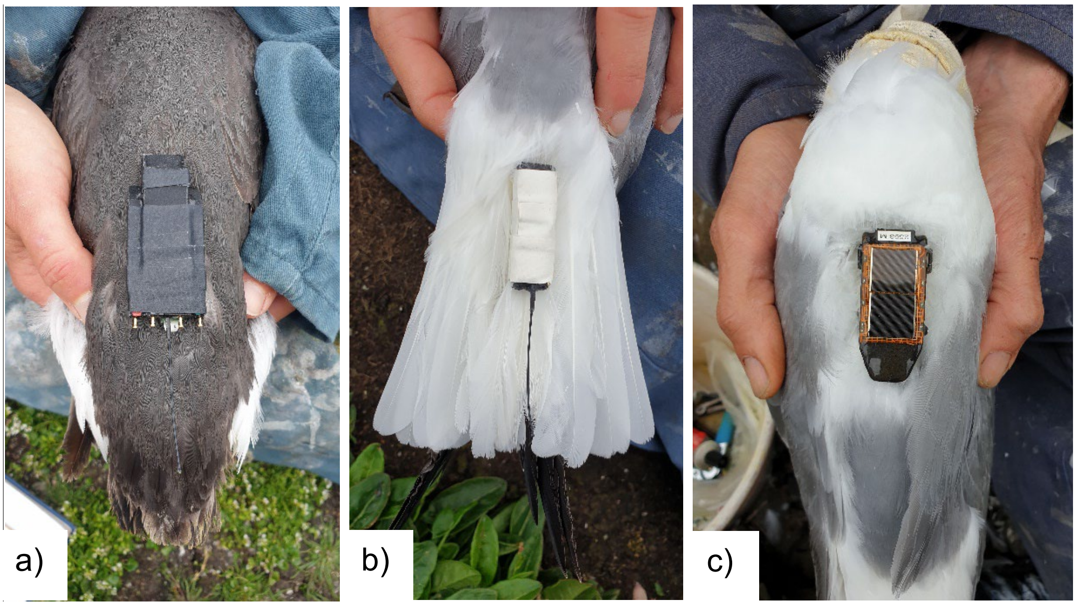


Figure 2. Completed deployment of data logger on a) guillemot, b) kittiwake (Pathtrack) and c) kittiwake (UvA).

As in previous years, data from the Pathtrack loggers were automatically downloaded to fixed base stations positioned in line of sight of nest sites each time the logger was within range (Bogdanova et al. 2020). Across all colonies and species, the base stations successfully received data from 146/154 loggers (94.8%). The data stored in the base stations were then downloaded daily onto a computer. Data from the UvA loggers were transmitted via several relay stations back to a fixed base station (Bogdanova et al. 2021). Across all colonies, data were successfully received from 39/40 of these loggers (97.5%).

The GPS sampling interval was set at 5 min for guillemot and razorbill and 10 min for puffin to maximise deployment duration while retaining sufficient resolution to estimate behaviours. Temperature, pressure and depth data for guillemots and razorbills were recorded every 4 sec when depth was below 1 m. In kittiwakes, the sampling interval was 5 min for Pathtrack loggers but varied in UvA loggers: 16 sec (Fowlsheugh, St Abb's Head) or 64 sec (Isle of May) when the bird was within a wind farm footprint, 5 min when it was at sea but outside the footprints and 15 min when it was at the colony. Triaxial acceleration was recorded at 19.8Hz.

The average length of deployments on auks was 3.8 days for guillemot (range: 0.9, 10.1), 5.0 days for razorbill (range: 2.1, 10.3) and 5.5 days for puffin (range: 2.1, 9.3). Across the three colonies average deployment length in kittiwakes was 5.2 days for Pathtrack (range: 0.2, 10.0) and 15.4 days for UvA loggers (range: 0.4, 39.3).

2.1.2 Data processing

2.1.2.1 Initial processing and trip identification

The data processing involved several steps. First, the raw data were cleaned by removing GPS fixes recorded before the loggers were fitted to the birds, duplicate records and fixes with low accuracy (where signal from ≤ 4 satellites was received or where speed between subsequent fixes was too high to be biologically plausible). Second, locations recorded at the colony (within 500m of the nest site) were also removed from the data set as we were interested in the birds' behaviour and distribution at sea. The remaining tracking data from each colony was split into distinct foraging trips using the track2kba package in the R environment (Beal et al. 2021). Trips were defined as periods of ≥ 30 minutes away from the colony. These thresholds were based on prior knowledge of species biology at the sites studied, and typical GPS sampling rate. In some cases, particularly when using UvA loggers we observed large time gaps between recorded GPS fixes. Therefore, prior to trip identification we segmented tracking data using a value of 60 minutes to define when a large gap occurred between successive fixes. This value was chosen as a compromise between maximising records of complete foraging trips, whilst minimising the use of data during larger data gaps where bird location and behaviour are increasingly uncertain (Fig. 3). The same value was used for all species for consistency. Trip identification was conducted only within segments of continuous tracking data meaning that the maximum time between locations within a trip was 60 minutes. Consequently, we occasionally identified foraging trips without a defined start or end due to gaps in GPS coverage and missing fixes.

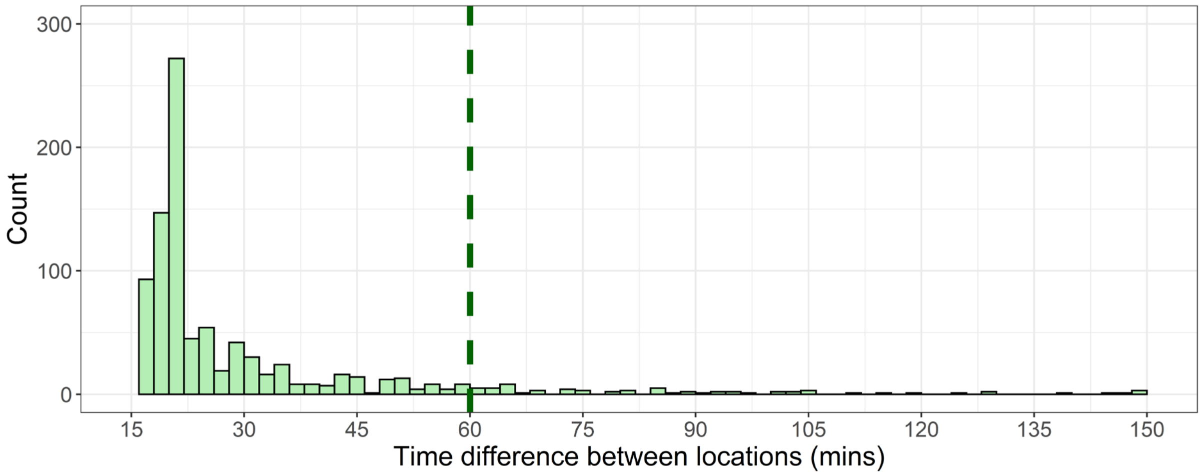


Figure 3. Data frequency of recorded kittiwake GPS locations, for locations where the time gap between locations was between 15 minutes and 150 minutes (2.5 hours). Shown are the time differences between at-sea locations only (histogram) and chosen maximum allowable time gap for segments (dashed line).

2.1.2.2 Behavioural classification

At-sea GPS data were then categorised in three key behaviours: commuting flight, foraging and resting at sea (except in puffins where separating foraging from resting is more challenging) as the potential impacts of offshore wind farms on seabirds are likely to differ during flight (when collision and barrier effects are expected to be more important; Desholm & Kahlert 2005; Searle et al. 2014, 2018) and during foraging and resting (when displacement is expected to be more relevant; Masden et al. 2010; Searle et al. 2014, 2018).

The behavioural classification procedure differed between species, depending on whether additional bio-logging data were obtained.

Behavioural classification using time-depth and speed data

For guillemot and razorbill fixes away from the colony that were within 150 seconds of a diving record (as indicated by TDR data) were classified as 'foraging'; the remaining fixes were classified based on distribution of speeds between subsequent fixes as 'commuting' (if speed exceeded 5.6 m/s) and as 'resting at sea' (if speed was less than 5.6 m/s).

Behavioural classification using hidden Markov Models

For puffins and kittiwakes we used hidden Markov models (HMMs; Zucchini et al. 2016) to classify recorded locations into different behavioural states. In general, HMMs can be used to cluster measurements of animal movement data into a discrete number of groups ("states"), which may represent different types of behavioural activity (Langrock et al., 2012). Details are provided below.

1) General methodology

For each species, data were analysed using each of the individual movement segments determined in the initial processing, and therefore included return foraging trips as well as partially recorded trips. A minimum of three consecutive recorded locations was required to estimate behavioural state (and calculate the corresponding step length and turn angles), and so movement segments containing less than three locations were removed. For kittiwakes, locations received from each of the three colonies were combined into a single dataset. For puffins, locations which were recorded using less than five satellites were removed, as it was observed that these locations (~1% of the data) often had higher error in recorded locations and measurement error was not explicitly accounted for in the movement models. For kittiwakes, locations using less than five satellites were not visually determined to have higher error, and so were not removed from the data. As HMMs model data in discrete time, the raw tracking data were interpolated so that locations were regularly spaced in time. Locations were linearly interpolated within each complete segment of movement data, ensuring that previously identified gaps between segments were maintained (i.e. gaps in locations of >1 hour were not interpolated). Locations were interpolated to a 5-minute resolution for kittiwakes (as 94% of raw at-sea locations within segments were <6 minutes in frequency), and a 12-minute resolution for puffins (as 91% of raw at-sea locations within segments were <13 minutes in frequency) (Fig. 4).

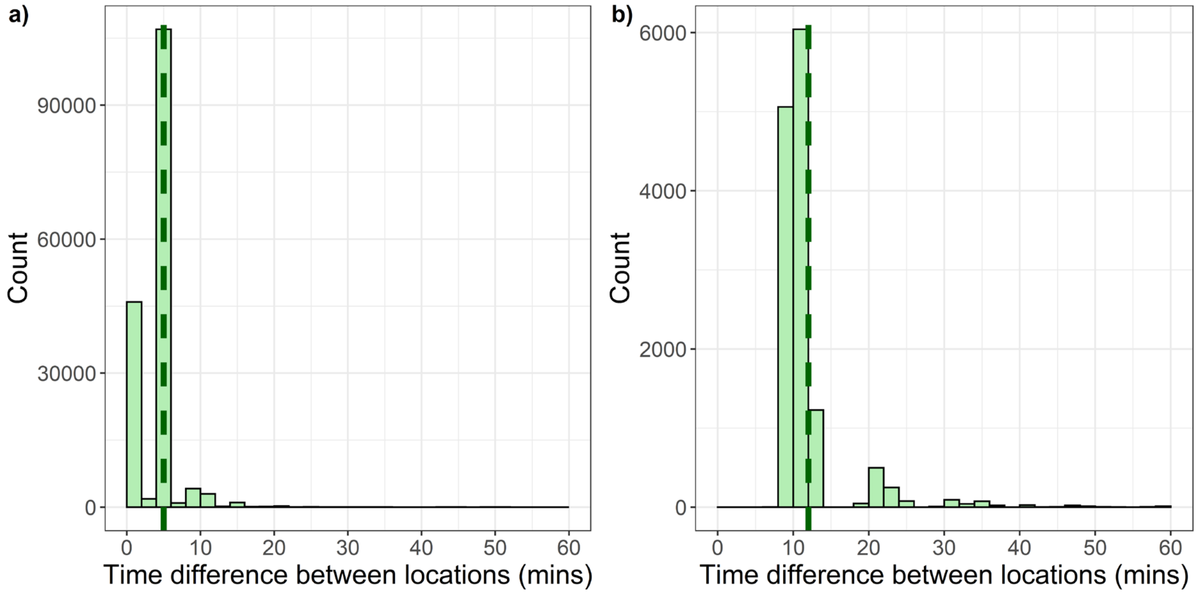


Figure 4: Data frequency of recorded GPS locations for (a) kittiwakes and (b) puffins. Shown are the time differences between at-sea locations within segments (histogram), with the chosen resolution for interpolation: 5 min for kittiwake and 12 min for puffin (dashed line).

After interpolation, the step length (metres) and turn angle (degrees) between interpolated locations were calculated. These two metrics were used as inputs in the HMMs, where step length was assumed to be Gamma distributed, and turn angle was assumed to follow a von Mises distribution with a mean turn angle of zero. By assuming a mean turn angle of zero, the aspect of movement being modelled is how variable the animal's heading is, irrespective of the absolute heading direction (e.g. Grecian et al., 2018). For each species, models were fitted to all individuals at the same time (termed "complete pooling"; Langrock et al., 2012); this approach assumes that model parameters are identical across individuals but ensures that the behavioural states identified are comparable across individuals and colonies. The different HMMs applied to each species are detailed in the following sections.

All models were evaluated by examining the predicted movements tracks classified within each state, as well as inspection of QQ-plots and pseudo-residuals. AIC (Akaike Information Criterion) was used to compare the different fitted models within each species; a lower AIC (with a difference of >2) was considered to indicate a model with more support (Burnham and Anderson, 2002). Models were also compared by evaluating the decoded states estimated for each location from the different models. A variety of starting values for the models were tested, to ensure that the models were not sensitive to the starting values chosen. All models were fit using the R package "momentuHMM" (McClintock and Michelot, 2018).

2) Application to kittiwakes

For kittiwakes, two null models were fitted: a two-state model and a three-state model. Both models produced distinct states which were biologically interpretable. As the three-state model provided more detailed inference on behavioural classification and was favoured by AIC (ΔAIC =71019), the three-state model was chosen for use in further modelling. Following this, a suite of three-state HMMs were fitted with different covariates influencing the probability of birds switching between behavioural states. Three different covariates were considered: (1) logger type deployed on bird (Pathtrack or UvA; categorical covariate); (2) colony of bird (Fowlsheugh, Isle of May or St Abb's; categorical covariate); and (3) hour of the day (as a continuous cyclical smooth; for more information see Appendix 2). All combinations of covariates were fitted as main effects, as well as a null model with no covariates. Models were compared by AIC, and the Viterbi algorithm was used to estimate the most-likely behavioural states at a 5-minute resolution using the best-fitting model.

All of the HMMs identified the same three distinct states. State 1 consisted of large step lengths and concentrated turn angles (Table 2, Fig. 5). These high speeds and directed movements, often seen while birds were travelling to and from the colony (Fig. 6), suggests that this state most likely represents commuting behaviour. State 2 consisted of intermediate step lengths and wide turning angles (Table 2, Fig. 5). State 2 was mostly observed once birds had travelled further out to sea and resembles area-restricted search patterns, suggesting that this state may represent foraging behaviour (Fig. 6). State 3 consisted of small step lengths and highly concentrated turn angles. Locations classified as state 3 were often closely associated (temporally and spatially) with state 2 (Fig. 6), and transitions between state 1 and 3 were rare (<1% in either direction, Table S2). State 3 most likely represents resting behaviour, where birds remain relatively stationary at the sea surface but may drift in a given direction with water currents. Overall, these three states are consistent with the behaviours observed in previous studies of kittiwake at-sea activity budgets (Chivers et al., 2012; Christensen-Dalsgaard et al., 2018; Daunt et al., 2002).

State	Step length (mean (SD))	Turn angle (concentration parameter)	Proportion of locations in each state (at-sea)	Possible biological interpretation
1	2,859 m (SD=977)	14.9	0.27	Commuting
2	543 m (SD=658)	0.8	0.56	Foraging
3	104 m (SD=55)	34.4	0.17	Resting

Table 2: Estimated state-dependent parameter values, proportion of locations classified as being in each state, and the likely biological interpretation of each state for kittiwakes. Values are from the full model. Step length was assumed to be Gamma distributed, and turn angle was assumed to follow a von Mises distribution with a mean of 0.

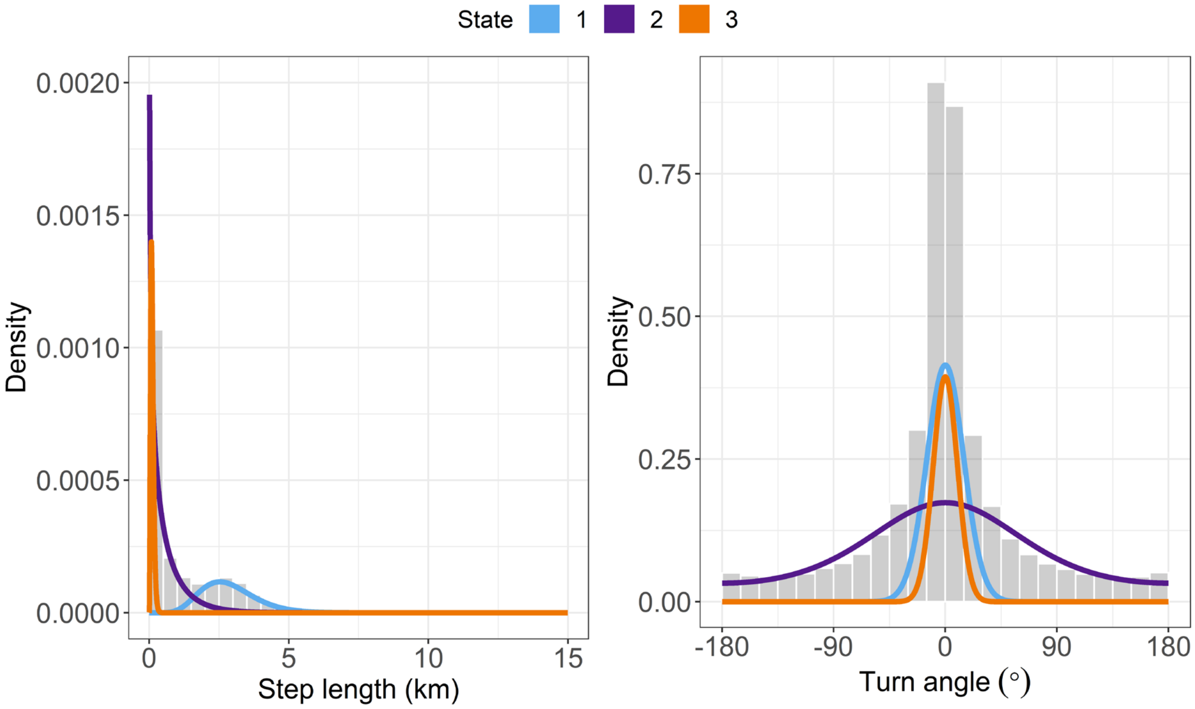


Figure 5: Fitted state-dependent distributions of step length and turn angle for kittiwakes. Shown are the estimates from a three-state HMM (full model).

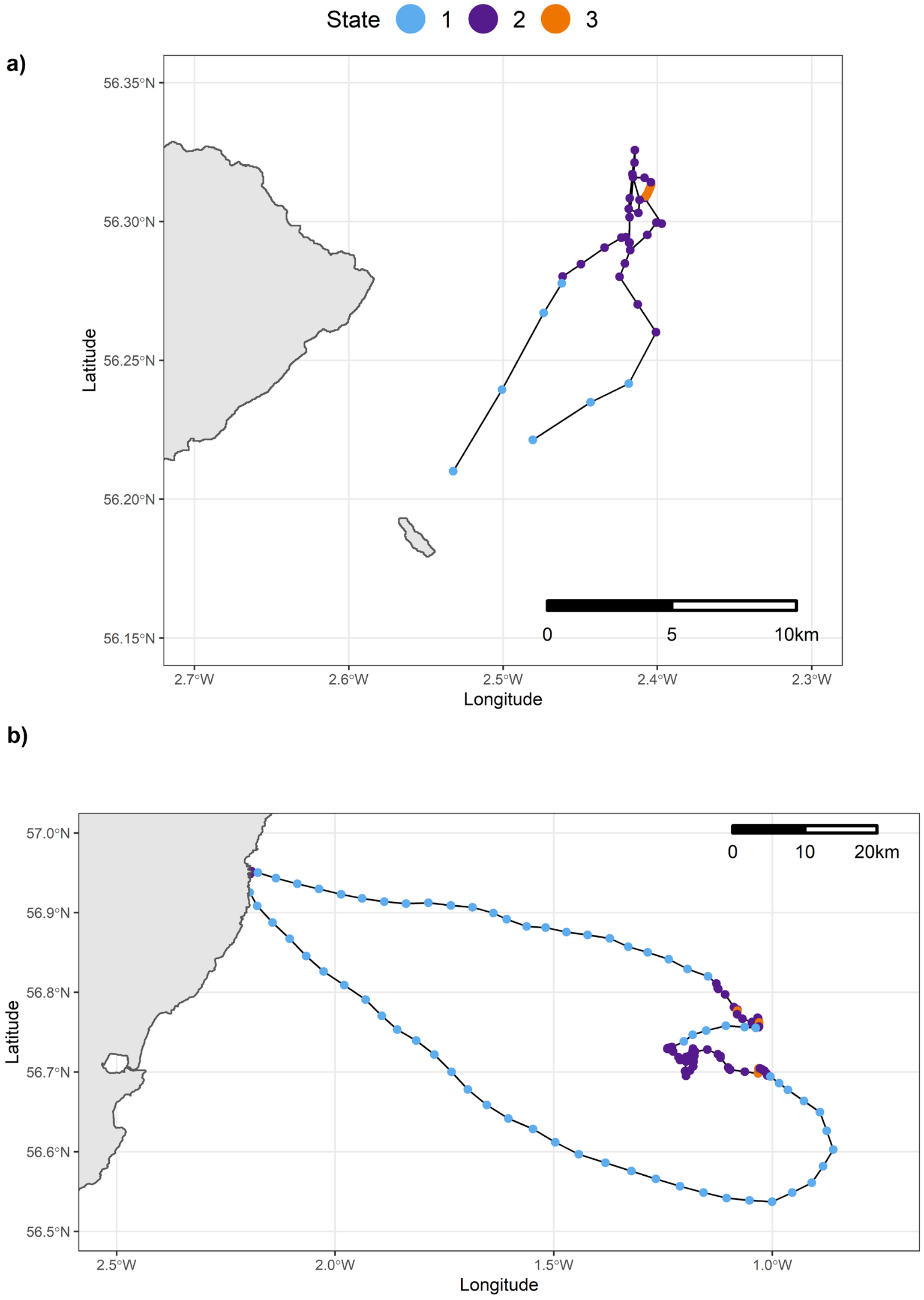


Figure 6: Two example tracks of kittiwakes with estimated behavioural states. The possible biological interpretations of each state are (1) commuting, (2) foraging, and (3) resting. Track (a) is from a tagged individual from the Isle of May; Track (b) shows a tagged individual from Fowlsheugh.

Comparison across the fitted models showed AIC support for the inclusion of all the candidate variables in influencing the transitions between behavioural states (the "full model"; Table S2). This suggests that there was evidence of differences between tag types and colonies in behaviour-switching (Fig. 7a,b). In particular, hour of the day appeared to be strongly favoured in model selection (Table S2), and the occurrence of different behaviours showed a diurnal pattern (Fig. 7c). Resting behaviour (state 3) was less likely to occur during the day and was more likely at night, which is consistent with the findings of Daunt et al. (2002).

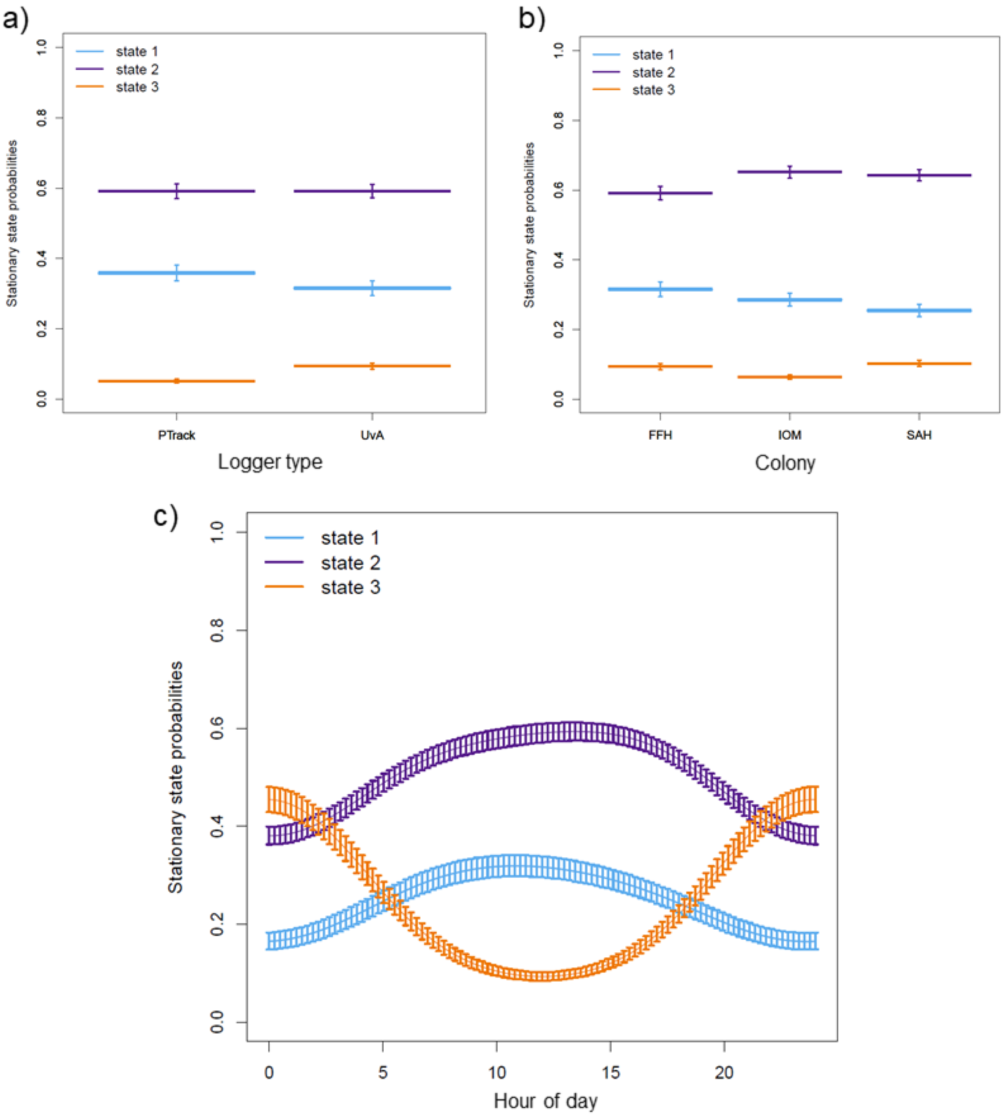


Figure 7: Probability of occupying each behavioural state (with 95% CI) as a function of (a) logger type, (b) colony and (c) hour of the day. Shown are the estimated stationary state probabilities from the full three-state HMM for kittiwakes.

All the fitted models converged and consistently identified the same three states; however, there was evidence of residual autocorrelation in the step lengths within each model. Thinning data inevitably results in data loss, and so regularising locations to a particular time interval is a balance between preserving data integrity and removing autocorrelation to an acceptable level. Further investigation may be required to examine how this autocorrelation could affect the inference made from these models. As behavioural classifications may be used differently in future stages of analysis, predictions of the most-likely behavioural state (and associated uncertainty) were made using both the best-fitting model ("full model") and the null model (Table S2). More than 99% of locations were estimated to be of the same behavioural state from either model, suggesting that the covariates had a limited effect on the resulting estimated behavioural state.

3) Application to puffins

For puffins, two initial null models were fitted: a two-state model and a three-state model. In the two-state model, state 1 consisted of large step lengths and wide turning angles (Table 3, Fig. 8). The high speeds observed, often seen when birds were travelling to and from the colony (Fig. 10a), suggest that this state most likely represents flight behaviour. State 2 consisted of small step lengths and highly concentrated turn angles (Table 3, Fig. 8). State 2 most likely represents non-flight behaviours. In this state the GPS tracks show slow horizontal movement; however, birds may be sitting at the water surface or diving underwater. It is possible puffins could also be drifting along with currents within the Forth (e.g. Bennison et al., 2019). In the three-state model, state 3 was equivalent to state 2 within the two-state model (~98% match; Table S5) and likely continues to represent non-flight behaviours (Table 3, Fig. 9). In the three-state model, state 1 consisted of large step lengths and concentrated turn angles, and state 2 consisted of intermediate step lengths and wide turning angles (Table 3, Fig. 9). State 1 most likely represents commuting flight, where birds travel fast with directed movements to different destinations (Fig. 10b); however, the biological interpretation of state 2 is not as clear. State 2 was often, but not always, associated (temporally and spatially) with state 3 (Fig. 10b). Birds were more likely to switch to state 2 than to switch between states 1 and 3 directly (Table S4), suggesting that state 2 is potentially a hybrid of behaviours that occurs between directed flight (state 1) and non-flight (state 3). It could potentially represent the transition period between flight and non-flight; however, further work is required to have more confidence in this biological interpretation. It should be noted that the behavioural classification of puffins presented here is an initial analysis, with a limited scope, and that further work is required. Given the challenges in biologically interpreting the observed movement behaviours, and the limited difference in behavioural classification observed in the kittiwake analysis when covariates were included, the effects of covariates were not examined in this initial analysis.

State	Step length (mean (SD))	Turn angle (concentration parameter)	Proportion of locations in each state (at-sea)	Possible biological interpretation
Two-state model:				
1	2,008 m (SD=2,887)	1.4	0.35	Flight
2	221 m (SD=104)	14.3	0.65	Non-flight
Three-state model:				
1	5,672 m (SD=4,641)	12.4	0.13	Commuting flight
2	582 m (SD=716)	0.7	0.21	Hybrid behaviour
3	220 m (SD=109)	14.4	0.66	Non-flight

Table 3: Estimated state-dependent parameter values, proportion of locations classified as being in each state, and the likely biological interpretation of each state for puffins. Values are shown for both the two-state and three-state models. Step length was assumed to be Gamma distributed, and turn angle assumed to follow a von Mises distribution (mean = 0).

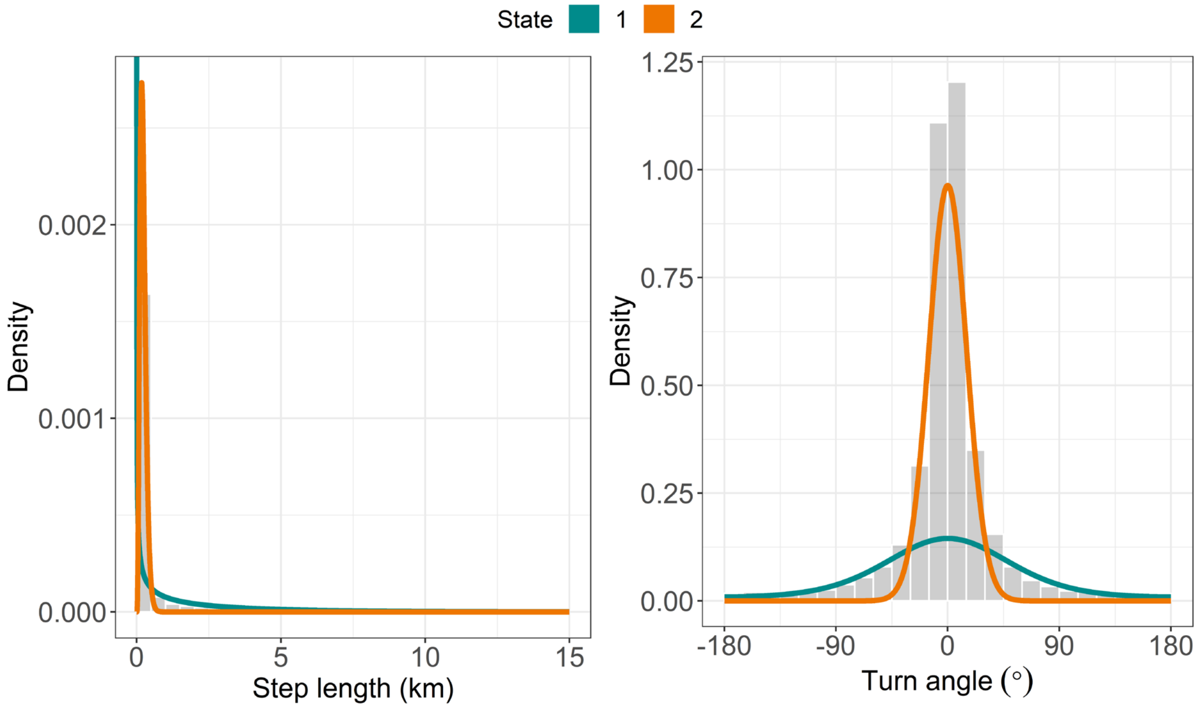


Figure 8: Fitted state-dependent distributions of step length and turn angle for puffins. Shown are the estimates from a two-state HMM.

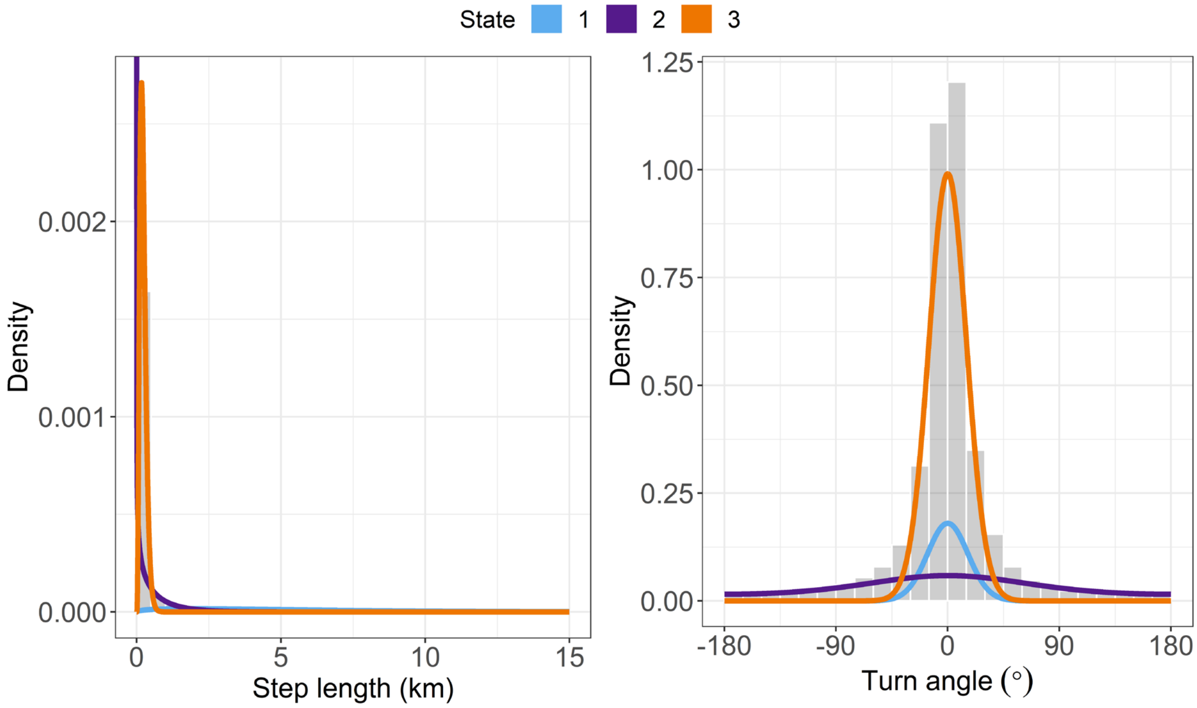
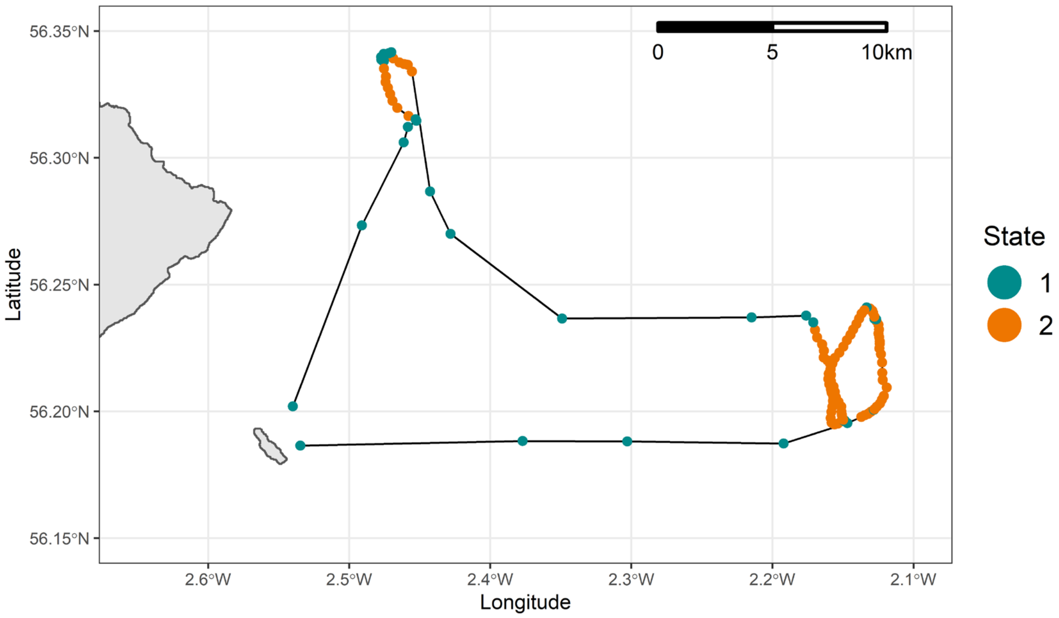


Figure 9: Fitted state-dependent distributions of step length and turn angle for puffins. Shown are the estimates from a three-state HMM.

a)



b)

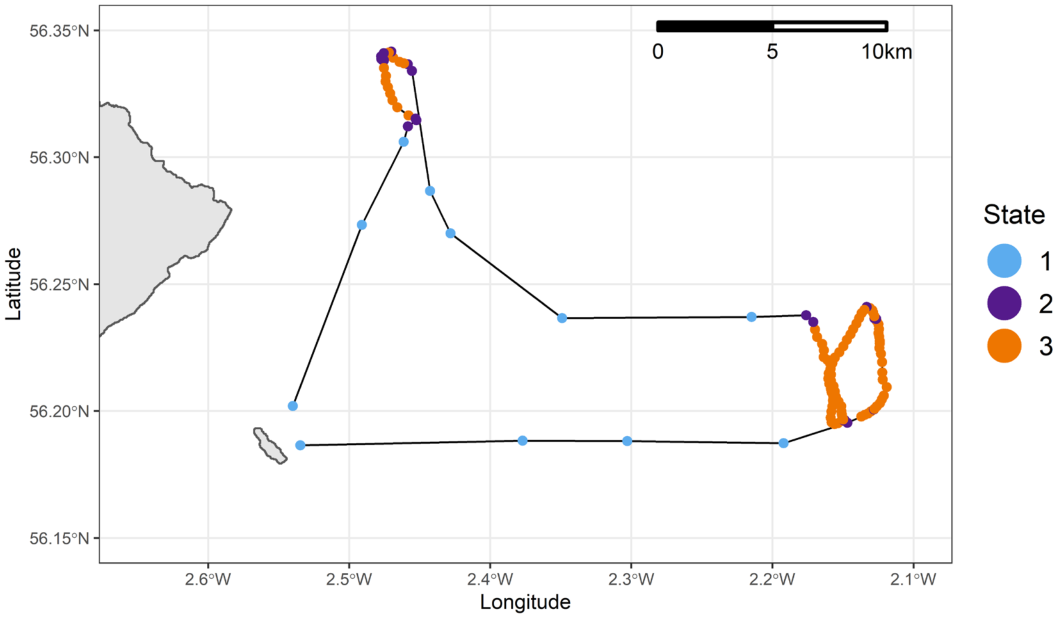


Figure 10: An example puffin track from the Isle of May with estimated behavioural states. The same track is shown in both panels, for comparison between the estimated behavioural states of the (a) two-state HMM and (b) three-state HMM. The possible biological interpretations of each state in the two-state HMM are (1) commuting, (2) non-commuting (foraging and resting); interpretation of states in the three-state model is uncertain at this stage.

All of the fitted models converged; however, there was some evidence of residual correlation in the step lengths within each model. The Viterbi algorithm was used to estimate the most-likely behavioural states (and associated uncertainty) at a 12-minute resolution, using both the two-state and three-state model. Due to challenges in interpreting the biological function of some of the modelled states (in particular state 2 of the three-state model), predictions from the two-state model were chosen for use in further analyses.

2.1.2.3 Flight altitude in kittiwake

A sample of kittiwakes at all three colonies were tracked using multi-sensor loggers which incorporated GPS units, altimeters and accelerometers (UvA loggers). Accelerometer data provided a means of classifying bird behaviour and crucially distinguishing periods when birds were floating on the water for the calibration of altimeters. Consequently, we were able to derive simultaneous estimates of kittiwake flight height from both GPS signals and altimeter pressure readings allowing for a direct comparison of these measures.

GPS measures of altitude

UvA loggers recorded a direct measure of altitude based on in-built GPS sensors. Altitude was recorded in two ways by the logger. Firstly, altitude measures were recorded as the altitude above a geoid approximating global mean sea level. Secondly, altitude above ground level was reported as the GPS altitude minus the terrain elevation based on a Digital Elevation Model (DEM). However, altitude above sea-level and altitude above ground level only differed when birds were on land. Therefore, for a project focussing on seabird movements at sea the same altitude values were reported in either case. To aid with the assessment of measurement error the logger also reported the number of satellites used to fix each GPS observation, the positional DOP (3-d DOP, which combines both horizontal and vertical DOP) and separate measures of both horizontal and vertical accuracy. DOP is generally thought to be higher in the vertical dimension than the horizontal because all the satellites that provide a signal are above the receiver (Langley 1999, Ladetto et al. 2000). Ideally, GPS coordinates (including altitude) would be based on signals from multiple satellites spaced evenly apart. In addition, faster GPS sampling rates may increase the precision (reduce DOP) of location estimates due to the greater time GPS sensors are switched on, increasing the number of satellites available for a fix. In particular, the UvA GPS module is powered constantly up to a sampling frequency of 16 seconds (Thaxter et al. 2018). Consequently, GPS location estimates, including altitude, are expected to be most precise when sampling at intervals of 16 seconds or less. Knowledge of the DOP or other measures of GPS error can also be used to model error-generating processes using techniques such as state-space modelling (e.g. Ross-Smith et al. 2016, Peron et al. 2020). In this way an estimate of uncertainty around altitude estimates can be made and errors associated with negative values or poor DOP can be incorporated within the modelling process.

Barometric pressure and altitude

UvA loggers recorded a measure of air pressure (Pascals) associated with every GPS fix. Air pressure values recorded by UvA loggers are subject to some initial processing by the logger itself. Specifically, air pressure was sampled at 10 Hz in 3 second bursts generating approximately 60 samples of pressure per GPS fix. Pressure data was then filtered by removing the minimum and maximum pressure values before calculating the median value across all remaining samples.

We used the barometric formula (Berberan-Santos et al. 1997; Wallace & Hobbs 2006) to estimate height *z* (m) above sea level:

$$z = -\frac{kT}{mg}\ln\left(\frac{P}{P_0}\right), \tag{1}$$

where P_0 and P are the atmospheric pressures (Pascals) at sea level and at height z (m), respectively; k is the universal gas constant for air ($8{\cdot}31432$ N m mol^{-1} K^{-1}); m is the molar mass of air ($0{\cdot}0289644$ kg mol^{-1}); g is the acceleration due to gravity (m s^{-1}); and T is the temperature (K) of the atmospheric layer between z_0 and z. The accuracy of altitude estimates depends on how P_0 is defined and the accuracy with which it is measured along with the accuracy of measures of P from a barometric sensor (for more details see Section 7.2, Appendix 1). UvA-bits loggers provide us with direct measures of P and T required for equation 1 but an estimate of the reference sea-level pressure (P_0) is also required. The average pressure at mean sea-level in the International Standard Atmosphere (ISA) is 1013.25 hPa). However, in practice pressure drifts constantly due to local weather conditions and even atmospheric tide fluctuations, making a single, fixed value of P_0 unsuitable. There are multiple ways in which P_0 could be assessed including using data from bird-borne barometric loggers or relying upon weather stations or remote sensing data. These approaches are likely to have different pros and cons and the exact altitude estimated may vary across methods. Therefore, we estimated P_0 using four different methods: 1) Estimating P_0 from periods within a single logger deployment when the bird was judged to be floating on the sea; 2) Estimating P_0 from periods when birds were adjudged to be floating on the sea but incorporating data from all loggers deployed during the study; 3) using ERA5 remote sensing data on mean sea-level pressure (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels, accessed: 23/03/2022); 4) using weather data obtained from a floating offshore buoy collected in 2020.

Estimating P_0 within a single logger deployment

To estimate suitable values of P_0 during a single deployment we identified periods when birds were floating on the water using accelerometery data. Accelerometery data were processed using existing UvA algorithms that have previously been used for tracking gull species and can identify a range of behavioural categories (W. Bouten, pers. Comms.). Using this approach, each GPS fix was assigned to a specific behavioural category. For the purposes of the current work, we are only interested in identifying periods when birds were floating on the sea for calibration purposes. Therefore, although it may be possible with accelerometery data to distinguish different flight modes (e.g. soaring versus flapping flight) this is not pursued further here.

Once periods when birds were floating on the water were identified we then calculated a unique reference sea-level pressure for each GPS location fix during a given deployment. To do so we selected all observations within a 1-hour time window centred on a focal GPS fix and selected those observations which were classified as 'floating'. Provided at least 5 observations of the bird floating on the water were recorded within the 1-hour time window we then calculated a value for P_0. P_0 was calculated as the mean pressure value across these observations with greater weighting given to observations that were closer in time to the focal GPS fix. The exact weighting scheme we used was 1 / (distance in time to focal GPS fix)^2, which represents an inverse-distance weighting (IDW) scheme. IDW assumes that closer values are more related than values further apart. If there were not at least 5 floating observations within the specified 1-hour time window we increased the window size to 2 hours, 3 hours, 4 hours up until 12 hours. If there not sufficient observations within the 12-hour window, then no estimate of altitude was attempted.

Estimating P_0 across all logger deployments

We used the same approach as outlined above to estimate P_0 when using data from all UvA logger deployments. However, when identifying floating observations within a specified time window around a focal GPS fix, we searched across all deployed loggers simultaneously. In theory, this should provide more observations of birds floating for calibration, but we also had to account for the distance in space between observations from different loggers. Therefore, we excluded all observations greater than 5 km from the focal observation when calculating P_0. As before, if the number of instances of birds recorded as floating within the initial 1-hour time window surrounding the focal GPS was less than five then we gradually increased the length of this time window until this threshold was met up to a maximum window size of 12 hours. A reference value for P_0 was then calculated as the average pressure across all observations of floating weighted by the distance in time and the distance in space from the focal GPS fix. We used a spatio-temporal inverse distance weighting scheme calculated via the idwST() function from the geosptdb R package (Melo et al. 2022). Using this approach weights are defined as:

$$\lambda_i = \frac{d_{i0}^{-P_w}}{\sum_{i-1}^{n} d_{i0}^{-P_w}} \qquad (2)$$

Where the weight is controlled by a factor (P_w) with each increment of distance set at a value of P_w = 2. d_{i0} is the distance between a focal position and each of its neighbours. Note that as we restricted calculations to observations within a 5 km radius of a focal point the weight of any observations outwith this radius is 0. The spatio-temporal distance between points is calculated as:

(3)

$$d_{i0} = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + C \times (t_i - t_0)^2}$$

Where x, y and t correspond to spatio-temporal coordinates. The exact scaling factor of C was determined using cross-validation to determine which out of a set of candidate C values ranging from 0.5 to 3 at intervals of 0.05 minimized the root-mean-square prediction error. Ultimately, C was set at 1.025 at St Abb’s Head, 1.01 at Fowlsheugh and 1,015 at the Isle of May.

Estimating P_0 using ERA5 data

ERA5 provides hourly data on sea-level pressure across the globe at a spatial scale of 30 km^2. This provides relatively fine-grained temporal information on pressure, though at a coarser spatial scale then could be achieved using logger data. We downloaded hourly data from ERA5 to cover the period and spatial extent covered by the available tracking data. To examine the suitability of the downloaded ERA5 data we compared estimates of surface pressure derived from ERA5 over the period May – July 2020 with simultaneous data obtained from a weather buoy that was deployed in the Firth of Forth (56° 17' 41.5896'' N, 1° 26' 31.4196'' W) and recording pressure during the same period. Overall, pressure from ERA5 closely tracked pressure values recorded by the weather buoy over time and pressure values were highly correlated (Fig. 11). However, while initial comparison of pressure and altitude estimates based on ERA5 data or UvA logger processing showed they were highly correlated, ERA5 data often appeared to be offset from UvA logger estimates which resulted in nonsensical estimates of flight height (i.e. negative numbers of a large magnitude). Closer inspection revealed this issue was most prevalent in spatial tiles around the coast which were not 100% covered by water. In spatial tiles located purely offshore the mismatch of ERA5 and UvA logger flight height estimates showed much less discrepancy. It should be noted therefore that the comparison between ERA5 data and weather buoy data in 2020 was based on an ERA5 tile located entirely offshore.

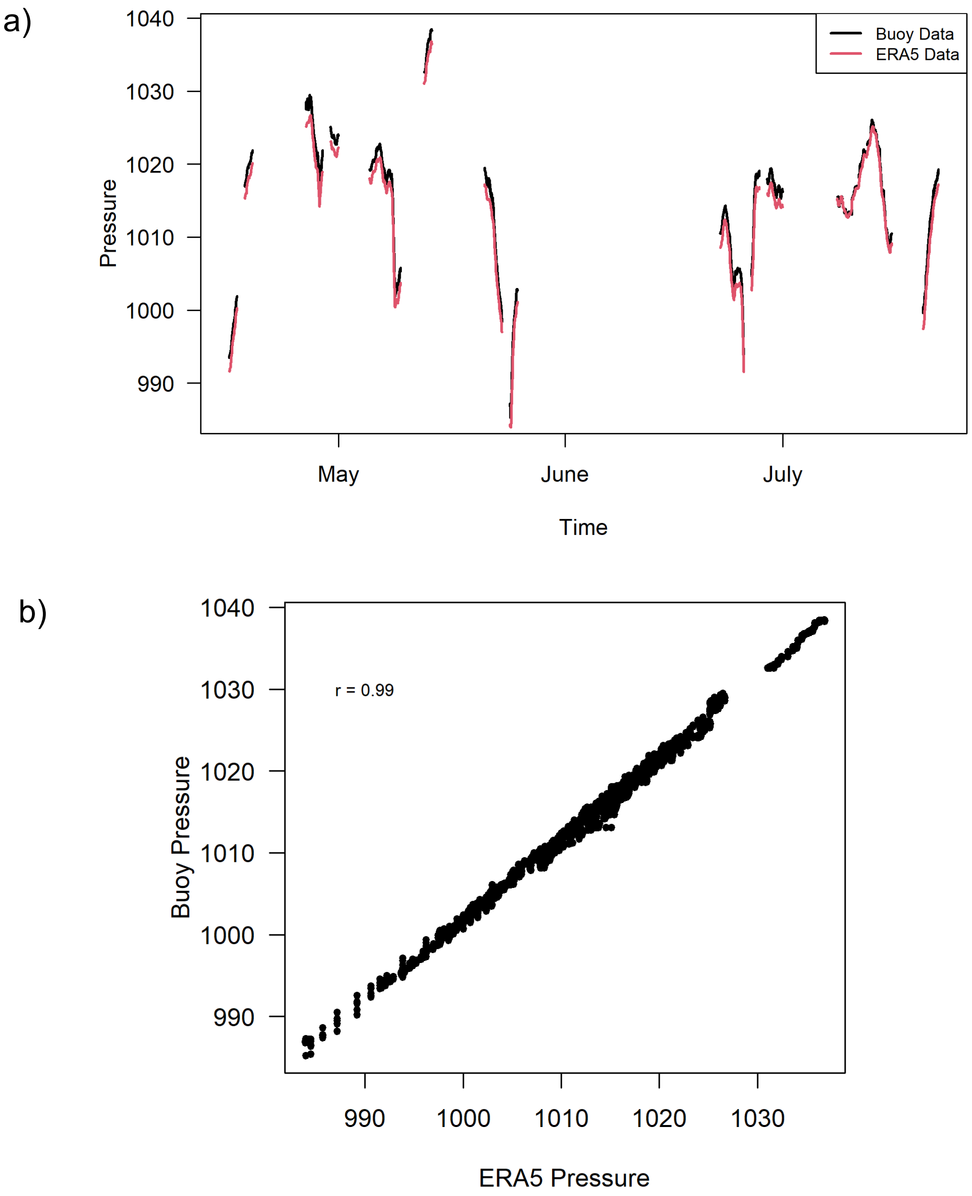


Figure 11: a) Relationship between surface pressure obtained from ERA5 and that recorded by a weather buoy in the Firth of Forth during 2020. Note gaps in the time-series below reflect gaps in the data available from the weather buoy. b) Correlation between measured surface pressure from ERA data and weather buoy data.

To fully utilise the ERA5 data, we sought to calibrate pressure estimates from ERA5 with estimates from UvA loggers. To do this we first subsetted the UvA tracking dataset and selected only those observations where birds were classified as floating on the sea. We then used generalised additive models (GAM) to build a model in which UvA-logger derived pressure was the response variable and ERA5 derived pressure was a covariate fitted using a cubic shrinkage spline. In these models we also included latitude and longitude as a 2-dimensional thin-plate spline and included an additional smoother for the day of the year. The outputs from these models were then used to predict a reference sea-level pressure based on ERA5 data but calibrated to the sea-level pressure readings recorded by UvA loggers in the relevant spatio-temporal neighbourhood for each GPS fix. This approach meant we could combine sea-level pressure data from ERA5 with pressure data recorded by UvA loggers to estimate flight altitude.

Estimating P_0 using a weather buoy

As well as being used to validate the use of ERA5 data information from weather buoys could also be used to directly assess sea-level pressure in the surrounding area. While the current work is largely focussed upon tracking data from 2021, tracking data using UvA loggers with both barometric and GPS sensors was available for kittiwakes from the Isle of May in 2020 for which buoy data was available. Occasionally birds from this colony do forage in the vicinity of the previously mentioned weather buoy. Consequently, for some observations it is possible to compare altitude estimates based on GPS and barometric pressure readings from UvA loggers with those derived from a suitable weather buoy. To achieve this, we first identified tracking observations that were close in space and time to pressure readings obtained from the weather buoy. We used a threshold distance of 10 km from the weather buoy and a threshold time of 1 hour from the nearest pressure observation recorded by the buoy. While a range of different thresholds could be used, we found that shorter thresholds quickly reduced the number of tracking observations available for comparisons as the buoy is towards the edge of Isle of May foraging ranges and had temporal gaps in recording. We then calculated altitude using equation 1 above, with the reference sea-level pressure set by pressure recorded by the weather buoy. We then compared altitude estimates based on using weather buoy data to set P_0 to those obtained via GPS or using the single logger calibration approach outlined above.

2.1.3 Data analysis

2.1.3.1 Species utilisation distribution (UD)

Utilisation distribution at sea was determined for each species by calculating the kernel density of locations recorded away from the colony. Locations were projected in Lambert azimuthal equal-area projection and bivariate normal kernel was calculated in R (R development core team, 2022; package adehabitatHR, Calenge 2006), using a cell size of $500m^2$ and a smoothing parameter *h* identified with the *ad hoc* (reference bandwidth) method (Worton 1989). For each species, density contours (50, 70 and 90%, the former representing the core area used, the latter – the overall area used) were extracted in R (package adehabitatHR) and mapped in ArcGIS 10.7.1 (ESRI). For guillemot, razorbill and kittiwake, separate maps were generated based on all at-sea locations and on foraging and resting locations. For puffin, maps were generated based on all at-sea locations and on non-flight locations (representing foraging and resting behaviours).

2.1.3.2 Horizontal flight trajectories

Individual commuting flights were extracted from the processed tracking dataset and horizontal flight trajectories were mapped in ArcGIS 10.7.1. On the maps, breaks in the lines at sea represent periods when the birds were engaged in non-flight behaviours. Also, due to the relatively large GPS sampling intervals only single locations were recorded for some of the shorter commuting flights. These are not shown in the flight line maps but were included in calculations of number of flights passing through the wind farm footprint.

2.1.3.3 Minimum adequate sample size

To establish whether the sample size of tracked individuals was adequate to estimate the at-sea area used by the population of each species during the sampling period, we examined the relationship between overall area used (area of the 90% UD contour) and number of individuals using a resampling procedure. This procedure was performed in R using a bespoke code, and involved creating 1,000 datasets for each sample size of birds, ranging from 1 to n (where n denotes the total number of birds for which we had data), by choosing birds randomly without replacement (Manly, 2009). Resampling without replacement was used to avoid systematic underestimation of the overall area used by the birds. A UD estimate was then derived from the pooled data from all individuals within each resample (using the adehabitatHR package within R) and the area of the 90% UD contour calculated. The distribution of these areas across the 1,000 resampled datasets was used to quantify the typical at-sea area used for a given sample size of birds and to provide a partial assessment of the uncertainty associated with estimating this area.

2.1.3.4 Overlap with Forth/Tay wind farm footprints

To quantify overlap between the core and overall areas used by each species and the Neart na Gaoithe, Seagreen and Berwick Bank wind farms, we calculated the proportion of 50% and 90% UD contours lying within each wind farm footprint. For guillemots, razorbills and kittiwakes, overlap was estimated separately for foraging and resting behaviours; for puffins overlap was estimated for non-flight behaviours combined. To assess the extent to which commuting birds travelled through the wind farms we calculated the proportion of birds and flights passing through each footprint. Overlap metrics (UD area overlap and % UD overlap with wind farm; % birds and % flights passing through wind farm) were extracted in R using the raster package (Hijmans 2019).

2.1.3.5 Flight altitude in kittiwake

Modelling of kittiwake altitude data using state-space models

Even after initial processing it is apparent that errors in estimated altitudes remain - whether using GPS- or barometer-based estimates - such as negative altitude values or occasional, extreme outliers. Therefore, recent recommendations (Peron et al. 2020) are that flight heights be analysed using a state-space framework that can model the changes in a state variable (here true flight height) even when that variable is imperfectly observed (recorded flight height). In this way we can model both a state process and an observation process. Crucially, it enables us to model factors known or suspected to alter the errors in our flight height estimates such as sampling frequency, quality of GPS fix etc. as part of the observation process model. In addition, the state model can be set-up in such a way that is must stay above a set value (such as 0 in the case of flight heights). As a result, when negative flight altitudes are recorded, the model recognizes such values are errors and can adjust accordingly. State-space models have been used previously for estimation of GPS-derived flight altitudes specifically within the context of windfarm impact assessments (Ross-Smith et al. 2016). Here, we construct separate state-spaces for estimates of flight height from GPS and barometric sensors. These models provide estimates of flight height at each point along a movement trajectory along with corresponding estimates of the standard deviation and upper and lower 95% CRI of such estimates.

State-space modelling of flight altitude from barometers

Flight heights observed using GPS or barometric sensors were recorded as irregular time-series (y_{t_j}) where t_j is an increasing sequence of observation times ($t_1<t_2<\ldots<t_n$) for each foraging trip and *n* denotes the number of observations in a given foraging trip. State process models of flight height were constructed using a truncated normal distribution with the lower value limited to value of -1. In theory, the distribution could be limited at 0, but the action of processes such as waves means that small negative values could reasonably expected hence the lower bound of -1. Models were fit in R (R version 4.1.2, R Development Core Team 2021) using the JAGS Bayesian statistics program (Plummer 2003) via the rjags package (Plummer 2021).

Our latent (expected) flight height process was written as:

$$z_{t_j} \sim Norm\left(\mu_{t_j}, \sigma_{sp}^2\right);\ min = -1 \tag{4}$$

$$\begin{aligned}\mu_{t_j} = \beta_1 \times z_{t_j - t_{j-1}} + \beta_2 \times Lag_{t_j} + \beta_3 \times Vertical\,Velocity_{t_j} + \beta_4 \times Sampling\,Frequency_{t_j} + \\ \beta_5 \times z_{t_j - t_{j-1}} \times Lag_t + \beta_6 \times Sampling\,Frequency_{t_j} \times Vertical\,Velocity_{t_j}\end{aligned} \tag{5}$$

where σ_{sp}^2 represents the variance in the state process part of the model. The mean of the flight height distribution at observation time t_j (μ_{t_j}) is estimated based on the estimated flight height in the preceding record ($z_{t_j - t_{j-1}}$), the difference in time (Lag_{t_j}) between observation times t_j and t_{j-1}. Following suggestions in Peron et al. (2020) we included a term for vertical velocity between times t_j and t_{j-1} in our model. Specifically, Peron et al. (2020) recommend including a term for vertical velocity at higher sampling frequencies (≤ 1 minute). As the sampling frequency used in the current study is irregular, we created a variable that identified whether sampling frequency was greater than 1 minute (1) or not (0) and included an interaction between vertical velocity and whether sampling frequency was ≤ 1 minute or not. For the same reason, we included an interaction between height measurements taken at observation time t_{j-1} and the length of the time lag between observation times t_j and t_{j-1}.

The observation part or our model was defined as

$$y_{t_j} \sim Norm(z_{t_j}, \sigma_{t_j}^2) \tag{6}$$

where y_{t_j} is the altitude recorded by the GPS logger at each observation time, t_j. The observed altitude was assumed to be normally distributed around the true altitude (z_{t_j}) with a variance of $\sigma_{t_j}^2$. Initially, the observation variance at time t_j was modelled on the log-scale as a function of time since the end of the most recent calibration period when the bird was on the water. Time since last calibration period was first fitted as a continuous variable, but this specification often led to problems with model convergence. To address this issue time, since calibration was fitted as a four-level factor with the levels being 0 – 1 hours since calibration, 1-2 hours since calibration, 2-3 hours since calibration and >3 hours since calibration. However, we found no clear evidence that time since the most recent calibration influenced the precision of altitude estimates so these terms were removed. Instead, we included a predictor denoting whether the GPS sampling frequency was ≤ 60 seconds as visual inspection of the data suggested less dispersion of altitude estimates at higher sampling frequencies and has also been reported previously in similar studies using barometers to estimate avian flight heights (Cleasby et al. 2015). Thus, our eventual model for the observation variance was

$$\log(\sigma_{t_j}^2) = a_{d,0} + \beta_{d,1} \times Sampling\ Frequency\ \leq 60\ seconds_{t_j} \tag{7}$$

Where α represents the model intercept and the subscript *d* denotes that coefficients refer to a model of the variance or dispersion of a normal distribution.

State-space modelling of flight height from GPS

State process models of flight height based on GPS estimates were constructed in a similar way to those for barometric estimates of flight height. We used the same equation for the state process model as described above in equation 5 using GPS estimates of flight altitude. However, for the observation process the variance at time t_j was modelled as a function of whether the sampling frequency was less than 16 seconds or not and as a function of the vertical accuracy recorded by the device at Fowlsheugh and St Abb's Head. For the Isle of May data we did not observe instances of sampling frequency less than 16 seconds so this term was not included for trips from this colony.

$$\log\left(\sigma_{t_j}^{2}\right) = a_{d,0} + \beta_{d,1} \times 16\, Second\ Sampling\ Frequency_{t_j} + \beta_{d,2} \times Vertical\ Accuracy_{t_j} \tag{8}$$

Note that a term indicating whether sampling frequency was 16 seconds or less was included because manufacturers' specifications suggested that GPS loggers would run continuously at this frequency, leading to the expectation that observation precision would be enhanced (Thaxter et al. 2018). Similarly, higher values of vertical accuracy denote greater uncertainty around estimates, resulting in lower precision of altitude estimates.

Bayesian Model Set-up

Initially we ran state-space models for multiple trips and individuals as part of a single modelling process for the Fowlsheugh, Isle of May and St Abb's Head 2021 datasets. However, issues surrounding computing time and memory meant that we instead ran separate models for each foraging trip for which we had sufficient data (at least 10 measures of altitude from both barometric pressure and GPS and at least one observation of the bird floating on the sea). To calculate population-level coefficients for each colony we therefore used a two-step approach. We averaged trip-level coefficients across all trips at a given colony weighting by the standard deviation associated with each parameter for a given trip.

State-space models were run for each foraging trip for an initial 30,000 MCMC iterations with a burn-in of 10000 for three separate chains. Model convergence was then established by visual inspection of MCMC trace plots and rank plots (Vehtari et al. 2021) to ensure MCMC chains had mixed. In instances where convergence had not occurred, further iterations were run until model convergence was attained. We used diffuse priors for all the coefficients in our model which were assumed to be normal distributed with a mean of 0 and a large variance. We used uniform priors for the variance of the state process (Equation 2) set with a lower limit of 0.1 and upper limit of 20.

2.2 Device effects

2.2.1 Data collection

2.2.1.1 Puffin

As in previous years, a large number of active burrows in the study area were marked prior to deployment using flags attached to small canes. These were then assigned to a treatment or control group. At treatment burrows one member of the pair was tagged with a GPS logger and a unique combination of colour rings or with a combination of colour rings only, whereas in control burrows birds were not captured. To further aid subsequent identification of individuals, a letter was drawn onto the Tesa tape used to attach the logger to each bird. Birds equipped with colour rings only were held for a similar amount of time as the GPS-tagged birds. The two treatment groups were included to investigate the potential negative effects of handling and device deployment. Due to the large number of burrows needed, two deployment sessions took place at two adjacent sites approximately 30m apart. This allowed us to avoid any burrow overlap between the deployments. The deployments took place towards the later part of the puffin breeding season when the chicks were older and more robust.

Only adult birds bringing fish into the burrow were caught at the burrow entrance using purse nets. The nets (size: *ca.* 50cm^2, mesh size *ca.* 4x4cm) were pegged down and laid over the entrance to the burrow. This method allowed us to catch incoming adults and instantly allocate them to their burrow, and also eliminated the risk of catching both members of the pair.

From the day following each deployment, feeding watches were undertaken for 8 hours a day (5 am to 1 pm) shared by five observers from a canvas hide set up in the study site prior to deploying the loggers. The watches were carried out for a period of 3 days after each deployment. Every feed into a marked burrow was recorded, and for treatment burrows, also which individual fed (tagged or partner). In addition, the behaviour of treatment birds was recorded (e.g. colony attendance, entering the burrow without fish). Observation dates and sample sizes of treatment and control burrows in each deployment session's feeding watches are summarised in Table 4.

Deployment session	Site	Deployment date	Observation dates	Logger	Colour ring only	Control
1	1	2 July	3 – 5 July	13	6	35
2	2	6 July	7 – 9 July	12	6	29

Table 4. Sample size of puffins within each treatment category and deployment session.

Due to the decrease observed in the provisioning rates of tagged birds in previous years (Bogdanova et al. 2020, 2021), supplementary feeding of puffin chicks with whitebait and seabird supplements took place from the 1st day after the feeding watches were completed (4th day after deployment) until fledging. Chicks were not fed during the first three days after deployment to ensure the provisioning rates of adults were not affected at the time when GPS data were being collected. Once feeding commenced, all chicks were given 50g of fish over three meals each day. Meals comprised of defrosted sprat in whitebait. The chicks were also given an 1/8th of a Aquaminivits tablet every second day to replace the nutrients lost from the freezing of the fish. All fish were dropped into the burrows and hand feeding only took place if a chick did not take the fish themselves. All treatment chicks and a subsample of control chicks were weighed and measured once every 4 days until they fledged to monitor their condition and survival. Body condition was calculated as weight/(wing length)^3 and condition values were scaled (multiplied by 1000) to avoid very small decimal numbers.

2.2.1.2 Kittiwake

At-colony behaviour

Within each colony, a sample of study nests that were observable from a distance without disturbance were selected. Observations of treatment and control nests focused on recording change-over rates of adults at the nest, from which a proxy of foraging trip duration was later calculated. Treatment nests had one member of the pair tagged with a Pathtrack or an UvA logger, whereas control nests had no birds tagged. The treatment birds were given a BTO metal ring, a GPS logger and were weighed and measured. Prior to release the birds were marked with yellow dye which is not known to have any negative effect on the birds but allows distinguishing them from their mates during observations. Observations were carried out from hides for three days, 8 hours a day (5am–1pm). On the Isle of May the observations commenced the day after deployment, at Fowlsheugh between day 3 and 5 after deployment and at St Abb's Head between day 1 and 10 after deployment. Observation dates and sample sizes per treatment category are shown in Table 5. Note that within each colony observations were conducted only at sites where both logger types were deployed which is why not all Pathtrack nests were observed.

Site	Deployment date	Observation dates	Pathtrack	UvA	Control
Isle of May					
Clett	28 June	29 June - 1 July	5	7	19
Colms	1 July	2 - 4 July	7	6	27
Fowlsheugh					
Turturra LL	1 - 3 July	6 - 8 July	5	9	7
Turturra LR	2 - 3 July	6 - 8 July	9	5	11
St Abb's Head					
Cleaver Bay 2	28 June - 4 July	8 - 10 July	6	7	8
Cleaver Bay 3	4 – 7 July	8 - 10 July	5	6	6

Table 5: Sample size of kittiwake nests within each colony, site and treatment category where observations of parent change-over rates were conducted.

During the observations, all events involving two adults at any nest were recorded (time the birds spent together at the nest, whether changeover occurred, identity of arriving and leaving bird – this was possible only at treatment nests, where the mates could be distinguished). A change-over was defined as any event in which the leaving bird at a specific nest was not the same as the arriving bird. The status of each nest (number of chicks) was recorded daily during the watches.

We also carried out checks of chick attendance at the study nests (at least one adult present or both parents absent). On the Isle of May, checks were conducted daily at all study sites, for 21 days starting 6-9 days after logger deployment. At Fowlsheugh attendance checks were also carried out at all sites, approximately every 3 days starting 5-7 days after deployment for a period ranging from 17 to 37 days. At St Abb's Head attendance was recorded at one site where only Pathtrack loggers were deployed. At this colony checks were carried out at variable intervals, starting 4-12 days after deployment, for 26-40 days. Observation dates and sample sizes per treatment category are shown in Table 6.

Site	Observation dates	Pathtrack	UvA	Control
Isle of May				
Clett	7 - 28 July	9	7	30
Colms	7 - 28 July	14	6	35
Fowlsheugh				
Turturra LL	26 June - 14 Aug	5	7	60
Turturra LR	26 June - 14 Aug	9	5	79
Trelung Ness	26 June - 14 Aug	12	0	60
St Abb's Head				
Cleaver Bay 2	NA	0	0	0
Cleaver Bay 3	NA	0	0	0
West in Thirle Bay	26 June - 10 Aug	10	0	63

Table 6: Sample size of kittiwake nests within each colony, site and treatment category where nest attendance observations were conducted.

On the Isle of May only, accessible chicks from treatment and control nests were weighed and measured (wing length) on the day of logger deployment and 7 days later, so that change in body condition could be assessed (Table 7). Where possible, treatment adults were also recaptured 7-16 days after logger deployment (70% recapture rate) and their body mass and size (wing length) taken, to estimate body condition (Table 7). Loggers were not removed to ensure that the maximum amount of data was collected per individual, except for 4 Pathtrack loggers that had stopped transmitting by the time the birds were recaptured. Body condition was calculated in the same way as for puffins.

Site	Pathtrack			UvA			Control		
	Nests	Chicks	Adults	Nests	Chicks	Adults	Nests	Chicks	Adults
Clett	3	5	6	2	3	3	14	26	0
Colms	3	5	11	4	6	3	24	42	0

Table 7: Number of kittiwake chicks and adults in each site and treatment category that were weighed and measured to assess body condition.

At-sea behaviour

For the analyses of at-sea behaviour we used the processed GPS tracking data (see section 2.1 above).

Breeding success

Breeding success (number of chicks fledged) of all study nests at each colony was recorded.

2.2.2 Data analysis

2.2.2.1 Puffin

We investigated effects of treatment, site and time since deployment on chick feeding rates using linear mixed models. The analysis of feeding rates was conducted at the individual and at the pair level. At the individual level, we tested whether the proportion of daily feeds delivered by birds equipped with a GPS logger (and colour rings) or with colour rings only, differed, and therefore to what extent partners would need to compensate for any reduction in feeding rates by the tagged birds. Direct testing for effects of treatment on individual feeding rates was not possible because we could not distinguish mates in unmarked control pairs. At the pair level, we tested whether the number of feeds delivered daily differed in 'colour-ring' and 'logger' pairs compared to controls, and therefore whether compensation by the partner of the tagged bird is effective. One 'logger' burrow at site 2 was removed from the dataset as the chick fledged during the feeding watches.

In the analysis at the individual level, we used generalised linear mixed models (GLMM) with binomial error distribution, where the response was the proportion of feeds per observation day delivered by the tagged bird. Treatment (logger vs colour ring) and site were fixed effects, day since deployment was a covariate, and nest identity and day were random effects in the models. Day since deployment was included in the analysis to test whether potential device effects persisted or whether the birds habituated to the loggers. Site was included to account for potential differences between the two study sites. We also included the treatment by site interaction to test for any differences in device effects between study plots. Day was included as a random effect in the models to account for random variation between days common to all individuals.

In the analysis at pair level, we used GLMMs with Poisson error distribution, where number of feeds per observation day was the response, treatment (logger vs colour ringed vs control) and site were fixed effects, day since deployment was a covariate, and nest identity and day were random effects. Again, we included the interaction between treatment and site in the models. Day since deployment (coded as 0, 1 or 2) was only relevant to 'treatment' nests so for all control nests it was set to 0. Coding it this way meant that it was not necessary to estimate a treatment by day since deployment interaction, as the main effect of "day since deployment" is effectively already an interaction (representing the impact of "day since deployment" in the treatment group only).

To investigate effects of logger deployment on chick fledging condition and survival we used generalised linear models (GLMs) with Gaussian and binomial error distributions, respectively. Treatment, site and the interaction treatment by site were main effects in the models.

2.2.2.2 Kittiwake

At-colony behaviour

As kittiwakes do not feed their chicks immediately after returning from a foraging trip, it is extremely challenging to record feeding rates directly in this species. We therefore used a proxy of foraging trip duration as a measure of provisioning effort and compared this between treatment and control nests. The proxy of trip duration was derived using the number of observed parent changeovers, the time the pair members spent together at the nest and the duration of the observation watch. Initial data exploration and comparison to trip durations obtained from the GPS tracking indicated that some changeovers may have been missed at Fowlsheugh and St Abb's Head, resulting in data from the three colonies not being directly comparable. We therefore, for this analysis only, modelled the colonies separately which allowed us to compare trip duration between the treatment groups within each colony. The number of changeovers was the response in the models and we included an offset for hours spent away from nest, meaning we modelled number of changeovers per hour away from the nest which is equivalent to the inverse of trip duration (but, since it is a count, is easier to model in a way that explicitly accounted for the effect of sample size). As in previous years (Bogdanova et al. 2021) we integrated the data for the three observation days for each nest and fitted GLMs with treatment (Pathtrack vs UvA vs control), site and treatment by site interaction as fixed effects. Two nests were excluded from this and all subsequent analyses, one UvA nest at Fowlsheugh which was predated during deployment and one Pathtrack nest at St Abb's Head where both adults were tagged by accident.

To ease comparison with the remaining analyses, for the analysis of chick attendance we included only sites where both GPS logger types were deployed (two on the Isle of May and two at Fowlsheugh; no data from the respective sites at St Abb's Head were available). We subsampled the data from the Isle of May to include every 4th observation per nest and subsampled the data from Fowlsheugh to cover a period of three weeks, to make the resolution and duration comparable in the two colonies. In addition to the two nests removed from all analyses (see above), two UvA and four control nests that failed soon after deployment or after attendance checks started were also removed from the dataset. In order to simultaneously model colony and site effects in a way that is readily interpretable, we created a new variable 'newsite' to distinguish sites within colonies: for each colony the site where attendance was lower than the colony-level average was assigned a value of -1 and the site where attendance was higher than the colony average was assigned of value of 1, so the effect of the "newsite" parameter represents the difference between the site-level and colony-level mean. As a final step, the raw data were aggregated to avoid issues due to potentially high levels of auto-correlation between days and data sparsity. We then fitted GLMs with binomial error distribution where the proportion of days nests were attended was the response, treatment, colony and brood size were fixed effects. We also included treatment by colony, colony by newsite and treatment by colony by newsite interactions in the analysis.

To test for device effects on adult condition we fitted GLMMs where condition was the response, treatment (Pathtrack vs UvA), measurement day, site, and the interactions treatment by measurement day and treatment by site were the explanatory variables and nest identity was a random effect. Of particular interest was the treatment by measurement day interaction as this was the main variable used to estimate device effects. Device effects on chick condition were investigated using GLMMs where treatment (Pathtrack vs UvA vs control), brood size, measurement day and site, as well as interactions involving treatment were the explanatory variables; nest and individual identity were random effects. Again the treatment by measurement day interaction was the main variable of interest. Models that did not include measurement day were not considered as they were not relevant to this analysis. The response variable (chick condition) was square-root transformed to achieve approximate normality. As data on condition were only collected on the Isle of May, colony was not included in these analyses.

At-sea behaviour

The analyses of at-sea behaviour constituted comparison of trip metrics and utilisation distributions of birds that were tracked with Pathtrack and with UvA loggers.

To avoid biasing the data due to colony-specific differences, the processed GPS tracking data and the extracted trip summaries that were derived at the data-processing stage (see section 1.1.2.1 above) were subsetted to only include data from sites where Pathtrack and UvA loggers were co-deployed. As deployment durations of UvA loggers were longer than those of Pathtrack loggers and we expected a change in trip characteristics over the breeding season, UvA tracking data were subset to only include data from the same time window that Pathtrack deployments covered. Similarly, trip characteristics are expected to change just before and after nest failure and therefore data from birds whose nests failed were excluded.

Utilisation distributions use a smoothing parameter which can be determined using different estimators and methodologies dependent on focus of the study including the *ad hoc* method that was used to calculate utilisation distributions for the full data set (see section 1.1.3 above). We followed the method outlined by Lascelles et al. *(*2016) to derive utilisation distributions, which uses the mode of optimal search radii of all trips calculated by the *varlogfpt* function in the adehabitatLT package as smoothing parameter (3.8km). Afterwards the *kernelUD* function in the adehabitatHR package was used to calculate 50 and 95% density contours for each of the logger type data sets across all colonies and for colony-specific data sets to a 1 km^2 resolution. Most commonly used contours are 50 and 95% which, for data from full trips, are commonly associated with highly used areas (e.g. foraging locations) and the overall distribution/home range, respectively. Therefore, the lower the percentage the higher the probability to encounter a bird in the area within the contour.

1) Trip metrics

To test whether foraging trip characteristics differed between birds tagged with different logger types, trip metrics (duration, distance and foraging range) were statistically compared using mixed effect models (function *lmer*, package lme4 in R*).* The full models comprised the trip metric as response and treatment, colony, treatment by colony interaction and julian day as explanatory variables (3 full models in total). Bird ID was included as a random effect to account for pseudoreplication. The response variables were log transformed to achieve approximate normality.

2) Utilisation distributions

To further examine differences in spatial distributions between the two logger types, the percentage overlap for the 50 and 95% contours was calculated across all colonies (data from the three colonies pooled) and at colony level (data from each colony analysed separately). Note that we would not expect a complete overlap of distributions due to individual variability and sampling error, and a reduction in the percentage of overlap from the 95% contour to the 50% contours is normal due to the reduction of area the contours cover.

Breeding success

As with the analysis of chick attendance, we included only sites where both GPS logger types were deployed and created a new variable 'newsite'. Device effects on breeding success were investigated using 1) GLMs with binomial error structure, where the response was the proportion of fledged chicks from those present at deployment and 2) GLMs with Poisson error structure, where the response was the number of fledged chicks per nest. Treatment, colony, and the interactions treatment by colony, colony by newsite and treatment by colony by newsite were the main effects in the models.

2.2.2.3 Model selection

For each analysis, our candidate set included a 'null' model containing only the random effect(s) and no fixed effects, models testing for each of the main effects separately and in combination, and a 'full' model containing all main effects and relevant interactions between them. Support for different candidate models was assessed using Akaike's information criterion adjusted for small sample size (AICc). The model with the lowest AICc value was considered best supported. Models were deemed strongly supported if they differed from the best model by less than 2 AICc units (Burnham & Anderson 2002), unless they were otherwise identical to the best model but contained one more parameter, in which case this rule is not appropriate (Burnham & Anderson 2002) and the more complex models were disregarded on the grounds of parsimony. Marginal coefficient of determination (R^2m, representing the variance explained by the fixed effects) and conditional coefficient of determination (R^2c, representing the variance explained by both fixed and random effects; Nakagawa & Schielzeth 2013) were calculated for the best model in each candidate set. Analyses were performed in R, using packages lme4 (Bates et al. 2019) and MuMIn (Bartoń 2019).

3 Results

3.1 Species utilisation distribution in 2021

3.1.1 Isle of May

For three of the species (guillemot, razorbill and kittiwake) where we were able to separate foraging and resting at sea behaviours, we present maps of all locations at sea, foraging locations only and resting at sea locations only, as well as utilisation distributions based on those (guillemot: Fig. 12a-f, razorbill: Fig. 13a-f, kittiwake: Fig. 15a-f). For puffin, where we were unable to reliably distinguish foraging from resting behaviour, we present maps of all locations at sea and of non-flight locations only, as well as utilisation distributions based on those (Fig. 14a-d). In 2021, clear differences in distributions were apparent among the four species. Guillemots were concentrated inshore within the Firth of Forth and to the north of the Isle of May reaching St Andrews Bay (Fig. 12a,b). The distribution of razorbills included both inshore and offshore areas, with core areas located around the Isle of May, within the Firth of Forth and to the east of the colony (Fig. 13a,b). Puffin distribution was concentrated around the Isle of May and in an offshore area spanning north-east to south-east of the colony, with core areas overlapping with the Neart na Gaoithe footprint (Fig. 14a,b). Kittiwakes mainly used areas located offshore, to the north, north-east and east of the colony, although they also visited inshore waters within the Firth of Forth and St Andrews Bay (Fig. 15a,b). Several individuals also travelled to more distant areas offshore, located >100km of the colony (Fig. 15a). The mean maximum range (± SE) from the Isle of May was 26.5 ± 2.2 km in guillemots, 36.0 ± 2.3 km in razorbills, 40.5 ± 1.8 km in puffins and 71.3 ± 3.6 km in kittiwakes.

Within species there were some differences in distributions associated with different behaviours. In guillemots, core foraging and resting areas largely overlapped (Fig. 12d,f), in contrast to findings in 2020 (Bogdanova et al. 2021). In razorbills, there was substantial overlap however core resting areas extended slightly further offshore than foraging ones (Fig. 13d,f). In kittiwakes, core foraging areas included an area north of the colony near the coast whereas core resting areas were located mostly offshore (Fig. 15d,f). We are currently unable to establish whether puffins use specific areas at sea for foraging or resting but are in the process of developing more sophisticated behavioural classification methods for this species with the aim to distinguish between these two behaviours in GPS tracking data.

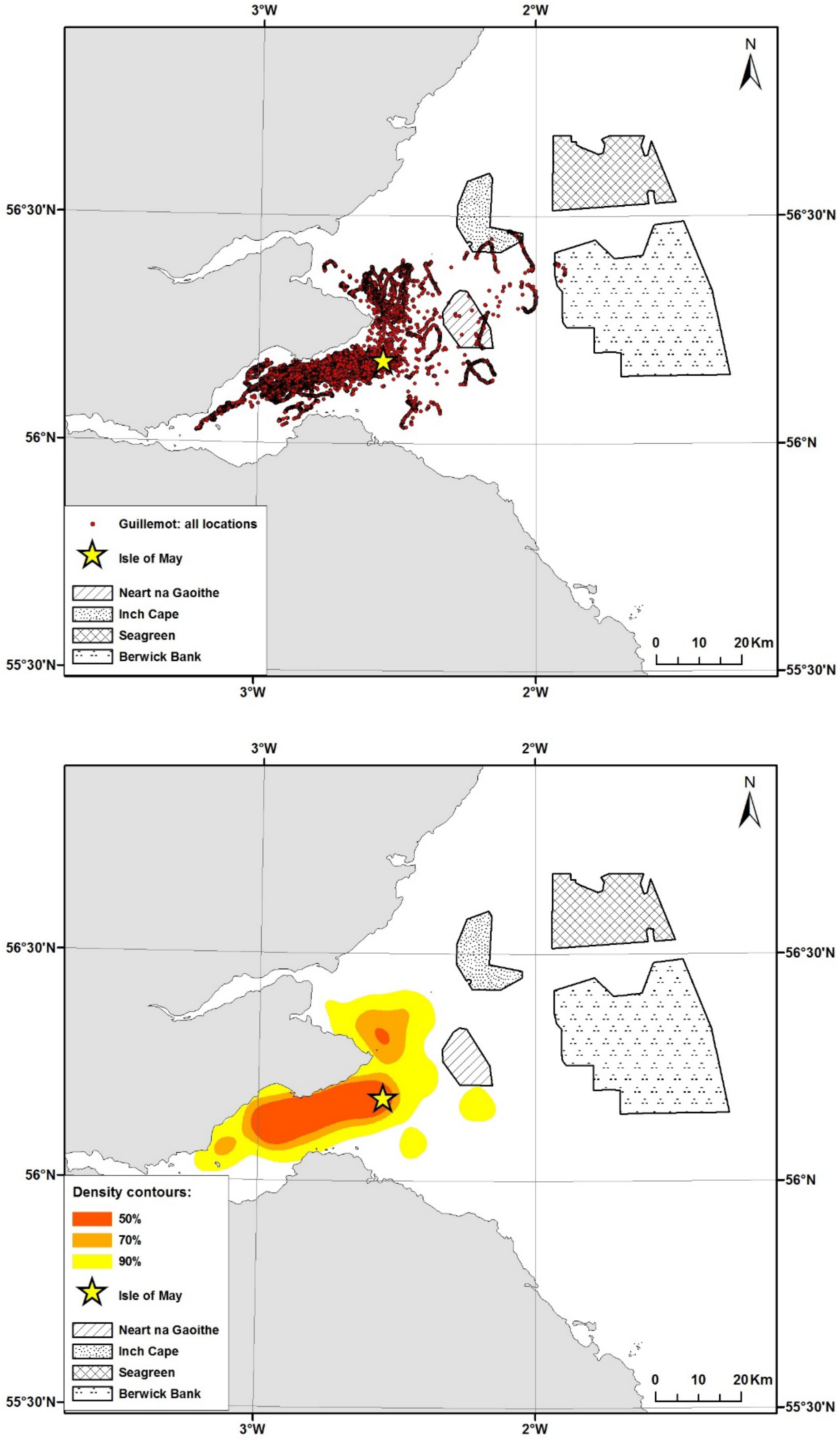


Figure 12: a) Individual GPS fixes and b) utilisation distributions (50%, 70%, 90% contours) for guillemot in 2021 for all behaviours combined.

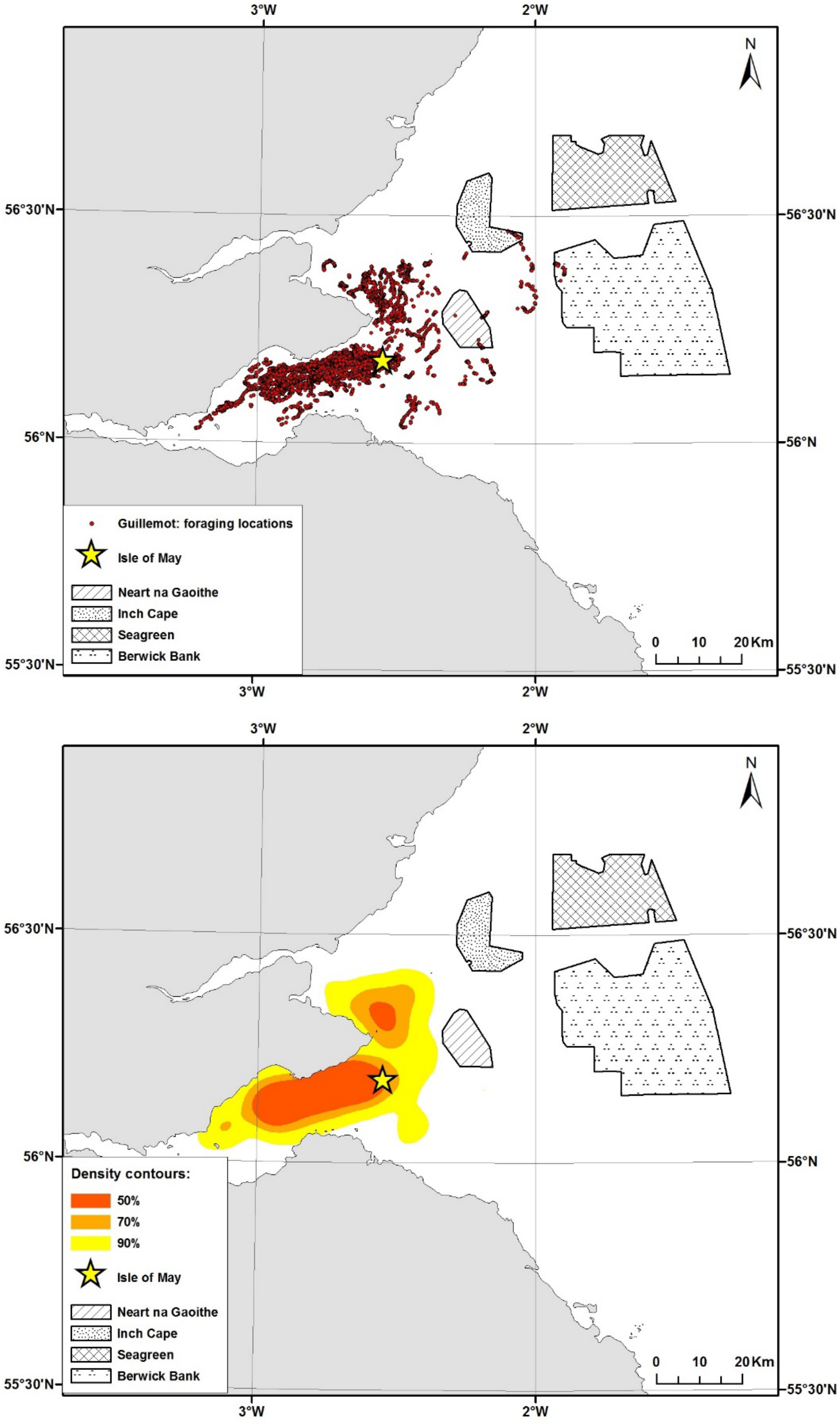


Figure 12 (cont.): c) Individual GPS fixes and d) utilisation distributions (50%, 70%, 90% contours) for guillemot in 2021 for foraging behaviour only.

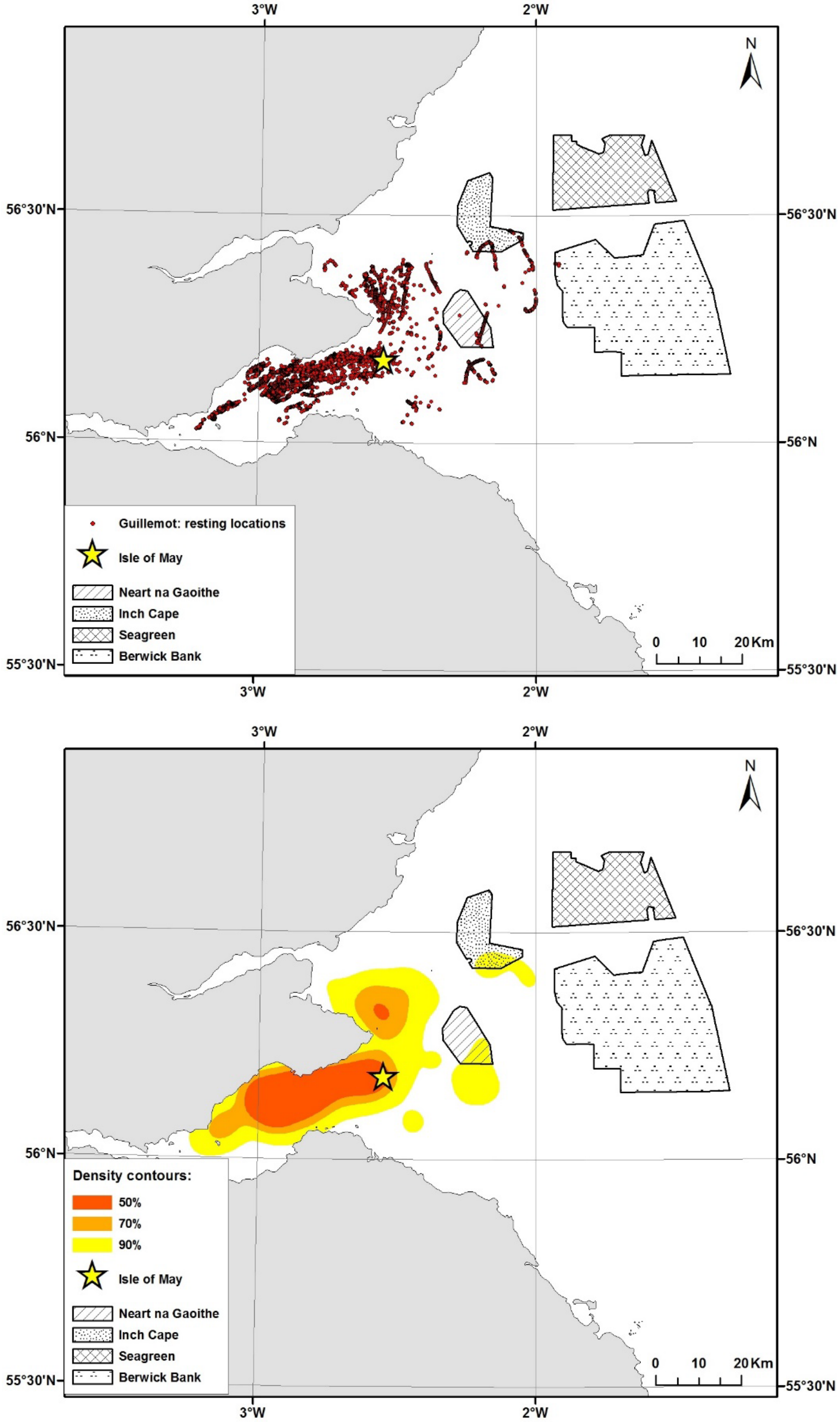


Figure 12 (cont.): e) Individual GPS fixes and f) utilisation distributions (50%, 70%, 90% contours) for guillemot in 2021 for resting behaviour only.

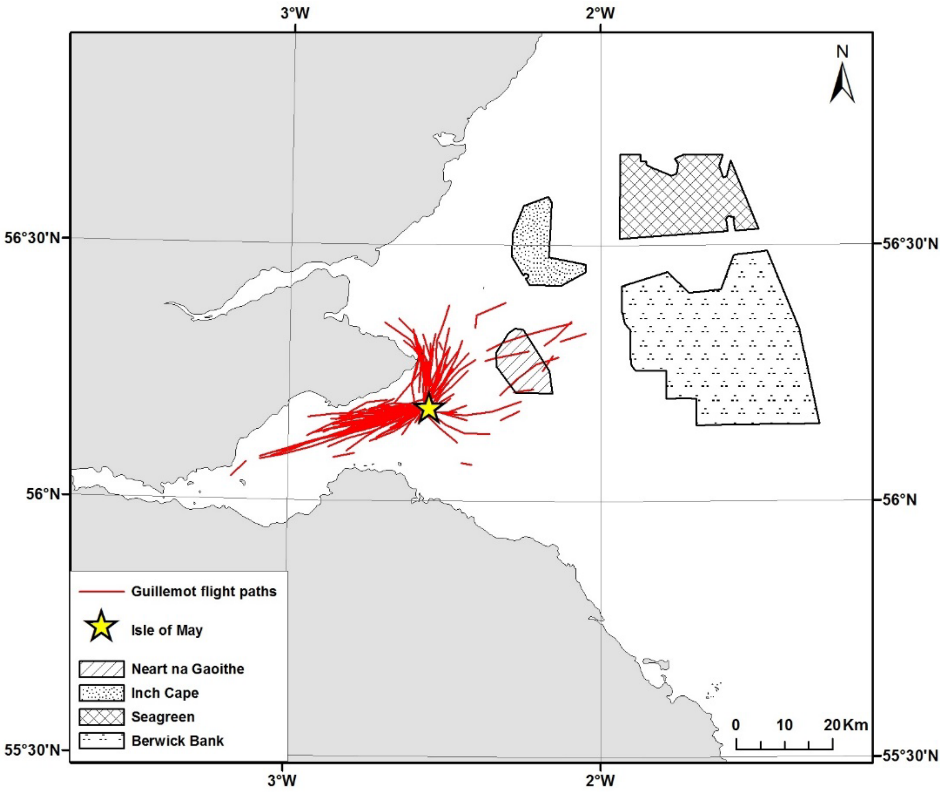


Figure 12 (cont.): g) Horizontal flight lines for guillemot in 2021.

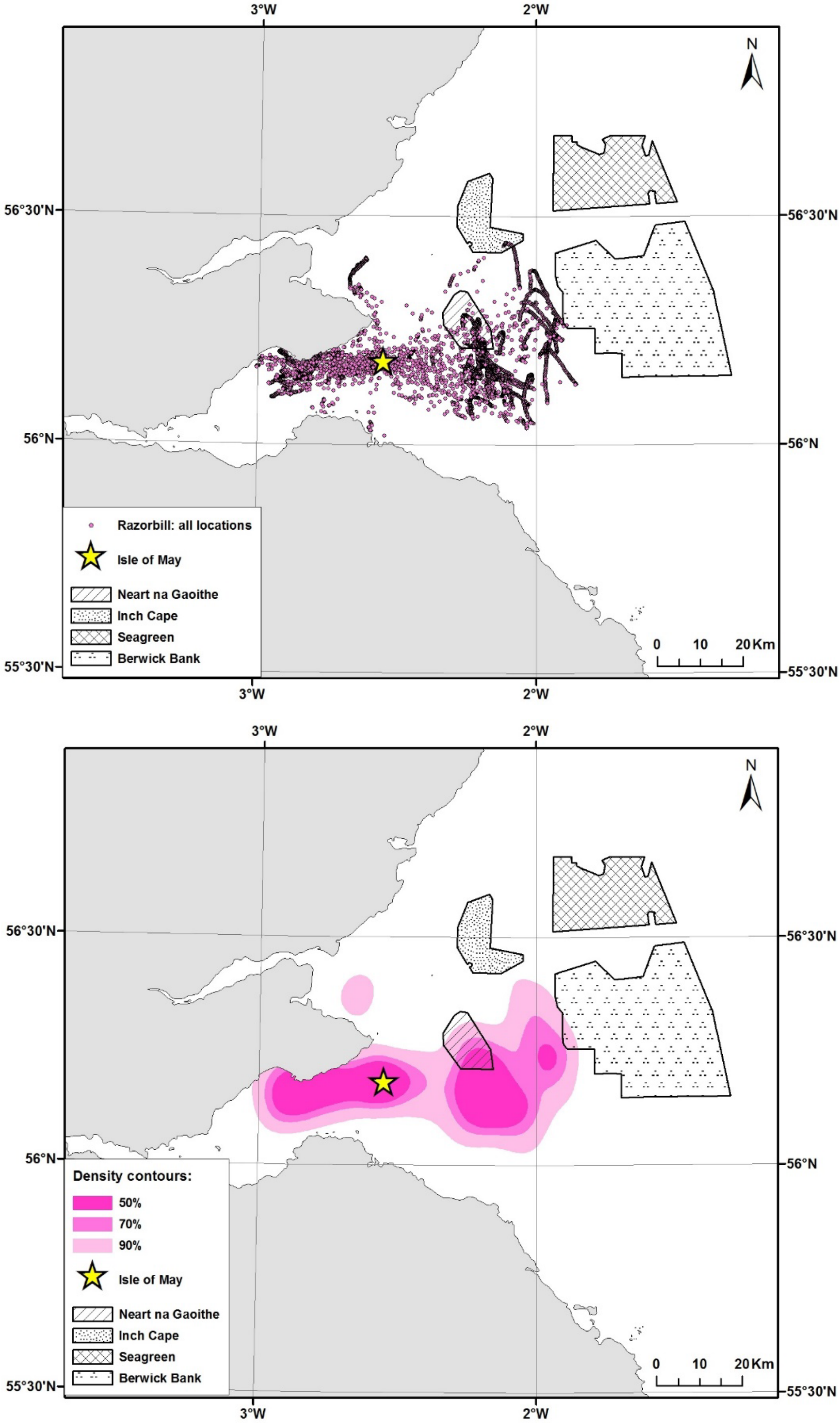


Figure 13: a) Individual GPS fixes and b) utilisation distributions (50%, 70%, 90% contours) for razorbill in 2021 for all behaviours combined.

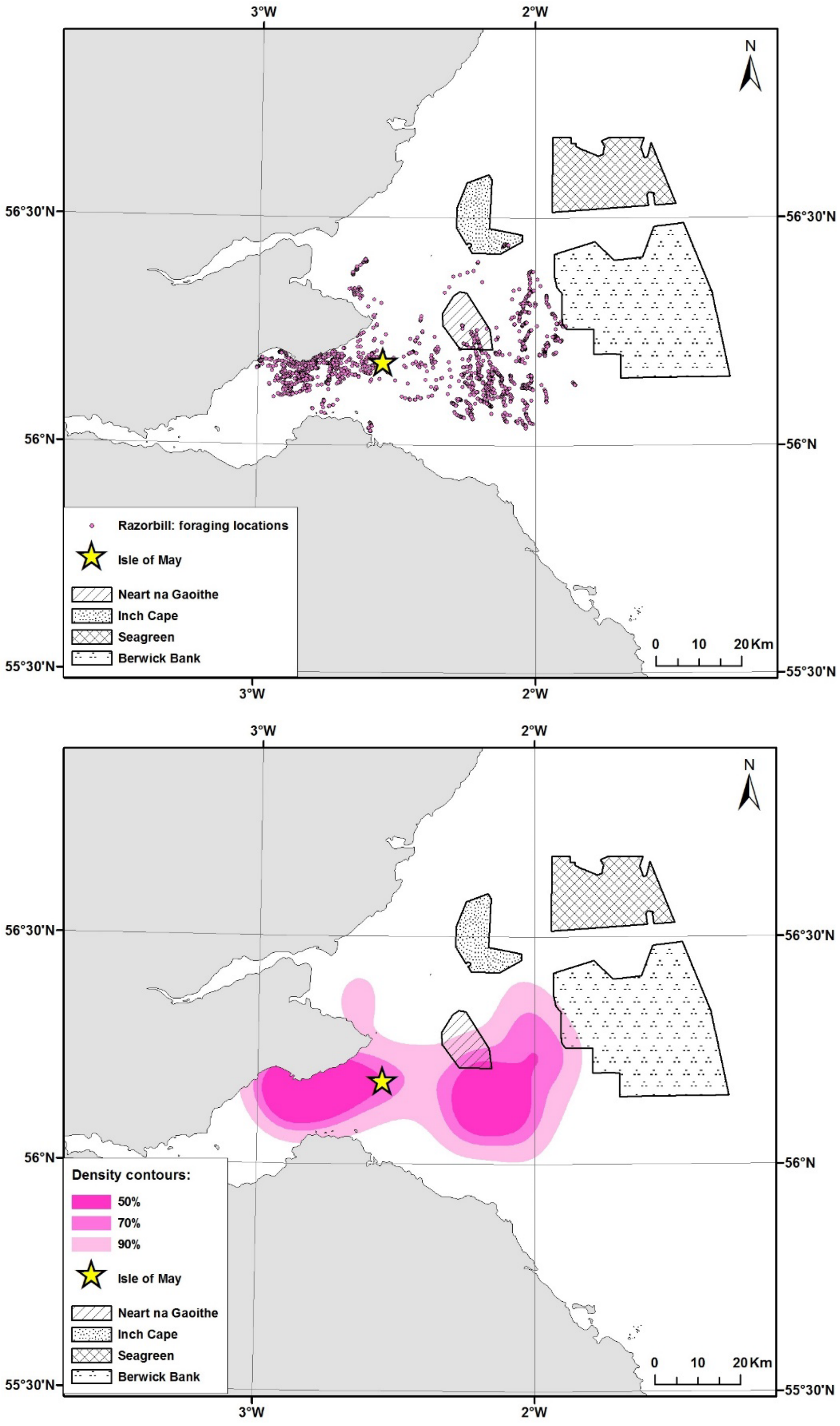


Figure 13 (cont.): c) Individual GPS fixes and d) utilisation distributions (50%, 70%, 90% contours) for razorbill in 2021 for foraging behaviour only.

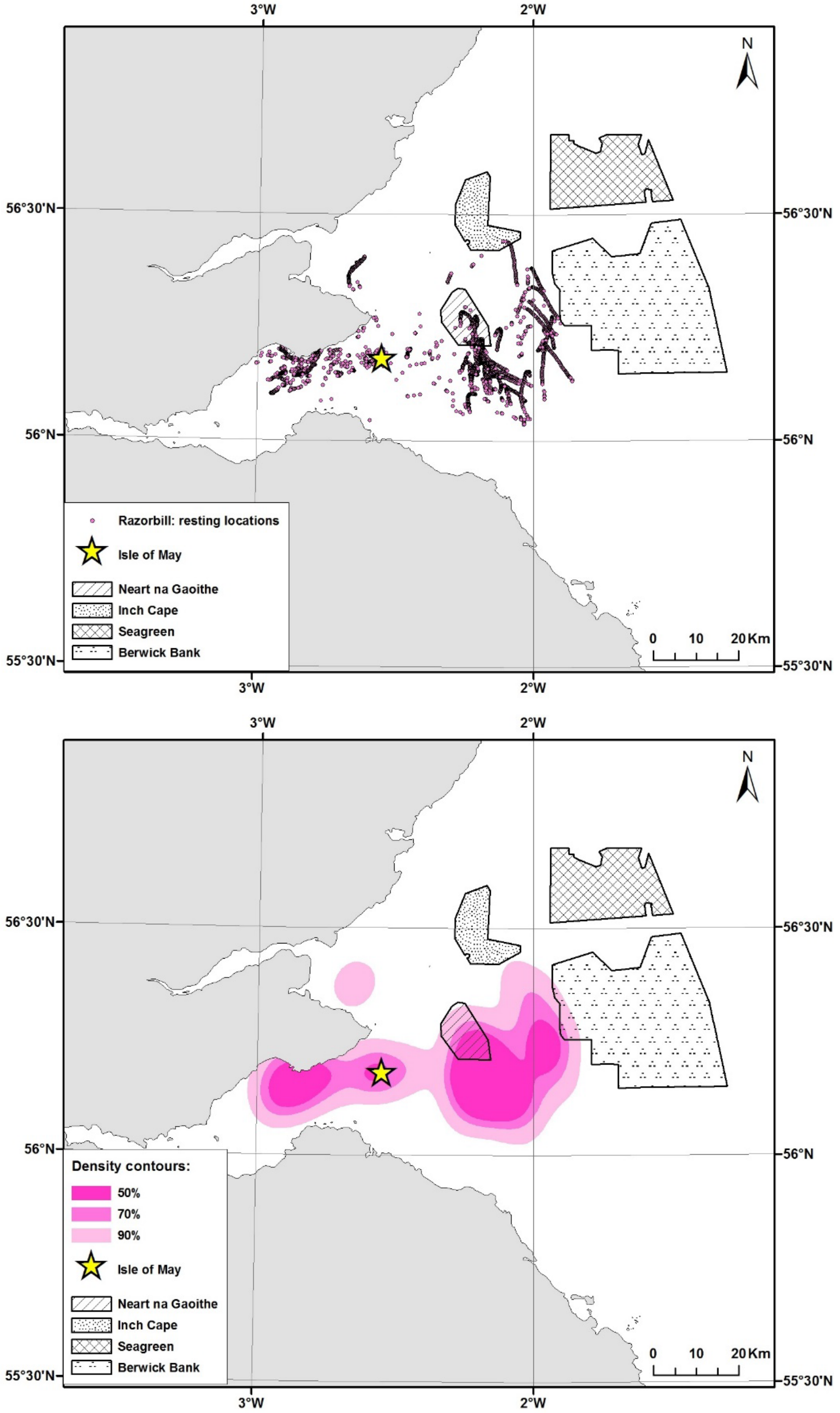


Figure 13 (cont.): e) Individual GPS fixes and f) utilisation distributions (50%, 70%, 90% contours) for razorbill in 2021 for resting behaviour only.

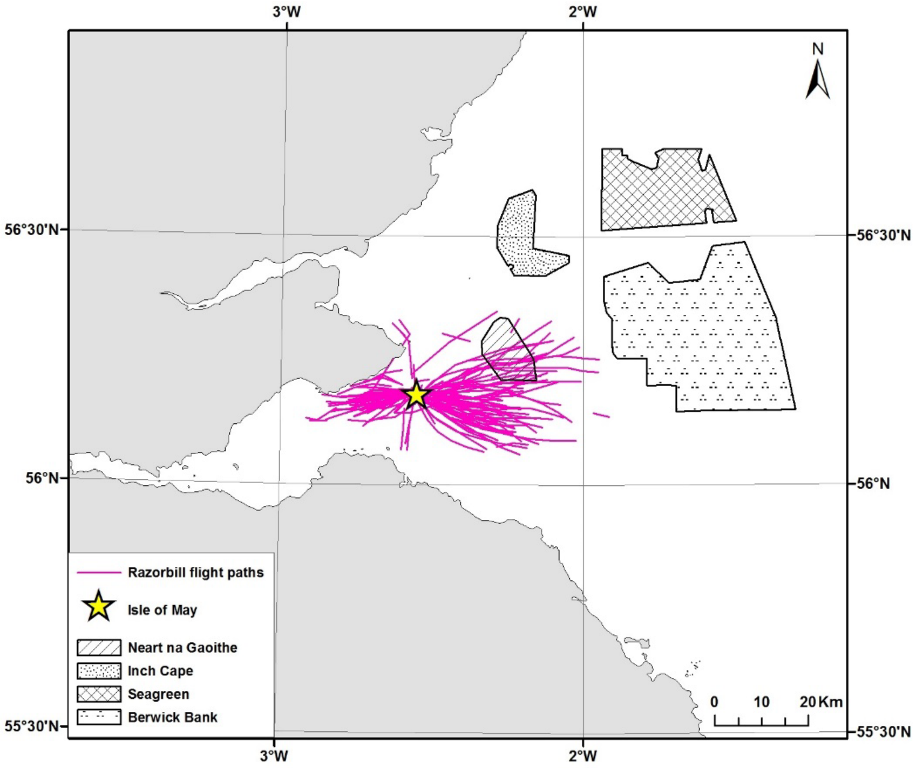


Figure 13 (cont.): g) Horizontal flight lines for razorbill in 2021.

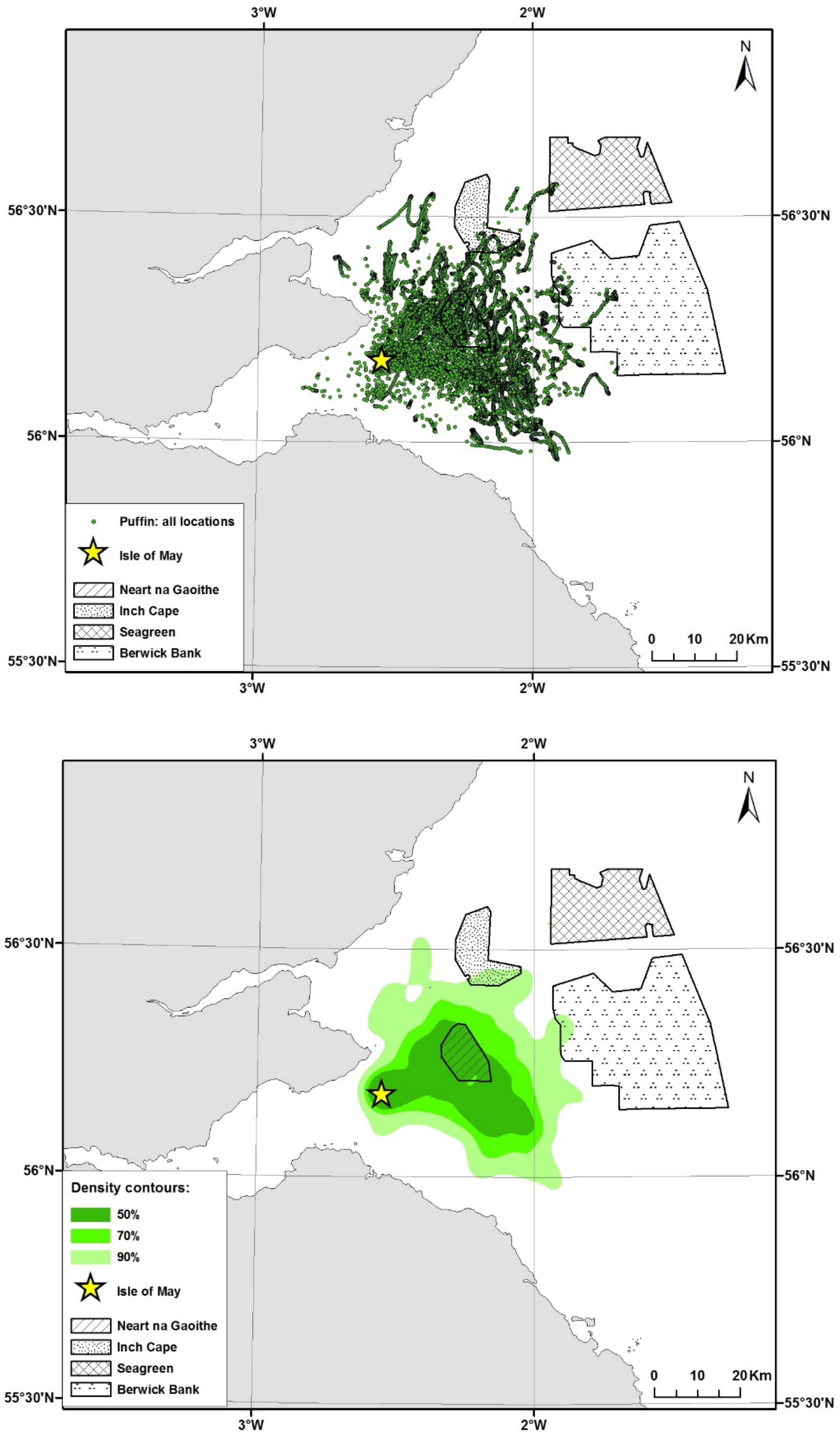


Figure 14: a) Individual GPS fixes and b) utilisation distributions (50%, 70%, 90% contours) for puffin in 2021 for all behaviours combined.

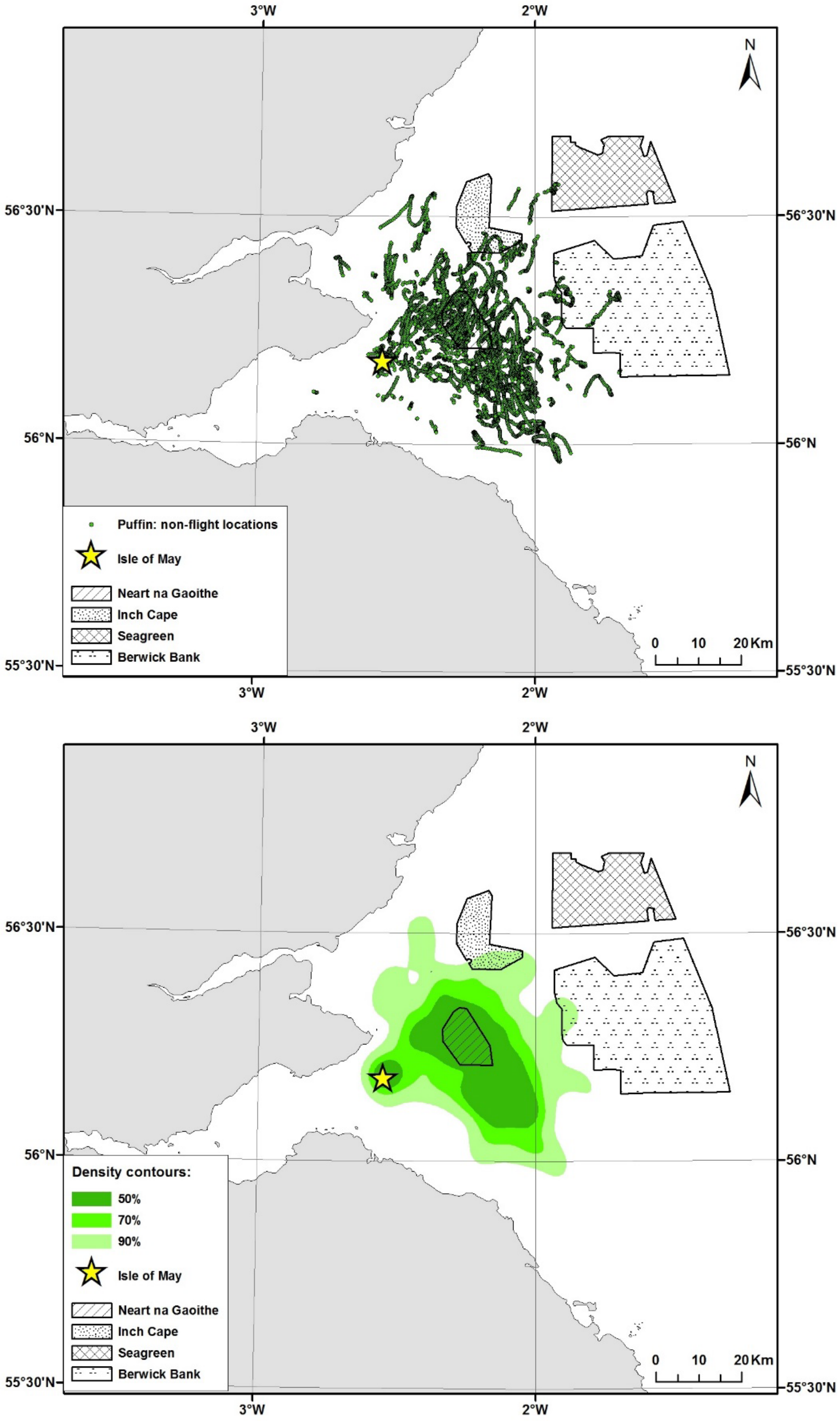


Figure 14 (cont.): c) Individual GPS fixes and d) utilisation distributions (50%, 70%, 90% contours) for puffin in 2021 for non-flight behaviours only.

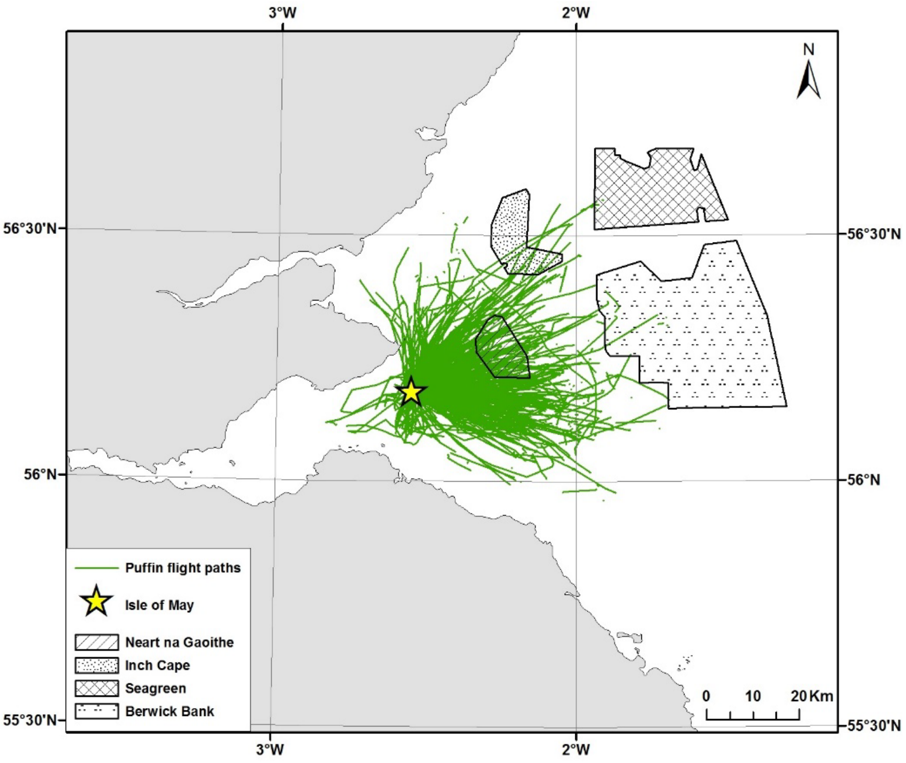


Figure 14 (cont.): e) Horizontal flight lines for puffin in 2021.

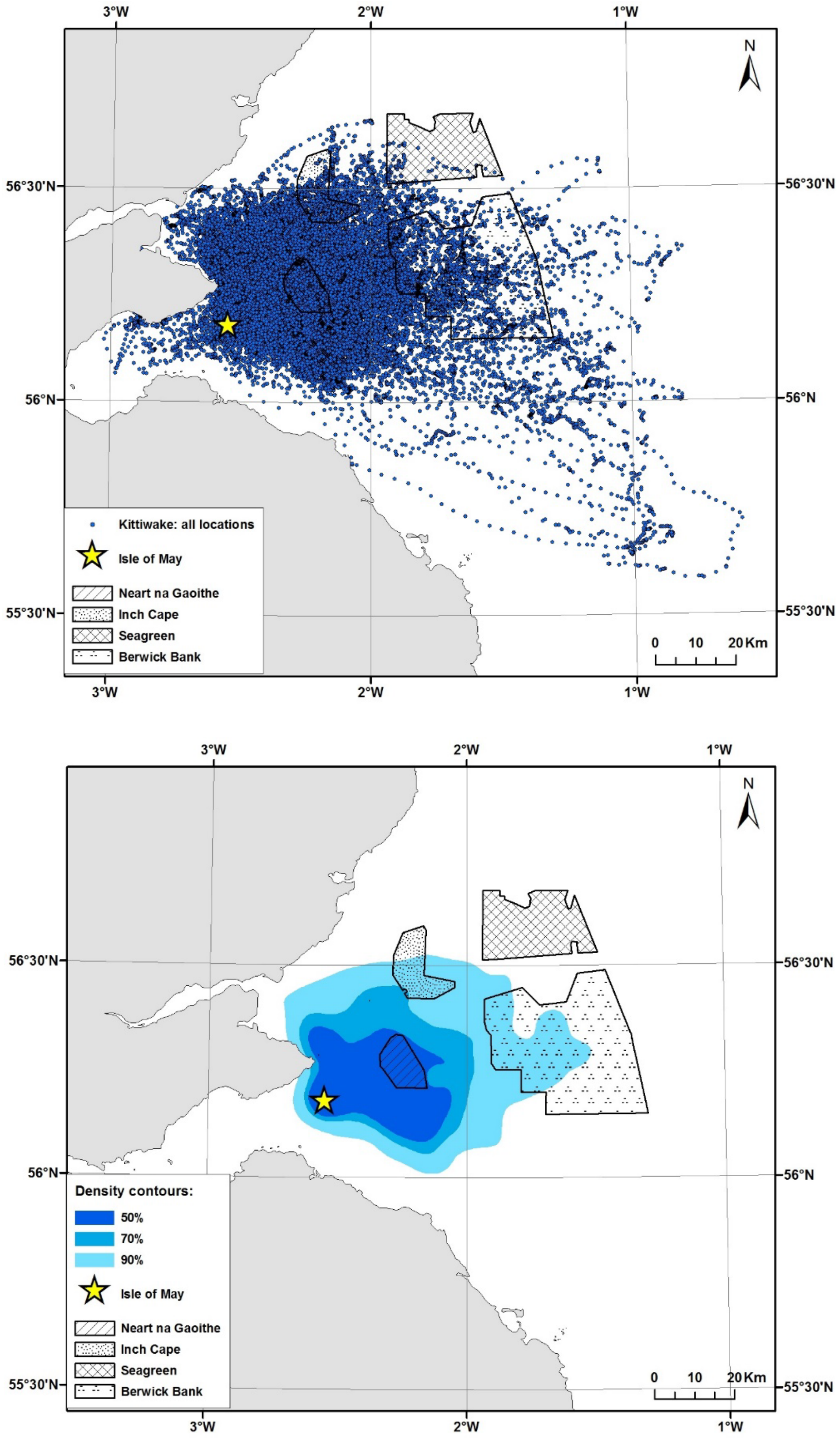
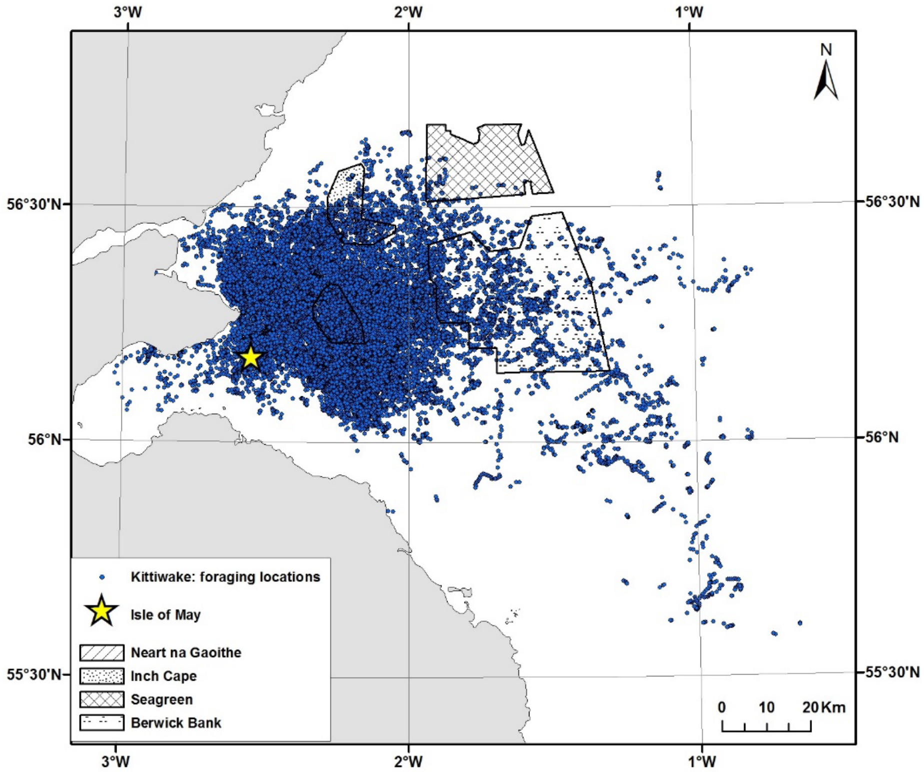


Figure 15: a) Individual GPS fixes and b) utilisation distributions (50%, 70%, 90% contours) for kittiwakes from the Isle of May in 2021, all behaviours combined.



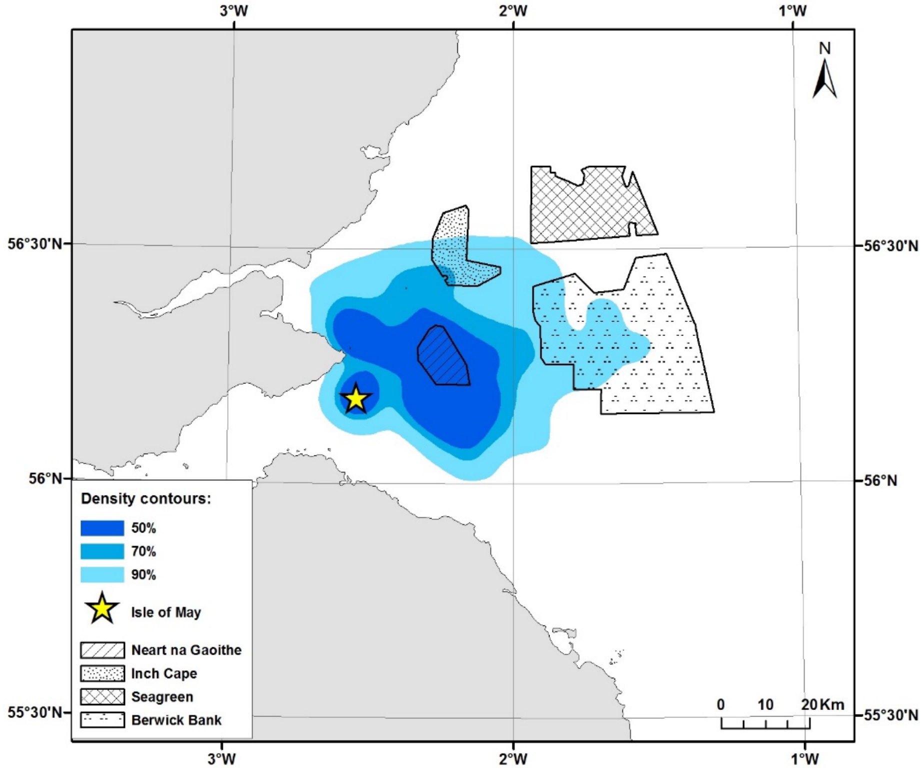


Figure 15 (cont.): c) Individual GPS fixes and d) utilisation distributions (50%, 70%, 90% contours) for kittiwakes from the Isle of May in 2021, foraging behaviour only.

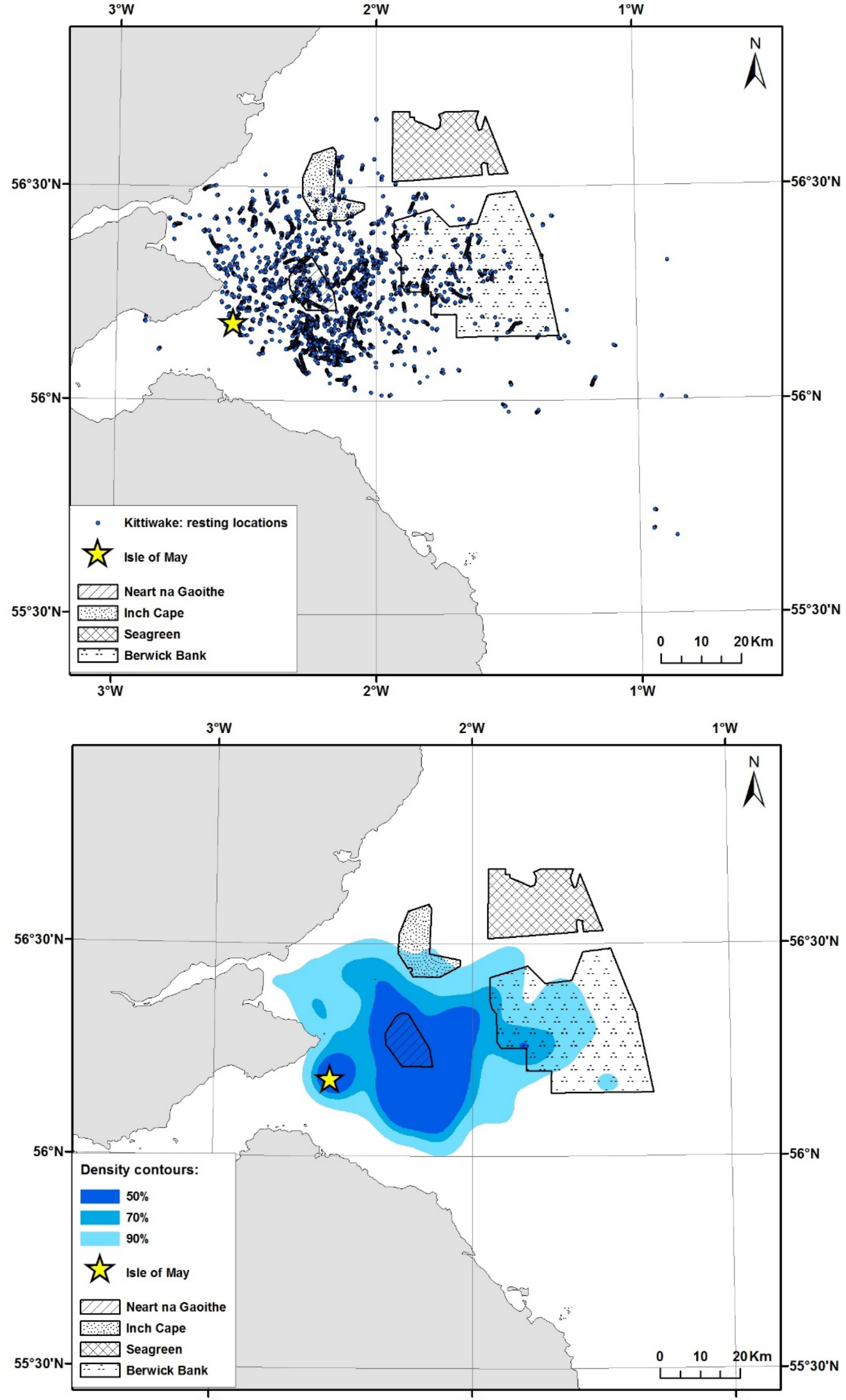


Figure 15 (cont.): e) Individual GPS fixes and f) utilisation distributions (50%, 70%, 90% contours) for kittiwakes from the Isle of May in 2021, resting behaviour only.

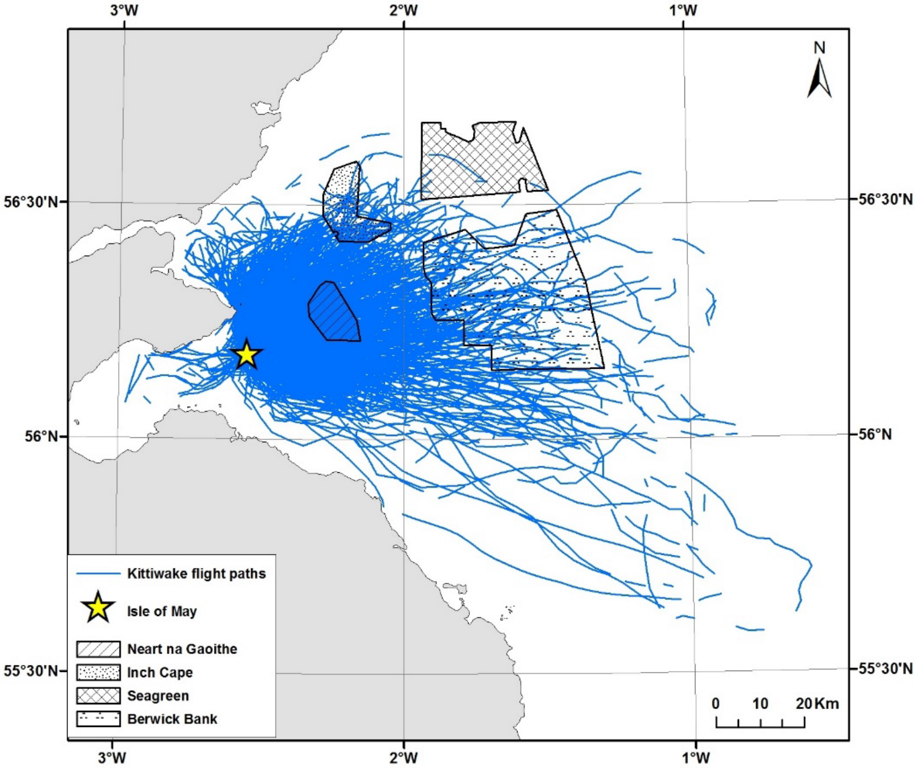


Figure 15 (cont.): g) Horizontal flight lines for kittiwakes from the Isle of May in 2021.

3.1.2 Fowlsheugh

As for the Isle of May, we present maps of all kittiwake locations at sea, foraging locations only and resting at sea locations only, as well as utilisation distributions based on those (Fig. 16 a-f). Birds from Fowlsheugh used areas ranging from north-east to south-east of the colony, with core areas located mostly to the east of it. The mean maximum range for this colony was 81.9 ± 3.8 km, with some individuals undertaking trips to more distant areas (including a bird that travelled to St Abb's Head). The areas kittiwakes used for foraging and resting were largely overlapping (Fig. 16 d,e).

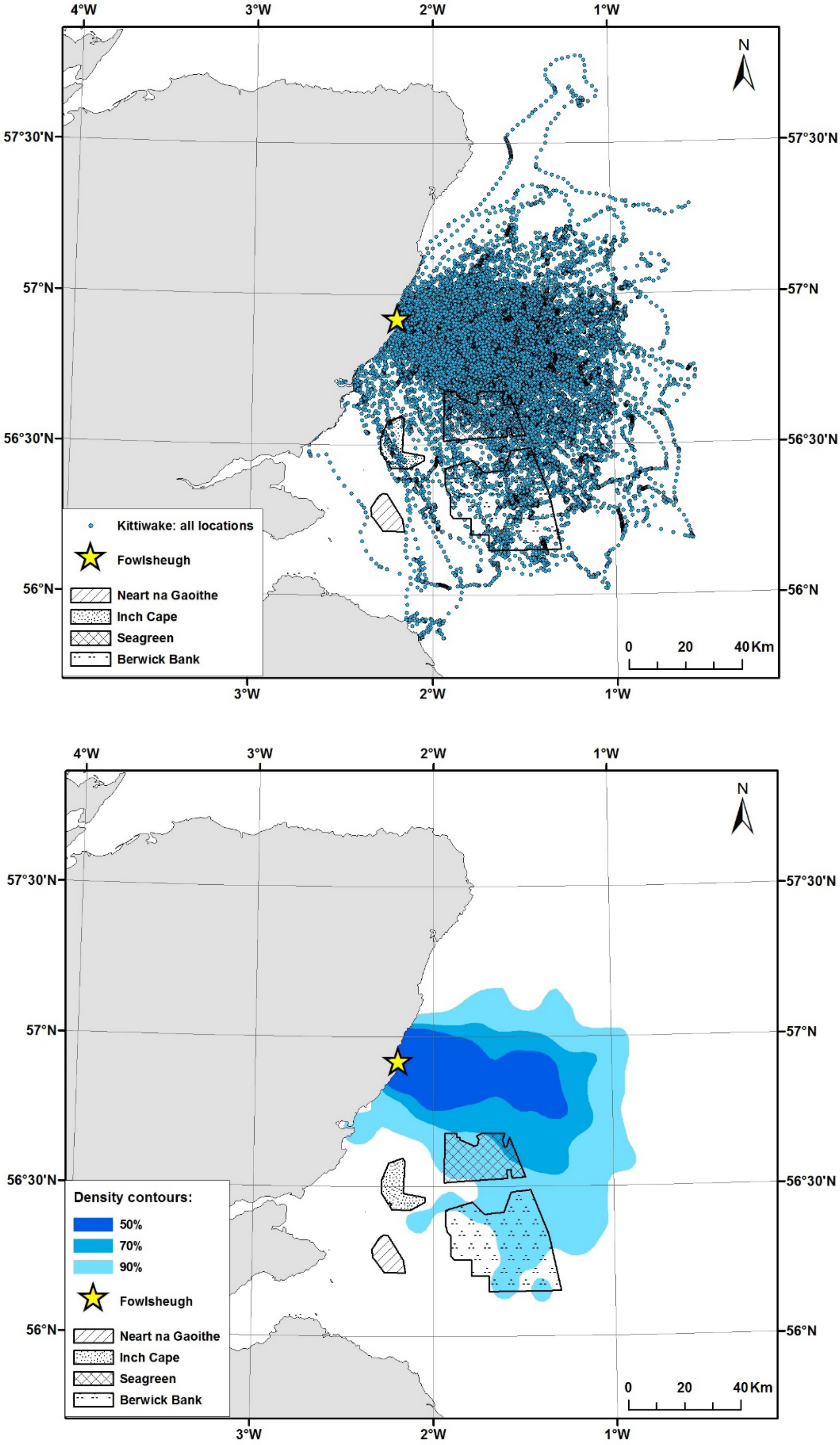


Figure 16: a) Individual GPS fixes and b) utilisation distributions (50%, 70%, 90% contours) for kittiwakes from Fowlsheugh in 2021, all behaviours combined.

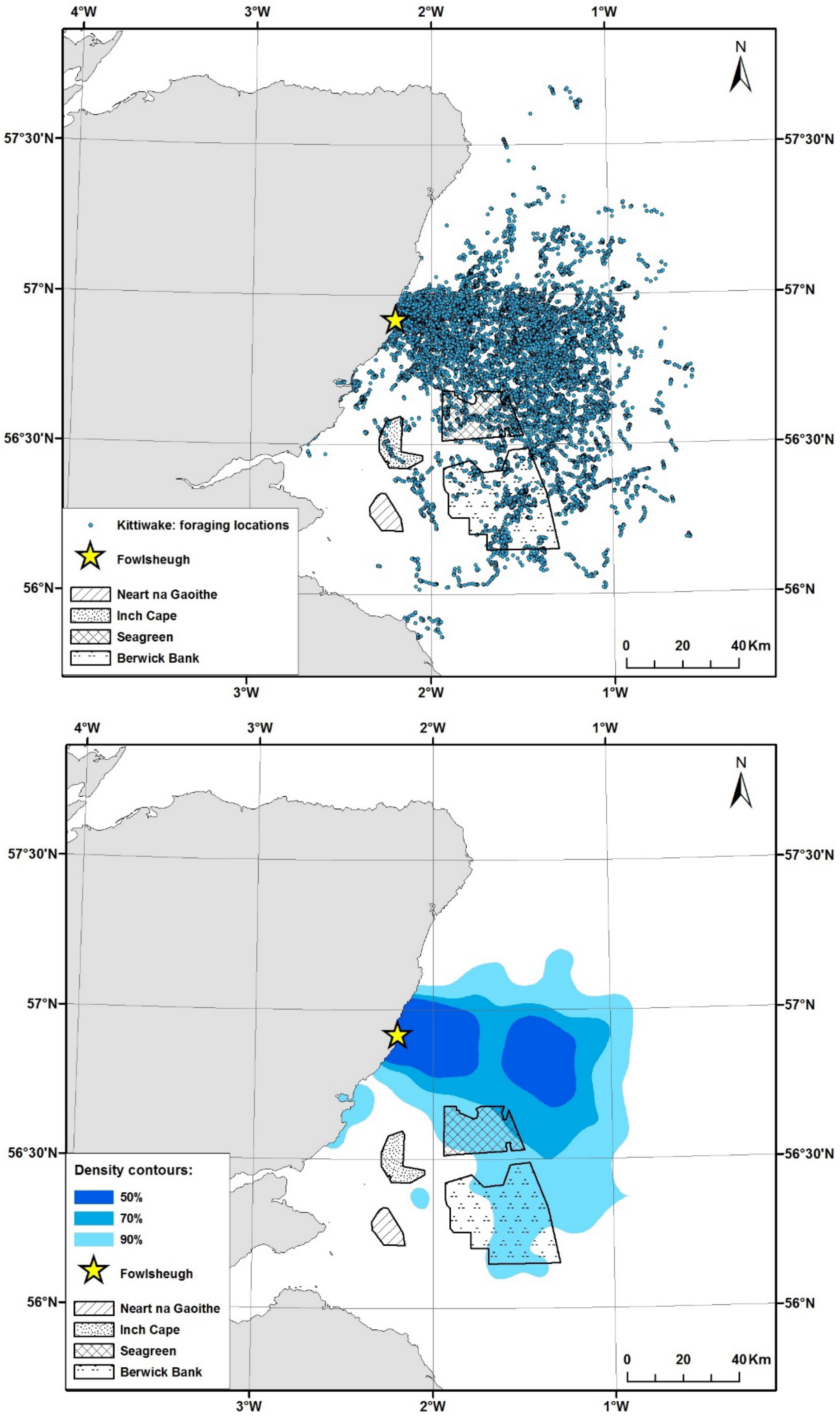


Figure 16 (cont.): c) Individual GPS fixes and d) utilisation distributions (50%, 70%, 90% contours) for kittiwakes from Fowlsheugh in 2021, foraging behaviour only.

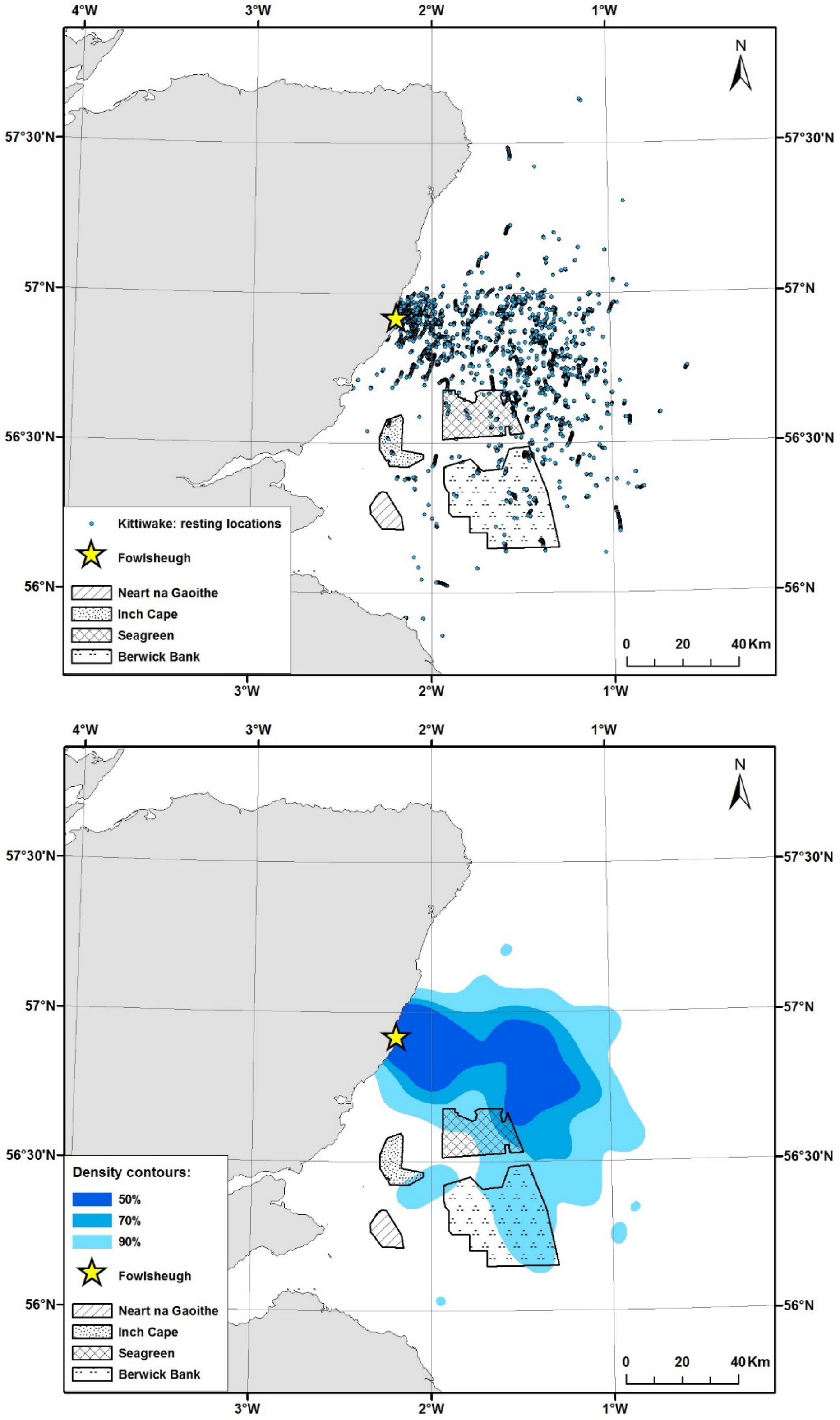


Figure 16 (cont.): e) Individual GPS fixes and f) utilisation distributions (50%, 70%, 90% contours) for kittiwakes from Fowlsheugh in 2021, resting behaviour only.

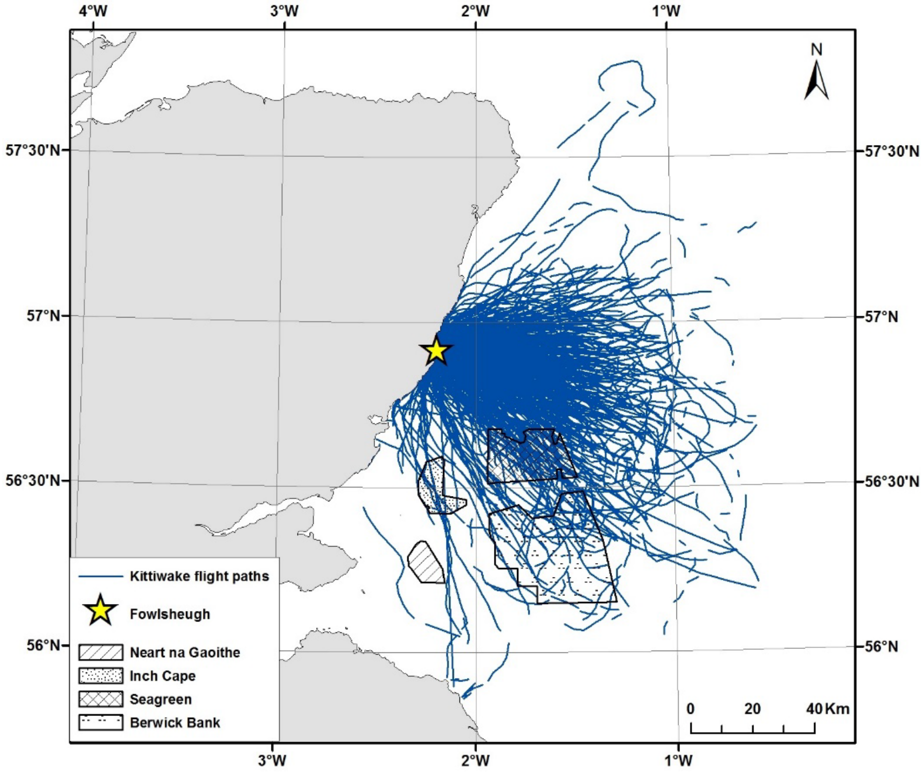


Figure 16 (cont.): g) Horizontal flight lines for kittiwakes from Fowlsheugh in 2021.

3.1.3 St Abb's Head

Maps of all locations at sea, foraging locations only and resting at sea locations only, as well as utilisation distributions based on those for kittiwakes from St Abb's Head are presented in Fig. 17 a-f. The tagged birds used offshore areas spanning north to east of the colony, as well as coastal areas to the north-west and south-east of the colony. Core areas were located mostly to the north-east of it. The mean maximum range of kittiwakes at this colony was more limited compared to the Isle of May and Fowlsheugh (63.7 ± 4.8 km). However, a few individuals undertook more distant trips, including a bird that travelled to the coast between Aberdeen and Peterhead. Core foraging and resting areas largely overlapped, although the overall area used for foraging extended further to the north-east of the colony compared to the overall area used for resting (Fig. 17 d,e).

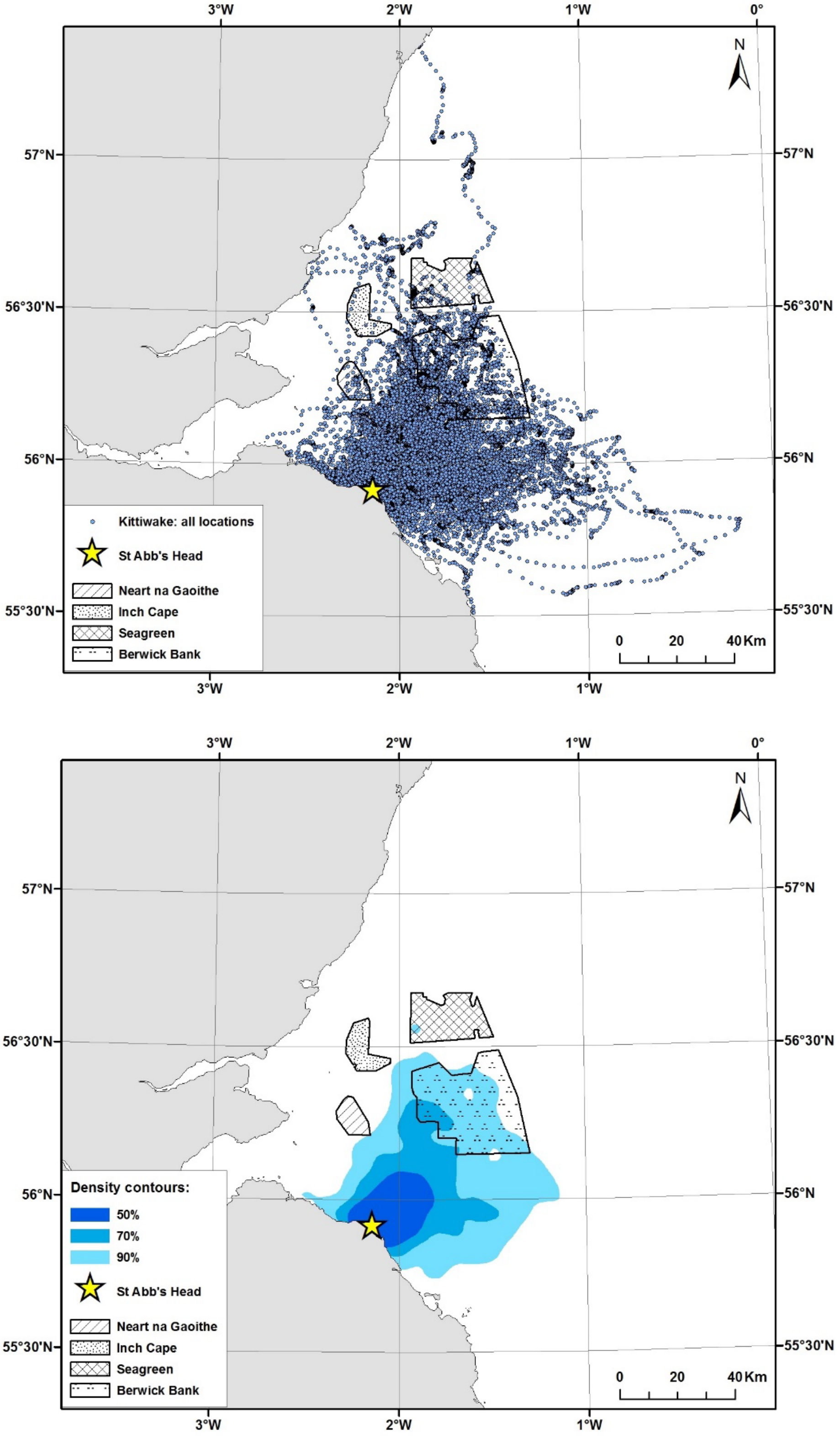


Figure 17: a) Individual GPS fixes and b) utilisation distributions (50%, 70%, 90% contours) for kittiwakes from St Abb's Head in 2021, all behaviours combined.

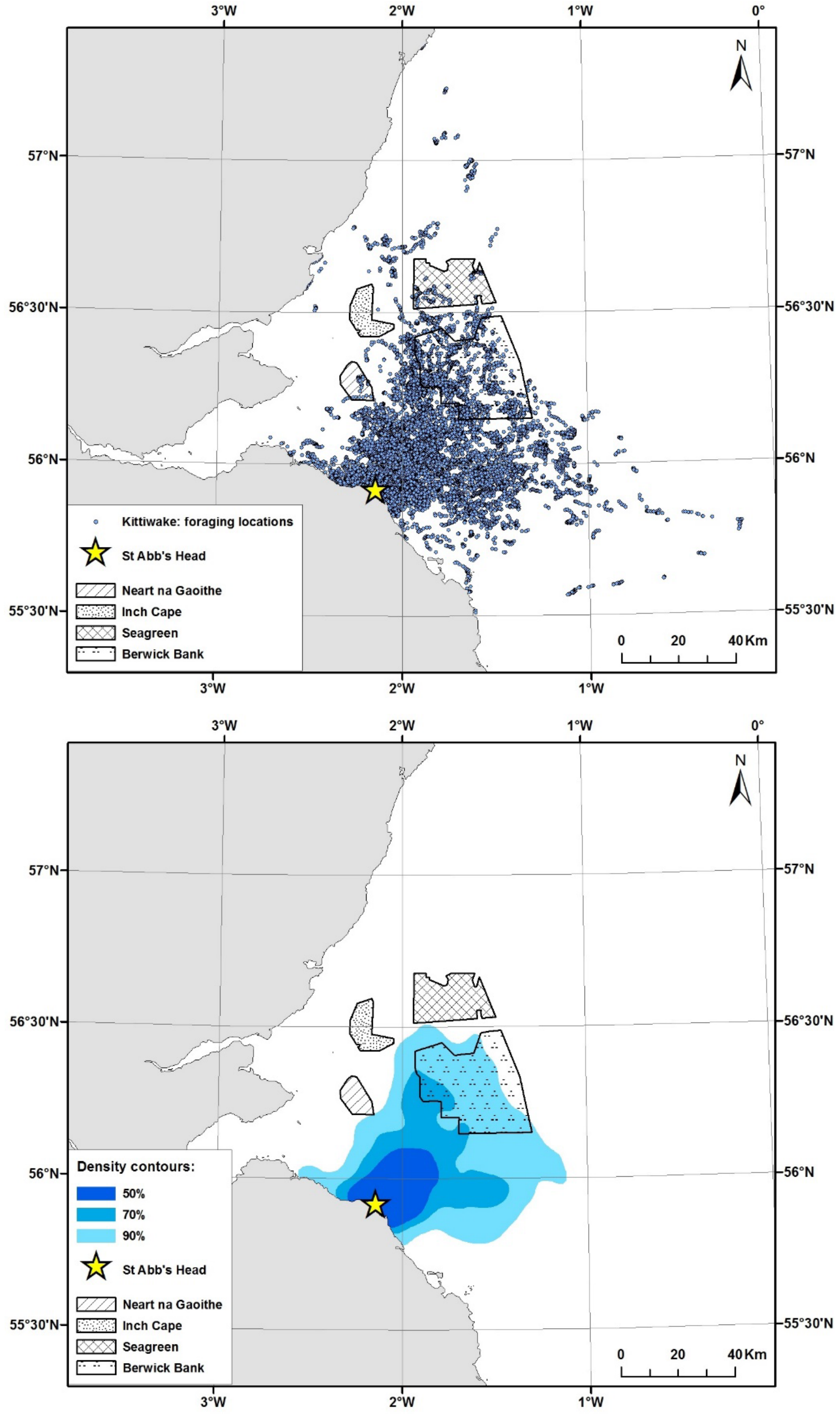


Figure 17 (cont.): c) Individual GPS fixes and d) utilisation distributions (50%, 70%, 90% contours) for kittiwakes from St Abb's Head in 2021, foraging behaviour only.

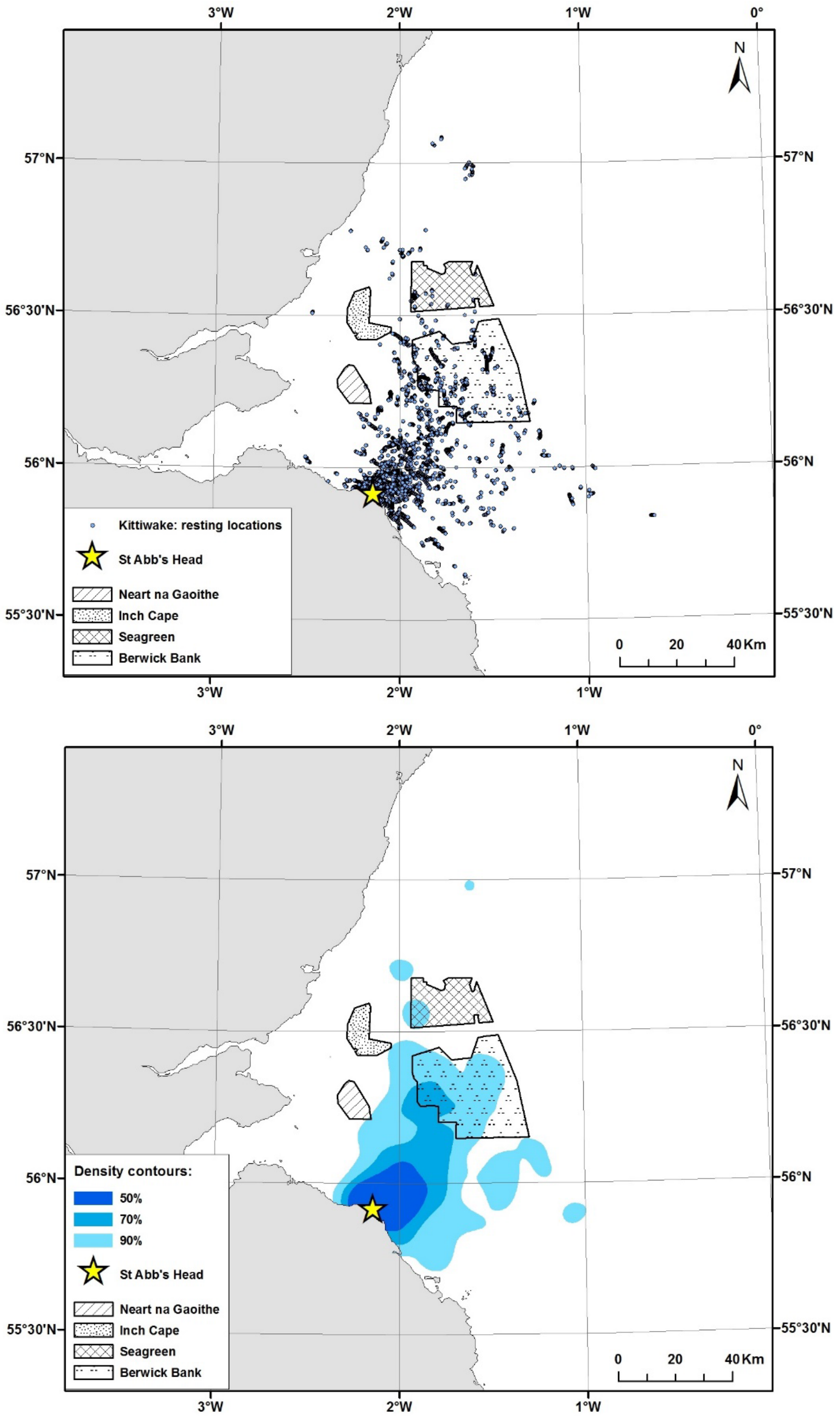


Figure 17 (cont.): e) Individual GPS fixes and f) utilisation distributions (50%, 70%, 90% contours) for kittiwakes from St Abb's Head in 2021, resting behaviour only.

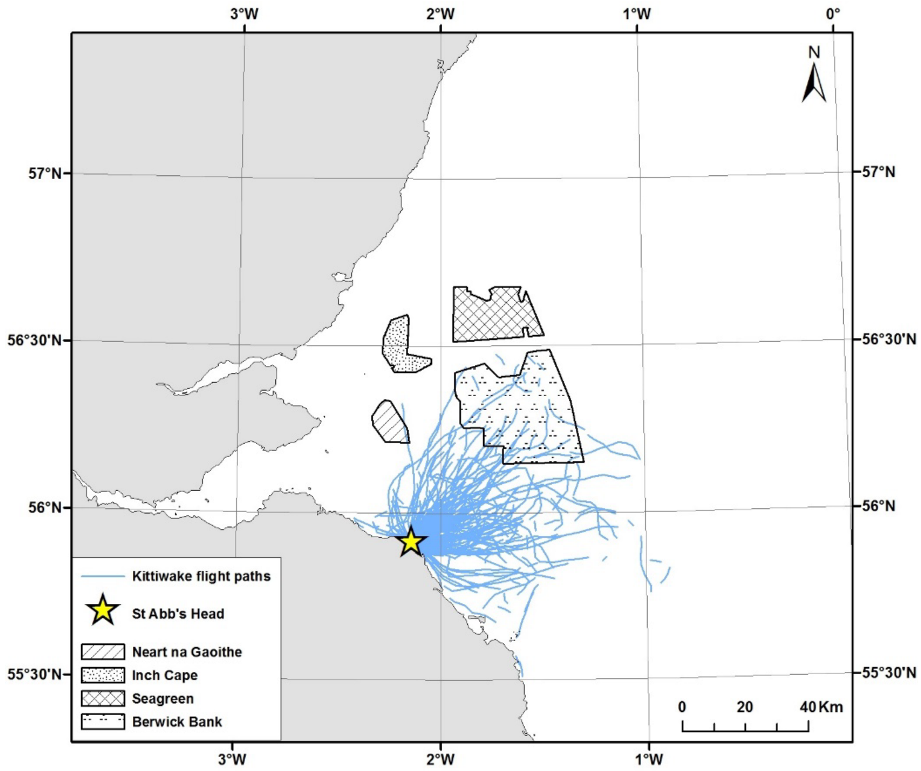


Figure 17 (cont.): g) Horizontal flight lines for kittiwakes from St Abb's Head in 2021.

3.2 Horizontal flight trajectories in 2021

Maps of horizontal flight trajectories are shown in Fig. 12g (for guillemot), Fig. 13g (for razorbill), Fig. 14e (for puffin) and Figs. 15-17g (for kittiwake). As expected, the distribution of flight trajectories generally matched closely with the UD distributions. Guillemots departed from and returned to the colony on bearings ranging from south-west and north/north-west (for inshore foraging trips) to north-east and east (for offshore trips), with inshore trips being more common. In razorbills, most bearings were to the west or to north-east to south-east from the Isle of May, and the proportion of commuting flights to offshore areas was higher in this species. Flight bearings of puffins spanned from northerly to south-easterly directions from the colony. Bearings of kittiwake flights from the Isle of May were predominantly from north-west to south-east of the colony, with few flights recorded inshore within the Firth of Forth (Fig. 15g).

Flights of kittiwakes from Fowlsheugh were mostly offshore, with bearings ranging from northeasterly to southerly direction from the colony, although some coastal trips were also observed (Fig. 16g). Kittiwakes from St Abb's Head ranged less widely compared to birds from the other two colonies, with flight bearings mainly to the north/north-east and east of the colony (Fig. 17g).

To ease comparison, combined maps of utilisation distributions and of flight trajectories of kittiwakes from the three study colonies are presented in Figure 18.

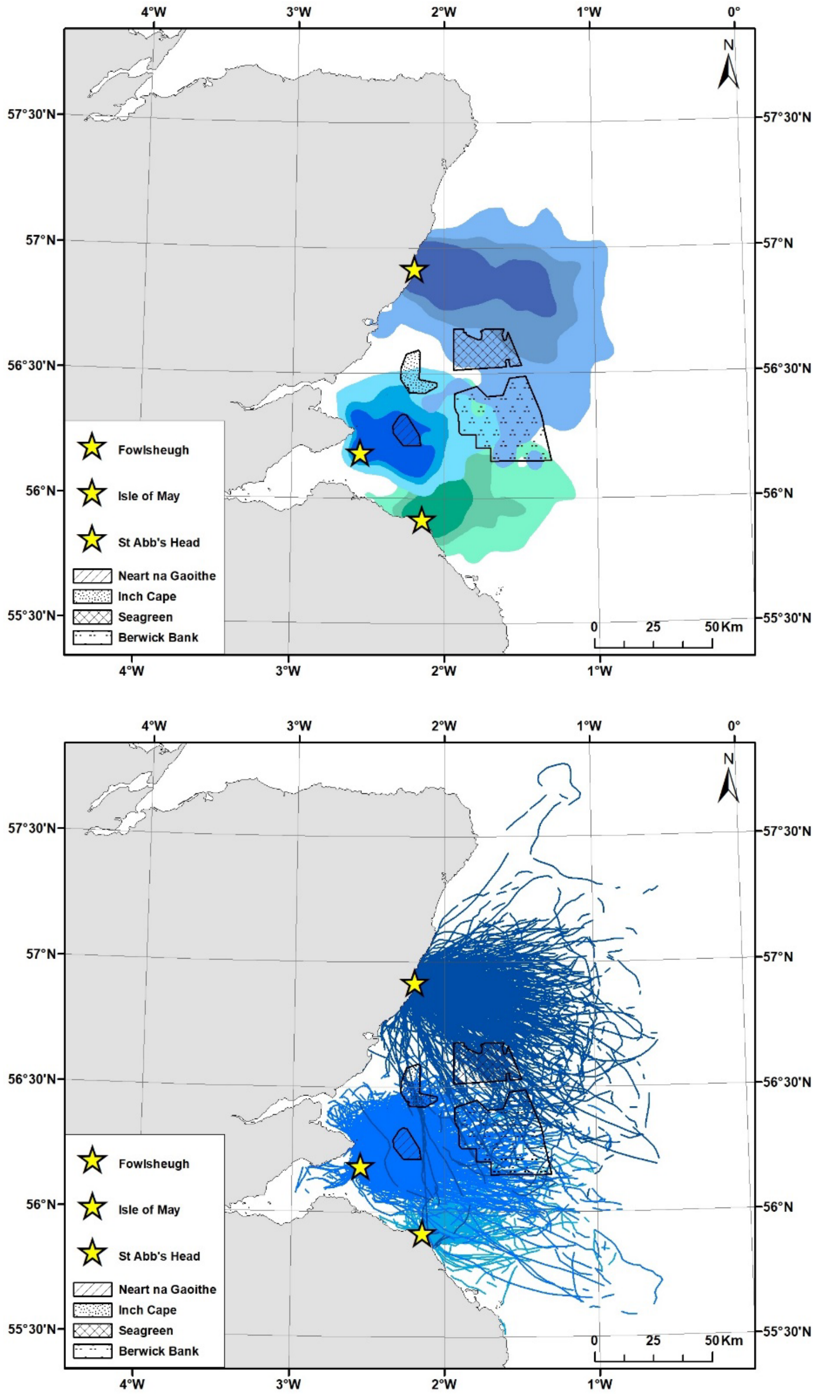


Figure 18: a) utilisation distributions (50%, 70%, 90% contours) and b) horizontal flight lines of kittiwakes from the three study colonies in 2021.

3.3 Minimum adequate sample size

In guillemots, the resampling procedure using 90% density contours indicated a relatively substantial increase of at-sea area used with sample size up to around 11 tagged birds, after which the increment with each additional tagged bird was smaller and eventually the area size appeared to plateau at a sample size of around 19 birds (Fig. 19). Randomized samples of 11 birds captured 91.7% and samples of 19 birds captured 99.1% of the area identified using all study birds.

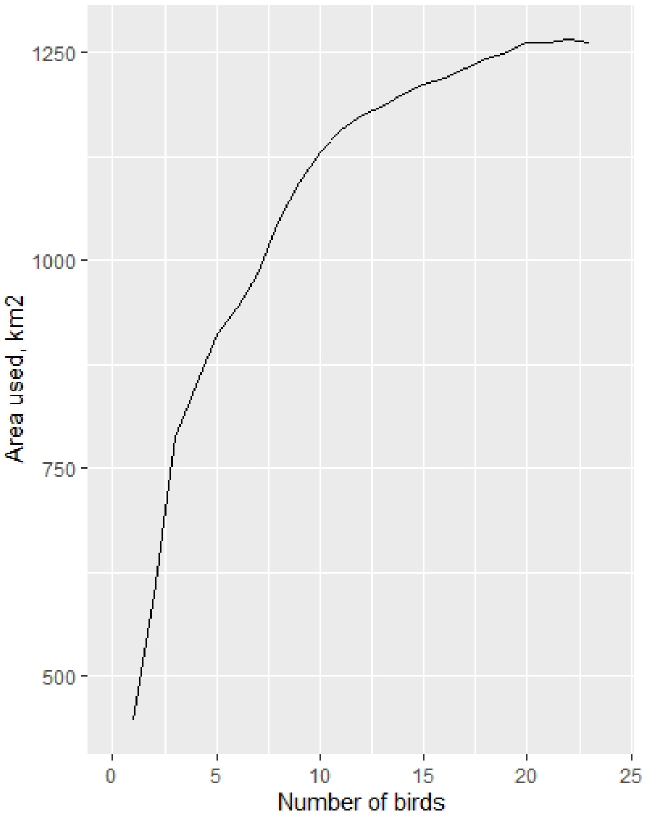


Figure 19: Relationship between at-sea area used and sample size of birds estimated from a resampling procedure in guillemots from the Isle of May. Median area shown for each randomized sample size.

In razorbills, a substantial increase of area used was observed only up to a sample size of 3 birds, after which there was little change in area size (Fig. 20), suggesting a high degree of consistency in at-sea distribution between tagged individuals. This pattern was reflected in the cumulative percentage of area used, with randomized samples of 3 birds covering 98.8% of the area identified using all study birds.

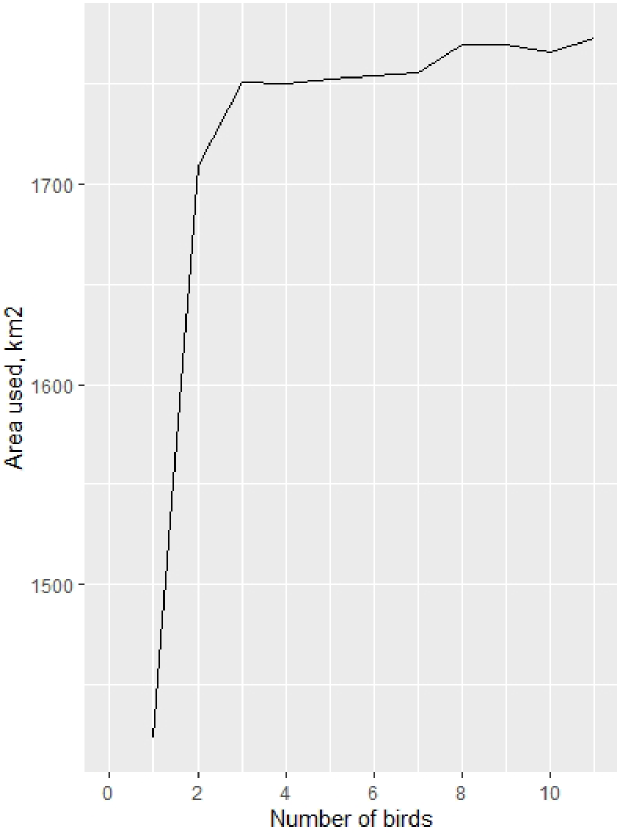


Figure 20: Relationship between at-sea area used and sample size of birds estimated from a resampling procedure in razorbills from the Isle of May. Median area shown for each randomized sample size.

In puffins, area used increased substantially up to a sample size of 11 tagged birds, after which the increment with each additional bird was small and eventually the area size appeared to plateau at a sample size of around 20 birds (Fig. 21). Randomized samples of 11 birds captured 95.7% and samples of 20 birds captured 99.4% of the area identified using all study birds.

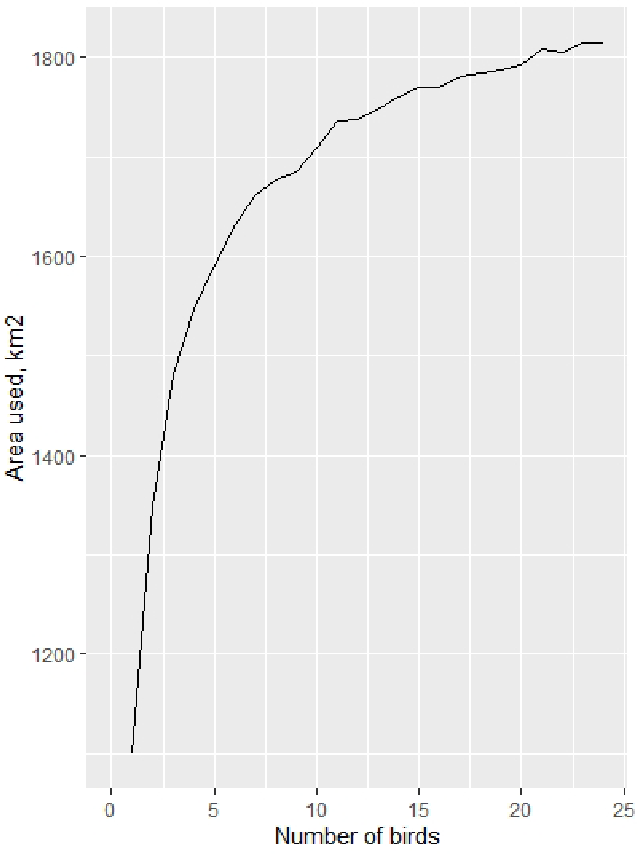


Figure 21: Relationship between at-sea area used and sample size of birds estimated from a resampling procedure in puffins from the Isle of May. Median area shown for each randomized sample size.

In kittiwakes, some colony differences in minimum adequate sample size were apparent. On the Isle of May, area used increased more substantially up to a sample size of 11 tagged birds, thereafter the increment with each additional bird was negligible (Fig. 22a). Randomized samples of 11 birds captured 98.5% of the area identified using all study birds. At Fowlsheugh and St Abb’s Head, area used increased more substantially up to a sample size of 19-20 birds, with a small increment thereafter up to a sample size of 27-29 birds after which the area size plateaued (Fig. 22b, c). At Fowlsheugh, randomized samples of 19 birds captured 95.8% and samples of 29 birds captured 99.6% of the area identified using all study birds. Corresponding figures for St Abb’s Head were 96.5% (20 birds) and 99.0% (27 birds).

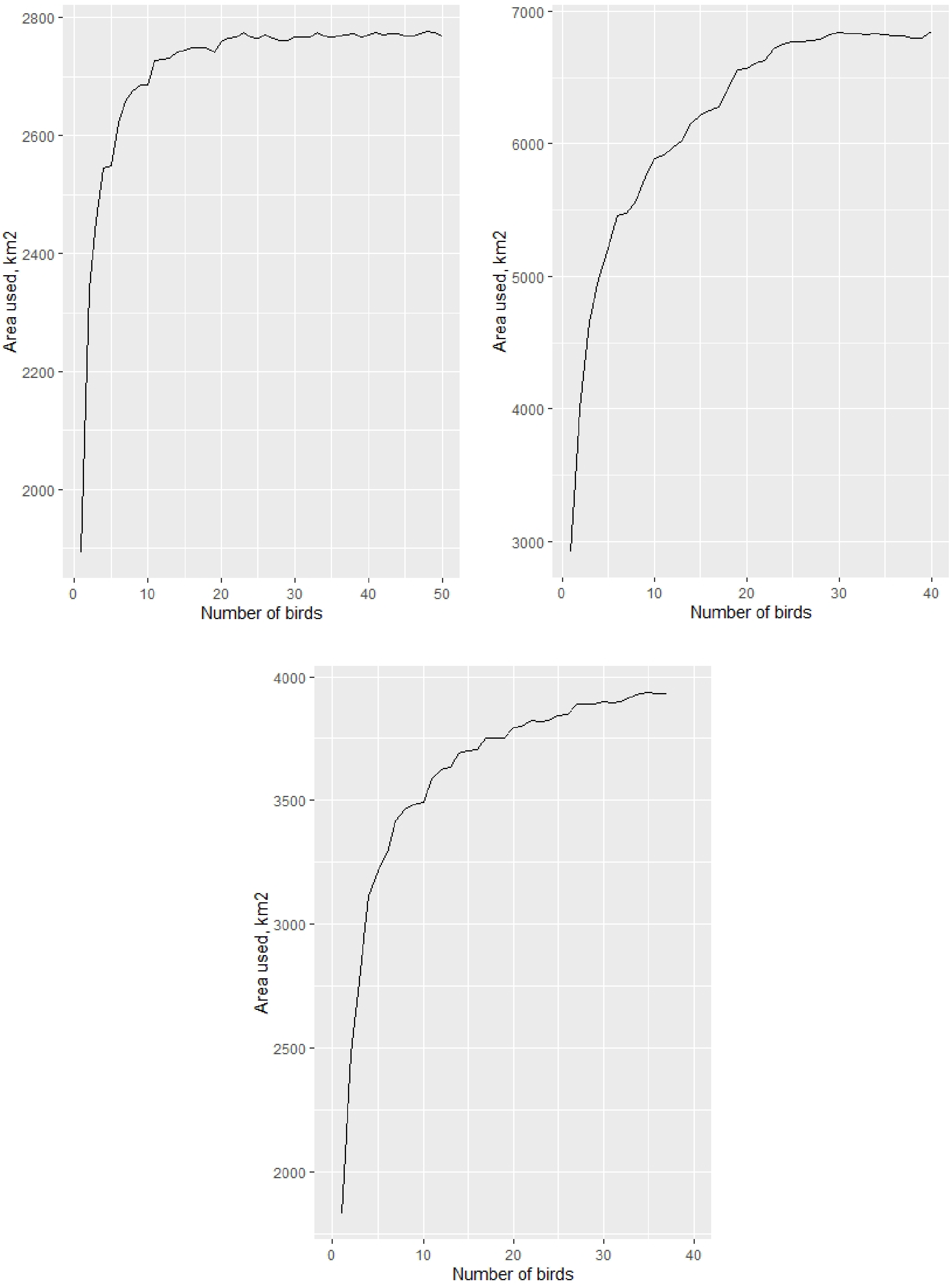


Figure 22: Relationship between at-sea area used and sample size of birds estimated from a resampling procedure in kittiwakes. Median area shown for each randomized sample size for a) the Isle of May, b) Fowlsheugh and c) St Abb's Head.

In summary, our results indicate that the sample sizes of tracked birds were adequate to estimate at-sea area used by the study populations of all four species during the deployment period. There was some variation between species and between populations of the same species (kittiwake only) in the sample sizes at which the area size appeared to plateau. For the auks, the minimum adequate sample size was around 20 birds, except in razorbills where it was substantially lower, likely reflecting high consistency in distributions between individuals of this species. For kittiwakes, the minimum adequate sample size was lower on the Isle of May (<20 birds) than at Fowlsheugh and St Abb’s Head (*ca*. 28 birds), suggesting higher consistency in individual distributions at this colony.

3.4 Overlap with the Forth/Tay wind farm footprints

3.4.1 Utilisation distribution contours

3.4.1.1 Neart na Gaoithe

The area and percentage overlap of UD contours associated with foraging and resting behaviours (or non-flight behaviours in the case of puffins) with the Neart na Gaoithe footprint are shown in Table 8. In guillemots, core areas (50% UD contours) associated with foraging and resting did not overlap with the wind farm footprint and the overall area used (90% UD contours) overlapped with it to a small extent (Table 8a). In razorbills, overlap of foraging areas (both core and overall) with Neart na Gaoithe was low but overlap of areas used for resting was slightly higher (Table 8b). Highest overlap among the study species (nearly 20%) was recorded in puffins, for core areas associated with non-flight behaviours (Table 8c). The areas used for foraging and resting by kittiwakes from Fowlsheugh did not overlap with the wind farm footprint (Table 8d). For kittiwakes from the Isle of May, overall areas used for foraging and resting overlapped little with Neart na Gaoithe, whereas core areas overlapped to a greater extent (Table 8d). There was virtually no overlap between the areas used by birds from St Abb’s Head and the wind farm (Table 8d). However, note that the entire footprint fell within the 50% and 90% UD contours of puffins and kittiwakes from the Isle of May, and large parts of the footprint fell within the 90% UD contours of razorbills from this colony.

Species UD	UD area (km^2)	Neart na Gaoithe	
		UD overlap (km^2)	UD overlap (%)
a) Guillemot			
Foraging			
50% contour	318.5	0	0
90% contour	1282.9	0	0
Resting at sea			
50% contour	332.4	0	0
90% contour	1499.3	27.4	1.8
b) Razorbill			
Foraging			
50% contour	631.1	7.1	1.1
90% contour	2137.8	71.6	3.4
Resting at sea			
50% contour	545.2	45.8	8.4
90% contour	1826.8	99.9	5.5
c) Puffin			
Non-flight behaviours (foraging and resting combined)			
50% contour	543.7	105.2	19.4
90% contour	1834.8	105.2	5.7
d) Kittiwake			
Foraging (Fowlsheugh)			
50% contour	1553.4	0	0
90% contour	7087.7	0	0
Resting at sea (Fowlsheugh)			
50% contour	1594.1	0	0
90% contour	6258.9	0	0
Foraging (Isle of May)			
50% contour	776.4	105.2	13.6
90% contour	2813.0	105.2	3.7
Resting at sea (Isle of May)			

50% contour	832.9	105.2	12.6
90% contour	2847.0	105.2	3.7
Foraging (St Abb's Head)			
50% contour	629.3	0	0
90% contour	4065.7	0.1	0.002
Resting at sea (St Abb's Head)			
50% contour	483.3	0	0
90% contour	3500.7	0.9	0.03

Table 8. Overlap between seabird utilisation distributions associated with key non-flight behaviours and Neart na Gaoithe OWF, expressed as area of overlap and % of the UD area covered by the windfarm footprint. Results are presented for core areas and overall area used (50% and 90% UD contours).

3.4.1.2 Seagreen

There was no overlap between the at-sea distributions (50% and 90% UD contours) of any of the species breeding on the Isle of May and the Seagreen footprint (Table 9a-d). The areas used for foraging and resting by kittiwakes from Fowlsheugh overlapped little with the wind farm (Table 9d). For kittiwakes from St Abb's Head foraging areas and core resting areas did not overlap with Seagreen, and overlap with the overall resting areas was low (<2%; Table 9d).

Species UD	UD area (km^2)	Seagreen	
		UD overlap (km^2)	UD overlap (%)
a) Guillemot			
Foraging			
50% contour	318.5	0	0
90% contour	1282.9	0	0
Resting at sea			
50% contour	332.4	0	0
90% contour	1499.3	0	0
b) Razorbill			

Foraging			
50% contour	631.1	0	0
90% contour	2137.8	0	0
Resting at sea			
50% contour	545.2	0	0
90% contour	1826.8	0	0
c) Puffin			
Non-flight behaviours (foraging and resting combined)			
50% contour	543.7	0	0
90% contour	1834.8	0	0
d) Kittiwake			
Foraging (Fowlsheugh)			
50% contour	1553.4	0	0
90% contour	7087.7	374.0	5.3
Resting at sea (Fowlsheugh)			
50% contour	1594.1	2.3	0.1
90% contour	6258.9	288.8	4.6
Foraging (Isle of May)			
50% contour	776.4	0	0
90% contour	2813.0	0	0
Resting at sea (Isle of May)			
50% contour	832.9	0	0
90% contour	2847.0	0	0
Foraging (St Abb's Head)			
50% contour	629.3	0	0
90% contour	4065.7	0	0
Resting at sea (St Abb's Head)			
50% contour	483.3	0	0
90% contour	3500.7	55.5	1.6

Table 9. Overlap between seabird utilisation distributions associated with key non-flight behaviours and Seagreen OWF, expressed as area of overlap and % of the

UD area covered by the windfarm footprint. Results are presented for core areas and overall areas used (50% and 90% UD contours).

3.4.1.3 Berwick Bank

There was no overlap between the at-sea distribution (50% and 90% UD contours) of guillemots from the Isle of May and the Berwick Bank footprint (Table 10a). For razorbills and puffins from the Isle of May, core areas (50% UD contours) did not overlap with the wind farm footprint and the overall areas (90% UD contours) overlapped with it to a limited extent (Table 10b,c). There was no or virtually no overlap between the core areas of kittiwakes from all three colonies and Berwick Bank, however the overall areas used overlapped with it to some extent, with highest overlap (~20%) recorded in birds from St Abb's Head (Table 10d).

Species UD	UD area (km^2)	Berwick Bank	
		UD overlap (km^2)	UD overlap (%)
a) Guillemot			
Foraging			
50% contour	318.5	0	0
90% contour	1282.9	0	0
Resting at sea			
50% contour	332.4	0	0
90% contour	1499.3	0	0
b) Razorbill			
Foraging			
50% contour	631.1	0	0
90% contour	2137.8	28.1	1.3
Resting at sea			
50% contour	545.2	0	0
90% contour	1826.8	36.4	2.0
c) Puffin			
Non-flight behaviours (foraging and resting combined)			

50% contour	543.7	0	0
90% contour	1834.8	27.4	1.5
d) Kittiwake			
Foraging (Fowlsheugh)			
50% contour	1553.4	0	0
90% contour	7087.7	626.4	8.8
Resting at sea (Fowlsheugh)			
50% contour	1594.1	0	0
90% contour	6258.9	601.1	9.6
Foraging (Isle of May)			
50% contour	776.4	0	0
90% contour	2813.0	364.2	13.0
Resting at sea (Isle of May)			
50% contour	832.9	2.9	0.4
90% contour	2847.0	476.5	16.7
Foraging (St Abb’s Head)			
50% contour	629.3	0	0
90% contour	4065.7	804.5	19.8
Resting at sea (St Abb’s Head)			
50% contour	483.3	0	0
90% contour	3500.7	682.5	19.5

Table 10. Overlap between seabird utilisation distributions associated with key non-flight behaviours and Berwick Bank OWF, expressed as area of overlap and % of the UD area covered by the windfarm footprint. Results are presented for core areas and overall areas used (50% and 90% UD contours).

3.4.2 Horizontal flight trajectories

The proportion of individuals and flights passing through the Forth/Tay wind farm footprints is shown in Table 11.

3.4.2.1 Neart na Gaoithe

In terms of number of birds, three of the species breeding on the Isle of May used the wind farm extensively, in particular puffins and kittiwakes where 96% and 100% of the study birds, respectively, passed through the footprint at least once. In comparison, this figure was 13% for guillemots. At the level of individual flights, overlap was similarly higher in Isle of May kittiwakes and puffins compared to the other two auk species. The lower overlap of guillemot and razorbill flight activities with Neart na Gaoithe most likely reflects their partially inshore distribution.

There was no overlap of at-sea areas used by kittiwakes from Fowlsheugh and the Neart na Gaoithe footprint. Overlap of kittiwakes from St Abb's Head with the footprint was limited, with a single flight by one individual recorded passing through (Table 11d).

3.4.2.2 Seagreen

There was no overlap between the at-sea distributions of guillemots and razorbills from the Isle of May and the Seagreen footprint. Puffin flight activities overlapped little with the Seagreen footprint (Table 11c). Similarly, overlap with flight activities of kittiwakes from the Isle of May was low (Table 11d). In contrast, kittiwakes from Fowlsheugh used the wind farm more extensively, with 58% of the study birds passing through it at least once. At the level of individual flights, overlap was similarly higher in kittiwakes from this colony compared to birds from the other two colonies (Table 11d). There was no overlap between the at-sea areas used by kittiwakes from St Abb's Head and the Seagreen footprint.

3.4.2.3 Berwick Bank

No commuting flight activities of guillemots and razorbills from the Isle of May were recorded within the Berwick Bank footprint. Puffin flight activities overlapped relatively little with the planned wind farm (Table 11c). Kittiwakes from all three colonies used the wind farm area, with highest proportion of birds and flights from the Isle of May passing through it, followed by birds/flights from Fowlsheugh and St Abb's Head (Table 11d).

Category	n	% passing through Neart na Gaoithe	% passing through Seagreen	% passing through Berwick Bank

a) Guillemot				
Bird	23	13	0	0
Flight	278	2	0	0
b) Razorbill				
Bird	11	64	0	0
Flight	272	6	0	0
c) Puffin				
Bird	24	96	4	17
Flight	713	16	0.6	2
d) Kittiwake				
Fowlsheugh				
Bird	40	0	58	35
Flight	1364	0	7	5
Isle of May				
Bird	50	100	8	72
Flight	2271	18	0.3	8
St Abb’s Head				
Bird	37	3	0	11
Flight	1336	0.1	0	3

Table 11. Percentage of birds and flights passing through the Forth/Tay windfarm footprints for each species and colony.

Figure 23 shows the distribution of commuting flights that overlapped with the Forth/Tay windfarm footprints. The correspondence between flight directions and the location of the breeding colonies is apparent for all species and colonies.

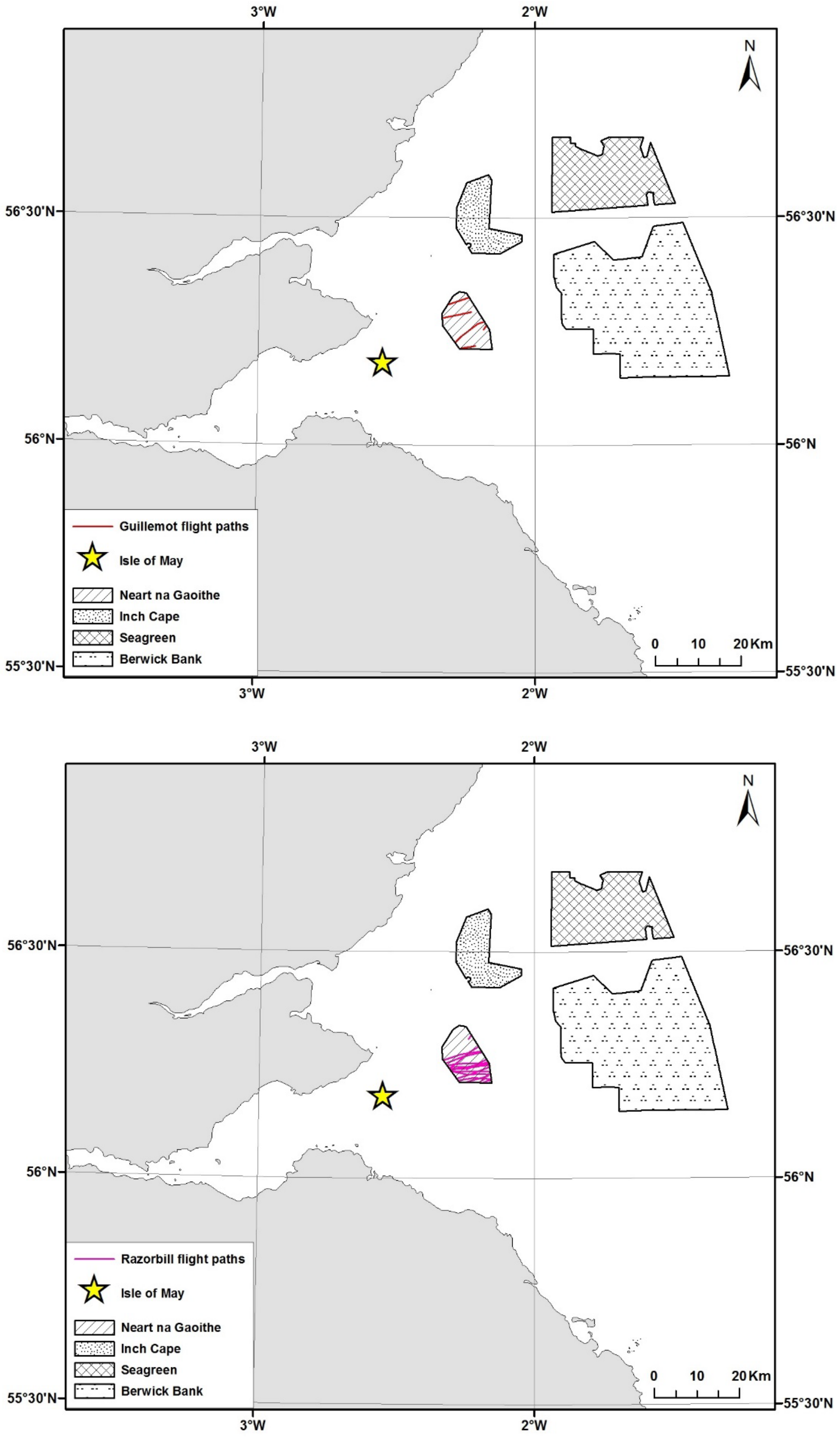


Figure 23: Flights passing through the Forth/Tay wind farm footprints for a) Isle of May guillemots and b) Isle of May razorbills.

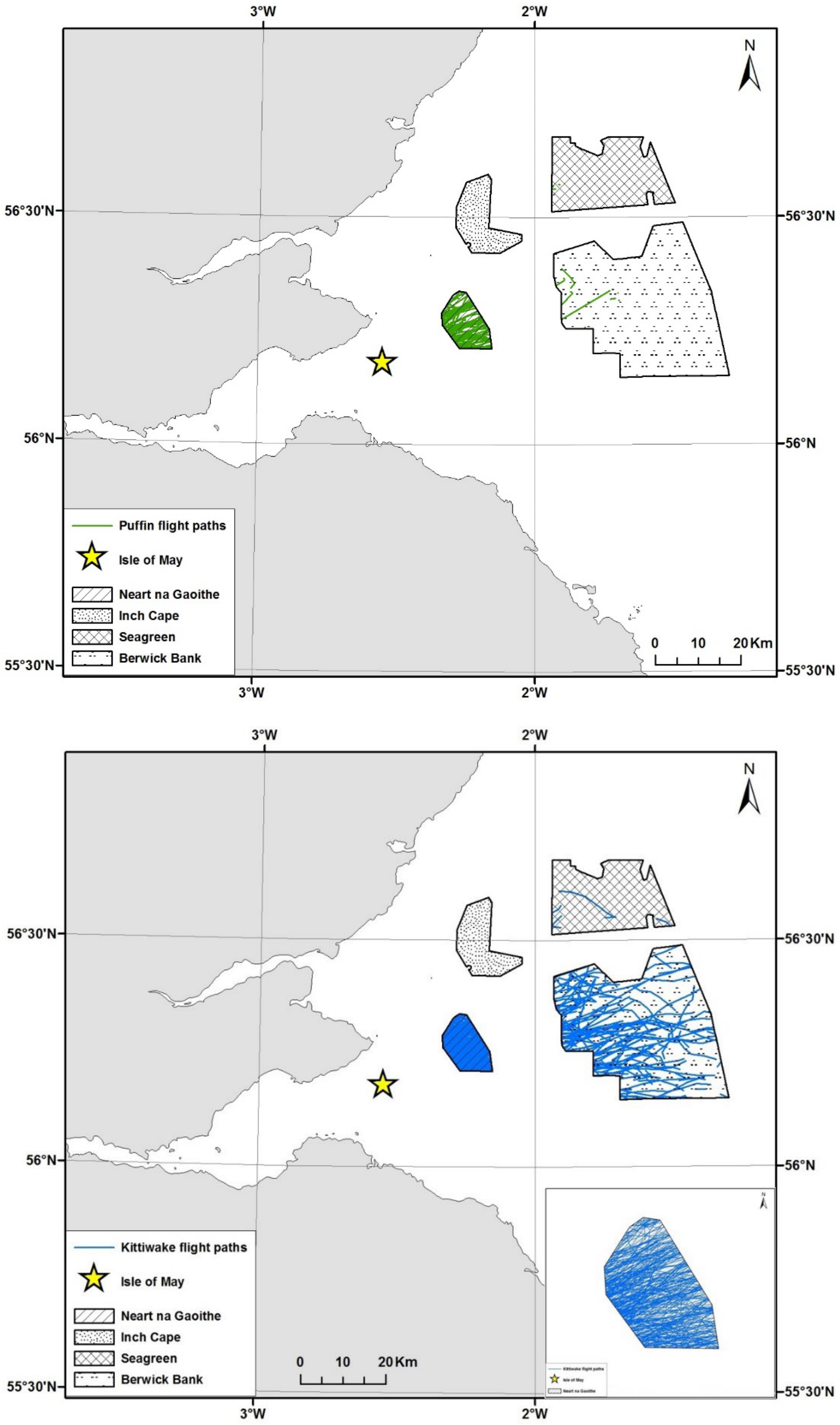
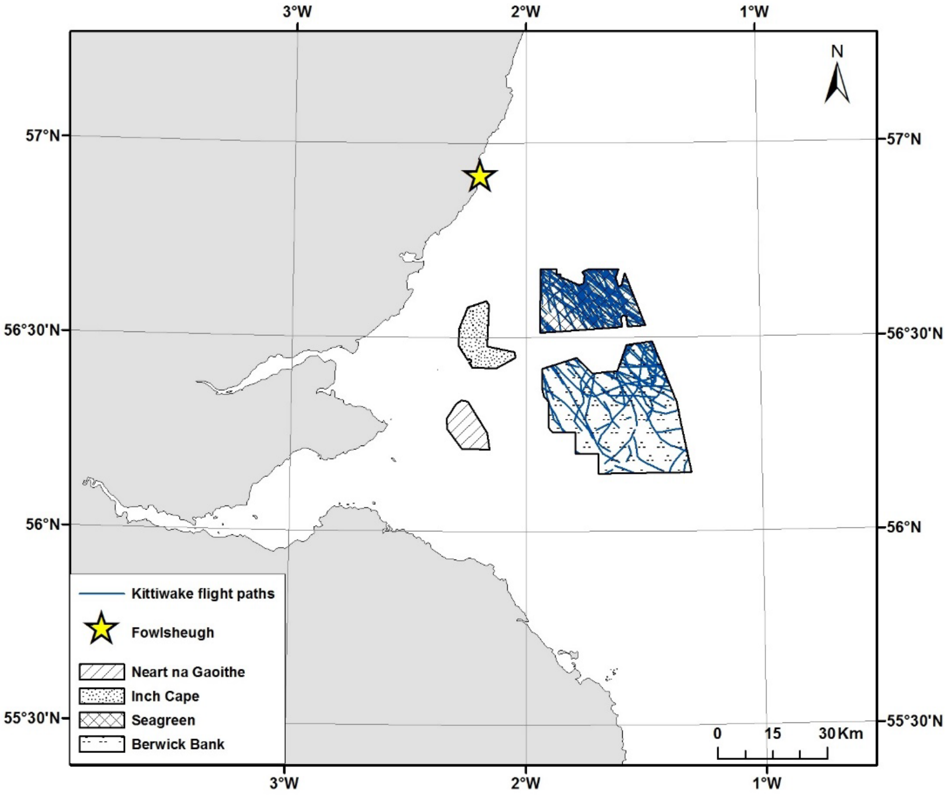


Figure 23 (cont.): Flights passing through the Forth/Tay wind farm footprints for c) Isle of May puffins and d) Isle of May kittiwakes. Inset: flights passing through Neart na Gaoithe shown in more detail.



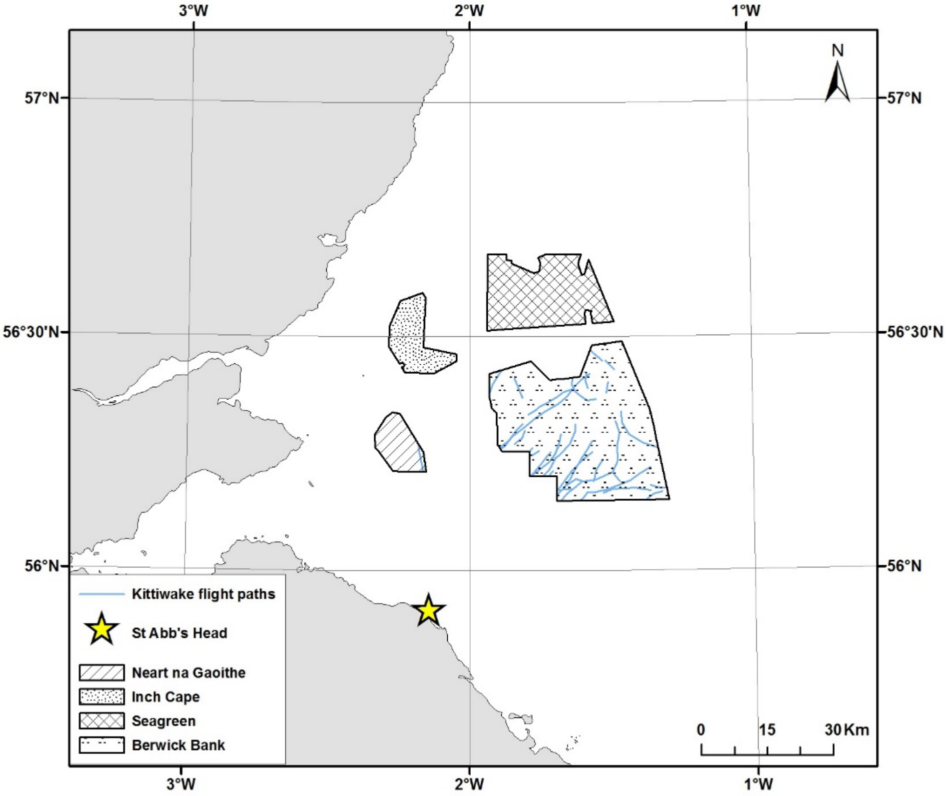


Figure 23 (cont.): Flights passing through the Forth/Tay wind farm footprints for e) Fowlsheugh kittiwakes and f) St Abb's Head kittiwakes.

3.5 Kittiwake flight height

3.5.1 Correlations between kittiwake flight height estimates

At the Fowlsheugh, Isle of May and St Abb's Head colonies, estimates of kittiwake flight height based on sea-level pressure were moderately to highly correlated across all samples and across a subset of samples in which sampling frequency was ≤ 16 secs (Figs 24, 25 & 26). This included estimates of pressure based on calibration of tags using records of surface pressure from UvA tags while birds were classified as floating and those based on pressure data from ERA5. Correlations between GPS-derived and sea level pressure-derived estimates of altitude were slightly smaller; however, results also differed between the colonies examined. For example, the correlation between GPS and barometric-derived altitudes was generally greater at St Abb's Head than either Fowlsheugh or the Isle of May. However, at Fowlsheugh, the correlation between GPS and barometer derived flight heights was moderate when examining all recorded estimates but increased when focussing on estimates obtained during high sampling frequencies. In contrast, at St Abb's Head, the correlation between GPS and barometric flight heights was greater than that observed at Fowlsheugh when considering all estimates (St Abb's Head, r=0.65 vs. Fowlsheugh, r=0.38). However, the correlation between GPS and barometer derived flight height estimates was weaker when focussing on high frequency samples at St Abb's Head, the opposite pattern to that observed at Fowlsheugh. Sampling frequencies of ≤ 16 seconds or less were not observed at the Isle of May so we have no data on how this high sampling rate influenced the correlation between GPS and barometer-derived altitudes at this colony.

As well as observing different patterns of correlations between GPS and barometric flight heights across colonies there was also evidence of variation in the correlations among birds tracked at the same colony. Table S1 demonstrates that while GPS and barometric flight height (using the single logger approach) were highly correlated in some birds, estimates were less correlated in others.

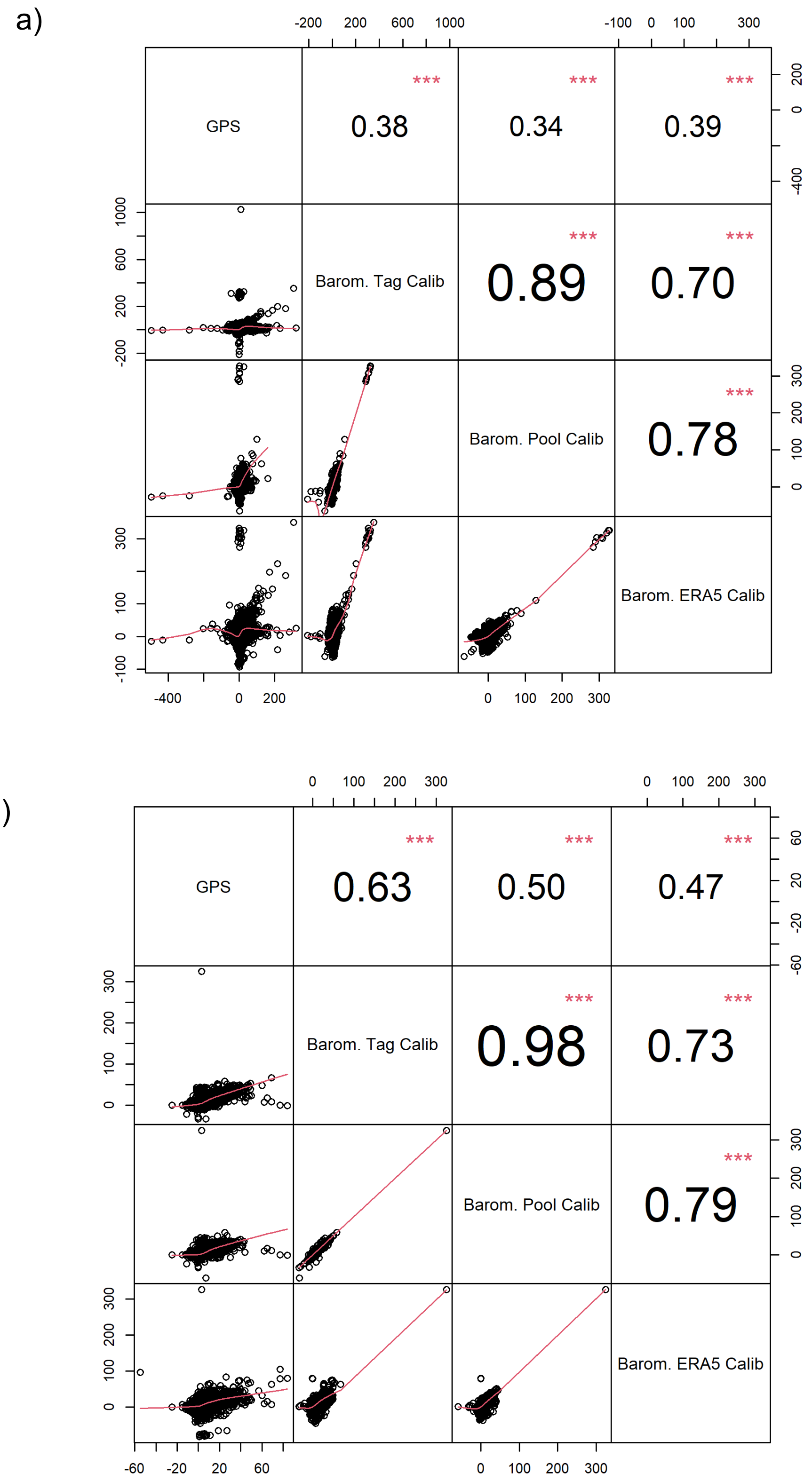


Figure 24: Correlations between kittiwake flight height estimates from GPS and barometric sensors from birds tracked at Fowlsheugh. Plots show correlations at different sampling frequencies: a) all samples; b) sampling frequency ≤ 16 secs.

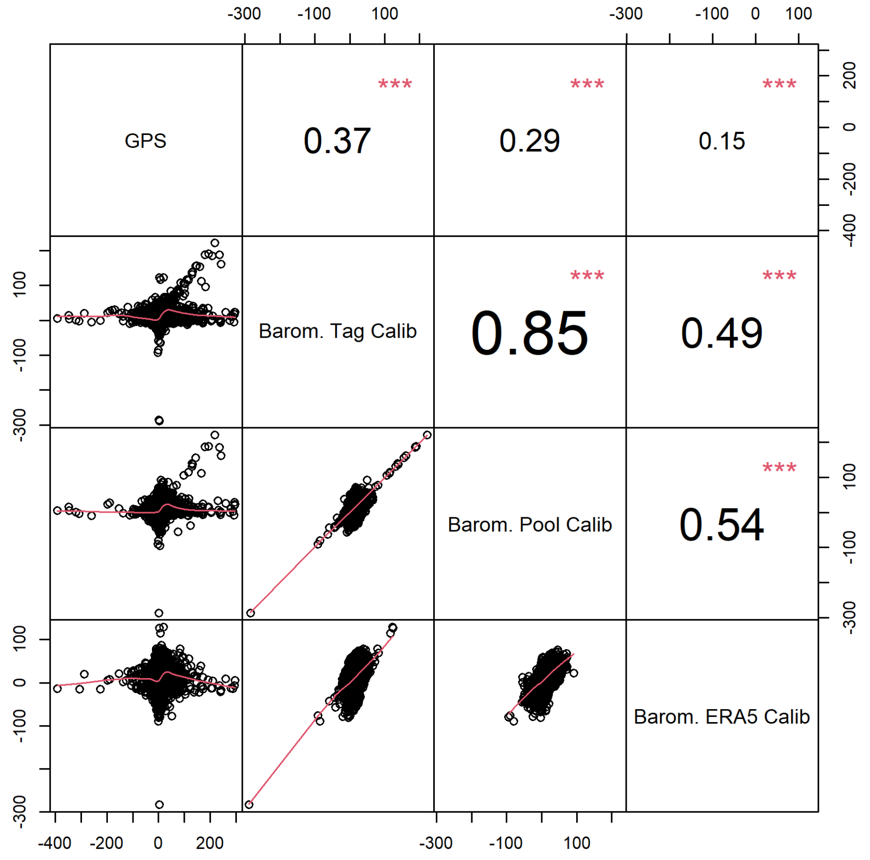


Figure 25: Correlations between kittiwake flight height estimates from GPS and barometric sensors from birds tracked at the Isle of May, 2021. Plots show correlations across all samples. Note that at the Isle of May sampling frequencies ≤ 16 seconds were not observed.

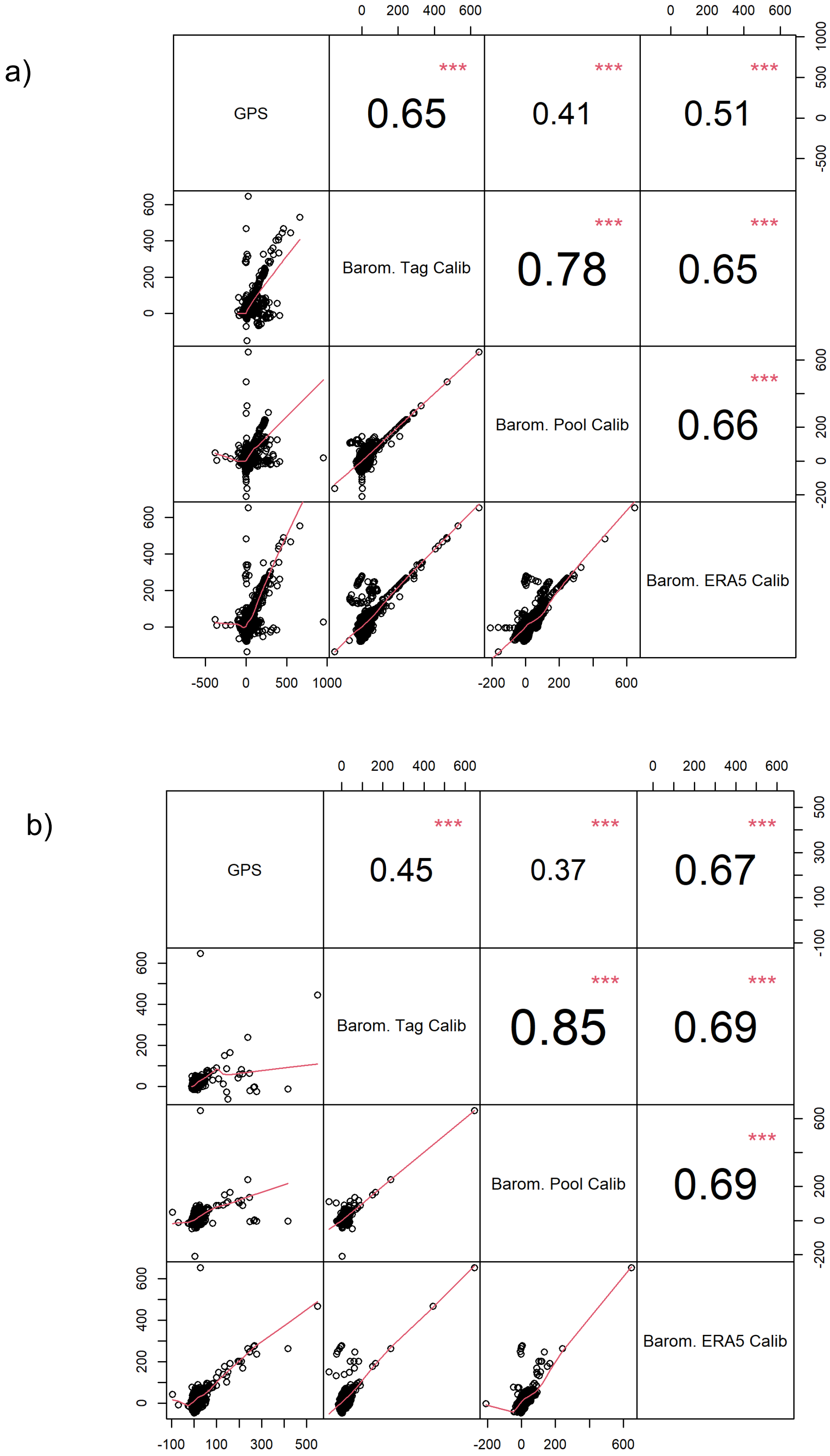


Figure 26: Correlations between kittiwake flight height estimates from GPS and barometric sensors from birds tracked at St Abb's Head. Plots show correlations at different sampling frequencies: a) all samples; b) sampling frequency ≤ 16 secs.

3.5.2 Comparing flight height distributions

Estimated altitude distributions appeared similar whether they were based on GPS or barometric methods at all colonies with the potential exception of estimates based on ERA5 data (Figs 27, 28 & 29). Distributions based on ERA5 data tended to be more evenly distributed than the other methods, particularly at the Isle of May and St Abb’s Head. In all cases the bulk of estimates fell with the range -10m – 50 m (Tables 12, 13, 14). Similar patterns are also seen when focussing on altitude estimates which were classified as floating, though in this case estimates are tightly distributed around 0 (Fig. 30), as expected given the assumed behaviour. The one exception to this was estimates based on ERA5 pressure calibration at St Abb’s Head, where the median estimate for flight altitude during floating periods was 3.38 metres. Similarly, altitude distributions based purely on observations classified as in-flight appeared similar across each colony (Fig. 31), though estimates based on barometric loggers tended to be slightly higher than those based on GPS. In addition, extreme values of altitude were occasionally observed leading to minimum and maximum values of large magnitude regardless of which method was used to estimate altitude.

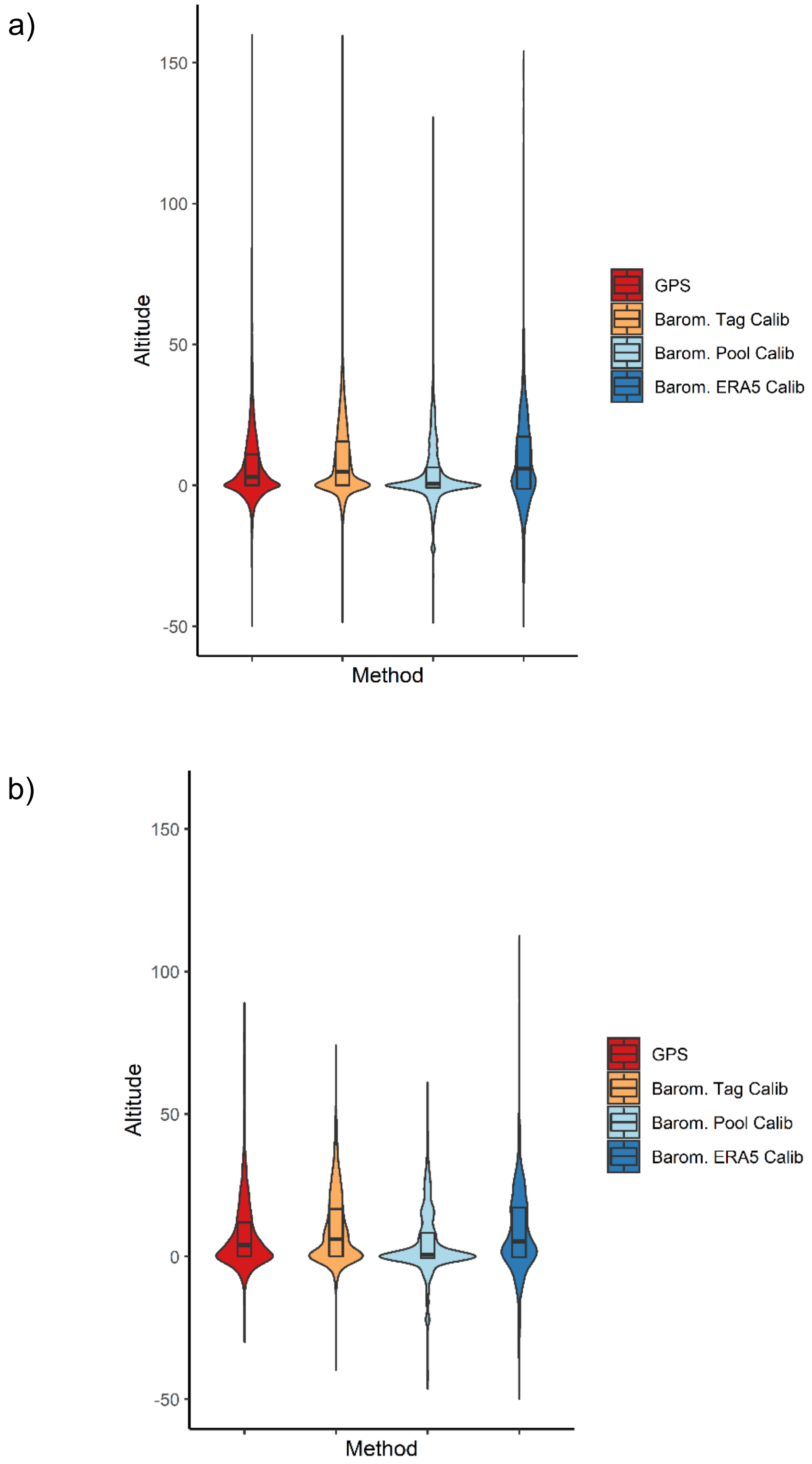


Figure 27: Violin plots of estimated kittiwake altitude distributions based on both GPS and barometric measurements at Fowlsheugh based on a) all observed estimates and b) estimates observed when sampling frequency was ≤ 16 seconds. Plots includes a box denoting the inter-quartile range and median value.

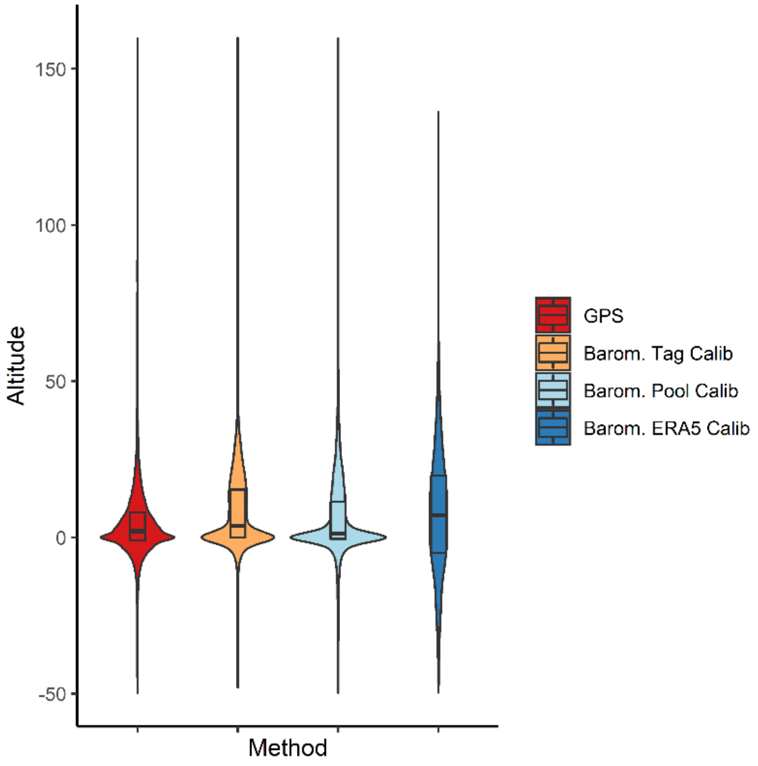


Figure 28: Violin plots of estimated kittiwake altitude distributions based on both GPS and barometric measurements at the Isle of May based on all observed estimates. Plots includes a box denoting the inter-quartile range and median value.

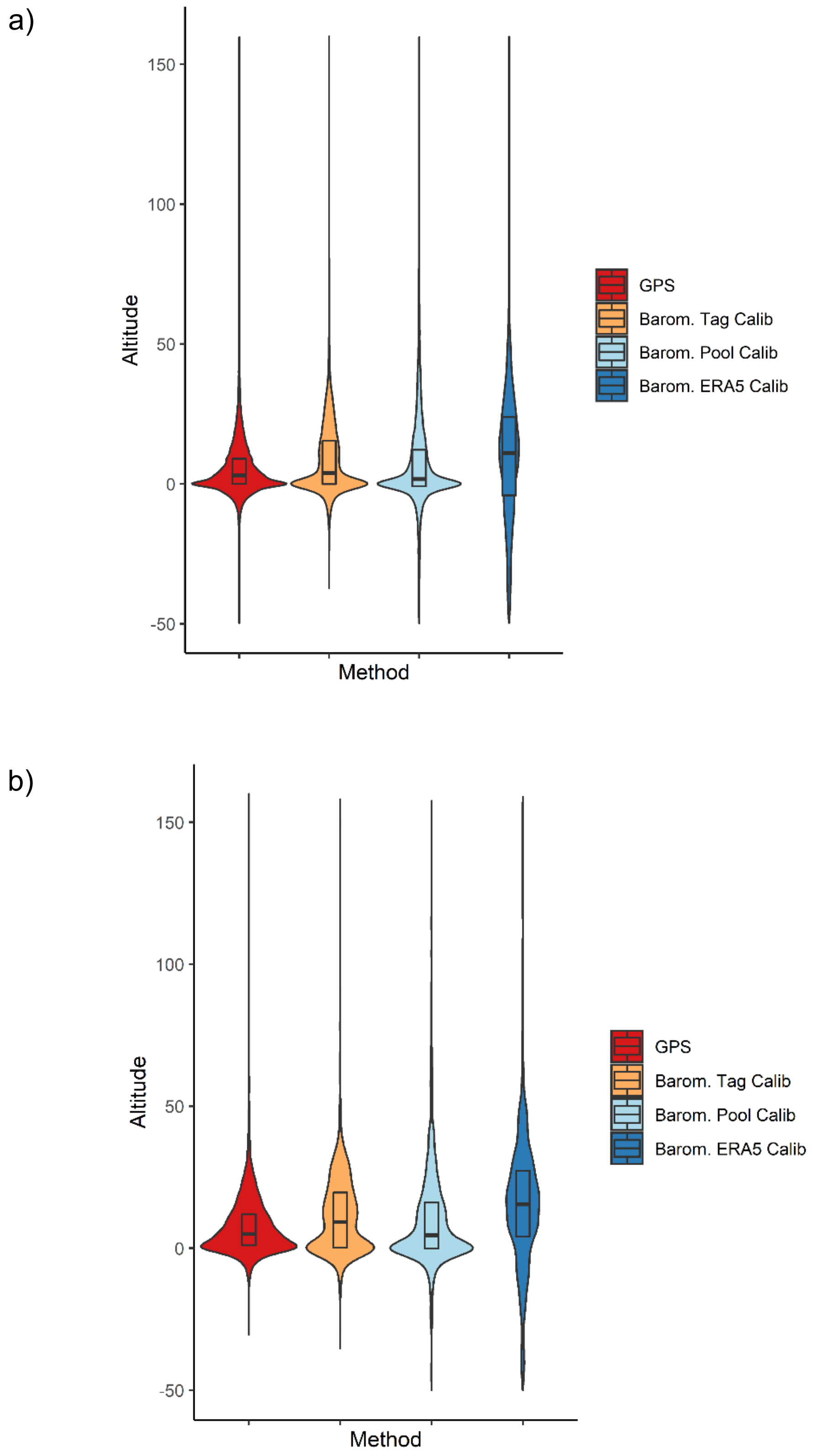


Figure 29: Violin plots of estimated kittiwake altitude distributions based on both GPS and barometric measurements at St Abb’s Head based on a) all observed estimates and b) estimates observed when sampling frequency was ≤ 16 seconds. Plots includes a box denoting the inter-quartile range and median value.

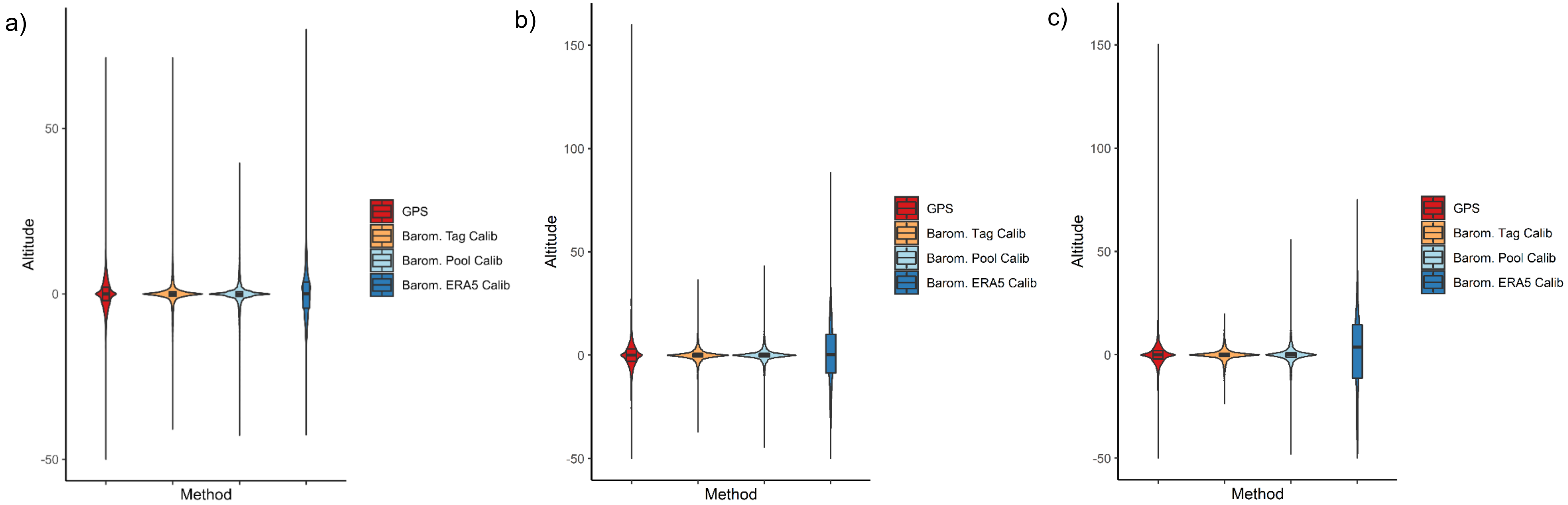


Figure 30: Violin plots of estimated kittiwake altitude distributions based on both GPS and barometric measurements at a) Fowlsheugh, b) Isle of May, and c) St Abb's Head based on all estimates calculated when birds were classified as floating. Plots includes a box denoting the inter-quartile range and median value.

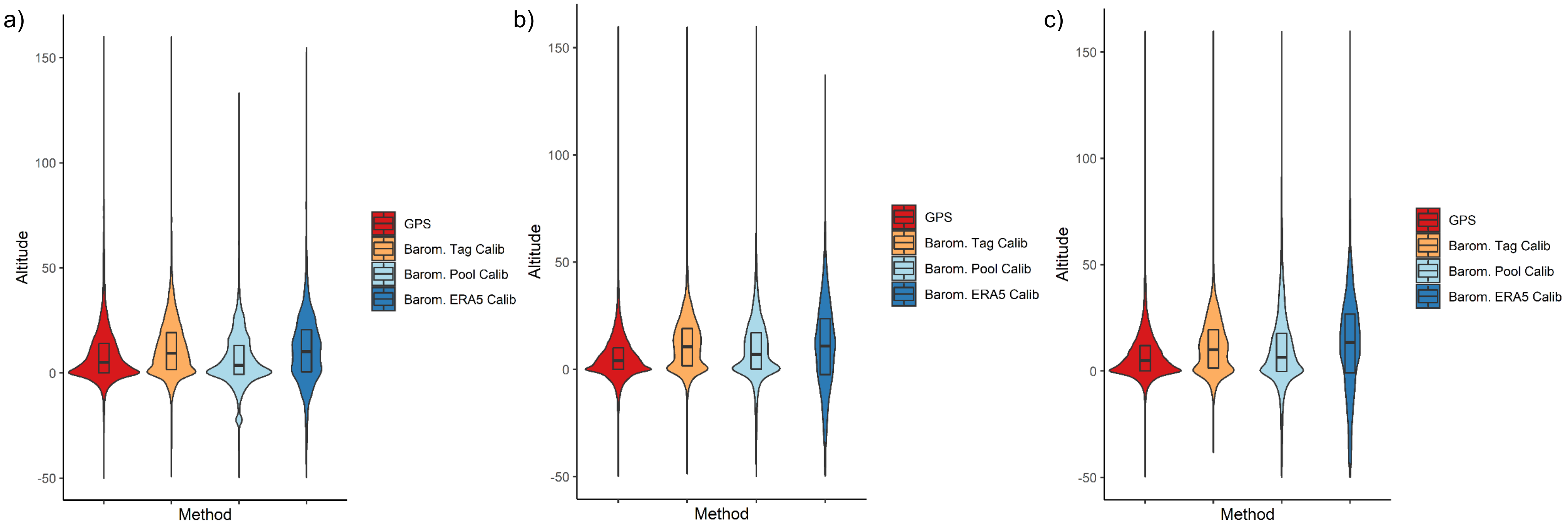


Figure 31: Violin plots of estimated kittiwake altitude distributions based on both GPS and barometric measurements at a) Fowlsheugh, b) Isle of May, and c) St Abb's Head based on all estimates calculated when birds were classified as 'in flight'. Plots includes a box denoting the inter-quartile range and median value.

Method	Behaviour	Min.	1st Quantile	Median	3rd Quantile	Max.
GPS	Floating	-50.00	-2.01	0.001	2.02	70.01
Barom. Tag Calib		-210.19	-0.69	-0.0101	0.66	324.92
Barom. Pool Calib		-46.56	-0.91	-0.0092	0.92	326.92
ERA5 Calib		-57.29	-4.28	0.024	3.59	326.35
GPS	In flight	-495.00	0.02	5.01	14.03	320.00
Barom. Tag Calib		-103.84	1.66	9.38	19.20	925.34
Barom. Pool Calib		-64.83	-0.69	3.59	13.15	325.37
ERA5 Calib		-94.07	0.44	10.05	20.50	350.46

Table 12. Flight altitude estimates for kittiwakes tracked at Fowlsheugh. Table shows estimates for floating and in-flight behaviour separately. Note birds assumed to be in flight if behaviour was classified as any behaviour other than floating.

Method	Behaviour	Min.	1st Quantile	Median	3rd Quantile	Max.
GPS	Floating	-320.01	-3.02	0.001	3.03	231.05
Barom. Tag Calib		-91.89	-0.90	0.0001	0.86	35.89
Barom. Pool Calib		-90.92	-0.85	-0.015	0.85	42.67
ERA5 Calib		-76.49	-8.69	0.14	10.14	81.57
GPS	In flight	-393.00	-0.01	4.02	10.03	296.00
Barom. Tag Calib		-288.19	1.69	10.43	19.07	221.61
Barom. Pool Calib		-286.58	0.13	6.93	17.06	220.89
ERA5 Calib		-283.74	-2.54	10.86	23.59	129.36

Table 13. Flight altitude estimates for kittiwakes tracked at the Isle of May. Table shows estimates for floating and in-flight behaviour separately. Note birds assumed to be in flight if behaviour was classified as any behaviour other than floating.

Method	Behaviour	Min.	1st Quantile	Median	3rd Quantile	Max.
GPS	Floating	-49.02	-2.02	0.011	2.03	275.01
Barom. Tag Calib		-26.79	-0.92	-0.0083	0.87	288.17
Barom. Pool Calib		-111.17	-1.17	-0.097	1.06	287.43
ERA5 Calib		-78.38	-12.05	3.38	14.35	280.22
GPS	In flight	-823.00	-0.01	5.01	12.02	951.00
Barom. Tag Calib		-149.33	1.76	10.26	19.14	645.13
Barom. Pool Calib		-208.59	-0.19	6.53	17.78	643.53
ERA5 Calib		-135.84	-12.056	13.42	14.36	653.18

Table 14. Flight altitude estimates for kittiwakes tracked at St Abb's Head. Table shows estimates for floating and in-flight behaviour separately. Note birds assumed to be in flight if behaviour was classified as any behaviour other than floating.

While flight altitude distributions appeared broadly similar and estimates of altitude were often highly correlated, there remained point-wise differences in altitude estimates at the same focal timestamp. When birds were classified as floating, point-wise differences between GPS and barometric altitude were generally tightly distributed around 0, although in some isolated cases differences could be substantial (Figs 32, 33 & 34). When birds were not floating (and therefore assumed to be flying) differences between GPS and barometric based estimates were less tightly distributed and were centred on positive values, which suggested a tendency for barometric based estimates of altitude to be greater than corresponding GPS estimates. In general, the median difference between GPS and barometric estimates of flight height when birds were in flight was 3.85 metres at Fowlsheugh, 5.79 m on the Isle of May, and 4.19 m at St Abb's Head (Table 15). Similar patterns were evident when focussing only on estimates from high frequency sampling periods at Fowlsheugh and St Abb's Head. However, it can also be seen that the maximum differences observed between GPS and barometric estimates were smaller at higher sampling frequencies (Figs 32 & 34).

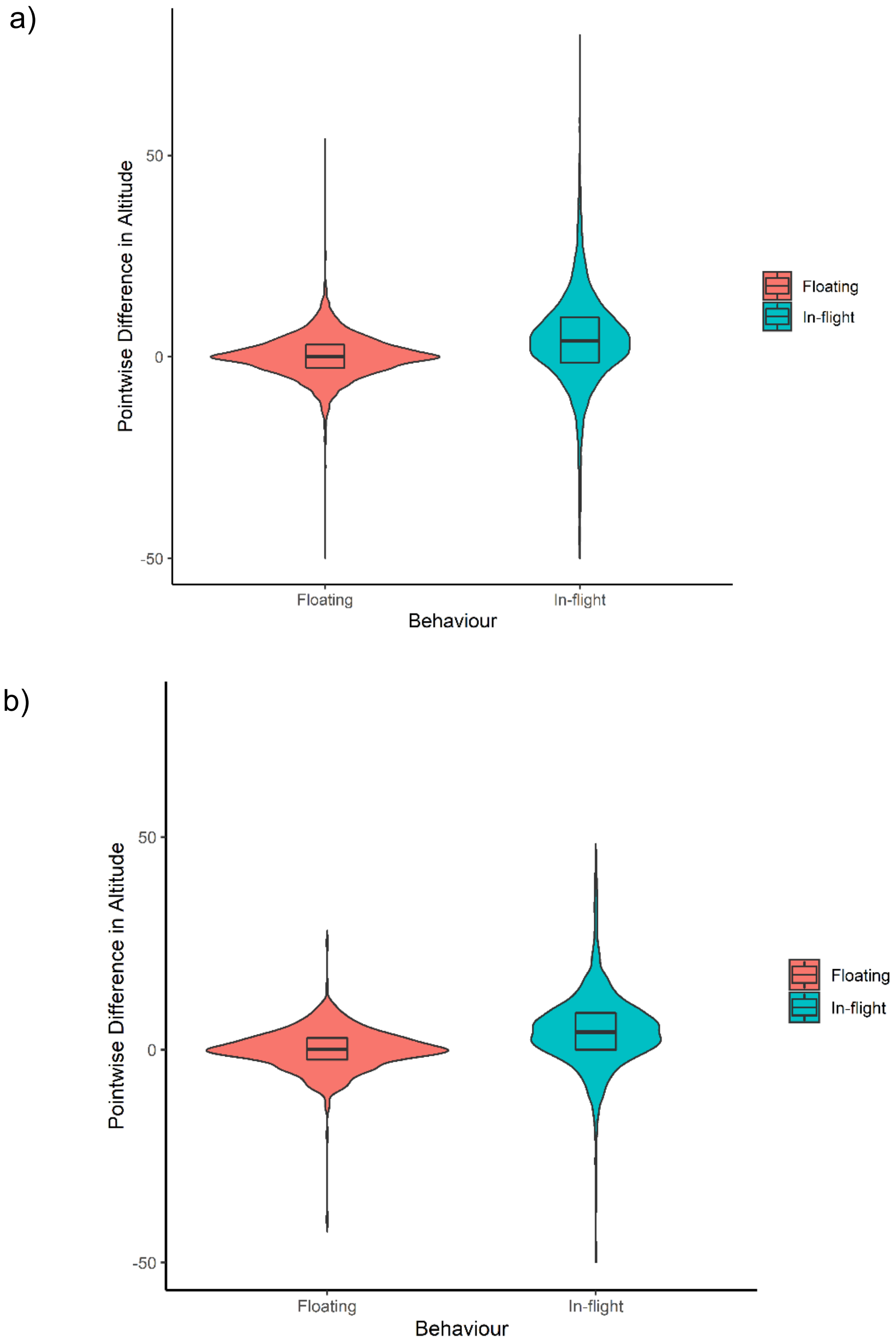


Figure 32. Violin plots showing the difference in point-wise estimates of kittiwake altitude at Fowlsheugh based on GPS derived measures of altitude and those based on barometric pressure using a single logger to calibrate surface pressure. a) all estimates; b) estimates when sampling frequency ≤ 16 seconds. Plots include a box denoting the inter-quartile range and median value.

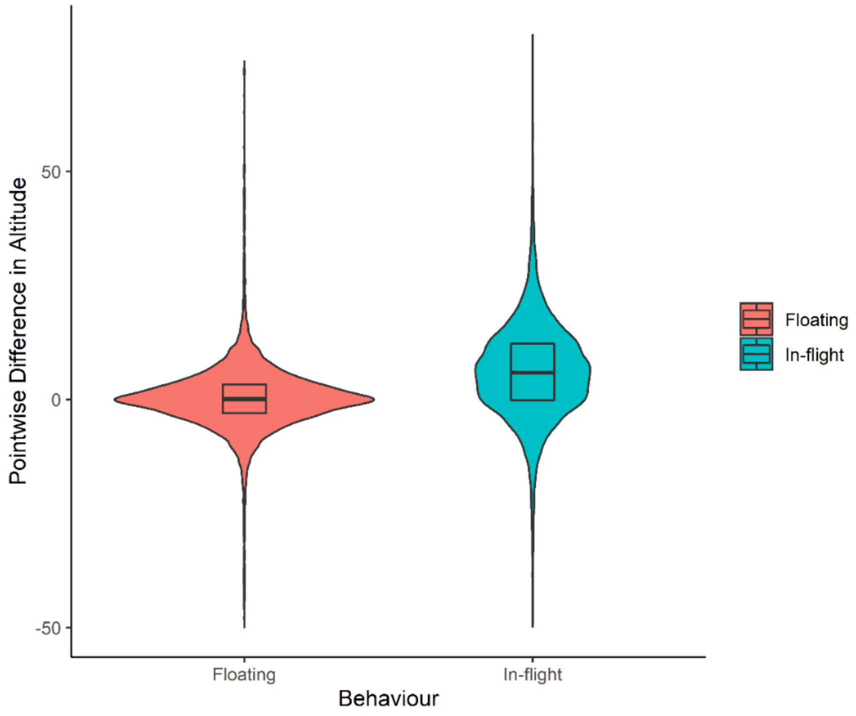


Figure 33. Violin plots showing the difference in point-wise estimates of kittiwake altitude at the Isle of May based on GPS derived measures of altitude and those based on barometric pressure using a single logger to calibrate surface pressure for all estimates.

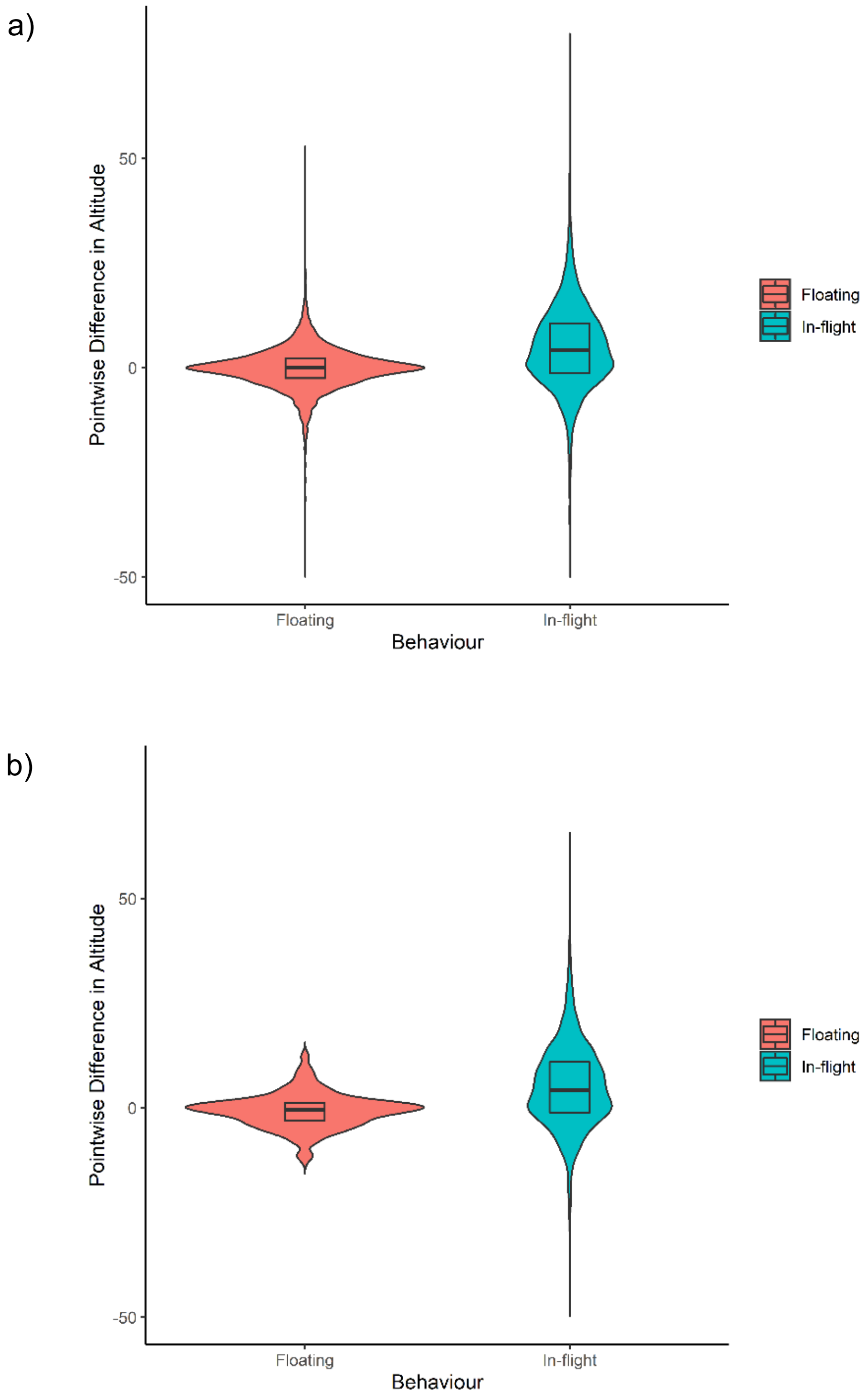


Figure 34: Violin plots showing the difference in point-wise estimates of altitude at St Abb's Head based on GPS derived measures of altitude and those based on barometric pressure using a single logger to calibrate surface pressure. a) all estimates; b) estimates when sampling frequency ≤ 16 seconds. Plots include a box denoting the inter-quartile range and median value.

Colony	Behaviour	Min.	1st Quantile	Median	3rd Quantile	Max.
Fowlsheugh	Floating	-210.19	-1.65	0.089	9.73	322.79
	In flight	-306.11	-2.68	3.85	2.98	916.35
Isle of May	Floating	-228.28	-3.03	0.053	3.32	321.21
	In flight	-298.23	-0.30	5.79	12.27	398.88
St Abb's Head	Floating	-211.46	-2.47	0.022	2.50	51.28
	In flight	-926.71	-1.19	4.19	10.32	845.33

Table 15. Point-wise differences in flight altitude estimates based on either GPS or barometric pressure from a single logger. Table shows estimates for floating and non-floating behaviour separately.

Focussing on point-wise differences for individual loggers, differences when birds were classified as floating were centred at 0 for all individuals though there was a greater spread around this central value in some individuals (Figs 35, 36, & 37). Barometric derived estimates of altitude also tended to be slightly higher than corresponding GPS estimates across all individuals at each colony. Thus, even across loggers and colonies barometric altitudes consistently report higher altitude estimates. The exact magnitude of this difference, however, did vary widely across individuals.

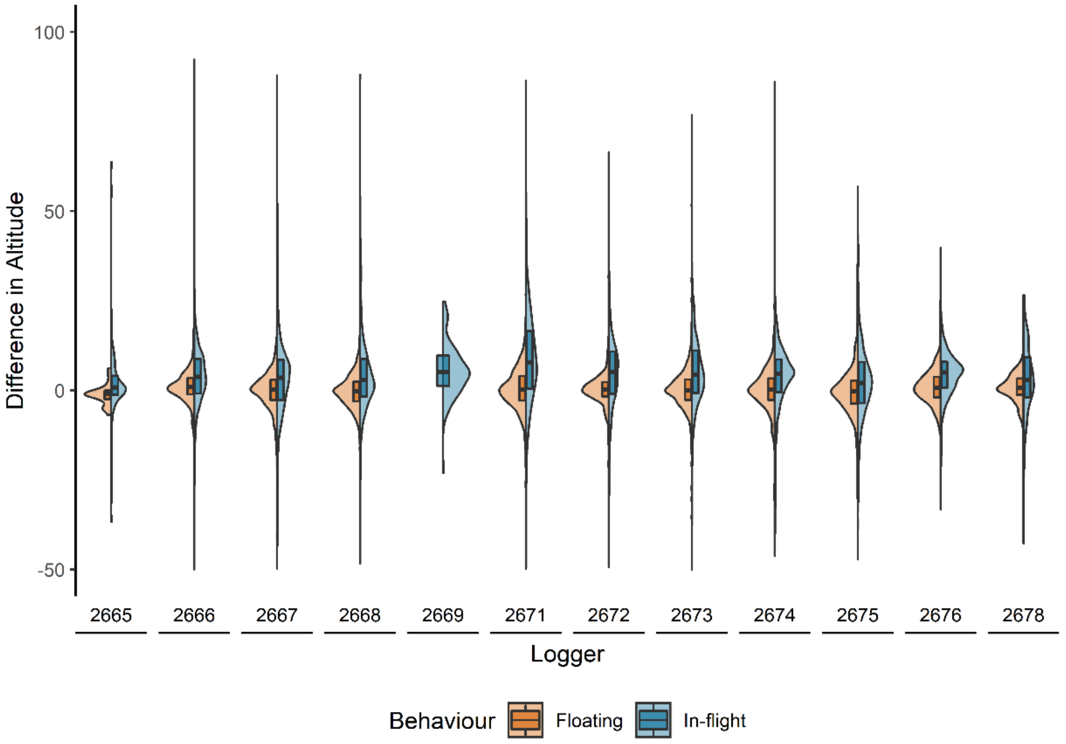


Figure 35: Split violin plots showing the distribution of the difference in point-wise estimates of kittiwake altitude at Fowlsheugh based on GPS derived measures of altitude and those based on barometric pressure using a single logger to calibrate surface pressure. Data is split by individual logger and by bird behaviour (floating versus not floating). Plots include a box denoting the inter-quartile range and median value. For individual logger 2669 there was only a small sample size of floating observations observed (n = 4) hence the lack of a corresponding data distribution.

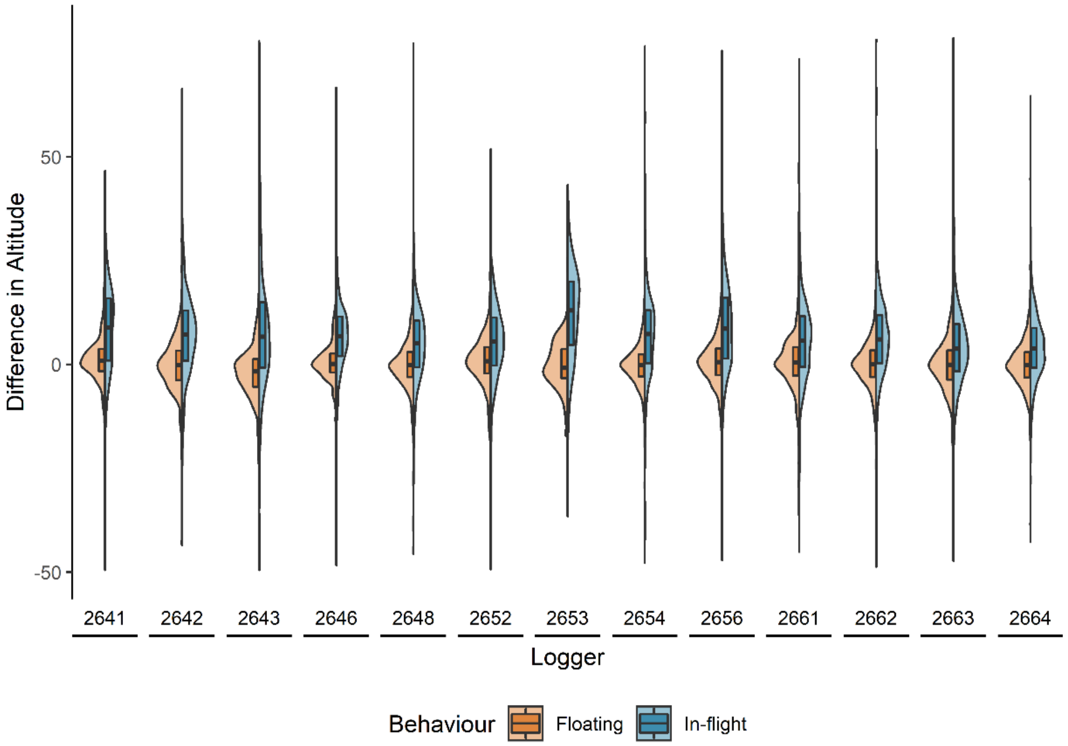


Figure 36: Split violin plots showing the distribution of the difference in point-wise estimates of kittiwake altitude at the Isle of May based on GPS derived measures of altitude and those based on barometric pressure using a single logger to calibrate surface pressure. Data is split by individual logger and by bird behaviour (floating versus not floating). Plots includes a box denoting the inter-quartile range and median value.

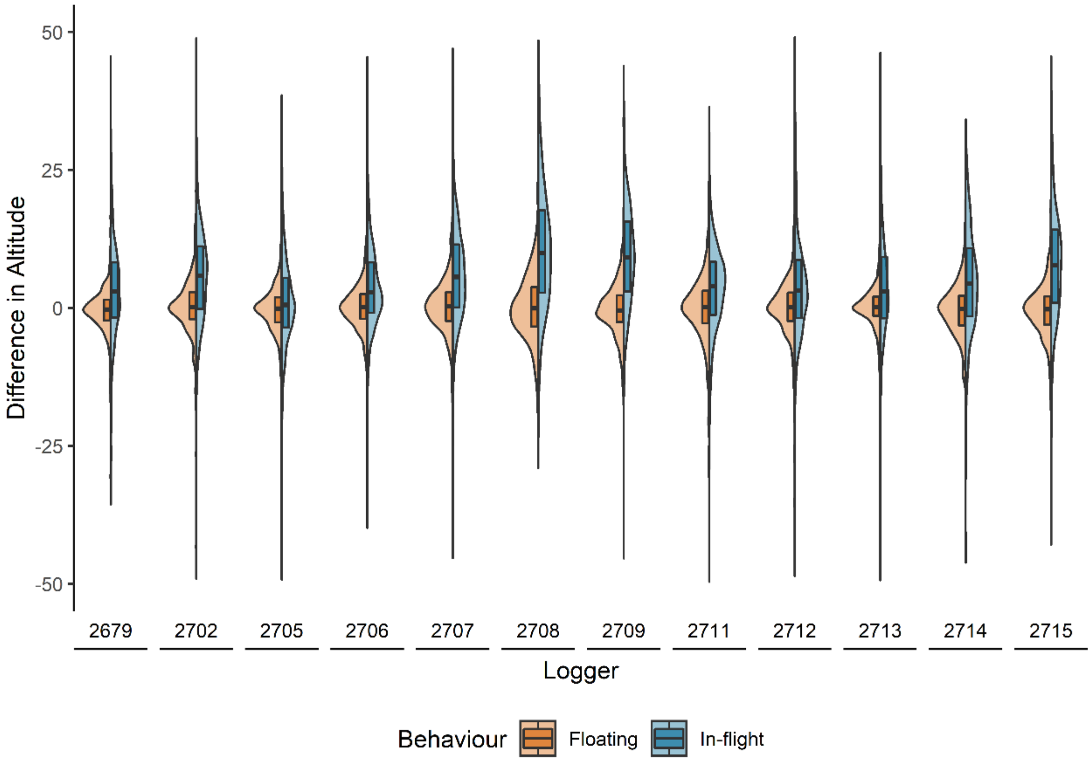


Figure 37: Split violin plots showing the distribution of the difference in point-wise estimates of kittiwake altitude at St Abb's Head based on GPS derived measures of altitude and those based on barometric pressure using a single logger to calibrate surface pressure. Data is split by individual logger and by bird behaviour (floating versus not floating). Plots includes a box denoting the inter-quartile range and median value.

3.5.3 Calibration with weather buoy data

Overall, only a small set of tracking locations were close enough in space and time to be considered suitable for buoy data to be used for sea-level calibration. Values of pressure recorded by the weather buoy were highly correlated with measure of pressure recorded by UvA loggers when birds were classified as floating (i.e. at the same altitude as the buoy). However, despite this correlation UvA loggers also tended to record pressure values that were slightly lower than corresponding estimates from the weather buoy (Fig. 38).

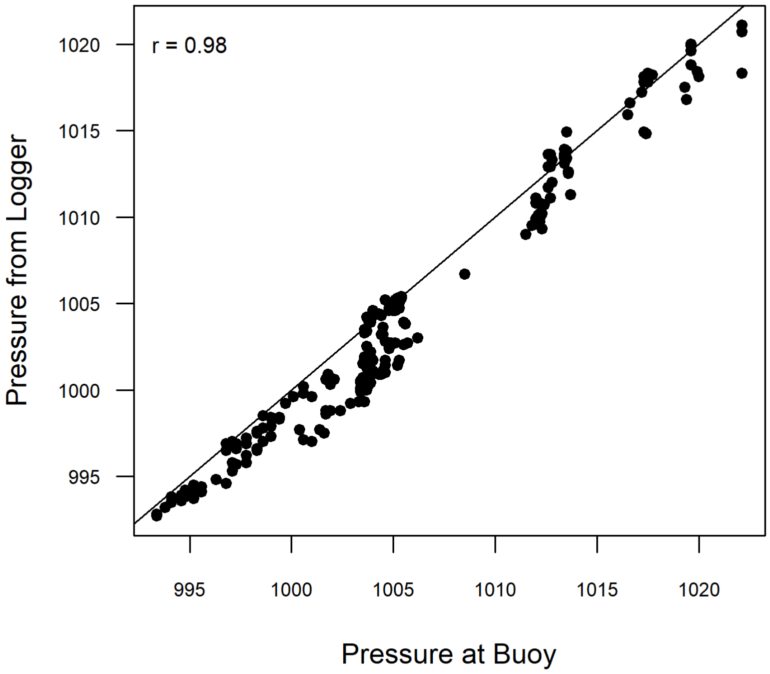


Figure 38: Correlation between pressure recorded by a weather buoy operating in the Firth of Forth and pressure readings from UvA loggers in vicinity of weather buoy when birds classified as floating. Diagonal line on graph shows reference line if recorded pressure was identical across data sources.

Altitude estimates based on using weather buoy data as sea-level reference pressure were weak to moderately correlated with altitude estimates based on barometric data from UvA loggers when birds were classified as floating and slightly more strongly when birds were classified as not floating (Fig. 39). In contrast, estimates of altitude based upon GPS were less strongly correlated with estimates based on pressure data from the weather buoy, and when birds were classified as floating this correlation was negative (Fig. 40). Plots of estimated altitude also demonstrate that altitude estimates based on weather buoy data are typically greater than those based on either GPS- or barometric- derived estimates from UvA loggers. Indeed, while estimates of altitude when birds were classified as floating were centred around 0 when based on GPS or barometric readings from UvA loggers this was not the case when using buoy data (Table 16).

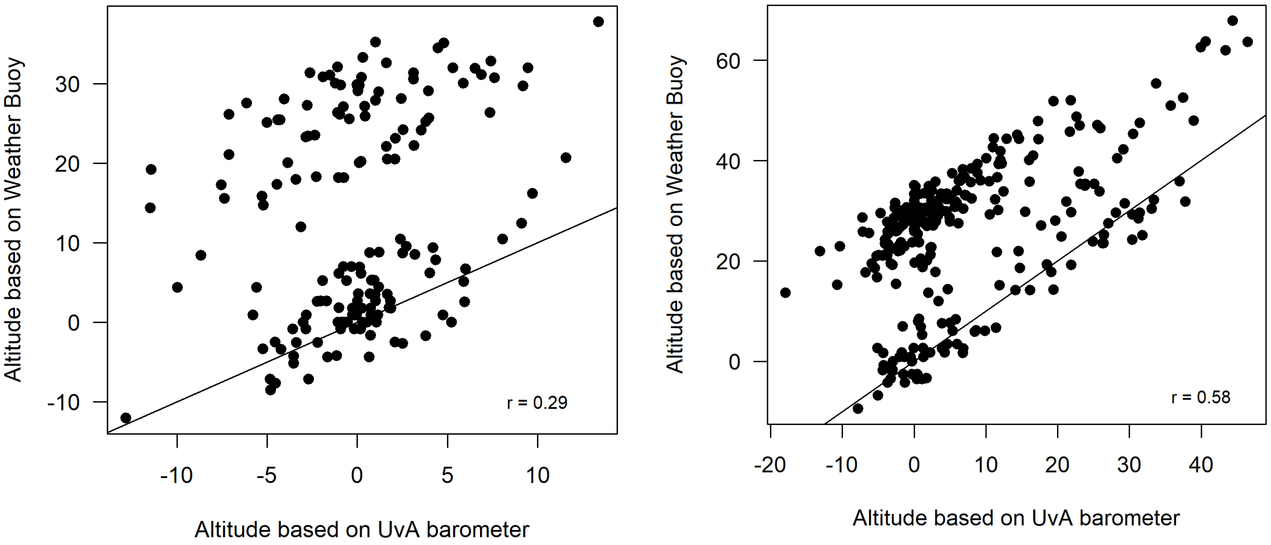


Figure 39: Correlation between altitude estimates based on sea-level pressure readings from a weather buoy operating in the Firth of Forth versus those based on pressure readings from UvA barometric sensors. a) birds classified as floating; b) birds classified as not floating. Diagonal line on graph shows reference line if recorded altitude was identical across approaches.

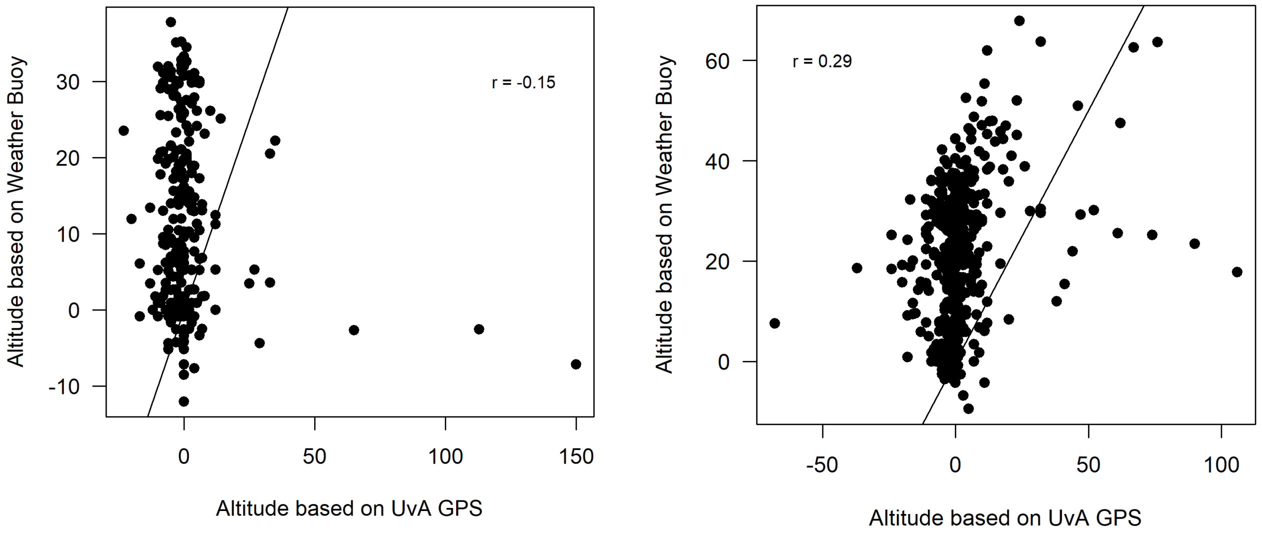


Figure 40: Correlation between altitude estimates based on sea-level pressure readings from a weather buoy operating in the Firth of Forth versus those based on estimates from UvA GPS sensors. a) Birds classified as floating; b) birds classified as not floating. Diagonal line on graph shows reference line if recorded altitude was identical across approaches.

Method	Min.	1st Quantile	Median	3rd Quantile	Max.
UvA GPS	-23.0	-2.98	0.021	3.05	150.01
UvA barometric pressure	-12.82	-2.84	0.016	2.25	13.38
Weather buoy pressure	-12.10	1.18	11.23	20.73	37.87

Table 16. Summary of flight altitude estimates when birds classified as floating using different methods.

3.5.4 State-Space Modelling of Altitude

State-space models of altitude based on both GPS-derived estimates and those based on barometric pressure from a focal UvA logger were highly correlated at Fowlsheugh, and St Abb's Head but less so at the Isle of May (Figs 41, 42 & 43). In addition, correlations between GPS and barometer altitudes were high regardless of whether sampling rate was ≤ 16 seconds or not. The weaker correlations between altitude estimates from the Isle of May may have been driven by a number of estimates of GPS altitude that were much higher than corresponding altimeter estimates. This pattern can be seen on the left-hand side of each graph but was more apparent at the Isle of May than at the other colonies.

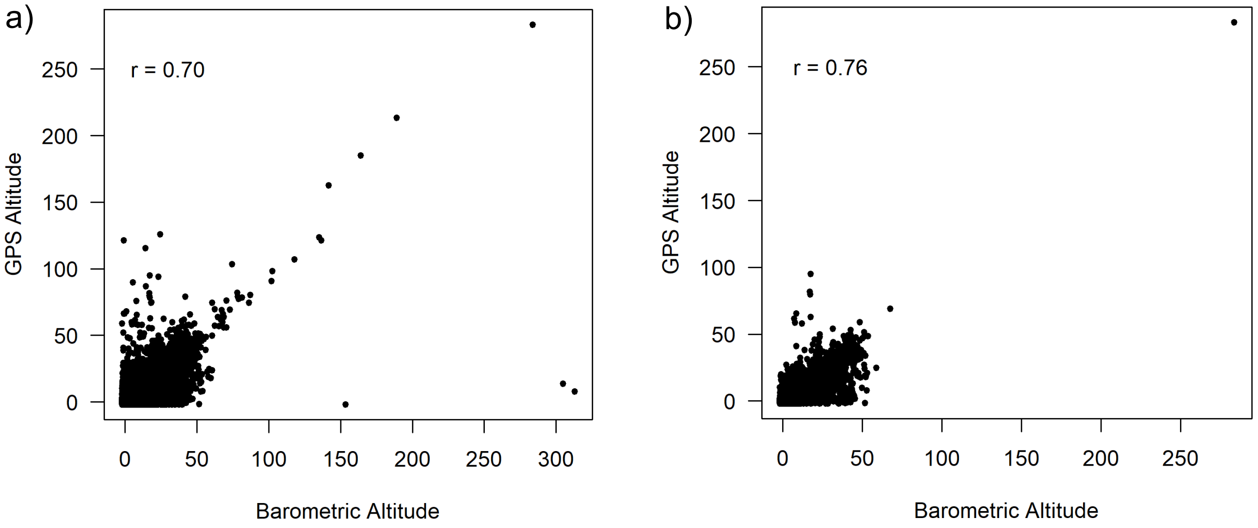


Figure 41: Correlation between altitude estimates based on GPS or barometric pressure from a UvA logger after processing via a SSM approach at Fowlsheugh. Points represent the median altitude for each observation calculated from a posterior distribution using a Bayesian approach. a) All observations; b) Observations when sampling frequency ≤ 16 seconds.

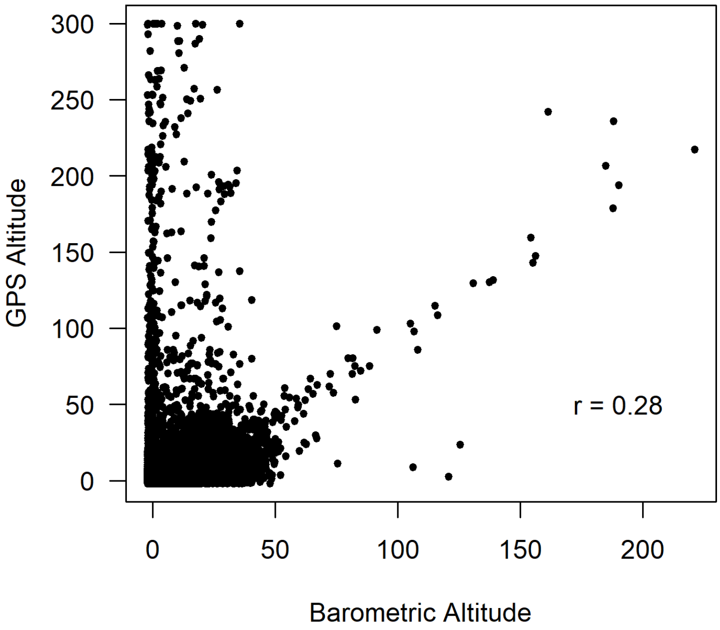


Figure 42: Correlation between altitude estimates based on GPS or barometric pressure from a UvA logger after processing via a SSM approach on the Isle of May. Points represent the median altitude for each observation calculated from a posterior distribution using a Bayesian approach.

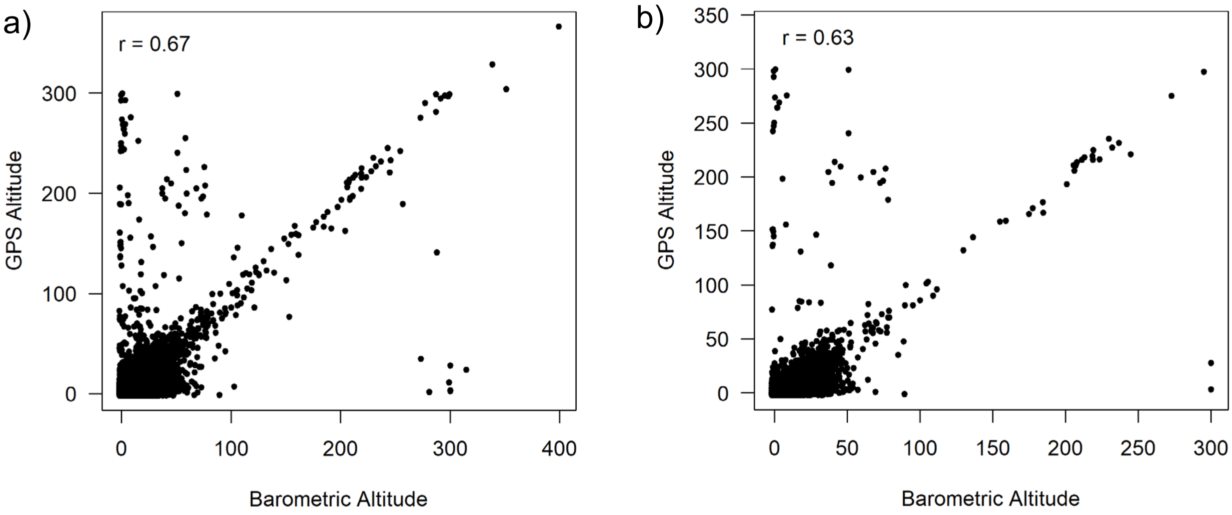


Figure 43: Correlation between altitude estimates based on GPS or barometric pressure from a UvA logger after processing via a SSM approach at St Abb's Head. Points represent the median altitude for each observation calculated from a posterior distribution using a Bayesian approach. a) All observations; b) Observations when sampling frequency ≤ 16 seconds.

Examples of state-space estimates are provided for a randomly selected foraging trip at both Fowlsheugh and St Abb's Head (Figs 44 & 45). Generally, estimates from both sources are well aligned and the inter-quartile ranges for flight altitude over-lapped extensively across colonies (Fig. 46). However, altitude estimates based on barometric pressure consistently exceeded those based on GPS when birds were in-flight. In contrast, when birds were classified as floating, distributions of estimated altitude were broadly similar and centred on 0 for both GPS- and barometer-derived estimates.

Population-level coefficients from state-space modelling were broadly similar regardless of which sensor was used to estimate altitude and which colony was analysed (Tables 17 & 18). For example, when using barometric pressure to estimate altitude, we found that the time lag between measures was negatively associated with altitude at each colony. In contrast, vertical velocity was positively associated with altitude. At Fowlsheugh, there was a two-way interaction between vertical velocity and sampling frequency, such that positive association between vertical velocity and estimated altitude was slightly reduced at higher sampling frequencies (≤ 1 minute). However, at the Isle of May or St Abb's Head, there was little evidence for the same two-way interaction. In the dispersion part of the model, we found that higher sampling frequency was associated with lower dispersion (higher precision) across trips at each colony.

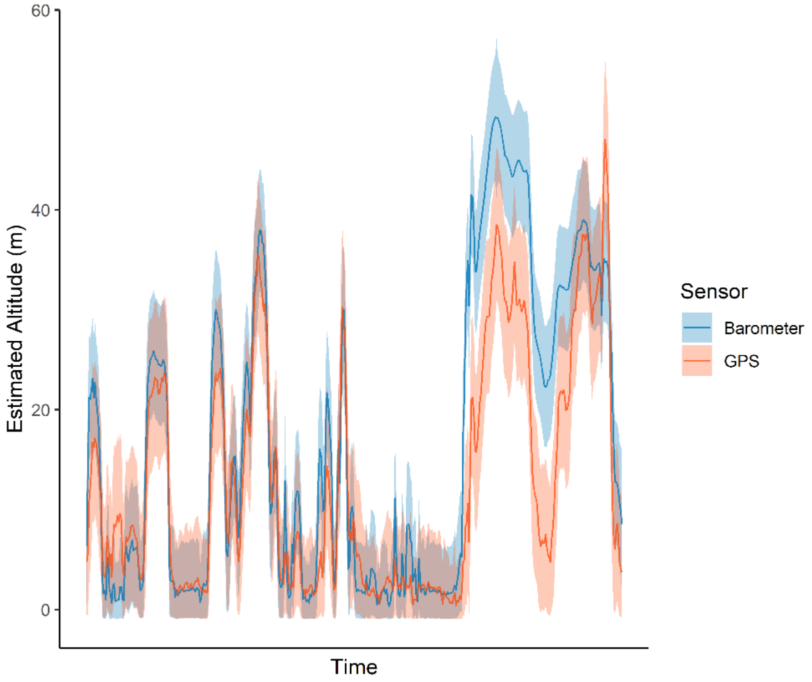


Figure 44: State-space estimation of flight altitude for a single kittiwake foraging trip conducted at Fowlsheugh. Solid lines show median estimate of altitude and opaque ribbon denotes 95% CRI of estimated altitude.

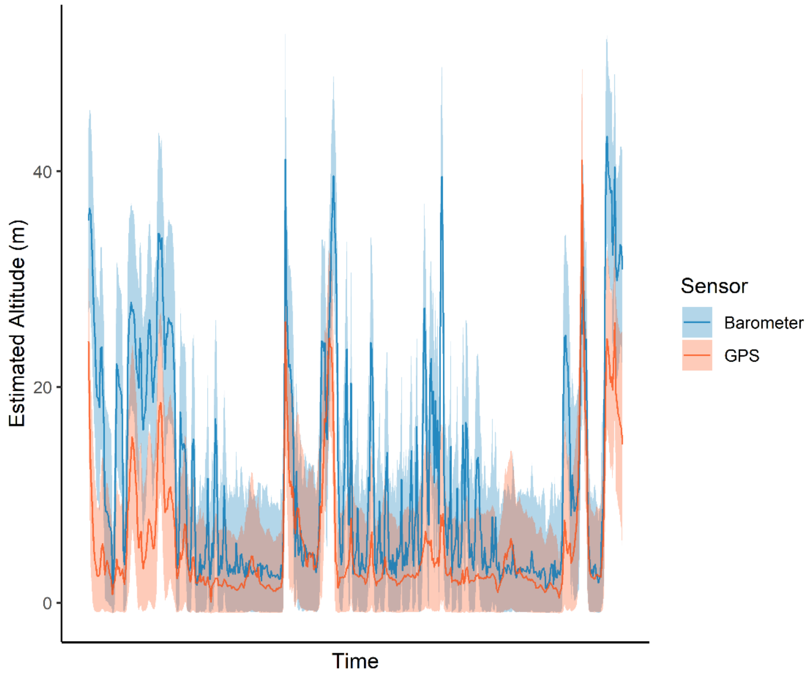


Figure 45: State-space estimation of flight altitude for a single kittiwake foraging trip conducted at St Abb's Head. Solid lines show median estimate of altitude and opaque ribbon denotes 95% CRI of estimated altitude.

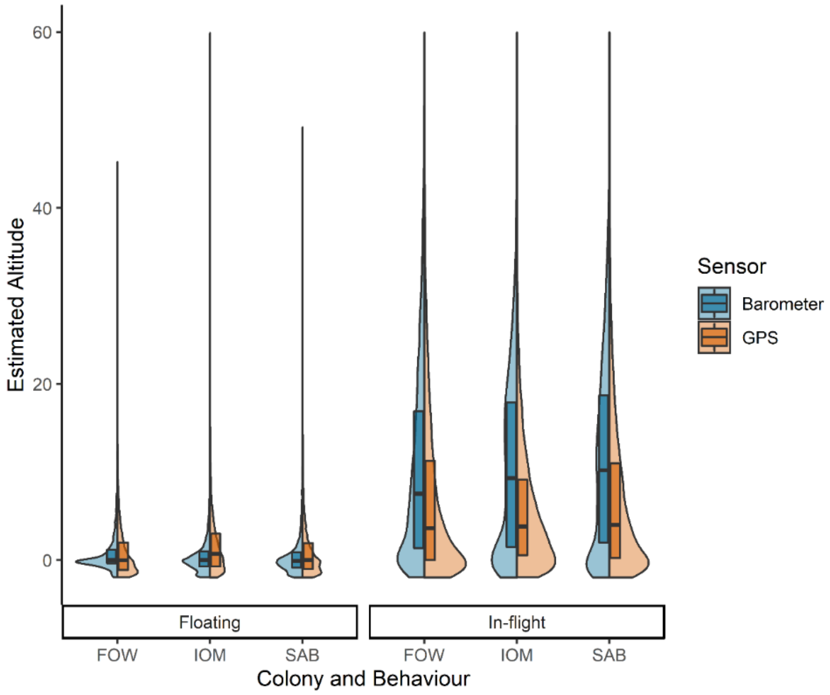


Figure 46: Split violin plot showing distribution of kittiwake altitudes at each colony (FOW = Fowlsheugh, IOM = Isle of May, SAB = St Abb's Head) estimated via SSM using either GPS signals or barometric pressure. Data also split by behaviour into a floating and in-flight category. Plots includes a box denoting the inter-quartile range and median value.

Sensor	**Colony**					
	Fowlsheugh		**Isle of May**		**St Abb’s Head**	
Barometer	$\beta_1 \times z_{t-1}$	1.02 (0.99 – 1.06)	$\beta_1 \times z_{t-1}$	0.99 (0.98 – 0.99)	$\beta_1 \times z_{t-1}$	1.01 (0.99 – 1.03)
	β_2 × Lag	-5.15 (-5.74 - -4.49)	β_2 × Lag	-4.96 (-5.25 - -4.66)	β_2 × Lag	-4.71 (-5.11 - -4.27)
	β_3 × Vertical Velocity	6.21 (6.14 – 6.28)	β_3 × Vertical Velocity	7.83 (7.80 – 7.86)	β_3 × Vertical Velocity	7.76 (7.72 – 7.82)
	β_4 × Sampling Frequency	-0.27 (-0.38 - -0.16)	β_4 × Sampling Frequency	-0.24 (-0.70 – 0.19)	β_4 × Sampling Frequency	-0.19 (-0.31 - -0.062)
	$\beta_5 \times z_{t-1}$ × Lag	0.10 (-0.028 – 0.23)	$\beta_5 \times z_{t-1}$ × Lag	0.25 (0.21 – 0.31)	$\beta_5 \times z_{t-1}$ × Lag	0.15 (-0.021 – 0.26)
	β_6 × Vertical Velocity × Sampling Frequency	-0.27 (-0.49 – 0.036)	β_6 × Vertical Velocity × Sampling Frequency	-0.048 (-0.78 – 0.69)	β_6 × Vertical Velocity × Sampling Frequency	0.023 (-0.19 – 0.25)
	σ_{sp}	1.20 (1.18 – 1.22)	σ_{sp}	0.59 (0.57 – 0.62)	σ_{sp}	0.59 (0.57 – 0.61)
	d_0	1.89 (1.87 – 1.92)	d_0	0.95 (0.92 – 0.98)	d_0	1.11 (1.09 – 1.15)
	d_1 × Sampling Frequency	-1.47 (-1.70 – -1.26)	d_1 × Sampling Frequency	-1.30 (-2.07 - -0.49)	d_1 × Sampling Frequency	-1.04 (-1.37 - -0.74)

Table 17. Population-level coefficients from state-space modelling of kittiwake flight altitude at Fowlsheugh, Isle of May and St Abb’s Head colonies using barometric pressure estimates from UvA loggers. Population level estimates derived from a weighted average of initial trip-level coefficients. 95% Confidence intervals supplied in brackets and based on a boot-sampling procedure.

Sensor	**Colony**					
	Fowlsheugh		**Isle of May**		**St Abbs**	
GPS	$\beta_1 \times z_{t-1}$	1.01 (0.99 – 1.02)	$\beta_1 \times z_{t-1}$	0.96 (0.95 – 0.97)	$\beta_1 \times z_{t-1}$	1.01 (1.01 – 1.02)
	β_2 × Lag	-7.88 (-8.45 - -7.37)	β_2 × Lag	-8.81 (-9.47 - -8.12)	β_2 × Lag	-6.72 (-7.18 - -6.27)
	β_3 × Vertical Velocity	7.48 (7.43 – 7.55)	β_3 × Vertical Velocity	9.72 (9.65 – 9.78)	β_3 × Vertical Velocity	8.36 (8.31 – 8.41)
	β_4 × Sampling Frequency	-0.37 (-0.45 - -0.25)	β_4 × Sampling Frequency	-0.87 (-1.72 – 0.061)	β_4 × Sampling Frequency	-0.28 (-0.37 - - 0.19)
	$\beta_5 \times z_{t-1}$ × Lag	0.25 (0.11 – 0.38)	$\beta_5 \times z_{t-1}$ × Lag	0.16 (0.024 – 0.29)	$\beta_5 \times z_{t-1}$ × Lag	0.11 (-0.02 – 0.24)
	β_6 × Vertical Velocity × Sampling Frequency	0.11 (-0.071 – 0.26)	β_6 × Vertical Velocity × Sampling Frequency	0.45 (-0.56 – 1.48)	β_6 × Vertical Velocity × Sampling Frequency	0.24 (0.071 – 0.38)
	σ_{sp}	1.11 (1.09 – 1.12)	σ_{sp}	1.32 (1.29 – 1.35)	σ_{sp}	1.12 (1.09 – 1.14)
	d_0	-0.062 (-0.065 - -0.059)	d_0	2.85 (2.82 – 2.89)	d_0	-0.059 (-0.13 – 0.027)
	d_1 × 16 Second Sampling Frequency	-0.71 (-0.90 – -0.49)	d_1 × 16 Second Sampling Frequency	NA	d_1 × 16 Second Sampling Frequency	-0.74 (-0.55 - -0.95)
	d_2 × Vertical Accuracy	2.26 (2.18 – 2.34)	d_2 × Vertical Accuracy	0.52 (0.49 – 0.55)	d_2 × Vertical Accuracy	1.38 (1.33 – 1.44)

Table 18. Population-level coefficients from state-space modelling of kittiwake flight altitude at Fowlsheugh, Isle of May and St Abb's Head colonies using GPS-based estimates from UvA loggers. Population level estimates derived from a weighted average of initial trip-level coefficients. 95% Confidence intervals supplied in brackets and based on a boot-sampling procedure. Note that at the Isle of May no instances of ≤ 16 second sampling frequency were observed, therefore this term was not included in models for this colony.

When using GPS-based estimates of altitude we found that the time lag between measures was negatively associated with altitude at each colony, mirroring the result we saw with barometric pressure. However, there was some evidence that the dependence of altitude estimates at time t_j on preceding values at time t_j-*1* increased slightly with longer time lags at the Isle of May, with less evidence that the same pattern occurred at Fowlsheugh and St Abb’s Head. Vertical velocity was positively associated with altitude at Fowlsheugh and St Abb’s Head, but not the Isle of May. At St Abb’s Head a two-way interaction between vertical velocity and sampling frequency occurred, such that positive association between vertical velocity and estimated altitude was slightly increased at higher sampling frequencies (≤ 1 minute). The same relationship, albeit weaker, may also have occurred at Fowlsheugh. The absence of a two-way interaction between vertical velocity and sampling frequency on the Isle of May may have been due to a lower rate of sampling at ≤ 1 minute and a subsequent lack of sample size. In the dispersion part of the model we found that higher sampling frequency (≤ 16 seconds) was associated with lower dispersion (higher precision) across trips at both Fowlsheugh and St Abb’s Head. No instances of (≤ 16 second were observed at the Isle of May, hence this term was not estimated for this colony. Vertical accuracy was positively associated with dispersion (lower precision) across all colonies. Note that smaller values of vertical accuracy denote greater accuracy so positive associations with dispersion were expected.

3.6 Device effects in puffins

3.6.1 Chick feeding rates

We were not able to test directly for effects of treatment on individual feeding rates because we could not distinguish mates in unmarked control pairs. The proportion of feeds per observation day delivered by puffins tagged with GPS loggers was similar to that in puffins that received colour rings only (Fig. 47). Pair members normally share chick provisioning duties equally (Harris & Wanless 2011), therefore if the tagged birds were unaffected they would be expected to provide on average 50% of the daily feeds. Colour-ringed birds provided a median of 25% and GPS tagged birds 13% of the daily feeds, suggesting there were some negative effects of device deployment at the individual level.

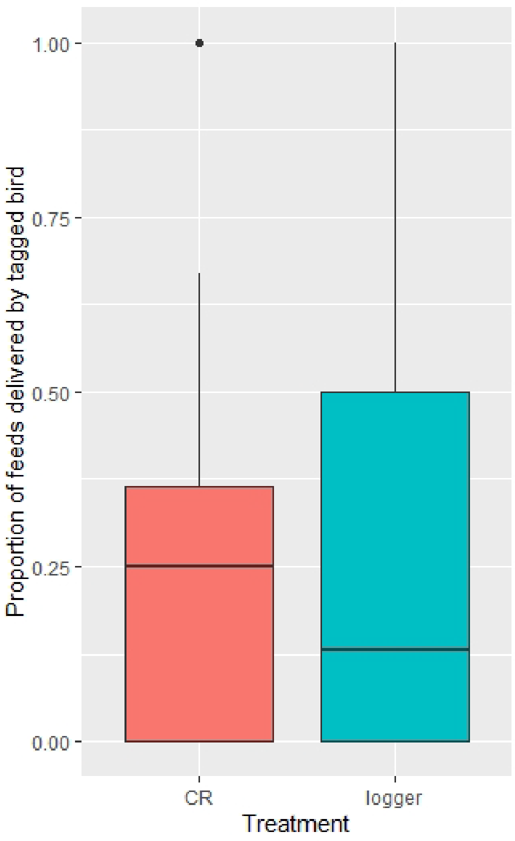


Figure 47: Proportion of feeds delivered to puffin chicks per observation day in relation to treatment (colour-ringed (CR): n=12; logger: n=24 pairs).

There was no evidence for a significant difference in provisioning rates between the study sites. The proportion of daily feeds delivered by individuals in the two treatment groups did not change within the first four days since the tag was deployed/colour ring fitted.

Since none of the explanatory variables we considered had a substantial effect on chick provisioning rates at the individual level, the most parsimonious model was the null (intercept only) model. The random effects of 'nest identity' and 'day' explained relatively small amount (19%) of variation in the proportion of daily feeds.

Similar to the result at individual level, chick provisioning rate at the pair level was unaffected by treatment, site or time since logger deployment/ring attachment (Fig. 48). Again, the most parsimonious model was the null model. The random effects of 'nest identity' and 'day' explained moderate amount (38%) of variation in the proportion of daily feeds.

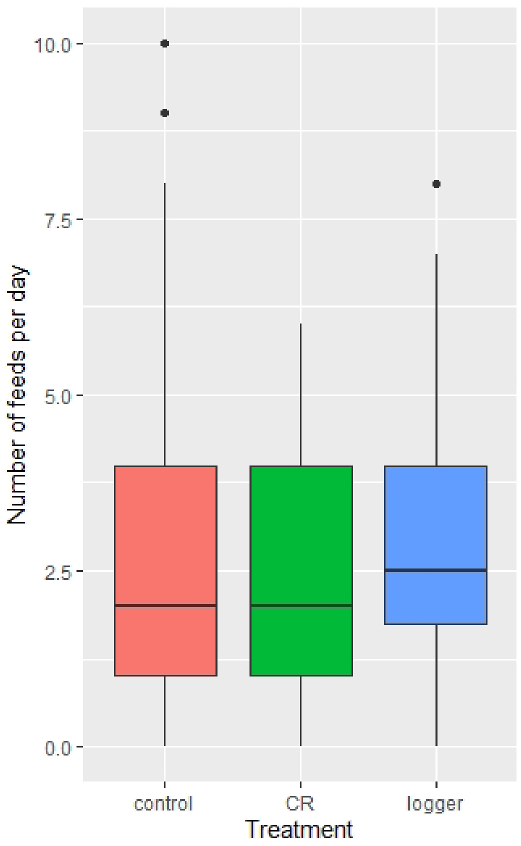


Figure 48: Number of feeds per observation day delivered by puffin pairs in relation to treatment (control: n=64, colour-ringed (CR): n=12, logger: n=24 pairs).

3.6.2 Chick fledging condition and survival

Chick fledging condition was higher in logger (but not in colour-ring) nests compared to control nests (Fig. 49, Table 19). It is possible that, since chick provisioning rates in treatment nests were not reduced compared to controls, the supplementary feeding we undertook has resulted in treatment chicks fledging in better condition. There was no evidence for a significant difference in chick condition between the study sites (Table 19).

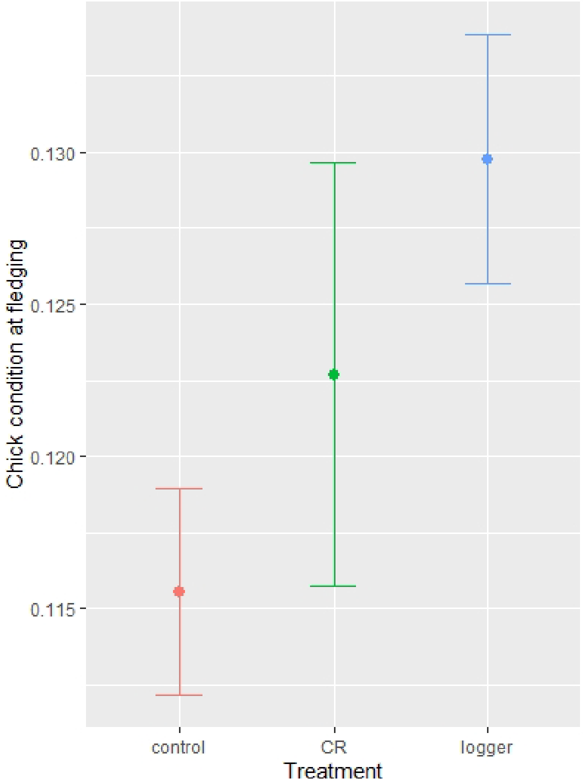


Figure 49: Puffin condition at fledging in relation to treatment (control: n=19, colour-ringed (CR): n=12, logger: n=24 pairs).

Model	AICc	N param.	Parameter estimate (± SE)			R^{2}m
			Treatment	Site	Treatment × site	
M3	**-246.2**	**4**	**0.007 ± 0.007** **0.014 ± 0.006**	**-**	**-**	**0.10**

* Parameter estimates for each treatment level are relative to the control group (first row: colour-ringed, second row: logger)

Table 19. Generalised linear mixed models testing for effects of treatment and site on chick fledging condition.

There was no difference in chick survival between treatment and control nests or between sites. Puffin chicks in all three study groups had a high survival rate (control: 100%, n=19; CR: 100%, n=12; logger: 92%, n=24).

3.7 Device effects in kittiwakes

3.7.1 At-colony behaviour

3.7.1.1 Foraging trip duration

Fowlsheugh

At this colony, the number of changeovers recorded was low overall, suggesting kittiwakes made long-lasting foraging trips. Fewer changeovers were observed at nests where one adult was tagged with a UvA logger compared to controls and nests where one bird was tagged with a Pathtrack logger (Fig. 50, Table 20). There was no significant difference in changeover rates between study sites.

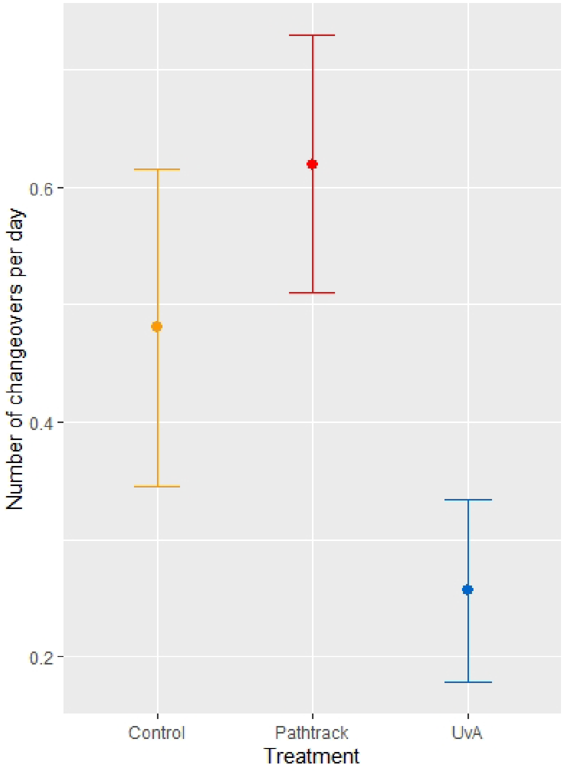


Figure 50: Number of parental changeovers per observation day in kittiwakes at Fowlsheugh in relation to treatment. Sample sizes are presented in Table 5 (one UvA nest was removed from this analysis; see Methods, section 2.2.2.2).

The best model included treatment only. The null model, however, was within 2AICc units of the best model and treatment explained 9% of the variation in parental changeover rates, suggesting a relatively weak effect (Table 20, R^2m).

Model	AICc	N parameters	Parameter estimate (± SE)			R^2m
			Treatment	Site	Treatment × site	
M3	**141.0**	**3**	**0.20 ± 0.28** **-0.59 ± 0.37**	**-**	**-**	**0.09**
M5	141.7	1	-	-	-	0

* Parameter estimates for each treatment level are relative to the control group (first row: Pathtrack, second row: UvA)

Table 20. Generalised linear models testing for effects of treatment and site on parental changeover rates in kittiwakes at Fowlsheugh. The best model (in bold) and models within 2 AICc units of the best model are presented. The conditional coefficient of determination (R^2c) was not calculated as no random effects were included in the models (see methods for details).

Isle of May

There were fewer parent changeovers per observation day (indicative of longer foraging trips) at nests where one adult was tagged with a UvA logger compared to controls and nests where one bird was tagged with a Pathtrack logger (Fig. 51a, Table 21). Also, fewer changeovers were observed at nests at site 2 (Clett) than at site 1 (Colms; Fig. 51b, Table 21).

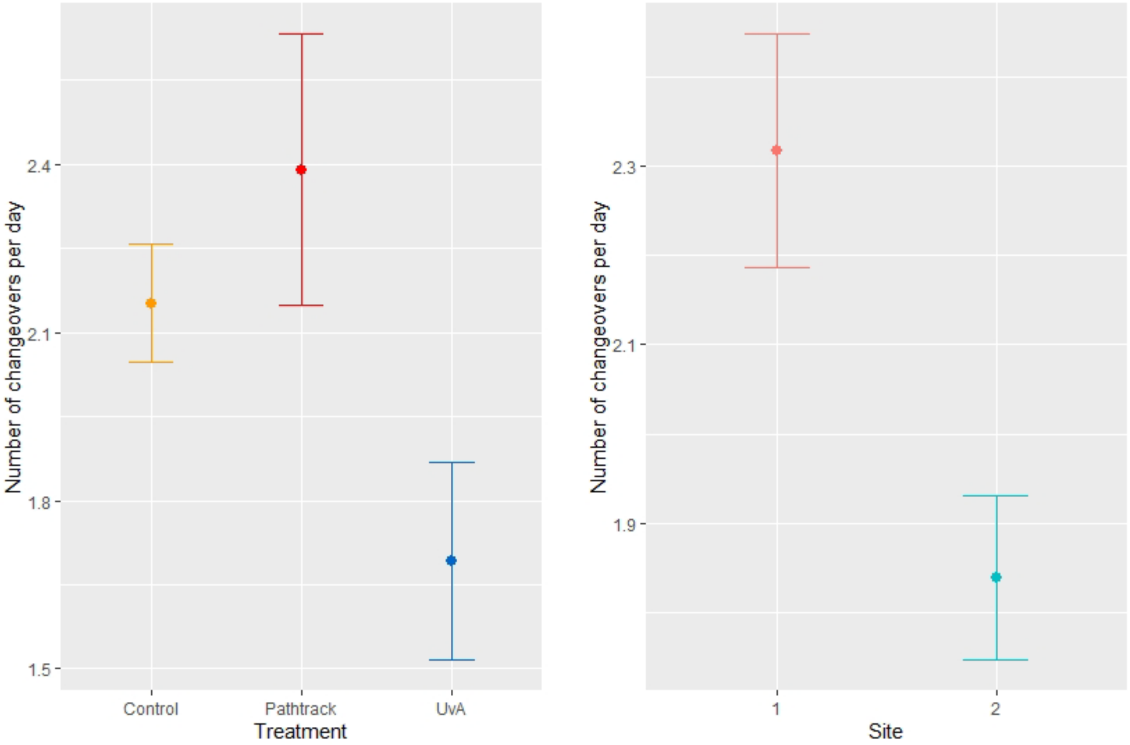


Figure 51: Number of parental changeovers per observation day in kittiwakes on the Isle of May in relation to a) treatment and b) site. Sample sizes per treatment and site are presented in Table 5.

The best model included treatment and site. The model with site only was within 2AICc units of the best model. The main effects explained 13% of the variation in parental changeover rates (Table 21, R^2m).

Model	AICc	N parameters	Parameter estimate (± SE)			R^2m
			Treatment	Site	Treatment × site	
M2	**312.5**	**4**	**0.11 ± 0.12** **-0.22 ± 0.14**	**-0.22 ± 0.10**	**-**	**0.13**
M4	312.7	2	-	-0.24 ± 0.10	-	0.08

* Parameter estimates for each treatment level are relative to the control group (first row: Pathtrack, second row: UvA)

Table 21. Generalised linear models testing for effects of treatment and site on parental changeover rates in kittiwakes on the Isle of May. The best model (in bold) and models within 2 AICc units of the best model are presented. The conditional coefficient of determination (R^2c) was not calculated as no random effects were included in the models (see methods for details).

St Abb's Head

Fewer changeovers were observed at nests where one adult was tagged with a UvA logger than at control nests (Fig. 52, Table 22). At this colony, nests where one bird was tagged with a Pathtrack logger had on average slightly reduced changeover rate compared to controls although the difference was not statistically significant (Fig. 52, Table 22). There was no significant difference in changeover rates between study sites (Table 22).

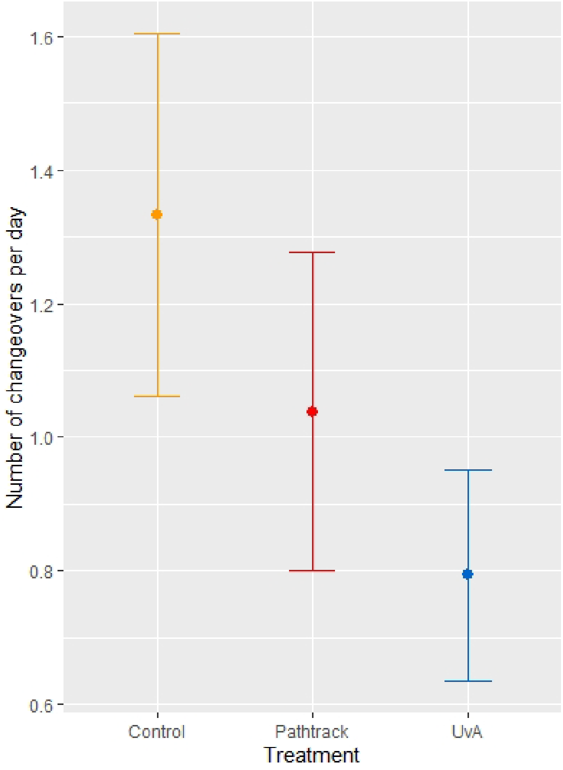


Figure 52: Number of parental changeovers per observation day in kittiwakes at St Abb's Head in relation to treatment. Sample sizes are presented in Table 5 (one Pathtrack nest was removed from this analysis; see Methods, section 2.2.2.2).

The best model included treatment only. The model containing treatment and site, and the null model were within 2AICc units of the best model; however, the model with treatment and site was not considered further as it was otherwise identical to the best model but contained an extra parameter (see Methods, section 2.2.2.3). Treatment explained 12% of the variation in parental changeover rates (Table 22, R^2m).

Model	AICc	N parameters	Parameter estimate (± SE)			R^2m
			Treatment	Site	Treatment × site	
M3	**162.7**	**3**	**-0.28 ± 0.23** **-0.52 ± 0.22**	**-**	**-**	**0.12**
M5	163.8	1	-	-	-	0

* Parameter estimates for each treatment level are relative to the control group (first row: Pathtrack, second row: UvA)

Table 22. Generalised linear models testing for effects of treatment and site on parental changeover rates in kittiwakes at St Abb's Head. The best model (in bold) and models within 2 AICc units of the best model are presented. The conditional coefficient of determination (R^2c) was not calculated as no random effects were included in the models (see methods for details).

3.7.1.2 Chick attendance

Chick attendance was generally high at both Isle of May and Fowlsheugh (Fig. 53b; no data available for St Abb's Head). However, chicks at UvA nests were left unattended more often compared to those at control nests, whereas chicks at Pathtrack nests were left unattended less often than controls (Fig. 53a, Table 23). Lower chick attendance was observed on the Isle of May compared to Fowlsheugh (Fig. 53b, Table 23) and for larger broods compared to single chicks (Fig. 53c, Table 23).

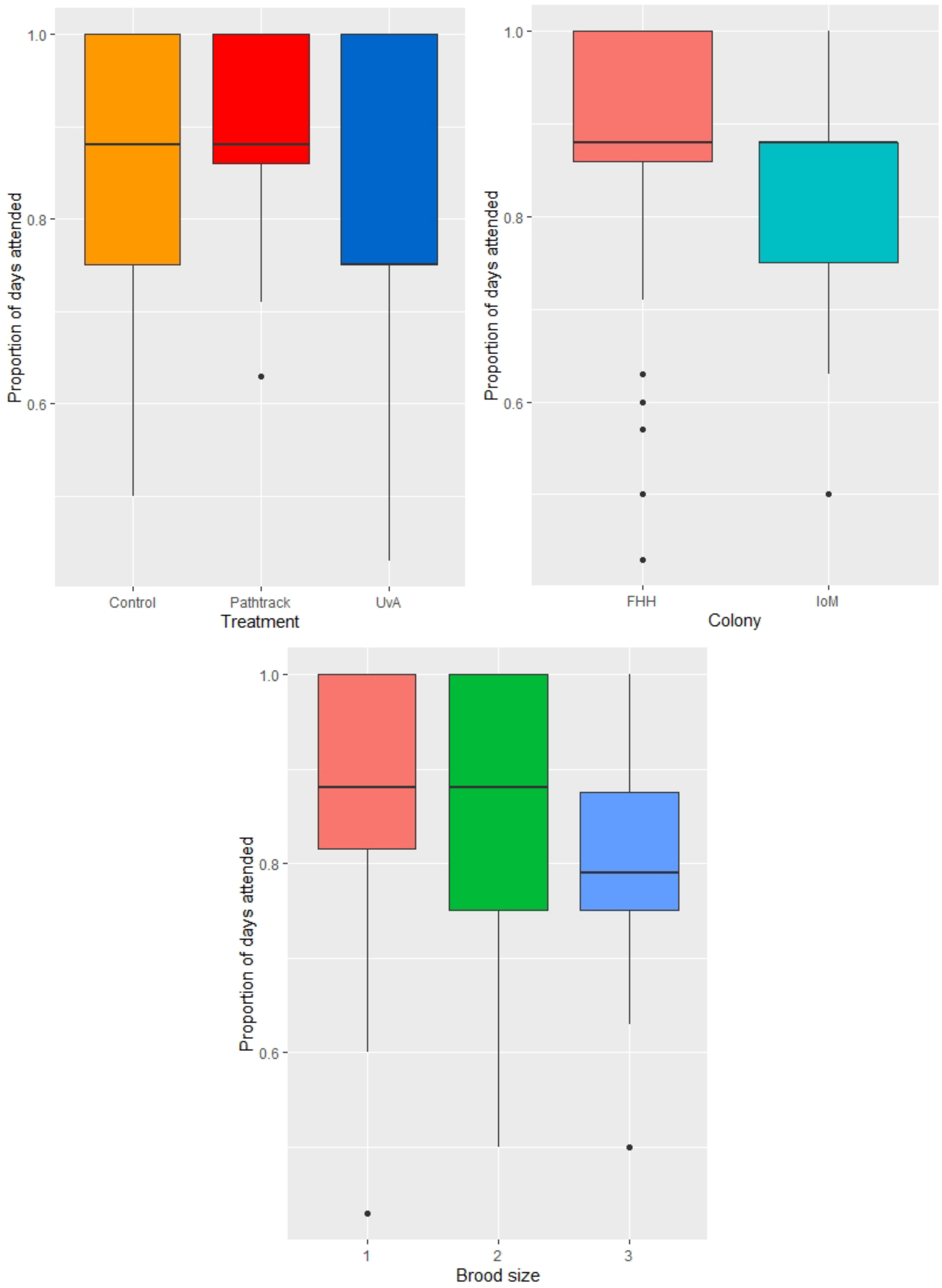


Figure 53: Chick attendance of kittiwakes in relation to a) treatment, b) colony and c) brood size. Sample sizes are presented in Table 6 (three UvA and four control nests at Fowlsheugh were removed from this analysis; see Methods, section 2.2.2.2).

The best model included treatment, colony and brood size. The model with treatment and colony was within 2 AICc units of the best model. The fixed effects explained 24% of the variation in chick attendance (Table 23, R^2m).

Model	AICc	N param.	Parameter estimate (± SE)					R^2m
			Treatment	Brood size	Colony	Treatment × colony	Colony x 'newsite'	
M8	**654.8**	**6**	**0.41 ± 0.22** **-0.35 ± 0.21**	**-0.32 ± 0.21** **-0.80 ± 0.36**	**-0.62 ± 0.14**	**-**	**-**	**0.24**
M10	655.8	4	0.48 ± 0.22 -0.34 ± 0.21	-	-0.62 ± 0.14	-	-	0.21

* Parameter estimates for each treatment level are relative to the control group (first row: Pathtrack, second row: UvA); parameter estimates for brood size are relative to single-chick brood (first row: 2-chick brood, second row: 3-chick brood)

Table 23. Generalised linear models testing for effects of treatment, colony, site and brood size on chick attendance in kittiwakes. The best model (in bold) is presented; no models were within 2 AICc units of that. The conditional coefficient of determination (R2c) was not calculated as no random effects were included in the models (see methods for details).

3.7.1.3 Chick and adult condition

Chick condition

There was no evidence of device effect (treatment by measurement day interaction) on chick condition, or difference in chick condition between study sites or brood sizes. Chicks from all treatment groups (control, Pathtrack and UvA) were in poorer condition at the second measurement compared to the initial one (Fig. 54, Table 24); however, this may reflect rate of skeletal growth relative to body mass gain.

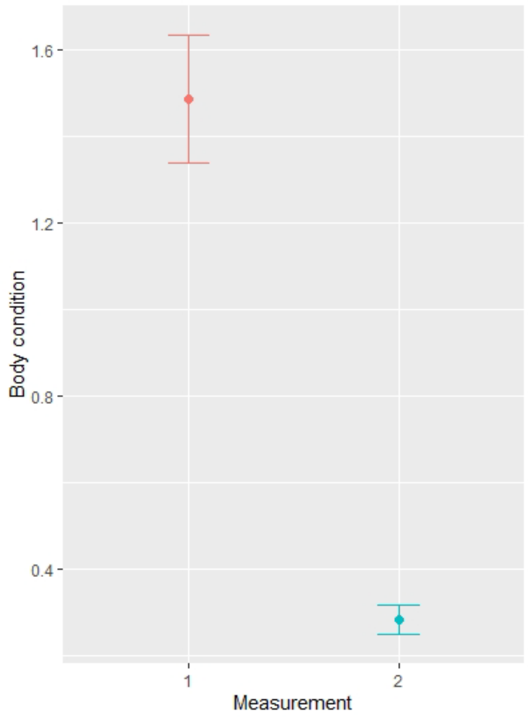


Figure 54: Kittiwake chick body condition in relation to measurement day. Sample sizes are presented in Table 7.

The best model included measurement day only. The models containing measurement day and site and measurement day and brood size or both were within 2AICc units of the best model; however, they were not considered further as they were otherwise identical to the best model but contained extra parameters (see methods for details). The fixed effects explained 44% and the fixed and random effects combined explained 79% of the variation in chick condition (Table 24).

Model	AICc	N param.	Parameter estimate (± SE)						R^2m	R^2c
			Treatment	Meas. day	Brood size	Site	Treatment × meas.day	Treatment × site		
M14	**78.0**	**5**	**-**	**-0.63 ± 0.03**	**-**	**-**	**-**	**-**	**0.44**	**0.79**

Table 24. Generalised linear mixed models testing for effects of treatment, measurement day, brood size and site on kittiwake chick condition. The best model (in bold) is presented.

Adult condition

Adults from both treatment groups were in poorer condition at the second measurement compared to the first (Fig. 55, Table 25); however, biologically the difference was small. Furthermore, since data were not available from control birds this result is difficult to interpret because body condition can change through the breeding season due to factors independent of GPS tagging. However, in a sample of 53 unmanipulated adults measured and weighed once during the season as part of the Isle of May long-term study, there was no significant difference in body condition between birds caught at approximately the same times as the first and second measurements of our study birds, suggesting that there may be a weak device effect on adult body condition.

There was also an interactive effect of treatment and site: at site 1 (Colms), kittiwakes tagged with Pathtrack loggers were in poorer condition than those tagged with UvA loggers, whereas at site 2 (Clett) the opposite pattern was observed. Note, however, that sample sizes of birds per treatment and site are small (see Table 7) so results should be treated with caution.

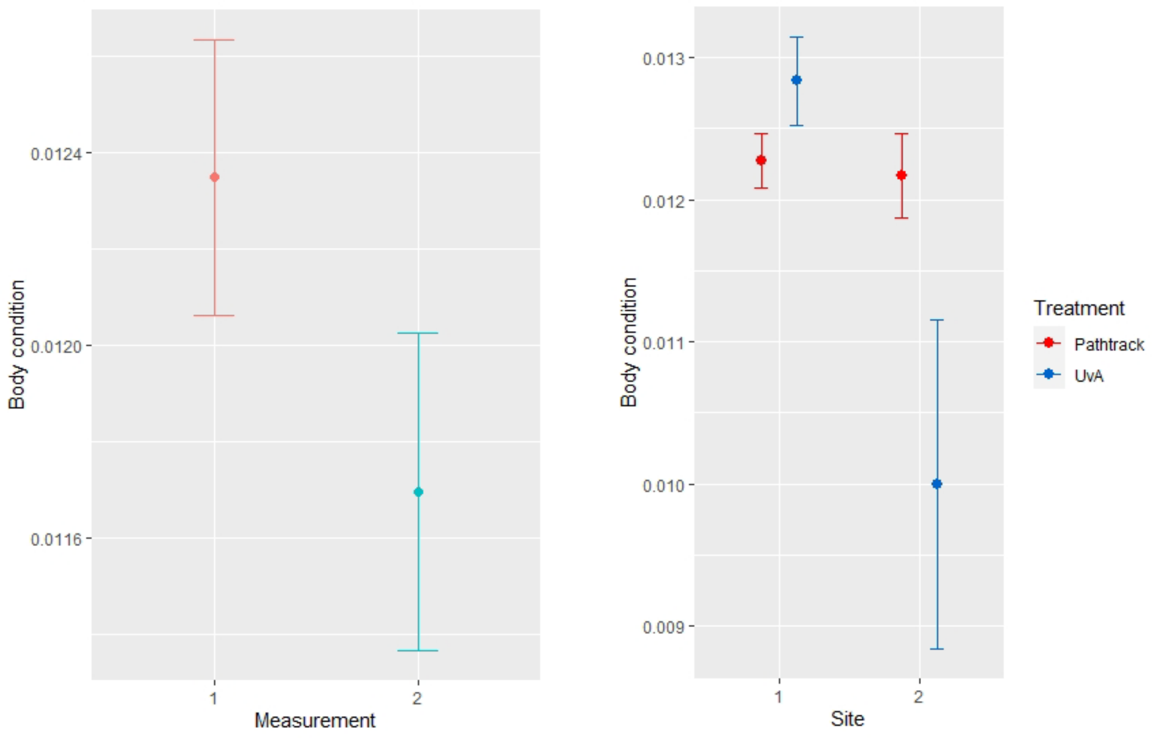


Figure 55: Kittiwake adult body condition in relation to measurement day, site and treatment. Sample sizes are presented in Table 7.

The best model contained treatment, measurement day, site and treatment by site interaction (Table 25). The full model was within 2AICc units of the best model but contained an extra parameter and was therefore disregarded on the grounds of parsimony (see methods for details).

Model	AICc	N param.	Parameter estimate (± SE)					R^2m	R^2c
			Treatment	Meas. day	Site	Treatment ×meas.day	Treatment × site		
M3	**487.5**	**7**	**0.001 ± 0.001**	**-0.001 ± 0.0002**	**-0.0001 ± 0.001**	**-**	**-0.003 ± 0.001**	**0.31**	**0.82**

Table 25. Generalised linear mixed models testing for effects of treatment, measurement day and site on adult condition in kittiwakes. The best model (in bold) is presented.

3.7.1.4 Breeding success

For both measures we explored (number of fledged chicks per nest and proportion of fledged chicks from the maximum brood size recorded during the chick-rearing period) the most parsimonious model was the null (intercept only) model, suggesting none of the explanatory variables we considered had a substantial effect on breeding success. However, data visualisation suggested that there may be a difference in breeding success between the treatment groups at Fowlsheugh that was not apparent from the analysis involving all three colonies (Fig. 56). We therefore conducted an additional analysis, where we tested for effects of treatment, site and treatment by site interaction at this colony. The analysis indicated that there was an effect of treatment (the best model within the candidate set included treatment only).

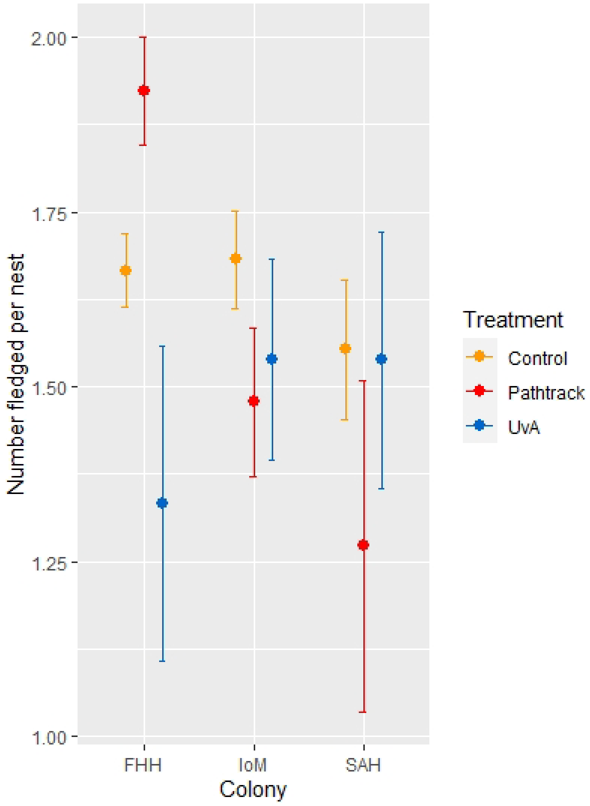


Figure 56: Breeding success of treatment (Pathtrack, UvA) and control kittiwake nests at the three study colonies.

3.7.2 At-sea behaviour

3.7.2.1 Trip metrics

Model selection indicated that the models without the Julian day term were the best fit for all three trip metrics based on a difference in AICc of ≥2 compared to the next best model. Patterns across all metrics were consistent with significant interactions between the treatment (logger type) and colony terms indicating that device effects were colony-specific (Table 26, Fig. 57-59).

	Trip duration			Trip distance			Foraging range		
	df	F	p	df	F	p	df	F	p
Treatment	1	5.619	0.021	1	1.241	0.269	1	1.250	0.267
Colony	2	22.792	<0.001	2	15.551	<0.001	2	15.439	<0.001
Treatment x Colony	2	8.629	<0.001	2	5.454	0.006	2	5.858	0.004
R^2m		0.124			0.129			0.134	
R^2c		0.210			0.307			0.322	

Table 26. Summaries for final trip metric models. Note that all three trip metrics are log-transformed.

At Fowlsheugh, trip metrics varied significantly between logger types, with UvA birds going on longer trips (0.646 ± 0.147, df=82.133, t=4.410, p<0.001; Fig. 57), travelling greater distances (0.789 ± 0.254, df=82.108, t=3.107, p=0.003; Fig. 58) and going to farther offshore areas (0.810 ± 0.249, df=80.946, t=3.250, p=0.002; Fig. 59). There were no differences between the trip metrics of birds tagged with Pathtrack and UvA loggers at either of the other colonies (all t < |1.5|, p>0.1).

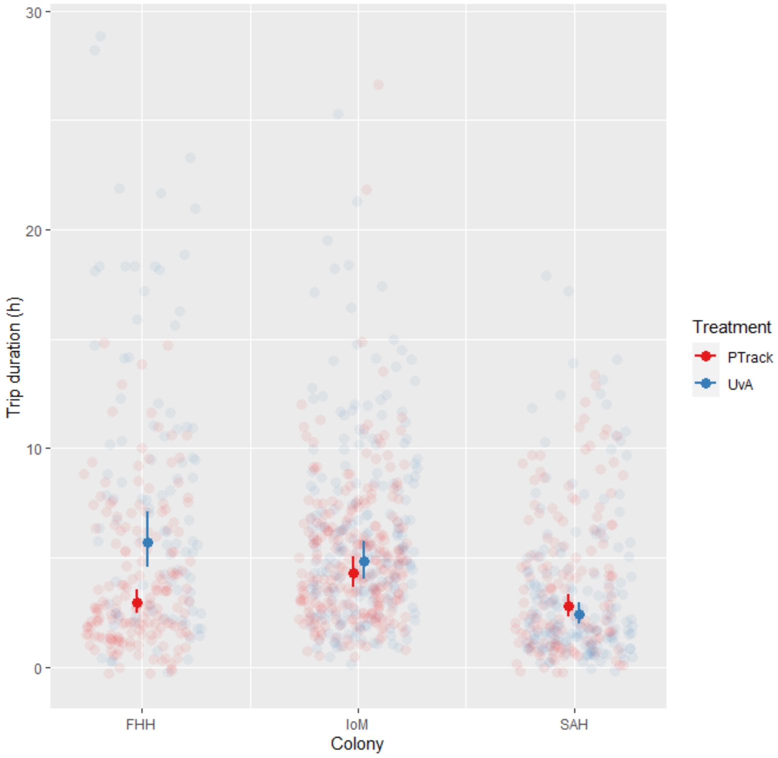


Figure 57. Predicted duration of foraging trips for kittiwakes tagged with Pathtrack and UvA loggers at each colony. The raw data are plotted in the background. FHH: Fowlsheugh, IoM: Isle of May, SAH: St Abb's Head.

Regarding inter-colony differences, Pathtrack tagged birds from the Isle of May made longer trips (Fowlsheugh: -0.404 ± 0.125, df=60.038, t=-3.246, p=0.002; St Abb's: -0.472 ± 0.128, df=62.648, t=-3.671, p<0.001; Fig. 57), travelling larger distances (Fowlsheugh: -0.663 ± 0.224, df=64.674, t=-2.960, p=0.004; St Abb's: -0.713 ± 0.230, df=66.068, t=-3.103, p=0.003; Fig. 58) and going further (Fowlsheugh: -0.656 ± 0.221, df=64.132, t=-2.971, p=0.004; St Abb's: -0.717 ± 0.226, df=65.440, t=-3.168, p=0.002; Fig. 59) than Pathtrack tagged birds at either of the other colonies, but there was no difference between birds from Fowlsheugh and St Abb's Head (all t < |0.5|, p>0.6).

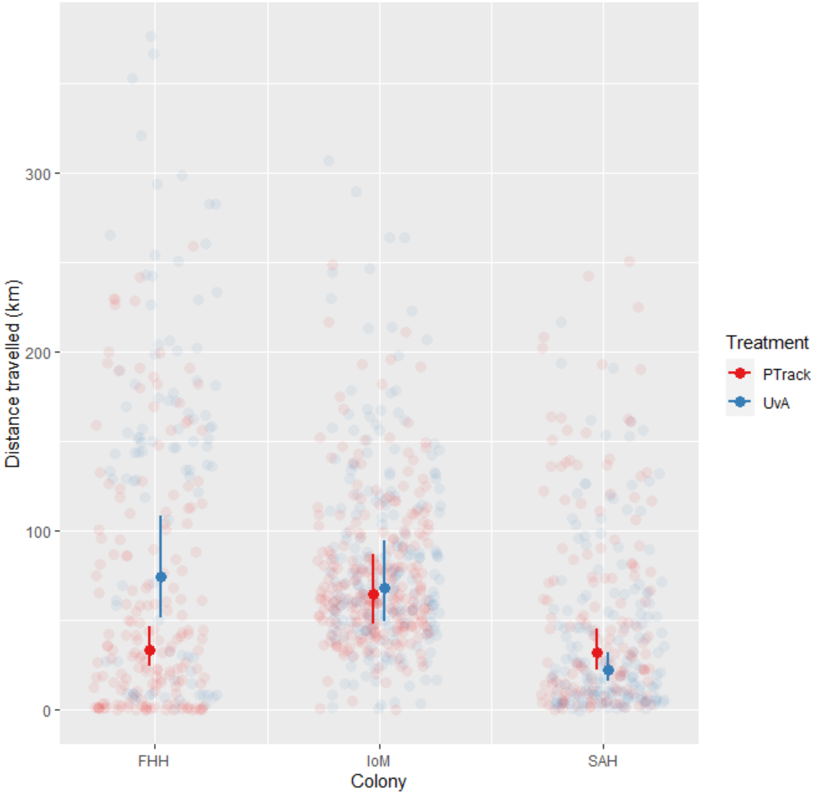


Figure 58. Predicted distance travelled during foraging trips of kittiwakes tagged with Pathtrack and UvA loggers at each colony. The raw data are plotted in the background. FHH: Fowlsheugh, IoM: Isle of May, SAH: St Abb's Head.

Among UvA tagged birds, those from St Abb's Head made shorter trips (Fowlsheugh: 0.893 ± 0.149, df= 76.134, t=5.987, p<0.001; Isle of May: 0.744 ± 0.133, df=55.395, t=5.577, p<0.001; Fig. 57), travelled shorter distances (Fowlsheugh: 1.207 ± 0.260, df=78.041, t=4.639, p<0.001; Isle of May: 1.129 ± 0.242, df=61.560, t=4.667, p<0.001; Fig. 58) and had smaller foraging ranges (Fowlsheugh: 1.228 ± 0.256, df=77.098, t=4.799, p<0.001; Isle of May: 1.081 ± 0.239, df=61.198, t=4.530, p<0.001; Fig. 59) compared to birds from the Isle of May and Fowlsheugh. There was no difference between UvA tagged birds from either of the other colonies for any of the trip metrics (all t < |1.5|, p> 0.3).

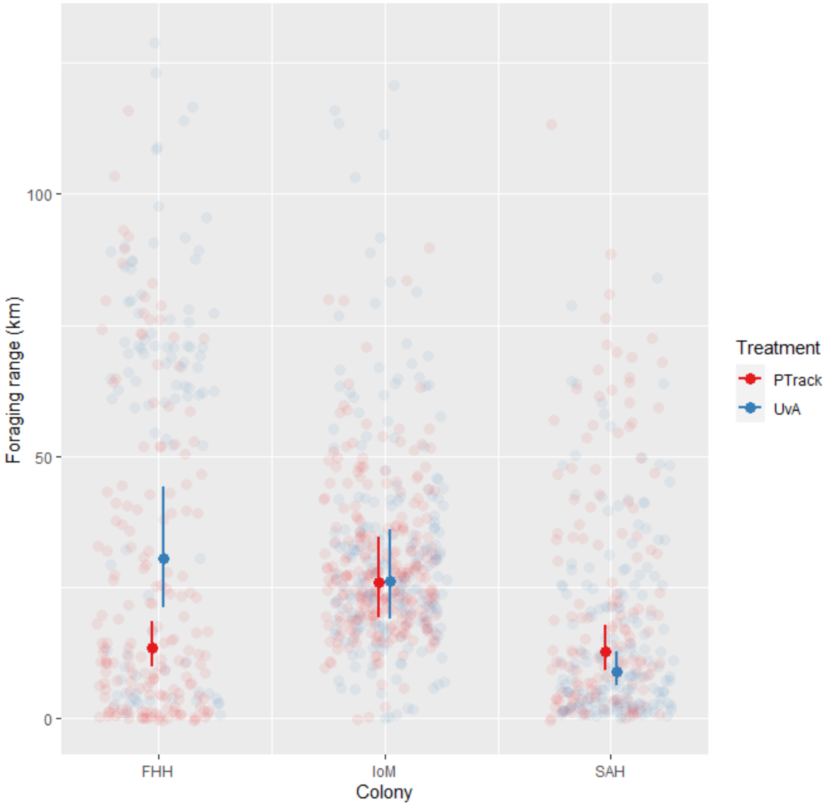


Figure 59. Predicted foraging range of trips of kittiwakes tagged with Pathtrack and UvA loggers at each colony. The raw data are plotted in the background. FHH: Fowlsheugh, IoM: Isle of May, SAH: St Abb's Head.

3.7.2.2 Utilisation distributions

Overall, there were no substantial directional differences between the logger-specific UD contours, and birds tagged with either of the logger types seemed to visit approximately the same areas at sea (Fig. 60), with high overlap between contours (Table 27). However, some differences were apparent at colony level. There were substantial differences in the 50% contours for Fowlsheugh, with birds tagged with UvA loggers accessing sites much further offshore (Fig. 61a). Also, the 95% UvA contour for the Isle of May extended directionally further offshore than the 95% Pathtrack contour (Fig. 62b). In contrast, the 50% contours for this colony showed a high overlap

(Table 27). No differences between logger-specific UDs were apparent at St Abb's Head (Fig. 63a, Table 27).

	Fowlsheugh		Isle of May		St Abb's		All	
	PTrack	UvA	PTrack	UvA	PTrack	UvA	PTrack	UvA
50%	56.182	19.569	71.553	95.924	86.009	64.991	66.650	89.573
95%	88.263	69.165	63.952	98.086	91.906	62.157	77.655	86.113

Table 27. Percentage overlap of 50% and 95% UD contours of birds tagged with UvA and Pathtrack (PTrack) loggers per colony and overall. Reference treatments are provided (i.e. PTrack means it is the proportion of the Pathtrack contour).

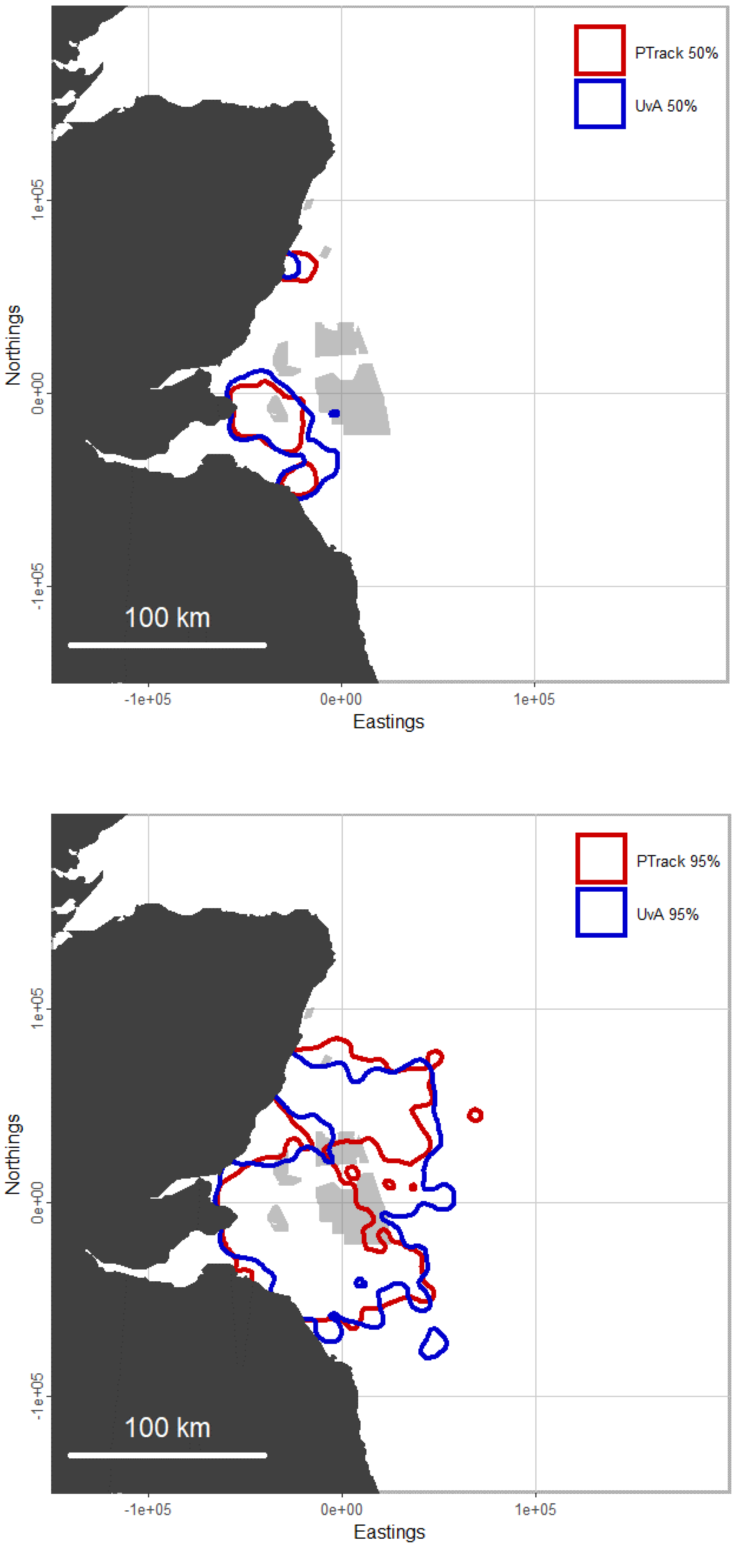


Figure 60. Kittiwake utilisation distribution by tag type for complete foraging trips from all three colonies (Fowlsheugh, Isle of May and St Abb's Head); a) 50% UD contours, b) 95% UD contours. Relevant windfarm footprints are shown in light grey. H=3.6km.

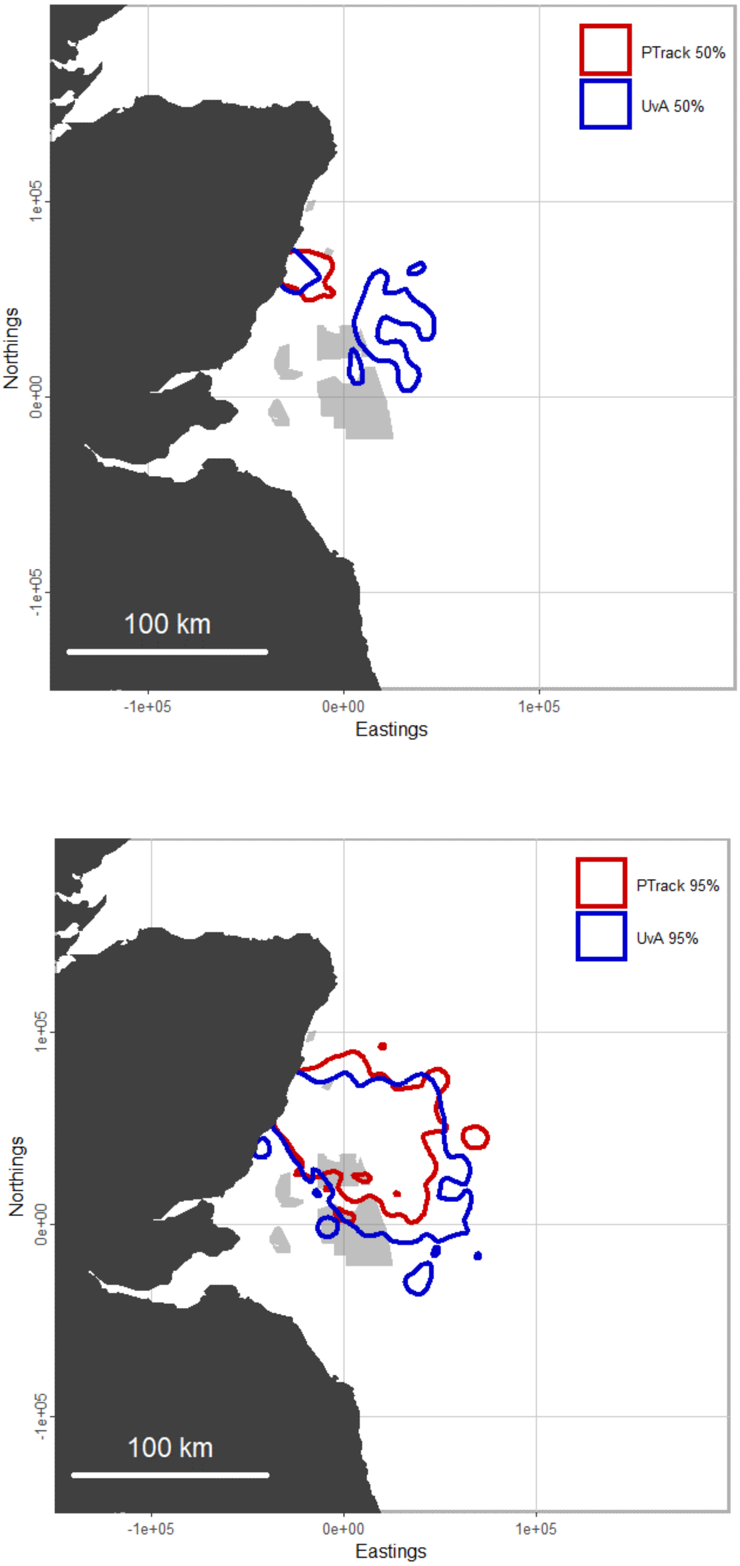


Figure 61. Kittiwake utilisation distribution by tag type for complete foraging trips from Fowlsheugh; a) 50% UD contours, b) 95% UD contours. Relevant windfarm footprints are shown in light grey. H=3.6km.

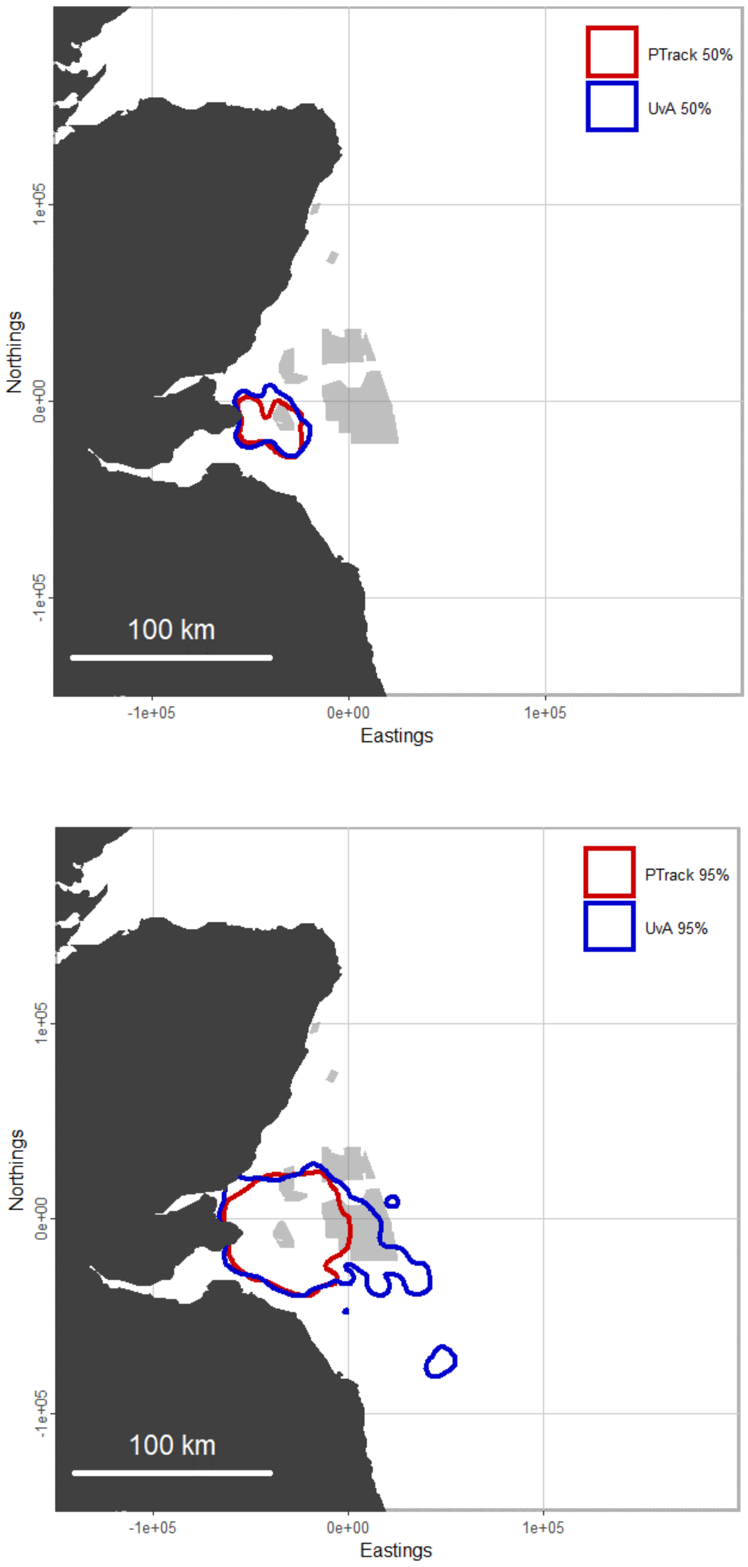


Figure 62. Kittiwake utilisation distribution by tag type for complete foraging trips from the Isle of May; a) 50% UD contours, b) 95% UD contours. Relevant windfarm footprints are shown in light grey. H=3.6km.

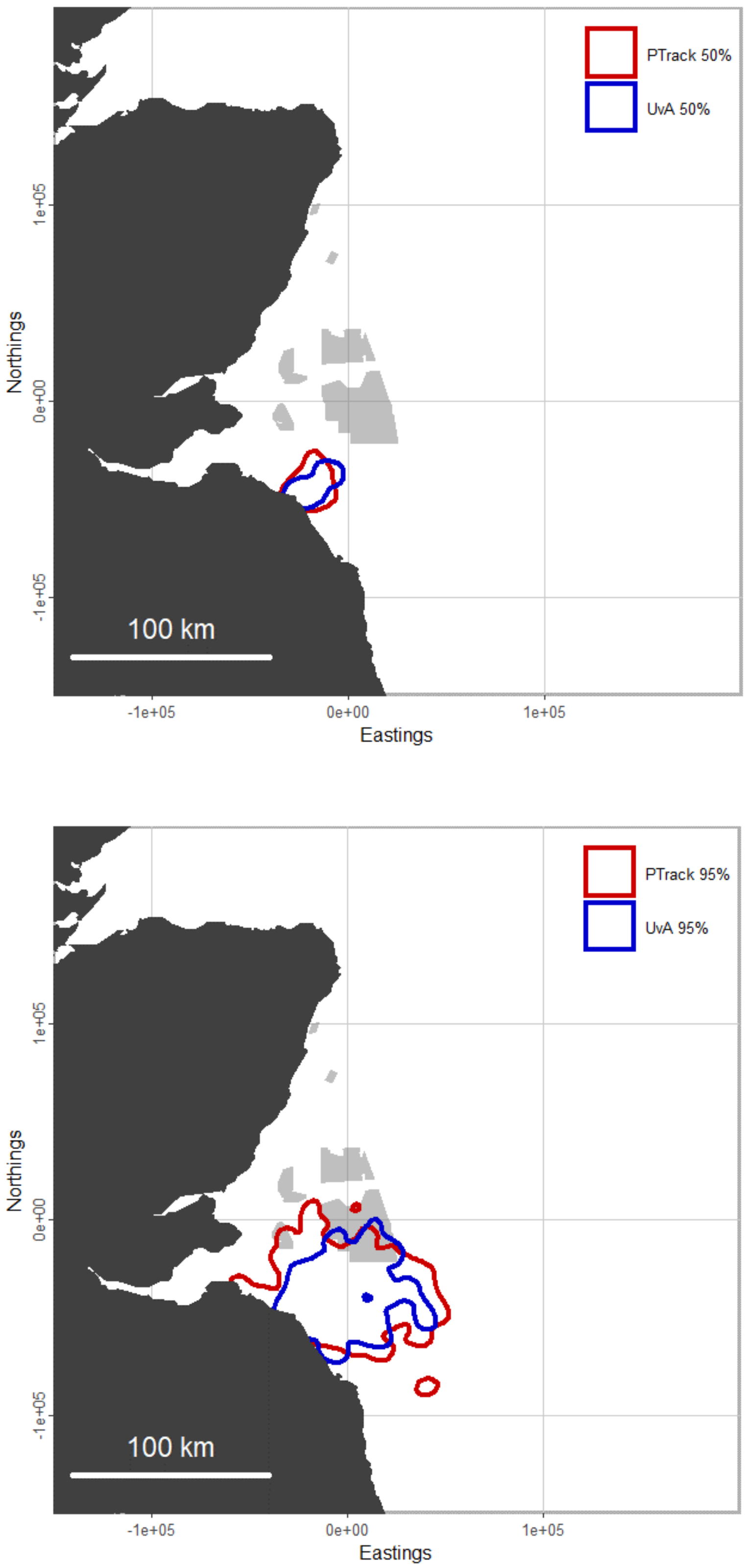


Figure 63. Kittiwake utilisation distribution by tag type for complete foraging trips from St Abb's Head; a) 50% UD contours, b) 95% UD contours. Relevant windfarm footprints are shown in light grey. H=3.6km.

4 Discussion

4.1 Data collection and processing

Overall, the remote-download GPS technology performed well. Locational data were obtained from 95% of study individuals (vs 93% in 2018, 98% in 2019 and 83% in 2020). In comparison, only 60% of individuals contributed data using archival loggers in 2010 (Daunt et al. 2011a) as not all individuals could be recaptured for logger retrieval and, in the case of kittiwakes, some of the loggers were removed by birds soon after deployment.

Extensive work was carried out to classify behaviours from GPS tracking data using hidden Markov models (HMMs). This method provided an objective approach to behavioural classification, whereby the modelling approach divided the movement behaviours into states and identified the parameters associated with each type of movement behaviour, as opposed to the user determining fixed quantitative rules to divide behaviours. We focused on two species (kittiwake and puffin) using GPS data alone; behavioural classification for the other two species (guillemot and razorbill) where we collect TDR (diving) data alongside GPS will be developed as part of the work planned for 2022/23.

For kittiwakes, a three-state model was favoured and distinct states were identified; these most likely represent commuting (state 1), foraging (state 2) and resting (state 3) behaviours (Chivers et al., 2012; Christensen-Dalsgaard et al., 2018; Daunt et al., 2002). Model selection showed support for the inclusion of all three candidate covariates in influencing the transition probabilities between states: logger type, colony, and time of day. Potential reasons for variability between logger types include differences between the individuals tagged, data recording regimes, and potential tag effects on individuals. Some or all of these factors may have contributed to the detected difference here; however, it should be noted that the effect is relatively small. The effect observed between colonies could be caused by differences between the geography of each colony location and local foraging availability. For example, if birds from one colony typically have to travel further to reach their foraging areas, they would be more likely to spend time in the commuting behavioural state. Also, if birds from different colonies visited prey patches of different densities or qualities, it may influence the amount of time they spend, or need to spend, in the foraging behavioural state. There was also a diurnal effect on behaviour switching, with resting behaviour more likely to occur at night, and commuting and foraging more likely to occur in daylight (Daunt et al., 2002). Whilst these effects are biologically plausible, it should be noted that these effects should be interpreted with caution. Residual autocorrelation within the models, suggesting that the behavioural patterns are not fully captured by the model, could cause artificial reduction in uncertainty from pseudo-replication. This effect can lead to AIC-based approaches artificially selecting the most complex models, resulting in the inclusion of covariates which may only have a marginal or insignificant effect. The limited difference in the decoded states observed between models, despite the large differences in AIC values (Table S2), suggests that this may be occurring here. Further work to evaluate whether autocorrelation influences the model selection undertaken here is required. This could be approached by investigating more advanced statistical models (e.g. continuous time models), or including auxiliary data (e.g. accelerometer data) to divide the data into more states and capture more complex sets of behaviours.

For puffins, both a two-state and a three-state model were considered. The two-state model identified two distinct states: flight (state 1) and non-flight (state 2). The three-state model sub-divided flight behaviour into periods of clear commuting flight, and periods when birds were potentially transitioning between different behaviours. Understanding the true biological function of the states within both models remains challenging, and thus the two-state model was chosen for behavioural classification in this report. However, conducting further work to understand the behaviours being observed in the data and models is an important next step. One way this could be approached is through examining the frequency of GPS fixes in the puffin data, to highlight times of potential diving activity (as GPS fixes are likely to be missed during underwater periods; Fayet et al. 2021) and improve behavioural interpretation. Another approach would be to compare puffin GPS tracking data with tracking data from razorbills for which both GPS and TDR records exist, to further examine how horizontal movement differs during different levels of diving activity. Further work could also be conducted to examine individual trips to sense-check the occurrence, sequence and location of different behaviours.

4.2 Utilisation distributions

The at-sea distributions of the four study species breeding on the Isle of May encompassed both inshore and offshore areas, as previously found at this colony (Daunt et al. 2011a, Harris et al. 2012, Bogdanova et al. 2018, 2020, 2021) and at other UK breeding colonies (Robertson et al. 2014, Shoji et al. 2016, Wakefield et al. 2017). Differences among the species were apparent, with guillemots and razorbills using coastal as well as offshore areas, whereas puffins and kittiwakes concentrated mainly in offshore areas. These differences most likely reflect variation in foraging strategies, including factors such as flight costs, foraging effort, foraging mode and diet (Thaxter et al. 2013, Wanless et al. 2018). The core foraging areas of all species included waters around the Isle of May, suggesting that food resources were available in the vicinity of the colony. Horizontal trajectories of commuting flights showed the predicted directional movement to and from the colony, both for foraging trips inshore and offshore. At the foraging grounds, flight bearings were more variable, likely reflecting movements between foraging patches. Foraging range was more restricted compared to previous years in all four species, and compared to foraging ranges at other colonies (Woodward et al. 2019), indicating that the birds undertook shorter foraging trips. Also, breeding success was higher than in 2018-2020 in all species except razorbill, and generally high in puffins and kittiwakes relative to the long-term mean for the Isle of May. Together, these patterns may suggest that local environmental conditions and/or prey availability in 2021 were more favourable than in the previous years.

At-sea distributions of kittiwakes from Fowlsheugh and St Abb's Head were mostly offshore, although the birds from St Abb's Head also used coastal waters around the colony. The core areas used by kittiwakes from the three colonies were largely segregated. Spatial segregation in distribution of neighbouring colonies has been observed in other seabird species and is likely driven by density-dependent competition for prey (Ashmole 1963, Cairns 1989, Wanless & Harris 1993, Ainley et al. 2004, Wakefield et al. 2013).

Among the three species for which we were able to separate foraging from resting at sea, distributions associated with these non-flight behaviours were generally overlapping. However, in razorbills and kittiwakes from the Isle of May there were some differences, with resting areas extending further offshore. It is possible that areas closer to the colony were utilised to forage for the chicks, whereas offshore areas may be used primarily for self-feeding and resting.

The resampling analysis suggested that the sample size of individuals we tracked was adequate to estimate the at-sea area used by the study populations of all four species during the deployment period. It is therefore likely that we have captured the key areas used by seabirds from the three colonies for both flight and non-flight activities at that time. There was some variation in minimum adequate sample size between species and between the three kittiwake populations. Notably, the minimum adequate sample size was substantially lower in razorbills compared to the other three species, however similar results were obtained in previous years (Bogdanova et al. 2018, 2020, 2021), suggesting higher consistency in distributions between individuals of this species. These findings should be interpreted with substantial caution because of the challenges in visually detecting stabilisation of values and because of the possibility that sampled individuals and time periods may not be representative of the whole population or season. In addition, there was evidence for device effects of UvA logger deployment on at-sea distribution of kittiwakes from Fowlsheugh therefore distributions in this group may not be fully representative of the distribution of unmanipulated birds. We discuss this issue further in Section 4.5.2.

It is important to note that, with the exception of kittiwakes tagged with UvA loggers, the period over which the birds were tracked was relatively short (average deployment duration was around 5 days; range: <1 day to 10 days) so caution is required if interpreting the distributions presented in this report as representative of periods outside the study period.

A comparison of the distributions of Isle of May seabirds in 2021 to earlier years (Daunt et al. 2011a, Harris et al. 2012, Bogdanova et al. 2018, 2020, 2021, Appendix 3 to this report) shows that there is significant inter-annual variation within each of the species. The inter-annual differences are likely due to variation in environmental conditions among years, particularly the distribution and availability of prey. Adult lesser sandeels are one of the main prey species of the seabirds breeding on the Isle of May (Wanless et al. 2018) and tend to be closely associated with sandy substrates (Wright et al. 2000), so areas where the birds forage on these (and hence overlap with sandy benthic habitats) can be expected to be relatively consistent/predictable among years. However, during chick rearing (when logger deployments took place) most species switch to feeding on the young of the year (0 group) sandeels that are not so closely associated with sandy habitats (Wright et al. 2000). Furthermore, other processes such as climate warming and fishing have resulted in dramatic changes in the North Sea over the last few decades (Beaugrand et al. 2008). As a result, the abundance and quality of lesser sandeels has declined and, linked to that, the proportion of other prey species in the diet, notably Clupeids, has increased (Wanless et al. 2018). Such changes in diet, with an increasing focus on alternative prey to adult sandeels, are likely to result in inter-annual differences in foraging distributions.

4.3 Connectivity with Forth/Tay wind farms

4.3.1 Neart na Gaoithe

For birds from the Isle of May, distributions associated with non-flight behaviours overlapped to a variable but generally small extent with the Neart na Gaoithe wind farm. Lowest overlap was observed in guillemots (<2%), followed by razorbills (8%), reflecting their more extensive use of inshore waters compared to the other two species. In puffins and kittiwakes, overlap was higher, in particular for core areas (14% and 19%, respectively). Note also that the entire wind farm footprint fell within the core and overall areas used by puffins and kittiwakes, and large parts of it fell within the overall area for razorbills from this colony. In terms of commuting flight activities, all species breeding on the Isle of May except guillemot used the wind farm extensively, in particular puffins and kittiwakes where 96% and 100% of the study birds, respectively, passed through the footprint. Overlap in guillemots was substantially lower (13%) reflecting the predominantly inshore distribution of this species. These results suggest that Neart na Gaoithe may potentially pose a higher risk for puffins and kittiwakes than for razorbills and guillemots, depending on the degree of displacement following construction. We found little evidence for connectivity between kittiwakes breeding at Fowlsheugh and St Abb's Head and Neart na Gaoithe. The areas used for foraging and resting by Fowlsheugh birds did not overlap with the wind farm footprint and no flights were recorded passing through it, although some trips in close proximity were recorded. Similarly, there was virtually no overlap between the at-sea areas used by birds from St Abb’s Head and only a single flight from this colony was recorded passing through the footprint. However, since environmental conditions and consequently seabird at-sea distributions vary between years, and the wind farm is located within the foraging range of kittiwakes from these colonies, some level of connectivity with the wind farm area in other years may occur.

4.3.2 Seagreen

Overlap between foraging and resting distributions and the Seagreen wind farm was low overall across the four species and three study colonies, likely reflecting its location further offshore, in particular relative to the Isle of May and St Abb’s Head. The foraging and resting areas of birds from all four species tracked on the Isle of May did not overlap with the Seagreen footprint, and those of kittiwakes from the other two colonies overlapped to a small extent (between 0 and 5%). Similarly, commuting flights of guillemots and razorbills did not overlap with the Seagreen footprint. Overlap in flight activities of puffins and kittiwakes from the Isle of May with this wind farm was low. Kittiwakes from Fowlsheugh used Seagreen more extensively, with 58% of the study birds passing through it at least once, likely reflecting the closer proximity of this colony to the wind farm. No kittiwakes from St Abb’s Head were recorded flying through the Seagreen footprint. Our results, therefore, suggest that this wind farm may pose a higher risk to kittiwakes from Fowlsheugh than birds from the other two colonies.

4.3.3 Berwick Bank

Overlap between core foraging and resting areas of birds from all species and colonies and the Berwick Bank wind farm was negligible (<0.5%) or zero. Overlap with the overall areas used varied between species and colonies, from no or low (≤2%) overlap in the auks to nearly 20% in kittiwakes from St Abb’s Head. Commuting flights of guillemots and razorbills did not overlap with Berwick Bank, although some foraging trips reached the western edge of the footprint. Puffin flight activities overlapped relatively little with this proposed wind farm (17% of birds and 2% of flights passed through it). Kittiwakes from all three colonies used the proposed wind farm footprint , although to a variable extent, with highest proportion of birds from the Isle of May passing through it (72%), followed by birds from Fowlsheugh (35%) and St Abb’s Head

(11%). As with Neart na Gaoithe, our findings suggest that Berwick Bank may pose a higher risk for puffins and kittiwakes than for razorbills and guillemots which have more inshore distributions.

The impacts of offshore wind farms on seabirds can be positive or negative (Inger et al. 2009). Recent post-construction studies in European waters (reviewed in Dierschke et al. 2016, see also Vanermen et al. 2020, Peschko et al. 2020, 2021) demonstrate that responses of seabirds to offshore wind farms can vary substantially, ranging from strong avoidance to strong attraction, with some species showing little change in behaviour. Guillemots and razorbills were among the species showing avoidance, whereas kittiwakes showed mixed responses at different wind farm sites; data on puffins were lacking. Furthermore, the strength of the response can differ among populations of the same species, most likely linked to factors such as local food availability and distance of the development from the colony (Dierschke et al. 2016). Given the extent of variation in seabird distributions and responses to offshore wind farms (both among and within species), to gain a robust understanding of the potential effects proposed offshore developments may have on local seabird populations, ideally tracking, energetics and demography data should be collected over several years spanning before, during and after construction from multiple relevant breeding populations. This is an approach we are currently following within this study, with a particular focus on kittiwakes for which such data are more readily obtained (Daunt et al. 2021).

4.4 Estimation of flight altitude in kittiwake

Estimation of kittiwake flight altitude proceeded by two basic steps. Firstly, an initial processing and estimation step to calculate flight altitude values that subsequently fed forwards into a second, state-space modelling step. The level of processing in this first step was more involved when using barometric pressure than GPS estimates. In the case of GPS, all processing was done by the tag itself, providing measures of altitude that we used for state-space modelling without further work on our part. In contrast, we developed multiple different approaches to estimate altitude using the pressure values recorded by UvA tags. Such approaches were either based on identifying periods when birds were floating on the water to estimate pressure at sea-level or by combining tracking data with remote sensing datasets or data collected from a local weather buoy.

4.4.1 Comparison of altitude estimates from GPS and barometric sensors

Overall, we found that estimates of altitude at Fowlsheugh and St Abb’s Head were generally moderately to strongly correlated regardless of the exact approach we used. Crucially, such correlations included those between measures of altitude based on barometric pressure and those based on GPS signals. The correlation between GPS and barometric estimates of altitude also increased after application of state-space models for these colonies. In part this may occur because restricting negative estimates using truncated distributions reduced the influence of more extreme negative estimates of altitude observed during the initial processing which are more likely to be outliers. In addition, estimates of flight altitude during periods defined as floating were centred around 0 regardless of which approach was used to estimate altitude. While we lack true ground truth values to calibrate our results (i.e. observations where we could categorically state birds were floating on the water) it was nonetheless reassuring that locations classified as floating using accelerometry data gave reasonable altitude values. At the Isle of May correlations between barometric- and GPS-based estimates of altitude were weak or moderate. This pattern persisted even after the application of state-space models to altitude estimates. We believe the most likely explanation for this result is that the Isle of May data contained less instances of high-resolution sampling when we expected both sensors to give better altitude estimates. For example, no periods of ≤ 16 second sampling were observed at the Isle of May and sampling intervals ≤ 1 minutes were also less common than at Fowlsheugh or St Abb’s Head.

When birds were classified as in flight (not floating), we observed a tendency for results based on barometric pressure to give slightly higher altitude estimates than those based on GPS. The magnitude of this difference varied slightly across colonies and individuals, but barometer estimates had a median altitude 3.5-6.0 metres higher than corresponding GPS altitude. Such a difference could have important implications for the results of flight height collision risk modelling. Why such a difference occurs is currently unclear. One potential explanation is that as methods based on barometric pressure require constant re-calibration when birds are on the sea to determine P_0, there is a tendency for altitude estimates to drift when tags have not been recalibrated during long periods of flight. The degree of miscalibration will also depend to some extent on prevailing weather conditions. This might also explain why we observe no consistent difference in altitude estimates when birds are floating as in this case barometers can be calibrated easily. However, regardless of time since calibration plots of raw data suggested barometric altitude estimates were higher than GPS estimates and inclusion of time since calibration as a covariate within our state-space models did not change our predictions. Moreover, while drift and miscalibration are expected to bias results from barometers to some extent, it is striking how consistently barometric pressure led to higher estimates of altitude on average than GPS but never lower estimates which might also be expected due to drift.

In addition to differences in estimated altitudes between barometric and GPS-based approaches we found that altitude distributions based on ERA5 pressure data had a greater spread than other approaches. This is most apparent in the greater extent of the interquartile range (IQR) when using ERA5 data to measure P_0. Currently, hourly ERA5 data is only available at a relatively coarse spatial scale (0.25° grid) preventing more fine-scale estimation of P_0. Consequently, the same value of P_0 is used for many observations over a large area but is likely to be better calibrated for some points than others. A further issue we encountered when using ERA5 data was the poor estimation of altitude from spatial tiles that included a significant proportion of land, a situation that occurred frequently in coastal areas. In such cases, while altitude estimates based on ERA5 and other approaches were correlated, ERA5 pressure was consistently offset from logger pressure, sometimes by large amounts, leading to nonsensical estimates of altitude. In theory, data on sea level pressure based on remote sensing data, such as ERA5, or from local weather buoys could obviate the need to calibrate based on such floating behaviour. However, in practice, ERA5 pressure data was calibrated to pressure readings from UvA barometers, necessitating the identification of floating behaviour as a reference. For example, while ERA5 pressure data was calibrated to pressure data recorded by UvA loggers using a modelling process, estimated altitudes for birds at the Isle of May in 2020 are based on raw weather buoy data. In this case, altitudes based on weather buoy data are typically higher than those based on pressure recorded by data loggers. Most pertinently, estimates of altitude while birds were classified as floating had a median value of 11 m based on buoy data. In general, barometric pressure needs to be calibrated between different devices (Xia et al. 2015, Manivannan et al. 2020). Consequently, we would not recommend the use of ERA5 or weather buoy pressure data without suitable comparative data from a bird-borne logger for calibration purposes. Such calibration seems less of an issue when pooling pressure data from multiple co-deployed UvA loggers which was achieved here without calibration between different loggers. However, this could potentially be an issue in other data sets.

Even within individual UvA loggers, re-calibration of P_0 will be important as pressure drifts throughout deployment due to changes in environmental conditions. Moreover, calibration of loggers may drift over time even in calm conditions so loggers may require occasional re-calibration with themselves (W. Bouten, pers comms). Here, we used periods of floating behaviour to re-calibrate UvA loggers with suitable P_0 values. However, this raises two issues. Firstly, a behavioural annotation method is required to identify floating behaviour. We used data from accelerometers to achieve this, exploiting an existing UvA algorithm that had been developed for large gulls, but where it was thought that floating behaviour would appear similar regardless of species. Other methods of behavioural annotation are available such as hidden Markov modelling (HMMs, Langrock et al. 2012) or Expectation-maximization binary clustering (EmBC, Garriga et al. 2016) among others which could be used in the absence of accelerometery data. Indeed, the current report demonstrates the utility of HMMs in this regard. However, we used data from accelerometers because it allowed behaviour to be sampled at the same, irregular frequency as GPS spatial coordinates. As such it could deal with the irregular nature of tracking time-series. In contrast, HMMs require tracking time-series to be at regular intervals and so information about behaviour at fine temporal scales would be lost. Regardless of the annotation method used, accurate identification of floating behaviour is crucial (Poessel et al. 2018). In the worst-case scenario, many instances of birds in flight may be erroneously classed as floating which would subsequently bias estimates of P_0 above true sea-level. Classifying floating behaviour incorrectly as flight would also be problematic as valuable information on sea-level pressure would be lost by discarding useful records but this should not itself lead to biases in P_0. Unfortunately, in most cases we lack a genuine ground-truth assessment of P_0 as we do not observe birds in situ during tracking. As an alternative, we relied on assessing flight altitude during floating periods assuming that it should be tightly distributed around 0. In most cases our estimates of altitude do indeed centre around 0, though when using ERA5 data at St Abb’s Head median altitude during floating was ~ 3 metres flagging a potential issue in this case. Secondly, periods of floating are not observed systematically, and we are reliant upon when and where birds engage in this behaviour for our calibration. Consequently, the quality of any calibration will vary both between loggers and throughout the course of a single deployment. Foraging trips with few observed instances of floating will not be as well calibrated as trips in which floating was observed frequently. Likewise, long periods of flight may result in some observations being poorly calibrated because they are not close in time and / or space to a suitable calibration period. Given that seabird behaviours are unlikely to be uniformly distributed throughout their foraging range, this might result in certain areas being less well calibrated when using barometer data than others. For example, if birds rarely engage in floating behaviour while commuting through certain areas of their range, estimates of flight altitude here may be affected.

Our analyses also show that the different approaches to determine altitude are likely to vary between different colonies and between different individuals. For example, correlations between GPS and barometer derived altitudes differed between birds tracked at Fowlsheugh, Isle of May, and St Abb’s Head. Similarly, while for some individuals GPS and barometer altitudes were highly correlated, for others this was not the case. Currently the exact reasons for this are unclear but nevertheless it highlights the difficulty in developing an approach that will work consistently across logger deployments.

The algorithms we developed to extract altitude estimates based on barometric pressure involved setting many different parameters which could be examined further via sensitivity analysis. For example, our choices of both the number of floating observations required within set time windows (set at 5 currently) as well as weighting schemes could all be adjusted. For instance, here we relied on inverse distance weighting by time (single logger calibration) or spatio-temporal distance (calibration across multiple loggers) as it was thought the simplest approach to implement. Moreover, when calibrating across UvA loggers we also used a spatial buffer of 5 km to determine when data from one logger could be used to calibrate another. More restrictive choices could be made regarding these parameters. However, while these could lead to more reliable calibration and hence better altitude estimations they will also result in more missing data as fewer observations meet these stricter criteria. In addition, we currently perform no additional filtering on the pressure data collected by UvA loggers, though some is done by the logger itself. Additional filtering steps could be used to remove the more extreme outliers we observe in pressure data but would also increase the number of parameters and processing decisions made by practitioners including the choice of which filtering algorithm to use.

In contrast to barometric pressure, GPS estimates of altitude did not require constant recalibration and included fewer missing altitude values. Missing altitude values when using barometric pressures occurred because some observations could not be calibrated using the current parameterisation of our processing algorithms. For example, some points were not within 12 hours of a suitable period of floating behaviour. In addition, UvA loggers also occasionally failed to record pressure for some observations meaning that no estimate of altitude was possible. While some GPS altitude fixes were also missing, the rate of missing altitude fixes was lower than the number of missing pressure values. Thus, one advantage of GPS estimates is a more complete dataset. Some GPS altitude fixes were not of particularly high quality based on reported DOP values or vertical accuracy but could still be used in our state-space models. GPS estimates also required less initial processing on our part than using barometric pressure which is advantageous in terms of computing time and reduces the need to set user-defined parameters to calibrate loggers. As with barometric pressure, GPS estimates could be filtered prior to state-space modelling but this requires practitioners to decide upon a preferred filtering algorithm and parameterise it accordingly.

4.4.2 State-space modelling of altitude

State-space models of altitude were developed for the Fowlsheugh, Isle of May and St Abb's Head colonies based on both GPS altitude and altitude estimates from barometric pressure using the single logger calibration approach. State-space models could also be utilised to estimate altitude using the pooled logger approach or ERA5 data, but such models were not conducted here due to time constraints. SSMs were carried out on a trip-by-trip basis after a larger model incorporating random effects failed to converge. Depending on the application it is worth considering if the additional effort of random effects modelling is required (McClintock 2021), but this may be an extension worth further consideration if there is an interest in specifying the degree of between-individual and between-trip variation in flight altitudes. Currently, we rely on a two-step estimation approach to aid in our interpretation of state-space models at the population-level.

The process equation component of our state-space models included covariates for altitude during the preceding time step and the vertical velocity. This approach was designed based upon recommendations in Peron et al (2020). However, Peron et al. (2020) suggested only using covariates for vertical velocity at high sampling rates ≤ 1 min. In the current study the sampling rate was variable shifting from high sampling rates ≤ 1 min to rates of 5 to 10 minutes. To address the irregular nature of altitude time-series we therefore included an additional covariate for the time lag between successive observations as well as the two-way interaction between time lag and both the previous altitude measure and vertical velocity. Across all colonies and both altitude estimation techniques state-space coefficients were generally similar in both their magnitude and directions. The one exception was the direction of the two-way interactions between sampling frequency and vertical velocity which was positive at St Abb's Head for both sensors but was negative in Fowlsheugh for barometric estimates of altitude and essentially 0 for estimates based on GPS. However, coefficients for two-way interactions were generally small and did not have a great impact on altitude estimates.

Modelling the variance in the observation process, we found that the precision of estimates based upon barometric data was greater when the sampling frequency was ≤ 1 minute. This difference was also apparent in plots of raw altitude estimates and supports findings that higher resolution tracking data results in greater precision of altitude estimates (Cleasby et al. 2015). Similarly, we found that GPS altitude estimates were more precise when the sampling rate was ≤ 16 seconds in line with our previous expectations based upon manufacturers details concerning the operation of UvA loggers. More broadly, a higher GPS sampling rate may be expected to increase precision of altitude estimates but the thresholds at which this increase occurs will depend in how loggers are set-up and operate. As a result, while a sampling frequency of ≤ 16 seconds improves precision from UvA loggers, other loggers may perform differently, and as such observation error should be modelled accordingly. Alongside sampling frequency, we also found that precision decreased with the extent of vertical accuracy such that less accurate estimates were more imprecise as expected. While we used vertical accuracy here, preliminary examinations using positional DOP (also provided by UvA logger) in place of vertical accuracy yielded similar results, namely that precision decreased as DOP increased as expected.

There remains scope to extend the state-space modelling approach further. For example, by including weather-based predictors such as wind speed of bird's flight speed (Lane et al. 2019). In addition, depending on the exact goals of any analysis we could also examine differences in flight altitude in different behaviours by including categorical predictors of behaviour based on accelerometery data or another behavioural annotation technique such as the HMMs provided here (see also: Ross-Smith et al. 2016). Such an approach was not pursued here but would provide a means of assessing flight altitude during different flight modes (e.g. commuting versus foraging or inbound versus outbound flight). However, the success of such an approach will rely on the quality of any behavioural classification. Alternatively, flight altitude could be included as an additional data stream to classify bird behaviour alongside more traditional movement measures such as step length and turn angle. Such an approach might be more straightforward using GPS-based estimates as barometric-based estimates of altitude require behavioural classification of floating behaviour prior to their estimation so flight behaviour would be classified twice. We also noted that, regardless of the exact set-up of our state-space models, estimated altitude closely resembled the original altitude values. The main exception being that lower estimates were restricted to -1 m or above. As such, there is probably more scope to influence altitude estimation by tweaking the initial processing algorithms for barometric pressure data than during the state-space modelling procedure. A further consideration is whether to further limit the range of altitudes permitted by the state-space model. Currently, a lower bound of -1 metre altitude is included within the modelling framework but it would be feasible to include an upper bound as well. Deciding upon an appropriate upper bound to altitude estimates will be more difficult than defining a lower bound but may help reduce the influence of occasional extreme values via the processing of GPS or barometer data.

4.4.3 Comparison of kittiwake altitude estimates with other data sources

In the current work kittiwake altitude distributions did not appear to differ widely between the three colonies sampled when using the same processing method. For example, when using barometric pressure and the single logger calibration method the median altitude of birds classified as in-flight was 9.38 m at Fowlsheugh, 10.43 m on the Isle of May and 10.26 m at St Abb's Head. Inter-quartile ranges (IQR) were also similar across colonies. As noted, GPS altitude estimates tended to be lower than barometric estimates, but even comparing across barometric and GPS-based methods there was a high overlap between IQRs (Tables 13-15). The SSM altitude distributions presented in Fig. 46 also look similar across colonies and in each case the bulk of in-flight altitude estimates were below 40 metres. Other sources of kittiwake altitude data are also available. For example, Johnston et al. (2014) provided modelled flight height distributions for many species of seabird based on a collation of different observer-based surveys (i.e. not based on tracking data). These flight height distributions have subsequently been incorporated into collision risk models as generic species-level flight height distributions in many cases (e.g. Masden et al. 2015). A comparison of SSM altitude estimates with the altitude distributions provided by Johnston et al. (2014) shows that the current estimates, based on telemetry data, typically result in higher proportions of birds on or close to the sea surface (Fig. 64). Because we included information of behaviours classified as floating and in-flight when producing these curves this may also reflect the amount of time birds spend floating on the water. The proportion of birds flying at altitudes ~ 5 – 25 m was slightly higher based on the Johnston et al. (2014) data, but after this point there was little difference between altitude frequency distributions and few birds were expected to fly higher than 40 metres across each data source.

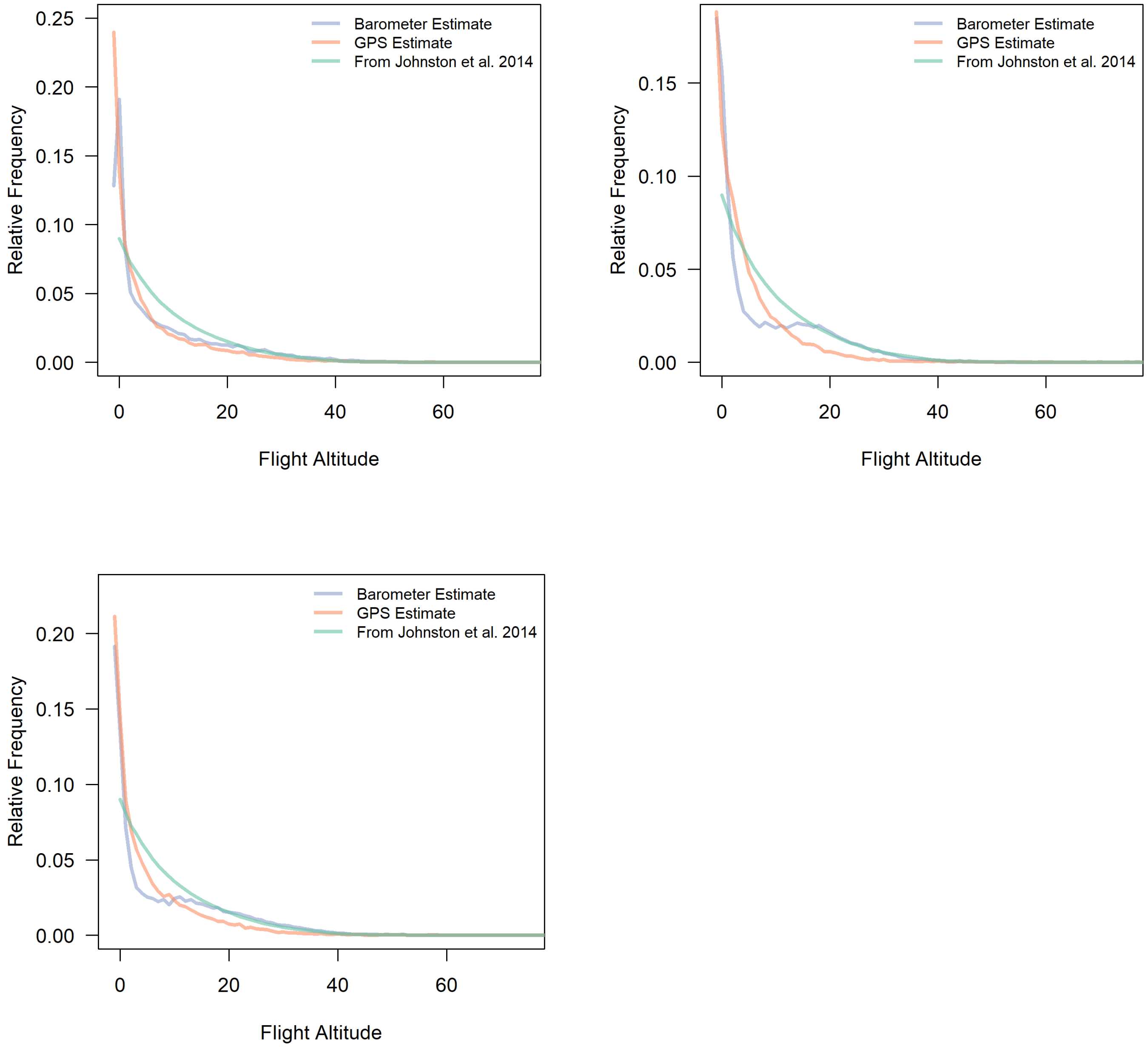


Figure 64. Comparison of flight height distributions at Fowlsheugh (a), Isle of May (b) and St Abb's (c) based on estimates GPS- or barometric-based SSMs with data provided by Johnston et al. (2014). The y-axis shows the proportion of measures taken in each 1 metre altitude band. Data for Johnston et al. (2014) was sourced from Masden (2015). SSM curves based on behaviours classified as floating and in-flight.

Declines of kittiwake flight heights above 40 metres were also noted by radar tracking studies in the Aberdeen Offshore Wind Farm (Tjørnløv et al. 2021). In this study, low numbers of kittiwakes were recorded flying between ≥ 40 metres and turbine hub height (105 metres). In addition, 56% of kittiwakes were also recorded flying at altitude below rotor height (25 metres). Here, we observed higher proportions of birds flying below 25 metres as can be seen by the IQRs reported in Tables 13-15 or Figure 46. However, the radar study collected cross-sectional data across a localised area close to wind turbines which seabirds may react to and over a longer time period spanning incubation and/or chick rearing, therefore results may not be expected to align exactly with tracking data.

Altitude distributions obtained using LiDAR (Light Detection and Ranging) have previously found to differ from those obtained via aerial or boat-based surveys (Cook et al. 2018). Using LiDAR data to estimate kittiwake flight altitudes in the Outer Firth of Tay, Cook et al. (2018) reported that across three surveys conducted throughout September the proportion of altitude estimates falling within a 20 m – 120 m risk envelope was 0.043. During another LiDAR-based survey at Hornsea, kittiwakes were not observed flying higher than 30 metres and were most frequently observed flying at 5 metres or below. Moreover, the proportion of birds flying within a defined risk envelope (20 m – 120 metres) was 0.11 (4 / 34 observations), though note the overall sample size here was relatively low (NIRAS 2018). Corresponding values from the current study based on estimates from SSMs were slightly higher (Proportion of birds between 20 – 120 m at Fowlsheugh – GPS-based = 0.083, Barometer-based = 0.14; Proportions at Isle of May – GPS-based = 0.047, Barometer-based = 0.13; Proportions at St Abb’s – GPS-based = 0.071, Barometer-based = 0.16).

Overall, different approaches to data collection are likely to explain some of the differences between studies. For example, birds may not be tracked at the same time and at the same location creating variation between datasets. More generally, the nature of the data collected will vary between methods. Tracking data typically generates a densely sampled time-series of auto-correlated measures from a single deployment whereas survey techniques such as LiDAR or radar are cross-sectional. Similarly, tracking data are often limited to breeding individuals whereas other techniques may include immature and non-breeding individuals. Observer-based studies also tend to be more limited in their spatial extent than tracking data. Perhaps as a consequence, observer-based studies have reported important levels of spatial variation in estimated altitudes (Johnston et al. 2014, Johnston & Cook 2016, Cook et al. 2018). Here, we did not observe clear differences between colonies in flight height distributions, but this does not preclude fine-scale spatial variation in flight altitude across a given colonies foraging range, which has been reported in tracking studies from northern gannets (Cleasby et al. 2015).

4.4.4. GPS or barometric measures of flight altitude?

Ultimately, both GPS and barometric sensors were able to provide estimates of flight height. Moreover, both sensors appeared to perform better when sampling at higher resolutions (≤ 1 minute). Thus, whichever sensor is to be used, we would recommend high sampling resolutions, but this will come at the cost of battery life and therefore the length of time a logger will collect data. In the case of UvA loggers, sampling at ≤ 16 seconds should result in the best performance if using GPS altitudes, but to spare some battery life it might be possible to switch from ≤ 16 second sampling to ≤ 1 minute sampling rates in pre-defined areas without drastic declines in precision.

When properly calibrated, barometric sensors are generally thought to give more precise estimates of altitude (Peron et al. 2020). However, calibration can be challenging and needs to be regularly updated. The methods used here rely upon identifying when birds are floating. Hence, using barometers required a larger tag incorporating a GPS-unit, a barometer, and an accelerometer. Thus, there may be an increase in ethical concerns about device effects related to logger mass and design using this approach (Bodey et al. 2018, see also: Section 4.5). In contrast, most GPS loggers record altitude alongside locational data and do not require the identification of floating behaviour for calibration. Ideally, GPS units would also report DOP or other values related to GPS error and accuracy which can then be incorporated within SSMs. The barometric sensors used here did not report such convenient measures of error.

One caution with using GPS, even at higher resolutions, is the occasional recording of altitude estimates that seem extreme or erroneous. For example, in Figures 42 and 43 the left-hand side of the plots highlights regions in which barometric altitude is low but GPS altitude is high. We believe this is caused primarily by GPS error because some of the observations highlighted were classified as floating behaviour and while barometric estimates for these observations seemed reasonable and were centred on 0 as expected, GPS estimates varied widely and were centred on altitudes greater than zero. Altitude estimates based on barometric pressure also result in occasional extreme examples but appeared to do so less frequently. Finally, it should be noted that the different altitude estimation approaches used here may not perform in the same manner in different locations, across different species, or on data loggers from other manufacturers. Therefore, we would recommend that where possible studies initially use both barometric and GPS estimates of altitude to cross-check estimates. In some species, the larger size of tags incorporating barometers may prove prohibitive, and in this case our results demonstrate that smaller GPS-only loggers could be used to estimate altitude, ideally using high sampling rates.

4.5 Device effects

4.5.1 Puffin

As in 2019 and 2020, we used a capture and logger deployment protocol that aimed to minimise negative effects of the associated disturbance on the study birds. We used only the smallest available loggers, captured the birds at their burrows to ensure only one adult per pair was tagged and the location of all chicks of instrumented birds was known, and carried out deployments when the chicks were older and more robust. This study design also meant that we could undertake supplementary feeding of all chicks of tagged birds as the burrows of all tagged birds were known.

In contrast to previous years, we did not find evidence of substantial negative effects of handling or GPS logger deployment on chick provisioning rates in puffins. There was some evidence for a reduction in feeding rates by both colour-ringed and GPS-tagged individuals compared to the expected if pair members shared chick provisioning duties equally as is typical in this species (Harris & Wanless 2011). However, no difference was observed between treatment and control burrows, indicating that the unmanipulated partners were able to compensate for their mate's reduced provisioning rate, a phenomenon that has been observed in other puffin studies (Harris & Wanless 2011, Symons & Diamond 2019). As we carried out supplementary feeding of all chicks at treatment burrows (based on findings in previous years), the chicks from treatment burrows fledged in better condition compared to controls. Chick survival was high in all study groups. The lack of evidence for negative device effects, coupled with high breeding success (in puffins and other seabirds species on the Isle of May) suggests that local food availability may have been higher in 2021 compared to previous years, which may have allowed the birds to maintain better condition and buffer against stressors.

The lack of evidence for device effects in 2021 provides us with the opportunity to compare the puffin distribution in this year to distributions in years where negative device effects were apparent, and gauge whether logger deployment may have affected the representativeness of the data. The areas used by the birds in 2021 were within the same general direction from the colony as in previous years, although the overall distribution was more restricted, resulting in higher overlap of core areas with Neart na Gaoithe (Appendix 3). It is possible that individuals experiencing negative device effects tend to be distributed further offshore on average than undisturbed birds, and that shorter trips to locations closer to the colony, often associated with foraging for the chick (Harris & Wanless 2011), have been under-represented in the 2018-2020 data. However, the more restricted distribution in 2021 may also reflect more favourable foraging conditions in this year.

4.5.2 Kittiwake

Our previous studies of device effects in kittiwakes on the Isle of May and at Flamborough and Filey (Bogdanova et al. 2020, 2021; Wischnewski et al. 2018) suggest that device effects may only be apparent for some logger types and in some years or at some locations. Therefore, in order to separate device effects from potentially confounding interannual and between-colony variation, in 2021 we conducted a more extensive study involving simultaneous deployments of the two logger types used previously (Pathtrack and UvA) on the Isle of May, at Fowlsheugh and St Abb's Head. This approach allowed us to compare the effects of the two logger types on key measures recorded at the colonies in tagged and control birds (parent changeover rates, chick attendance, adult and chick body condition and breeding success). It also enabled us to compare key measures of at-sea behaviour between the two logger types (foraging trip duration, distance travelled and at-sea distribution).

However, it was not possible to tease apart logger and attachment effects since Pathtrack loggers were taped to tail feathers and UvA loggers were glued to back feathers. Using the same attachment method for both logger types was not feasible: UvA loggers need to be glued to feathers to ensure that the solar panel is unobstructed and maximum data volumes are obtained, so tape attachment would not be appropriate for those. Similarly, it is not appropriate to glue Pathtrack loggers as this would result in birds carrying them for considerable time after they have stopped collecting data, which should be avoided for welfare reasons.

We found small but measurable negative effects of UvA logger deployment on kittiwake at-colony behaviour. At all study colonies, there were fewer parent changeovers per day (indicative of longer foraging trips) at nests where one bird was tagged with a UvA logger than at control nests. Chicks of UvA-tagged birds were left unattended more often than those at control nests, although chick attendance was generally high in all groups. No difference in changeover rates or chick attendance was found between Pathtrack and control nests. Data on chick and adult condition were available for the Isle of May only and indicated that chick condition was not affected by deployment of either logger type. There was some indication of a weak negative device effect on adult body condition, for both logger types; however, it was difficult to determine if this was a genuine effect due to the lack of comparable data from control birds, which comprised cross-sectional data across birds, not repeated measures of the same individuals. Breeding success at UvA and Pathtrack nests was similar to that at control nests, except at Fowlsheugh where UvA nests fledged fewer chicks and Pathtrack nests fledged more chicks than controls. Device effects on at-sea behaviour were found at Fowlsheugh only, where UvA-tagged birds made longer foraging trips (duration, total distance and range) compared to Pathtrack-tagged birds. This was reflected in the at-sea distributions for this colony, with core areas of UvA birds being further offshore than those of Pathtrack birds. There was no significant difference in foraging trip metrics or distributions between the two logger types on the Isle of May and St Abb’s Head.

In summary, there was variation between colonies in the magnitude of device effects on the key measures considered, for reasons that are difficult to establish at this stage. However, the findings from 2020 and 2021 indicate that UvA loggers (attached with glue) are having a negative impact on the birds unlike Pathtrack loggers (attached with tape). This raises concerns about animal welfare and representativeness of the data obtained by UvA loggers that should be considered for future tracking work.

4.6 Conclusions

In this project, we undertook GPS tracking of kittiwakes, guillemots, razorbills and puffins on the Isle of May, and of kittiwakes at Fowlsheugh and St Abb's Head in the 2021 breeding season. Sample sizes were sufficient to obtain at-sea distributions representative for these populations over the deployment period. The technology was successful, ensuring that data were obtained from most individuals. Substantial advances were made also with data processing, in particular with respect to behavioural classification and estimation of flight heights from GPS and associated sensor data. These steps are of key importance for assessing collision risk and potential displacement effects associated with wind farms in the Forth/Tay region.

Kittiwake flight heights estimated via GPS- or barometric-sensors were generally moderately or strongly correlated, although there was some variation both between sites and between different individuals. Altitude estimates based on barometric sensors also tended to be greater than those from GPS sensors (~ 3 – 6 m higher) when birds were assumed to be flying but not when birds were classed as floating on the sea surface. Across the three colonies tracked distributions of flight altitude appeared to be similar. In addition, regardless of the method used, kittiwakes were rarely estimated to fly at altitudes above 40 metres, which supports previous findings in this species.

There was considerable variation among species in at-sea distribution, and consequently in the extent of overlap with the Forth/Tay wind farms. Further, our results highlight the marked variation in distribution among years within species breeding on the Isle of May, both in terms of directionality and range. Given the extent of interannual variation in at-sea distributions, during a period when the North Sea is experiencing significant environmental change, the GPS data we have obtained are extremely valuable in quantifying at-sea distribution of breeding birds in the absence of a wind farm. This is an important step within a structured before-during-after monitoring protocol, as is the inclusion of additional energetics and demographic parameters collected at the breeding colony, at the individual level, to maximise the power to detect effects and thus quantify population-level impacts of wind farms in the study region. Device effects in puffins and kittiwakes remain an issue that warrants further investigation and careful interpretation due to their potential to affect the representativeness of the collected data.

5 Acknowledgements

We thank Polly Tarrant, Nick Brockie and Andrew Logie for support throughout the project. We are grateful to Ellie Owen and the RSPB Fowlsheugh Reserve team for help with fieldwork planning and logistics. We are also grateful to NatureScot for access to the Isle of May and to the National Trust for Scotland for access to St Abb's Head NNR. The work was conducted under licence from the British Trust for Ornithology and NatureScot.

6 References

Ainley DG, Ribic CA, Ballard G, Heath S, Gaffney L, Karl BJ, Barton KJ, Wilson PR, Webb S (2004) Geographic structure of Adélie penguin populations: overlap in colony-specific foraging areas. Ecol Monogr 74: 159–178.

Ashmole NP (1963) The regulation of numbers of tropical oceanic birds. Ibis 103: 458-473.

Band, W (2012) Using a collision risk model to assess bird collisions for offshore windfarms. Report Commissioned by Strategic Ornithological Support Services (SOSS). https://www.bto.org/sites/default/files/u28/downloads/Projects/Final_Report_SOSS02_Band1ModelGuidance.pdf. Last accessed: 02/05/2022.

Bartoń K (2018) MuMIn: multi-model inference, R package version 1.42.1 https: //cran.r-project.org/web/packages/MuMIn/.

Bates D, Maechler M, Bolker B & Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67: 1−48.

Beal M, Oppel S, Handley J, Pearmain EJ, Morera-Pujol V, Carneiro APB, Davies TE, Phillips RA, Taylor PR, Miller MGR, Franco AMA, Catry I, Patrício AR, Regalla A, Staniland I, Boyd C, Catry P & Dias MP (2021) track2KBA: An R package for identifying important sites for biodiversity from tracking data. Methods Ecol Evol 12: 2372-2378.

Beaugrand G, Edwards M, Brander M, Luczak C & Ibanez F (2008) Causes and projections of abrupt climate-driven ecosystem shifts in the North Atlantic. Ecol Lett 11: 1157-1168.

Bennison, A., Quinn, J. L., Debney, A., and Jessopp, M. (2019) Tidal drift removes the need for area-restricted search in foraging Atlantic puffins. Biol Lett 15: 20190208.

Berberan-Santos, M.N., Bodunov, E.N. & Pogliani, L. (1997) On the barometric formula. Am J Phys 65: 404–412.

Bodey TW, Cleasby IR, Bell F, Parr N, Schultz A , Votier SC & Bearhop S (2018) A phylogenetically controlled meta-analysis of biologging device effects on birds: Deleterious effects and a call for more standardized reporting of study data. Methods Ecol Evol 9: 946–955.

Bogdanova MI, Gunn C, Langlois Lopez S, Phillpot P, Newell M, Harris M, Wanless S & Daunt F (2018) GPS tracking of common guillemot, razorbill, Atlantic puffin and black-legged kittiwake on the Isle of May, summer 2018. Contract report to EDF Renewables.

Bogdanova, M.I., Langlois Lopez, S., Gunn, C., Tremlett, C., Newell, M., Harris, M.P., Searle, K., Wanless, S. & Daunt, F. (2020) GPS tracking of common guillemots, razorbills, Atlantic puffins and black-legged kittiwakes on the Isle of May in 2019 in relation to the Neart na Gaoithe offshore wind farm. Contract report to EDF Renewables.

Bogdanova, M.I., Gunn, C., Newell, M., Benninghaus, E., Bouten, W., Searle, K. & Daunt, F. (2021) GPS tracking of common guillemots, razorbills, Atlantic puffins and black-legged kittiwakes on the Isle of May in 2020 in relation to the Neart na Gaoithe offshore wind farm. Contract report to EDF Renewables.

Burnham, KP, Anderson, DR (2002) *Model Selection and Multimodel Inference: a Practical Information-Theoretic Approach*. Springer-Verlag, New York.

Cairns DK (1989) The regulation of seabird colony size: A hinterland model. Am Nat 134: 141.

Calenge C (2006) The package 'adehabitat' for the R software: a tool for the analysis of space and habitat use by animals. Ecol Model 197: 516−519.

Chivers, L. S., Lundy, M. G., Colhoun, K., Newton, S. F., Houghton, J. D., and Reid, N. (2012) Foraging trip time-activity budgets and reproductive success in the black-legged kittiwake. Mar Ecol Prog Ser 456: 269-277.

Chivers LS, Hatch SA & Elliott KH (2016) Accelerometry reveals an impact of short-term tagging on seabird activity budgets. Condor 118: 159-168.

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

Cleasby, I. R., Wakefield, E. D., Bearhop, S., Bodey, T. W., Votier, S. C., & Hamer, K. C. (2015) Three-dimensional tracking of a wide-ranging marine predator: flight heights and vulnerability to offshore wind farms. J Appl Ecol 52: 1474-1482.

Cook, A. S. C. P. (2014) The avoidance rates of collision between birds and offshore turbines. BTO Research Report.

Cook, A. S., Humphreys, E. M., Bennet, F., Masden, E. A., & Burton, N. H. (2018) Quantifying avian avoidance of offshore wind turbines: current evidence and key knowledge gaps. Mar Env Res 140: 278-288.

Cook, A. S. C. P., Ward, R. M., Hansen, W. S. & Larsen, L. (2018) Estimating Seabird Flight Height using LiDAR. Scottish Marine and Freshwater Science, 9, 14.

Daunt F, Benvenuti S, Harris MP, Dall'Antonia L, Elston DA & Wanless S (2002) Foraging strategies of the black-legged kittiwake *Rissa tridactyla* at a North Sea colony: evidence for a maximum foraging range. Mar Ecol Prog Ser 245: 239–247.

Daunt F & Bogdanova MI (2022) Quantifying the population-level effects of offshore renewable developments on seabirds in the Forth/Tay region using GPS tracking: empirical and analytical approaches to maximise insights from puffin data given known device effects in this species. Report to the Ornithology subgroup of the Forth and Tay Regional Advisory Group (FTRAG-O).

Daunt F, Bogdanova MI, McCluskie A & Wischnewski S (2021) Quantifying the population-level effects of offshore renewable developments on seabirds in the Forth/Tay region using GPS tracking: approach for 2021 breeding season work to the Ornithology subgroup of the Forth and Tay Regional Advisory Group (FTRAG-O). Report to FTRAG-O.

Daunt F, Bogdanova MI, Newell M, Harris M & Wanless S (2011a) GPS tracking of common guillemot, razorbill and black-legged kittiwake on the Isle of May in 2010. Contract report to the Forth and Tay Offshore Developer Group.

Daunt F, Bogdanova MI, Redman P, Russell S & Wanless S (2011b) GPS tracking of black-legged kittiwake and observations of trip durations and flight directions of common guillemot at Fowlsheugh and St Abb's Head, summer 2011. Report to Forth & Tay Offshore Wind Developers Group (FTOWDG)

Desholm M & Kahlert J (2006) Avian collision risk at an offshore wind farm. Biol Lett 1: 296-298.

Dierschke V, Furness RW & Garthe S (2016) Seabirds and offshore wind farms in European waters: Avoidance and attraction. Biol Cons 202: 59–68.

Drewitt AL, & Langston RHW (2006) Assessing the impacts of wind projects on birds. Ibis 148: S29–42.

Enstipp MR, Daunt F, Wanless S, Humphreys E, Hamer KC, Benvenuti S & Gremillet D (2006) Foraging energetics of North Sea birds confronted with fluctuating prey availability. In: Top predators in marine ecosystems: their role in monitoring and management. (Eds I.L. Boyd, S. Wanless & K. Camphuysen). Cambridge University Press, Cambridge, pp191-210.

Fayet, A. L., Clucas, G. V., Anker-Nilssen, T., Syposz, M., and Hansen, E. S. (2021) Local prey shortages drive foraging costs and breeding success in a declining seabird, the Atlantic puffin. J Anim Ecol 90: 1152-1164.

Finney SK, Wanless S & Harris MP (1999) The effect of weather conditions on the feeding behaviour of a diving bird, the Common Guillemot *Uria aalge*. J Avian Biol 30: 23-30.

Garriga, J., Palmer, J. R., Oltra, A., & Bartumeus, F. (2016) Expectation-maximization binary clustering for behavioural annotation. PloS one 11: e0151984.

Garthe, S., Guse, N., Montevecchi, W. A., Rail, J. F., & Grégoire, F. (2014) The daily catch: Flight altitude and diving behavior of northern gannets feeding on Atlantic mackerel. J Sea Res 85: 456-462.

Grecian WJ, Inger R, Attrill MJ, Bearhop S, Godley BJ, Witt MJ & Votier SC (2010) Potential impacts of wave-powered marine renewable energy installations on marine birds. Ibis 152: 683-97.

Grecian, W. J., Lane, J. V., Michelot, T., Wade, H. M., and Hamer, K. C. (2018) Understanding the ontogeny of foraging behaviour: insights from combining marine predator bio-logging with satellite-derived oceanography in hidden Markov models. J Roy Soc Interface 15: 20180084.

Harris M, Bogdanova MI, Daunt F & Wanless S (2012) Using GPS technology to assess feeding areas of Atlantic Puffins Fratercula arctica. Ringing Migr 27: 43-49.

Harris MP & Wanless S (2011) *The Puffin*. T & A D Poyser, London, 256pp.

Heggøy O, Christensen-Dalsgaard S, Ranke PS, Chastel O, Bech C (2015) GPS-loggers influence behaviour and physiology in the black-legged kittiwake *Rissa tridactyla*. Mar Ecol Prog Ser 521: 237–248.

Hijmans RJ (2019). raster: Geographic Data Analysis and Modeling. R package version 3.0-7. https://CRAN.R-project.org/package=raster.

Inger R, Attrill MJ, BearhopS, Broderick AC, Grecian WJ, Hodgson DJ, Mills C, Sheehan E, Votier SC, Witt MJ, Godley BJ (2009) Marine renewable energy: potential benefits to biodiversity? An urgent call for research. J Appl Ecol 46: 1145–1153.

Johnston A, Cook ASCP, Wright LJ, Humphreys EM & Burton NHK (2014) Modelling flight heights of marine birds to more accurately assess collision risk with offshore wind turbines. J Appl Ecol 51: 31–41.

Lane, J. V., Spracklen, D. V., & Hamer, K. C. (2019) Effects of windscape on three-dimensional foraging behaviour in a wide-ranging marine predator, the northern gannet. Mar Ecol Prog Ser 628: 183-193.

Langrock, R., King, R., Matthiopoulos, J., Thomas, L., Fortin, D., & Morales, J. M. (2012) Flexible and practical modeling of animal telemetry data: hidden Markov models and extensions. Ecology 93: 2336-2342.

Langton R, Davies IM & Scott BE (2011) Seabird conservation and tidal stream and wave power generation: information needs for predicting and managing potential impacts. Mar Policy 35: 623-30.

Largey, N., Cook, A. S., Thaxter, C. B., McCluskie, A., Stokke, B. G., Wilson, B., & Masden, E. A. (2021) Methods to quantify avian airspace use in relation to wind energy development. Ibis 163: 747-764.

Larsen JK & Guillemette M. (2007) Effects of wind turbines on flight behaviour of wintering common eiders: implications for habitat use and collision risk. J Appl Ecol 44: 516-522

Lascelles BG, Taylor PR, Miller MGR, Dias MP, Oppel S, Torres L, Hedd A, Le Corre M, Phillips RA, Shaffer SA, Weimerskirch H & Small C (2016) Applying global criteria to tracking data to define important areas for marine conservation. Div Distr 22: 422–431.

Manly BFJ (2009) *Randomization, bootstrap and Monte Carlo methods in biology*. Chapman Hall.

Manivannan, A., Chin, W. C. B., Barrat, A., & Bouffanais, R. (2020) On the challenges and potential of using barometric sensors to track human activity. Sensors 20: 6786.

Masden EA, Haydon DT, FoxAD & Furness, RW (2010) Barriers to movement: Modelling energetic costs of avoiding marine wind farms amongst breeding seabirds. Mar Pollut Bul 60: 1085–1091.

Masden, E. A. (2015) Developing an avian collision risk model to incorporate variability and uncertainty. Scottish Marine and Freshwater Science.

Masden, E. A., & Cook, A. S. C. P. (2016) Avian collision risk models for wind energy impact assessments. Env Impact Assess Rev 56: 43-49.

Masden, E. A., Cook, A. S., McCluskie, A., Bouten, W., Burton, N. H., & Thaxter, C. B. (2021) When speed matters: The importance of flight speed in an avian collision risk model. Env Impact Assess Rev 90: 106622.

McClintock, B. T. (2021) Worth the effort? A practical examination of random effects in hidden Markov models for animal telemetry data. Methods Ecol Evol 12: 1475-1497.

McGregor, R. M., King, S., Donovan, C. R. & Webb, A. (2018) A Stochastic Collision Risk Model for Seabirds in Flight. Marine Scotland. https://tethys.pnnl.gov/sites/default/files/publications/McGregor-2018-Stochastic.pdf

Melo, C. E., Melo, O. O. & Melo, S. E. (2022) Spatio-temporal radial bias functions with distance-based methods (Optimization, prediction and cross-validation). https://CRAN.R-project.org/package=geosptdb

Mendel, B., Kotzerka, J., Sommerfeld, J., Schwemmer, H., Sonntag, N. & Garthe S (2014) Effects of the offshore test site alpha ventus on distribution patterns, behaviour and flight heights of seabirds. In: Federal Maritime and Hydrographic Agency and Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (eds.) Ecological research at the offshore windfarm alpha ventus—challenges, results and perspectives. Springer Fachmedien Wiesbaden. doi:10.1007/978-3-658-02462-8_11, pp 95–110.

Nakagawa S & Schielzeth H (2013) A general and simple method for obtaining R^2 from generalized linear mixed effects models. Methods Ecol Evol 4: 133−142.

Péron, G., Calabrese, J. M., Duriez, O., Fleming, C. H., García-Jiménez, R., Johnston, A., Lambertucci, S. A., Safi, K. & Shepard, E. L. (2020) The challenges of estimating the distribution of flight heights from telemetry or altimetry data. Anim Biotelem 8: 1-13.

Peschko V, Mercker M & Garthe S (2020) Telemetry reveals strong effects of offshore wind farms on behaviour and habitat use of common guillemots (*Uria aalge*) during the breeding season. Mar Biol 167:118.

Peschko V, Mendel B, Mercker M, Dierschke J & Garthe S (2021) Northern gannets (*Morus bassanus*) are strongly affected by operating offshore wind farms during the breeding season. J Env Mnt 279: 111509.

Plummer, M. (2003) JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd international workshop on distributed statistical computing. 124 pp. 1-10.

Plummer, M. (2021) rjags: Bayesian Graphical Models using MCMC. R package version 4-12. https://CRAN.R-project.org/package=rjags.

Poessel, S. A., Duerr, A. E., Hall, J. C., Braham, M. A., & Katzner, T. E. (2018) Improving estimation of flight altitude in wildlife telemetry studies. J Appl Ecol 55: 2064-2070.

Robertson GS, Bolton M, Grecian WJ & Monaghan P (2014) Inter- and intra-year variation in foraging areas of breeding kittiwakes (*Rissa tridactyla*). Mar Biol 161: 1973–1986.

Ross-Smith VH, Thaxter CB, Masden EA, Shamoun-Baranes J, Burton NHK, Wright LJ, Rehfisch MM & Johnston A (2016) Modelling flight heights of lesser black-backed gulls and great skuas from GPS: a Bayesian approach. J Appl Ecol 53: 1676–1685.

Rodway MS, Montevecchi WA & Chardine JW (1996) Effects of investigator disturbance on breeding success of Atlantic puffins *Fratercula arctica*. Biol Cons 76: 311-319.

Searle K, Mobbs D, Butler A, Bogdanova MI, Freeman S, Wanless S & Daunt F (2014) Population consequences of displacement from proposed offshore wind energy developments for seabirds breeding at Scottish SPAs (CR/2012/03). Report to Scottish Government.

Searle KR, Mobbs DC, Butler A, Furness RW, Trinder TN & Daunt F (2018) Finding out the Fate of Displaced Birds. Scottish Marine and Freshwater Science Vol 9 No 8.

Shamoun-Baranes, J., Van Loon, E., van Gastere, H., van Belle, J., Bouten, W., & Buurma, L. (2006) A comparative analysis of the influence of weather on the flight altitudes of birds. Bull Am Meteorol Soc 87: 47-62.

Shoji A, Aris-Brosou S, Owen E, Bolton M, Boyle D, Fayet A, Dean B, Kirk H, Freeman R, Perrins C & Guilford T (2016) Foraging flexibility and search patterns are unlinked during breeding in a free-ranging seabird. Mar Biol 163: 72.

Shamoun-Baranes J, Bouten W, van Loon EE, Meijer C & Camphuysen CJ (2016) Flap or soar? How a flight generalist responds to its aerial environment. Phil Trans Roy Soc B 371: 20150395.

Symons SC & Diamond AW (2019) Short-term tracking tag attachment disrupts chick provisioning by Atlantic Puffins Fratercula arctica and Razorbills Alca torda. Bird Study 66: 53-63.

Thaxter CB, Daunt F, Gremillet D, Harris MP, Benvenuti, S,cWatanuki, Y, Hamer, KC & Wanless, S (2013) Modelling the effects of prey size and distribution on prey capture rates of two sympatric marine predators. PLOS ONE 8: e79915.

Thaxter CB, Daunt F, Hamer KC, Watanuki Y, Harris MP, Grémillet D, Peters G & Wanless S. (2009) Sex-specific food provisioning in a monomorphic seabird: nest defence, foraging efficiency or parental effort? J Avian Biol 40: 75-84.

Thaxter CB, Ross-Smith VH, Clark JA, Clark NA, Conway GJ, Marsh M, Leat EHK & Burton NHK (2014) A trial of three harness attachment methods and their suitability for long-term use on Lesser Black-backed Gulls and Great Skuas. Ring & Migr 29: 65-76.

Thaxter CB, Wanless S, Daunt F, Harris MP, Benvenuti S, Watanuki Y, Grémillet D & Hamer KC (2010) Influence of wing loading on trade-off between pursuit-diving and flight in common guillemots and razorbills. J Exp Biol 213: 1018-1025.

Thaxter CB, Ross-Smith VH, Clark JA Clark NA, Conway GJ, Masden EA, Wade H., Leat EHK, Gear SC, Marsh M, Booth C, Furness RW, Votier SC & Burton NHK (2017) Contrasting effects of GPS device and harness attachment on adult survival of Lesser Black-backed Gulls *Larus fuscus* and Great Skuas *Stercorarius skua*. Ibis 158: 279-290.

Thaxter, C. B., Ross-Smith, V. H., Bouten, W., Masden, E. A., Clark, N. A., Conway, G. J., Barber, L., Clewley, G. D. & Burton, N. H. (2018) Dodging the blades: new insights into three-dimensional space use of offshore wind farms by lesser black-backed gulls Larus fuscus. Mar Ecol Prog Ser 587: 247-253.

Vandenabeele SP, Shepard EL, Grogan A & Wilson RP (2012) When three per cent may not be three per cent: device equipped seabirds experience variable flight constraints. Mar Biol 159: 1–14.

Vandenabeele SP, Grundy S, Friswell MI, Grogan A, Votier SC & Wilson RP (2014) Excess baggage for birds: Inappropriate placement of tags on gannets changes flight patterns. PLOS One 9: e92657.

Vanermen N, Courtens W, Daelemans R, Lens L, Muller W, Van de walle M, Verstraete H & Stienen EWM (2020) Attracted to the outside: a meso-scale response pattern of lesser black-backed gulls at an offshore wind farm revealed by GPS telemetry. – ICES J Mar Sci 77: 701–710.

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P. C. (2021) Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with Discussion). Bayesian Analysis, 16, 667-718.

Wallace, J.M. & Hobbs, P.V. (2006) Atmospheric Science: An Introductory Survey. Academic Press, San Diego.

Wakefield ED, Bodey TW, Bearhop S, Blackburn J, Colhoun K, Davies R et al. (2013) Space partitioning without territoriality in gannets. Science 341: 68–70.

Wakefield ED, Owen E, Baer J, Carroll MJ, Daunt F, Dodd SG, Green JA, Guilford T, Mavor RA, Miller PA, Newell MA, Newton SF, Robertson GS, Shoji A, Soanes LM, Votier SC, Wanless S & Bolton M (2017) Breeding density, fine-scale tracking, and large-scale modelling reveal the regional distribution of four seabird species. Ecol Appl 27: 2074–2091.

Wanless S & Harris MP (1993) Use of mutually exclusive foraging areas by adjacent colonies of blue-eyed shags (Phalacrocorax atriceps) at south Georgia. Colon. Waterbirds 16: 176-182.

Wanless S, Harris MP, Newell MA, Speakman JR & Daunt F (2018) Community-wide decline in the occurrence of lesser sandeels *Ammodytes marinus* in seabird chick diets at a North Sea colony. Mar Ecol Prog Ser 600: 193–206.

Wischnewski S, Adlard S, McCluskie A & Wright LJ (2018) Seabirds and Wind Farms: Kittiwake and gannet tracking in the Flamborough and Filey Coast SPA. Fieldwork report 2018. RSPB Centre for Conservation Science Report to Ørsted.

Woodward I, Thaxter CB, Owen E & Cook ASCP (2019) Desk-based revision of seabird foraging ranges used for HRA screening. BTO Research Report No. 724. British Trust for Ornithology, Thetford.

Worton BJ (1989) Kernel methods for estimating the utilization distribution of home range studies. Ecology 70: 164–168.

Wright PJ, Jensen H & Tuck I (2000) The influence of sediment type on the distribution of the lesser sandeel *Ammodytes marinus*. J Sea Res 44: 243-256.

Xia, H., Wang, X., Qiao, Y., Jian, J., & Chang, Y. (2015) Using multiple barometers to detect the floor location of smart phones with built-in barometric sensors for indoor positioning. Sensors 15: 7857-7877.

Zuur AF, Ieno EN, Walker NJ, Saveliev AA, Smith GM (2009). *Mixed Effects Models and Extensions in Ecology with R.* Springer, New York.

7 Appendices

7.1 Appendix 1: Estimating flight altitude from bird-borne loggers

Bird-borne loggers allow time-series of geographical and vertical height to be combined within a single movement trajectory. However, altitude data collected by GPS loggers or via barometers do contain errors, most notably the occasional recording of negative values. More specifically, the altitude measured by a GPS unit can be affected by both user equivalent range error (UERE) and the vertical dilution of precision (VDOP). UERE is a key performance parameter of GPS. It combines multiple different sources of GPS error within a single term including diffusion and defraction of GPS signals in the atmosphere, multipath effects (where signals may have been reflected by local objects), and clock errors. The exact arrangement of satellites in the sky also influences the accuracy of GPS positioning. VDOP quantifies the effect of the available satellite network on the precision of GPS records in the vertical dimension. Some GPS units report VDOP directly but others may report the more generic Dilution of Precision (DOP) or other similar DOP measures. Ideally, GPS coordinates would be based on signals from multiple satellites spaced evenly apart. When satellites are clustered close together in the sky then the dilution of precision increases. DOP multiplies the uncertainty associated with UERE hence larger DOP values are undesirable but may sometime be unavoidable.

In contrast to GPS-derived measures of altitude, flight heights calculated using altimeters do so by using a standard barometric equation to convert recorded pressure values into altitude. However, this formula only holds when the atmosphere is at equilibrium. This equilibrium can be influenced by factors such as changes in temperature, air pressure and air composition. Therefore, local weather conditions can often have a large impact on altimeter-derived estimates of flight height. When altimeters can be suitably calibrated to local conditions, they can be more accurate than GPS loggers, particularly over short periods when weather conditions could be considered relatively consistent (Peron et al. 2020). However, for longer term tracking deployments barometer data will require constant recalibration as weather conditions change and birds move throughout their foraging range. Ideally, altimeters would be regularly re-calibrated using direct observations of flight height combined with accurate measures of air pressure at sea-level. However, this data is often difficult to obtain and instead re-calibration has often been achieved by identifying periods when tracked birds are sat on the water to provide a measure of reference sea-level pressure (Cleasby et al. 2015, Lane et al. 2019). In the absence of direct observations of tracked birds floating on the sea surface, such floating behaviour must be identified by researchers, typically via a behavioural classification algorithm of some kind. This approach will introduce some calibration error, most pertinently when birds are classified as on the water but are in fact in-flight. Secondly, some inflight observations may not be close in space or time to a suitable re-calibration period when the tracked individual was on the water. In extreme cases a bird may complete a foraging round trip while spending little to no time on the water making calibration and therefore flight height estimation difficult. To address these issues, ERA5 remote sensing data (European Centre for Medium-Range Weather Forecasts, accessed: 14/04/2022) can provide sea-level pressure data at hourly intervals for barometer calibration (Manola et al. 2020), albeit at a relatively coarse spatial scale (0.25° × 0.25°). Reference sea-level pressure data can also be obtained from already deployed weather buoys or local weather stations (Garthe et al. 2014), although this may only be suitable for calibration within the immediate vicinity of a weather stations. On the other hand, data from these sources has the advantage that they do not rely on observations of birds floating on the sea.

Correlation between GPS and barometric flight heights among birds tracked at the same colony. While GPS and barometric flight height were highly correlated in some birds, estimates were less correlated in others (Table S1).

Colony	Logger ID	Correlation GPS vs. barometric flight height
Fowlsheugh	2665	0.31
	2666	0.56
	2667	0.37
	2668	0.31
	2669	0.19
	2671	0.42
	2672	0.36
	2673	0.30
	2674	0.44
	2675	0.64
	2676	0.10
	2678	0.56
Isle of May	2641	0.44
	2642	0.26
	2643	0.19
	2646	0.56
	2648	0.29
	2652	0.45
	2653	0.59
	2654	0.23
	2656	0.37
	2661	0.31
	2662	0.52
	2663	0.36
	2664	0.45

St Abb's Head	2679	0.53
	2702	0.69
	2705	0.56
	2706	0.80
	2707	0.81
	2708	0.83
	2709	0.22
	2711	0.59
	2712	0.24
	2713	0.26
	2714	0.53
	2715	0.87

Table S1. Correlation between GPS derived estimates of flight height and those based on barometric pressure using calibration within a single tag deployment.

References

Cleasby, I. R., Wakefield, E. D., Bearhop, S., Bodey, T. W., Votier, S. C., & Hamer, K. C. (2015). Three-dimensional tracking of a wide-ranging marine predator: flight heights and vulnerability to offshore wind farms. Journal of Applied Ecology, 52, 1474-1482.

Garthe, S., Guse, N., Montevecchi, W. A., Rail, J. F., & Grégoire, F. (2014). The daily catch: Flight altitude and diving behavior of northern gannets feeding on Atlantic mackerel. Journal of Sea Research, 85, 456-462.

Ladetto, Q., Merminod, B., Terrier, P. & Schutz, Y. (2000) On foot navigation: when GPS alone is not enough. Journal of Navigation, 53, 279– 285.

Lane, J. V., Spracklen, D. V., & Hamer, K. C. (2019). Effects of windscape on three-dimensional foraging behaviour in a wide-ranging marine predator, the northern gannet. Marine Ecology Progress Series, 628, 183-193.

Langley, R. B. (1999). Dilution of precision. GPS world, 10, 52-59.

Manola, I., Bradarić, M., Groenland, R., Fijn, R., Bouten, W., & Shamoun-Baranes, J. (2020). Associations of synoptic weather conditions with nocturnal bird migration over the North Sea. Frontiers in Ecology and Evolution, 328.

Péron, G., Calabrese, J. M., Duriez, O., Fleming, C. H., García-Jiménez, R., Johnston, A., Lambertucci, S. A., Safi, K. & Shepard, E. L. (2020). The challenges of estimating the distribution of flight heights from telemetry or altimetry data. Animal Biotelemetry, 8, 1-13.

Ross-Smith, V. H., Thaxter, C. B., Masden, E. A., Shamoun-Baranes, J., Burton, N. H., Wright, L. J., Rehfisch, M. M.& Johnston, A. (2016). Modelling flight heights of lesser black-backed gulls and great skuas from GPS: a Bayesian approach. Journal of Applied Ecology, 53, 1676-1685.

7.2 Appendix 2: Behavioural classification

Supplementary description of the time-of-day covariate in the kittiwake HMMs

In the kittiwake HMMs for behavioural classification, time of day was included as a candidate covariate in influencing the transition probabilities between behavioural states. For each interpolated GPS location, time of day was calculated as a decimal hour of the day (i.e. the number of hours since midnight, ranging from 0 to 24). Decimal hour of the day was then included as a cyclical covariate, with a 24-hour periodicity. This was assumed to follow a sinusoidal pattern, where a cyclical pattern was included by adding two variables: cos(2π x hour of day / 24) and sin(2π x hour of day / 24). Including both covariates is equivalent to including a sinusoidal pattern with unknown phase. This was implemented in the R package "momentuHMM" using the "cosinor" function.

Model	Covariates on the transition probabilities			ΔAIC
	Hour of day	Colony	Logger type	
1 ("full model")	Y	Y	Y	0
2	Y	-	Y	155
3	Y	Y	-	284
4	Y	-	-	578
5	-	Y	Y	1541
6	-	-	Y	1731
7	-	Y	-	1743
8 ("null model")	-	-	-	2028

Table S2: Comparison of kittiwake three-state HMMs by AIC. Models were fit with different combinations of covariates on the transition probabilities between the behavioural states. Models are ordered by ΔAIC (the difference in AIC between the model with the lowest AIC and all other models).

		To state:		
		1	2	3
From state:	1	0.85	0.15	<0.01
	2	0.07	0.88	0.05
	3	<0.01	0.18	0.82

Table S3: Estimated average transition probability matrix between the three states for kittiwakes (predicted from the null model with no covariates).

		To state:	
		1	2
From state:	1	0.84	0.16
	2	0.10	0.90

Table S4: Estimated average transition probability matrix between states for the two-state model for puffins.

		To state:		
		1	2	3
From state:	1	0.69	0.26	0.05
	2	0.18	0.53	0.28
	3	<0.01	0.12	0.87

Table S5: Estimated average transition probability matrix between states for the three-state model for puffins.

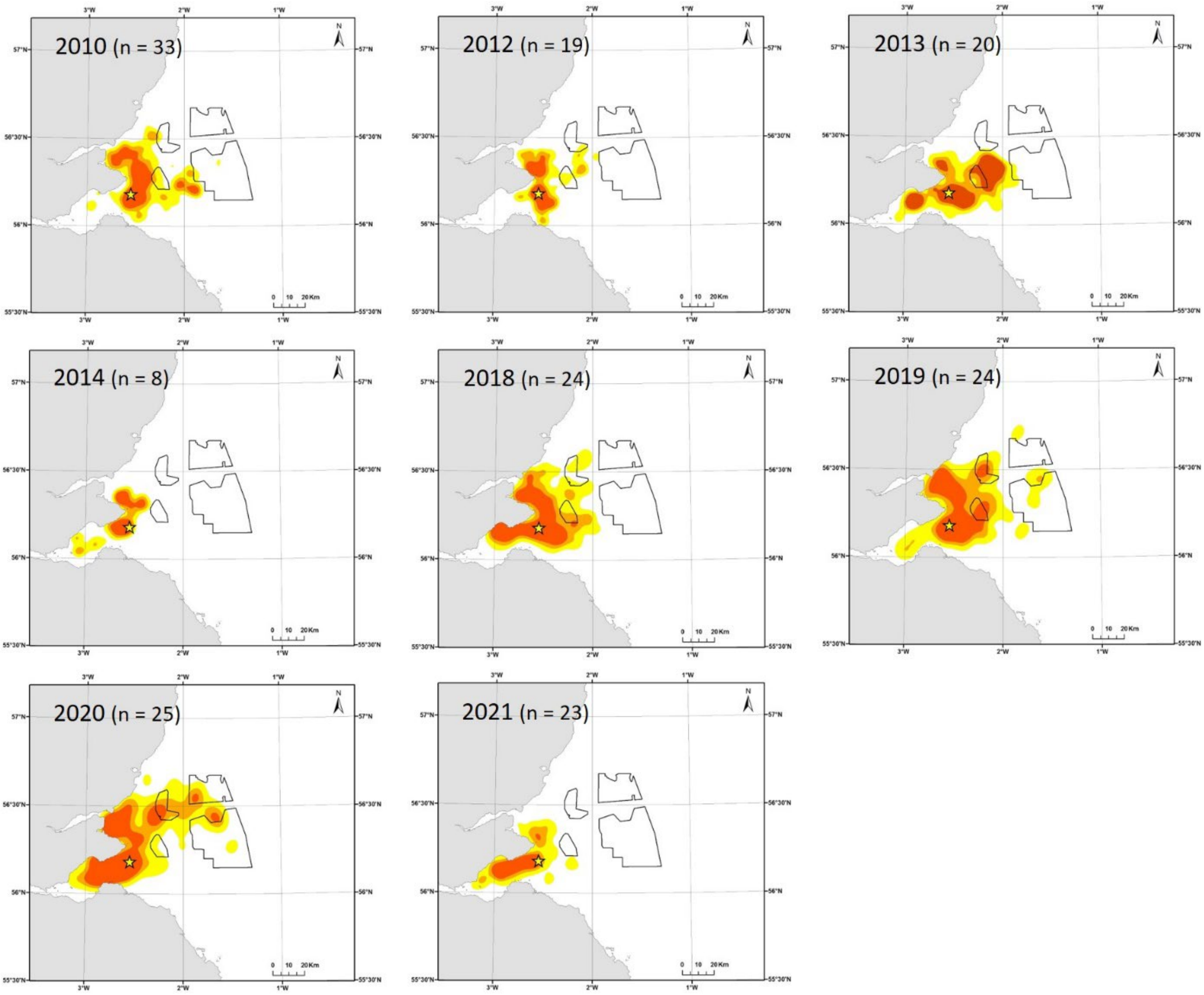
State assigned in 2-state model	Percentage of locations assigned to each state in the 3-state model		
	1	2	3
1	36%	58%	6%
2	<1%	<1%	98%

Table S6: Comparison of state assignment between the two-state and three-state models for puffins.

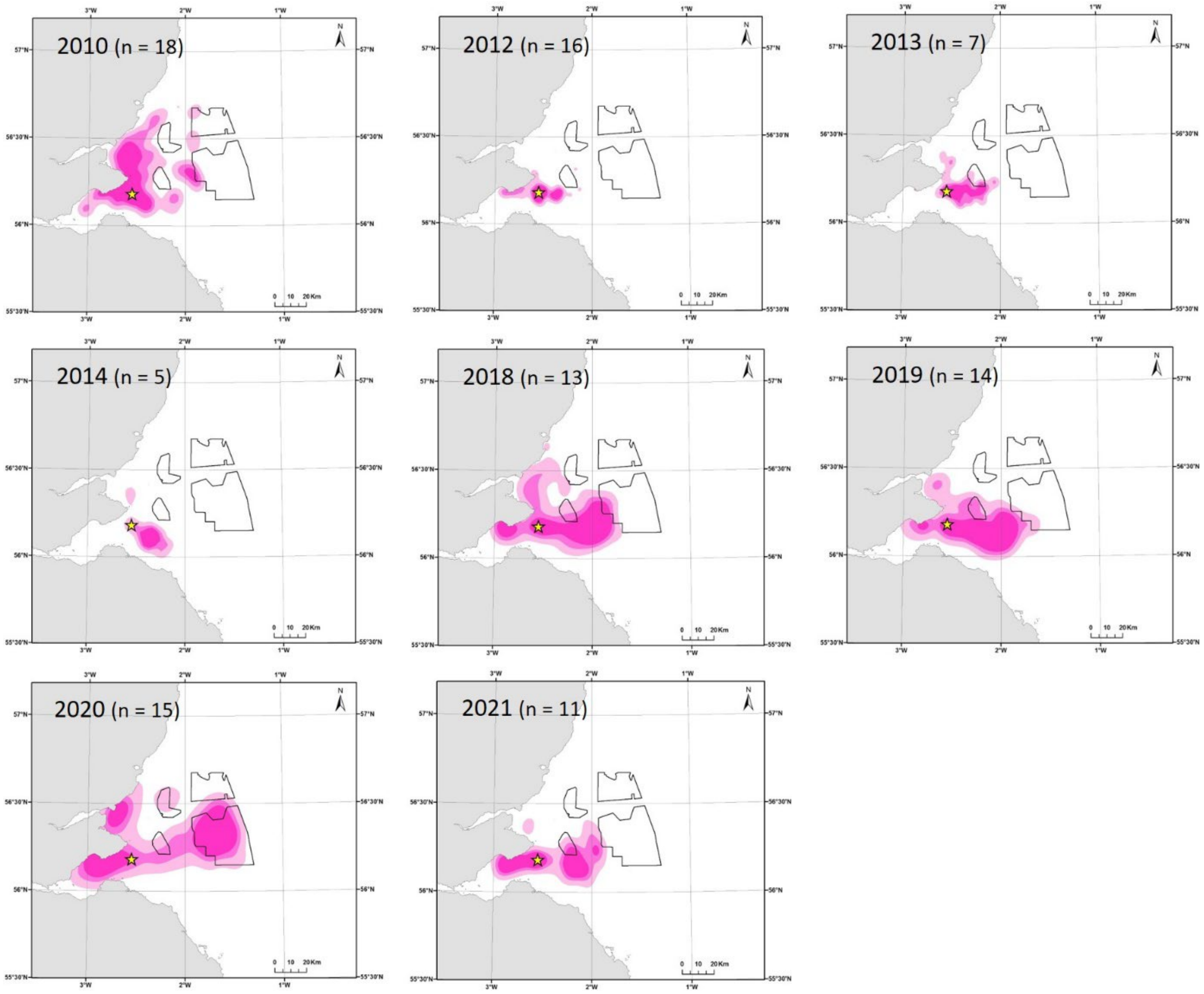
7.3 Appendix 3: Inter-annual variation in at-sea distribution of four seabird species breeding on the Isle of May

50%, 70% and 90% UD contours are presented, yellow star denotes the location of the breeding colony.

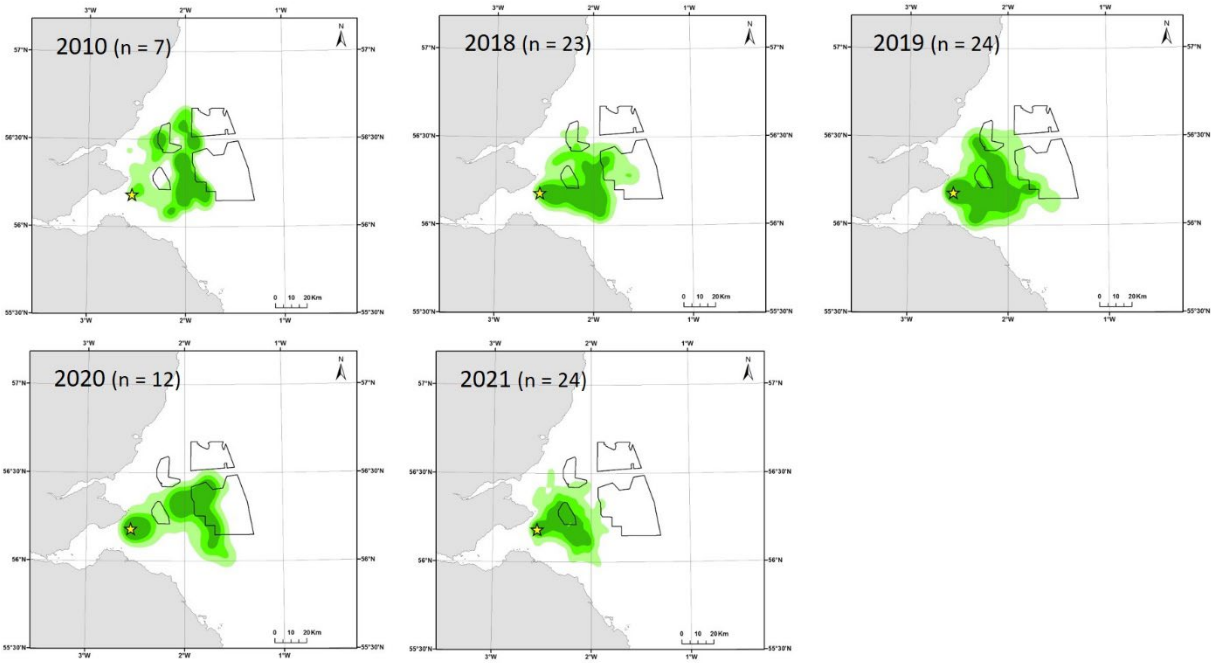
Guillemot



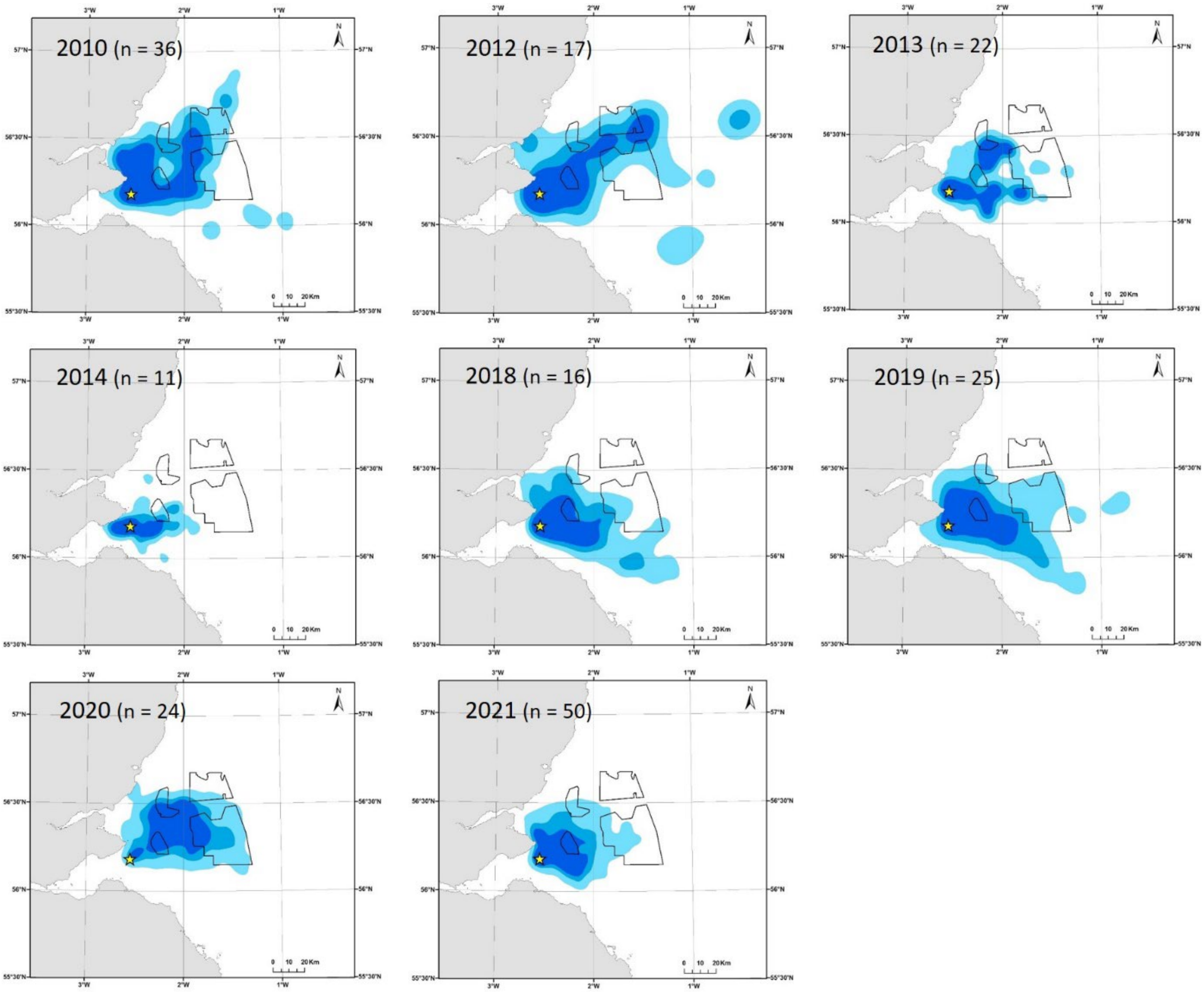
Razorbill


2010 (n = 18)
2012 (n = 16)
2013 (n = 7)
2014 (n = 5)
2018 (n = 13)
2019 (n = 14)
2020 (n = 15)
2021 (n = 11)
N
3°W
2°W
1°W
57°N
56°30'N
56°N
55°30'N
0 10 20 Km

Puffin


2010 (n = 7)
2018 (n = 23)
2019 (n = 24)
2020 (n = 12)
2021 (n = 24)
3°W
2°W
1°W
57°N
56°30'N
56°N
55°30'N
N
0 10 20 Km

Kittiwake


2010 (n = 36)
2012 (n = 17)
2013 (n = 22)
2014 (n = 11)
2018 (n = 16)
2019 (n = 25)
2020 (n = 24)
2021 (n = 50)
3°W
2°W
1°W
57°N
56°30'N
56°N
55°30'N
N
0 10 20 Km

BANGOR
UK Centre for Ecology & Hydrology
Environment Centre Wales
Deiniol Road
Bangor
Gwynedd
LL57 2UW
United Kingdom
T: +44 (0)1248 374500
F: +44 (0)1248 362133

EDINBURGH
UK Centre for Ecology & Hydrology
Bush Estate
Penicuik
Midlothian
EH26 0QB
United Kingdom
T: +44 (0)131 4454343
F: +44 (0)131 4453943

LANCASTER
UK Centre for Ecology & Hydrology
Lancaster Environment Centre
Library Avenue
Bailrigg
Lancaster
LA1 4AP
United Kingdom
T: +44 (0)1524 595800
F: +44 (0)1524 61536

WALLINGFORD (Headquarters)
UK Centre for Ecology & Hydrology
Maclean Building
Benson Lane
Crowmarsh Gifford
Wallingford
Oxfordshire
OX10 8BB
United Kingdom
T: +44 (0)1491 838800
F: +44 (0)1491 692424

enquiries@ceh.ac.uk

www.ceh.ac.uk